

**ANALYSING ALLEGEDLY EXCESSIVE
PRICES CHARGED BY TRAIN
OPERATING COMPANIES**

**A Final Report for the Office of the
Rail Regulator**

Prepared by NERA

March 2001
London

**Project Team:
John Dodgson
Simon Baker
Derek Ridyard
Ian Jones
Andrea Coscelli
James Cameron**

n/e/r/a

National Economic Research Associates
Economic Consultants

15 Stratford Place
London W1N 9AF
Tel: (+44) 20 7659 8500
Fax: (+44) 20 7659 8501
Web: <http://www.nera.com>

An MMC Company

TABLE OF CONTENTS

EXECUTIVE SUMMARY	i
1. INTRODUCTION	1
2. BACKGROUND	2
2.1. The Competition Act	2
2.2. The Standard Approach	2
2.3. Problems in the Standard Approach	3
2.4. Additional Complications in UK Railway Cases	4
3. IDENTIFYING DOMINANCE	5
3.1. The Concept of Dominance	5
3.2. Market Definition	6
3.3. Direct Evidence of Dominance	13
3.4. Market Definition in Train Operation	16
4. ANALYSING ALLEGATIONS OF EXCESSIVE PRICING	18
4.1. Precedent and Guidance on Excessive Pricing	18
4.2. Implications of Common Costs	19
4.3. Estimating a Normal Rate of Return for a Trading Business	23
4.4. Assessing Excessive Prices when there is Competition for the Market through Franchising	30
5. CONCLUSIONS	38
5.1. Introduction	38
5.2. The 1998 Competition Act	38
5.3. Market Definition	38
5.4. Direct Evidence of Dominance	38
5.5. Implications of Common Costs	39
5.6. Estimating a Normal Rate of Return for a Trading Business	39
5.7. Excessive Pricing with Franchising	40
APPENDIX A. TERMS OF REFERENCE ON CONSULTANTS REPORT ON FARES POLICIES	41
A.1. Context	41
A.2. Contents of Report	41
A.3. Market Definition	41
A.4. Assessment of Abuse	43
A.5. Empirical Case Studies	44

APPENDIX B. INTER-MODAL COMPETITION, CONSTRAINTS ON TOC PRICING, AND PREVIOUS MARKET DEFINITION ASSESSMENTS	45
B.1. Introduction	45
B.2. The Own –Price Elasticity of Demand for Rail Travel	45
B.3. Evidence on Inter-Modal Competition and Cross Elasticities	54
B.4. Review of Previous Studies of Market Definition in the Rail Sector	57
B.5. A Template for Excessive Pricing Cases	65

EXECUTIVE SUMMARY

Introduction

1. This report has considered the issue of the identification of excessive prices charged by Train Operating Companies (TOCs). The purpose of the project was to design a framework for market definition and the analysis of the abuse of dominant position, as these concepts arise under Chapter II of the 1998 Competition Act. The report sets out the relevance of market definition in the context of abuse of dominance cases. It considers the specific issues that would arise in assessing dominance in the case of a TOC, including: the issues arising in assessing whether individual fares might be judged to be excessive in the context of the allocation of common costs between passengers on the same train; the assessment of a normal return for “trading enterprises” such as TOCs; and the issues that arise because TOCs have competed for the right to operate their services through the franchising process.

The 1998 Competition Act

2. The 1998 Competition Act contains two general prohibitions. The Chapter I prohibition prohibits agreements which have “...the objective or effect of preventing, restricting or distorting competition in the UK”. The Chapter II prohibition prohibits conduct that “...amounts to the abuse of a dominant position...”. It is not the existence of dominance that is prohibited under Chapter II, but abuse of that dominance. This report is concerned with Chapter II prohibitions.

Market Definition

3. A pre-condition for abuse must be that the firm has a dominant position to abuse. Under European competition law the courts have consistently ruled that the Commission must define a market before it can reach conclusions on the market position of the firm or firms under investigation. As a consequence the first stage of any standard analysis under Article 82, on which Chapter II is based, is to identify the market within which the allegedly dominant firm operates. Once a market has been defined, it is necessary to determine whether the firm under investigation holds a dominant position in that market.
4. The standard market definition test was developed in merger cases, and relates to the ability of the merged firm to raise prices by a “small but significant amount”. However, this test is not appropriate in the case of abuse of dominance cases, where firms will already be setting prices that maximise their profits.

Direct Evidence of Dominance

5. The alternative approach to the one based on market definition is to consider whether the existing prices are above the competitive level, thus providing direct evidence of abuse. This can be done either by assessing the overall profitability of the firm, to see whether the firm is earning more than a normal rate of return, or by assessing the relationship between the prices of individual goods or services, to see whether these prices exceed costs by a margin greater than that required to earn a normal return. This approach is consistent with the OFT's guidance on the evaluation of excess pricing allegations.
6. While this approach does not depend on identification of the appropriate market (since the questions are simply whether profits are excessive or whether individual prices are excessive in relation to unit costs), market definition will still need to be addressed in any formal investigation for legal reasons. An appendix to this report reviews market definition issues and precedent with regard to rail services, including evidence on fare elasticities and competition between modes.
7. There may be more direct evidence of dominance. One approach would be to consider whether specific prices were high in relation to the prices of closely-related products produced under similar cost conditions: in the railway context this would involve comparisons with fares charged by TOCs providing similar services to the one under investigation. Another source of direct evidence of dominance would be evidence of cross-subsidies within franchises, which meant that a TOC had the ability to finance losses on some routes from profits on others. This type of evidence would be particularly compelling where the franchisee had been prepared to pay positive sums for the right to operate the franchise. In this case, the franchisee must be expecting to set fares that enable it to cover the costs of running services **and** to generate enough revenue to meet the franchise payments.

Implications of Common Costs

8. A major complication in the rail industry is that a large proportion of the total costs of operating trains on a route cannot be directly attributed to a particular fare type because there are typically few costs that could be avoided either in the short run or the long run by the withdrawal of that fare type. This means that there is a large cost that cannot be attributed to any single fare type in any simple and unequivocal way, cannot be avoided by withdrawal of any single fare type, but that must be recovered from the overall operation of the service through the setting of prices across all fares that in aggregate will cover those common costs. This makes it difficult to assess profitability of individual fare categories: this profitability will depend on the cost attribution method used.

Estimating a Normal Rate of Return for a Trading Business

9. The normal methodology for assessing company profitability by focusing on the achieved return on capital employed and comparing it with a normal rate of return for the type of activity concerned cannot be readily applied to a trading business like a TOC with few or no requirements to finance capital expenditure. Our review of regulatory precedents in this area shows that:
- in contrast to the position in other regulated industries, there is as yet no consensus as to an appropriate conceptual approach to assessing a reasonable operating margin for such businesses; and
 - previous experience suggests that, because of this, competition authorities will be extremely reluctant to reach conclusions on whether a particular level of operating margin constitutes excessive profitability.

Excessive Pricing with Franchising

10. The complication that arises with franchising is that TOCs will have competed for the right to operate services through a competitive franchising process. If this process is efficient, it should extract the value of the ex ante expected profits from the system. It might be regarded as a reasonable expectation on the part of the TOCs that they would then be able to recoup payments made, either those made directly to OPRAF/SRA or those made within their franchise to cross-subsidise loss-making services, by means of fares charged.
11. In assessing this, an important issue is that of how far TOCs are constrained in the fares they can charge for individual services, given that they knew in advance that some fares would be regulated, and some would not. The crucial consideration is that of whether competition law imposes constraints on the levels that can be set for those fares that were **not** to be subject to explicit regulation. Investigation of such fares would require allocation of costs between passengers, and such allocation would be dependent on the allocation rule used.
12. As a matter of public policy, NERA's own view is that if the public authorities wish to control fare levels, it is much more efficient to do so through explicit regulation rather than through use of Chapter II investigations.
13. Profits earned by a franchise might also be regarded as excessive if there had been completely unanticipated increases in the demand for rail services that led to "windfall" profits. An investigation would need to show that such increases could not have been expected at the time bids were made. Since bids would likely have been made on the basis of a range of forecasts, proof of this would depend on actual traffic levels being outside the range of traffic forecasts made.

14. Finally, a TOC might be able to earn excessive profits if the franchising process had not been efficient, and so did not extract the value of ex ante anticipated profits. Such a line of investigation would have to reveal that there had been serious flaws in the conduct of the original OPRAF franchise bidding process.

1. INTRODUCTION

This report is concerned with the analysis of excessive prices charged by Train Operating Companies (TOCs). The purpose of the project on which it was based was to design a framework for market definition and the analysis of the abuse of a dominant position, as these concepts arise under Chapter II of the new UK Competition Act (the Act), and may be applied to train operating companies (TOCs). NERA was required to assess the practicality of applying this framework. The Office of the Rail Regulator (ORR) is particularly interested in the analysis of the identification of excessive prices. Our Terms of Reference are set out in Annex A.

The report is structured as follows. Section 2 briefly discusses the background to the project, including a short summary of the 1998 Competition Act, the standard approach to the enforcement of the European law on which it is based, and a review of the difficulties with the application of that approach. In addition, we note those aspects of the operation of the TOCs in the UK that may introduce particular difficulties into the application of the standard approach in this case.

Section 3 considers the identification of dominance. Market definition is traditionally central to any assessment of dominance and the section is primarily focussed on the problems associated with the application of the usual market definition framework in cases involving allegedly dominant firms. However, in addition, the section considers how the different markets within which the TOCs typically operate might be defined, exploring the possibility of distinguishing rail markets by fare type. Importantly, the section also addresses the type of evidence that may be used to support a finding of dominance independently of any analysis of market definition and market share that may be undertaken.

The analysis of allegedly excessive prices in the context of UK TOCs is explored in Section 4. Specifically, the section explores the precedent and guidance that had been issued in relation to the application of the new Act, as well as any sector-specific guidance or statements made about the regulation of pricing and the application of the Act to the TOCs. The section also looks at the difficulties for the measurement of excess prices caused by the high level of common costs that exist in the operation of train services across different fare types. It then considers difficulties of assessing a normal return in a trading business such as a TOC. Finally Section 4 considers the difficulties arising because of the existence of a competitive franchising process.

Section 5 sets out the main conclusions of the report.

2. BACKGROUND

2.1. The Competition Act

The Competition Act (1998) (the Act) came into force in March 2000. It is largely modelled on Articles 81 and 82 of the Treaty of Amsterdam and has replaced the Restrictive Trade Practices Act (1976), the Resale Prices Act (1976) and most elements of the Competition Act (1980).

The Act contains two general prohibitions. The Chapter I prohibition, based on Article 81, prohibits agreements which have "...the objective or effect of preventing, restricting or distorting competition in the UK...". This project is not concerned with Chapter I and it is not considered further in our report.

The Chapter II prohibition, based on Article 82, prohibits conduct that "...amounts to the abuse of a dominant position...". It should be noted that it is not the existence of dominance that is prohibited under Chapter II, but the anti-competitive exploitation or "abuse" of that dominance.

There is no comprehensive list of behaviour that will be found to be abusive. However it is possible to construct a list of those forms of behaviour that may be found to be abusive based on the illustrative abuses listed in European legislation and on the past practice of the European competition authorities and the courts. Charging excessive prices, engaging in price discrimination, predatory pricing and refusal to supply are all activities that could be construed as abusive when undertaken by a dominant firm.

2.2. The Standard Approach

Although the Act is a relatively new piece of legislation, its prohibitions are closely modelled on those that have operated in European competition law for many years. As a consequence there is a substantial body of legal and analytical precedent to guide those seeking to apply the new prohibitions. This legal and administrative precedent has led to the emergence of a more-or-less standard approach to the assessment of dominant firm behaviour, which is described below.

Notably, under European competition law the courts have consistently ruled that the Commission must define a market before it can reach conclusions on the market position of the firm or firms under investigation. For example, in *Continental Can Co Inc.* (Case 6/72 (1973) ECR 215) the European Court held that a market must be defined before a dominant position can be found. As a consequence the first stage of any standard analysis under Article 82 is to identify the market within which the allegedly dominant firm operates.

Once a market has been defined it is necessary to determine whether the firm under investigation holds a dominant position in that market. The Courts have defined dominance

as a firm's ability to behave "...to an appreciable extent independently of its competitors and customers and ultimately of consumers". The economic interpretation of this legal definition has been to view dominance as synonymous with the possession of significant and enduring market power, such that a firm is dominant if it has the ability to set prices that are above those that would prevail in a competitive market.

Recent economic theory has begun to cast serious doubt over the presumed linkage between a high share of a particular market and the possession of market power. Nonetheless, as a practical matter, it is the market share of the allegedly dominant firm that has usually been used as an initial indicator of dominance in investigations under Article 82. Although neither the European Court nor the Commission have been dogmatic over the level of market share at which dominance arises, shares above 40 per cent will usually be suspected of conferring dominance and shares in excess of 50 per cent are very likely to be regarded as doing so. Only if the other relevant evidence on dominance, such as the ease of new entry into the market, is extremely strong will the presumption of dominance that flows from a high market share be overturned. The perceived analytical value of the market share of the allegedly dominant firm therefore provides a second reason for defining a market, alongside the legal requirement to do so.

Once dominance has been established, the final step of the standard approach to the assessment of an allegation of abuse of dominance is to evaluate the alleged abuse. On occasion it has appeared that certain practices, when undertaken by dominant firms, have the status of *per se* offences. In other cases the identification of the abuse has involved a more sophisticated economic approach or rule-of-reason. The application of the rule of reason has typically involved trying to identify the economic consequences of the allegedly abusive behaviour and, when done properly, has involved seeking to identify the harm to consumers that is immediately or ultimately likely to flow from the behaviour of the dominant firm.

2.3. Problems in the Standard Approach

The standard approach to the investigation of dominance would appear to be straightforward. However, there are a number of problems in the application of the standard approach. The most well known and potentially significant issue raised by the standard approach is that there is a profound conceptual problem with the application of the standard market definition framework in situations in which it is believed that there is pre-existing market power. This problem is known as the "cellophane fallacy" and is discussed in greater detail in Section 3.2.2 of this report.

A second problem with the application of the standard approach has been in its application to alleged instances of excessive pricing. The European Court has confirmed that it is an abuse to charge prices that "bear no relation to the economic value of the product", but have provided no guidance as to how to apply this benchmark in practice. As a result of the difficulties of identifying whether or not a price bears any relation to its economic value

there have been no cases where the Commission has successfully prosecuted firms for excessive pricing under Article 82.

2.4 Additional Complications in UK Railway Cases

In the case of the UK train operating companies (TOCs) there are a number of additional difficulties in applying Chapter II, as well as the more general problems outlined above. The most significant issue in the context of the evaluation of the market position and the behaviour of TOCs is that they hold the market positions that they do as a result of a competitive auction of monopoly rights. Thus, whilst it seems highly likely that several TOCs possess significant and enduring market power, in relation to at least some of the routes that they operate, there was competition for the market at the time of the initial franchising.

3. IDENTIFYING DOMINANCE

3.1. The Concept of Dominance

The concept of dominance finds legal expression in the decision of the European Court of Justice in *United Brands*, where it stated that dominance is the ability:

“...to prevent effective competition being maintained on the relevant market by affording it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers.”

As previously mentioned, the European Commission has stated that dominance can generally be taken to exist when a firm has a market share greater than 40 per cent of the relevant market.

In economic terms, the only firms that can prevent the maintenance of effective competition and behave independently of their competitors are those that possess significant market power.¹ The economic identification of dominance therefore amounts to the identification of firms with significant market power.

Market power can be defined in economic terms as the ability to set prices that are above the levels that would prevail in a competitive market. Market power therefore exists whenever the competitive process fails to place a firm under sufficient competitive pressure to cause that firm to price at the competitive level.²

On this basis dominance cannot exist where there is an effective competitive process capable of constraining all firms to price at competitive levels. However this raises the issue as to whether franchisees who have been subject to competition at the bidding stage can legitimately be regarded as dominant firms. In other words, is the existence of effective competition for the right to operate in a market sufficient for one to conclude that a dominant position cannot be held in that market, even if the ultimate outcome of that competitive process is the award of monopoly rights?

It seems to us that merely because a position of significant market power has been acquired in a competitive bidding process it does not necessarily follow that that position of market

¹ In fact, it is rather difficult to reconcile the idea of independent action with any sensible economic interpretation of dominance. This is because all firms, not matter how dominant, will be constrained from raising their prices yet further by something. This may be direct competition from an identifiable rival or it may be willingness on the part of customers simply to cease to buy at an even higher price. Either way the firm is constrained by at least some competitors or customers.

² A more narrow definition of market power holds that firms possess market power whenever they have the ability to set prices in excess of short run marginal cost or face downward sloping demand. Whilst this definition has theoretical attractions, as a benchmark for competition policy intervention it is seriously flawed because it would lead almost all firms to be regarded as in possession of market power.

power should not be regarded as conferring dominance. For the lifetime of the franchise the winning bidder may face no effective competition within the market, may have the ability to set prices that exceed those that would have prevailed in a more competitive environment and may be able to engage in activities that have an exclusionary effect. Based on the economic interpretation of dominance described above and in the light of the economic rationale for competition laws to control dominant firm behaviour we see no reason why one should automatically conclude that a firm that has acquired its market power in a competitive bidding process cannot be dominant.

Indeed, as we shall discuss later, the existence of a competitive bidding process in which participants were prepared to bid significant positive sums for the right to operate the franchise actually provides compelling evidence in support of the contention that the winning bidders hold dominant positions in at least some of the markets in which they operate. However, we nevertheless argue that, whilst the existence of such a process is not incompatible with a finding of dominance, it creates considerable difficulties for any attempt to judge whether the prices charged by the franchisee are excessive.

