THE ECONOMIC AND ENVIRONMENTAL EFFECTS OF
LOCAL BUS SERVICE DEREGULATION IN BRITAIN

VOLUME I

Peter J. Romilly

A thesis submitted in partial fulfilment of the requirements of the University of
Abertay Dundee for the degree of Doctor of Philosophy.

October 2000

I certify that this thesis is the true and accurate record of the thesis approved by the
examiners.

Signed: [Signature]
(Director of Studies)

Date: 27 Oct 2000
# CONTENTS

## VOLUME I

### Table of Contents

**Acknowledgements** ................................................................. vii  
**Abstract** ................................................................................. viii

### CHAPTER 1 INTRODUCTION ......................................................... 1  
1.1 Background and Aims of the Thesis ............................................. 1  
1.2 Thesis Structure ......................................................................... 3  
1.3 Appendices ................................................................................. 6

### CHAPTER 2 REGULATION OF BUS SERVICES 1930-1985 ................... 7  
2.1 The Background to the 1930 Road Traffic Act ................................ 7  
2.2 The Provisions and Implementation of the Road Traffic Act 1930 ....... 15  
2.3 The Effects of the 1930 Road Traffic Act ..................................... 20  
2.4 The Post-war Bus Industry 1945-1979 ......................................... 24  
2.5 Fundamental Change .................................................................. 28  
2.5.1 Effects of the Transport Act 1980 ............................................ 28  
2.5.1.1 Express Service Fares and Welfare Changes ......................... 28  
2.5.1.2 Entry Deterrence .................................................................. 30  
2.5.1.3 Trial Areas .......................................................................... 31  
2.5.1.4 Licensing Changes ............................................................... 32  
2.5.1.5 Fare Controls ..................................................................... 33  
2.5.1.6 Summary ........................................................................... 33  
2.6 Conclusions ............................................................................ 33

### CHAPTER 3 THEORETICAL PERSPECTIVES I: MARKET FAILURE AND THE CASE FOR REGULATION ................................. 35  
3.1 Pareto Efficiency and Welfare ..................................................... 35  
3.2 Market Failure and its Causes ..................................................... 39  
3.3 Instances of Market Failure ........................................................ 42  
3.3.1 Monopoly ............................................................................ 42  
3.3.2 Externalities .............................................................. 45  
3.3.3 Wasteful Competition .......................................................... 54  
3.3.3.1 Natural Monopoly .............................................................. 54  
3.3.3.2 Product Differentiation ....................................................... 55  
3.3.3.3 Loss of Integration ............................................................ 63  
3.3.3.4 Unstable Competition ......................................................... 67  
3.3.3.5 Unfair Competition and the Problem of Second Best .......... 73  
3.3.4 Public Goods ...................................................................... 77  
3.3.5 Common Goods ................................................................. 81  
3.4 Summary and Conclusions ......................................................... 83

### CHAPTER 4 THEORETICAL PERSPECTIVES II: REGULATORY FAILURE AND THE CASE FOR A MARKET SOLUTION ....................... 85  
4.1 The Public Choice Approach ....................................................... 86  
4.1.1 Introduction .............................................................. 86  
4.1.2 Regulatory Capture ............................................................. 88  
4.1.3 Coalitions .......................................................... 89  
4.1.4 The Durability of the Regulated System ................................ 91  
4.1.5 Regulatory Failure ............................................................ 93  
4.2 The Austrian Approach to Market Efficiency ............................... 95  
4.2.1 Introduction .............................................................. 95
4.2.2 Monopoly, Mergers and Restrictive Practices. .................................................. 96
4.3 Contestable Markets ................................................................................................. 99
4.3.1 Introduction ......................................................................................................... 99
4.3.2 Natural Monopoly ............................................................................................... 100
4.3.3 Imperfect Contestability: the role of actual and potential competition .......... 102
4.4 Entry and Exit Barriers in the Bus Service Industry .............................................. 104
4.4.1 Entry Barriers: the Bain approach .................................................................... 104
4.4.2 Entry Barriers: alternative approaches .............................................................. 105
4.4.3 Sunk Costs ......................................................................................................... 108
4.5 Innocent Entry Barriers in the Bus Industry: the empirical evidence ................. 111
4.5.1 Economies of Scale ............................................................................................. 112
4.5.2 Economies of Density ........................................................................................ 115
4.5.3 Economies of Scope ........................................................................................... 118
4.5.4 Cost Economies, Network Effects and Mergers .............................................. 120
4.5.5 Economies of Experience .................................................................................. 123
4.5.6 The Extent of Innocent Sunk Costs in the UK Bus Industry ............................ 125
4.6 Strategic Entry Barriers ......................................................................................... 127
4.6.1 Excess Capacity .................................................................................................. 128
4.6.2 Raising Rivals' Costs ......................................................................................... 134
4.6.3 Predatory Pricing ............................................................................................... 137
4.7 Summary and Conclusions .................................................................................... 141

CHAPTER 5 DEREGULATION OF LOCAL BUS SERVICES .................................. 145
5.1 Local Bus Service Deregulation in Britain: the background ................................. 146
5.1.1 The Case for Deregulation ................................................................................ 146
5.1.2 The Case against Deregulation ......................................................................... 148
5.1.3 Privatisation ....................................................................................................... 154
5.1.4 Subsidy .............................................................................................................. 156
5.1.4.1 External Subsidy ............................................................................................ 156
5.1.4.2 Internal Subsidy ........................................................................................... 161
5.1.4.3 Legislation and Subsidy Control 1983 - 88 ................................................... 162
5.2 Effects of the 1985 Transport Act: a framework for analysis ............................... 163
5.3 Short to Medium Term Effects of the 1985 Transport Act ..................................... 167
5.3.1 Overall Changes 1985/86 - 88/89 ................................................................. 167
5.3.2 Vehicle Seating Capacity and Trends in Vehicle Size .................................... 171
5.3.3 Profit Margins .................................................................................................. 175
5.3.4 Return on Capital Employed ........................................................................... 177
5.3.5 Labour Productivity ......................................................................................... 179
5.3.6 Commercial and Subsidised Services ............................................................. 180
5.3.7 Competitive Tendering .................................................................................... 181
5.4 The Longer Term Effects of the 1985 Transport Act ........................................... 183
5.4.1 Changes in Market Structure .......................................................................... 183
5.4.2 Profitability ...................................................................................................... 184
5.4.3 Costs and Employment .................................................................................... 185
5.4.4 Vehicle Seating Capacity and Trends in Vehicle Size .................................... 187
5.4.5 Investment in the Bus Industry ....................................................................... 188
5.4.6 Market Entry and Contestability ..................................................................... 190
5.4.7 Subsidy ............................................................................................................ 191
5.4.8 Bus Travel Demand Factors .......................................................................... 192
5.5 Summary and Conclusions ................................................................................... 194

CHAPTER 6 SUBSIDY AND LOCAL BUS SERVICE DEREGULATION IN BRITAIN: A RE-EVALUATION ..................................................... 196
6.1 Introduction ............................................................................................................ 197
6.2 Data, Methodology and Model Specification ...................................................... 199
6.2.1 Data .................................................................................................................. 199
6.2.2 Methodology ................................................................................................... 200
6.2.3 Model Specification and Estimation Methodology ......................................... 202
6.3 Model Estimation ................................................................................................. 204
6.3.1 Estimation Methodologies .............................................................................. 204
List of Tables

2.1: Number of road vehicles and rail passengers in Britain: selected years 1904-34.............................. 8

3.1: Number of registrations accepted by traffic area offices........................................................................ 72

4.1: Greater Glasgow peak period buses (Union Street south of St. Vincent Street)........................................ 129

4.2: Number of complaints made to the OFT about anti-competitive practices by bus operators 1986-94 .................................................................................................................. 140

5.1: Bus subsidy to consumers (£ per capita) 1982/83 .................................................................................. 159

5.2: Short to medium term effects of the 1985 Transport Act....................................................................... 168

5.3(i): Vehicle seating capacity (000s) 1985/86.......................................................................................... 172

5.3(ii): Vehicle seating capacity (000s) 1988/89 ...................................................................................... 173

5.4: Costs, revenues and operating profit 1985/86-1988/89 .................................................................... 175

5.5: Registration of NBC services (000s bus miles) .................................................................................... 180

5.6: Changes in market share (%) by turnover in the bus industry 1989-95 .................................................. 184

5.7: Pre-tax profit (as % of turnover) by operator type 1988-94 ................................................................. 185

5.8: Average earnings of bus and coach drivers (£ per hour at April 1994 prices) ........................................ 186

5.9: Labour costs per employee (£) and employment changes ....................................................................... 187

5.10: Bus supply by bus and seat miles ....................................................................................................... 188

5.11: Capital investment in the bus industry (£m) 1989-94 .......................................................................... 189

5.12: Investment in new buses (number) 1988-94 ...................................................................................... 190

5.13: Subsidy per passenger (pence, current prices) by area ........................................................................ 192

5.14: Effects of bus travel demand factors 1987-93 ................................................................................... 193

5.15: Stagecoach fare decreases ................................................................................................................. 194

6.1: Deregulation (with and without subsidy reduction) and regulation: actual and forecast values for Britain excluding London ........................................................................................................ 217

6.2: Deregulation (with and without subsidy reduction): actual and forecast values for the English metropolitan areas ................................................................................................................. 219

6.3: Deregulation (with and without subsidy reduction): actual and forecast values for the non-metropolitan areas .................................................................................................................. 219

6.4: Welfare gains from deregulation (with subsidy replacement) (£m) ........................................................... 226

7.1: Urban exhaust emissions (gm per km) from buses and cars in Britain ................................................. 239

7.2: Changes in urban exhaust emissions (kg) from a reduction of 100 cars .................................................. 241
7.3: Emission costs from UK road transport 1991 and 1993 ............................................................... 244
7.4: Emission costs for the Los Angeles (1992, £000s per ton of emission)........................................ 245
7.5: Bus and car urban exhaust emission costs 1995: unit, net change and total costs. ....................... 248
7.6: Overall urban cost changes at different bus load factors (£000s)................................................. 260

LIST OF FIGURES

3.1: A comparison of perfect competition and monopoly................................................................. 42
3.2: Divergence between private and social cost.................................................................................. 46
4.1: A comparison of perfect competition and monopoly................................................................. 88
4.2: A hub and spoke system .............................................................................................................. 121
6.1: Bus fares with subsidy reduction, subsidy replacement and regulation........................................ 201
Acknowledgements

The author would like to thank the following people for their advice and assistance during the preparation of this thesis:

- Dr. Pamela A. Siler, Head of the Division of Economics, School of Social and Health Sciences, University of Abertay Dundee, for her role as Director of Studies;

- Dr. John Preston, Director and Reader in Transport Studies, Transport Studies Unit, University of Oxford, for his role as Chief External Supervisor;

- Dr. Haiyan Song, Reader in Tourism Studies, School of Management Studies for the Service Sector, University of Surrey, and Mr. Alan McGillivray, formerly Head of the Division of Economics, University of Abertay Dundee, for their roles as Second Supervisors;

- Mr. Graeme Fraser, formerly Head of Transport & Special Projects, Roads & Transport Division, Perth & Kinross Council, for his help in providing background information in the formative stages of this thesis;

- Dr. Martin Higginson, Economic Adviser, Confederation of Passenger Transport, London, for his encouragement in providing a grant from the CPT towards the costs of research.
Abstract

This thesis has three aims. First, to develop the theoretical perspectives necessary for the evaluation of local bus service deregulation in Britain (excluding London). Second, to identify the effects of deregulation if external subsidy to the bus industry had not been reduced. Third, to estimate a number of the externalities, particularly the environmental effects, associated with substitution of bus for car travel.

Criticism of local bus service deregulation has focused on the welfare loss arising from wasteful competition “on the road”. This view is based on the public interest theory of regulation, which assumes that market failures such as wasteful competition can be remedied by regulators acting to increase social welfare. But the public interest theory is not sufficient to provide an appropriate evaluation framework. Alternatives are public choice theory, Austrian theory and contestable market theory. The thesis provides a full discussion of these alternatives and their relevance to the local bus service market. It concludes that the local bus service market is relatively contestable, although the existence of strategic and non-strategic entry barriers implies a continuing supervisory role for the Office of Fair Trading and the Monopolies and Mergers Commission.

A major empirical problem in the evaluation of local bus service deregulation is that government macroeconomic policy caused significant reductions in external subsidy to the bus industry concurrent with deregulation itself. The thesis uses an econometric model to disentangle the long term effects of subsidy reduction and deregulation on bus fares and patronage for both the metropolitan and non-metropolitan areas. The focus on the long-term effects is an important addition to the research literature on local bus service deregulation, since previous quantitative studies have been able to deal with the short to medium-term effects only. The econometric model shows that, if subsidy reduction had not occurred, deregulation would have resulted in a decrease in bus fares and an increase in bus travel. The associated overall welfare gain is in the region of £6 to £10 million per annum for the period 1986-97.

Part of the increase in bus travel would come from people substituting bus for car travel. Given the well-publicised and researched external costs of car use, it is important to know the extent to which these costs are reduced as people transfer from car to bus. The thesis uses a simulation model developed by the author to provide a number of monetary estimates in this under-researched area. Particular attention is given to the environmental effects, that is air pollution and global warming, arising from bus for car travel substitution. The simulation model shows that, for bus load factors in the range of 35 to 55 per cent, the total cost reductions from changes in the externalities vary from £640 to £1,234 per annum (1995 prices) for every car ceasing to travel. Exhaust emission costs increase, however, because of the large amounts of particulate matter emitted by bus diesel engines. The policy implication is that bus for car substitution should be encouraged, but particulate emissions from bus engines reduced by improved engine design and the use of low-sulphur fuels.
VOLUME II: APPENDICES

APPENDIX 1 LEGISLATIVE AND STRUCTURAL CHANGES IN LOCAL BUS SERVICES 1947-1985 ................................................................. 1

APPENDIX 2 MEASURES TO PREVENT STRATEGIC ENTRY DETERRENCE ......................................................................................... 35

APPENDIX 3 EXHAUST EMISSION VALUATION: METHODOLOGICAL DETAILS ................................................................................... 45

APPENDIX 4 THE DATA SET ......................................................................................................................................................... 61

APPENDIX 5 AN ALTERNATIVE ESTIMATION STRATEGY ................................................. 74
CHAPTER 1

Introduction

1.1 Background and Aims of the Thesis

The deregulation of local bus services in Britain (excluding London) in October 1986 created considerable controversy. There was general agreement amongst those involved in the deregulation debate that the regulated system introduced by the 1930 Road Traffic Act needed reform, but considerable disagreement about the means of achieving this aim. This disagreement should not be surprising, given the unique characteristics of the local bus service market. Bus travel is, to use the economist's unfortunate terminology, an "inferior" good. Real disposable incomes in Britain in the post-war period increased substantially, with the result that car travel increased at the expense of bus travel. The decline in the demand for local bus services locked the industry into a spiral of falling revenue, increasing subsidies from the taxpayer, rising fares and further falls in demand.

Bus operators were also required by the Traffic Commissioners to provide a comprehensive service network of profitable and unprofitable services. For many people living on low incomes or in rural areas, bus travel is the only travel option. In effect, the industry had a public service obligation to meet as well as the usual commercial criteria. The Traffic Commissioners overcame this dilemma by granting bus operators territorial monopolies from which they could cross-subsidise from profitable to unprofitable services. The basic problem with this system, at least as administered by the Traffic Commissioners, was that it stifled competition and innovation. Operators wishing to start a service in competition with an incumbent operator found it very difficult to obtain the requisite Road Service Licence.

From the economist's viewpoint, however, the debate is not just about whether local bus services should be provided on a competitive or non-competitive basis. There are wider
issues that should also be addressed. A key issue is the theoretical perspective from which competition (or the lack of it) is evaluated. A number of arguments against deregulation centre on the welfare loss from wasteful competition. If competition is viewed as a process, however, the concept of "wasteful" competition becomes problematic. Similarly, traditional approaches that measure the extent of competition solely in relation to the structure of the industry may not be appropriate if the market is contestable. There is also a practical issue to resolve. Fare increases after deregulation were not caused primarily by a lack of competition but, as this thesis argues in subsequent chapters, by the very large reductions in external subsidy to the industry. Any attempt to evaluate "deregulation" must therefore disentangle the effects of deregulation per se and those of subsidy reduction.

Deregulation was also accompanied by privatisation of the industry. This was not privatisation en bloc, but a process of splitting up the larger companies into smaller ones. Subsequently, a number of companies merged to create an industry that was oligopolistic in structure. This fragmentation and subsequent merger activity raises the issue of what happened to economies of scale, scope and density in the industry. This is important not only in terms of company cost structures but also in terms of market contestability, since scale and other economies may create entry barriers.

The local bus service market is subject to a number of externalities. Attention has tended to focus on the Mohring effect, where increased service frequencies generate external benefits for passengers. Less attention has been given to the externalities associated with the substitution of bus for car travel that might have occurred if deregulation had not been accompanied by subsidy reduction. These externalities include the costs of congestion, exhaust emissions, noise, fuel and road accidents. The monetary valuation of exhaust emission costs is particularly difficult in view of the different types of exhaust emissions from different types of engine, and the interaction effects between these emissions.

Accordingly, the aims of this thesis are to:
• Develop the theoretical perspectives necessary for the evaluation of deregulation
• Establish the outcome of deregulation without subsidy reduction
• Estimate the externalities associated with substitution of bus for car travel.

1.2 Thesis Structure

The thesis develops the theoretical perspectives in the context of the chronological development of the local bus service market. This historical approach is adopted because it is sometimes easier to understand the rationale for policy changes if one has some knowledge of the preceding policy regime. Policy is rarely made in a vacuum, and policy evaluation must take account of the economic and non-economic constraints faced _ex ante_ by policy makers.

Chapter two describes the development of the highly regulated system inaugurated by the 1930 Road Traffic Act, and the three principles of priority, protection and public need that guided the Traffic Commissioners in their administration of this system. The way in which the Commissioners interpreted the principle of public need was particularly important since it lead to the provision of profitable and unprofitable services, the latter cross-subsidised from the monopoly profits earned on the former. This system had to operate against the backdrop of rapidly increasing car use and decreasing fare revenues, so that cross-subsidy had eventually to be augmented by rapidly increasing external subsidy. The 1980 Transport Act deregulated the express coach or long-distance market, and introduced deregulation of local bus services in a small number of trial areas. It marked the beginning of the transition to further deregulation of local bus services under the 1985 Transport Act.

Chapter three discusses the theoretical rationale for the regulated system described in Chapter two. This rationale is based on the public interest theory of regulation, which assumes that the effects of market failure can be remedied by regulators acting to increase social welfare. The concept and measurement of social welfare is discussed in the context
of Pareto efficiency. Market failures in local bus service provision are identified and discussed, with particular attention given to the idea of wasteful competition.

Chapter four discusses the deficiencies of the public interest theory of regulation, and introduces alternative theoretical frameworks within which to analyse the local bus service market. These alternatives are public choice theory, Austrian theory and contestable market theory. The idea of a contestable market is particularly important, since it shifts the emphasis away from market structure towards the existence of entry and exit barriers. If these barriers do not exist, then the re-emergence of local monopolies, together with increased market concentration via mergers in the deregulated bus industry, is not incompatible with a competitive market outcome. The Buses White Paper (1984) claim that the industry is highly contestable is given extensive discussion, since this claim is central to the policies of deregulation and privatisation introduced in the 1985 Transport Act.

Chapter five moves the thesis away from theoretical perspectives and towards the specific background and actual effects of the 1985 Transport Act. These effects differed substantially between the metropolitan and non-metropolitan areas. One reason for this difference is the relatively large reduction in external subsidy that occurred in the metropolitan areas. A number of performance indicators are analysed to illustrate the different outcomes in these areas. These outcomes are set within an analytical framework that distinguishes between the short, medium and longer-term effects of deregulation. A number of short to medium term issues specific to local bus service deregulation in Britain are discussed. These include changes in market structure, profitability, costs and employment, vehicle seating capacity and size, investment, contestability and subsidy.

The issue of subsidy is crucial to the evaluation of local bus service deregulation. If subsidy had not fallen so sharply, the long run outcome of deregulation might have been substantially different. Chapter six develops and estimates an econometric model to identify what might have happened in this counter-factual situation, as well as the possible outcomes associated with a continuation of the regulated system. An important
issue to be resolved is the direction of causation between fare increases and subsidy reductions: did fare increases enable subsidies to be reduced, or did subsidy reductions force operators to increase fares? The model is estimated on time series data for 1953 to 1997, and thus allows for the long run deregulation effects identified by the analytical framework developed in the previous chapter. Time series estimation has become increasingly complex because of the desire to avoid the problem of spurious regression when dealing with trended variables. This problem can be reduced by the use of cointegration methodology, which forms the basis for the estimation procedures in Chapter six. The estimation of the long run demand curve for bus travel enables a welfare evaluation of deregulation with subsidy replacement to be made.

The results of Chapter six imply that deregulation with subsidy replacement would have caused an overall reduction in bus fares over the period 1985 to 1997. Although this decrease is small it is not unreasonable to assume that bus travel would have increased, and that some of this increase would have occurred as a result of car users transferring to bus. Since car use creates a number of external costs, bus for car substitution will generally increase social welfare because of the reduction in these costs. Chapter seven identifies and places a monetary value on these costs. The case of exhaust emission costs is particularly interesting, since these costs depend not only on monetary valuations of the direct effects, but also on interaction effects, load factors, type of vehicle and emission standards. The treatment of the external effects from bus for car substitution is an area that appears to be under-researched, and the results of Chapter seven, whilst certainly not providing a definitive analysis, do at least open up new avenues for research. These results are not just relevant to the area of bus deregulation, but also to the continuing public debate about the benefits of bus as opposed to car travel.

Chapter eight provides an overview of the main conclusions from Chapters two to seven.
1.3 Appendices

Appendices are contained in a separate volume. Although not essential to the thesis arguments developed in Volume I, they provide further details of the local bus service market that may be of interest to the reader. Appendix one provides an account of the legislative and structural changes in the regulated local bus service market from 1947 to 1985. Appendix two contains discussion of the author’s suggestions regarding the prevention of strategic entry deterrence in the form of schedule matching and predatory pricing in the deregulated market. The details of the methodological procedures used to derive the monetary valuations of exhaust emissions contained in Chapter seven are given in Appendix three. This also contains an example of the spreadsheet model CarBus developed by the author to facilitate the calculations of the monetary valuations. Appendix four presents the details of the variables used in the econometric analysis of Chapter six, together with the observations for the data set from 1953 to 1997. Finally, Appendix five contains the results of an alternative econometric analysis to the one used in Chapter six. The purpose of this alternative modelling strategy is to use Granger causality tests to establish the direction of causation between bus fares and external subsidy.
CHAPTER 2

Regulation of Bus Services 1930-1985

Before 1930 regulatory control of the omnibus service industry was fragmentary. The industry developed quickly, as did the number of accidents associated with it. The 1930 Road Traffic Act introduced systematic quality and quantity regulation of the industry, and this system survived virtually intact until the 1980 and 1985 Transport Acts. This chapter describes the background to the 1930 Road Traffic Act, and the ways in which the regulatory system developed under the administration of the Traffic Commissioners. The reasons for the eventual demise of this system are discussed, together with the effects of the partial deregulation introduced by the 1980 Transport Act. Section 2.1 discusses the background to the 1930 Road Traffic Act, in particular the concerns over safety and the subsequent widening of these concerns to cover unfair competition. The provisions and implementation of the 1930 Act are described in Section 2.2, especially the way in which the Traffic Commissioners interpreted the principle of "public need" to provide profitable and unprofitable services through internal cross-subsidy. The effects of the 1930 Act are discussed in Section 2.3, whilst Section 2.4 provides an account of the development of the post-war bus industry over the period 1945-79. Section 2.5 describes the changes to the industry introduced by the 1980 Transport Act, which deregulated coach services and introduced partial deregulation for local bus services in the form of the Trial Areas.

2.1 The Background to the 1930 Road Traffic Act

The dominant feature of the transport sector in Britain in the early twentieth century was the rise of road motor transport, both passenger and freight. Its evolution started with the development of the first internal combustion engine, attributed to Jean Lenoir of Paris in 1860, and progressed to the announcement by Henry Ford in 1909 that he intended to produce a motor car for the masses - the model "T". In Britain the Motor Car Act of 1903
gave legal recognition to the motor car, introduced a maximum speed limit of 20 miles per hour, and set regulations on licensing, driving standards, vehicle identification and maximum weight.

The first motor omnibuses appeared in Britain about 1898 and, despite suffering from mechanical unreliability, expanded rapidly in number, particularly in London. In 1905 there were twenty motor buses in operation in London, by 1911 there were 3000. Road freight transport also expanded rapidly, particularly after the First World War, when a large increase in railway rates in September 1920 coincided with the sale of some 20,000 surplus vehicles at very low prices by the Ministry of Munitions. The growth of the railway industry, virtually uninterrupted throughout most of the nineteenth century, slowed and was then reversed by the competition from road motor transport. From 1923 to 1930 railway receipts fell by 17%. A similar story applies to the tramway companies.