3.2. Market Definition

Market definition has come to be a central part of almost all competition law enforcement. There are both legal and economic reasons for the core role that market definition has come to play in competition analysis. The legal rationale for undertaking a market definition exercise in any case is that the courts have consistently ruled that a market must be defined before a position of dominance can legally be established under the European competition law that UK law now largely mimics.³

The economic rationale for market definition is that the act of market definition can:

- focus the competitive analysis on the main competitive constraints on a firm and permit the elimination from the analysis of irrelevant or misleading considerations at an early stage;
- permit the calculation of market shares, giving an initial indication of the extent of possible market power; and
- identify the set of products into which entry must be contemplated in establishing the nature of the relevant entry barriers.

In order for market definition to play this substantive role in the analysis it is necessary that the market is defined in a manner such that a high market share is likely to confer the ability to raise prices, whilst a low share is unlikely to confer such ability. If the market is not defined in such a way then any inferences about market power and dominance that are

³ See for example *Continental Can Co Inc.* (Case 6/72 (1973) ECR 215).

drawn from the resultant market shares or from the analysis of entry barriers into that market, as defined, will be of very limited value.

3.2.1. The basic market definition framework

Whilst the courts have consistently ruled that markets must be defined they have provided no detailed guidance on how this should be done. The approach that has now come to be standard in competition law enforcement is the approach set out in the Commission Notice on market definition. This is also broadly the approach set out in the OFT/ORR guidance on market definition. However the origins of this approach can be found in the 1992 US merger guidelines, where the basic market definition test that has now been almost universally adopted is described in some detail. It is significant that the origins of the test are in merger guidelines and the use of essentially the same framework in non-merger contexts is problematic and contentious, as we will discuss in more detail below.

However, the basic market definition test, commonly referred to as the “5 per cent” test, defines a market as the smallest set of products over which a single supplier would need to have control in order to be able to raise prices by a small, but significant amount, normally regarded as 5-10 per cent. In essence, the practical application of the test proceeds by seeking to identify all of those products whose existence materially constrains the price of the product under investigation from profitably being increased by 5-10 per cent. These may be all of the products to which customers would turn in significant numbers were the price of the product of interest to rise by such an amount (*demand-side substitutes*). It may also include all of the products whose suppliers could readily commence production of the product of interest and secure meaningful sales of that product without the need to invest significantly in any new productive assets, including advertising, distribution and so on (*supply-side substitutes*).⁴

3.2.2. Market definition in non-merger contexts

The basic concept of a market as consisting of those products that are demand and supply-side substitutes is valid in all cases. However, the standard market definition test was developed in the context of merger analysis and a number of serious difficulties arise in seeking to give it practical application in many non-merger contexts.

Specifically, the difficulties with implementing the test in many non-merger inquiries arise from the benchmark against which to apply the hypothetical price increase. In merger cases, the competitive concern is whether the merger will result in prices increasing from their current level and the relevant issue is whether the merger will make matters worse.

⁴ A more detailed discussion of demand and supply-side substitution, as well as a more complete explanation of the market definition framework, can be found in the European Commission *Notice on the Definition of the Relevant Market for the Purposes of Community Competition Law*, OJ 1997 C372/03. Other relevant discussions of market definition can be found in the Department of Justice and Federal Trade Commission Horizontal Merger Guidelines, April 2, 1992, as well as in OFT 403, Market Definition, Office of Fair Trading, March 1999.

Therefore for the purposes of a merger case what one is interested in is the identification of competitive constraints at prevailing prices. Market data can frequently provide useful information on the identity of those products that are currently demand and supply-side substitutes for the products of the merging parties. Such data can consequently generate a market definition that has great value in understanding how the proposed merger might change the existing competitive constraints faced by the merged firms.

But in non-merger cases, it is no longer the case that prevailing prices necessarily provide the appropriate benchmark against which to assess the profitability of a further price rise. The problem arises because profit-maximising firms will always set their prices at a level at which a further price increase would be unprofitable.⁵ In other words, the degree of substitution between two products depends in large part upon the current relative prices of the products concerned. Consequently, the mere fact that at its current price an allegedly monopolised product has significant substitutes does not prove that the firm producing the product does not already possess significant market power.

The alleged monopolist may only face demand substitutes at existing prices because it has already elevated its prices above the competitive level to the point at which those products become substitutes. Or, as was noted by the US courts, if the price of the monopolised product is high enough even inferior substitutes will begin to look attractive to consumers. In other words, the identification of substitute products at existing prices does not mean that those products would necessarily have been substitutes for the products under investigation at the competitive price. The identification of substitutes at prevailing prices does not mean that a firm does not possess market power or a dominant position. It may simply be that the significant market power that it possesses is already being fully exercised through the setting of excessive prices.

This potential problem is known in competition policy analysis as the *cellophane fallacy* after the celebrated *Du Pont* case. In that case, *Du Pont* argued that cellophane was not a separate relevant market since it competed directly and closely with flexible packaging materials such as aluminium foil, wax paper and polyethylene. But, as many commentators have since noted, these products can only be regarded as providing an effective competitive constraint on cellophane if the price at which they are identified as substitutes is the competitive price. The US Supreme Court failed to recognise that a high own-price elasticity and the identification of effective substitutes at prevailing prices may merely be the result of the exercise of existing market power.

A simple illustration using rail travel as an example may clarify the point. Suppose, for simplicity, that there is a only one fare type, and that there is on-track competition between two different train operators between towns A and B which has led to both train operators setting a price of £30 return on the route. Further suppose that there is a coach service in

⁵ Technically, prices will be set at a point at which demand for the product is elastic.

operation between A and B with return prices at £15 and a budget airline in operation between airports close to A and B charging £40. If a merger between the two train operators were contemplated, it would be appropriate to consider whether the relevant market for the analysis of that merger was that for train services between A and B or whether it was a broader market that included airlines and coaches too.

In considering this issue one would be very interested in the degree to which a price increase by one train company led to higher sales by the other train company and to what extent those price increases might lead to higher sales by the coach and airline companies. If the evidence suggested that when one train company raised its price most lost sales went to the other train company then it would be appropriate to define the market as that for train travel between A and B. On that basis the two train companies would have, say, 50 per cent each and the merger would create a monopoly of the train market. Coaches and airlines should not properly be considered as part of the relevant market because they do not constrain one train company's price as tightly as the other train company does.

Nonetheless, suppose that the merger was permitted and the train companies merged to create a single train monopolist. Moreover, suppose that following the elimination of all train-train competition train prices were raised from £30 return to £35 return. In other words, suppose the merger permitted the train monopoly to raise prices by £5 (17 per cent) above the competitive level. This merger would then have created a dominant firm, in the sense that it is a firm that is no longer constrained to price at the competitive level, but is free profitably to set a significantly higher price than the competitive level.⁶

Suppose that the merged train company were now to be investigated for an alleged abuse of a dominant position. One can again attempt to use market definition as a tool to aid in the analysis. So, with the single train company charging £35 return, one can now ask whether a, say, 10 per cent (£3.50) price rise would be profitable for all train services. In other words, what are the consequences of raising train prices from £35 to £38.50? In this situation the train company may well be able to show that if prices were raised in this way significant numbers of customers would switch to coach travel and significant numbers would switch to the airline. This would appear to justify a broad market definition of train, coach and airline travel between A and B.

If, on the basis of passenger miles, the train would have 35 per cent, the airline 35 per cent and the coach 30 per cent, one would be tempted to conclude that, with just 35 per cent, the train company did not have market power. However, we know (by construction of the example), that the train company has market power because we know that the competitive price is £30 and it is already charging £35 - significantly more than the competitive price level. Coaches and airlines are only substitutes in this situation because the rail monopoly is already being fully exploited.

⁶ Where the competitive level is defined as the level that prevailed when there was on-track competition.

In this example, we know that market power exists because we know that current prices (£35) are already above competitive prices (£30). Hence we know that the substitution that we observe between trains and other modes is subject to the cellophane fallacy, other modes are only substitutes at current prices because an existing monopoly is being exploited and that consequently we should not define the market broadly. However, in most real-world cases the competitive price level is unknown. In these cases, when substitution is observed we do not know whether it is substitution at the competitive price level (meaning that those substitutes should properly be included in the market) or whether it is substitution at an excessive price level (meaning that they should not be included for the purposes of identifying dominance).⁷

In consequence, mindful of the cellophane fallacy, if the train company came with evidence of inter-modal substitution at current prices, one would be forced to conclude that such evidence neither proved nor disproved the validity of the broad market definition.

3.2.3. “Solutions” to the cellophane fallacy

The cellophane fallacy is clearly a serious problem for market definition in non-merger cases, because it means that evidence of substitution between products at current prices has no value in discriminating between narrow and broad market definitions. The seriousness of this problem has led to several attempts to avoid or minimise the problems that the cellophane fallacy causes for market definition.

One possible approach is not to apply the test on the basis of price increases from the current price level, but rather to consider the profitability of a price increase from the competitive level. In the example above, rather than ask if the train monopolist could raise prices above the prevailing level - which it cannot due to competition from other modes - we should ask if it could price above the competitive level - which we know (by construction) it can. Thus, in this example, using the competitive price level rather than the prevailing level to undertake the market definition test, one would correctly identify the narrower of the two markets as the appropriate one to adopt (i.e. trains). In this market the single company's share of the train market is 100 per cent and consequently one would correctly conclude that it was dominant.

However, whilst the use of the competitive level as the appropriate benchmark has some conceptual attractions, undertaking market definition in this way is either impossible or pointless. First, in the example above, we know that the competitive price is £30 and

⁷ It should, of course, be noted that even in the world in which rail fares were already set at excessive levels the relevant market definition for consideration of a rail-coach merger would be the broad market. This is because although the train market has already been monopolised and price raised to the level of the next constraint, that constraint may now be coach travel and the elimination of that constraint may permit yet greater price increases. In other words, just because a narrow market is justified for considering the allegedly abusive actions of a train monopolist in a narrow train market, does not mean that a broader definition should not be adopted for a merger case.

therefore we know that the monopoly train company has the ability to raise prices above this level because it is currently charging £35. However, in almost all real-world cases the competitive price is unobservable, even if a satisfactory definition of the competitive price could be determined even in theory. Thus, in most cases estimation of the competitive price level is not feasible. Second, if calculation of the competitive price were feasible, the entire exercise of market definition would be pointless since one could simply compare the prevailing price to the estimated competitive price. If the former significantly exceeded the latter, one could **directly** conclude that a position of dominance existed without the need to go through the **indirect** analytical step of market definition and market share calculation, which would be wholly irrelevant, other than to satisfy the legal requirement to define a market.

A second argument as regards the cellophane fallacy is that it does not affect all non-merger cases and that in some abuse cases the standard approach to market definition remains valid. Werden eloquently sets out the argument in some detail in a recent article. In it he states:

“When a firm has already raised the price substantially above competitive levels, it makes no sense to determine whether that firm is a monopoly by asking if it could increase its profits through further price increases. Competition of some sort constrains even a monopolist’s prices. Hence, in monopoly cases, relevant markets should not include products or areas that are reasonably close substitutes only because price is already substantially above the competitive level. This is the lesson of the Cellophane case. Nevertheless, the Guidelines’ hypothetical monopolist paradigm can play several useful roles in many Section 2 cases.

The Guidelines’ approach to market definition is useful in some Section 2 cases because they present precisely the same market-power questions as does the typical merger case. In these cases, the Guidelines’ approach, complete with the use of prevailing prices as a benchmark, works for the same reasons it does in merger cases. The Cellophane fallacy simply does not arise.

In most merger cases, the issue is whether a proposed merger would create or enhance market power. The prevailing price is the logical benchmark for market delineation because the question posed by Section 7 of the Clayton Act is whether the merger would lessen competition – not how much competition there currently is. Similarly, the issue in many cases arising under Section 2 of the Sherman Act is whether ongoing or threatened conduct, if left unchecked, would create monopoly power – not whether the defendant already possesses monopoly power.

Section 2 plaintiffs commonly allege that a rival has (recently) embarked on a course of conduct that constitutes an unlawful “attempt to monopolize” because there is a “dangerous probability” that the conduct, if not enjoined, would create monopoly power. Pleading the case this way raises an issue much like that in a merger case:

Assuming the conduct has the alleged exclusionary effect, would the likely result be the creation of monopoly power.”⁸

However, it will be noted that Werden’s approach does not solve the problem as it typically arises in a European law context. In these cases market definition is not typically used to evaluate the likely consequences of the alleged abuse, but rather it is used to try to test for the pre-existence of market power – exactly the role Werden claims that it cannot play.

His approach is to treat certain prospective Section 2 breaches of the Sherman Act like mergers because the concern in the law in the US in these cases is not whether market power exists, but whether it would be created by the exclusionary conduct. There is an analogue to this in European law, which is the analysis of the effects of the (exclusionary) abuse on the market. In analysing the consequences of an allegedly exclusionary abuse on competition it may well make sense to undertake an analysis similar to that suggested by Werden to see if the exclusion of the competitor would be likely to create a situation in which a firm’s market power may be enhanced and prices be increased from the current level.

Unfortunately, that is not how market definition has tended to be used in European law. The main purpose of market definition has been to permit the calculation of market shares with the express purpose of trying to identify whether a firm already holds a position of dominance. The kind of market definition exercise needed to identify pre-existing dominance is very different to the market definition exercise that Werden is suggesting for analysing the competitive effects of prospective exclusionary abuses. It has no equivalent role in purely exploitative abuse in the form of excessive pricing.

In short, market definition applied in the traditional way identifies those firms that are effective competitors at prevailing prices. This is very valuable for analysing the effects of future changes to market structure through merger or exclusion. *However, the identification of the current competitive constraints on a firm is of no value in identifying whether that firm already holds a position of market power and dominance.*

3.2.4. A defensible role for market definition

In those circumstances where the objective of the analysis is to try to establish whether a firm has pre-existing market power and dominance, as is the case when dominance is being established in the first step of an abuse of dominance case, market evidence on the extent of substitution between products at current prices is not helpful. In these cases – given the inability to discriminate between rival market definitions on the basis of market evidence – the role of market definition is more limited and will normally be confined to ensuring that all hypothesised market definitions are consistent with the basic concepts of what constitutes a market (i.e. a set of products that are all either demand or supply-side

⁸ G. Werden, “Market delineation under the merger guidelines: monopoly cases and alternative approaches”, *Review of Industrial Organisation*, 16: 211-218, 2000

substitutes for other products in that set). It may also be worthwhile to engage in an analysis of product characteristics that may give some insight into the innate substitutability of two products without recourse to market evidence tainted by the cellophane fallacy.

However, it may often be possible to identify two or more market definitions that are both consistent with the basic market concepts of supply and demand-side substitution and find support from a detailed analysis of the characteristics of the product. The result may be that the market definition stage of the analysis provides very little substantive assistance in the identification of dominance (although it may do so in the analysis of the effects of the abuse, if the abuse is an exclusionary one, as described by Werden). If this is the case, as it may well be in many non-merger cases, market definition may simply be a step that it is necessary to undertake after the analysis of the abuse in order to fulfil the legal requirement to define a market. In support of the market definition which one uses to meet that legal requirement it may be possible to use evidence on the identified effects of the abuse to infer dominance and give support to the market definition which supports that dominance finding.⁹

3.3. Direct Evidence of Dominance

Of course, merely because in many cases the market definition stage of the analysis is rendered impossible or meaningless by cellophane fallacy considerations, does not mean that there is no scope for additional work to help to establish dominance. First, there may be evidence from prior abuses that would suggest that the firm has a dominant position.

Unusually, in the case of rail franchisees there may be yet more direct evidence of the existence of a dominant position from the nature of the franchise process and the limitations on entry that protect many franchisees. Specifically, where loss-making routes have been bundled with other routes to create a financially viable franchise package it is implicit that the operator will be dominant on at least some of those routes – in the sense of having the ability to set prices that are above the competitive level and so leads to super-normal profits – in order to be able to finance the loss-making routes.

In a 1998 book chapter, Swift explicitly acknowledges that the packaging of the initial franchises means that there are cross-subsidies within the system minimising total subsidy payments.¹⁰ This indicates that it was known by OPRAF at the moment of franchising that some fare types on some routes would generate excessive profits in order for the franchisee to be able to fund the loss-making elements of the franchise package. Paradoxically therefore the identification of loss-making routes within a franchise may, if the total subsidy paid to the franchisee is less than the capitalised value of those losses, actually provide

⁹ Of course, this is a purely legalistic step, since once the analysis of the effects of the abuse is undertaken and those effects are determined then it is substantively no longer necessary to revisit the market definition.

¹⁰ John Swift (1998), “Competition in the rail industry”, in *Competition in Regulated Industries*, edited by Dieter Helm and Tim Jenkinson, Oxford University Press.

strong evidence that the franchisee has the ability to set prices that are in excess of the competitive level on certain other fare types for certain other routes within the package. In other words, unless the losses on the loss-making part of a franchise are fully funded through a subsidy payment it is likely that the franchisee will be enjoying significant market power on at least some of the other routes in the franchise package.

This kind of direct evidence is particularly compelling where the franchisee has been prepared to pay a very significant positive sum for the franchise. In effect, this payment represents the price to be paid for the acquisition of market power and, in an effectively functioning franchise system, would be the capitalised value of the expected excess returns to be had from the franchise over its lifetime. In other words, if one defines the competitive outcome as that which would emerge in a situation of free entry, any payments above zero that a franchise is willing to make to operate the franchise must be the value of the net revenues that it anticipates earning on the package of routes after payment of all costs associated with running the franchise.

Evidence of this kind can provide better and more direct evidence of market power than the indirect and unreliable evidence provided by market definition and market share analysis. This line of reasoning would suggest that careful examination of the composition of the franchise and the magnitude of the bid is warranted in order to ascertain the likely excess revenues that have been anticipated for the non-loss-making routes. This will be the sum of the losses that have been estimated for the loss-making components of the franchise and the size of the payment or subsidy made by the franchisee. Franchisees must have been expecting to make significant excess revenues from their non-loss-making routes if they:

- incur significant losses on certain loss-making elements of their franchise; and
- were nonetheless prepared to pay a significant sum for the franchise package as a whole.¹¹

In addition, contractual guarantees against the entry of on-track competitors in some cases mean that barriers to entry into the relevant market are complete. This would further support the idea that the franchise winner may have acquired a dominant position through the franchise process since there can be no entry mechanism in place to erode the excess returns that may be made on selected routes.

Specifically, Swift explicitly recognises in the same chapter that

“Some operators may require protection from competition in the longer term in order to provide sufficient certainty to support investment in network enhancements or improved rolling-stock. For example, I have agreed that the main flows from London

¹¹ Even where bidders believe that they could make efficiency savings, they would still have to expect to have the market power to prevent competition from bidding these savings away.

to stations on the West Coast Main Line should be protected from competition for the whole of the 15-year franchise period, provided the planned upgrade in services – involving both infrastructure enhancement and investment in a new fleet of tilting trains – is taken forward.”¹²

This provides further direct proof of the dominant positions that certain franchisees hold on certain routes. As the quote above makes clear, contractual protection was offered to some operators to “...provide sufficient certainty...” to support required investments. But, this must mean that in the absence of protection there must have been a risk that competitive entry would have occurred and reduced returns to be made from the franchise. In other words, only through the prevention of competition could the franchisee be certain that the expected revenues would be high enough to fund investments. However, if competitive entry would have reduced expected revenues to levels below those that would otherwise have prevailed it must be the case that the actual expected revenues from the franchise were higher than the competitive level of expected profits. In this case, it is absolutely clear that OPRAF was selling monopoly rights and that the franchisee was therefore buying market power on the protected routes – if it were not there would have been no need to protect them.

In summary, it seems to us that given the profound complications inherent in any attempt to define a market for substantive purposes, as opposed to legal purposes, the substantive analysis can, in this case, proceed by analysis of the direct evidence of the existence of significant and enduring market power. This evidence may consist of:

- the bundling of loss-making routes with excessively profitable routes to minimise subsidy payments;
- the bidding of large positive sums for some of the franchises, even though they are contain some significantly loss-making elements; and
- contractual protection from competition, which not only creates absolute barriers to entry into the activity, but also suggests that both OPRAF and the franchisee anticipated that expected profits with entry (i.e. in the presence of competition) would be significantly lower than profits without entry (i.e. in the absence of competition).

Of course, the potential existence of direct evidence of dominance does not mean that market definition can be avoided altogether for the legal reasons outlined earlier. However, the same evidence of dominance can just as well be used to substantiate the existence of the narrow market definition(s) necessary to support a finding on dominance on the usual basis of high share within a narrow market. However, it should at all times be recognised that market definition undertaken in these circumstances does not provide an independent

¹² Ibid p.250.

analytical path and that market definition is being inferred from the existence of dominance, and not the other way round. Market definition is being undertaken solely to go through the legal hoop that is required by *Continental Can*.

3.4. Market Definition in Train Operation

This section considers the definition of the relevant markets within which TOCs may operate, focusing in particular on how those different markets can be reconciled with the basic concepts of demand and supply-side substitution that underpin the standard market definition test.

As we have discussed, much of the empirical evidence on the extent of substitution at existing prices will not be relevant to the definition of the market for the purposes of Chapter II because of cellophane fallacy concerns. However, ORR has nonetheless expressed an interest in understanding the competitive constraints that exist on TOC pricing at current price levels. This information may well have considerable value in market definition in intra-modal (e.g. TOC/TOC) and inter-modal (e.g. TOC/coach) mergers, as well as in identifying the impacts of various possible exclusionary practices, and a detailed review of the empirical evidence in this area is contained in Annex B.

The most basic issue to address in considering market definition in this area is whether the operation of a TOC on a particular route constitutes operation in one relevant market or more than one relevant market. In order to answer this question it is necessary to recall that markets must be defined by reference to the degree and identity of demand and supply-side substitutes. The basic building block of market definition must be the different fare types that a TOC offers. The issue then becomes what are likely to be the competitive constraints that influence the setting of the level of prices for each fare type.

Traditionally there have been a number of different types of fares offered on a particular route. This include:

- full first class fares that do not need to be booked in advance and are available on all trains for travel in first class accommodation;
- full standard class fares that do not need to be booked in advance and are available on all trains for travel in standard class accommodation;
- saver and supersaver fares that do not need to be booked in advance, but are only available on off-peak trains with conditions of availability that vary by route;¹³

¹³ Supersaver fares were not available on Fridays. In some cases supersaver and saver tickets have now been combined so that only a saver ticket is available.

- various forms of advance purchase tickets with lower fares that are generally only available for specified outward and return services and are not available on other trains.

These different fares might be said to serve different markets, in which different competitive constraints apply. Thus on a particular route, i.e. in a particular geographic market, the lower price tickets may serve a market for leisure travel and be constrained by competition from coach fares. Full fare tickets may primarily serve a market for business travel in that market, and be constrained by air fares in that market. In a merger investigation, one would want to consider how the merger would affect the ability to raise each of these different types of fare. In a dominance case, the relevant market is more difficult to determine, since the relevant issue is whether the firm is charging fares above the competitive level - however, formulation of the competitive benchmark must take into account the existence of common costs and the need to allocate them between different fare classes.

4. ANALYSING ALLEGATIONS OF EXCESSIVE PRICING

At the direction of the ORR this report primarily considers the issues that would be raised in any investigation into allegations of excessive pricing by a TOC. However since price discrimination plays a central role in the pricing policies of all TOCs it is impossible to analyse allegations of excessive pricing without understanding the motivation for and the effects of price discrimination across the various markets served by a TOC. This section therefore considers pure excessive pricing issues and price discrimination, in so far as it may be a device for allowing a TOC to charge excessive prices to one or more classes of passenger.

In addition to exploitative abuses, the most obvious of which is the setting of excessive prices, the European Commission and the courts have additionally identified a range of other potential abuses, the aim of which is to exclude competitors from the market.

4.1. Precedent and Guidance on Excessive Pricing

In *United Brands* the European Court of Justice stated that prices would be considered excessive where the profit margin earned by the dominant firm was excessive in comparison to the costs of providing the good or service.¹⁴ Specifically, it stated that prices would be considered excessive where they do not bear a:

“...reasonable relation to the economic value of the product supplied.”

This formulation of the test for excessiveness appears to place the emphasis on the relationship between the prices charged for a product and the costs of supplying that product. This view is reflected in recent draft guidance issued by ORR where, in relation to the application of the Chapter II prohibition to the rail industry, it argues that the:

“...Regulator will consider the relationship between prices and costs when deciding whether prices are excessive...”

However, application of a cost-based criterion for the determination of excessive pricing raises a number of complex issues. The most obvious issue raised by the application of such a criterion is the identification of what the appropriate relationship between costs and prices should be in any particular instance. Inevitably therefore one must seek to ascertain what relationship between costs and prices might be expected to prevail in a competitive environment and use this as the benchmark against which to judge the relationship observed between the costs and prices of the dominant firm. This is reflected in the statement that:

¹⁴ Case 27/76 *United Brands v Commission* [1978] E.C.L.R. 207; [1978] 1 C.M.L.R 429

“...the Regulator may consider that a dominant position has been abused by the introduction or maintenance of prices that led to profits being sustained at a higher level than in a competitive market.”

The Office of Fair Trading (OFT) Competition Act guidance states that:

“An undertaking’s prices in a particular market can be regarded as excessive if they allow the undertaking to sustain profits higher than it could normally expect to earn in a competitive market. ... All undertakings clearly need to earn some level of profits in order to remain in business and to provide a sufficient return to shareholders and lenders, on whom they depend for capital. The return required by shareholders (the equity cost of capital) is at least the return they could earn elsewhere, having regard to the relative risks incurred by investing in the particular undertaking. Prices within a particular market can be excessive if they allow shareholders to yield a rate of return that is significantly and persistently greater than this equity cost of capital. Similarly, lenders will have their own cost of capital and, by analogy, prices can be excessive if lenders’ returns persistently exceed this level. ...”¹⁵

Guidance appears to suggest two approaches to the evaluation of excess pricing allegations. These are to:

- compare prices with costs; and
- compare profits with those that would be made in a competitive environment.

We believe that given the particular characteristics of the railway industry, both approaches present considerable difficulties. This is because the nature of the cost conditions in railway operation means that it is difficult to determine the relationship that any one price (i.e. fare type) would bear to costs in a competitive market. This in turn means that the first approach to the issue suggested by the guidance – comparison of prices with costs – cannot meaningfully be undertaken. It only makes sense to consider excessive pricing in the overall context of a set of prices that in aggregate generate excessive profits in consideration of the aggregate costs of engaging in the activity of train operation.

However, the direct analysis of profits is also rendered extremely difficult by the importance of risk in the assessment of profitability and the existence of a bidding process in which all expected excess profits would have been extracted. All of these issues are discussed below, beginning with a discussion of the implications of common costs.

4.2. Implications of Common Costs

A major complication in the rail industry is that a large proportion of the total costs of operating trains on a route cannot be directly attributed to a particular fare type because

¹⁵ Office of Fair Trading, *Assessment of Individual Agreements and Conduct*, September 1999 (paragraph 2.6).

there are typically few costs that could be avoided either in the short run or the long run by the withdrawal of that fare type. This means that there is a large cost that cannot be attributed to any single fare type, cannot be avoided by withdrawal of any single fare type, but that must be recovered from the overall operation of the service through the setting of prices across all fares that in aggregate will cover those common costs.

Since it is usually the case that very few costs can be avoided by not carrying one extra passenger, the TOC will always be better-off carrying one extra passenger at a low fare than not carrying that passenger. This is because the extra passenger imposes almost no extra costs on the system as a whole, but the revenues from the sale will help contribute to the overall recovery of the common costs of the system. Thus, it will always be worthwhile introducing a new lower fare type to attract extra customers, who are not prepared to travel at current prices, as long as the sale of tickets to those customers does not reduce sales of other higher ticket types because of:

- capacity constraints; or
- trading down to the lower fare type by customers who would otherwise have paid a higher fare type.

These realities mean that all rail operators, in common with almost all transport operations (e.g. airlines, coaches, ferries, etc.), engage in price discrimination. This manifests itself in some customers making a greater contribution to common costs than some others. Nonetheless, this mode of behaviour is efficient because without the ability to make sales at lower prices to more price-sensitive customers, the prices to all other customers would have to be higher, for any given level of costs, if all costs are to be covered.

The economic theory suggest that almost all firms in imperfectly competitive markets – almost all real markets - will wish to engage in price discrimination and may well be able to do so to some degree. Technically, firms will seek to charge different prices to different customers according to the different degrees of elasticity in demand in various segments of their markets. Price discrimination of this kind is used by suppliers ranging from law firms to cement producers, theatre ticket agencies to record companies, and is certainly not confined to dominant firms. Moreover, the same theory suggests that such practices are not necessarily detrimental to competition or overall economic welfare.

As we have seen, by permitting firms to price discriminate, some consumers will be able to make purchases that they would not otherwise have made if the supplying firm were constrained to charging a flat price and the firm will earn revenues that it would not otherwise have earned, enabling it to set lower prices to other customers. In its guidance the OFT argues that “...in general, price discrimination will not be an abuse if it leads to a

higher level of output in the relevant market(s) than could be achieved if all customers were charged the same price.”¹⁶.

Nonetheless, there are two basic concerns according to both the OFT and the ORR when a dominant TOC price discriminates. These are that it might have exclusionary intent or that it might be a device to charge excessive fares. For example, the ORR states “...the charging of discriminatory prices will be considered to be an abuse where there is evidence that it is a device to charge excessive prices, or where it is used to reduce competition significantly”.

In this report, we are not primarily concerned with exclusion. In any event, for those TOCs that have been awarded protection from on-track competition for particular flows and for a particular period exclusion of potential on-track competitors is not an issue. Under the current regulatory framework exclusionary behaviour for those TOCs given on-track protection can relate solely to the possible exclusion of indirect rail competitors and non-rail competitors, such as competing coach services.

Alternatively, as noted by both the OFT and ORR, price discrimination might be abusive if it allows a TOC to exploit its market power by charging excessive prices to a captive group of customers. However, whether these prices are excessive cannot be judged against a cost standard since we know that the different fare types will each bear a different relationship to costs because price discrimination is inherently a demand-down rather than a cost-up pricing philosophy. Whether a TOC is charging excessive prices on one or more of its fare types can only be determined by reference to the overall level of profits for the activity as a whole. Moreover, given that pricing is undertaken on the basis of different customers making different contributions to costs that in aggregate will aim to cover all costs and generate a profit, it is not clear that any one fare type can be labelled excessive and therefore solely responsible for any excessive profits that are ultimately identified. No single fare type would alone generate sufficient revenue to cover all costs and alone generate excessive profits. Since costs are common, it only makes sense to consider all revenues in aggregate and to make judgements about the overall level of revenues.

For example, for simplicity suppose that there are just two fare types – first class and second class, both of which are unregulated. Suppose that there are costs that are completely common to the operation of both first and second-class services that amount to 800 each period and marginal costs that are equal to 5 per first class passenger and 4 for each second-class passenger. Moreover, suppose that the TOC charges 20 for first class tickets and 10 for second-class tickets. It therefore makes a contribution to common costs and profits of 15 for each first class passenger, but just 6 from each second-class passenger. Suppose that at these prices it sells 20 first-class tickets and 100 second-class tickets.

On this basis the TOC makes a 300 per cent mark-up on first class passengers and a 150 per cent mark-up on second-class. The net contribution of all first-class passengers is 300 (i.e. 20

¹⁶ Ibid, paragraph 3.9

passengers x 15 contribution) and that of second-class passengers is 600 (i.e. 100 passengers x 6 contribution). Total contribution is therefore 900 and with common costs of just 800 the TOC makes a profit of 100. If these profits were regarded as excessive, which fare class could be said to be the excessive prices contributing to the excessive profits?

The first-class fare has a higher mark-up, but this would be expected even in a situation of non-excessive profits. The second-class passengers make a greater overall contribution, 600 to 300, but this is because there are five times more of them than first-class passengers. Moreover, neither fare class is covering the common costs of the TOC in isolation and therefore neither class would even be normally profitable without the contribution of the other. In our view, in this situation it makes no sense to say that either price is individually excessive, rather all that can be said is that taken together the first and second-class fares are excessive because they together generate excessive profit.

Of course, even if there were no excessive profits, such that price discrimination across the combination of all fare types generated only sufficient revenues to cover costs and normal profits, the Regulator might wish to intervene by choosing a different fare structure that was equivalent in terms of profits. The Regulator could decide how the operator's common costs should be recovered from groups of customers and judge charges accordingly, either for individual routes or for individual operators. The problem is that the Regulator is likely to have substantially less information than the franchisee and the consequent fare structure is likely to be less efficient than that set by the TOC.¹⁷

The economic literature has recognised that one of the main advantages of franchising is that it leaves flexibility to the private sector operator in between franchise auctions to operate independently subject only to the limits explicitly written in the franchise contract. For instance, Cave (1995) has pointed out that "a supposed strength of franchising lies in its allowing private sector providers to be the judges of consumer and market preferences. Insofar as service specification involves the making of judgements by franchise authorities, and insofar as the franchise authority selects the best menu of services for the consumer, this advantage of franchising diminishes and franchising approximates to a system of regulation". It seems therefore difficult for the Regulator to mandate change to TOCs fare structures. A better approach would be to base intervention on grounds of excessive prices solely on the basis of the identification of excessive profits and, in circumstances where price discrimination is the standard method of pricing the services, to request that the overall fare structure is adjusted by the TOC to bring profits back into line with those deemed normal by the Regulator.

However, even this approach requires that the Regulator be able to identify excessive profits. This is not simple, particularly in the context of the regulatory framework of railway

¹⁷ Armstrong and Vickers discuss the issue of regulating the structure of prices in a regulated firm, given that the firm will have better information about its own costs than will the regulator. M. Armstrong and J. Vickers "Multiproduct price regulation under asymmetric information" *Journal of Industrial Economics* XLVIII, June 2000, 137-160.

operation in the UK. The complexities of identifying excessive profits in general and in the specific circumstances of TOC operation are discussed in Section 4.3 below.

4.3. Estimating a Normal Rate of Return for a Trading Business

4.3.1. Assessing normal rates of return

Assessments of company profitability by competition policy agencies and regulators normally focus on the return on capital employed (ROCE), and compare the achieved rate of return with a “normal” or reasonable rate of return for the activities concerned. The achieved rate of return is typically measured as the profits before interest and taxation (PBIT) divided by the sum of the depreciated fixed assets plus net current assets employed in the business. Evidence that actual returns had been persistently above normal or reasonable returns, to an extent not explained by factors such as superior efficiency or innovatory performance, would constitute prima facie evidence of abusive or monopolistic behaviour.

In principle, the normal rate of return should be determined by reference to the weighted average cost of capital (WACC) for the activities in question, and much time and effort has been expended both by industry regulators and regulated businesses in estimating the WACC. Despite these efforts, the methodology used to derive the equity component of the WACC remains controversial. In the absence of robust estimates of equity betas for segmental activities, the UK competition policy agencies have tended to follow a more pragmatic approach to deriving a benchmark rate of return, based on estimates of achieved sectoral or even economy wide rates of return. Possibly in recognition of the difficulties attending measurement of achieved rates of return in reference activities, and of the lack of robustness of the usual benchmark measures, competition authorities have generally only been willing to identify a wide and persistent gap between achieved and benchmark rates as evidence of abusive behaviour.

This methodology cannot be readily applied to "trading businesses" with few or no requirements to finance investment in specific or illiquid assets such as plant, machinery and buildings. "Trading businesses" may have some requirements for working capital, but these might be small relative to turnover; cost structures are dominated by labour and material inputs, with any capital inputs usually purchased through leasing arrangements or via access charges levied by a network provider.