The motor omnibus industry developed in London in the early 1900s, in other cities such as Birmingham from about 1912, and in rural areas from the early 1920s. The structure and performance of the industry during this period is difficult to establish precisely, since before 1930 there were no adequate statistics available for buses and coaches classified according to size of operator. The overall trends in the road and rail sectors are summarised in Table 2.1:

**Table 2.1: Number of road vehicles and rail passengers in Britain: selected years 1904-34**

<table>
<thead>
<tr>
<th>Year</th>
<th>Hackneys, Buses, Coaches (a)</th>
<th>Goods Vehicles</th>
<th>Private Cars</th>
<th>Railway Passengers (millions) (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1904</td>
<td>5345</td>
<td>4000</td>
<td>8465</td>
<td>1142.3</td>
</tr>
<tr>
<td>1914</td>
<td>51167</td>
<td>82000</td>
<td>132015</td>
<td>1199.3</td>
</tr>
<tr>
<td>1924</td>
<td>94153</td>
<td>203156</td>
<td>473528</td>
<td>1236.2</td>
</tr>
<tr>
<td>1934</td>
<td>45689</td>
<td>413320</td>
<td>1308425</td>
<td>829.7</td>
</tr>
</tbody>
</table>

Source: Basic Road Statistics (various), British Road Federation, columns 2, 3 and 4. Railway Returns (various), Railway Companies (Returns and Accounts Act 1911) column 5.

Notes: (a) Figures from 1904 - 1924 arc for hackneys, and include taxi-cabs. The figure for 1934 is for buses and coaches, including trolleybuses.
(b) Figures exclude season ticket holders: the first two figures shown arc for 1900 and 1913 respectively, and arc for Britain; the remaining two arc for the UK.
The rapid growth of road traffic was accompanied by a rise in the number of road accidents, some of which were associated with omnibus travel. There is evidence to suggest that conduct in the industry, particularly in the early phase of development, was characterised by some dubious practices. In a contemporary account of services outside London and the municipalities, Garcke (1923) distinguishes four types of omnibus operator: firstly, the provincial owner with a sizeable fleet and offering a stable set of services; secondly, smaller companies, perhaps owned on a family or individual basis, allegedly seeking short term profits which were often at the expense of customer goodwill; thirdly, owners pursuing "blackmail" tactics by running services in the territory of an established operator with the intention of forcing the operator into a takeover; finally, "pirate" owners who took advantage of established operators' service networks by running buses just ahead of their scheduled times.

Competition in the rapidly expanding omnibus industry also resulted in "on-the-road" problems such as "racing", where rival drivers would speed against each other in order to reach a bus stop first, and "nursing", where one bus would closely accompany a rival bus in order to divert passengers away from it. A variant of this was "blanketing", where two buses would nurse the rival bus, one in front and one behind. If the blanket were tucked tightly enough, the rival bus would be unable to pull out and overtake. Passenger demand in the pre-1930s industry tended to be very dense, so that the number of buses on a particular route was usually very high: this gave considerable opportunity, at least by contemporary standards, for these practices. Birch (1971) provides an account of London omnibus operations in the 1920s, and Jenson (1971) describes the development of omnibus services in Birmingham in the latter half of the nineteenth century. An indication of the problems associated with the development of the (horse drawn) omnibus is given in the following extract from the 1836 prospectus of the Midland Omnibus Company:
"The glaring misconduct of both drivers and conductors together with the terrifying effect and often fatal consequences arising from careless and furious driving, prevent many (especially females) from availing themselves of those facilities which, otherwise, they might feel desirous to embrace..."

(Jenson, op. cit., p. 115)

Occasionally competition manifested itself in more overtly dangerous forms: it was not unknown, for example, for drivers to attempt to force their rivals into a convenient roadside ditch (Foster and Golay, 1986). Although the extent of these practices is open to dispute, the number of accidents involving omnibuses rose rapidly between 1918 and 1930 (Glaister and Mulley, 1983). The problem of dangerous driving was not confined solely to omnibuses, however, but to all vehicles, especially in London. In the minutes to the Royal Commission on London Government (1923) there are numerous references to driving behaviour which, by current standards, would probably be illegal. Compounding the general lack of control on driver behaviour was the growing problem of urban road congestion, caused by the increase in freight as well as passenger traffic.

Prior to 1930 local authorities were empowered to regulate omnibuses under the Town Police (Clauses) Acts of 1847 and 1889. The first Act applied to hackney carriages, whilst the second extended these powers to (horse-drawn) omnibuses. The provisions of these Acts were also applied to the motorised omnibuses introduced in Britain at the turn of the century, but they were not amended to take into account the greater speeds and distances of these motorised vehicles. Local authorities could license vehicles, drivers and conductors, and pass bye-laws to regulate vehicle safety standards, stopping places, and the conduct of owners and their omnibus crews.

Although the potential for regulatory control was considerable, there was no statutory obligation on the local authorities to use these powers. Consequently, there was little uniformity in the extent to which local authorities exercised them. Implementation was complicated by the exemption of privately hired vehicles from the scope of the Acts, so
that vehicles which were unable to satisfy the conditions for omnibus licensing could simply be transferred to alternative uncontrolled uses in the private hire trade. Additionally, vehicle licensing applied only to the vehicle itself, so that no restrictions could be placed within the licence on fares or service routes and frequencies.

Rural authorities were also required to obtain a special Order under section 276 of the Public Health Act 1875 before they could use their powers under the Town Police Clauses Acts: this probably explains why, in the great majority of rural areas, it was possible to start an omnibus service without any licence at all. Of the total 644 rural district authorities, only 65 possessed licensing powers. Where authorities did exercise their licensing powers, there was no requirement to observe impartiality: Birmingham Corporation, for example, reached an agreement with the Midland Red Company in 1914 whereby its own buses were given a monopoly of local services within Birmingham, whilst those of Midland Red were given a near monopoly of services from outside into the city. An account of the pre-1930 licensing system is given in the Report of the Committee on the Licensing of Road Passenger Services ("Thesiger" Report, 1953). Overall, it appears that the degree of competition varied considerably from place to place: in some areas freedom of entry was complete, and competition was sometimes fierce; in others, freedom of entry was impossible and competition non-existent.

Another dimension to the problem was the attitude of the railway companies, who frequently argued that road users were not paying the full costs of road use, in contrast to rail users. The growing competition between these two modes was thus unfairly biased in favour of road transport. This argument was supported by the cost to revenue ratio for road finances in the period up to 1930: in 1911 total government spending on the road network was some fourteen times higher than its revenue from fuel tax and licences, although by 1930 the ratio had fallen to 1.6:1. From 1933 onwards, however, revenues began to exceed costs. The question of inter-modal equity was examined in the Report on the Conference of Rail and Road Transport ("Salter" Report, 1932). Meanwhile the government had attempted to assist the railway companies by the Railway (Road Transport) Acts of 1928, which gave them powers to operate their own road services. By
1928, however, the Conservative government had decided to appoint a Royal Commission on Transport, under the chairmanship of Sir Arthur Griffith Boscawen, with the following terms of reference:

"To take into consideration the problems arising out of the growth of road traffic and, with a view to securing the employment of the available means of transport in Great Britain (including transport by sea coastwise and by ferries) to the greatest public advantage, to consider and report what measures, if any, should be adopted for their better regulation and control, and, so far as is desirable in the public interest, to promote their co-ordinated working and development."

(Royal Commission on Transport, Final Report, 1930, Cmnd. 3751, p. vii)

The Commission published three reports between 1929 and 1931. The second report of the Royal Commission was published in October 1929 and its recommendations formed the basis of parts IV and V of the 1930 Road Traffic Act. In its second report, the Commission unanimously came to the conclusion that:

"... as modern road passenger transport has grown to such importance ... the present chaotic system of licensing must disappear and be replaced by an entirely different system more suited to present-day needs".

(Royal Commission on Transport, Second Report, 1929, Cmnd. 3416)

The 1930 Road Traffic Act inaugurated a comprehensive and remarkably durable regulation of the British bus service industry. The reasons given for the quantity licensing aspect of this piece of legislation have generally been the prevention of wasteful competition and pressure from railway and tram interests to protect their markets from motor vehicle competition. Mulley (1983) has questioned this view, arguing that the original cause was purely concern for public safety beginning in the early 1920s. This reinterpretation stems from the availability of new sources of information, namely the
original Ministry of Transport files of both the Hackney Vehicle Committee and the Royal Commission, as well as the Cabinet Papers of the 1920s. Whilst examining the case for regulation on safety grounds, the discussions were subsequently extended to cover "unfair" competition, both within the road transport sector and between it and the railway and tram interests. The progression from quality to quantity licensing occurred during one of the Sub-Committee meetings of the Hackney Vehicle Committee (convened in 1922), where it was suggested that quantity licensing should be used to compensate incumbent operators for the expense of higher safety standards by protecting them from competition. The rationalisation of quantity licensing on the grounds of "unfair competition" came from later submissions.

Mulley (1983) argues that the Ministry of Transport's evidence to the Royal Commission was heavily influenced by the report of the Hackney Vehicle Committee, which in turn was considerably influenced by Frank Pick who represented the interests of the larger incumbent omnibus operators. The evidence to the Commission about the undesirability of competition was based on:

"... the predominant fallacy of the era - that "unfair" and "free" competition were synonymous."

(Mulley, 1983, p. 7)

The resulting legislation, based on quantity as well as quality licensing, tended to protect these vested interests rather than promote the public interest, which would:

"... undoubtedly have benefited from the wider choice of services under the unregulated system at lower fares than might otherwise have been offered."

(Mulley, 1983, p. 16)

Pursuing this line of argument, Mulley observes that:
"No real evidence was put before the Royal Commission ... to justify the quantity licensing subsequently incorporated in its second report and which formed the basis of the Road Traffic Act."

(Mulley, 1983, p. 17)

and concludes that:

"The most remarkable aspect of the 1930 quantity licensing legislation is the way in which it lasted, despite the faulty premises underpinning it, almost without change for fifty years."

(Mulley, 1983, p. 17)

Presumably the "real evidence" referred to would be evidence based on an economic analysis of competition, particularly the benefits of competition from the smaller operators. Whilst the argument that the Commission did not hear such evidence may be correct, it would be dubious to conclude that it was unaware of the economic debate over the costs and benefits of competition. Sir Arthur Boscawen was familiar with American legislation against unrestricted competition and the reasons for it, as an earlier quote demonstrates (p. 12). Academic debate on unrestricted competition, particularly in the USA, was well advanced. Hotelling's classic article on the problems associated with competition, for example, was published in 1929. Whilst the members of the Commission may not have been practising economists, a lack of contact with the academic literature would not preclude them from a general knowledge of the economic issues involved.

Nonetheless, the success of the incumbent operators in extending regulation from safety to entry conditions does indicate that regulatory "capture" (see Section 4.1.2) occurred at an early stage in the deliberations of the Royal Commission. This early capture is not surprising, given the potential rewards to incumbent operators from a system of quantity licensing. As Yarrow (1986) has observed:
"... while it is possible to conceive of institutional arrangements that impede capture by interest groups, the point at which the regulatory process is most susceptible to these pressures is when institutions and policies are first being established. At that moment, politicians are necessarily the decision makers and the stakes are high for all concerned."

(Yarrow, 1986, p. 349)

The Thesiger Report (1953) later identified the main objectives which the architects of the 1930 Road Traffic Act sought to achieve in a licensing system for road passenger services: firstly, licensing should be service, rather than vehicle, based; secondly, road passenger services should be classified into types, with no competition between types; thirdly, the licensing authority should be impartial; fourthly, a licensing area should be large enough to allow for a review of public transport as a whole; fifthly, municipal authorities should not have to seek special powers from Parliament whenever they wished to make small changes to their services; finally, there should be uniformity of practice between the different licensing authorities. The next section describes how these objectives were translated into the 1930 Road Traffic Act.

2.2 The Provisions and Implementation of the Road Traffic Act 1930

The Act replaced the existing licensing legislation and divided Britain into thirteen Traffic Areas. There were eleven areas for England and Wales (reduced to ten on January 1st 1934) and two for Scotland. Each Area had three Traffic Commissioners (apart from the Metropolitan Area, which had one). The Chairman was appointed by the Minister of Transport and was a full-time, paid official. He was assisted by two part-time officials who were appointed by the Minister of Transport from a list of names supplied by local authorities. There were three types of licence which the Commissioners were empowered to grant: the Public Service Vehicle (PSV) Licence, the Drivers' or Conductors' Licence, and the Road Service Licence.
Every public service vehicle was required to have a Certificate of Fitness and a PSV licence. The former was granted subject to the vehicle satisfying, after an inspection by a qualified examiner, detailed requirements under the Act relating to its construction and suitability. Once the Certificate of Fitness had been obtained, the PSV licence could be issued subject to the vehicle owner being a "fit and proper" person to hold such a licence. Details of these procedures are provided by Chester (1936). Drivers were required to pass a driving test before gaining their Drivers' Licence. The Act also placed a limit on the number of hours which drivers of buses and coaches (as well as goods vehicles) could remain continuously on duty, and incorporated a "Fair Wages Clause" for bus crews.

Most services on which separate fares were charged required a Road Service Licence (RSL). By granting or withholding these licences the Traffic Commissioners could regulate entry into the bus service industry. An RSL was required for stage carriage and express carriage services (where the legal distinction between them was that the latter charged a minimum fare of one shilling or more), excursions and tours. Contract carriage services did not require a road service licence: these were services where either no separate fare was paid or, if it was, the vehicle was used to transport a private party on a special occasion. This definition, however, was unclear in the 1930 Act and, in spite of clarification in the 1934 Road Traffic Act, remained difficult to interpret in a legal context (Thesiger Report, 1953, Chapter X and Appendix II). In the granting of a licence, the Act required the Traffic Commissioners to take into account:

(a) the suitability of the routes on which a service may be provided under the licence;

(b) the extent, if any, to which the needs of the proposed routes or any of them are already adequately served;

(c) the extent to which the proposed service is necessary or desirable in the public interest;
(d) the needs of the area as a whole in relation to traffic (including the provision of adequate, suitable and efficient services, the elimination of unnecessary services and the provision of unremunerative services), and the co-ordination of all forms of passenger transport, including transport by rail.

(Road Traffic Act, 1930, Part IV, Section 72 (3))

They were also given wide powers to attach conditions to the licence. The nature of these conditions was determined by the four criteria above and additionally that:

(a) the fares shall not be unreasonable;

(b) where desirable in the public interest the fares shall be so fixed as to prevent wasteful competition with alternative forms of transport, if any, along any route or any part thereof, or in proximity thereto;

(c) copies of the time-table and fare-table shall be carried and be available for inspection in vehicles used on the service;

(d) passengers shall not be taken up or shall not be set down except at specified points or shall not be taken up or shall not be set down between specified points; and generally for the securing of the safety and convenience of the public.

(Road Traffic Act, 1930, Part IV, Section 72 (4))

The application for a road service licence was heard by the Traffic Commissioner at a public sitting in a semi-judicial Traffic Court. The applicant was normally expected to provide evidence of public need for the service. Objections to the application, as well as other representations, were also heard in the Court. Where objections were raised, the onus of proof lay with the applicant. Most applicants employed legal experts to present their case, often provided by their Trade Association. It was customary (although not
statutory) for road service licences to be granted for three years, after which a new application had to be submitted. The licence could be revoked or suspended if the Commissioner judged that a serious breach of the conditions attached to it had occurred.

Appeals against the Commissioners' decisions could be made to the Minister of Transport, who would normally appoint an Inspector to conduct a hearing into the appeal. The Inspector would then provide a report to the Minister, who would take the final decision as to whether the appeal be upheld or not. The Inspector's task was to judge, on the basis of the evidence heard by the Commissioners, whether they had reached the right decision. Accordingly, new evidence was only allowed in exceptional circumstances. The close relationship between the Minister and the Commissioners, together with the difficulty of introducing new evidence, provoked criticism of the appeals procedure on the grounds that it contained a bias towards the rejection of appeals. Between 1931 and 1934, the Minister upheld nearly 40% of appeals, and rejected nearly 40%. Some 23% of appeals were withdrawn prior to a decision, although it could be argued that these were in anticipation of a rejection (Chester, 1936, p. 80).

In granting the licences, the Traffic Commissioner had to take into account a number of local conditions that were specific to that service. Consequently it was difficult to formulate general principles within the Act which could guide the Commissioner in specific decisions. Nevertheless, given the decisions of the Commissioners and the appeal hearings in the years immediately following the Act, Chester (1936) identified three principles on which these decisions appeared to be based: priority, protection, and public need.

The principle of priority meant that existing operators of a service were given preference to new applicants in the granting of a licence, particularly if they could demonstrate a satisfactory record in the running of the service. There were problems in the implementation of this principle in the years immediately following the Act because of the large number of operators who had set up in business shortly before the Act became effective. Where more than one operator had been running a service, and where
the Commissioner considered that all of them had an equal claim to running the service, the solution was to allow each operator to continue to provide the service on a rota basis. In general, however, the Commissioners adopted the approach of "one operator, one route". This rationing device was particularly important in the 1930s since the Great Depression was causing a fall in the demand for bus services and a consequent over-capacity in the industry.

The principle of protection was applied at an intra and inter-industry level. Within the industry, protection was given to local bus services against long distance services. This was usually in the form of restrictions on picking up and setting down points for long distance services operating within the territory of a local service, or the stipulation of a (relatively high) minimum fare for the local section of the long distance service. Local services were also protected against excursion services, usually by fare, scheduling or vehicle number restrictions on the excursion service. Outwith the industry, protection was given to alternative forms of transport, namely local tramways, trolleybuses and railway services. Long distance railway services, for example, were usually protected against competition from long distance express coach services by restricting their duplication at peak time periods of travel. The inter-industry protection enabled the Commissioners to fulfil their duty under section 72 of the 1930 Act to consider the "co-ordination of all forms of public transport".

The principle of public need came directly from section 72 of the 1930 Act. Applications could be divided into three categories: the continuation of an existing service, the alteration of an existing service, and the provision of a new service. In the first category an operator had to provide statistical evidence of vehicle loadings, costs and profits. In the absence of other factors, poor vehicle loadings or unprofitability might be interpreted by the Commissioners as a lack of public need for the service. For the other two categories such statistical evidence might be more difficult to provide: instead the operator could gather support from other sources such as local residents, ratepayers' associations, local authorities and local societies in order to prove public need. Proof was
needed before the licence was granted, and in all applications the Commissioners were required to consider the implications for alternative forms of transport.

Although there was no statutory requirement on them to do so, the Commissioners later commonly interpreted "public need" to mean that operators should provide unprofitable as well as profitable services (Thesiger Report, 1953, p. 39). Operators were granted a road service licence on condition that they supplied both types of service, using part of their profits to provide a subsidy to the loss-making parts of the network. In this way the practice of internal cross-subsidisation became a fundamental part of the post-1930 bus industry. In order to prevent competition from eroding profits and undermining the system of cross-subsidisation, operators were usually given a territorial monopoly by the Traffic Commissioners. The system had the advantages of simplicity and cheapness, since the operator was in effect used as a surrogate tax collector.

2.3 The Effects of the 1930 Road Traffic Act

The quality licensing in the 1930 Act secured an improvement in safety and operating standards that was generally welcomed by both operators and passengers. The average age of PSVs fell as older vehicles were less likely to obtain a Certificate of Fitness, and driving standards improved. In the first years after the Act, some 25% of applicants failed to obtain a driving licence (Chester, 1936).

Changes in the structure of the industry as a result of quantity (i.e. road service) licensing are more difficult to assess because of the lack of comprehensive data on buses and coaches before 1930 and the effect of the Depression on the industry. Between 1930 and 1933 there was a continuous fall in the number of buses and coaches in Britain, from 52,648 to 45,656. The number of PSV operators fell from 6,486 in 1931 to 4,789 in 1937. This was mainly as a result of the 1930 Act, which forced less reliable operators out of business as well as creating entry barriers for potential operators, although the Depression was also a contributory factor. The average number of vehicles per operator rose from 7.1 in 1931 to 10.4 in 1937. Over the same period, the percentage of buses and coaches
owned by operators with less than 10 vehicles fell from around 29 to 22, whilst the percentage owned by operators with 100 or more vehicles rose from around 47 to 61 (Savage, 1966).

The increase in industry concentration after 1930 occurred because many small operators went out of business or were taken over by larger firms, and also because of the increase in amalgamations between the larger firms. This tendency towards large scale operations cannot be attributed solely to the 1930 Act, since it provided protection to both small and large operators. Arguably, without the Act, the tendency would have been even more pronounced. Although amalgamation was sometimes encouraged by the Commissioners as a means of promoting the co-ordination of services, there were no provisions in the Act which allowed them to enforce amalgamation. The railway companies played a significant part in this process, since they used their powers under private Acts to buy a financial interest in some 40% of the largest bus companies, and then implemented a policy of amalgamation and rationalisation (Chester, 1936). Another factor was the London Passenger Transport Act of 1933, which created the London Passenger Transport Board. The Board was an amalgamation of all the transport agencies in the London area. Prior to the creation of the Board, passenger transport services were provided in its area by a combination of (at least) 5 railway companies, 14 municipally owned tramway undertakings, 3 company-owned tramway undertakings, and 61 companies, firms or individuals running bus services (Savage, 1966). The purpose of the Board was:

"To secure the provision of an adequate and properly co-ordinated system of passenger transport for the London Passenger Transport Area ... and ... while avoiding the provision of unnecessary and wasteful competitive services, to take from time to time such steps as they consider necessary for extending and improving the facilities for passenger transport in that area in such a manner as to provide most efficiently and conveniently for the needs thereof."

(London Passenger Transport Act, Section 3)
The process of amalgamation and takeovers created a situation where route monopolies conferred by the Traffic Commissioners were converted into territorial monopolies. By 1939 over two-fifths of the British bus industry was controlled by three groups: Thomas Tilling, Scottish Motor Traction and British Electric Traction (Savage, 1985). As administered by the Commissioners, the 1930 Road Traffic Act created a system of territorial monopolies in which preference and protection were given to established operators, in return for improved safety standards and the provision of a service network underpinned by internal cross-subsidisation. Although the Act did not explicitly require them to do so, the Commissioners generally tried hard to implement uniform fare scales both within and between Traffic Areas (Glaister and Mulley, 1983). In addition to its intrinsic simplicity for administrative and operational purposes, standard charging was often favoured by the Commissioners as a method of establishing "proper" competition between competing operators on the same route. The problem with the system was that competition and innovation were stifled by the entry barriers (Hibbs, 1975). The extent of these barriers is noted by Chester (1936), who observes that the licensing system made it:

"... extremely difficult for new concerns to enter the industry."

(Chester, 1936, p. 181)

Opinions on the effects of the 1930 Act differ. Chester (1936) argues that:

"The direct effects of the Road Traffic Act have probably been of a negative rather than of a positive character ... There have been a few new entrants ... but certainly not as many as there would have been without the licensing system."

(Chester, 1936, pp. 181-82)

and concludes that the Act benefited incumbent small operators:
"On the whole, however, the small operator, especially where he operates regular services, has gained by the Road Traffic Act. It has given him a security which he did not previously possess. The large operator, by virtue of his greater financial resources and his capacity to flood the route with vehicles, could and did eliminate many smaller operators prior to 1931. Now, if he desires to obtain a route, almost the only method is by buying out the existing operator at a price at least consistent with the value of his service."

(Chester, 1936, p. 184)

Savage (1966) concludes that:

"The consequences for the travelling public of the 1930 system of public regulation of road passenger transport were, on balance, probably beneficial. On the one hand, safer, more reliable services ... replaced the spasmodic unreliable services .... Stricter supervision of fares and facilities, together with greater stability, was achieved in the industry as a whole."

(Savage, 1966, p. 162)

There is a note of caution, however:

"On the other hand, the protection of existing operators in their little monopolies, the barring of entry to new-comers, the necessity of establishing the existence of public need before a new service might be run instead of allowing the normal criterion of profitability to operate - all these things tended to benefit established operating interests and to that extent worked against the public advantage. ... it can be argued that the Act over-emphasised stability at the expense of new enterprise."

(Savage, 1966, p. 162)

Dyos and Aldcroft (1969) are less equivocal:
"... when account is taken of the improvement in facilities which occurred in the 1930's it is difficult to reach any other conclusion than that the public secured a net gain from the licensing system."

(Dyos and Aldcroft, 1969, p. 360)

Glaister and Mulley (1983) take a different view, arguing that the effects of the Act were detrimental because:

"From the economic and social point of view the principles on which it was based were, in our view, not carefully considered nor were they entirely sound. This was compounded by the strong line that the Traffic Commissioners took in the early days in interpreting the legislation."

(Glaister and Mulley, 1983, p. 134)

Hibbs (1975) argues that the Act benefited the bus operators (large and small), as well as the railway companies, to the detriment of the consumer by the stifling of the competitive process:

"Thus, we may conclude that the main beneficiaries of the route licensing system were the transport operators, including the railways... What the system did was virtually to prohibit the entry of new capital and initiative to the field of regular service operation."

(Hibbs, 1975, p. 54)

2.4 The Post-war Bus Industry 1945-1979

Sections 2.4 and 2.5 describe the development of the post-war bus industry in terms of two contrasting time periods. The first period from 1945-79 (Section 2.4) is characterised by the attempts of Labour governments to increase co-ordination and planning in the
transport sector. The second period from 1980-85 (Section 2.5) is dominated by the laissez-faire attitudes of the newly elected Thatcher government of 1979, which deregulated both the bus and coach sectors during this period.

Road passenger transport in the period 1945-79 was affected by four major influences: the increase in car ownership, the decrease (from the mid 1950s) in local bus travel, the desire by Labour governments to exert greater public control over the transport sector, and the reorganisation of local government under the 1972 Local Government Act. This section examines the role of the first two influences: the legislative and structural changes are described in Appendix one.

Bus Use and Car Ownership

The 1930 Road Traffic Act was implemented in a period of rapid growth in demand for motorised road transport, particularly bus and coach services. Although the number of buses and coaches in Britain declined in the period 1930-1933 as a result of quantity licensing and the Depression, from 1934 to 1953 their numbers increased continuously from 45,689 to nearly 80 thousand. Although one in five public transport journeys by road in 1950 were still by tram, the use of this mode of transport was declining rapidly (Pickup et al., 1991).

From its highpoint in 1953 the bus (and coach) industry entered a period of long term decline which lasted until the mid-1980s. The number of passenger kilometres travelled by bus and coach users in Britain declined in both absolute and relative terms: from 1953 to 1983 they more than halved from 82 billion to 40 billion, whilst the industry's share of total passenger transport slumped from 42% to 8%. This was in a period when total passenger kilometres travelled by all modes increased by 154%. Bus fares rose at a much faster rate than the costs of private transport: between 1972 and 1983 local bus service fares increased by roughly 30% in real terms, compared to an increase in motoring costs of only 3%. Over a similar period, expenditure on subsidies and grants at current prices rose from £71 million to £897 million, an increase of 238% in real terms (Buses White
This occurred because of the rise in costs, particularly labour costs. Excluding fuel tax, the direct costs of bus operation per vehicle mile increased by some 3% per annum in real terms during the 1960s and early 1970s. The bus industry is relatively labour-intensive, and the rise in unit costs reflected the inability of operators to secure increases in labour productivity sufficient to offset the rise in real wages (Gwilliam and Mackie, 1975). The industry became locked into a cycle of falling demand and higher fares, with the gap between revenue and cost filled by subsidy from the public purse.

The main reason for the decline in bus use was the rise in car ownership. Between 1961 and 1971 the number of passenger kilometres travelled by car more than doubled. Car ownership has continued to increase, albeit at a slower rate, since 1971. There are two key economic factors that account for the increase in car ownership: the increase in real personal disposable incomes after 1945 and the decrease in the relative price of car to bus travel. Between 1953 and 1987 personal disposable income per capita in Britain rose from £239 to £4,860, an increase of 118% in real terms. The change in relative price has been referred to in the previous paragraph. In addition, technological improvements and the exploitation of significant economies of scale in car manufacturing were instrumental in the process of containing unit cost increases in the car industry.

Social factors have also influenced the shift from bus to car travel in the post-war period, although the extent to which they influenced, rather than were influenced by, increasing car use is difficult to determine. Some people have preferred to live further from their place of work, increasing their journey to work distances and making bus travel less attractive. Residence in some rural areas may necessitate car use for the journey to work if there are no suitable bus services. There has also been a tendency for people to shop, as well as reside, further from the city centre. The growth of large out-of-town retail developments has generally encouraged car rather than bus use. Leisure time has increased because of shorter working hours, and this has changed social habits:
“The 1950s pattern of social visits at weekends and evening trips to theatre, cinema and bingo has been largely replaced by a wider variety of car-based journeys. Demand for cinema seats has been superseded by television thus reducing demand for evening bus and train services…”

(Cole, 1998, p. 27)

It has been estimated that in 1964, taking into account other factors such as the change in average household size, each additional car reduced bus patronage by about 380 bus trips per annum. By 1976, when car ownership was higher and bus patronage lower, the figure was still in the region of 300 bus trips per annum (Transport and Road Research Laboratory (TRRL), 1980, p. 95). The substitution of car for bus travel occurred amongst higher income groups, leaving bus travel as a predominantly lower income transport mode and thereby making it more difficult for bus operators to raise revenue by increasing fares. There were other important socio-economic changes associated with this substitution. In 1965, 83% of total car mileage was driven by men of working age, whilst 73% of bus passenger miles were travelled by women, the young and the old (Department of Transport, 1967). This age and gender imbalance was another factor constraining operators from raising fares, since women, the young and the old were less likely to have paid employment. This gender imbalance has persisted into the 1990s: excluding central London, only 5% of bus passengers on average are males between 20 and 65 years old (Cole, 1998, p. 201).

It is sometimes argued that car users misperceive the full costs of car and bus travel, where such costs would include not only the running costs of motoring (petrol, oil, maintenance etc.) but also indirect or implicit costs such as depreciation and travel time. Quarmby (1967), for example, noted that characteristics of the choice between bus and car use made sense if car users were only taking into account their petrol costs. The greater the extent to which car users fail to perceive the full costs of their travel, the greater the substitution of car for bus travel over the period.
The decline in the demand for bus travel also had significant implications for the system of cross-subsidisation that had developed in the post-1930 industry. Arguably, the system had worked reasonably well when the demand for bus travel was generally healthy from 1930 to the mid-1950s: after this period, declining demand caused a squeeze on profits which rendered cross-subsidisation increasingly difficult.

2.5 Fundamental Change

The election of the Conservative government under Margaret Thatcher in 1979 marked a radical departure from the "Keynesian consensus" economic policies of previous post-war Labour and Conservative governments. The failure of Keynesian demand-management policies to deliver economic growth and price stability in the 1960s and 1970s was instrumental in creating a shift towards "supply-side" policies designed to increase competition and efficiency. Product and factor markets were to be set free from restrictions by a range of policies including privatisation, deregulation, cuts in direct taxation and legislative changes curtailing trade union powers. One manifestation of this new approach was the Transport Act of 1980, which deregulated the coach (or express) sector and introduced important changes to local bus services. Details of the purpose and provisions of the Act are given in Appendix one, Section 1.7.

2.5.1 Effects of the Transport Act 1980

2.5.1.1 Express Service Fares and Welfare Changes

Evidence suggests that fares on express services fell in real terms by an average of 40% between 1980 and 1983, and that 700 new services were introduced (Buses White Paper, 1984, p.11). A significant proportion of this increase in demand, however, perhaps as much as 40%, took place at the expense of British Rail (Robbins and White, 1985). Douglas (1987) concludes that, if the loss of rail revenue is taken into account, the first-year effects of deregulation caused an overall fall in social welfare in the region of £1.5m to £2.0m. This figure changes dramatically, however, if it is assumed that deregulation
caused the subsequent decrease of rail fares via the introduction of the "London Saver". The resulting increase in rail user benefit was estimated at £36m, producing an overall increase in social welfare of roughly £34.0 to £34.5m. Although the White Paper uses the experience of the long distance market to support deregulation of local services, there are important differences. The price elasticity of demand is significantly lower for local services, subsidy is much greater, and rail competition much less. Ceteris paribus, the expected benefits from local service deregulation will be less.

Kilvington and Cross (1986) also conclude that, overall, there were net benefits from the 1980 Act. Their evaluation of costs and benefits is qualitative rather than quantitative, and shows that both coach and rail users experienced an overall benefit from the 1980 Act, mainly from the lower average fares which resulted from the competition between National Express and British Rail on the major trunk routes. National Express was a subsidiary of the publicly owned National Bus Company. In general, this competition was won by National Express. From 1980 to 1982, National Express patronage increased by 75% and its profit rose from £0.1 million to £1.6 million. Bleasdale (1983) estimates that British Rail losses on its Inter-City network were in the region of £12 million in 1981, increasing to £15 million in 1982 (although this should be placed in the context of an annual turnover of roughly £450 million). The performance of independent coach operators was generally disappointing, especially on express services to London where the larger service network of market leader National Express and fierce price competition created a difficult operating environment: the most notable casualty in this group was the British Coachways Consortium. A significant exception was the performance of the Perth-based Stagecoach group, which expanded rapidly after the 1980 Act. Some independent operators preferred to remain with their existing markets and avoid competition, and some entered into joint agreements with National Express. The only sector in which coach users suffered a net disbenefit was that of secondary route express services, mainly because of the National Express strategy of concentrating resources on the major trunk routes.
2.5.1.2 Entry Deterrence

Further analyses of the post-deregulation express sector are given in Davis (1984), Jaffer and Thompson (1986), and Vickers and Yarrow (1988, pp. 372-7). Their conclusions are similar to those above. In particular, they focus on the use of market power by National Express to deter competition from other express operators. This market power derived from its existing advantages in national marketing, the development of computerised booking systems, predatory pricing, and its control over terminals or hubs such as Victoria Station. National Express refused entry to new operators to these travel interchanges, and also barred operators who had been using them prior to 1980 if they were considered to be in competition with National Express. Its market behaviour is summarised as follows:

"Given the nature of its incumbent advantages, it is not surprising that National Express responded aggressively ... before its rivals could build up goodwill and customer awareness. The entrants' pockets could not withstand the effects of the incumbents' sharp price cutting strategy for very long and National Express had good reason to believe that short-run revenue losses would soon be recouped by the return of its market dominance."

(Vickers and Yarrow, 1988, p. 374)

These practices had important implications for the subsequent deregulation of local bus services, the main one being that deregulation (i.e. the removal of road service licensing) does not necessarily lead to competition. As Beesley (1992) comments:
"In summary, the case of the UK's long-distance bus operations has further lessons for the analyst of deregulation ... the quantity restrictions were indeed inhibiting the market, but their removal did not lead to a successful challenge to the incumbent. On my interpretation, the incumbent used other, underlying impediments to entry to maximise sales, with scarce attention to profit making."

(Beesley, 1992, p. 204)

2.5.1.3 Trial Areas

The evidence for the "trial areas", which covered mainly rural areas within the local authorities concerned, is more difficult to interpret. Although fares generally fell, there is a suspicion of predatory pricing in some cases. In Hereford, for example, Midland Red West, a subsidiary of the National Bus Company and the largest single operator in the City of Hereford, lowered fares and then raised them once rival operators had ceased to provide a competing service. Other problems associated with this company were schedule matching and, on occasion, the provision of free services. In a case-study of the Hereford area, Evans (1988) concludes that the company's post-deregulation strategy was characterised by: first, the provision of high frequency services to deter route entry by potential competitors or to force out existing competitors; second, increasing fares where it had a route monopoly, but not on competitive routes: third, a reduction in operating costs to improve competitiveness. In general, the lesson from the experience of Hereford is that "Deregulation is likely to give higher frequencies and higher fares than are optimal." (Evans, 1988, p. 305). This conclusion is reinforced by the results of the theoretical analysis of route competition in Evans (1987).

There were also problems concerning the quality standards of operators competing in Hereford (Buses White Paper, 1984). Fairhead and Balcombe (1984) conclude that, whilst fares fell and service frequencies increased in Hereford and on some inter-urban

---

1 See the report by J. Tucker in the Sunday Times, 17.2.85. Note that Hereford is the largest town in the trial areas.
routes, deregulation appeared to have little effect in the rural areas. The 1984 White Paper reaches a similar conclusion for the trial areas, but adds an important caveat in anticipation of the 1985 Transport Act:

"Deregulation has not been sufficient to halt the overall decline in rural bus services, but neither has it worsened the situation ... The effects of deregulation are more discernible in urban areas and on inter-town routes ... Under present conditions, deregulation in itself may not be sufficient to allow small operators, however efficient, to compete successfully with established operators with greater resources."

(Buses White Paper, 1984, p. 79)

2.5.1.4 Licensing Changes

The White Paper also discusses the effects of the licensing changes introduced in the 1980 Transport Act on the provision of local bus services. It concludes that, although in theory the Act made it easier for operators to obtain a road service licence, in practice this was not generally the case. The first problem concerns the interpretation of "public interest" by the Traffic Commissioners, in the case of objections from established operators that the granting of a licence would allow revenue abstraction from their routes and thus reduce their ability to cross-subsidise socially desirable but loss-making services. In the early months after the Act, the Commissioners tended to attach greater importance to the interests of those gaining from the cross-subsidy than to other interests such as tax and ratepayers and those gaining from the proposed new services. This emphasis in the interpretation of "public interest" made it difficult for new operators to obtain an RSL in response to an objection by an established operator. This emphasis was re-appraised by the Commissioners, who began to take a more balanced view of the "public interest" and, in particular, required objectors to support their case in "a more analytical and quantitative manner" (Buses White Paper, 1984, p. 73). The second problem was that, in the small number of cases where operators succeeded in obtaining the licence to enable them to compete with established operators, they were faced with

2 The section on Trial Areas in the 1984 White Paper is based on the preliminary findings of the Fairhead
"formidable" resistance in the form of predatory pricing, increased service frequencies and other practices. As a result of these problems, operators wanting to provide new services were restricted to the relatively small number of profitable routes not covered by existing operators.

2.5.1.5 Fare Controls

The 1980 Act removed the power of the Commissioners to control fares for local bus services except in special circumstances but, again, the change did not generally cause significant variations in fares. In particular, by January 1983 there were no cases in which the Commissioners had intervened to prevent overcharging, and only a few where they had intervened to control the terms of competition.

2.5.1.6 Summary

The 1980 Act produced net benefits in the express service sector, although the dominance of National Express meant that the full benefits of deregulation were not realised. In the local bus service sector the Act, in spite of its intentions, did not encourage significant entry to the industry. The main lesson for policy makers was that deregulation via the removal of quantity licensing was not, in itself, sufficient to produce a competitive industry. Other measures, such as reducing the market power of incumbent operators, were necessary. With these objectives in mind, the Conservative government produced the 1984 "Buses" White Paper that culminated in the Transport Act of 1985.

2.6 Conclusions

The 1930 Road Traffic Act was introduced in response to problems over safety, unfair competition and co-ordination between road, rail and tram interests. It initiated a fifty year period of comprehensive regulatory control in the bus service industry, underpinned by the creation of area monopolies and cross-subsidisation. This underpinning was not a necessary consequence of the Act but a result of the way in which the Traffic and Balcombe (1984) study, which was published shortly after the White Paper.
Commissioners chose to administer the system of road licensing. The theoretical rationale for this regulated system is discussed in Chapter three.

The system that developed was criticised because it stifled competition and innovation, although the Thesiger Report (1953) concluded that there should be no change in the licensing system. But reform was made inevitable by the post-war rise in car ownership, coupled with changes in social conditions, which caused a sustained decline in the demand for local bus services. Between 1953 and 1980 local bus service passenger journeys in Britain fell by 60%, whilst fares increased by 86% in real terms. But costs rose faster than revenue, so that the industry became increasingly dependent on subsidies and grants. In 1969, the year in which revenue support and concessionary fares were introduced, total subsidies and grants were £24 million. Eleven years later this figure had climbed to £570 million, an increase of 511% in real terms (inflation adjustment based on the consumer price index). Comparisons of performance measures indicate significant differences between public sector operators, and between the public and private sectors. In particular, labour productivity appears much higher for private sector operators, although this conclusion is based on data for both bus and coach operations.

The experience of the 1980 Transport Act indicated net benefits from the deregulation of coach services, and encouraged the Conservative government to extend deregulation to all local bus services in Britain (excluding London) in the 1985 Transport Act. The background to the 1985 Act, together with its provisions and effects, is discussed in Chapter five.
The purpose of this chapter is to provide a theoretical framework within which to analyse the economic rationale for the regulatory system discussed in Chapter two. In particular, it develops a classification of the ways in which market failure can occur in the local bus service market. The discussion is based on the public interest theory of regulation, which assumes that market failure can be corrected by regulators acting to increase social welfare. This is in contrast to public choice theory, which emphasises that regulation may occur for reasons other than the public interest. The implications of using public choice theory as a basis for analysis are examined in Chapter four, along with other refinements to the basic theoretical framework developed in this chapter.

The structure of this chapter is as follows. Section 3.1 discusses the meaning of social welfare and the Pareto efficiency conditions on which it is based. Sections 3.2 and 3.3 deal with the causes and instances of market failure, and relate the theory to specific occurrences in the local bus service market. Section 3.4 provides a summary and conclusions.

3.1 Pareto Efficiency and Welfare

Welfare economics is a branch of general equilibrium theory that endeavours to determine whether one economic situation is better than another. Vilfredo Pareto (1848-1923) regarded social welfare as a function of individual utility levels \( W = f(U_1, \ldots, U_n) \), where the individual is the only legitimate evaluator of his or her own welfare. This implies that the individual cannot determine whether a change leaves him better or worse off than someone else, or whether it increases or decreases someone else's utility. If
dW/dU_i > 0, according to the evaluation of individual i, then social welfare will rise. This approach marked a movement away from the use of interpersonal utility comparisons in welfare economics.

A situation is defined as Pareto-efficient if it is impossible to make at least one person "better off" without making someone else "worse off". The quotation marks are intended to emphasise the difficulties associated with words such as better and worse, which embody value-judgements that some economists would prefer to avoid. Pareto efficiency depends on achieving technological, distributional and allocative efficiency. These "first order" conditions must be satisfied simultaneously, and have received extensive discussion (Reder, 1947, Bator, 1957, Little, 1957, Winch, 1971). The concept of Pareto efficiency is developed within the neoclassical tradition of a static framework that allows comparisons to be made between alternative equilibrium positions. A perfectly competitive equilibrium, for example, is Pareto-efficient, whereas a monopoly equilibrium is not. Monopoly, in the neoclassical tradition, is an example of market failure in need of government intervention to correct this failure. An alternative approach, that of the Austrian School, stresses the importance of the dynamics of the market process and the role of the entrepreneur in discovering new opportunities. The Austrian approach, together with other refinements, is discussed in Chapter four.

The Pareto criterion implies that social welfare increases when a Pareto improvement is made, that is when someone is made better off without making anyone else worse off. Since many proposed changes in real life would not satisfy this criterion, it has been criticised as being ideologically non-neutral:

"The Paretian approach must be viewed as conservative, offering intellectual support to those who would maintain the status quo."

(Peacock and Rowley, 1975, p.10)
Peacock and Rowley develop this theme, arguing that Paretian welfare economics is incompatible with liberalism, which is defined as freedom from coercion (op. cit., p. 78).

The criterion also embodies an individualistic welfare ethic, since it is based on the assumption that the individual is the "best" judge of his or her welfare. Conventionally, the Pareto-efficiency conditions have been derived on a number of assumptions, including that of no externalities in production or consumption. One implication is that the individual's utility function must be independent of those of others i.e. interdependent utility functions are ruled out. This results in a situation of ethical solipsism, where individuals have no regard to the welfare of others:

"Each individual's preference ordering over states of the economy is characterised by selfishness ... on this assumption [i.e. no externalities in consumption] neither malevolence nor benevolence is recognised as having any relevance for individual welfare."

(Peacock and Rowley, 1975, p. 11)

Hochman and Rodgers (1969) argue that, even if there are externalities in consumption, a Pareto-efficient outcome can still be achieved, although it may require a redistribution of income. If the rich are concerned about the welfare of the poor, for example, there may be a case for progressive taxation in order to effect an appropriate income transfer from rich to poor. Interestingly, this conclusion mirrors that of Pigou and some other welfare economists, although their results are based on the cardinalist notion of the diminishing marginal utility of income, which requires interpersonal utility comparisons to be made. The welfare arguments for and against progressive taxation are discussed in Blum and Kalven (1953). The importance of the Hochman and Rodgers result is that it enables Paretian welfare economics to escape from the "straightjacket" of ethical solipsism. It can be shown, given certain assumptions, that income redistribution from rich to poor will leave some better off and no-one worse off. In effect, the concept of vertical equity is contained within the Paretian concept of efficiency.
There are problems with these results, however. First, it can be shown that it is the preferences of the rich which determine the Pareto-efficient income transfer, whilst the preferences of the poor are ignored. Second, the analysis assumes that the institutional framework precludes the possibility of the free-rider problem, where individuals may intentionally undervalue their utility from income transfers. The absence of the free-rider problem allows the analysis to be generalised from a two-person economy. Musgrave (1970) criticises this assumption, arguing that, in practice, the free-rider problem is likely to increase as the number of people increases. The existence of a free-rider problem implies that a Pareto-efficient income transfer cannot be achieved.

Given these limitations, one might reasonably ask why the Pareto criterion is used at all. There are two main reasons. First, it avoids the problems of making interpersonal utility comparisons that characterised the "old" welfare economics. Interpersonal utility comparisons were strongly criticised by Lord Robbins (1932) as being inconsistent with scientific methodology. Second, it can be shown that perfect competition (subject to certain assumptions such as no externalities in production or consumption, non-increasing returns to scale and no indivisibilities) will lead to a Pareto-efficient outcome. This is an appealing property since perfect competition can then be used as a benchmark for the analysis of market structure, conduct and performance.

It should be noted that perfect competition is a sufficient, but not a necessary condition for Pareto-efficiency. Socialist systems (Lange, 1936, Lerner, 1944) or perfectly discriminating monopolists (Pigou, 1932) can result in a Pareto-efficient outcome. Hurwicz (1960) has constructed a model in which a "greed process" can also lead to a Pareto-efficient outcome, even in the presence of indivisibilities and increasing scale returns.

The outcome generated by perfect competition will depend on the initial income distribution, which in turn depends on the initial distribution of human abilities, human capital and financial wealth. This implies that, given a perfectly competitive economy, the role of government could be confined to expressing the value judgements of society.
by an appropriate redistribution of income and wealth to achieve society's preferred Pareto-efficient outcome. Pareto efficiency would be promoted by the workings of perfectly competitive markets, whilst equity would be determined by the government. Governments could alter the distribution of income and wealth by taxes and welfare benefits, but should not intervene in the competitive process itself.

This argument depends on the assumption that there is no market failure. If markets fail, then there is a case for government intervention. This view is sometimes referred to as the public interest theory of regulation. In this view, government intervention to correct for market failure is justified as a means of promoting the public interest. Essential to this view is the role of the regulators, who are regarded as serving the public interest, rather than their own self-interest. The next section identifies the causes of market failure.

3.2 Market Failure and its Causes

Market failure occurs when the market outcome is Pareto-inefficient. This section deals with the underlying causes of market failure and its manifestations. Market failure is traditionally analysed in terms of the existence of monopoly, externalities and public goods (Bator, 1958). It can be argued that these represent instances of market failure, rather than its underlying causes. The existence of monopoly, for example, is not in itself sufficient to produce a Pareto-inefficient outcome, since the monopolist could practise perfect price discrimination. A more fundamental cause of market failure is related to problems over property rights (Coase, 1960, Demsetz, 1964).

A market can be viewed as a set of arrangements through which buyers and sellers exchange not only goods and services, but also the property rights to use them in certain ways. A car can be purchased and driven on the public highway, but not through your neighbour's garden. Labour can be hired to undertake certain tasks over a certain time period, although there is no (legal) right to specify slavery as one of its uses. Markets determine both the ownership and use of resources, where use is specified in the property rights associated with the resource. Market failure represents a Pareto-inefficient outcome
in which it is possible to leave at least one person better off without leaving anyone else worse off: the causes of this failure can then be interpreted as the reasons why such a change is not implemented. These reasons depend on the nature of the property rights involved. They involve three considerations: insufficient control, information costs and the indeterminacy of the bargaining process. Their existence may prevent potentially advantageous trades or production from taking place.

Insufficient control over property rights might arise because of imperfect excludability or imperfect transferability. Imperfect excludability occurs when property rights are vested in a group of individuals, rather than in a single individual. Common grazing land is an example of such an arrangement, and the problem with it is that there will tend to be over-grazing (Demsetz, 1967). Exclusion of individuals who over-graze may be legally impossible, or may involve prohibitively high exclusion costs. Even if property rights are vested in a single individual, and exclusion costs are low, there may be restrictions on the transfer of the property rights. An extreme example would be that of squatters, who might have the right to use unoccupied land or property without the consent of the owner, but do not have the right to sell it. In the labour market, contracts can incorporate agreements on hours and pay, but not on slavery i.e. the individual can sell his labour, but not himself. The existence of regulations in areas such as price and trading hours can also reduce transferability.

The existence of information costs will tend to reduce the number of advantageous exchanges. Decision-making generally requires information on the price, availability and quality of goods, as well as the reputation of buyers and sellers. It may also require validation of the property rights attached to an asset. In some cases the advantages of the exchange may be outweighed by the costs of acquiring information on it.

Finally, it may not be possible to agree on the terms of exchange because of a multiplicity of potential exchanges. Even if a potential exchange leaves both parties better off, the distribution of the gains between the parties depends on the outcome of the bargaining process. This process can be very complex, depending on the number of parties involved,
the characteristics of those parties, and the uncertainties regarding their reactions to proposals and counter-proposals. This process may be further complicated by the problem of sequential acceptances when more than two parties are involved (Olsen, 1970). Bargains involving an open-access resource (such as international fisheries) and sequential acceptance will produce a free-rider problem, which is an example of the more general problem of opportunism in contracting (Williamson, 1979).

Although this section has concentrated on the role of property rights in market failure, another possibility is to view the existence of imperfect property rights as creating increased transactions costs. It is then the existence of transactions costs that may give rise to market failure, rather than property rights per se. Coase (1960) has shown that, if transactions costs are zero, the assignment of property rights does not affect the Pareto-efficient outcome. If transactions costs are not negligible, then the outcome will be affected by the property rights assignment. The view that transactions costs are the generic factor in causing market failure is argued by Arrow (1970a), where transactions costs are defined as the costs of running the economic system.