Such trading businesses, nonetheless, expose their backers to risks. However, the question of what reward is sufficient to justify the risks taken on cannot be directly addressed by the usual technique. Instead, regulators must assess whether the achieved profit margin, equal to the difference between revenues and operating expenses, is excessive given the risks of the business.

A TOC is well characterised by the concept of a trading business. Working capital is typically negative and exceeds the value of fixed assets. Although larger capital commitments are being taken on, these will generally take the form of providing a revenue stream for specific projects which others - Railtrack, train leasing companies - will fund. Should TOCs start to take on capital-intensive operations, then the rewards to such activities should be evaluated separately, using a conventional cost of capital approach.

The problem of determining a reasonable operating margin for trading businesses has been faced by UK regulators, and the following section (**Error! Reference source not found.**) examines the precedents offered by regulatory decisions in respect of the gas and electricity supply businesses (as opposed to transmission) with similar characteristics to those of rail and franchise operations. Centrica, the "British Gas" supply business formerly called British Gas Trading (BGT), is of particular interest because of the risks it holds. For example, its income is closely related to winter temperatures, and it holds fixed price "take or pay" gas supply contracts which expose it to fluctuations in the market price of gas. This section also draws conclusions from the review of existing regulatory practice for the measurement of TOC profitability.

4.3.2. Precedents in gas and electricity supply

4.3.2.1. Introduction

In the most recent gas supply¹⁸ (November 1996) and electricity supply¹⁹ (October 1997) price reviews, the relevant regulators (Ofgas and Offer) each sought to determine a reasonable operating margin for what we have characterised as trading businesses.

The reviews covered the supply of energy to small users, which we can colloquially call the domestic market. In each case, the supply companies purchase energy and distribution services wholesale. Their main physical assets are limited to those required to supply marketing, billing and administrative services. Their balance sheets are dominated by working capital, since energy is usually supplied on credit. Moreover the working capital can be highly seasonal, reflecting seasonality in energy use.

In each case, the regulator decided that a target margin on sales of 1.5 per cent was adequate to reward the taking of certain risks. It is important to understand that the margin is a target. The margin was provided as a reward for taking on risk, and in poor outcomes returns would be below the target. However as we shall see, there were other risks in the business which were treated differently.

¹⁸ 1997 Price Control Review, Supply at or below 2,500 therms a year - British Gas Trading, The Director General's final proposals, Ofgas, November 1996.

¹⁹ The Competitive Electricity Market from 1998: Price Restraints, Proposals, Offer, October 1997.

4.3.2.2. Gas supply

Ofgas' general approach to British Gas Trading's (BGT) profit margin is as follows:

"Ofgas derived a profit margin using a regulatory asset value approach, supplemented to recognise turnover-related risk. 'Turnover-related risk' may be thought of as a risk which is caused by BGT's assets forming a relatively small asset base, when taken against the financial consequences of loss of market share as competition is introduced."²⁰

Competition is not the only reason that BGT has volume risk - weather is a major factor in domestic gas consumption - but one can consider such other risks as compounded with the market share risk.

The introduction of competition to final consumers also exposes BGT to risk from variations in the field price of gas. BGT's portfolio still includes contracts negotiated when BG was in the public sector, which were struck at a price well above the current field price. Although Ofgas has allowed BGT to pass through the weighted average cost of its gas purchases (rather than the lower market price), this does not protect it from price risks, as the company may be required to match new entrant prices reflecting the current field price. The less closely BGT matches new entrant prices, the more rapidly will its market share be eroded.

Ofgas' approach to setting the margin in the face of price and volume risks was through scenario analysis. It first estimated BGT's weighted average cost of capital (WACC) by use of the capital asset pricing model (CAPM) methodology. Although BGT was at this date still part of British Gas, it was apparently able to estimate an equity beta and gearing for the domestic supply business. This gave a range for the real WACC, and 9 per cent was selected.

The second stage of the analysis was to construct scenarios for BGT's market share retention. These were then analysed to compute the net present value of the cash-flow generated in the different scenarios according to the margin allowed. The results were presented as shown in Table 4.1.

It is not clear to us how the scenarios were constructed, in particular, how Ofgas addressed the non-independence of price and volume risks, or what kind of probability was attached to them. The first three scenarios were first introduced in the *initial proposals*, though the valuations had changed when reported again in the *final proposals*. The last two scenarios were introduced in the *final proposals*. The "alternative middle" was based upon experience in the South West where competition had been introduced into the domestic market as a

²⁰ 1997 Price Control Review, supply at or below 2,500 therms a year - British Gas Trading, The Director General's final proposals, Ofgas, November 1996, para 8.1a.

pilot scheme. At least the Director General did not base his decision on the dubious "average" which is shown.

Table 4.1
Ofgas' Valuation of British Gas' Domestic Supply Business

	Net present values in £m at margin of				
	0.5%	1.0%	1.5%	2.0%	2.5%
Low market share scenario	500	600	700	850	950
Middle market share scenario	700	900	1100	1300	1500
High market share scenario	900	1150	1400	1650	1900
Alternative middle scenario	700	900	1100	1300	1500
Transfer Scheme	550	700	850	1000	1150
Average	650	850	1000	1200	1400

Notes

- 1 *Figures shown to nearest E50 million*
- 2 *Low - assumes 37 per cent loss of market share by 2000 and 75 per cent by 2005*
Middle - assumes 21 per cent loss of market share by 2000 and 37 per cent by 2005
High - assumes 3 per cent loss of market share by 2000 and 10 per cent by 2005
Alternative middle - assumes 25 per cent loss of market share by 2000 and 37 per cent by 2005.
Transfer Scheme - assumes 42 per cent loss of market share by 2000 and 60 per cent by 2005

Source: 1997 Price Control Review, Supply at or below 2,500 therms a year - British Gas Trading, The Director General's final proposals, Ofgas, November 1996, para 8.1d Table 13.

The final stage of the determination was based on the observation that the regulatory asset value of BGT at the opening of the period would be £725 million, which is more or less the value of its balance sheet net assets. Ofgas chose 1.5 per cent on the basis that

"This would, under a wide variety of circumstances, ensure that the value of British Gas' domestic supply business would be close to, or above, the value of its balance sheet net assets at the start of the new price control period."²¹

On the face of it this is quite generous. Apparently, there is practically no volume-risk downside for BGT under this determination. However other elements of uncertainty have been suppressed in this analysis. For example, the WACC was determined as a range, but only the central value was used to construct the valuations.

²¹ *1997 Price Control Review, Supply at or below 2,500 therms a year - British Gas Trading, The Director General's Final proposals, Ofgas, November 1996, para 8.2.*

British Gas criticised this approach in a number of ways.²² Putting aside the familiar contention that the cost of capital allowed by Ofgas was too low, its main criticism was that the discount rate used to value the cash-flows should have been higher than the cost of capital to reflect the effect of the volume risk. Although it is difficult to judge the validity of this argument without sight of British Gas' detailed submissions, it appears to confuse the concepts of systematic and specific risks. The systematic risk in BGT's business is captured in the estimated WACC. Investors are, in principle, able to diversify away from the specific risks in BGT's business arising from uncertainty as to the rate of market share erosion.

A more serious potential shortcoming of the Ofgas approach is the fact that not all of the outcomes illustrated are equally probable. Ofgas has in the past measured the success of the introduction of competition into a market by the extent to which BGT lost market share, indeed setting targets for market share loss. Further, it promoted loss of market share through artificial regulatory restraints on BGT which created cream-skimming opportunities for new entrants. It would therefore appear that Ofgas has created an environment in which outcomes better than the middle scenario are most unlikely, not only on the evidence of the South West pilot, but also because on past form Ofgas would take action to prevent them. In other words, one should have a realistic view of the true range of outcomes, and in the BGT case concentration on the low scenario is appropriate.

British Gas proposed several alternative methodologies for estimating a margin on sales:-

- a 'capital-equivalent' approach, in which some of the trading risks faced by BGT were capitalised in the rate base (for example, using option valuation techniques), to which an appropriate cost of capital is applied;
- a 'risk-to-reward ratio' approach, which calculated appropriate profits for BGT by applying the relationship between profit variability and average profit levels for the whole UK economy to a calculation of the risks faced by BGT. This approach is unpromising, because BGT faces very specific risks which are unlike those faced by the economy as a whole;
- comparisons with the actual return on turnover achieved by other competitive trading companies. The problem here is finding suitable comparators for which separate accounting information is available. Major retailers, for example, have a large capital base in shops and stock over which to spread risk. There are traders who do not finance stock and who have relatively few fixed assets, for example traders on financial markets, who might be comparable. However most of the traders in this sector are integrated into a larger financial institution, which defeats the point of the comparison. It is also likely that such traders survive on a very small margin, and have additional sources of profit;

²² *Response to Ofgas: The Director General's Initial Proposals*, British Gas, 1 July 1996.

- an analysis of the return allowed for the RECs (Regional Electricity Companies) in Offer's 1993 electricity supply price control review, adjusted for differences between the RECs at that time and BGT at the present time. As we note below, Offer's 1993 settlement was not intellectually well founded, and would come to appear rather generous in comparison to later settlements, e.g., the 1.5 per cent margin allowed in the 1997 public electricity supply settlement. BGT could live in hope, but it was unlikely this comparison would be taken seriously.

British Gas claimed that these approaches all supported a margin of at least 3 per cent, but without sight of the material used to justify this claim, it is hard to know to what extent it can be supported. Ofgas did not publicly comment on these alternative approaches, merely asserting the primacy of its own approach. Of the four approaches mentioned, the first appears to have the best prospect of offering more concrete conclusions. Had BGT published its analysis, it could have been most interesting to see how British Gas sought to value some of the risks it faced.

4.3.2.3. *Electricity supply*

In the electricity supply decision, two significant risks were identified, and their treatment was rather different. In this decision, Offer determined a maximum supply price for domestic electricity for the first time without permitting pass-through of actual energy purchase costs, and additionally the domestic market was being opened to competition. Public Electricity Suppliers (PESs) now faced both price risk for their electricity purchases, and volume risk in the sense of possible loss of domestic customers to competitors. As in gas, there is also volume risk due to economic conditions, weather, increasing use of energy-saving devices, etc, but this appears not to be considered.

For price risk, Offer had available to it a market test of the value of the risk. PESs can hedge their price risk from the pool by trading Contracts for Differences²³ in a market. Whilst a PES can decide to leave its price risk unhedged, in practice most choose to hedge, because with about 50 per cent of their income going on electricity purchase and a large fraction on use of system charges;²⁴ there is little margin for error in the final customer price. The Director General calculated that the weighted average premium for hedging a load curve typical of domestic supply is 6 per cent. He therefore allowed a 6 per cent uplift on average pool prices prevalent at the time as the fair cost of engaging in Contracts for Differences.

As regards the volume risk, Offer suggested that hedging was available, but it did not examine the price of it. Although no reason was given, one can perhaps understand why.

²³ Under a Contract for Differences, the contractor in effect offers to sell electricity at a fixed price schedule, but it takes the form of a hedge or insurance contract. The purchaser takes electricity from the pool at the prevailing price, and the contractor pays or receives the difference between the pool price and the contracted price.

²⁴ Although the PESs are part of the same company which provides local distribution services, the distribution service is separated for regulatory purposes, and a common tariff is offered to all.

Whilst a company might be offered a hedge for volume risk outside its control, e.g., weather-related, it seems unlikely that insurance would be offered against the risk of losing business to competitors. Offer therefore followed Ofgas in allowing a margin on gross turnover for the risk.

First, Offer sought a benchmark for a return on the capital employed in an electricity supply business.

"For Hydro-Electric and other energy supply companies where the market is not open to competition, the MMC has proposed or endorsed price controls involving a margin of 1/2 per cent on turnover to provide a return on the capital involved."²⁵

This was made to up 1.5 per cent by allowing one per cent for the volume risk arising from the introduction of competition in additive fashion, as was suggested in an earlier consultation paper:

"1 per cent for volume risk, in addition to 1/2 per cent for return on capital employed."²⁶

The value of one per cent for volume risk appears to be asserted rather than derived from analysis.²⁷ The precedent of the British Gas Trading 1.5 per cent decision is repeatedly stated. When the Director General finally states:

"Previous consultation papers concluded that a 1 1/2 per cent margin on turnover was also appropriate for PESs"²⁸

the only statement suggesting that such a conclusion had in fact been drawn is the following:

"There are certainly differences in risk between gas and electricity, but they do not all point in the same direction."²⁹

Rather, it appears that the conclusion of the consultation had in fact been the following:

²⁵ *The Competitive Electricity Market from 1998: Price Restraints, Proposals*, Offer, October 1997, para 6.19.

²⁶ *The Competitive Electricity Market from 1998: Price Restraints, Fifth Consultation*, Offer, August 1997, para 6.22. In fact Offer had previously allowed supply companies a return of 1 per cent on turnover in a situation without this volume risk due to competition, so one might say that it was reducing the element for return on capital, and thus taking away with one hand much of what it gave with the other.

²⁷ See especially *The Competitive Electricity Market from 1998: Price Restraints, Fourth Consultation*, Offer, July 1997, para 4.58.

²⁸ *The Competitive Electricity Market from 1998: Price Restraints, Proposals*, Offer, October 1997, para 6.19.

²⁹ *The Competitive Electricity Market from 1998: Price Restraints, Fifth Consultation*, Offer, August 1997, para 6.24.

"At this stage of the electricity market development it is not possible or appropriate to prescribe a particular profit margin with any degree of confidence."³⁰

It therefore appears that the decision to allow a one per cent margin on turnover as an appropriate reward for volume risk (in addition to the 0.5 per cent margin for a return on capital), had no analytical basis. On the other hand, the approach to the price risk, which takes a market estimate of the value of the risk, and builds this into the costs of the business rather than the profit margin, appears reasonable and well supported.

4.3.2.4. *Conclusions*

We draw the following conclusions from the discussion of UK regulatory precedent.

- First, and most important, **there is as yet no consensus as to an appropriate conceptual approach to assessing a reasonable operating margin for trading businesses such as TOCs.** This contrasts strongly with the position in other regulated industries, where there is at least a degree of consensus on methodology, and dispute focuses on the empirical estimates, especially of the cost of equity.
- Second, previous experience suggests that, **given the conceptual uncertainties, competition authorities will be extremely reluctant to reach conclusions on whether a particular level of operating margin constitutes excessive profitability.**

4.4. **Assessing Excessive Prices when there is Competition for the Market through Franchising**

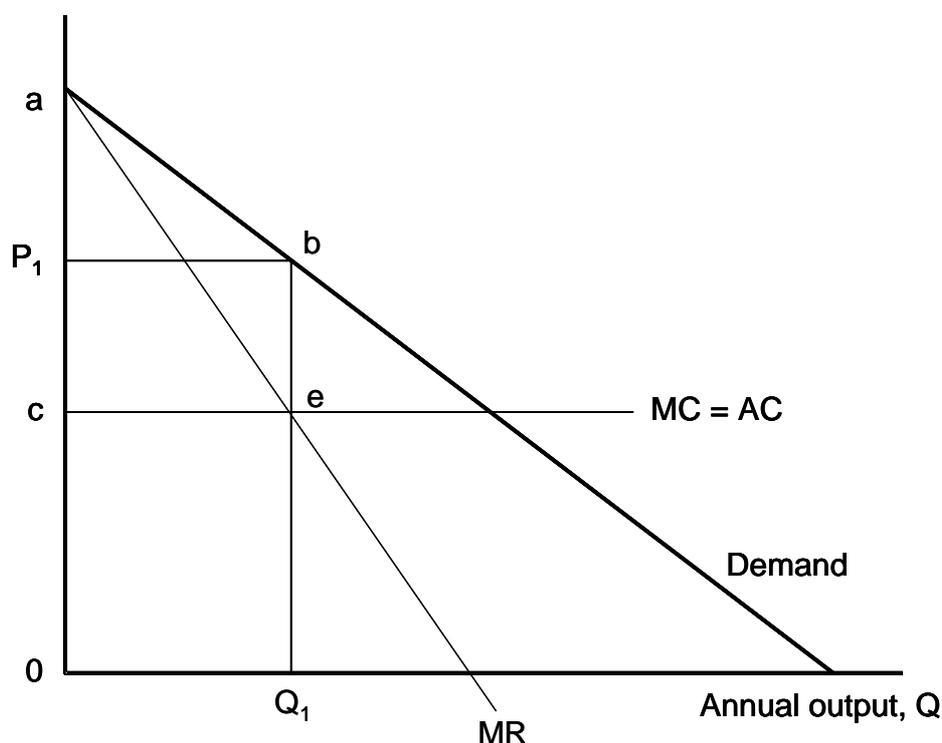
4.4.1. **The impact of franchising on excessive profits**

Franchising adds an additional complication to the issue of the determination of excessive prices and excessive profits. The nature of this complication is explained in this section.

A single price charged for a service is excessive if that price exceeds the average cost of supplying the service, where that average cost includes a normal return. This is the classic monopoly case where, as shown in Figure 4.1, price exceeds average cost (AC) at the (annual) output level at which profits are maximised, that is, at which marginal cost (MC) equals marginal revenue (MR). The monopolist earns a super-normal, or “excessive”, profit equal to area $c P_1 b$.

³⁰ Ibid para 6.28.

Figure 4.1
Excessive Price in the Simple Monopoly Case



Now suppose that the monopolist is to organise a franchise contest for the right to provide this service. If there were perfect information about future market demand and cost levels bidders would be prepared to bid a sum up to cP_1b per annum for the exclusive right to provide the service. The “excessive” profits pass from the franchisee, who sells the service at a price of P_1 per unit, to the franchiser.

This complicates the attempt to judge whether TOCs are charging excessive prices. As in the simple example described in the two previous paragraphs, the franchise payment that the TOC makes to the SRA can largely be thought of as a payment for monopoly rights, net of any cross-subsidies that must also be paid to cover losses on non-profit-making routes in the franchise. It provides direct evidence of dominance and provides a very direct estimate of the net level of excess profits that the franchise was expecting. In other words, for those franchisees that have bid a positive sum we know that there are excess profits being earned from their monopoly rights that are expected to be at least as great as the payment that was made. They may be higher if there are loss-making routes within the overall package of routes covered by the franchise. However, the franchise has paid for the right to earn such profits by means of the payment made to the rail franchising authority.