The transactions cost approach has subsequently been extended by Williamson (1975), in conjunction with the concepts of bounded rationality (Simon, 1947) and opportunism in the presence of uncertainty, to explain how hierarchies (firms) develop in response to organisational failure (market failure in particular). The firm emerges because of its ability to reduce transactions costs associated with the market. A precursor to the transactions cost approach is Coase (1937), although Coase used the term "marketing costs". More recently, Coase (1991) has adopted the term "transactions costs" and used it to explain the relationship between law and the economy.
3.3 Instances of Market Failure

3.3.1 Monopoly

Suppose that a perfectly competitive industry is transformed into a multi-plant monopolist, and that there are no economies of scale resulting from this. If we also assume that the industry is subject to constant costs, it is possible to compare the price and output combinations of the two types of market structure in Figure 3.1 below:

![Graph showing the comparison of perfect competition and monopoly](image)

Figure 3.1: A comparison of perfect competition and monopoly

The perfectly competitive industry will supply $Q_c$ units at a price of $P_c$ per unit. If the industry is monopolised, the multi-plant monopolist will supply $Q_m$ units at a price of $P_m$ per unit. The area $P_macP_c$ represents the amount of consumer surplus which is appropriated by the monopolist in the form of supernormal profit, whilst the area abc represents the amount of consumer surplus which is lost as a result of the decreased output. Area abc can be used to measure the loss of welfare arising from the monopolisation of a perfectly competitive industry, subject to the qualifications noted in the preceding paragraph. This measure of welfare loss forms the basis for the analysis of the welfare effects of bus deregulation in Chapter six.
The monopoly situation is Pareto-inefficient, since there are potential gains to both consumers and producers from changing to perfect competition. Consumers could contract with the monopolist to produce output $Q_c$ at a price of $P_c$. If this change occurs, consumers gain the area $P_m ab P_c$, whilst the monopolist loses the area $P_m ac P_c$. By striking an appropriate bargain, consumers could return area $P_m ac P_c$ to the monopolist, plus some part of area abc. Both parties would be better off.

This potentially advantageous exchange might not occur because of the property rights problems discussed in Section 3.2. The parties may be unable to agree on the distribution of area abc, or how the burden of the lump-sum payment $P_m ac P_c$ to the monopolist should be shared between the consumers. There may be a free-rider problem with this lump-sum payment, since it might not be possible to exclude those consumers who refuse to contribute to it from enjoying the benefits of the lower output price. Even if exclusion is possible, the costs might be prohibitively high.

The above analysis assumes that there are no economies of scale available to the monopolist. If there are, then the monopoly argument has relevance to the pre-1930 omnibus industry in that the production technology of the industry may give rise to the characteristics of a natural monopoly. Because of these economies which accrue to a monopoly producer, but not to a number of smaller producers, competition will not be sustainable. In this case, public policy could allow the monopoly to continue in order to take advantage of lower costs, but regulate it in order to prevent the exploitation of monopoly power. There is some evidence that, at least from the USA, this view was influential in determining public policy towards the regulation of urban bus transit:

"In the past public policy has treated urban mass transit as if it were a natural monopoly. The roots of this policy can be traced to the consolidation of competing transit systems within a city in the late 1800s and early 1900s."

(Windle, 1988, p. 122)
It is unlikely that economies of scale were a significant source of cost economies in the early days of the motorised omnibus, since as Windle notes:

"The motor bus was self powered and required no large capital investment. Despite these facts, the regulatory framework of monopoly was retained to the present day."

(Windle, 1988, p. 123)

An explanation of this anomaly is the existence of significant economies of density in the provision of bus services. It is the density, rather than scale, economies that give bus services on individual routes the declining unit cost characteristic of a natural monopoly. But Windle argues (1988, p. 135) that this does not justify regulation, since density economies arise through the existence of fixed costs that are not sunk. As long as the market is contestable, a single operator on a given route will earn only normal profit.

The assumption of contestability weakens the argument for regulation on the grounds of natural monopoly, irrespective of whether the monopoly arises from economies of scale or density. If a market is perfectly contestable, regulation is unnecessary since the threat of entry will force the incumbent to earn normal profit only. In the absence of contestability, regulation can be justified as a means of reducing monopoly profits. The consideration of issues such as sunk costs and contestability, together with the distinction between scale and density (and scope) economies, open up a much wider area of discussion which is considered more fully in Chapter four.

Monopoly is also relevant to the post-1930 industry because, under the 1930 Road Traffic Act, the Traffic Commissioners were obliged to take into account the traffic needs of their area, including the provision of unremunerative services. Operators were given a road service licence on condition that they would provide unprofitable as well as profitable services through a system of cross-subsidisation. In order to protect profits from the effects of competition, the operator was effectively granted a territorial monopoly by the Traffic Commissioners. Whether or not the bus industry is a natural
monopoly, the post-1930 industry became, by virtue of the way in which the Traffic Commissioners administered the system, a protected monopoly.

Monopolies, especially protected ones, may also generate another form of welfare loss that Liebenstein (1966) referred to as “X-inefficiency”. Lack of the “competitive stick” implies that managers do not have to minimise costs, so that the cost curve for the monopolist in Figure 3.1 could be higher than the one drawn. In effect, there will be an additional welfare loss equal to the area bounded by the vertical axis, the two cost curves and the demand curve. This area could be substantially larger than the welfare loss triangle abc arising from allocative inefficiency. The provision of substantial external subsidies to the regulated local bus service industry, and the associated problem of “subsidy leakage” (see Section 5.1.4.1), would be another source of “X-inefficiency”. Liebenstein (1980) argues that a “back of the envelope” calculation suggests that the production sector in the USA is only 70 per cent as efficient as it could be. Stigler (1976) and De Alessi (1983) dispute this calculation, arguing that some of the “X-inefficiencies” identified by Liebenstein can be explained as the result of institutional and other constraints and transactions costs.

As noted earlier, the analysis of the welfare effects of monopoly presented above is based on a comparative static framework. It does not take account of dynamic factors such as the reasons for the evolution of the monopoly, or the justification of monopoly profit as a necessary reward for entrepreneurial initiative. These issues are also deferred until Chapter four.

3.3.2 Externalities

Perfect competition fails to produce a Pareto-efficient outcome if the costs and benefits of producers and consumers are not reflected in market prices. This can occur if there are external costs (benefits) which the producer (consumer) does not pay for. The concept of an externality forms the basis for the welfare analysis of the substitution of bus for car travel in Chapter five. It is possible to stipulate a further condition, namely that the
externality should be incidental, although this condition may create ambiguities in certain cases (Mishan, 1971, p. 2). The definition of an externality is not unproblematic, since it can be approached from a number of viewpoints (Papendreou, 1994). Following the analytical framework developed by Pigou, an externality creates a divergence between the private and social costs of production, where social costs consist of both private and external costs. The outcome is Pareto-inefficient, as shown in Figure 3.2 below:

![Figure 3.2: Divergence between private and social cost](image)

Assuming, for simplicity, that the good is produced in a perfectly competitive market, firms set price equal to the marginal private cost (MPC) of production and produce $Q_1$ units at a price of $P_1$ per unit. If firms pay for both the private and external costs of production, they set price equal to the marginal social cost (MSC) of production, producing $Q_2$ units at a price of $P_2$ per unit. In order to do this, Pigou argued that the government could levy a tax on producers in such a way that their marginal private costs become equal to the marginal social costs of production.
It follows that, if externalities are not taken into account, price is too low in relation to the socially optimal price $P_2$ and output (and pollution) is too high. The welfare loss arising from the price-output combination $P_1, Q_1$ can be measured by the area $abc$, since output levels above $Q_2$ have a marginal social cost greater than their marginal benefit. The outcome $P_1, Q_1$ is Pareto-inefficient, since it is possible to increase price to $P_2$ and distribute the resulting gain in welfare to producers and consumers in such a way that both are better off.

If property rights can be established, Coase (1960) showed that voluntary bargaining between the parties concerned might also achieve the Pareto-efficient outcome, irrespective of the allocation of property rights. This has the advantage of avoiding government intervention in the form of the Pigovian tax solution mentioned earlier in this section. One implication of the Coase Theorem is that the role of government should be limited to facilitating the establishment of property rights in the activities concerned. The inadequacy of existing property rights as a cause of externalities, rather than "market failure", had been noted by von Mises (1949a). Following the Austrian School of thought in which von Mises played a prominent role, Block (1979, 1980) has argued that the problem of road congestion could be remedied by private, rather than public, ownership of roads.

But the Pareto-efficient outcome might not arise for reasons similar to those discussed in the case of monopoly. The parties may not be able to agree on the distribution of the welfare gain, or it may be difficult to exclude consumers who do not pay the higher price from enjoying the benefits of reduced pollution. Property rights may not be clearly defined e.g. do smokers have a right to smoke in a restaurant, or do non-smokers have a right to clean air? Establishing and enforcing such property rights by legal process may be very costly.

In terms of the pre-1930 omnibus industry, an important form of externality would be that of inadequate safety. This could arise through deficiencies in vehicle maintenance or from dangerous driving practices (see Section 2.1). Both factors could cause harm not
only to the fare paying passengers, but also to parties not involved in the initial transaction. A loose handbrake, or wheel-nuts, or poorly trained drivers, for example, could all result in accidental injury to other road users or pedestrians. Even where injuries are confined to the driver and passengers, there would be external costs arising from the delay to other road users. It may be argued that all parties should have insurance that covers this type of contingency, and that these "external" costs would be effectively internalised by the insurance market. But the insurance market is subject to the problems of moral hazard and adverse selection, both of which make it harder and more expensive to obtain insurance (Arrow, 1970b, Pauly, 1974).

Given these considerations, the problem of safety can also be analysed as a form of market failure occurring through imperfect information. In particular, where suppliers have more information than consumers, the "lemon principle" (Akerlof, 1970) may cause inferior quality products to be offered to consumers. As an example, if consumers do not know the quality of a good until after they have purchased it, then there may be no incentive for firms to offer a high-quality product. High-quality products cost more to produce, but if consumers cannot detect whether or not the product really is of high quality, then the firm has an incentive to provide a low-quality product at a low cost. It seems reasonable to assume that most bus passengers would be unaware of anything other than the most basic safety defects of the vehicle concerned, and so this argument may have some relevance in the short term. Over time, however, one would expect passengers to become more aware of these problems.

Bargaining between passengers and bus operators over appropriate safety standards would be possible but unlikely, given that only a single bus may be provided for passengers with many different views on safety standards, and that passengers may not be able to co-ordinate their views effectively. Transactions costs are likely to be lower in these circumstances if safety regulation is imposed by the government rather than negotiated between passengers and operators. One possible problem with safety regulation is that its introduction e.g. compulsory car seatbelts, may actually increase accidents because drivers think they can go faster more safely (Peltzman, 1975). Another
problem is that safety regulation, as opposed to the use of Pigovian taxes or subsidies, will force firms to adopt inefficient price and output combinations (see Section 4.1.1).

Mulley (1983) argues that it was concern for safety which was the original cause not only of the safety regulations introduced in the 1930 Road Traffic Act, but also the quantity licensing regulations. Between 1918 and 1930 the number of omnibus accidents rose rapidly (Glaister and Mulley, 1983), against a background of increasing urban road congestion (see Section 2.1).

Congestion and pollution are other forms of externalities. Prior to 1930, increasing urban road congestion, as a result of rising volumes of passenger and freight transport on a slowly expanding road network, was another source of market failure. Although not a major factor in the deliberations of the Royal Commission set up in 1928, increasing congestion would have provided a further impetus towards regulation.

As far as pollution from vehicle exhaust fumes was concerned, there is no evidence to suggest that this formed any part in the deliberations of the Royal Commission. With the rise of car ownership and travel from the 1950s, however, the situation regarding pollution from buses has changed. Whilst a pre-1930s motorised omnibus was undoubtedly more harmful to the environment in terms of toxic exhaust emissions than, say, a hackney, the volume of bus emissions relative to those from car traffic has decreased over time. Bus travel may offer environmental (and other unpriced) benefits when compared with the effects of transporting the same number of people by car, although this conclusion depends on the monetary valuations attached to exhaust emissions and other parameters such as load factors. These issues are analysed in detail in Chapter seven.

The preceding discussion has focused on external costs. The provision of bus services is subject to external benefits, since increasing service frequencies in response to increased demand will reduce waiting times for existing bus users (Mohring, 1972). Given the existence of these external user benefits, there is a case for subsidising bus services to
achieve Pareto-efficiency. These benefits appear greatest at low to medium passenger densities (Gwilliam, 1987, p. 12; Mackie and Preston, 1996, chapter two).

There is another externality that exists in the production of services, since the provision of a service by one operator may raise or lower the value of another operator’s service. If these externalities cannot be internalised in a free market, then market failure will occur. Nash (1982) uses the following example:

"... suppose that the fare from Bradford to Leeds is 30p and from Leeds to London is £5. A single operator of two connecting services will pay close attention to the needs of through passengers, since their loss represents a loss of £5.30. An independent operator from Bradford to Leeds, on the other hand, would regard these passengers as no more important than local passengers, since all pay him 30p. Unless the Leeds-London operator makes a side payment to the local operator to improve connections, journeys involving interchange may be hampered. The situation would be even worse if the Leeds-Bradford operator also operated a Bradford-London direct service. He would have a strong incentive to prevent connections into and out of the Leeds-London service, for fear of passengers diverting from his own long distance service. Common control of services, or a regulatory body to co-ordinate timetables, would seem to have considerable justification in such cases."

(Nash, 1982, pp. 69-70)

A number of comments can be made about this example. First, it seems unlikely that the independent Bradford-Leeds operator would be indifferent to the needs of through passengers. The closer he schedules his service to that of the Leeds-London operator, the more passengers will travel with him and the greater his revenue. Secondly, it seems likely that bargaining over a side payment between the operators would reach a satisfactory outcome, since only two parties are involved. Thirdly, if the Bradford-Leeds operator runs a direct Bradford-London service, it is difficult to see why anyone would travel from Bradford to Leeds in order to travel to London, given that Bradford and Leeds
are roughly equal distances from London. The problem of passenger diversion would simply not occur. If Bradford and Leeds were significantly different distances from London, then the two direct services would not be in competition with each other so that the problem of passenger diversion again becomes irrelevant. The overall point to be made here is that competition does not necessarily lead to a loss of co-ordination between services: indeed, competition may actually promote co-ordination as operators seek ways of increasing profitability (Ponsonby, 1969).

An example is the provision of through (or integrated) ticketing schemes such as travelcards, which may not be provided by competitors even if social benefits exceed costs. Although there was a reduction in the number of these schemes in the short to medium-term post-deregulation period, there are indications that in some areas operators are co-operating to provide this facility:

“There are now signs of improved inter-operator co-operation making it possible to relaunch travelcards in other conurbations (such as Manchester).”

(White, 1997, p. 15)

The problem can occur not only within a transport mode, but also between them (Nash, 1988, pp. 114-5). As in the case of integrated services, however, it is not entirely clear why the market will fail to provide such a facility. If social benefits, which consist mainly of time savings to existing users and consumer surplus to generated and diverted passengers, are significant, then there is no reason why operators should not increase the price of travelcards in order to internalise these external benefits. The argument that price increases will reduce patronage and benefits and therefore tend to prevent the introduction of travelcards (Nash, 1988, p. 115) is only partly true: if benefits exceed costs at the higher price, and those benefits can be internalised, then a travelcard will be introduced by competing operators.
Thus a free market may fail to take account of the external benefits associated with connecting services and through ticketing. In the case of bus services, however, the transactions costs associated with the internalisation of these externalities do not seem particularly significant. If transactions costs do prevent internalisation, then a planned network may be able to co-ordinate (or integrate) services in such a way that these external benefits are realised. Cottham (1986) estimates that the benefits of integrated ticketing, information and route planning in West Yorkshire were in the order of £15 million per annum, or 20 per cent of operating costs. Research in this area is comparatively limited, however, and it is not clear to what extent the problem of regulatory failure may affect this type of result.

Integration or co-ordination also avoids the problem of wasteful competition, which is discussed in the following section. Co-ordination was an important part of the terms of reference of the 1928 Royal Commission: implicitly, its inclusion was an acceptance that the unregulated nature of the pre-1930 system had failed to achieve the most desirable outcome. Although the pre-1930 system may have suffered from a lack of co-ordination, it is not necessarily the case that the post-1986 bus service industry need also suffer from the same problem. Operating conditions were different, and the pre-1930 industry was at a much earlier stage of development. It is interesting, then, that the lack of co-ordination via externalities argument was expressed more recently by the Royal Commission on Environmental Pollution (1994):

"Although competition may help to stimulate innovation and efficiency, the choices to be made cannot be left entirely to market forces to determine because of the external costs unavoidably generated by transport, and because of the benefits lost if networks fragment under unrestrained competition."

(Royal Commission on Environmental Pollution, 1994, para. 13.8)

This is not a recommendation for central government planning, however, but a recognition that service provision must be devolved to the appropriate level:
"The choices are also too complex to be decided by central government acting alone. A more fruitful model would be for decisions to be taken, openly and following full consultation, as close to the point of use as possible but within a framework, set by central government, which enables proper account to be taken of the full complexity of transport issues."

(Royal Commission, op. cit., para. 13.8)

The Royal Commission recommends the creation of additional PTAs and PTEs, including one for London, to "ensure the provision of a co-ordinated and efficient system of public transport throughout their areas" (op. cit., para. 13.24). This recommendation contrasts with the philosophy of the 1985 Transport Act, which removed the powers of PTEs and local authorities outside PTA areas to provide bus services and co-ordinate fares and services.

But the Act does empower them (although not oblige them) to promote the availability of public transport (sections 57(2) and 63(6)). This includes the publication of timetables and the administration of through ticketing (e.g. travelcard) schemes. In fact, since deregulation, all-operator travelcard schemes existed in every former metropolitan county in England. Five of these schemes were administered by the Passenger Transport Executive, and one in Tyne and Wear by a specially formed company of operators. This latter development is particularly significant since it demonstrates that private sector operators can internalise the external benefits associated with these schemes, as suggested earlier in this section. Multi-operator travelcard schemes also exist in a number of English shire counties. These schemes must ensure fair competition between operators: an individual operator cannot be unfairly excluded from such a scheme, or allowed to participate only on unfair terms. Timetable information is also provided by local authorities, although there is a considerable variation in quality (see Section 3.3.4).

To conclude this section, there is a strong case for regulation on safety grounds because it is difficult to internalise the externalities associated with safety problems. As far as
externalities arising from the provision of related services and through ticketing are concerned, the evidence for regulation is less compelling.

3.3.3 *Wasteful Competition*

The concept of "wasteful" competition is a troublesome one. Although it can be given a clear interpretation in the market failure framework discussed in this chapter, the concept becomes problematic if the dynamics of the competitive process are considered. Indeed, competition can never be wasteful for some Austrian School economists. This view is typified by Schumpeter (1950), who referred to the competitive process as "a perennial gale of creative destruction". This approach to competition is discussed in Chapter four. The emphasis in this chapter is on the static efficiency conditions associated with the neoclassical school, and the ways in which market outcomes might violate these conditions. Within this static framework there appear to be three contexts in which economists, either explicitly or implicitly, use the term "wasteful" competition: these are natural monopoly, product differentiation and loss of integration.

3.3.3.1 *Natural Monopoly*

A natural monopoly exists if economies of scale are so large in relation to market demand that only one supplier can profitably serve the market. The existence of two or more suppliers would not be sustainable, and would in any event involve higher average costs than those of a single supplier. The situation is rather more complex in the case of bus services, since economies can also arise from scope and density (see Sections 3.3.1, 4.5.2 and 4.5.3) although the same conclusion would apply.

But it is possible to think of a situation in which, even for a natural monopoly, competition may be a necessary part of the formation of the monopoly and, for that reason, cannot be thought of as wasteful. Suppose the production technology is new, as in the case of the introduction of the motorised omnibus in the early 1900s, and firms do not
have accurate information as to how the costs of production vary with output. A number of these firms decide to take the risk of entering the market and compete with each other. After a time it becomes apparent that significant economies exist, and firms either merge or leave the market. Incomplete information (a type of market failure) makes competition a part of discovering the costs of production and the eventual establishment of a monopoly supplier.

Once the costs of production are known accurately, however, then further competition would be difficult to justify if there are significant cost economies. Since the motorised omnibus started in London in the early 1900s, and was operating even in rural areas by the early 1920s, there would appear to be sufficient time to have ascertained the extent of these economies prior to the regulation introduced in the 1930 Road Traffic Act. If, de facto, the industry is subject to significant cost economies, then regulation to avoid wasteful competition is justified in terms of the natural monopoly argument, subject to the caveat that the market is not contestable.

3.3.3.2 Product Differentiation

In order to clarify the concept of wasteful competition in terms of product differentiation for bus services, it is necessary to distinguish between horizontal and vertical product differentiation. In the first case, services are differentiated in terms of departure times: ceteris paribus, passengers will choose the bus which arrives closest to their preferred departure time. In the second, differentiation occurs in terms of journey time: ceteris paribus, passengers will choose the fastest service. An example of vertical differentiation would be the provision of faster or more frequent minibuses in competition with slow or infrequent double decker buses.
Horizontal Product Differentiation

The seminal article of horizontal product differentiation is that of Hotelling (1929) which uses Cournot's (1838) original example of two sellers of mineral water to demonstrate that, given spatial differentiation in a competitive duopoly, the sellers will eventually cluster together at the centre of the market. Using total transport costs as a proxy measure for social disutility, the cluster equilibrium maximises transport costs and thus social disutility. A uniform distribution of sellers would minimise transport costs and social disutility. Given the assumptions of Hotelling's model, competition is wasteful because social welfare is not maximised. Although the model uses spatial differentiation, temporal differentiation would produce the same result. The general implication is that competition, whether spatial or temporal, leads to a restricted choice for consumers:

"Buyers are confronted everywhere with an excessive sameness ... So general is this tendency that it appears in the most diverse fields of competitive activity, even quite apart from what is called economic life."

(Hotelling, 1929, p. 54)

The Hotelling result implies that, in the context of temporal differentiation such as bus timetabling, two operators will tend to schedule their rival services close together. In particular, one operator will always have an incentive to timetable his service before that of his rival in order to increase market share, since passengers will usually catch the first bus to arrive. This practice is referred to as "headrunning", and was one of a number of practices characteristic of the pre-1930 omnibus market. This form of competition will be wasteful because a uniform temporal distribution of passengers (analogous to the uniform spatial distribution of consumers in the Hotelling model) implies that passengers would prefer buses to be uniformly spaced in time rather than arriving together. The analogy between spatial and temporal differentiation is not exact, however, since location on either side of one's rival can be profitable in spatial competition, but not in temporal: arriving before your rival's bus will increase profits, but arriving after it will not.
The assumptions of the Hotelling model, however, are very restrictive. In particular, the assumptions of totally inelastic demand and the non-consideration of a rival's response to a relocation decision are very strong. Smithies (1941) shows that clustering is unlikely if these two assumptions are relaxed. If the analysis is extended to more than two sellers there is a tendency to disperse rather than cluster, since each seller will try to avoid being caught as "piggy in the middle" (Eaton and Lipsey, 1975, Rothschild, 1976). If the market is modelled as a competitive duopoly serving a circular area rather than a straight line, dispersal will occur in order to provide sheltered markets for the sellers (Devletoglou, 1965). Given the assumptions of the Hotelling model, d'Aspremont et al. (1979) show that there is a logical inconsistency in the result claimed by Hotelling. Additionally, if transport costs are modelled as rising with the square of distance (as opposed to the linear assumption in Hotelling's model) then the principle of minimum differentiation is replaced by that of maximum differentiation: in order to maximise profit, duopolists locate as far apart as possible (d'Aspremont et al., 1979, Capozza and Van Order, 1982). Some studies have sought to justify Hotelling's result in terms of consumers' imperfect information and a corresponding desire to economise on search costs (Stuart 1979, Wolinsky, 1983). The basic idea is that suppliers locate together (e.g. clothes shops in shopping centres) to reduce consumers' search costs for clothes and thereby increase the probability that consumers buy clothes rather than some other product.

Foster and Golay (1986) use a model of temporal, rather than spatial, differentiation to examine the nature of competitive equilibrium in bus timetabling. The model builds on that of Novshek (1980), where the concept of modified zero conjectural variations (MZCV) is developed. Zero conjectural variations (ZCV) are where "each firm views the strategies of all other firms as fixed, regardless of its own strategy": this is the assumption used in Hotelling's model. MZCV are where "Each firm views the strategy of other firms as fixed so long as its own strategy does not cause its delivered price to match or undercut any other firm's price at that other firm's own location ... and that other firms will reduce price if they are undercut at their own locations" (Novshek, 1980). The effect of the MZCV assumption is to reduce the likelihood of direct competition (because of
retaliation) and increase the likelihood of deriving a competitive equilibrium that is symmetric i.e. where frequencies and fares are uniform. The process of entry into the market is modelled on that of Hay (1976), where sequential entry occurs and, because of the spatial (or temporal) location of the entrants, further entry is discouraged.

Where demand is reasonably predictable and there is no substantial congestion, Foster and Golay conclude that headrunning will not occur. Given a market characterised as the circumference of a circle, the competitive equilibrium will be symmetric: fares and headways will be the same for all buses. As noted in the preceding paragraph, this conclusion is unsurprising given the assumption of MZCV: operators will be much more likely to schedule buses at even intervals between their rivals' services, since headrunning would invite retaliation. Even if there is unpredictable demand and congestion, a timetable which is spaced will still be more profitable than one which is subject to clustering, provided that "gaming" behaviour (e.g. cutting in) by bus drivers can be prevented.

Fares may be higher under competition than under monopoly because the higher service frequency implies that load factors will be lower. But the loss of consumer surplus from higher fares will be more than offset by the increase in consumer surplus arising from the reduction in rescheduling disutility. Foster and Golay emphasise that "This is in principle the main economic benefit of bus deregulation" (1986, p. 204), although no evidence is provided to support this statement. It may be the case, for example, that the benefits of reduced costs or innovation arising from competition are greater than those from the reduction in rescheduling disutility.

Although this competitive equilibrium avoids the wasteful clustering which characterises the Hotelling model, Evans (1987) demonstrates that it is not an optimal equilibrium. If fares and service frequencies are set by omniscient planners, rather than the free market, it is possible to increase net economic benefit. In effect, lack of integration in the setting of fares and frequencies causes a loss of social welfare. This view is discussed in the following section.
Vertical Product Differentiation

Wasteful competition under vertical product differentiation occurs when consumers are offered an excessive choice between different qualities of a given product. Excessive choice implies a loss of social welfare arising from a provision of different product qualities which are not the most preferred by consumers. Shaked and Sutton (1982) develop a model of a noncooperative oligopolistic market in which firms decide sequentially whether to enter, what quality to produce, and what price to charge for a given quality. Consumers have different incomes, with higher income consumers choosing higher quality products. The range of consumer incomes determines the number of firms in the market: if maximum income is not greater than four times minimum income, then there will be at most two firms in the market; if the maximum is not greater than eight times the minimum, there will be at most three firms in the market, and so on. Shaked and Sutton show that, provided income differences are sufficiently large, more than one quality will be produced in the noncooperative equilibrium. In the two-firm case, given certain restrictions on relative quality, the first entrant will produce the lowest possible quality level and the second entrant the highest. In general, both firms earn supernormal profit, and the high-quality firm earns more profit than the low-quality firm.

Empirical results for city-centre bus services in London and Aberdeen respectively (Glaister, 1985a, 1986a) support the view that competition (arising from deregulation) will increase vertical product differentiation. In particular, small buses would have an important role to play: they would tend to offer higher frequencies at higher fares than large buses, and carry a higher proportion of passengers with high time values. In high flow areas, large buses would offer lower fares but lower frequencies than small buses. Thus vertical product differentiation in high flow areas would allow both high and low value of time passengers to achieve their desired fare/quality combination. In low flow areas, however, the number of low fare large buses might be "very substantially reduced" (Glaister, 1985a, p. 79) so that welfare for lower income groups would decrease. In some circumstances, only small buses would serve low flow areas.
Nash (1985) questions some of the model parameter values used in Glaister (1985a): in particular, he argues that the assumption of a mean large bus load of 40 is too high so that, even with high frequencies, many buses are full and waiting times correspondingly long in the regulated system. Lower load factors would reduce the strength of the argument in favour of deregulation: "... the model presented in Glaister's paper is so unrealistic that its predictions should not be taken seriously" (Nash, 1985, p. 316). In a rejoinder Glaister (1985b) argues that the model parameter values are not unrealistic and that, in a more detailed study of the Aberdeen area (not then published), conclusions similar to those of his 1985 study had been reached.

Galvez (1986) criticises the modelling of the boarding and alighting of passengers in Glaister (1985a), arguing that the boarding process significantly over-estimates the number of passengers choosing small buses: this implies that, for stops at the end of a high flow route and with assumed breakeven fares at loads of 40 and 7 passengers for large and small buses respectively, large buses must make losses and small ones must make profits. The alighting process over-estimates the average trip length and causes an upward bias in the number and frequency of both large and small buses. This bias is even greater if buses miss out some stops and, since small buses are more likely to do this, the bias significantly increases the relative profitability of small buses. The overall effect of these modelling procedures is to over-emphasise the importance of small buses.

Using amended modelling procedures for boarding and alighting, Galvez (1986, p.104) concludes that, in the case of unregulated competition (i.e. competitive costs and zero subsidy) between large and small buses on high flow routes, the competitive equilibrium contains only 92 large buses. This is completely different from Glaister's (1985a, p.73) result of a mixture of large and small buses (40 and 222 respectively). Although small buses can displace large buses on low flow routes, there is a reduction in welfare for nearly all the users (mainly because of higher fares) rather than the gains predicted in Glaister (1985a). In a rejoinder, Glaister (1986a) asserts that his conclusions are similar to those of Galvez (1986), namely that high flows tend to favour large buses and vice versa. In other words, the differences in the model results are quantitative rather than
qualitative. A further study (Glaister, 1986b) using detailed data on bus services in Aberdeen, and incorporating refinements to the modelling procedures, generally confirmed the results in Glaister (1985a). In a study of the potential effects of minibus competition against existing London Transport services, Bly and Oldfield (1986) also concluded that small buses would generally be able to compete against large buses, although the modelling procedures were different to those in Glaister (1985a, 1986b).

It is the view of the Buses White Paper that deregulation would encourage innovation (i.e. the introduction of higher quality services) and increase social welfare. The opposing view is that the benefits of innovation were overstated, and that many passengers, particularly the poorest, would be worse off through lower quality, but higher priced, bus services (Gwilliam et al., 1985a, 1985b, Nash 1985). Dodgson and Katsoulacos (1988a, 1988b) use the Shaked and Sutton (1982) model of quality competition to examine the welfare implications of bus deregulation. Welfare comparisons are made between a regulated monopolist and a deregulated competitive bus industry, using appropriately specified cost and utility functions. The impact of deregulation on welfare depends on the weight given by the regulated monopolist to retained profits relative to consumer surplus, and the behaviour of total unit costs relative to changes in consumers' willingness to pay for higher quality services.

If the regulated monopolist acts as a welfare maximiser i.e. a weight of zero is given to profit maximisation, then low income consumers may lose welfare after deregulation, since they are likely to face lower quality services at a higher fare. High-income consumers will gain welfare if total unit cost increases at a slower rate than their willingness to pay. In this case the welfare effect is equivocal since, without weightings for low and high-income consumers' welfare changes, it is not possible to know whether the welfare gain by high-income consumers outweighs the welfare loss of low-income consumers. If total unit cost increases more rapidly than willingness to pay of high-income consumers, however, then high-income consumers lose welfare as well. In this case, the welfare effect is unequivocal: all consumers suffer a welfare loss as a result of deregulation. By analogy, it is possible for low-income consumers to gain from
deregulation and for high-income consumers to lose, or for all consumers to gain from deregulation.

The results of Dodgson and Katsoulacos (1988a, 1988b) are qualitatively similar to those of Glaister (1985a, 1986b): a competitive long run equilibrium can exist in which different quality services are provided at different prices, but where the welfare effects may be equivocal. It should be noted that the specification of the Dodgson and Katsoulacos (op. cit.) model is less favourable to deregulation, since it does not allow for cost reductions as a result of competition. The empirical evidence shows that there was a very rapid growth of minibus services in the period immediately after deregulation (Gomez-Ibanez and Meyer, 1987, pp. 71-86), although it was usual for competition to take place at similar fare levels (Dodgson and Katsoulacos, 1988b, p. 280).

Unfortunately, the coexistence of small and large buses with similar fare levels is incompatible with the models of vertical product differentiation discussed above. A possible explanation is that the theoretical models of Shaked and Sutton (op. cit.) and Dodgson and Katsoulacos (op. cit.) depend on the existence of Nash market equilibria, where firms determine equilibrium prices on the assumption that all other firms' prices remain fixed. Given the existence of an oligopolistic market, however, it seems more likely that any form of price determination will depend on the reaction of rival firms, so that price competition becomes less likely. In other words, the actual outcome of deregulation is consistent with the traditional prediction of oligopoly theory that competition, if it occurs, will be non-price competition.

Implicitly, the concept of wasteful competition is applied to actual competition. The introduction of potential competition adds a minor complication to the wasteful competition argument. It seems reasonable to enquire whether potential competition can ever be wasteful. In general, the answer would appear to be "no". If the actual competition is not wasteful, it is hard to think of reasons why any associated potential competition should be wasteful. If actual competition were wasteful, as in the case of the natural monopoly argument above, potential competition would force the monopolist to
behave as if it were in a competitive market. Once again, it is difficult to find a reason why this form of competition should be wasteful.

To summarise, there can be no presumption that competition is necessarily wasteful in either the natural monopoly or product differentiation cases. In the latter case the existence of wasteful competition under horizontal product differentiation depends on a number of restrictive assumptions, particularly the assumption that firms do not consider the reactions of rivals to any change in product differentiation. If rival reactions are taken into account, a competitive equilibrium characterised by clustering is much less likely to occur. For vertical product differentiation, competition can result in welfare gains or losses, depending on the nature of the cost and utility functions. In the case of natural monopoly, competition can be justified as a transitory stage in the discovery of the nature of the industry's cost structure: if, and when, significant cost economies are identified, further competition implies that average costs of production will be higher than under a single supplier. Actual competition will then be wasteful because of the technological inefficiency it creates. In the case of a known natural monopoly, therefore, there is a much stronger case for the argument that actual competition is wasteful. But regulation of the natural monopoly may not be necessary if the industry is contestable.

3.3.3.3 Loss of Integration

Section 3.3.2 dealt with problems arising from externalities in the provision of bus services. Some of these externalities imply a loss of integration. Although integration (or co-ordination) is commonly thought of in terms of connecting timetables and through ticketing arrangements, it is possible to take a much broader view:

"...the argument over public transport integration really is the traditional one of central planning versus the market. In a system which is not integrated, individual routes, modes and services are planned separately by a variety of operators and authorities ... By contrast, in an integrated system, a single authority takes an overview of fares and
services throughout the area, and seeks to plan these to achieve its objectives best subject to the constraints it faces."

(Nash, 1988, p. 99)

Evans (1987) compares the equilibrium fares and frequencies under four alternative possibilities: competition, breakeven maximum net economic benefit (MNEB), monopoly, and unconstrained MNEB. Net economic benefit is measured as social surplus i.e. consumer plus producer surplus. In terms of the discussion about integration, and the losses arising from not achieving it, the first two possibilities are of interest. It is assumed that firms have identical cost functions, buses have unlimited capacity, passengers are uniformly distributed with identical demand functions along an infinitely long day, and that all agents have full information. Using an exponential demand model in order to make tractable the estimation of consumer surplus, the demand function of a representative bus in an infinitely long day is given by:

\[
q = \frac{2Lv}{c} e^{-f_1/v} \left( 1 - e^{-(f_2-f_1+ch)/2v} \right)
\]

where \(q\) = number of actual passengers
\(L\) = number of potential passengers
\(v\) = mean valuation of passenger journeys
\(c\) = cost of rescheduling time
\(f_1\) = fare of representative bus
\(f_2\) = fare of next bus
\(h\) = headway (or service frequency)

\(L, v\) and \(c\) are determined exogenously. In perfect competition, \(f_1, f_2\) and \(q\) are determined simultaneously in order to derive the competitive equilibrium. The operator of the representative bus chooses its departure time and fare in order to maximise profit, on the assumption that the departure times and fares of other buses are given. The departure time will occur midway between neighbouring buses, and entry will occur until profit is reduced to normal (or breakeven). Under an integrated transport system,
however, a single authority is assumed to set $f_1$ and $f_2$ equal to each other, prior to deriving the maximisation conditions, in such a way that MNEB is achieved subject to a breakeven constraint. This latter objective is by definition the optimal combination of fare and frequency to achieve normal profit. The conclusion is that, for levels of demand that are high enough to avoid subsidy, competition results in higher fares and frequencies than the breakeven MNEB, so that competition must be sub-optimal by comparison. Competition causes excessive service frequencies, so that the declining load factors result in excessively high fares. The reduction in net economic benefit from competition relative to breakeven MNEB is in the order of 10 to 12 per cent over a wide range of parameter values.

There are a number of problems with this comparison. Firstly, and probably less critically, there is the analytical problem of how to specify the length of the market in temporal competition. In the case of a finite line, it may not be possible for normal profits to be made: a finite number of buses have to be fitted into a finite timetable, so that fares and frequencies will not be continuously divisible. Supernormal profit may exist with $n$ buses, but $n+1$ buses may cause losses to be made. Given a solution with integer values only, supernormal profit may continue to exist, so that a long run competitive equilibrium becomes impossible to achieve. Another difficulty is that there must always be a first and a last bus, so that a symmetric equilibrium cannot be achieved. These problems are overcome if the market is modelled as an infinite line, but then it is difficult to give a realistic interpretation of this assumption in the context of bus departure times.

Evans (1987) considers the less unrealistic case of a finite line with the end joined to the beginning, thereby circumventing the first and last bus problem. For MNEB, the finite line assumption does not affect the results qualitatively when compared to the infinite line case. But competition is associated with supernormal profits which, on average, can be up to 22 per cent of costs. Offsetting this gain in net economic benefit is a further loss of consumer surplus relative to the maximum under the infinite line case: typically, it is of the same order of magnitude as the 10 to 12 per cent mentioned above. Additionally, at low levels of demand there will be a tendency towards monopoly rather than competition:
at very low demand levels the finite time cycle results in a natural monopoly, as opposed to the competition which would have existed in the infinite line case. As Evans concedes, there are significant differences between the finite and infinite line competitive equilibria:

"The analysis and results for competition are qualitatively different ... It would take another paper and perhaps a different, game-theoretic, approach to deal with it properly."

(Evans, 1987, p. 27)

Secondly, the assumption of identical cost curves rules out the possibility that producer costs will be lower under competition. The Buses White Paper, for example, argues that costs would fall by up to 30 per cent as a result of competition, and this prediction is supported by the fact that costs have fallen significantly since the deregulation of local bus services. Beesley and Glaister (1985) argue that the cost reductions, together with the innovation associated with competition, will outweigh the welfare losses from sub-optimal fares and frequencies.

Thirdly, the existence of full information assumes away one of the fundamental problems with an integrative agency, namely the costs of obtaining such information. In competition decision-making is decentralised so that operators are more likely to know the characteristics of their particular market, and consequently the costs of obtaining information will be lower than those associated with a (presumably) centralised integrative agency. In the market for taxis, for example, it is argued that the information requirements for an integrated approach are so demanding that regulators will not improve on the market outcome (Beesley and Glaister, 1983). In addition, there is no guarantee that the agency will actually operate in the public interest. There is a substantial literature, generally referred to as public choice theory, which suggests that regulators may be strongly influenced by factors other than the public interest. In these circumstances, it may be the case that the costs of regulatory failure outweigh those associated with the market. These theories are discussed in Chapter four. Kay and Vickers (1990) conclude that:
"Although the case for regulation to restrict undesirable or destructive competition cannot always be dismissed out of hand, the private interest in it is almost invariably far greater than the public interest. There is a very heavy burden of proof on those who seek to advance this case, and they should be treated with great suspicion."

(Kay and Vickers, 1990, p. 228)

3.3.3.4 Unstable Competition

Although Hotelling's (1929) model is often used to illustrate the concept of wasteful competition, its primary aim is to determine whether the competitive equilibrium of the Cournot duopoly is stable or not. In non-spatial economic analysis, a stable market is one in which the market price converges to its equilibrium value. If there are multiple price equilibria, a particular equilibrium price may be unstable, although the market itself may be stable. A market is locally stable if it tends to equilibrium from a starting point close to that equilibrium, and globally stable if it tends to equilibrium from wherever it starts. But even if a market is globally stable in the sense above, there may be a perception that the fluctuations in equilibrium price are excessive because they produce too much uncertainty for market participants.

There are a number of additional complications in the case of bus services. Both fares and frequencies are decision variables for operators so that, given an objective such as breakeven, there will be a number of different fare and frequency combinations which satisfy it (Nash, 1978). For global stability, both price and service quality must converge to their respective equilibria. In spatial, or analogously, temporal competition, the location of suppliers must tend towards an equilibrium. This implies that, for bus departure times, the timetable must remain fixed in equilibrium.

It follows that a market can be unstable for a number of reasons. Firstly, shifts in the demand or supply curves can be rapid and unpredictable, causing significant variations in equilibrium price. An example is the market for foreign exchange where, although market
clearing is very fast, fluctuations in the equilibrium exchange rate cause uncertainty for traders and a case for central bank intervention to smooth, or remove entirely, such fluctuations. Thus, even if a market is globally stable, the existence of price fluctuations \textit{per se} may be considered an undesirable aspect of competition. The Bretton Woods system of managed exchange rates lasted nearly thirty years, and was an attempt to overcome the uncertainties associated with freely floating exchange rates. But whilst exchange rates can fluctuate considerably in a floating system, there seem to be few, if any, parallels with the price of bus services. Whatever the criticisms of the pre-1930 omnibus market, price instability does not appear to be one of them. Neither does price instability seem to be a feature of the post-1986 deregulated market: indeed, where actual competition has occurred, the extent of price changes has been limited (Evans, 1990).

An alternative source of uncertainty in bus operations could arise where high frequency services mean that passengers will arrive randomly at bus stops. Failure to catch a particular bus would not result in an unacceptably long wait for the next one. An operator scheduling her bus at 8.00 hrs might find that every passenger caught the rival bus at 8.10 hrs. If arrivals are random then, \textit{ceteris paribus}, both operators would expect to take half the market on average over time. The practical problem from the operators' view is that, on any given day, supply may not be matched with demand: a large bus may run empty, a small bus may have to turn passengers away. A lack of competition manifested in low service frequencies would not produce the same problem of uncertainty. If services run every two hours, rather than ten minutes, passengers would be most unlikely to arrive randomly because of the possibility of very long waiting times if they narrowly miss a bus.

Foster and Golay (1986) examine the competitive stability of situations in which demand for bus services is either predictable or unpredictable. A stable equilibrium is defined as one where "... the buses operating a route, the timetables and the fares change only if there are corresponding changes in the external environment: in demand, relative prices, bus technology or bus costs." (op. cit., p. 192). Where demand is predictable, and there is no congestion, there is a stable equilibrium. Where demand for bus services is dense and
unpredictable i.e. passengers arrive at random, there will be a tendency towards a stable equilibrium if certain "curious old practices" such as missing stops and hanging back are allowed. If a bus cannot carry all the passengers at a particular stop, it will simply miss out further stops until some passengers wish to alight and thereby allow new ones to board. No-one is worse off from this practice, since there was no space anyway for those passengers left behind. The passengers who did board the bus are better off, since they will arrive earlier. There seems to have been little objection to this procedure from either passengers or the Traffic Commissioners: the practice of missing out bus stops, if the bus was full, was common and accepted practice both before and after the 1930 Road Traffic Act.

Where demand is less than expected, so that buses are running below their expected load factor, there is an incentive for buses to hang back until a peak in the traffic distribution is reached, after which hanging back will reduce or stop. It is argued that this will be an efficient way of adjusting to changes in the pattern of demand because "Individuals may lose through having to wait longer than they planned, but more on balance will wait a short time than if the bus service had not adjusted." (op. cit. p. 210). But the fact that some individuals do lose from this practice casts doubt on the existence of a stable equilibrium, since the timetable is effectively being rewritten by the bus drivers to the detriment of some passengers. Foster and Golay acknowledge this difficulty by arguing that it may be inefficient to insist on a fixed timetable in these circumstances. But this would violate the stability condition above, so that the case of hanging back would appear to be neither stable nor an equilibrium.

Evans (1987) concludes that competitive instability is more likely when the market is characterised by a finite, rather than an infinite, line. Potential entrants base their entry decision on the post-entry equilibrium fares and frequencies. In the case of a finite line market, the entry of a new bus causes step, rather than infinitesimal, changes in these variables. The difference between pre and post-entry fares and frequencies may thus be substantial, so that the attainment of a competitive equilibrium requires greater foresight.
by potential entrants than in the infinite line case. If this foresight is lacking, there can be no guarantee that the market will converge to equilibrium.

Where competition results in greater vertical product differentiation e.g. small and large buses on a given route, then the instabilities associated with horizontal differentiation on that route will tend to be reduced (Shaked and Sutton, 1983). Given the distinct quality characteristics of vertically differentiated products, their substitutability of demand will be reduced, thus increasing the likelihood of a stable equilibrium under horizontal differentiation (Foster and Golay, 1986).

Another case of market instability is where a time lag on the price variable is assumed to exist in the demand or supply functions. Although the functions themselves may be stable, the existence of a time lag causes price to fluctuate either towards or away from the equilibrium price. A commonly used example is that of the "cobweb model" which can be specified as:

\[
\begin{align*}
D_t &= a - bP_t \\
S_t &= -c + dP_{t-1} \\
D_t &= S_t \\
(a, b, c, d > 0)
\end{align*}
\]

If \(d > b\) i.e. the supply curve is steeper than the demand curve, then the price adjustment path will not converge to the equilibrium price at the intersection of the demand and supply curves. The market will be unstable, and some form of regulation may be justified. It is sometimes argued (e.g. Gravelle and Rees, 1981, p. 278) that the expectations formation process implicit in the model is naive, since it implies that producers expect a high price in the current period to prevail in the next. But this view seems to miss the essential justification of this model specification: where the product is non-storable, as in many agricultural markets, the entire output must be placed on the market in a given time period. It is not so much the naivety of producers, but the nature of the product, which justifies the supply lag. The model thus has some applicability to the provision of bus services, since empty seats on a given service cannot be stored.
In order to assess the degree of applicability of the model it would be necessary to have information on the slope parameters $d$ and $b$. But research work typically gives estimates of own-price demand, rather than supply elasticities and, in any event, elasticity values are generally not the same as slope values. There may also be a shortage of suitable data over which to estimate the model, since the assumption of market clearing would be valid only for the pre-1930 and post-1986 periods. As a very rough procedure, suppose own-price (compensated) demand elasticities of -0.35 and -0.87 are assumed for the peak and off-peak periods respectively (Glaister and Lewis, 1978), and that these values are used as proxies for the slope parameter $b$. The peak value suggests that peak demand is relatively inelastic i.e. the peak demand curve is fairly steep over the relevant price range. In this case, it is less likely that the supply curve will have a steeper slope and thus more likely that a stable equilibrium exists. The off-peak value of -0.87 is less inelastic, as economic theory would suggest, and it is more likely that the supply curve will have a slope which is greater than the off-peak demand curve. Overall, the very tentative conclusion is that the likelihood of a stable competitive equilibrium is greater in the peak than in the off-peak period, although there is no particular reason to suppose that there will be an unstable equilibrium in the off-peak.

Overall, the conclusion from the theoretical literature discussed so far is that there are no compelling reasons to suggest that competition between bus operators will necessarily be subject to instability. But the validity of this conclusion depends critically on the assumption of modified zero conjectural variations, which implies that an operator will not compete directly with a rival because of retaliation. Chapter two provides evidence to show that direct competition did occur in the pre-1930 bus market, and Chapter five also provides similar examples for the post-deregulation period. The Royal Commission on Environmental Pollution (1994) identified instability of bus services as a problem after deregulation. In particular it found little evidence to suggest that changes in services were a result of the market responding to changes in demand. Instability was a problem because:
"... stability matters, since people's decisions about where to live, work and shop, and how to travel, depend on confidence that the bus service they enjoy now will still be there in six month's time."

(Royal Commission on Environmental Pollution, 1994, para. 150)

But the problem of instability may be overstated. Inspection of the registration figures from traffic area offices (see Table 3.1) shows that, after an initial period of instability, the pattern of new and altered registrations becomes relatively stable:

Table 3.1: Number of registrations accepted by traffic area offices

<table>
<thead>
<tr>
<th>Year</th>
<th>New</th>
<th>Variations</th>
<th>Cancellations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985/86</td>
<td>14987</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85/87</td>
<td>14981</td>
<td>11794</td>
<td>1945</td>
</tr>
<tr>
<td>87/88</td>
<td>7412</td>
<td>22860</td>
<td>5653</td>
</tr>
<tr>
<td>88/89</td>
<td>6761</td>
<td>17775</td>
<td>5771</td>
</tr>
<tr>
<td>89/90</td>
<td>7605</td>
<td>17096</td>
<td>7288</td>
</tr>
<tr>
<td>90/91</td>
<td>6030</td>
<td>15734</td>
<td>5503</td>
</tr>
<tr>
<td>91/92</td>
<td>6621</td>
<td>15726</td>
<td>5127</td>
</tr>
</tbody>
</table>

Note: Figures are for Britain excluding London.

Given that the number of new and changed registrations in May and June 1987 was approximately double the corresponding figures for 1984 (Balcombe et al., 1988), it is clear that post-deregulation stability was not as great as that of the pre-deregulation period. This is hardly surprising, given the scale of the structural changes taking place in the industry. Nonetheless, the figures for 1988/89 onwards are not indicative of an industry suffering from significant instabilities, although they are not conclusive because there is no comparable data for the pre-deregulation period.
3.3.3.5 Unfair Competition and the Problem of Second Best

The Office of Fair Trading (OFT) has investigated a number of cases of alleged unfair competition in the bus service market. Some cases, such as problems involving operator access to bus terminals, are relatively easy to identify and rectify. Others, particularly predatory pricing, are much more problematic. Over the five year period following deregulation, the OFT received 105 complaints relating to predatory behaviour in the bus service market, but investigated only four of them. Two complaints were justified, but the OFT dismissed the other two (Elliot, 1991). Predatory pricing is often associated with situations in which a large company attacks a smaller company, using its greater financial reserves to sustain a loss-making price: the so-called "long purse" hypothesis (Benoit, 1984). A full discussion of predatory pricing is deferred until Section 4.6.3, where it is treated as one of a number of entry deterrence strategies. Although predatory pricing has occurred in the bus service industry, it has done so against the background of a general reluctance on the part of operators to compete in terms of price: instances of price "wars" in the industry have tended to be spasmodic and isolated.