4.4.2. Using the franchise payment as a direct estimate of net excess profits

One approach to estimating the value of the excess profits would be to use the ex ante franchise payment as the best estimate for the excess profits in the system and to demand

price reductions that, over the lifetime of the franchise, would lead to a revenue loss equal to the amount bid. This might well lead to pricing during the lifetime of the franchise that was closer to the static optimal level, but is almost certain to leave the franchisee making a loss on the operation of the franchise once its payment to OPRAF/SRA is taken into account. This approach may also not be regarded as reasonable since the franchisees would surely have had the legitimate expectation of at least being given the opportunity to recoup their franchise payment over the lifetime of the franchise. Moreover, whilst in one sense regarding the franchise payment as capitalised excessive profits makes sense, society, in the form of OPRAF/SRA received the benefits of those profits via the franchise fee at the time of franchise. If society were again to seek the benefits of those profits via a regulated price reduction of equal magnitude society would gain twice and the franchisee would be a net loser. **This would represent a dynamic distortion that would fundamentally change the nature of any future franchising process.**

If franchisees see that all positive bids will be extracted via imposed price reductions at a later stage in the franchise agreement than there will be no positive bids. Franchisees will have to bid over other elements of the package such as the level of prices and service quality. This may well be a better form of franchise than the model adopted so far, but achieving it in this way would introduce a considerable amount of regulatory risk into the process and the returns that the franchisees would need in order to make a (regulatory) risk-adjusted normal return would rise, as franchisees discount into their bids the possibility of ex-post re-regulation of their prices.

Therefore the size of the positive bid made for a franchise is in some senses the best estimate of the net excess profits that the franchisee expected from the operation of the franchise.

However, simply to mandate price reductions whose overall revenue impact would be equal to the size of the franchise payment would ignore the excess profits used to subsidise loss-making routes, as well as being unfair and opportunistic. Perhaps more importantly, it would represent a fundamental change to the rules of franchising and would create serious dynamic distortions and inefficiencies that would need to be paid for later, ultimately by taxpayers and passengers through lower bids and higher fares. If the prime concern of policy makers is fare levels, then a more appropriate franchise bidding process is some type of Chadwick-Demsetz auction, in which bidders bid on fare levels subject to a fixed subsidy level.

We would therefore suggest that any assessment of excessive profitability must take into account the payment made by the franchisee for the right to operate that franchise. In order to explain the effects of the franchise payment on the analysis, it is useful initially to ignore uncertainty and risk, although these are key elements of the later analysis, and focus on the payments in a world of certainty.

4.4.3. The role of franchise payments under certainty

If it were known with certainty at the outset that a certain level of excessive profits would flow from the operation of the franchise, a competitive franchise process, with a series of equally efficient bidders, would extract all of those excess profits in the form of the franchise payment.

This can be illustrated with a simple example. Suppose a franchise worth £15 million a year is auctioned off, while the normal rate of return on the capital employed to run the franchise implies that just £10 million would be sufficient to induce the market to provide capital for the operation of the franchise. In other words, £10 million is the competitive rate required to operate the franchise given the amount of capital required.³¹ If any bidder's bid were any less than £5 million, a rival bidder would be able to obtain some supra-normal profits by increasing its bid to supply the service. This process would continue until there were no more supra-normal profits left to be won by lowering a bid yet further. Therefore, a competitive bidding process would force the winner to pay to the auctioneer £5 million a year for the right to supply the service and the winner would then make the normal rate of return on its investment. Consequently, the competitive bidding process drives down profits in the industry to the normal competitive level.

Thus, the franchise payment would be the capitalised value of all of the monopoly profits that accrued to the franchisee over the lifetime of the franchise. In other words, the franchise payment would be the capitalised value of the excessive profits available. If the franchise payments are disregarded for the purposes of calculating the level of profits then it is clear that in this situation one would identify excess profits equal to the level of the franchise payment. However, as we have argued, it cannot be appropriate to ignore the franchise payment when calculating whether profits are excessive or not.

However, if one takes the franchise payment into account then there will only ever be normal profits available to the franchisee. The competitive franchise process will have ensured that these profits, although earned by the franchisee over the lifetime of the franchise will, nonetheless, be rebated to the taxpayer in the form of the franchise fee.

Cave (1995) discusses the relationship between price controls and competition when franchisees obtain the rights to operate in network industries through an auction process.³² He shows that if no price regulation is contemplated in the franchise contract, the franchisee will set the price that equals marginal revenues to marginal operating costs and the effect of the franchising process will be to shift the monopoly profits from the franchisee to the franchiser. In other words, there will be monopolistic distortions from the franchise, but no

³¹ In reality the return would be risk-adjusted, so that the return reflected not only the amount of capital required to run the business for the life of the franchise, but also the level of risk associated with the expected returns.

³² Martin Cave (1995), "Franchise auctions in network infrastructure industries", Presented at the OECD Conference on Competition and Regulation.

monopoly profits for the franchisee. Under these circumstances it is difficult to see how an accusation of sustained excessive profits can be made robust.

Of course, the franchise auction was not undertaken under conditions of absolute certainty and in reality it will be necessary to consider the role of risk and uncertainty in the franchise process.

4.4.4. The role of franchise payments under risk and uncertainty

In reality there will have been significant uncertainty about the levels of costs and revenues associated with the franchises. The potential franchisees will have bid for the right to offer the specified services, subject to a series of contractual obligations as regards the quality and quantity of services, ex-ante price caps on selected fare types and certain specified investment obligations. The contracts that set out the rights and obligations will have been incomplete, as not all contingencies could have been specified in the contract. No guarantees over the level of revenues will have been given and there were no caps on the total level of revenues or profits that could be made from any particular franchise.

The potential bidders would assume that outside the letter of the contract they would simply be subject to a range of commercial and regulatory risks. Each bidder will have had a different probability distribution on each of these commercial and regulatory risks, such as regulatory intervention, passenger demand growth and the potential for cost savings. As each franchisee will have anticipated different risks and opportunities arising from the franchise and, possibly had different tolerances to the acceptance of certain risks, they will have formulated different bids for the same franchise.

For example, one franchisee for a particular franchise may have developed a worse-case scenario of losing £30 million which it believed would occur with a 20 percent probability. It may have had a central-case of making £30 million that it believed would occur with a 70 percent probability. Finally, it may have had a best-case scenario of making £100 million that it believed would occur with a 10 percent probability. In this case, it would have expected its profits to be £25 million ($(£100 \text{ million} \times 0.1) + (£30 \text{ million} \times 0.7) - (£30 \text{ million} \times 0.2)$). In this case we would expect such a bidder to bid around £25 million for the franchise.

Other bidders will have developed their own bids in a similar way. Once each bidder has developed a bid based on their expected profits, then the analysis proceeds much as for the situation described above, with complete certainty. The only difference in this case is that the franchise process will not extract all ex-post monopoly profits that are actually earned, rather it will extract all ex-ante expected profits. The actual profits earned by the winning bidder may be higher or lower than those bid.

This introduces a huge complication for ex-post analysis. What we observe after the event is actual profits, but what was bid for would have been an expected profit that actually comprised a series of different profit streams, each with a different probability of occurring. For example, in the situation described above, assume that the bidder described won the

franchise by bidding its £25 million of expected profits. At the point of the bid, this bidder would have been expected, ex ante, to make only normal risk-adjusted profits from its investment. However, we know that one possibility was that it would make £100 million with a 10 percent chance. If we suppose that this is what actually happened, the bidder will have bid £25 million and made £100 million, giving it ex-post excess profits of £75 million. Although the profits made in this situation are clearly far greater than would have been necessary to attract capital into the activity, they cannot sensibly be characterised as excessive since at the moment the investment was made this outcome was merely one of many possible outcomes, many of which may have delivered far lower profits. In a situation of extreme uncertainty it only makes sense to evaluate profitability as it was perceived ex-ante, rather than it turned out ex-post.

Suppose that upon seeing profits of £100 million the Regulator capped prices so that those profits fell by £75 million (the extent of the “excessive profit”) to the £25 million that had been bid. This would return ex-post profits to their normal level. However, had this action been anticipated at the time of the franchise the franchisee would never have bid £25 million. Indeed, had it known that its best-case scenario would be regulated back until it was £25 million it would have bid very differently. Specifically, had it known that its profits from its two best scenarios would be limited to £25 million it would have bid just £14 million ($(£25 \text{ million} \times 0.1) + (£25 \text{ million} \times 0.7) - (£30 \text{ million} \times 0.2)$). However, if with its new lower bid, it had then achieved its £25 million, under the same distorted regulatory reasoning as before, it would then have been making £11 million of excess profits.

If prices are always capped at the level necessary to make ex-post profits normal and this is known to bidders then there would not be any profitable bid strategy since outcomes better than the bid will always be confiscated, whilst outcomes worse than the bid will always be left on the plate of the franchisee. If this intervention rule were known to bidders, none would bid, as there would be no profitable equilibrium.

4.4.5. What would bidders have assumed about controls on fares?

One important risk facing bidders is the nature of the price control to which they would be subject. One thing bidders did know was that some fares would be regulated, and some would not. For those fares that would be regulated, the franchisees knew the form of the regulation, which was to be set out in each franchise agreement.

There could be said to be less certainty about the other fares. The Government of the day had indicated publicly that it expected that the train operators’ ability to raise these unregulated fares would be limited by competition from other modes. The important question is then whether TOCs setting such fares had complete commercial freedom to set them. Were they limited only by the fact that a profit-maximising monopolist will have some restrictions on the price it can charge because of the elasticity of demand (that is, in other words, the fact that a fare set above the one on the demand curve - see Figure 4.1 above - at which marginal costs equal marginal revenue will lead to a reduction of profits)? Or did competition law, either the pre 1998 Competition Act version up to April 2000, or the

1998 Competition Act from April 2000 onwards, provide some restraint on TOCs pricing policy in regard to the unregulated fares? We think it very difficult for a TOC to draw the line this way in regard to an unregulated fare. We also note that, insofar as different bidders who did draw such a line might be expected to do so in different ways, other things being equal the winning bidder (and therefore the one requiring least subsidy/or providing a highest premium) would be that which drew it closer to, or ultimately at, the profit maximising limit.

NERA's own view on this matter is that if government agencies wish to regulate charges of services that they have a role in providing, it is more sensible to do so by direct regulation than to do so using competition legislation, especially in the context of a franchise competition where bidders are encouraged to put in high bids.

Moreover, such a price control mechanism appears to exist within the franchise agreements, viz:

The Franchising Director shall have the power at any time to alter the obligations of, and restrictions on, the Franchise Operator under this Schedule 5 for any Fare Year beginning after 31 December 2002 (whether by alteration of the value of "k" under paragraph 4 or otherwise) and the Franchise Operator shall accept any such alteration. In the event of such power being exercised, the Franchise Director may and shall, after consultation with the Franchise Operator, make such adjustment to the terms of the Franchise Payments payable hereunder as will reasonably ensure that the Franchise Operator suffers no Net Loss and make no Net Gain as a result of such alteration.

Nevertheless, if the legal advice that ORR receives in regard to the applicability of the 1998 Competition Act is that it does apply to TOCs, and ORR decides to bring a case, then that case would have to be based on the level of fares charged against the background of the franchising process and the subsidies received or payments made. One issue would likely be the extent to which the 1998 Competition Act imposes restrictions on the **way** in which fares could be set to recover franchise payments that TOCs had committed to make. In turn any economic support for this would have to be based on analysis of the "profitability" of individual fare categories, despite our reservations about the difficulty of doing this because of the difficulty of allocating common costs (see Section 4.2 above).

4.4.6. Could TOCs make excessive profits after allowing for risk?

In section 4.4.4 we discussed the difficulties in assessing the allowance for risk in the franchising process by considering ex post levels of profit. We indicated there that these actual profits may be higher or lower than those bid. However, one possibility is that **even though the franchising process did extract all ex-ante expected profits**, there were unanticipated changes that left a TOC with completely unanticipated windfall profits. It is always difficult to determine what is "completely unanticipated". One possibility would be if it could be shown that growth in rail traffic – and hence revenue – was much greater than

had been expected at the time of bidding and was not due to the success of the TOC itself in generating business. This would involve comparing actual traffic and revenue figures, which are of course available, with expectations of both OPRAF and the bidders at the time of bidding.

4.4.7. Did the franchising process extract all ex-ante anticipated profits?

The final issue we believe to be relevant is determining whether profits may be excessive is that of the question of whether the franchising process did extract all ex-ante anticipated profits. An efficient franchising process should do this, so the question can only be addressed indirectly by asking whether the franchising process was indeed conducted efficiently. This would involve asking questions such as:

1. What is an adequate number of bidders for this type of process?
2. Was there an adequate number of bidders?
3. Did the potential bidders have equal access to the information on existing traffic and revenue levels of the services for which they were considering bidding?
4. Were bids handled fairly by the franchising authority?

If a detailed investigation revealed that the franchising process had not met such detailed criteria, then this is evidence that the franchising process **might not** have extracted all ex-ante anticipated profits.

5. CONCLUSIONS

5.1. Introduction

This report has considered the issue of the identification of excessive prices charged by Train Operating Companies (TOCs). The purpose of the project was to design a framework for market definition and the analysis of the abuse of dominant position, as these concepts arise under Chapter II of the 1998 Competition Act. The report sets out the relevance of market definition in the context of abuse of dominance cases. It considers the specific issues that would arise in assessing dominance in the case of a TOC, including: the issues arising in assessing whether individual fares might be judged to be excessive in the context of the allocation of common costs between passengers on the same train; the assessment of a normal return for “trading enterprises” such as TOCs; and the issues that arise because TOCs have competed for the right to operate their services through the franchising process.

5.2. The 1998 Competition Act

The 1998 Competition Act contains two general prohibitions. The Chapter I prohibition prohibits agreements which have “...the objective or effect of preventing, restricting or distorting competition in the UK”. The Chapter II prohibition prohibits conduct that “...amounts to the abuse of a dominant position...”. It is not the existence of dominance that is prohibited under Chapter II, but abuse of that dominance. This report is concerned with Chapter II prohibitions.

5.3. Market Definition

A pre-condition for abuse must be that the firm has a dominant position to abuse. Under European competition law the courts have consistently ruled that the Commission must define a market before it can reach conclusions on the market position of the firm or firms under investigation. As a consequence the first stage of any standard analysis under Article 82, on which Chapter II is based, is to identify the market within which the allegedly dominant firm operates. Once a market has been defined, it is necessary to determine whether the firm under investigation holds a dominant position in that market.

The standard market definition test was developed in merger cases, and relates to the ability of the merged firm to raise prices by a “small but significant amount”. However, this test is not appropriate in the case of abuse of dominance cases, where firms will already be setting prices that maximise their profits.

5.4. Direct Evidence of Dominance

The alternative approach to the one based on market definition is to consider whether the existing prices are above the competitive level, thus providing direct evidence of abuse.

This can be done either by assessing the overall profitability of the firm, to see whether the firm is earning more than a normal rate of return, or by assessing the relationship between the prices of individual goods or services, to see whether these prices exceed costs by a margin greater than that required to earn a normal return. This approach is consistent with the OFT's guidance on the evaluation of excess pricing allegations.

While this approach does not depend on identification of the appropriate market (since the questions are simply whether profits are excessive or whether individual prices are excessive in relation to unit costs), market definition will still need to be addressed in any formal investigation for legal reasons. An appendix to this report reviews market definition issues and precedent with regard to rail services, including evidence on fare elasticities and competition between modes.

There may be more direct evidence of dominance. One approach would be to consider whether specific prices were high in relation to the prices of closely-related products produced under similar cost conditions: in the railway context this would involve comparisons with fares charged by TOCs providing similar services to the one under investigation. Another source of direct evidence of dominance would be evidence of cross-subsidies within franchises, which meant that a TOC had the ability to finance losses on some routes from profits on others. This type of evidence would be particularly compelling where the franchisee had been prepared to pay positive sums for the right to operate the franchise. In this case, the franchisee must be expecting to set fares that enable it to cover the costs of running services **and** to generate enough revenue to meet the franchise payments.

5.5. Implications of Common Costs

A major complication in the rail industry is that a large proportion of the total costs of operating trains on a route cannot be directly attributed to a particular fare type because there are typically few costs that could be avoided either in the short run or the long run by the withdrawal of that fare type. This means that there is a large cost that cannot be attributed to any single fare type in any simple and unequivocal way, cannot be avoided by withdrawal of any single fare type, but that must be recovered from the overall operation of the service through the setting of prices across all fares that in aggregate will cover those common costs. This makes it difficult to assess profitability of individual fare categories: this profitability will depend on the cost attribution method used.

5.6. Estimating a Normal Rate of Return for a Trading Business

The normal methodology for assessing company profitability by focusing on the achieved return on capital employed and comparing it with a normal rate of return for the type of activity concerned cannot be readily applied to a trading business like a TOC with few or no requirements to finance capital expenditure. Our review of regulatory precedents in this area shows that:

- in contrast to the position in other regulated industries, there is as yet no consensus as to an appropriate conceptual approach to assessing a reasonable operating margin for such businesses; and
- previous experience suggests that, because of this, competition authorities will be extremely reluctant to reach conclusions on whether a particular level of operating margin constitutes excessive profitability.

5.7. Excessive Pricing with Franchising

The complication that arises with franchising is that TOCs will have competed for the right to operate services through a competitive franchising process. If this process is efficient, it should extract the value of the ex ante expected profits from the system. It might be regarded as a reasonable expectation on the part of the TOCs that they would then be able to recoup payments made, either those made directly to OPRAF/SRA or those made within their franchise to cross-subsidise loss-making services, by means of fares charged.