Unfair competition is also interpreted as a situation in which a competitor does not pay the full costs of the service provided. This can arise in two ways: firstly, a government may decide to subsidise a domestic industry in order to protect it from international competitors or other domestic industries; secondly, historical or fiscal factors may result in a situation where a firm or industry avoids certain costs imposed on its competitors. It is the second case that is of relevance to bus services.

Section 2.1 has already mentioned the argument expressed by the railway companies in the pre-1930 period, namely their view that road users did not pay the full costs of road use so that competition between road and rail was conducted on unfair terms. Another example of differential treatment can be found in the revenue to cost ratios between bus and car use. The Royal Commission on Environmental Pollution (1994, p.100) estimated revenue to public cost ratios for 1994/95 of 1.0:1 for buses (including coaches) and 4.5:1 for cars (including light goods vehicles (LGVs)). Public costs do not include
environmental costs, so the Commission also provides an estimate of environmental costs. Taking public and environmental costs together, and estimating upper and lower bounds for environmental costs, the revenue-to-cost ratios become 1.0-1.5:1 for cars and LGVs. Unfortunately, on the basis of the figures given in the Royal Commission Report, it is not possible to derive precisely the corresponding ratio for buses and coaches, although it is likely to be slightly below the 1.0:1 figure above.

These revenue-to-cost ratios do not include congestion costs, which the Royal Commission argues are not external to road users as a group and should therefore not be included in the costs of road use:

"Although an individual user of the transport system may be said to impose external costs by contributing to congestion ..., the costs of congestion ... can be regarded as both imposed and borne by the relevant group of transport users;"

(Royal Commission on Environmental Pollution, 1994, p. 99)

Whilst this view may be true in relation to car users imposing congestion costs on other car users, it is difficult to see why it would apply to the situation where car users impose congestion costs on bus users, which is generally the case in peak period urban traffic. Since the costs of congestion caused by road transport in the UK in 1991 have been estimated as high as £17.5 billion (Peirson et al., 1994), it is likely that the congestion costs imposed by car users on bus users are of a considerable magnitude. Inclusion of these congestion costs, as well as environmental costs, would produce a revenue-to-cost ratio for cars and LGVs significantly below the 1.0-1.5:1 figure above. Although some congestion has been caused by buses after deregulation, the Traffic Commissioners are empowered under Section 7(4) of the 1985 Transport Act to impose traffic regulation conditions. This power has been infrequently used, with only twelve traffic regulation conditions imposed between 1986 and 1993.
Another way in which the relationship between car and bus use is distorted is the large subsidy given to company car users. This subsidy was estimated to be between £1000 and £2000 million (current prices) per annum in 1984, the latter figure being larger than the combined subsidies given to the PTEs, the bus industry and British Rail. These subsidies make car users insensitive to changes in the relative cost of bus and car use, and make it more difficult to persuade car users to switch to bus use. The problem was particularly acute in London, where some 72% of motorists travelling in Central London were financially assisted with their journey (House of Commons Transport Committee, Second Report, 1985, p. 13).

Thus it can be argued that bus services compete at an unfair disadvantage to car travel, since car users do not appear to pay the full cost of their car use. If this is the case, there will be market failure via allocative inefficiency since price will be below marginal social cost for car users. This raises the question of how to improve allocative efficiency. In a first best world, price would be equal to marginal social cost in every market. The first best solution to the above problem would thus be to bring price into line with marginal social cost for car users. But suppose this option is not available, perhaps because of political constraints or powerful pressure groups. Given the continuing existence of allocative inefficiency in a market, the problem of the second best is how to determine the allocation of resources in other markets. Intuitively, one might expect that the best policy is to achieve as many efficiency conditions as possible in these other markets. Surprisingly, second best theory shows that, once one efficiency condition is violated, a better outcome is achieved by violating other efficiency conditions (Lipsey and Lancaster, 1956).

Suppose that bus and car travel are substitutes, and that the price of car travel is less than its marginal cost i.e. $P_c < MC_c$. It can be shown that the second best pricing policy for bus travel is to set $P_b < MC_b$. Assuming that this condition is not satisfied by the bus industry creating its own externalities, then the implication is that bus use must be subsidised. In general, for two outputs i and j, and for small changes in the price of i, the second best pricing policy can be approximated as:
\[ P_i = MC_i - \frac{\partial Q}{\partial Q_i} (P_j - MC_j) \]

where \( \frac{\partial Q_j}{\partial Q_i} \) is a measure of the relationship between \( i \) and \( j \). If \( P_j = MC_j \) or if \( \frac{\partial Q_j}{\partial Q_i} = 0 \) (i.e. zero cross elasticity between \( i \) and \( j \)), then \( P_i = MC_i \). If \( \frac{\partial Q_j}{\partial Q_i} < 0 \) (i.e. \( i \) and \( j \) are substitutes) and \( P_j > MC_j \), then \( P_i > MC_i \), or if \( P_j < MC_j \), then \( P_i < MC_i \).

If \( \frac{\partial Q_j}{\partial Q_i} > 0 \) (i.e. \( i \) and \( j \) are complements) and \( P_j > MC_j \), then \( P_i < MC_i \), or if \( P_j < MC_j \) then \( P_i > MC_i \).

If markets are independent of each other \( (\frac{\partial Q_j}{\partial Q_i} = 0) \) then, if \( P_j \) does not equal \( MC_j \), the second best pricing policy is also the first best i.e. set \( P_i = MC_i \). In practice, however, markets such as bus and car travel are not independent, and some knowledge of the cross-elasticity would be needed in order to determine the second best price. If such knowledge is available, then it might be possible to devise a second best pricing policy for bus travel, bearing in mind that such a change would also have impacts on other markets related to bus travel. Pricing policy in these other markets would have to be changed, depending on the extent of the cross-elasticities. This "piecemeal" approach to the second best problem is developed in detail by Davis and Whinston (1965).

The importance of the relationship between buses and cars was stressed in the evidence given to the House of Commons Transport Committee (1993):

"... the effect of deregulation will be determined above all by the conditions of competition between buses and cars, not between one bus company and another. If it were the case that the relative price of buses and cars included the congestion and environmental costs of their use, and that the road space allocated to buses and cars reflected the efficiency of their use of that space, then buses would be faster and cheaper than cars. Passenger attitudes, and the outcome of competition, would be radically different."

But McKee and West (1981, 1984, 1987) argue that the second best problem is not really a problem at all, since a second best situation can be viewed as the outcome of public policy action rather than as an obstacle to it. To the extent that the outcome is a product of market processes, where these processes include the influence of politicians and other public policy makers, it can be regarded as first, rather than second best. As an example, if information is non-excludable it might reduce the incentive to innovate. The public policy response to this problem might be to introduce patent rights to protect innovators. Given that patent rights confer monopoly powers on their holders, it would be inappropriate to interpret this monopoly as a distortion requiring a second best solution. The monopoly outcome is part of a process in which public policy makers are pursuing efficiency improvements. In general, monopoly and other types of market failure are not necessarily distortions in need of offsetting corrections, but should be viewed "as part of the Paretian solution and their existence constitutes part of the first best allocation." (McKee and West, 1984, p. 250).

This view is representative of the public choice approach to economic problems, which takes a different perspective from that of the more traditional market failure approach discussed in this chapter. The next chapter examines this approach, together with other alternative perspectives to the traditional approach.

3.3.4 Public Goods

A pure public good has two characteristics: non-rivalry in consumption and non-excludability. It is doubtful that a pure public good exists, but impure examples might include national defence, lighthouses and radio and TV broadcast signals. Consumption of lighthouse services will be non-rivalrous up to a point, since consumption by one ship will not usually reduce the light available to another. It is possible to envisage a situation, however, where there are so many ships that visibility might be obscured. Similarly, whilst it might be very difficult to exclude some ships from consumption, it might not be impossible e.g. the extension of property rights further out to sea. In a sense, public goods can be regarded as a particular type of beneficial externality, since provision of a public
good for one person will make it available to others at no extra cost (Hirschleifer and Glazer, 1992, p. 461).

The characteristic of non-excludability gives rise to the "free-rider" problem i.e. it may be difficult to exclude consumers who do not pay for the public good from enjoying its benefits. If this is the case, then consumers may be deterred from paying anything at all, reasoning that as long as someone else pays for it then they will still be able to benefit from it. If all consumers think like this, a perfectly competitive market may not provide the good at all, even though its provision would bring net benefits. Allied to the free-rider problem is that of undervaluation in the revealing of consumer preferences. Attempts to discover whether the public good should be provided may founder on the incentive to understate the willingness-to-pay for the good, since some consumers may rely on others to pay for the good. Although various mechanisms have been developed to discover the "true" preferences (Clarke, 1971, 1972, Groves, 1970, 1973, 1976), a serious difficulty with such mechanisms is the lack of incentive to give any valuation at all in large number situations (Tideman and Tullock, 1977).

Information about bus services can be thought of as having public good characteristics. If a timetable is displayed at a bus stop, then the information service it provides will be non-rivalrous and non-excludable. Given the free-rider problem above, bus passengers may not be willing to pay for the benefits of this type of information and bus operators may not provide it or, if it is provided, the information may be incomplete or out-of-date. This latter problem of low quality information is likely to be particularly relevant to a competitive environment where services and frequencies may be changing rapidly. The free-rider problem may also extend to operators themselves, since one operator's publicity concerning services may generate benefits for other operators.

In a study of low-income families in Tyne and Wear, availability of information was considered to be the service characteristic that had deteriorated most after deregulation (Bradshaw and Holmes, 1988). There is also some evidence to suggest that the deterioration in information availability was a more general problem:
"In spite of the increased need for up to date, reliable information about services and fares, bus users have reported that there has been a decline, rather than general improvement in the already poor levels of information provided at the bus stop and on the bus."

(Buswatch, 1988, p. 107)

"Unfortunately, too many existing operators failed abjectly to publicise and promote their networks whilst many newcomers to the local bus market have been content to ride on the coat-tails of the larger companies ..."  

1 Quotation from "Deregulation – has the customer benefited?", paper by C. Moyes. Commercial Director, Go-Ahead Group plc., delivered to a seminar on the British Bus Industry: Ten Years of Change (1994).

The problem is further compounded by the decline in comprehensive timetable information after deregulation. Prior to deregulation, local authorities provided information to bus passengers relating to the services of all operators. In a competitive environment, an operator will be reluctant to provide information about rival services (unless he can arrange for compensating payments from his rivals). But the lack of comprehensive timetables is really a question of the loss of integration through competition, as discussed in Section 3.3.2, rather than public goods per se:

"Competition actually has negative effects on both commercial and supported services because it promotes a fragmentation rather than integration of information and services. This makes it difficult for the passenger to plan his or her journey without reference to more than one timetable and faretable and may involve changing buses that do not necessarily connect smoothly with each other."

(BusWatch, 1988, p. 109)
In a study of the metropolitan areas, Tyson (1990) concludes that co-ordination in terms of timetables, fares and passenger information decreased significantly after deregulation, particularly in its immediate aftermath. On the other hand, the study acknowledges that co-ordination has subsequently improved, although it estimates that this improvement was insufficient to offset the earlier welfare losses.

An interesting question concerns the limited appearance after deregulation of private sector companies specialising in the provision of services such as passenger information and through ticketing schemes\(^2\). If the types of information discussed above have private good characteristics, one would expect to observe the development of such specialist companies. The fact that they have not appeared on any large scale tends to reinforce the view that bus service information is, in part at least, a public good. Timetable and other information has been supplied by local authorities since deregulation, but the quality of such information has varied considerably between authorities since they have no statutory obligation under the 1985 Transport Act to provide it. As noted in Section 3.3.2, however, there are more recent indications that operators are beginning to co-operate to relaunch facilities such as travelcards (White, 1997, p. 15).

Samuelson (1955) demonstrated that the Pareto-efficient provision of a public good occurs when the "top-level" condition is reformulated as

\[
\sum_{i=1}^{n} \text{MRS}^{i}_{XY} = \text{MRPT}_{XY}
\]

where X is a public and Y a private good, and n is the number of consumers. This condition will not be satisfied if there is a problem with free-riding or undervaluation of consumer preferences, in which case a competitive market will provide less than the Pareto-efficient amount. It can then be shown that, by appropriate bargaining, it is possible to leave at least one consumer better off without leaving anyone else worse off.

\(^2\) Comment by W.J. Tyson in a paper delivered to the seminar on the British Bus Industry: Ten Years of Change (1994).
(Buchanan, 1968). As with monopoly and externalities, however, the costs of arranging and enforcing contracts may be prohibitively high, so that the market continues to provide a Pareto-inefficient amount of the public good.

It is extremely difficult to place a valuation on the extent of Pareto-inefficiency with respect to bus passenger information. Benefits and costs are not easily measurable, and there is also the problem of possible bias in bus passenger responses to "watch-dog" groups such as BusWatch: dissatisfied passengers are more likely to respond than satisfied ones. There is the additional problem that, even before deregulation, the provision of service information left much to be desired (BusWatch, 1988, p. 13). This may have caused bus passengers to form an unrealistically high expectation of the benefits which competition would bring in terms of improved information. Although information problems might be expected to decrease after the initial changes caused by deregulation, more recent studies of passenger attitudes towards the effects of deregulation confirm that a lack of information was still a problem (Steer Davies Gleave, 1991, Transport Advisory Service, 1992). In some areas, over 80% of residents were unaware of the basic details of their local bus service and more than 40% were uncertain about how to obtain information on bus services. Additionally, whilst county and regional councils usually published timetables about public transport services, only 50% of them published timetables about all services (Transport Committee First Report, 1995, Vol. III, p. 297).

### 3.3.5 Common Goods

A common good is one that is non-excludable but rivalrous in consumption. Access to the resource may be uncontrolled, as in the case of some ocean fisheries, or limited to a number of individuals who satisfy certain eligibility conditions. An example of the latter would be a lake in which only community members have the right to fish. The basic problem in both cases is that, in the absence of agreement to restrict fish catches, overfishing will occur because there will always be an incentive for an individual fisherman to increase his catch (Demsetz, 1967, Cheung, 1970). Essentially, if you don't
catch the fish, then someone else will. This outcome is Pareto-inefficient, since it is possible to show that social welfare could be increased by restrictions on the right to fish. There will also be other sources of inefficiency. Investment in the resource may not occur, or the resource may be used to extinction, since a single individual cannot prevent others from enjoying the benefits of investment or conservation even though they may contribute nothing towards them. Although total net benefits from such policies may be positive, the free-rider problem means that a competitive market will not provide the Pareto-efficient amount of investment or conservation.

Although the distinction between public and common goods is conceptually clear, in practice the boundaries are often blurred, as in the case of bus stations. Following access difficulties experienced by some operators after the 1980 Transport Act, the 1985 Transport Act made bus stations subject to the Fair Trading Act 1973 and the Restrictive Practices Act 1976. This makes it illegal to refuse access to rival operators, so that bus stations now possess a common good element. Given the right of access for all operators, together with rivalry in consumption\(^3\), they may be subject to the problems above. It could be argued that, since operators must pay a fee to the facility owner for their use of it, they are in fact private goods i.e. excludable and rivalrous in consumption. But access to a private good can be prevented, since ownership of the good includes the right to exclude others from consuming it. Access to a bus station cannot be legally prevented, since ownership no longer confers the right of exclusion.

But the owner of the bus station has a right to charge a (non-discriminatory) fee for the use of its services. If the fee is flat-rate rather than use-related, the common good problems of overuse and underinvestment may occur. If the fee is use-related, the problem of overuse will not be as great. Given non-excludability, however, there will still be a free-rider problem so that investment may continue to be sub-optimal. Rather than pursue these arguments further, however, it is worth noting that deregulation has not generally led to large numbers of new entrants to the industry so that, at an operational

\(^3\) Rivalry in consumption seems a reasonable supposition, given the limited space available in built-up areas where bus terminals tend to be located.
level, the common good problem associated with bus stations does not seem to have caused significant difficulties overall. A bigger common goods problem, however, may be associated with on-street bus stops where buses impose costs on each other and on other traffic.

3.4 Summary and Conclusions

Social welfare is maximised when the Pareto efficiency conditions are satisfied. The resulting Pareto optimum is one of a number of possible optima, depending on the initial distribution of income. Although the Pareto optimum is efficient, society may consider it inequitable. Market failure occurs when the Pareto efficiency conditions are not satisfied, so that social welfare is not maximised. Given market failure, the public interest approach argues that there is a case for regulation to increase social welfare. The fundamental cause of market failure is the existence of transactions costs that are sufficiently high to prevent the implementation of Pareto improvements. Instances of market failure include monopoly, externalities, wasteful competition, unstable competition, unfair competition, public goods and common goods. These categories are not always mutually exclusive.

The local bus service market is subject, in varying degrees, to each of these instances of market failure. In the case of safety, there is a general consensus that regulation is justified. The improved safety measures of the 1985 Transport Act met with little or no opposition from either pro or anti-market factions. Evidence shows that bus safety and maintenance standards have not fallen since deregulation (Astrop et al., 1991, Transport Committee First Report, 1995, Vol. III, p. 289).

It is much more difficult to evaluate the other market failures. Critics of the 1985 Transport Act have emphasised the wasteful and unstable competition aspects of deregulation, but neither the actual outcome of deregulation, nor the theoretical work on wasteful and unstable competition discussed in Section 3.3.3, presents a compelling case for significant market failure in these areas. The case becomes less compelling when alternative theoretical perspectives to the public interest approach are considered.
A problem often cited in relation to deregulation is the loss of co-ordination in areas such as timetable information and travelcards. Problems over comprehensive timetable information have increased after deregulation, although information provision was not particularly good before it. This problem has a relatively easy solution, however, which is to place a statutory obligation on local authorities to provide comprehensive information. The use of travelcards did reduce in the short to medium term after deregulation, but there are signs that operators are beginning to reinstate them. Another form of market failure which may be significant is the "Mohring effect", although this appears to diminish at relatively high service frequencies.

Even if one were to conclude that market failure is significant enough to justify regulation, however, it is important to remember that the analytical framework of this chapter has been used to provide a starting point for the discussion of local bus service deregulation. In particular, it has not considered the problem of regulatory failure. The fundamental question is not whether market failure exists but whether its costs outweigh those associated with regulatory failure. This chapter has focused only on the ways in which market failure affects the local bus service industry. The next chapter redresses this imbalance by considering the problem of regulatory failure in the industry. It also introduces some other refinements to the theoretical apparatus used so far.
CHAPTER 4

Theoretical Perspectives II: regulatory failure and the case for a market solution

Chapter three developed a theoretical framework built on the idea of Pareto efficiency. The general conclusion from this analysis is that where markets fail to provide a Pareto-efficient outcome, policy makers should intervene to correct for market failure. This traditional or "public interest" approach fails to take account of three important factors: first, it implicitly assumes that government intervention is unproblematic; second, it is based on a static analysis of efficiency; third, market structure is used as an indicator of market failure. This chapter discusses alternatives to the public interest approach, and examines their relevance to the UK local bus service market.

Section 4.1 discusses the public choice approach and its emphasis on the process of decision making, including the concepts of regulatory capture and regulatory failure. Section 4.2 discusses the Austrian school approach and its emphasis on the process of competition, in which monopoly can confer benefits through gains in dynamic efficiency, as well as its emphasis on entry and exit barriers to markets. The theme of entry and exit barriers is developed in Section 4.3, which deals with the concept of contestability. Entry and exit barriers within the context of the local bus service market are identified and evaluated in Section 4.4. Innocent entry barriers arising from economies of scale, density, scope, network effects and experience are discussed in Section 4.5. Deliberate attempts to deter entry, or strategic entry barriers, are dealt with in Section 4.6. Section 4.7 summarises and concludes.
4.1 The Public Choice Approach

4.1.1 Introduction

The public choice approach emphasises the process of economic decision-making, rather than its outcome. Buchanan (1986) has criticised the traditional approach of deriving efficiency conditions based on the assumption of constrained maximisation by economic agents. In particular, he argues that greater attention should be paid to the interactions between individuals in the process of exchange. The flavour of his views is contained in the following quotation:

"What should economists do? My 1962 as well as my 1982 response to this question was to urge that we excise the maximising paradigm from its dominant place in our tool kit, that we quit defining our discipline, our "science", in terms of the scarcity constraint, that we change the very definition, indeed the very name of our "science", that we stop worrying so much about the allocation of resources and efficiency thereof, and, in place of this whole set of ideas, that we commence concentrating on the origins, properties and institutions of exchange, broadly considered ... The approach to economics that I have long urged and am urging here was called "catallactics", the science of exchanges, by some nineteenth century proponents ... This approach to economics, as the subject matter for inquiry, draws our attention directly to the process of exchange, trade and agreement, or contract."

(Buchanan, 1986, p.20)

Wiseman (1985) characterises the traditional approach as one concerned with outcome efficiency, whilst the public choice approach is concerned with process efficiency. The central question is not whether the outcome is efficient, but whether the decision-making process is efficient i.e. could some other process have produced a better outcome? The role of the government and public sector agencies in this process is crucial. The traditional approach assumes that government intervention will reduce market failure
without creating problems of its own, whereas the public choice approach emphasises the
difficulties created for the decision-making process by the intervention of government
regulators who may have their own interests to consider. Indeed, regulatory failure may
be just as much a problem as market failure.

The differences in approach can be illustrated by reference to the problem of externalities
discussed in Section 3.3.2. Given the existence of an externality, traditional welfare
criteria suggest that taxes or subsidies should be used to correct this form of market
failure. In general, the use of taxation will be more efficient than regulation, since
regulation is a blunt instrument that affects all firms equally, irrespective of their
individual demand and cost conditions. The result is that regulation forces firms to
choose inefficient price and output combinations (Buchanan and Tullock, 1975, Shone,

But as Kay (1990, p.40) has noted, "whatever the conclusions of economic analysis,
policy makers have tended to prefer regulation to the use of fiscal policy." The public
choice approach can provide a rationale for this preference in terms of producer group
interests. Suppose that, in a perfectly competitive industry, each firm is constrained to
produce the same amount of output such that total output falls below the competitive
equilibrium. In terms of Figure 3.1 (reproduced below as Figure 4.1), the effect of the
quota regulation will be to restrict output from Q_c to Q_m, raise price from P_c to P_m, and
enable the industry to earn supernormal profit. Existing firms will not be in equilibrium
in this situation, since they have an incentive to increase output to increase profits, and
the regulator must monitor them to ensure that output quotas are not violated. The
regulator must also ensure that new firms do not attempt entry into the industry in
response to the existence of supernormal profit. The effect is that regulation allows
existing firms to act like a cartel and enjoy the benefits of monopoly power (Buchanan
and Tullock, 1975).
There may also be advantages for the regulator. Given the need to monitor and enforce the quota regulation, there may be an opportunity to expand the regulator's sphere of influence and prestige by increasing the agency's budget and number of staff. Empirical evidence collected by Eckert (1973) on the regulation of taxicabs supports this view. Such behaviour can be seen as part of a more general desire on the part of bureaucrats to enhance their own self-interest. Niskanen (1968, 1971) has developed models in which the economics of bureaucracy can produce a Pareto-inefficient outcome, although Bohm (1987) has questioned the view that bureaucracies are necessarily inefficient. Career advancement, for example, may be facilitated by reducing rather than increasing the size of the bureaucrat's fiefdom.

4.1.2 Regulatory Capture

Stigler (1975) analyses the phenomenon of "regulatory capture", whereby producer groups influence the regulator to promote the interests of the producer rather than the consumer. In this way, regulation may be welcomed by producers as a means of avoiding competition and creating monopoly power. Mulley (1983) argues that the development of the local bus service industry after 1930 was adversely affected by the strong influence
which the larger incumbent bus operators exerted in the evidence given to the Royal Commission on Transport set up in 1928. The resulting 1930 Road Traffic Act made it "extremely difficult" for new firms to enter the industry (Chester, 1936, p.181). Additionally, the way in which the regulators interpreted "public need" contributed to the creation of a system of local monopolies, even though there was no statutory requirement on the regulators to do so.

More generally, there is the view that regulation itself is subject to economic laws (Stigler, 1971). Regulation may be supplied by a government if its ministers gain utility proportionate to the amount of resources they control. Demand for regulation may come from firms who perceive benefits in comparison to an unregulated industry (Posner, 1974). In situations where consumer interest groups are large and diffuse, small and tightly knit groups (generally those of producers) will tend to dominate (Peltzman, 1976). If producers expend resources in trying to achieve regulation in order to capture monopoly rents (Tullock, 1967, Posner, 1975), the costs of regulation will be higher than the deadweight loss associated with monopoly. These extra costs are sometimes termed the "Tullock-Posner" (TP) costs. The implication is that the costs of a monopoly created by regulation may be much higher than those identified in the public interest approach discussed in Section 3.1. The differences between the "public interest" model and the "economic theory of regulation" model within the specific context of UK transport are discussed in Rowley and Mulley (1983).

4.1.3 Coalitions

The weakness of the capture theory, in particular Peltzman's (1976) formulation of it, is that it assumes there are no economies of scale and thus no reason for regulation in the public interest (Keeler, 1984). It also fails to account for the cycle of regulatory policy in the UK local bus service market, specifically the dismantling of the regulatory system in the 1985 Transport Act. But the public interest theory of regulation takes no account of the ways in which regulators may be influenced by producers. A more general model based on game theory (Sharkey, 1983) can incorporate both the capture and public
interest theories as special cases. In this model, producers and consumers form coalitions in order to achieve and influence the implementation of regulation. If the "winning coalition" consists only of producers, then the model reduces to that of the capture theory. If the winning coalition comprises only consumers, it reduces to the public interest theory. More generally, however, coalitions will consist of both producers and consumers in varying degrees, and the composition and motivation of these coalitions may change over time.

Button (1989) applies this model to the UK bus industry, where the following groups are identified:

1. original bus operators
2. railways
3. bus manufacturers
4. public transport unions
5. potential entrants to the industry
6. travelling public
7. tax/ratepayers
8. general public.

In the late 1920s only group 5 was opposed to regulation. Groups 1 to 4 perceived regulation to be in their self-interest, whilst groups 6 to 8 favoured regulation on the grounds of network stability, internal cross-subsidisation and improved safety respectively. This winning coalition managed to influence not only the 1930 Road Traffic Act but also the 1968 Transport Act, in spite of the latter's introduction of external subsidies. By the 1970s, however, the coalition began to fragment because of the rapid increase in these subsidies and a more general political trend towards supply-side policies that emphasised the benefits of competition. In the late 1970s the winning coalition changed to comprise groups 6 to 8 with some support from 3. The theory of coalitions can thus encompass not only the capture and public interest approaches, but also the reasons for the life-cycle of regulatory policy.
4.1.4 The Durability of the Regulated System

An interesting question not explored in detail by Button (1989) is why the regulated system survived for so long. There seems little doubt that its inception was strongly influenced by producer interests, and that many of the benefits of the entry (as opposed to safety) regulation introduced in the 1930 Road Traffic Act accrued mainly to incumbent operators (see Section 2.3). It is surprising, therefore, that consumers (i.e. groups 6 to 8) acquiesced for so long in a system of quantity licensing which promoted, or at least did not diminish, the interests of incumbent producers. Whilst the chaotic development of the omnibus industry in the 1920s may have been fresh in the minds of consumers in the 1930s and 40s, and provided a strong impetus to maintaining quantity licensing, there seems little justification for this view in the 1950s and 60s. The initial growth spurt of the industry had disappeared, and many routes no longer had the same high levels of passenger density as in the 1920s. Naturally, consumers may have been influenced by the views of producers, whether or not such views were justified. As late as the mid-1970s, a director of the National Bus Company was expressing the view that:

"... without the licensing provisions of the 1930 Road Traffic Act public transport, municipal, "nationalised" and privately owned could not have developed into the orderly system as we know it today..."¹

It is also surprising that the system of internal cross-subsidisation promoted by the Traffic Commissioners after the 1930 Act did not engender earlier criticism by consumers. The aim of the system was to provide a comprehensive service network by enabling operators to use profits from their protected monopolies to subsidise loss-making services. The system avoids the use of external subsidies to the industry, so that it would appear to be an attractive option to group 7. Yet the system is open to a number of strong objections. First, in the absence of both demand and cost interdependencies, the deviation of price from marginal cost on profitable and unprofitable services causes allocative inefficiency. Second, a protected monopoly may be liable to productive (or "X") inefficiency, as well

¹ Quoted in footnote 26 of Button (1989, p. 501)
as dynamic inefficiency i.e. a lack of innovation. Third, the direction of cross-subsidy may be inequitable if it runs from low to high-income consumers. Fourth, decisions on cross-subsidy are not accountable to the electorate via Parliament or local government, as would generally be the case with an external subsidy.

It was not until September 1970 that the National Bus Company (NBC) instructed its subsidiaries to close loss-making rural services, unless external subsidy was forthcoming. This decision followed from the provisions of the 1968 Transport Act, which required the NBC to act as a commercial enterprise. In spite of this, cross-subsidy did not disappear from the industry, and the NBC estimated that its cross-subsidies amounted to roughly twice its external subsidies (see Section 5.1.4.2).

The continued quiescence of consumers to these aspects of the quantity licensing system was due in part to regulatory capture, especially in the earlier stages of regulation, but more latterly and fundamentally to the pervasiveness of the "Keynesian consensus" within the UK. Keynes' work had articulated the case for government intervention in the economy, and both Labour and Conservative governments from the late 1930s to the late 70s were broadly in agreement on the efficacy of demand management policies as a means of achieving full employment. Naturally, the parties differed on the extent of desirable government intervention, but not its fundamental rationale. Whatever the original meaning of Keynes' "General Theory", the Keynesian message that permeated to policy-makers was that the market mechanism could not be relied upon to achieve full employment and stable growth.

The result was an emphasis on macroeconomic policy to achieve full employment and stable growth within a regulated environment. The use of microeconomic (or supply-side) policy to achieve these objectives via increased competition played a lesser role in economic policy-making. There followed the growth of a protected public sector in which transport, and transport policy, could be used as part of a range of public policy measures to influence employment and growth. The ethos of bus and rail transport as a "public service", rather than one that should be operated under competitive conditions,
gained a powerful hold on consumers and producers alike. It was not until the Thatcher and Reagan governments of the late 1970s and early 1980s that policymaking shifted decisively towards the supply-side approach. With this shift came a renewed emphasis on the interests of consumers rather than producers.

The public choice approach can be used to provide another explanation of the long duration of regulation in the UK bus market. Although the approach provides a strong argument against regulation, it also implies that, once regulation is in place, there will be a tendency for it to persist (McCormick *et al.*, 1984). This is because once the "Tullock-Posner" (TP) costs associated with rent-seeking behaviour are incurred, they become sunk. Although deregulation may recover the "Harberger" (H) costs (i.e. area abc in Figure 4.1) (Harberger, 1954), the TP costs can never be recovered. In effect they represent wasted resources. If the H costs are small, the incentive to deregulate will be weaker and the duration of regulation longer. This conclusion depends on whether the TP costs are a one-off expenditure or must be incurred on a continuing basis, as well as the size of these continuing costs. The greater the frequency and size of these continuing TP costs, the more difficult it will be for producers to preserve the regulatory system. Given the favourable policy background to producer interests discussed above, however, it is likely that these continuing costs for UK bus operators, if indeed any existed, were relatively small.

4.1.5 Regulatory Failure

Regulatory failure occurs when the outcome of regulation creates a lower level of social welfare than that associated with a market solution. One example is where regulators are "captured" by producers. Although there are no estimates for the UK bus industry, Lubulwa (1990) quantifies the costs of regulatory failure in the Australian rail and road freight industries using Posner's (1975) measure of welfare loss. If the regulated price is the same as the monopoly price, this welfare loss can be measured as the deadweight loss (= abc in Figure 4.1) plus the monopoly profit $P_m acP_c$. This procedure assumes that the waste of resources incurred in seeking monopoly rent can be measured by the amount of
monopoly profit earned. Rent seekers are prepared to compete for monopoly rent, so that the winner will pay an amount up to the area \( P_mP_c \) in Figure 4.1. The principle would be the same for any regulated price (= \( P_r \)) between \( P_c \) and \( P_m \), with the welfare loss approaching zero as \( P_r \) approaches \( P_c \). Assuming a linear demand curve, the welfare loss \( L \) can then be estimated as:

\[
L = R_r(1-k) + 0.5 R_r(1-k)e(1-k)
\]

where \( R_r = F_rQ_r \) = revenue under regulation

\( k = F_b/F_r \)

\( F_r = \) rail freight rate under regulation

\( F_b = \) benchmark freight rate i.e. the rate charged under competition

\( e = \) absolute value of price elasticity of demand evaluated at \( F_rQ_r \).

In the case of the regulated rail freight industry, Lubulwa (1990) concludes that the size of the welfare loss (particularly from wasteful rent-seeking activities) is generally sufficient to justify deregulation, although there are problems in obtaining appropriate data in some cases. The case of road freight deregulation is more ambiguous, since it is estimated within a general equilibrium framework where the gains and losses are unevenly distributed. The model also produces the counter-intuitive result that deregulation increases the real wage (Lubulwa, *op. cit.*, p. 92).

A problem with this TP approach is that it may overstate the welfare loss from monopoly since the total expenditure caused by rent seeking is likely to be less than monopoly profit. In particular, rent seeking is not a competitive activity, so that monopoly profit will not be completely dissipated by rent seeking firms (Fisher, 1985). Uncertainty may also cause rent seeking expenditures to be lower than monopoly profits. If rent seeking is conducted in a multi-time-period model, the firm has the problem of estimating the present value of a stream of uncertain future monopoly profits. Given the presence of uncertain profits, the firm will tend to reduce its rent seeking activities in comparison to the usual analytical case of a single time period model where profits are known with certainty (Jadlow, 1985). By contrast, Wenders (1987) argues that the TP approach
significantly underestimates the welfare costs of monopoly, since it does not include rent defending costs. If these costs are included, the welfare costs of monopoly may be more than double those associated with the TP approach.

To conclude, the public choice approach offers an important alternative framework to that of the public interest approach discussed in Chapter three. It can be used to explain not only the reasons for regulation, but also the reasons for deregulation. In particular, it stresses that the welfare costs arising from monopoly may be much higher than those associated with the static allocative inefficiency of Chapter three. This conclusion has considerable relevance for the post-1930 bus industry, where regulation resulted in the creation of protected territorial monopolies.

4.2 The Austrian Approach to Market Efficiency

4.2.1 Introduction

The neoclassical emphasis on static efficiency conditions was not shared by Adam Smith and other classical writers, who regarded competition as a process of rivalry which took place over time (McNulty, 1967). This approach to competition as a dynamic process has been developed in the work of the Austrian School, particularly the role of the entrepreneur in the competitive process (Menger, 1871, Hayek, 1948, von Mises, 1949b, Schumpeter, 1950, Kirzner 1973, 1979). Schumpeter's (op. cit.) characterisation of competition as "a perennial gale of creative destruction" emphasises both the dynamic nature of competition and the role of the entrepreneur in seeking out new opportunities that cause the growth of some firms and the extinction of others.

The distinctive approach of the Austrian School can be illustrated in its analysis of monopoly. Following the influential paper by Harberger (1954), the standard view of monopoly was that it generated a (static) welfare loss equal to the area abc in Figure 4.1. Although Harberger's estimate of the welfare loss from monopoly was very small, a subsequent study by Cowling and Mueller (1978) found significant welfare losses.
Following Tullock (1967) and Posner (1975), Cowling and Mueller included not only the welfare loss from the exercise of monopoly power, but also the welfare loss from the waste of resources involved in trying to achieve monopoly power.

The Austrian view would reinterpret the monopoly situation in Figure 4.1 as characterising the behaviour of an entrepreneur who has discovered and sold a new product (Littlechild, 1981). For the moment, he enjoys monopoly power since he is the only seller of this product. But this monopoly power should not be regarded as a welfare loss, since "the relevant alternative to his action is no product at all" (Littlechild, op. cit., p. 358). In other words, the provision of a new product generates a welfare gain equal to the monopoly profit $P_{mac}$ plus the consumer surplus $abc$. In so far as the entrepreneur enables the product to be made available at an earlier date than would otherwise have been the case, there is an additional welfare gain. As rivals enter the market and compete away the monopoly profit by expanding output to $Q_c$, society benefits from the earlier provision of the extra output $Q_c - Q_m$ by the amount of the consumer surplus on this output i.e. area $abc$. The development of new products i.e. innovation, is referred to as dynamic efficiency. The increase in dynamic efficiency from monopoly should be contrasted with the static efficiency losses discussed in Section 3.3.1.

4.2.2 Monopoly, Mergers and Restrictive Practices

The process above assumes that entry to the industry by rival firms will eventually appear, so that monopoly power is transitory. Given the possibility of competition, monopoly profit would be more accurately termed "entrepreneurial profit" by the Austrian School, since it results from the successful exploitation of a new idea. As long as competition is possible subsequently, monopoly per se should not be regarded as causing a welfare loss. It is the existence of (temporary) entrepreneurial profit that provides the incentive for entrepreneurs to undertake investment. This leads to the conclusion that the proper role of government is not to regulate the monopoly, but to ensure freedom of entry to the industry. The Austrian School would thus oppose the creation of protected monopolies (such as the system of local monopolies which the
Traffic Commissioners encouraged after the 1930 Road Traffic Act) and the granting of patent rights by the government. It would also tend to oppose policies directed against mergers:

"If a single firm in an industry is not undesirable per se, nor is a merger to create one. Indeed, merging is one obvious way of competing. To prevent mergers is to protect other firms from competition, perhaps to prevent the adoption of some new technique or mode of organisation and thereby to impede the competitive process."

(Littlechild, 1986, p. 50)

This observation is particularly relevant to the deregulated bus industry, where a number of significant mergers began to occur some three to four years after deregulation (see Section 5.4.1). By 1994 the industry was dominated by six major companies, including Badgerline and GRT (Grampian Regional Transport) Bus. In April 1995 these two companies announced that they hoped to merge to form FirstBus, and that this company would continue to expand by acquisition. One of the main reasons given by the management of FirstBus for the merger was that the merged companies would be able to benefit from its "critical mass" in centralised purchasing from its major suppliers.

This type of merger is a problem for the MMC in that, whilst the merger may be considered a valid part of the competitive process to the extent that unit purchasing costs are reduced, there is a danger that monopoly power may be created in particular areas. Badgerline had no operations in Scotland, which is the traditional base of GRT. GRT had no operations in south-west England, Wales or either side of the Pennines, all of which had a significant Badgerline presence. Although both companies owned subsidiaries in the Midlands, it was in the East Anglia area where the possibility of monopoly power was greatest. Badgerline and GRT owned Eastern National and Eastern Counties respectively, both of which were major companies in East Anglia.

---

2 Reported in the Times Newspaper, 5.4.95., p. 21.
But market concentration in a given area is not necessarily a good indicator of market power. If the market is contestable, the ability of First Bus to earn monopoly profits in East Anglia will be constrained, irrespective of market structure. The critical factor is thus not the merger itself, but the extent to which it might reduce contestability. The issue of contestability is discussed separately in the following section.

The Austrian approach to restrictive practices also emphasises the importance of entry conditions to the market:

"Restrictive practices pose a dilemma. On the one hand, many practices such as exclusive dealing between manufacturers and suppliers or distributors are undoubtedly intended to prevent new entry. They thereby impede the competitive process, and it would seem that current laws against them might be strengthened. On the other hand, it may be argued that this is their very merit. Protection from competition provides an inducement to make large and risky investments which otherwise would not be made. On balance, most Austrians would probably not favour laws against restrictive practices, relying for protection once again on the possibility that new entry would overcome such barriers."

(Littlechild, 1986, p. 51)

For the Austrian School, therefore, "market failure" does not derive from defects in the market mechanism per se, but rather from the defects in the legal and institutional framework within which the market operates. There is no reason to suppose that government can improve upon the market, since it is unlikely to have the information that the market provides automatically to its participants. Government economic policy should be directed not to replacing the market, but to the reduction of entry barriers and the development of private property rights in order to create the conditions within which the market will function efficiently.

The economic arguments discussed above should be seen in the wider context of a desire to remove government restrictions on human activity as far as possible. The Austrian
approach to economic policy, although not part of the mainstream neoclassical approach and its associated welfare economics, has enjoyed a significant revival in the USA and UK. Most notably, the election of the Conservative government under Margaret Thatcher in 1979 marked the beginning of an extensive programme of privatisation and deregulation. In 1980 the Transport Act introduced the first privatisation of a public enterprise (the National Freight Consortium) and the first deregulation of a state monopoly (the express coach sector). As Kay and Thompson (1991) observe, "Transport can be regarded as the cradle of privatisation - or more generally of regulatory reform - in the United Kingdom." (op. cit., p. 19).

To conclude, the Austrian approach takes a Darwinian perspective towards the competitive process. The evolution of a monopoly represents the survival of the fittest firm(s), as long as entry to the industry is possible. Unlike the public interest approach, which regards monopoly as an instance of market failure, the Austrian approach emphasises the benefits that accrue from monopoly. But protected monopolies, such as those created in the UK bus industry following the 1930 Road Traffic Act, interfere with the competitive process and cannot be justified. These conclusions depend on the assumption that entry to the industry is possible. The analysis of entry (and exit) conditions forms the basis of a more recent development in industrial economics. This is the theory of contestable markets, which is discussed in the next section.

4.3 Contestable Markets

4.3.1 Introduction

A market is contestable if entry (and exit) barriers are sufficiently low to enable firms to enter and exit quickly:

"The crucial feature of a contestable market is its vulnerability to hit-and-run entry."

(Baumol, 1982, p. 4)
More precisely:

"We define a perfectly contestable market as one that is accessible to potential entrants and has the following two properties: First, the entrants can, without restriction, serve the same market demands and use the same productive techniques as those available to the incumbent firms ... Second, the potential entrants evaluate the profitability of entry at the incumbent firms' pre-entry prices."

(Baumol et al., 1982, p. 5)

The main policy implication of contestability theory is that in order to achieve competitive market behaviour the authorities should seek to influence entry and exit conditions within a market rather than its structure. The existence of monopoly is not in itself an indicator that monopoly profits are being earned: if the industry is perfectly contestable, hit-and-run entry will force price and output to the levels associated with a perfectly competitive industry. It is the extent of potential competition for the market, rather than the number of firms within it, which should be the focus of public policy.

4.3.2 Natural Monopoly

Contestability theory implies that natural monopolies may be forced to price at competitive levels if the market is contestable (Baumol et al., 1982, p. 209). This insight was an important factor in the debate concerning airline deregulation in the USA (Waterson, 1988, p. 36). Although a natural monopoly can be defined in more than one way, the generally accepted definition is due to Baumol et al. (1977):

"By a natural monopoly we mean an industry whose cost function is such that no combination of several firms can produce an industry output vector as cheaply as it can be provided by a single supplier".

(Baumol et al., 1977, p. 350)
A necessary and sufficient condition for a natural monopoly is (strict) subadditivity i.e.

\[ C(Y) < \sum_{i} C(y_i); \quad Y = \sum y_i, \quad i = 1, 2, \ldots k; \quad k \geq 2. \]

If the natural monopoly is perfectly contestable, controls on price are superfluous: government policy should aim towards preserving free entry and exit. Weitzman (1983) argues, however, that the analysis of Baumol et al. relating to natural monopoly contains an internal inconsistency. In a perfectly contestable market, scale economies must arise through fixed rather than sunk costs, since the existence of sunk costs will create exit barriers that violate the assumption of perfect contestability. Weitzman claims that without sunk costs there cannot be fixed costs, and without fixed costs there cannot be scale economies. In the absence of sunk costs, market entry (however brief) can occur at minimum efficient scale and, provided the product is storable, all production will take place at minimum efficient scale. But if this is the case, returns to scale are in effect constant. An example would be a consortium of firms buying an optimal size computer and then time sharing its facilities. Weitzman’s argument depends on how sunk costs are defined, however. Some definitions do not preclude the existence of both zero sunk costs and substantial fixed costs. Baumol et al. (1982, pp. 280-1) discuss various definitions of sunk costs.

A further problem is that not all perfectly contestable natural monopolies are sustainable. Faulhaber (1975) shows, in the context of a natural monopoly arising from economies of scope in the supply of water, that freedom of entry can destroy the properties associated with the natural monopoly. In this case, a perfectly contestable natural monopoly is non-sustainable. Another example of non-sustainable natural monopoly is where a natural monopoly in the public provision of transport is undermined by competitive entry of private transport (von Weizsacker, 1984).

The model of contestable markets assumes that entrants are price takers, in the sense that their entry decision is made on the basis of a given market price. The entrant is not permitted to take account of incumbents' reactions, and must assess the profitability of
entry at the pre-entry market price. This rules out the possibility of allowance for strategic entry deterrence (e.g. predatory pricing) by the incumbent. The assumption of price-taking behaviour is plausible where there are a large number of sellers, but less so for the case of small numbers (Friedman, 1982, pp. 527-8). In industries subject to significant scale economies, such as the natural monopoly case above, contestability theory may have limited relevance. Baumol et al. argue that their model will still be relevant in situations where entrants are not price takers, if they can negotiate pre-entry long term contracts which protect them from price retaliation by incumbents. As Shepherd (1984) has noted, however, the option of pre-entry contracting will also be available to incumbents, who could defeat the entrants' strategy by promising to "undercut any legitimate price" (p. 576, fn. 12).

4.3.3 Imperfect Contestability: the role of actual and potential competition

If a market is perfectly contestable, then market structure will have no effect on market performance. The threat of entry will force even a (sustainable) natural monopoly to set price and output at perfectly competitive levels. If market structure does influence performance, then the industry cannot be perfectly contestable. A number of empirical studies of the deregulated USA airline industry have found a correlation between the number of actual competitors and performance (Graham, Kaplan and Sibley, 1983, Call and Keeler, 1985, Moore, 1986). Using similar methodologies, Bailey and Panzar (1981) conclude that in some cases there is no correlation, and attribute this to the presence of potential competition.

But contestability is not an all-or-nothing concept. It is possible for both potential and actual competition to influence performance. The difficulty is to provide some measure of potential competition. Without a measure of potential competition, studies which find a correlation between actual competition and performance may suffer from a specification error: the coefficient on the market concentration variable will be biased upwards if it captures the influence of potential as well as actual competition. Morrison and Winston (1987) specify a model of the deregulated USA airline market in which both
actual and potential competition contribute to changes in market performance and social welfare. The conclusion is that each additional actual competitor reduces the difference between optimal and actual welfare per passenger by $0.0044 per mile, whilst each additional potential competitor reduces it by $0.0015 per mile. Three potential competitors thus have roughly the same impact on welfare as one actual competitor. These results support the view that the deregulated USA airline market is imperfectly contestable.

The results are also consistent with the shift in opinion among proponents of contestability, who originally argued that the deregulated USA airline market was a close approximation to a perfectly contestable market. In 1981, for example, Bailey and Panzar observed that:

"... there is no reason, \textit{a priori}, to expect that economies of scale should lead to substantial barriers to entry in the airline industry because airline capital costs, while substantial, are not \textit{sunk costs} ... the major portion [i.e. aircraft] can be "recovered" from any particular market at little or no cost. ... Thus, despite substantial natural monopoly attributes, most airline markets are likely to be readily \textit{contested}. This fact ensures that, \textit{even if} actually operated by a single firm, their performance should approach the competitive norm ..."

\cite{BaileyPanzer1981}

By 1986 this view had been substantially modified:

"In the initial enthusiasm with which we described contestability analysis we ... more than once cited the airline industry as a case in point ... Reconsideration has led us to adopt a more qualified opinion on this score ... trucks and buses do not face the heavy sunk costs involved in the construction of airports or the shortage of gates and landing slots at busy airports ..."

\cite{BaumolWillig1986}
Nonetheless, it is clear from this quotation that the bus industry is regarded as highly contestable. The next section discusses the evidence for the assumption of contestability in the bus industry by examining the nature of its entry and exit barriers.

4.4 Entry and Exit Barriers in the Bus Service Industry

The dominant theme of Sections 4.2 and 4.3 is the importance of entry and exit barriers in determining market behaviour. For the Austrian school, monopoly profit is a reward for exploiting new opportunities, but freedom of entry will erode these profits. The contestable markets theory emphasises both entry and exit barriers as determinants of market behaviour. This section examines the nature of the entry and exit barriers that exist in the bus service industry generally, and the UK bus service industry in particular. Section 4.4.1 discusses the general nature of entry barriers, whilst Section 4.4.2 deals with a number of refinements to the concept of entry barriers. Section 4.4.3 deals with sunk costs and the asymmetry they cause between incumbent and entrant.

4.4.1 Entry Barriers: the Bain approach

Following Bain (1956, p. 10), an entry barrier exists when the incumbent has a cost advantage over entrants. This cost advantage can be measured by the difference between the competitive price and the entry-inducing price. The greater the cost advantage, the higher the incumbent can raise price above the competitive price without inducing entry. Entry barriers are determined by economies of scale, product differentiation and absolute cost advantages.

Economies of scale form an entry barrier if the minimum efficient scale of production is large relative to market demand. In the limiting case, the industry may consist of one seller only i.e. a natural monopoly. A potential entrant with access to the same production technology as the incumbent firm(s) has two unpleasant choices. First, entry could take place at a level of output sufficiently high to produce at minimum efficient scale, but this could depress the market price and may leave the entrant unable to make a profit. Large
scale entry may also incur retaliation from the incumbent. Second, entry on a small scale would leave the entrant operating at a cost disadvantage to the incumbent.

Product differentiation, via advertising and the promotion of brand loyalty for example, may act as an entry barrier because the unit costs associated with entry into a number of different sub-markets are likely to be higher than those for entry into the one market only. It should be emphasised that the product differentiation discussed here is not the same as that in Section 3.3.3.2, where differentiation occurs not through advertising but through genuine differences in either schedule times or product quality. An advertising campaign for a single undifferentiated brand will spread fixed costs over a relatively large output. This will not be the case where separate campaigns are needed for each of the relatively small differentiated product markets. Schmalensee (1978) cites and models the example of the USA ready-to-eat breakfast cereal market, where there is considerable product differentiation by incumbent firms, but where successful entry by new firms is rare. More recent models of product differentiation (Farrell, 1986, Bagwell, 1990) emphasise the role of imperfect information about different product characteristics in creating entry barriers.