In assessing this, an important issue is that of how far TOCs are constrained in the fares they can charge for individual services, given that they knew in advance that some fares would be regulated, and some would not. The crucial consideration is that of whether competition law imposes constraints on the levels that can be set for those fares that were **not** to be subject to explicit regulation. Investigation of such fares would require allocation of costs between passengers, and such allocation would be dependent on the allocation rule used.

As a matter of public policy, NERA's own view is that if the public authorities wish to control fare levels, it is much more efficient to do so through explicit regulation rather than through use of Chapter II investigations.

Profits earned by a franchise might also be regarded as excessive if there had been completely unanticipated increases in the demand for rail services that led to "windfall" profits. An investigation would need to show that such increases could not have been expected at the time bids were made. Since bids would likely have been made on the basis of a range of forecasts, proof of this would depend on actual traffic levels being outside the range of traffic forecasts made.

Finally, a TOC might be able to earn excessive profits if the franchising process had not been efficient, and so did not extract the value of ex ante anticipated profits. Such a line of investigation would have to reveal that there had been serious flaws in the conduct of the original OPRAF franchise bidding process.

APPENDIX A. TERMS OF REFERENCE ON CONSULTANTS REPORT ON FARES POLICIES

A.1. Context

The Competition Act 1998 contains two prohibitions, the first prohibiting agreements that prevent, restrict or distort competition, and the second prohibiting the abuse of a dominant position. These prohibitions came into effect on the 1 March 2000. The Rail Regulator is responsible, concurrently with the Director General of Fair Trading for enforcing the prohibitions within the railway sector.

Since the prohibitions came into force, it has become apparent to the Rail Regulator that he may wish to use his new Competition Act powers to investigate the fares policies of train operators in relation to fares that are not directly regulated by the Franchising Director via the franchise agreement.

Key elements of any such investigation would include the definition of the relevant market(s) and the economic assessment of prices, costs and profits, and whether or not they constitute an abuse of a dominant position. The Regulator would therefore like to have available an independently verified framework that can be used for assessing such matters. The purpose of this document is to define the terms of reference for a project that will deliver such a framework.

A.2. Contents of Report

The consultants will provide a report with two main sections. The first section will address the definition of the relevant market or markets which should be defined in cases where the ORR is applying the Competition Act 1998 to passenger transport markets.

The second section of the report should outline the principles for assessing whether the fares of a dominant operator should be considered abusive. The consultants will also be expected to undertake particular case studies assessing examples of actual fares policies and structures, using data provided by the ORR.

A.3. Market Definition

The consultants will need to set out the appropriate approach to defining the relevant markets for UK domestic passenger rail travel. This approach must be consistent with the Competition Act 1998 guidelines "Market definition", and "Application to Railway Services" and relevant EC jurisprudence.

The section of the consultant's report dealing with market definition will be expected to include the following elements:

- a discussion of the appropriate economic principles which should be applied in defining passenger transport markets;
- a survey of the relevant economic literature on defining passenger transport markets and assessing demand for transport services;
- an identification of, and discussion of the economic principles raised by, the relevant European Commission decisions and court cases which have dealt with market definition for passenger transport markets;
- a discussion of the application of appropriate quantitative techniques to passenger transport markets (for example, as set out in the LECG OFT Research paper), including identifying any particular problems or adjustments which would need to be made to standard techniques in this area;
- a brief discussion of the appropriate measures to use in terms of calculating and assessing market shares; and
- comments on the data required, and its availability, to apply the principles and techniques set out.

This section of the report will be required to consider the different dimensions of the market. It should include a discussion of the appropriate approach to be taken in dealing with the following factors:

- geographic market definition (the route or routes to be included in the relevant market);
- different transport modes (air, sea, coach, private car, rail);
- whether relevant markets should be defined according to journey purpose (e.g. business or leisure) or class (first or standard);
- the effect of ticket type and availability (e.g. whether the restrictions on use of certain ticket types delineate separate markets);
- the relevance of time of day or season and whether separate time markets may need to be defined (e.g. peak and off-peak);
- the relevance of input market definition to the assessment of passenger markets; and
- supply-side substitution.

A.4. Assessment of Abuse

The report will be expected to provide a framework for assessing whether the fares set by a dominant operator are abusive. This framework should be consistent with the Competition Act 1998 guidelines (especially "Assessment of individual agreements and conduct" and "Application to Railway Services") and relevant EC jurisprudence.

This section of the report will be expected to include:

- a discussion of the economic principles which should be applied to assess whether a particular fares policy or structure is anti-competitive or excessive;
- a description of the quantitative and analytical techniques which might be applied to assess whether fares are excessive, with a discussion of the appropriate data requirements;
- comments on the data required, and its availability, to apply the principles and techniques set out;
- a brief review of the relevant economic literature for assessing whether prices are anti-competitive or excessive; and
- an identification of, and discussion of the economic principles raised by, the relevant European Commission decisions and court cases which have assessed the pricing policy of passenger transport providers

It is expected that it will not be possible for the report to cover every possible potential abuse relating to fares. As a minimum the report will be expected to discuss anti-competitive bundling of services and excessive pricing. The ORR will expect to discuss the abuses to be covered in detail with the consultants shortly after their appointment.

The report will be expected to set the general economic principles by which fares could be assessed and to discuss the treatment of particular issues specific to the UK rail industry. In particular the analytical treatment of the following factors should be addressed:

- fares regulation;
- effect of major investment (for example in new rolling stock);
- provision of public subsidy to the industry and franchise premium payments;
- the existence of cross subsidies, both between TOCs (through the sSRA) and with TOCs (between routes).

A.5. Empirical Case Studies

It is also probable that the report will be required to contain a separate annex that applies the principles and techniques described in the main body of the report to specific case studies. This will involve suggesting the appropriate relevant market for particular point-to-point flows and an assessment of the fares in those markets.

The case studies will be selected by the ORR following discussions with the consultants. They would be expected to apply the principles and relevant quantitative techniques set out in the main body of the report to define suitable relevant markets and assess whether particular sets of fares were abusive.

The consultants will be expected to discuss data availability and requirements with the ORR shortly after the commencement of the project. They will also be required to provide input to assist ORR personnel in drafting information requests to appropriate industry bodies. The consultants will not be expected to undertake independent data collection exercises.

These case studies, and the general reports, will be considered confidential.

APPENDIX B. INTER-MODAL COMPETITION, CONSTRAINTS ON TOC PRICING, AND PREVIOUS MARKET DEFINITION ASSESSMENTS

B.1. Introduction

This section of the report provides a review of the empirical evidence on the extent of competitive constraints that exist on TOC pricing at current levels, and is intended to provide evidence to support the conceptual issues of market definition introduced in section 3.2. As we explained in Section 3.2, much of this evidence will not be relevant to the definition of the market for the purposes of chapter II investigations because of cellophane fallacy concerns³³, but the information may have considerable value in the definition of the market in intra-modal and inter-modal mergers, and for identifying the impacts of various exclusionary practices.

Some rail fares are regulated, so there is a presumption that the regulatory system will preclude the regulated fares from being excessive. Those fares left unregulated following privatisation, namely first-class journeys and a number of inter-city fares, were expected to be restrained by competitive pressure from coaches, airlines and private cars.³⁴ Empirical evidence of the competitive pressures on specific rail fares will centre around investigations of the price elasticity (both own-price and cross-price) of demand for rail travel.

We consider evidence on own-price elasticities of demand for rail travel in Section B.2, and evidence on inter-modal competition and cross-price elasticities in Section B.3. Section B.4 reviews previous studies of market definition in the rail transport sector. Section B.5 provides a broad framework for meeting the legal requirements to describe the market in an excessive fares case.

B.2. The Own –Price Elasticity of Demand for Rail Travel

B.2.1. The role of evidence on own-price elasticity

A crucial issue in regard to the ability to charge excessive fares is the elasticity of demand. NERA's 1993 report to the OFT on market definition notes: "it is own-price elasticity that underlies the approach to market definition under the 5 per cent test".³⁵

³³ Although we recognise that a market definition study will be a legal requirement in an investigation

³⁴ "Private train operating companies will be given a free hand to determine the fares for first class journeys and on inter-city routes where there is competition from coaches, airlines and the private car" *Financial Times*, May 8th 1995.

³⁵ OFT *Market Definition in UK Competition Policy* OFT Research Paper 1, 1993.

If demand is inelastic, a profit-maximising operator could increase revenue by increasing prices. However, if demand is inelastic **at existing prices** then the operator must not be maximising profits, since the profit maximising position involves elastic demand. We can conclude:

- If evidence from previous studies of rail demand elasticity shows demand to be price inelastic, then operators at the time the estimates were derived would have increased profits if they increased prices.
- If an operator is currently maximising profits, the point elasticity of demand at the current price will be elastic.³⁶ However, deriving this value would be difficult since the most satisfactory estimates of rail demand elasticity have been derived from time series analysis.

Since most estimates of rail fare elasticity are derived from time-series analysis, and **since there may be a supposition that previous fares were not excessive**, it is useful to look to see if there is evidence from sectors under investigation of inelastic demand in the past. Such evidence implies that fares could have been increased profitably in the future.

B.2.2. Evidence on own price elasticities

There are a number of sources of evidence on own-price elasticities for rail fares, usually within studies investigating the explanatory power of different economic variables on demand for passenger rail services. Within these studies, some of those that explicitly address the question of price elasticity are confidential, and some are in the public domain. The principal sources of evidence we have drawn on in this survey³⁷ are:

- The Passenger Demand Forecasting Handbook (PDFH); and
- NERA's 1999 report on the analysis of passenger rail demand for the sSRA.³⁸

Important main published sources on rail fare elasticity include:

- Owen and Phillips' 1987 study on the demand for InterCity services;³⁹

³⁶ Elastic demand is necessary but not sufficient for profit maximisation. In addition, profit maximisation is not necessary for prices to be excessive.

³⁷ Work has also been undertaken for the Association of Train Operating Companies (ATOC), but we have not had access to this..

³⁸ These results are quoted with sSRA permission, although NERA's views do not necessarily represent those of the sSRA.

³⁹ A. D. Owen and G. D. A Phillips "The characteristics of railway passenger demand" *Journal of Transport Economics and Policy*, September 1987.

- Wardman, Toner and Whelan's study on interactions between rail and car in the inter-urban leisure traffic market in Great Britain;⁴⁰
- A study by Nash, Fowkes and Whiteing at the Institute of Transport Studies at the University of Leeds on the demand for InterCity services, published in 1985.⁴¹

All of these studies investigate evidence on elasticities for different service types by two principle methodologies, that is evidence based on aggregate fare elasticities and evidence based on differential fare elasticities. Aggregate fare elasticity studies investigate the impact on different journey types of a rise in all types of fare, while differential fare elasticity studies investigate the effects of a change in some, though not all, rail fares change, thus altering the price differential between individual tickets. These are examined separately below.

B.2.3. Aggregate fares

Evidence on aggregate fares elasticities is the more consistently robust methodology, both in terms of statistical robustness and also the degree of consistency between the findings of different studies. The most comprehensive survey of the results of studies on aggregate fare elasticities is that reported in the Passenger Demand Forecasting Handbook (PDFH).

The PDFH is available to subscribers to the Passenger Demand Forecasting Subscription Service (PDFSS). It is produced by TCI Operational Research and the present version provides a summary of studies on passenger demand conducted up to 1997.

Table B.1 summarises the PDFH recommended fares elasticities for different types of average journey and flow.

⁴⁰ M. Wardman, J.P. Toner and G.A. Whelan "Interactions between rail and car in the inter-urban leisure travel market in Great Britain", *Journal of Transport Economics and Policy*, May 1997

⁴¹ A.S. Fowkes, C. A. Nash and A. E. Whiteing "Trends in inter-city rail travel in G.B", *Transportation Planning and Technology* 1985

Table B.1
Aggregate Fare Elasticities - PDFH Recommended Values

	First class	Commutin g	Business	Personal¹	Leisure	Total
InterCity – London-based ²	-0.5	- 0.5	- 0.5	-1.2	-1.5	-1.2
InterCity – non-London	-0.5	- 0.6	- 0.5	-1.0	-1.2	-1.0
South East – London-based ²	-0.5	- 0.3 to - 0.5	- 0.2	-0.6 to -0.7	-0.7 to -1.1	-0.3 to -1.1
South East – non-London ²	-0.5	- 0.3 to - 0.5	- 0.2	-0.6 to -0.7	-0.7 to -1.1	-0.3 to -1.1
Urban – non-London	-0.5	- 0.6	- 0.5	- 0.8	- 0.9	- 0.7
Other	-0.5	- 0.8	- 0.5	- 1.0	- 1.2	- 1.1

¹ *Personal travel is of a non-optional (and non-leisure) nature, but passengers usually pay their own fare.*

² *Lower elasticities apply to short distances, and higher elasticities for long distances.*

The elasticities given in Table B.1 are recommended for use on ‘average’ flows, that is flows without strong or weak competition. The PDFH concurs that elasticities may be up to 20 per cent higher for flows that face high levels of competition, for example from a direct motorway link or if the rail service is poor. Additionally, elasticities may rise if the fare is a high proportion of the generalised journey cost. Concurrently the elasticity may fall by 20 per cent if competition to rail is weak, for example if there are poor road services, or a strong rail service. Additionally the elasticity may fall if the fare is a low proportion of generalised journey cost.

Two distinct features can be seen in Table B.1. First, elasticities for business travel are lower than those for commuting, which in turn are lower than those for personal and leisure travel. Second, elasticities for long distance travel are generally greater than those for short distance travel, possibly reflecting the fact that fares account for a higher proportion of the total (generalised) journey cost for long distance travel. It is interesting to note that although research suggests that first class elasticities are very sensitive to the level of competition from air and car, PDFH estimates suggest an overall elasticity of -0.5 for all types of first class travel.

Two different types of analysis underpin the elasticities shown in Table B.1:

- the “total” elasticities, applying across all journeys on a particular type of flow, were generally derived from econometric analyses of actual data; and
- the elasticities for different types of journey were generally derived from stated preference studies, based on analysis of specially designed survey questions.

Different approaches are used because ticket and revenue data do not distinguish between different types of journey. However, the results of survey data analysis for different types of passenger generally yield elasticities which are consistent with the overall elasticities estimated from actual data.

Since publication of the most recent edition of the PDFH, a significant study has been carried out by TCI Operational Research.⁴² Although this more recent study primarily focused on the impact of external factors (GDP growth and time trends) fare elasticities for full and reduced fare journeys were derived. Results for four of the route types already examined in Table B.1 are shown in Table B.2 below. In the same study TCI OR also derived separate estimates for season ticket journeys. However, these proved more volatile, with some coefficients appearing to be too high.

Table B.2
TCI Operational Research (1997) Fare Elasticities for Full and Reduced Fare Journeys

Flow Type	Fare elasticity
InterCity – London-based	-0.73
InterCity – non-London	-0.95
South East - London-based	-0.22
South East – non-London	-0.66
Regional	-0.84

Despite the primary focus on the impact of external factors (GDP growth and time trend), this later TCI OR study derived estimated fares elasticities which are broadly consistent with the existing PDFH recommendations. Although a number of the results shown in Table B.2 above appear to be slightly lower than the PDFH previous estimations, TCI OR's elasticity estimations taking account of all journeys (as opposed to full and reduced fare journeys) are closer to the figures shown in Table B.1.⁴³

In 1999, NERA prepared an analysis of passenger rail demand for the sSRA, applying both time series (on both a short and long time scale) and panel data techniques to investigate the impact of external economic variables, including fare changes, on demand. Using aggregated data, NERA's resultant price elasticities are shown below.

⁴² TCI Operational Research, (1997), *Forecasting Rail Demand: The Effect of External Factors*.

⁴³ Specifically, TCI OR's estimated fare elasticities (all journeys) are: -0.94 for London-based InterCity flows; -0.95 for other InterCity flows; -0.51 for London-based South East flows; -0.62 for other South East flows; and -0.88 for other flows.

Table B.3
Results of NERA's Aggregate Price Elasticity Study

Method	Panel data	Time series	
		Short run	Long run
London-based inter-city flows (20)	-0.710	-0.036	-0.611
Short London journeys	-0.904	-0.636	-0.990
London/SE total	-0.906	-1.133	-0.951
Regional Inter-urban and Inter city	-0.680	-0.398	-0.096
Other regional	-0.630	-0.820	-1.060
PTEs	-0.730	-0.815	-1.498
Average	-0.783	-0.783	-1.008

The first noteworthy differences between the most recent PDFH and the NERA study is that the elasticity for London-based InterCity flows is rather lower than the PDFH average (TCI OR found a similar result for non-season ticket flows). This may reflect the impact of increasing road congestion on motorways and other major roads into and out of London, together with (for flows to London) the difficulty and cost of parking in central London. As a result, a higher proportion of those who could make their journeys by road or rail may be already travelling by rail, and unlikely to switch in response to relatively small price changes.

Secondly NERA found that local and regional flows in the London area may have a stronger than expected response to changes in price than that indicated in PDFH results.

Owen and Phillips used time series econometric analysis to examine the determinants of demand for 20 London-based InterCity flows⁴⁴ using four-weekly data for the period from 1973 to mid-1984. This period included two major economic downturns, the deregulation of long distance coach services and the introduction of High Speed Trains (HSTs) on some routes.

Owen and Phillips estimated separate equations for each of the 20 flows, based on a partial adjustment model which allowed them to estimate both short term and long term elasticities. Their main findings are summarised in Table B.4.

⁴⁴ These were to Bath, Birmingham, Bristol, Cardiff, Carlisle, Edinburgh, Glasgow, Leeds, Leicester, Liverpool, Manchester, Newcastle, Norwich, Nottingham, Plymouth, Preston, Stockport, Swansea, Swindon and York.

Table B.4
Owen and Phillips' Estimated Elasticities

	<u>Short-term elasticities</u>			<u>Long-term elasticities</u>		
	Total	1st class	2nd class	Total	1st class	2nd class
Fares	- 0.69	- 0.67	- 0.81	- 1.08	- 1.00	- 1.17

Source: Owen and Phillips (1997), Table 7

Though these figures represent the median results of the study, there was considerable dispersion around the figures, with only 9 out of 20 flow results for total journeys' elasticities statistically significant. The outcome for separately estimated 1st and 2nd class elasticities was more robust.

Wardman, Toner and Whelan, examining the interaction between rail and car travel in the inter-urban leisure market in Great Britain, derived both own price and cross price elasticities for the rail and car markets. This utilised data from the 1990 Trans-Pennine rail strategy study, estimating models solely according to those users who chose between rail and car travel. Elasticities were calculated according to a standard linear model, and a more accurate generalised, non-linear model that was perceived to be more accurate due to the better fit of results. Additionally, separate estimates were carried out for people travelling alone and those in a group, recognising that larger group sizes could be a deterrent from rail travel because of a larger total cost of travel. The results of the two models are shown in Table B.5 below

Table B.5
Wardman, Toner and Whelan's estimates of own price elasticities of train travel

	Generalised Model			Standard Model		
	Alone	Group	Total	Alone	Group	Total
Train cost	-0.57	-0.65	-0.59	-0.45	-0.65	-0.51

Source: Wardman, Toner and Whelan (1997), Table 5

As stated above, Wardman et al's models were only estimated according to those who chose between rail and car travel. Therefore they do not allow for any suppression or generation of trips, which the author's believe would act to reduce the elasticity values. Taking this into account, these own elasticities are largely consistent with the previously stated long-run evidence of own-price elasticities on inter-urban flows.

A recent study by the Centre for Economic and Business Research (CEBR)⁴⁵, examining the influence of a range of economic variables on the demand for London based flows in the South East, also derived some estimated price elasticities. However, these values are

⁴⁵ Centre for Economic and Business Research, (1998), *Rail Forecasts for the West End and the City*.

significantly different than those recommended for use in the PDFH. For season tickets/commuting tickets, the CEBR estimated a price elasticity of -1.5 , making these tickets more price elastic than reduced price tickets, whose elasticity was estimated as -0.9 , more in line with PDFH estimates.

B.2.4. Differential fares

The evidence and recommended values for fares elasticities presented above are based on the assumption that all types of rail fare rise or fall at the same time, and therefore significant switching between different types of ticket for the same journey is not expected. However, the PDFH also considers differential ticket prices when some, but not all, rail fares change, and thus the differential between individual ticket prices also changes.

Differential ticket prices can reflect a number of factors, including:

- the level of comfort enjoyed on a particular train (First class or Standard class);
- restrictions on which trains can be used, expressed in terms of times of day/week (for example, with Savers and Cheap Day Returns), the train operator (dedicated tickets) or the route (for example, “not via London”);
- requirements to book in advance (for example, with APEX tickets), often with limited numbers of tickets being available for each available train;
- restrictions on who is eligible for certain ticket prices, for example under certain railcard schemes.

For Saver and Cheap Day return tickets, the PDFH recommends that the normal fares elasticity be increased by between 3 per cent (for leisure journeys) and 70 per cent (for business travellers using Savers) to take account of switching to and from Open tickets in the event of a change in price differentials. These recommendations are based on modelling of passenger choices carried out by TCI OR and using a combination of PDFH recommendations and passenger survey data. From data on timetables, fares, train restrictions, journey purpose and demand profile, the model predicts the number of passengers purchasing different types of ticket, and can therefore be used to derive fares elasticities for the case where only one fare changes.

A similar treatment is suggested for First class: the fares elasticity should be doubled (from -0.5 to -1.0) if only First class (and not Standard class) fares change. This is based on cross-fare elasticities estimated by Owen and Phillips, though these authors report that (for First and Standard class together) the majority of estimated cross-fare elasticities were insignificant and a few appeared “unreasonable”. Similarly, the PDFH recommends that fares elasticities for travelcards should be increased by 50 per cent when normal season ticket prices do not change at the same time, based on a statistical analysis by TCI OR of South East to London flows between 1987 and 1994.

The PDFH also contains a set of market share assumptions in cases where a minor operator introduces a dedicated ticket. This distinguishes between full fare and reduced fare passengers and between different types of flow, and sets out different market share assumptions depending on the size of the fare differential and the relative speed and frequency of the minor operator's service. The assumptions are based on a study carried out by Accent Marketing & Research and Hague Consulting Group, which used the results of a stated preference passenger survey to construct a passenger choice model.

Other evidence summarised in the PDFH includes several studies using stated preference passenger surveys to examine how railcard holders' behaviour might change in the absence of those railcards. The PDFH suggests that railcards create a preference for rail and a psychological desire to use the card (having bought it), which may reduce railcard holders' fare elasticities by about one-third, though it also notes that railcards are often targeted at more elastic market segments.

A problem in assuming that traveller's behaviour may allow them to switch to alternative rail services given a rise in price of one type of fare is that in many cases, especially at peak times, such alternatives simply do not exist. Where they do, they may involve an increase in journey time that would deter peak rail travellers.

In his 1999 IEA Regulation Lecture, Ian Jones⁴⁶ reported on the results of analysis that NERA carried out on the impact on rail fares of the rather limited amount of on-rail competition that has occurred so far in Great Britain. This analysis compared changes in frequencies and fares on 19 competitive rail origin/destination flows⁴⁷ between 1995 and 1999 with network average changes over the same period.

- In regard to frequencies, the average increase in frequencies on the competitive flows was nearly 50 per cent greater than the average network-wide increase.
- **In regard to fares**, the analysis showed that average unregulated ordinary fares on the competitive routes only rose by two-thirds as much as revenue per passenger-km on ordinary (ie non-season) fares across the network as a whole.

Thus it appears that on-rail competition, where it exists, does provide a restraining impact on fares. However, on-rail competition has been limited in the rail passenger industry. First of all, it was restricted in order to encourage bidding for franchises. If franchisees were to face the possibility of competition on their routes, then the value of the franchise would be reduced, so bids for subsidy, and hence the call on the public purse, would be higher. Secondly, there is seen to be a role for centralised co-ordination of service levels and fares in

⁴⁶ I. S. Jones "Railway franchising: is it sufficient?" IEA Regulation Lectures, December 1999. See also I. S. Jones Transit 111, July 23rd 2000.

⁴⁷ Some of these involved competition on the same route, and some involved competition between services on different routes serving the same origins and destinations.

the rail industry, evidenced for example by frequent mention of the benefits of an “integrated service”.⁴⁸ Thirdly, competition uses up track capacity, and track capacity is increasingly scarce, so this limits the scope for future competition if the existing regime is relaxed. Fourth, some TOCs, particularly Virgin West Coast, are protected from on-track competition by the terms of their franchise agreement.

B.3. Evidence on Inter-Modal Competition and Cross Elasticities

An important set of exogenous factors affecting passenger rail demand is the extent of competition from other modes of transport. Changes in the nature of competition over a particular route will impact on the demand for rail services, and variations in the extent of competition between different routes may lead to differences, for example, in fares elasticities. The extent of inter-modal competition is summarised by the cross-elasticity of demand. Cross elasticities are relevant to the competitive constraints facing rail services as a significantly positive figure would imply a considerable element of competition between rail and another mode of transport where they serve common routes. This competition would be one factor that deterred fare increases. However, care has to be taken in interpreting relative cross-elasticity values, since these values will in part be determined by relative modal shares.

The impact of long run changes in the nature of competition, such as that caused by steadily increasing car ownership, may be captured in estimated time trends.⁴⁹ Beyond this, however, attempts to investigate the impact of competition on passenger rail demand have been affected by the difficulty of deriving a robust, quantitative measure of the extent of competition. Therefore there has been a scarcity of work on cross elasticities, especially outside an urban context. Most studies using actual data have focussed on quite large differences, either comparing demand “before and after” a specific major change, or comparing demand on otherwise similar routes which are subject to varying degrees of competition. Examples of such studies which are summarised in the PDFH include

- the studies by Owen and Phillips and Fowkes, Nash and Whiteing, described in the previous section, which introduced dummy variables to capture the impacts of long distance coach deregulation and increased competition from air (as a result of improved shuttle services and better surface links to the main London airports). Owen and Phillips found that introduction of shuttle services on three air routes (Glasgow, Edinburgh and Newcastle) had a significant impact on rail demand on those routes. Fowkes, Nash and Whiteing also found air shuttles impacted on rail demand, taking 8 per cent of BR traffic on the routes affected.

⁴⁸ The Regulator’s draft guidelines consider how agreements which are designed to secure integrated transport services or will otherwise secure transport benefits will be dealt with (ORR, 2000, paragraphs 5.10 to 5.15).

⁴⁹ Some studies have attempted to identify a separate impact of car ownership, though there is a danger that the findings from such studies will be affected by correlation between car ownership and economic variables.

- a study by Glaister of commuting from Bedford and High Wycombe to London,⁵⁰ over a period which included the opening of the M40 motorway (which reduced the demand for rail services from High Wycombe, but mainly affected cheap day rather than season ticket journeys);
- a study by ITS comparing fares and journey time elasticities for different non-London inter-urban flows, using the relative generalised cost of car and coach travel as a measure of the strength of competition for that flow.⁵¹ ITS concluded that fares elasticities could vary by ± 20 per cent where competition was stronger or weaker than average.

To overcome the problems with quantifying competitive factors and identifying the separate impact of competition from other modes, some studies have instead used stated preference techniques to assess the impact of changes in relative fares and service quality on passengers' choice of transport mode on a particular route. The PDFH summarises studies which have applied these techniques to examine competition between rail, bus and car on Regional Railways North East routes, between rail and car along the M1/A1 corridor and between rail and coach on the Midland Main Line. In addition, a study of competition between rail, car and coach on the TransPennine route used both actual data ("revealed preference") and survey data ("stated preference") to estimate particular coefficients.

Reflecting the results for both the ITS study and some of the stated preference studies, the PDFH's main recommendation is that the cross-elasticity of rail demand to the generalised cost of car, coach or air transport will be in the range 0.2 to 0.4. In addition, again reflecting the ITS findings, it recommends that fares elasticities on longer distance flows may be adjusted by up to 20 per cent to reflect strong or weak competition from other modes.

In the MMC report on the Midland Main Line franchise a wide range of evidence regarding the extent of competition between rail and coach services was discussed. Within this the MMC report considered a range of evidence regarding the price elasticities of demand for an between coach and rail. This issue was addressed in an MMC commissioned study carried out by the Institute of Transport Studies (ITS). ITS reviewed a large amount of empirical evidence from studies of rail and coach demand throughout Great Britain, the results of which suggested the elasticities for leisure passengers on MML that are shown in Table B.6.

⁵⁰ S. Glaister "Some characteristics of rail commuter demand" *Journal of Transport Economics and Policy* 17, 1983, 115-132.

⁵¹ M. Wardman "Inter-urban rail demand, elasticities and competition in Great Britain: evidence from direct demand models" *Transportation Research E* 33, 15-28. Generalised cost is calculated as the generalised journey time (see Section 2.3) multiplied by an appropriate value per minute of time, and added to the fare. The fare itself may also be adjusted in this calculation to take account of other aspects of service quality, such as hygiene and comfort (see Section 2.3.3).

Table B.6
ITS Suggested Elasticities for Leisure Passengers on MML Routes

	1 per cent increase in price of	
	Rail	Coach
<i>% Effect on demand for</i>		
Rail	-0.9*	0.14 ⁺
Coach	0.3	-1.1

* *This reflects an elasticity of -0.74 for travel to London and -1.06 for travel from London*

+ *At the time NERA argued that this result implies that 40 per cent of those who switched from coach as a result of a coach fare increase would switch to rail, which we regarded as "quite wildly implausible" (MMC, 1996, p.80).*

The level of cross price elasticity on a particular journey will depend on a number of factors, including the relative volume of rail and coach traffic on individual flows. In the discussion following the publication of the Midland Main Line report it was suggested that the impact of changes in rail prices on rail demand may be much greater than envisaged, and that elasticities are likely to be strongly route specific and varied through time.

The study of the interaction between rail and car in the inter-urban leisure travel market by Wardman et al, found that the cross-elasticity of demand for car travel with respect to rail was small (estimated value of cross elasticity 0.06). Therefore the effect on car use of improving conditions for train travel, and therefore also the effect on car use of worsening the conditions for train travel (including an increase in price) would be limited at the margin. However, because of rail's low share of the leisure market in the sample used for the estimates, rail demand is much more sensitive to variation in the cost of car travel (estimated value of cross elasticity 0.25). These values are approximately equivalent to those suggested by the ITS study.

A more general survey of cross-elasticity values in the UK transport sector is provided in Acutt and Dodgson.⁵² This paper reviews the general relationship between cross-price elasticities, own-price elasticities, market shares, and diversion factors. These are then used to derive estimates of the cross-elasticities of demand for different types of rail and bus travel in relation to the price of petrol, and the cross-elasticities of demand for car travel in relation to fares of different types of rail and bus service.⁵³ The paper shows that the cross-elasticity of a first mode in relation to the price of a second is equal to:

- [the own-price elasticity of the second mode] x
- [the share of the second mode divided by the share of the first] x

⁵² M.Z.Acutt and J.S.Dodgson "Cross-elasticities of demand for travel" *Transport Policy* vol 2, no 4, 1996, 271-277.

⁵³ The paper does not consider cross-elasticities between different public transport modes.

- [the diversion factor, in terms of the proportion of increased trips on the first mode as a result of a price decrease that are diverted from the second].

Results for rail services are as follows:

- Cross-elasticity of former InterCity services with respect to the price of petrol +0.094;
- Cross-elasticity of former Network SE services with respect to the price of petrol +0.041;
- Cross-elasticity of former Regional services with respect to the price of petrol +0.091;
- Cross-elasticity of Underground services with respect to the price of petrol +0.017.

These results indicate that petrol prices will have some influence on the demand for rail travel. However, in the context of the present report, petrol prices should be regarded as an exogenous factor determining levels of rail demand on particular routes. The increases in real petrol prices in recent years will have strengthened the demand for rail services, and will therefore have increased the ability of TOCs to increase fares charged for any specific level of rail traffic.⁵⁴

B.4. Review of Previous Studies of Market Definition in the Rail Sector

B.4.1. MMC merger investigations involving rail franchises

Market definition in the rail sector has also been considered in the three MMC reports into rail franchises, namely Midland Main Line, Central Trains and ScotRail.

B.4.1.1. Midland Main Line

The MMC's December 1996 report considered the acquisition of Midland Main Line (MMC) by the National Express Group (NEG).⁵⁵ NEG's wholly-owned subsidiary National Express Ltd (NEL) operates coach services within the UK. MML operates rail services mainly between South Yorkshire and the East Midlands and London St. Pancras. Five coach services operated by NEL overlap with MML's rail services between central London, and respectively Sheffield, Chesterfield, Derby, Nottingham and Leicester. Coach services are less frequent and slower than rail services, but fares are cheaper.

The Commission's Terms of Reference refer to "the supply of public transport services between London and, respectively, Derby, Sheffield, Chesterfield, Nottingham and

⁵⁴ The rise in petrol prices could also reduce the elasticity of rail demand at current rail fares. See above, section B.2.3.

⁵⁵ MMC *National Express Group plc and Midland Main Line Ltd: a Report on the Merger Situation* The Stationery Office, December 1996.

Leicester” and so the question of market definition was considered within a specific geographical area.

The chapter on the market describes the market within which the services operated, considering the nature of the rail and coach services operated by the merged organisations, the evidence on mode shares on different routes, and detailed fare levels and structures. The chapter also considers competition in the following forms:

- competition between coach and rail;
- competition between rail operators, primarily actual competition for MML from services on the parallel East Coast Main Line connecting to Nottingham, Chesterfield, and Sheffield,⁵⁶ and potential competition from new entrants on the MML, which would be constrained by the specific features of the moderation of competition regime applying to this route.
- competition between coach operators. There was only one minor actual competitor, so the MMC concentrated on potential competition. Although entry barriers to coach operation are low, the MMC believed that National Express’s ability to respond immediately to lower fare entrants reduced the possibility of the “hit-and-run entry” associated with contestable markets.
- competition from the car, where MMC concluded that there was competition in the corridor between rail and car, and between coach and car. However, this was not judged to be a sufficiently strong competitive constraint to prevent any increase in rail fares.⁵⁷

There were no air services operating in this corridor.

The results on competition between rail and coach are of particular interest for the present study. Do coach services and coach fare levels restrict the ability to raise rail fares (so that, in the context of a merger investigation such as this) rail fares might rise following a merger? The MMC considered in detail the passenger profiles of the rail and coach services, in terms of age and social class of travellers, and in terms of journey purpose (primarily differentiated between business and leisure travel).

Evidence on the extent of competition between rail and coach was presented under three headings:

⁵⁶ “While there is a limited degree of choice and competition, in particular for passengers who may drive to ICEC stations to user quicker ICEC services, we do not believe this represents such strong actual or potential competition to MML as significantly to constrain MML’s decisions with regard to fares or levels of service”. Op. cit., para. 2.5.1, p.17.

⁵⁷ Op. cit, para. 2.61, p. 19.

- The pricing behaviour of the two modes;
- Surveys of passengers' reasons for choosing mode and passengers' responsiveness to changes in prices; and
- Evidence of cross-price elasticities of demand.

The MMC found that coach fares had to be set some 30 per cent below competing rail fares to offset the poorer service quality of coach journeys, i.e. coach prices were constrained by the competing rail mode. A study of the impact of coach deregulation in the early 1980s had showed that, when coach fares were cut after deregulation, about half the newly-generated demand on some services in the East Midlands was from passengers transferring from rail.⁵⁸ British Rail's main response had been to introduce lower fares, primarily Saver fares, aimed at leisure passengers, so that coach competition had imposed a constraint on this type of rail fare. When BR introduced Apex fares in the early 1990s the number of coach passengers between London and Scotland fell sharply.

Survey evidence on passengers' reasons for choosing coach rather than rail showed that lower fares were a major factor. We have reviewed the results on cross-elasticities from the report commissioned by the MMC in section B.3: this derived values of the cross-elasticity of demand for coach services with respect to the price of rail travel of +0.3, and of the cross-elasticity of demand for rail travel with respect to the price of coach travel of +0.14.

For the rail cross-elasticity with respect to coach fares specific values of +0.46 for students and +0.29 for retired people were also derived. For coach cross-elasticity with respect to rail fares specific values of +0.90 for students and +0.60 for retired people were derived. The MMC correctly note that cross-elasticity values cannot be directly interpreted as measures of consumer substitutability as their actual values are affected by relative market shares. Given the number of passengers of different types the MMC's consultants concluded that any tendency of a joint owner of the coach and rail services to raise price would be stronger in the case of coach services. Nevertheless, in the context of our present investigation it is important to note that the MMC did believe that coach competition did impact on rail fares.⁵⁹

The MMC's main conclusion was as follows:

Around 90 per cent of NEL's passengers are travelling for leisure purposes. Most of MML's passengers are travelling for business or commuting; about 40 per cent are travelling for leisure purposes. We conclude from the full range of evidence presented

⁵⁸ R.P. Kilvington and A. K. Cross *Deregulation of Express Coach Services in Britain* Oxford Studies in Transport, Gower, 1986.

⁵⁹ "Although we accept that rail services would seem to provide stronger competition to coach than coach services to rail, we believe that in the light of all the evidence summarised above coach services can also be regarded as providing an element of competition to rail services". Op. cit., para. 2.48, p.16.

to us that there has been an element of competition between NEL's coach services and MML's rail services for the leisure passenger. This competition has been lost as a result of the merger. Because of the absence of other strong constraints on fares or services we conclude that the merger may be expected to lead, over time, to higher coach fares or higher fares on both coach and rail,⁶⁰ and/or a lower quality of coach services or a lower quality of both coach and rail services, than would have been the case had the merger not occurred.⁶¹

The MMC's solution was to propose a set of behavioural undertakings relating to coach fares and levels of service (but not rail fares, since MMC believed that the undertakings on coach fares would be sufficient to deter MML from increasing rail fares).⁶²

B.4.1.2. ScotRail

One of two MMC reports published in December 1997 investigated the merger in which NEG acquired ScotRail Railways Ltd (ScotRail).⁶³ At the time, NEG's subsidiaries Scottish CityLink Coaches and NEL were the two leading providers of Scottish long-distance coach services and cross border coach services respectively, primarily serving the leisure market. ScotRail was the principal operator of passenger train services in Scotland, also providing sleeper services between London and Scotland. Therefore the investigation was concerned with the level of competition between coach and rail services, and what competitive constraints would remain on the merged firm, especially in the leisure market.

The majority of services provided by Citylink, NEL and ScotRail services overlapped and the most important of these overlaps concerned inter-urban routes and rural routes in the West of Scotland. On inter-urban routes ScotRail's services were faster but less frequent than Citylink/NEL coach services, while on rural routes coach services were faster than rail services while frequencies were similar. Price differentials varied by route, but generally coach fares were considerably cheaper than rail fares.

Although the Commission's terms of reference refer to the supply of public transport passenger services within Scotland, the effects of the merger on cross border traffic with England were also considered.

As in the MML report the chapter analysing the market first considers the market in which ScotRail and NEG were operating, looking at the provision of both coach and rail services. Following this, the levels of competition between rail and coach services was examined,

⁶⁰ Although Saver fares were regulated, the MMC indicated why they believed that unregulated rail fares used for leisure journeys could rise. Op. cit., para. 2.69, p. 21.

⁶¹ Op. cit., para. 1.5, p. 3.

⁶² Op. cit., para. 2.80, p. 23.

⁶³ MMC *National Express Group plc and ScotRail Railways Ltd: a Report on the Merger Situation* The Stationery Office, December 1997.

detailing the characteristics of those routes where ScotRail and NEG services overlap. Finally, the market for cross-border services was considered.

When investigating competitive constraints on the merged operation, the MMC did not expect other TOCs operating in Scotland, that is Virgin and GNER, to provide a material increase in competition in the short to medium term on the behaviour of a merged ScotRail. In the coach sector, although there was significant competition on several individual routes, the prospect of further entry was not a strong competitive restraint on coach prices or service standards for the NEL/Citylink network as a whole. Finally, although some scheduled air services exist on inter-urban routes, and may be attractive to some business traffic, the potential market within Scotland was small and competition between air and rail services was unlikely to be a major constraint, especially on ScotRail Leisure fares.

Given that ScotRail was the principal provider of passenger train services in Scotland and Citylink was the principal provider of long distance coach services, the services of the two companies had considerable overlap, with Citylink providing services at much lower fares. Consistent with the MML report, that an element of competition exists between coach and train passengers especially for leisure passengers, the Commission came to the conclusion that competition had been lost as a result of the merger, and that remaining competition from public and private providers of transport services would not be sufficient to remove the incentive for a merged company to increase prices.

The MMC's verdict was that the divestiture of Citylink by NEG, so that it may be preserved as a strong competitive force in the market, would be a proportionate remedy to the expected adverse effects of the merger.

B.4.1.3. Central Trains

Also in December 1997 the Commission investigated the merger in which NEG acquired Central Trains Ltd. At the time of the report, NEG also owned the leading bus company in the West Midlands, West Midlands Travel Ltd (WMT). Within the West Midlands area, Central Trains' services comprised of an urban transport network centred on Birmingham and operated on behalf of the West Midlands PTE (Centro). Additional services included inter-urban services centred on the Midlands and urban/rural services covering a wide area of rural England and into Wales.

Due to the strong market position of WMT in the provision of bus services in the West Midlands, a predominant issue was whether common ownership of WMT and Central trains by NEG would give National Express sufficient market power to increase fares or reduce service levels. This was assessed in two separate geographic markets, the first examining the market for public transport within the West Midlands and the second considering specific routes outside the West Midlands where NEG operates both rail and coach services.

The Commission considered the two geographical markets in a consistent manner, first analysing all existing transport provision, before focusing on the characteristics of the rail

and bus markets. Subsequently competition between different modes was looked at alongside the extent of provision of overlapping services by the merged company.

Within the West Midlands, as well as controlling WMT, NEG would be the operator of the new Midland Metro between Birmingham and Wolverhampton. However, the extent of competition in the West Midlands between bus and Central Trains' rail services was found to be limited (and the same would be true of the competition between the Metro and the rail services when the Metro starts operating). Additionally NEG is under an obligation to deliver the level of rail, bus and multi-modal services that Centro has specified in detail in their franchise agreement, and fares in the West Midlands are set by the PTE. Therefore, the MMC was satisfied that most of these possibilities for NEG to benefit unfairly in the West Midlands through the merger were addressed in their franchise agreement or would be deterred by a possible referral to the franchise director or DGFT.

On individual routes outside the West Midlands, the study of the market revealed that competition from other TOCs would be a larger constraint on behaviour than competition from bus or coach operators⁶⁴ and therefore coach operations were not expected to be a constraint on current market practice.

Overall, the MMC concluded that although the merger would increase NEG's ownership of passenger transport services in the geographical market, it was not expected to operate against the public interest, primarily because of:

- the low perceived substitutability between existing bus, Metro and rail services;
- the regulatory and contractual controls exercised over Central Trains; and
- the assurances that NEG gave in Schedule 14 of the Franchising agreement, which bears on bus, train and metro transport.

Therefore, NEG was not expected to exercise additional market power and the merger was not contested.

B.4.2. MMC investigations involving bus mergers

The MMC has carried out a large number of investigations concerning bus company mergers. In that involving FirstBus and SB Holdings in Glasgow⁶⁵ the Commission considered whether rail provided competition to bus services. FirstBus had argued that the extensive rail network operated on behalf of Strathclyde Passenger Transport Executive provided a highly effective competitor to bus services in SBH's geographical market. While

⁶⁴ Ian Jones' work on on-track competition, reviewed in section B.2.4 above, confirms that TOC competition **does** provide a constraining impact on TOC fares.

⁶⁵ MMC *FirstBus plc and S B Holdings Limited: a Report on the Merger Situation* The Stationery Office, London, January 1997.

FirstBus claimed that 69 per cent of total SBH mileage was in competition with rail services, the MMC was sceptical. “We doubt that very many of those who use the bus would necessarily see these as close substitutes. More generally, given the different characteristics of rail and bus travel, particularly relative frequencies, journey times, relative prices and the location of bus stops and stations, it does not follow that passengers regard the two modes as competitive alternatives, therefore switching between them on the basis of fares. Indeed we have seen no convincing evidence of material price competition between bus and rail.”⁶⁶ The implication is that, partly as a result of its time series comparison of bus and rail fares, the MMC did not view rail and bus as forming part of the same market in this case.

In 1994 the MMC considered NEG’s acquisition of Saltire, the operator of Scottish Citylink coach services between Scotland and London.⁶⁷ The MMC noted in this report, as in their later Midland Main Line report, that coach fares were set at least 30 per cent lower than comparable rail fares in order to compete with rail’s shorter journey times and, for most passengers, added convenience. The MMC agreed with National Express’s view that the main constraint on the ability to put up coach fares on the Scotland – London routes was the pricing policy of the British Railways Board, and so took the view that competition from rail services was the main constraint on National Express’s ability to raise fares on their routes as a result of the merger. Consequently, the MMC concluded that the merger between Scottish Citylink and its main competitor, NEL, would not be against the public interest.

B.4.3. Studies of predatory behaviour in the bus industry

The issue of market definition in the transport industry was considered in a series of studies that the OFT conducted on whether local bus operators were engaged in predatory behaviour. The OFT adopted a three-stage approach in all these investigations. The three stages involved:

- Whether the structure and characteristics of the market are such as to make predation a sensible and feasible business strategy;
- Whether the alleged predator incurs losses from this course of conduct; and
- The intentions of the alleged predator, including relevant evidence on its behaviour in other markets.

The first stage is concerned with market definition, since predation is only a “sensible and feasible business strategy” if the predator can recoup the losses from his actions through earning excess profits from the subsequent prices charged as a result of successful predation, ie it has market power.

⁶⁶ Op. cit., para. 2.42-2.44, p.14. See also para. 2.92, p.25, and paras. 4.62 – 4.68, pp.66-67.

⁶⁷ MMC *National Express Group plc and Saltire Holdings Ltd: a Report on the Merger Situation* HMSO, Cmnd. 2468, February 1994.

In the case of local bus services in Southend-on-Sea, the Director-General of Fair Trading concluded that the bus services formed a distinct market, and that rail did not impose restraints on bus operators' ability to set fares independently.⁶⁸ In contrast, in his parallel investigation of coach operations between Southend and London, the DGFT concluded that "rail and coach travel are part of the same market for commuters between Southend and London".⁶⁹

In the West Yorkshire case of bus services in the Bradford to Skipton area, bus and rail were not judged to be part of the same market to the bus services under investigation, even though there was a parallel rail route. This was because "rail transport is relatively inflexible by virtue of its limited pick up/set down points".⁷⁰ The same conclusion was reached in the South Yorkshire case.⁷¹

B.4.4. Market definition in cross-channel rail services

Market definition was considered in the European Commission's decision (Decision 94/663/EC) on the grouping of European railway operators to provide European Night Services (ENS) services, and the consequent appeal to the Court of First Instance.⁷²

The Commission's decision distinguished two relevant service markets:

- the market for the transport of business travellers, for whom scheduled air travel, high-speed rail travel and the rail services to be operated by ENS are interchangeable modes of transport; and
- the market for the transport of leisure passengers, for whom substitute services may include economy-class air travel, train coach and possibly private motor car.

The Commission confined these markets to the four routes actually planned to be served by ENS, namely London-Amsterdam, London-Frankfurt/Dortmund, Paris-Glasgow/Swansea and Paris-Glasgow/Plymouth.

In their appeal the railways estimated market shares in these business and leisure markets as 2.4 per cent and 5 per cent respectively, though, in its defence to the Court, the Commission agreed that the market share should be assessed only in relation to late evening and early morning flights.

⁶⁸ OFT *Thamesway Ltd: the Operation of Local Bus Services Commencing In, Terminating In, or Passing Through Southend-on-Sea* 27 August 1993, paras.2.2 and 2.3, p.7.

⁶⁹ *Op. cit.*, para. 2.13, p.9.

⁷⁰ OFT *West Yorkshire Road Car Company Ltd: Fares Policy on Certain Routes between Bradford and Skipton* 15 August, 1989, para.5.11, p.20.

⁷¹ OFT *South Yorkshire Transport Ltd: the Registration and Operation of Service 74 between High Green and Sheffield* 13 October 1989, para. 5.10.p.18.

⁷² Judgement of the Court of First Instance of 15th September 1998.

The Court judged that it has been necessary for the Commission to provide evidence on market shares in its decision, but that the contested decision does not contain sufficient statement of reason on this matter. This was because “whereas in the present case horizontal agreements between undertakings reach or very only slightly exceed the 5% threshold regarded by the Commission itself as critical and such as to justify Application of Article 85(1) of the Treaty, the Commission must provide” such an adequate statement of reasons. Further, “since the contested decision did not contain the relevant analytical data concerning the structure and operation of the market on which ENS operates, the degree of competition prevailing in the market or therefore the nature and extent of the alleged restrictions on competition “ the Court judged that the Commission was not able to assess whether an exemption should be granted under Article 85(3).

For these, and other reasons, the Court annulled the Commission’s Decision.

B.5. A Template for Excessive Pricing Cases

This section of the annex provides a relatively brief description of the information needed to meet the requirements to include market definition in any investigation of excessive pricing in the rail sector. The guidance in this section would need to be expanded considerably in providing advice on any specific case.

For the purposes of this exposition, we presume that such an investigation is concerned with a specific TOC, and covers all the services (rather than a subset) operated by that TOC. We also presume that the investigation covers all the fare types offered by the TOC.

The first step will be to identify the routes operated by that TOC, and to describe the types of services operated, in terms of frequencies, journey times and other service characteristics, as well as full details of the types of fares offered, their levels and their conditions of availability. Such a description will also serve to identify the geographical area covered by the market to be defined.

Information will also be needed on the characteristics of the passengers using the TOC’s services, since evidence shows that rail travel markets are usually segmented by type of traveller. The TOC will be able to supply information on numbers of journeys of different types, broken down by ticket type and by period of time. In addition, the TOC should be asked to supply any information on journey purpose that it has collected from its own on-board or other surveys. The TOC may also have undertaken some analysis in terms of studies of factors determining choice of travel, including whether to travel by rail or other modes, and choice of ticket type. There should be a request for brief details of such studies to determine whether full copies should be requested. It will also be necessary to consider whether any special surveys need to be undertaken as part of the investigation, and also whether any stated preference studies of traveller choice need to be commissioned to assess possibilities of travellers in different fare classes switching to other modes or ceasing to travel if rail fares rise.

Alternative methods of travel need to be reviewed, route by route. Such a review should consider:

- On-rail competition, both on the same rail route, and on competing routes. The potential for future on-rail competition to limit fares may be restricted since an incumbent TOC can always reduce fares once entry has occurred, as long as this response might not be judged to be predatory.
- Competing air services. These need to be identified, and the routes, fares, frequencies and access conditions to airports need to be described. Traffic volumes will also be required from the relevant airlines.
- Competing coach services. These also need to be identified, and the routes, fares and frequencies described. Coach traffic volumes also need to be estimated. Operators should have some information, though not necessarily in terms of end-to-end flows.
- Parallel road routes. The main alternative road routes, together with average journey times and congestion conditions should be described. Road traffic volumes for end-to-end flows are not generally available. However, in this context it is worth noting that previous market definition studies in the rail sector have not generally identified road as part of the relevant market. This issue should not be pre-judged, but road market share data would not be needed if road were not part of the relevant market.

The previous studies of own- and cross-price elasticities summarised in sections B.1 and B.2 above should be reviewed to give a first indication of elasticity evidence for the types of rail service and rail traveller in the current investigation. Cross-elasticities from other services should only be used insofar as they are adjusted to take account of differences in modal splits.

The TOC should also be asked to supply results from any of its own studies which have investigated the elasticities of demand for its own services.

However, care will have to be exercised to ensure that elasticity values from other studies are appropriate.

A further part of the market definition description should consider what has happened to fares since franchising. As we indicated in Section B.1, it was argued that certain fares did not need to be regulated because they would be constrained by competition from other modes. The issue now is that unregulated fares have risen faster than regulated ones, and so the implication is either that competition did not, as had been anticipated, constrain certain fares (which has implications for market definition) or that other factors caused these fares to rise. The main other factor in theory might be cost increases (since a firm facing competition would still increase fares to cover cost increases as long as it were producing a differentiated product). Consequently, the fares charged by the TOC under investigation

should be tracked. Cost factors impacting on this TOC can be compared with those impacting on other TOCs, to see whether fare increase imposed by the TOC, under investigation which are higher than those imposed by other TOCs can be explained by cost differences.

One quantitative technique sometimes used in market definition investigations is to compare price levels in different products.⁷³ If they are correlated, then the inference is that the products form part of the same market. However, we do not recommend use of these techniques here:

- fare differences on public transport modes will reflect quality differences, so fare levels will need to be adjusted to take account of changes in quality levels over time.
- determining overall fare levels will be a time-consuming process, given the fare discrimination employed by air, coach and rail operators, and selection of typical fare levels for specific types of traveller will also be difficult and to some extent arbitrary.
- given the role we have identified for market definition in excessive pricing investigations, we cannot see that such an investigation will add value to the conclusions on market definitions, given that we would in any case expect the results of this type of investigation to be inconclusive for the reasons identified above.

⁷³ LECG *Quantitative Techniques in Competition Analysis* OFT, Research Paper 17, October 1999, p. 54.