Absolute cost advantages occur where the incumbent has a lower cost curve than the one facing the entrant. This could occur because the incumbent has patent rights, access to lower cost suppliers or sources of finance, or the advantage of the lowest cost location. In these circumstances, an entrant will always be at a cost disadvantage irrespective of the scale of entry.

4.4.2 Entry Barriers: alternative approaches

The problem with Bain's definition of an entry barrier is that it includes cases where entry could still be possible if incumbent and entrant have access to the same post-entry cost curve. In the case of economies of scale, for example, entry could occur at the same level of output as the incumbent so that both entrant and incumbent will have the same unit
costs. Even if entry were to reduce market price, this would affect both entrant and incumbent equally. Stigler (1968) prefers to restrict the definition of an entry barrier:

"A barrier to entry may be defined as a cost of producing (at some or every rate of output) which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry."

(Stigler, op. cit., p. 67)

Under this definition, economies of scale and product differentiation are only entry barriers if the entrant's post-entry cost curve is higher than the incumbent's. One reason why the entrant may have a higher post-entry cost curve than the incumbent is that entrants may have a systematically higher cost of capital. Incumbents have a trading record against which financial markets can assess the probability of success, whereas no such benchmark exists for entrants. In effect, financial markets are faced with a "lemons" problem (Akerlof, 1970), so that the cost of capital will be higher for entrants than incumbents. This differential cost of capital argument may be one reason for the tendency of operators with little or no track record to enter the subsidised part of the UK local bus service network in the post-deregulation period. If the contract is awarded on a full cost basis (see Appendix one, Section 1.9.4) the trading risk is borne by the tendering authority rather than the operator. The operator then avoids the lemons problem above.

Although it is not possible to reduce risk in this way in the competitive part of the market, the lemons problem may be overstated. This is because "new" entry often comes from experienced operators. At a national level, Stagecoach's switch from coach to bus operation is an obvious example. At a local level, an existing operator of schools contracts may enter the competitive sector or vice versa, or expand into another geographical area. Financial markets can thus make some assessment of trading risk for this type of entry.
Entry barriers from product differentiation via advertising may also be overstated, since advertising forms a relatively small part of total expenditure in the UK local bus service market. Similarly, absolute cost advantages are not likely to be a significant entry barrier. Road infrastructure is provided to all bus operators on equal terms, and operators have non-discriminatory access to bus terminals under the provisions of the 1985 Transport Act. Some operators, however, may have an absolute cost advantage in terms of depot location. This could be a significant factor in areas such as London, where potential entrants face very high land costs.

An implication of the Stigler definition is that fixed costs do not constitute an entry barrier where the post-entry cost curve is the same for entrant and incumbent, since the entrant can always achieve the same unit costs as the incumbent. In this case, the existence of fixed costs *per se* does not impose any extra costs on the entrant relative to the incumbent.

Demsetz (1982) argues that the traditional emphasis on the cost aspects of entry barriers is misleading, because their costs may not be defined correctly. Some entry barriers may be socially desirable if they represent costs that have arisen from situations such as building up a good reputation or bearing the risk of innovation:

"Licensure, trademark, copyright, patent, entitlement to the fruits of past investment, including the investment in an honorable history ... may or may not be desirable, depending on how these broad implications are valued."

(Demsetz, 1982, p. 56)

The concept of social desirability implies that both the costs and *benefits* of entry barriers should be taken into account. Von Weizsacker (1980) develops a model in which the welfare effects of entry barriers can be estimated from the specification of appropriate demand and cost curves. An entry barrier in this model is defined as a higher cost of
production incurred by an entrant but not by the incumbent (following Stigler's (1968) definition), and which implies "a distortion in the allocation of resources from a social point of view." (von Weizsacker, 1980, p. 400). Although fixed costs create a welfare distortion in this model, they are the same for entrants and incumbents and so do not qualify as an entry barrier. *Ceteris paribus*, it is the existence of *sunk costs* that creates an entry barrier i.e. those costs that cannot be recovered if a firm exits from the market. It can be shown that, using a Cournot equilibrium model, a dominant strategy for a potential entrant in a two-stage game where sunk costs exist is for no entry to occur (von Weizsacker, 1980, Martin, 1993, pp. 180-1)

4.4.3 *Sunk Costs*

The von Weizsacker model provides a rationalisation of the important asymmetry between incumbent and potential entrant in terms of sunk costs. In a market where no entry has yet occurred, all potential entrants must take into account the possibility that some of their costs may not be recoverable on exit. In this sense, all potential entrants face the same problem. Once entry occurs, the incumbent firm or group of firms incurs sunk costs which will no longer influence its future decisions. But this is not the case for firms that are still considering entry. The possibility (rather than the actuality) of sunk costs must still be taken into account in the entry decision. *Ceteris paribus*, potential entrants will always face higher costs than incumbents where the latter has incurred sunk costs. Another aspect of this asymmetry is that, having incurred sunk costs, the incumbent's incentive to stay in the industry is likely to be greater than the incentive to enter it:

"The difference between incumbent firms and entrants is that incumbent firms own plant and equipment specific to this industry and thereby are committed to continue in this industry, whereas this is not the case for a potential entrant. It is thus not just simple economies of scale which may cause a barrier to entry, but rather economies of scale in combination with irreversible capital commitments."

(von Weizsacker, 1980, p. 401)
The extent to which costs are "sunk" is determined by factors such as the transferability of capital equipment from one market to another (e.g. from short to long distance bus service work) and the existence of a developed second hand market in which that capital equipment can be sold. Interestingly, a study of USA manufacturing industry concluded that investment in machinery and equipment appears to be more "sunk" than investment in buildings (Kessides, 1990). This implies that investment in machinery and equipment is more likely to be industry specific than investment in buildings. The "fixity" of the investment, however, is not always a reliable guide to whether it is sunk or not: as always, much depends on the specific circumstances for which it is made. A supplier of water may have to undertake earthworks for which, in the event of exit from the industry, there may be few, if any, alternative uses (Waterson, 1988, p. 146). Sunk costs can include not only the costs of capital equipment, but also the costs of building up knowledge, reputation, goodwill and human capital in areas such as job training.

A development that involves building up a reputation and a corresponding sunk cost is that of quality partnerships in the local bus service market. The Transport Bill (Dec. 1st, 1999) states that:

"A local transport authority, or two or more such authorities acting jointly, may make a quality partnership scheme if they are satisfied that the scheme will to any extent implement the policies set out in their current bus strategy or strategies and will- (a) improve the quality of local services provided in the whole or any part of their area, or combined area, by bringing benefits to persons using those services, or (b) reduce or limit noise or air pollution."

(Transport Bill, Dec. 1st, 1999, Section 96)

Typically, quality partnerships involve agreement with operators on quality of service in exchange for local authority initiatives to improve the operating environment for buses, such as better bus stops and shelters and measures to give priority to bus traffic. At the time of writing, the exact nature of the partnership depends on the local authority and is non-statutory. The Transport Bill introduces various proposals, such as a minimum
period of five years for the partnership and the introduction of quality contracts, which will formalise to a greater extent the structure of such agreements. An important aspect of these arrangements is the synergy that such partnerships (or contracts) may provide: cooperation may provide a better outcome than that associated with local authorities and operators working independently of each other.

But these partnership (or contract) arrangements may reduce contestability. Once an operator has invested to meet the standards set in the partnership (or contract), then potential entrants face the asymmetric entry cost problem identified earlier in this section. In addition, smaller operators may not be able to afford the standards specified in the quality partnerships. This situation could arise because small operators cannot take advantage of the type of economies discussed in the previous section, or because large operators have implemented a deliberate strategy to raise their rivals' costs. This latter case is an example of a strategic entry barrier and is discussed further in Section 4.6.2.

The Buses White Paper considers that the bus service market is "highly contestable", citing the reasons that:

a) costs of entry are low
b) sunk costs are comparatively small in relation to operating costs
c) economies of scale are limited.

But these criteria seem to refer to non-strategic entry barriers. Even if the three conditions above are satisfied, a market may not be contestable if there is strategic entry deterrence on the part of incumbent firms. This form of incumbent firm behaviour is discussed in Section 4.6.

---

3 In contestability theory, economies of scale do not create an entry barrier provided the product is storable. Bus services are non-storable (i.e. seats unoccupied on a given journey cannot be resold at a later date) and economies of scale will constitute an entry barrier. For a discussion of this point, see M.C. Sawyer. The Economics of Industries and Firms, 2nd.ed., p. 251.
To conclude, economies of scale and product differentiation do not create entry barriers if entrant and incumbent have access to the same post-entry cost curve. This is an important point, because scale economies and product differentiation have often been cited as necessarily creating entry barriers. In practice, however, it is unlikely that both incumbent and entrant will have the same post-entry cost curves, so that these features may act to some extent as an entry barrier. On the other hand, there is no reason to suppose that, in the absence of sunk costs, the entrant's costs are very much higher than the incumbent's, so that the extent of entry barriers may not be significant. In any event, it is unlikely that product differentiation (via advertising) or absolute cost advantages are significantly large to act as entry barriers in the UK local bus service market.

The critical determinant of entry barriers is the existence of sunk costs. The existence of sunk costs introduces an asymmetry between incumbent and entrant that acts as an entry barrier. In effect, sunk costs raise the entrant's post-entry cost curve above that of the incumbent. Sections 4.5 and 4.6 review the evidence on non-strategic (or innocent) and strategic entry barriers respectively.

4.5 Innocent Entry Barriers in the Bus Industry: the empirical evidence

Salop (1979) distinguishes between innocent and strategic entry barriers. Economies of scale are an innocent entry barrier, since they are determined by the nature of a firm's production function rather than as a result of a deliberate act by a firm. Absolute cost advantages may be innocent: existing firms may already have located in the least-cost sites, for example. On the other hand, firms may have undertaken investment or research and development specifically aimed at the prevention of entry. Such costs may be sunk. If they are, this acts as a further deterrent to entry, since the incumbent firm has announced a commitment to stay in the industry. A strategic entry barrier is one that is deliberately created by a firm to deter entry. The remainder of this section discusses the various forms of innocent entry barrier that may exist in the bus industry. Strategic entry barriers are discussed in Section 4.6.
Section 4.4.3 concluded that economies of scale could act as an innocent entry barrier to the bus industry. Section 4.5 examines the evidence on economies of scale (and other types of cost economies that are possible entry barriers), and sunk costs in the bus industry.

4.5.1 Economies of Scale

In terms of scale economies, empirical studies of the bus transit industry provide a mixed set of results. McGillivray et al. (1980) find evidence of constant returns to scale in bus transit operations in the USA, a result supported by the findings of Williams (1981). Other studies (Viton, 1981, Williams and Dalal, 1981, Obeng, 1985) find evidence of scale economies. A major problem in comparing these studies is the differences which exist in their model specification, the measures of output used, the measurement of factor prices, the use of short or long run analysis, the econometric tests conducted, and the type of data used (Berechman, 1987). A review of developments in transportation cost function research is provided by Oum and Waters (1997), who emphasise the difficulties of discriminating between economies of scale and economies of density.

In a study of the Israeli bus transit sector Berechman (1987) concludes that the sector as a whole is subject to scale economies, although constant returns to scale may exist for specific segments and measures of output e.g. intra-urban trips using vehicle-kilometres as the output measure. A further finding of this study is that the price-elasticity of the demand for labour was low and that the amount of labour employed was unresponsive to the decline in total vehicle-kilometres. This implies that the possibilities for labour cost savings are relatively small in this industry. Given that the production technology of the bus transit industry is labour intensive, this result underlines the difficulties faced by bus operators in controlling costs.

The empirical evidence on the existence of scale economies in UK bus operations is also mixed. In a cross-sectional study of 44 UK municipal bus operators, Lee and Steedman (1970) reach the general conclusion that returns to scale were constant, although they did
not rule out the possibility of scale economies in larger undertakings such as the (proposed) Passenger Transport Authorities (PTAs). They argue that the main source of unit cost reductions lay in the extension of one person operation (OPO), rather than amalgamation via the creation of PTAs, because of the high proportion of labour to total costs in bus operations. According to Button (1989, p. 75), it would appear that the White Paper on Buses implicitly used the results of this study in arguing the case for limited scale economies in UK bus operation. The importance of OPO as a source of unit cost reduction is also supported in a study by Brown and Nash (1972), which estimated a 13.7 percent cost saving from its introduction. If the production technology of UK and Israeli bus transit is similar, then these arguments are contradicted by the results of Berechman's (1987) study.

Wabe and Coles (1975) find evidence of diseconomies of scale with respect to the size of bus fleet, although their study does not include variables on peak/off-peak differences, factor price differences, congestion and differences between types and sizes of vehicles. There may also be a problem of multicollinearity in the model specification, since bus fleet size and bus fleet mileage are both included as explanatory variables. Button and O'Donnell (1985) investigate the cost structure of British urban bus services using a flexible functional form (the transcendental or translogarithmic function) which does not impose restrictions on the shape of the firm's unit cost curve. This is in contrast to most of the earlier studies mentioned. Whilst there is evidence of scale diseconomies for some of the larger district bus undertakings, there is also evidence to suggest that smaller operators enjoyed scale economies.

Another source of economies of scale in bus operations is the size of vehicle. Ceteris paribus, the larger the vehicle the lower the cost per seat-mile. Webster (1968) estimates that increasing the size of vehicle from 10 to 20 seats reduces the cost per seat-mile by 40 to 50 percent. Further increases in vehicle size continue to yield cost reductions, although by less substantial amounts. This is unlikely to create an entry barrier, however, since

---

4 Strictly speaking, the economies are those of density rather than scale, since it is implicitly assumed that the size of the network remains fixed as vehicle size increases.
different sized vehicles are readily available to all operators: they can be purchased new or second-hand, leased, or transferred from other services.

More recently, Stagecoach Holdings has claimed that its costs are at least 20% lower than most operators because of economies of scale. One significant source is the access to cheaper equity finance following its listing on the London Stock Exchange in April 1993, as well as lower rates for bank borrowing, insurance, and bulk buying of fuel and spare parts. This claim was used to justify its "competitive" pricing policies, although some of its competitors claimed that it was practising predatory pricing (Transport Committee First Report, 1995, Vol. III, p. 358).

Mohring (1976) argues that the main economies associated with urban public transport are the savings in passenger waiting times that accrue from high service frequencies. These user cost reductions arise when service frequencies are so high that passengers do not try to catch a particular service, but arrive at a bus stop in the knowledge that any waiting time will be relatively short. Given a ten-minute headway and random arrival of passengers, expected waiting time per passenger will be five minutes. If service frequency and the number of passengers is doubled (thus maintaining a constant queue length) operators' unit costs will remain roughly constant, assuming constant returns to scale on the production side. But each pre-existing passenger will now have an expected waiting time of only two-and-a-half minutes, and so receives an external benefit from the increase in service frequencies. Since these user cost reductions do not affect production costs, however, the "Mohring effect" will not act as an entry barrier (although it is an instance of market failure).

To conclude, the evidence on scale economies is mixed. One highly plausible explanation for this indeterminacy is the failure to distinguish adequately between economies of scale and economies of density. The next section examines the evidence for economies of density in bus transit, and how the evidence for scale economies is affected if density economies are taken into account in the model specification.
4.5.2 Economies of Density

A problem with many studies on economies of scale in bus operations is that they fail to make an important distinction between economies of scale and economies of density. This failure has created an inconsistency between many of the results of empirical studies and the actions of public policymakers (Windle, 1988). Scale economies arise as a result of falling average costs when output increases and more of all factor inputs are employed (strictly speaking, the increase in factor inputs should be equi-proportionate). Economies of density are associated with falling average costs resulting from increased factor utilisation. In particular, testing for economies of scale using bus miles as a measure of output fails to control for changes in the density of service. Bus miles tend to be used as an output measure because of the difficulty in obtaining data on passenger miles.

Using passenger miles as a measure of industry output, returns to scale arise where the size of the network (i.e. route miles) increases equi-proportionately to passenger miles, whereas returns to density occur when passenger miles are increased within a fixed network. Economies of density can arise either from an increase in the number of passengers on existing buses and routes, or from an increase in the number of buses on existing routes (or, alternatively, increased capacity through larger buses). More precisely, for a cost function defined in terms of the following variables:

\[ C = f(Q, P_i, Z_k) \]

where \( C \) = total cost
\( Q \) = output (passenger miles)
\( P_i \) = input prices
\( Z_k \) = output characteristics (e.g. load factor, average vehicle speed, average length of trip, number of route miles)

then, assuming a translog functional form (Christensen, Jorgensen and Lau, 1973), returns to scale (RTS) can be defined as the relationship between costs and changes in passenger miles and network size, holding load factor, average vehicle speed, average trip length and input prices constant.
where \( M = \) route miles. Returns to network density (RND) are defined as the relationship between costs and passenger miles, where changes in passenger miles arise from changes in the number or capacity of buses, holding load factor, average vehicle speed, average trip length, input prices and network size constant:

\[
RND = \left[ \frac{\partial (\ln C)}{\partial (\ln Q)} \right]^{-1}
\]

Returns to passenger density (RPD) are defined as the relationship between costs and passenger miles, where changes in passenger miles result from changes in the number of passengers on a given number of buses and network size, holding average vehicle speed, average trip length and input prices constant:

\[
RPD = \left[ \frac{\partial (\ln C)}{\partial (\ln Q)} + \frac{\partial (\ln C)}{\partial (\ln Z_L)} \right]^{-1}
\]

where \( L = \) load factor. Evaluated at the sample means, Windle (1988, p. 130) estimates \( RTS = 1.03, \) \( RND = 1.25 \) and \( RPD = 3.13. \) These results imply constant returns to scale, significant economies of network density, and large economies of passenger density for urban bus transit in the USA.

The absence of significant scale economies is supported in a study by Hensher (1988), although this study had to use passenger revenue rather than passenger miles as the measure of output. Hensher (op. cit.) also finds that total operating costs for public sector operators were 64% higher than those of private operators, and that driver costs were over twice as great. These results are qualitatively similar to those of Windle (1988), which finds that Federal capital grant subsidies have created over-investment in capital
equipment and substantial excess capacity: in order to minimise costs a reduction of 50% in capital would be required.

The empirical finding of constant scale returns in bus operations implies that unit costs would not increase as the level of output falls. This is contradicted by the experience of the USA bus transit industry where, between 1945 and 1972, patronage fell by 55 percent and real fares rose by 200 percent. In view of the number of bus company failures over this period, it is likely that real unit costs rose even more rapidly than real fares. This experience is consistent with the view that the industry is characterised by economies of scale, and this is the view (whether valid or not) that has motivated public policy:

"The lack of credibility of the constant returns finding is reflected in public policy which has dealt with the bus industry as if it were a natural monopoly."

(Windle, 1988, p. 120)

Windle's results reconcile the anomaly between empirical results and public policy towards urban bus transit. Returns to scale are constant, but economies of density are not. As passenger numbers and bus miles fall at a faster rate than network size, unit costs will increase because of decreasing density of service. These results certainly have relevance to the British local bus service sector where, between 1953 and 1980, passenger journeys fell by 60% and bus miles by 18%. Although figures are not available for route miles i.e. network size, it is unlikely that they contracted as fast as passenger journeys, since the Traffic Commissioners generally attempted to maintain a network which included both profitable and unprofitable routes (see Section 2.2). Although fares over this period rose by 86% in real terms, unit costs rose by an even greater amount, as demonstrated by the increase in subsidies and grants of 511% in real terms from 1969 to 1980.

To conclude, the existence of economies of density implies that individual routes within a network have the declining unit cost characteristic of a natural monopoly. This characteristic results from the presence of costs that are fixed but not necessarily sunk. If
there are no sunk costs, then the route is contestable and regulation is unnecessary. The estimated values for economies of density, particularly that for passenger density, imply that long run equilibrium on an individual route will be associated with a small number of operators (possibly only one).

The existence of constant scale returns (i.e. where an X% increase in both passenger miles and network size will increase total costs by X%) provides further support for a market approach to urban bus transit. Public policy was based on the view that bus transit was subject to significant economies of scale, so that regulation was necessary to overcome the entry barriers created by these scale economies. This view was mistaken on two counts: first, there are no significant scale economies in bus transit (although there are significant economies of density); second, even if scale economies existed, it is the presence of sunk, not fixed costs which create the entry barriers.

4.5.3 Economies of Scope

The discussion so far has been based on the concept of a single product industry. The analysis needs to be extended in the case of bus services, where more than one product can be supplied by a firm. Economies of scope may arise from increasing the number of products that a firm provides (Panzar and Willig, 1975, Bailey and Friedlander, 1982).

An example of multi-output production would be the provision of local (bus) services and long-distance (coach) services by a single firm instead of two specialist firms. The importance of the links between bus and coach operation is shown in the Trial Areas, where entry into the local bus service market was invariably by established coach operators (Savage, 1985, p. 42). At the time of writing, at least one bus company was considering diversifying into the provision of privatised rail services in the UK. A bus operator could also provide subsidised services e.g. schools contract services, in conjunction with commercial services. Another example from the transport sector would be the provision of passenger and freight services by a railway company. Although

---

5 The acquisition of NBC subsidiaries by the Perth-based Stagecoach company is an illustration of this point. Stagecoach started in the express market and diversified into the bus market.
outward and return journeys are sometimes cited as multi-output production, they are really a case of joint supply where production of one output generally necessitates the production of the other. It would seem preferable to restrict the concept of economies of scope to the case of outputs that are voluntarily provided by the firm.

Economies of scope are conceptually distinct from economies of scale, which arise from the increase in output of a given product. Scope economies exist if the minimum cost (= \( C \)) of producing output quantities \( q_1 \) and \( q_2 \) of goods 1 and 2 is lower in combined rather than separate production i.e.

\[
C(q_1,q_2) < C(q_1,0) + C(0,q_2)
\]

Economies of scope arise because different outputs may have shareable inputs that can be used more efficiently by a multi-product firm. One of the most important sources of scope economies within bus operations is probably that of scheduling efficiency, since computerised techniques have been developed to allow operators to schedule services more efficiently over different routes e.g. profitable and unprofitable (Rousseau, 1985).

In order to test for economies of scope, a generalised translog function allowing for a Box-Cox transformation on the output variables can be used (Zarembka, 1974). The Box-Cox transformation allows for the inclusion of zero output levels in the cost function i.e.

\[
Y_i = Y_i^{\lambda_i} / \lambda_i \text{ for } \lambda_i \neq 0, \text{ and } \\
Y_i = \ln Y_i \text{ for } \lambda_i = 0.
\]

In a study of USA railroads, Kim (1987) concludes that it is subject to mild overall economies of scale, but significant product-specific economies of scale with respect to passenger and freight services. Surprisingly, there are diseconomies of scope with respect to their joint production. One reason is that the different speeds of passenger and freight services can cause serious scheduling problems. Another is that track must be heavy
enough to withstand freight traffic, and smooth and straight enough to ensure passenger comfort. Such track can be very expensive.

The existence of multiproduct firms will increase the degree of contestability for any particular market. Cairns and Mahabir (1988) argue that entry is most likely to occur from firms with sunk assets in related markets, and suggest that contestability theory is best suited to the analysis of competition among multiproduct firms. It should be noted that the analysis of a multiproduct natural monopoly is more complex than that of a single product monopoly. Achieving economies of scope may imply losing economies of scale for individual products. Additionally, there will be a number of ways in which a multiproduct firm can set price equal to marginal cost on average, and each combination of prices will have different implications for social welfare (Waterson, 1988, pp. 19-20).

4.5.4 Cost Economies, Network Effects and Mergers

Empirical studies on the existence of scope economies in UK bus operations seem rare (perhaps non-existent). Nonetheless, the development of service networks by a small number of large bus operators after deregulation (see Section 4.6.1) provides some evidence for the presence of scope, as well as density economies in the industry. The role of scope and density economies in the evolution of a service network is illustrated by the development of "hub and spoke" operations in the airline industry (Brown, 1991). In Figure 4.2 below, the hubs are the large cities of X and Y, and the spokes are the smaller cities of a, b, c, d and e. Flights between X and Y are profitable because of the high traffic densities. There are no direct flights between the spokes because the low traffic densities imply relatively high costs of operation. These costs may be sufficiently high to render some direct inter-spoke flights (e.g. a to b) unprofitable. By channelling this traffic via the hubs, where traffic densities are much higher, load factors can be increased. Alternatively, larger aircraft can be used, where their cost per seat mile is generally lower than that of smaller aircraft. This "feeder" system enables economies of scope and density to be obtained, so that a connecting flight from a to b via X, for example, may become viable within a network system. The hub acts like a warehouse in a retail distribution
system, collecting bundles of passengers from different places and matching them to destinations.

The increase in load factors provides economies of passenger density. The use of larger aircraft provides economies of network density. Economies of scope arise at the hubs from a more intensive use of sharable inputs by different types of traffic: one example would be the use of expensive computer systems for flight scheduling. Although the deregulated USA airline industry is often cited as an example of a hub-and-spoke system, it has appeared in other transport systems such as European shipping and UK inter-city bus services. In the latter case, Victoria Coach Station in London acts as the hub for the express coach network. The development of these service networks is not contingent on regulation: in each of the cases mentioned, the network has developed in a deregulated, competitive environment. Thus regulation is not necessary for the provision of an integrated service network, where integration here refers to the provision of connecting services (see Section 3.3.2).

![Figure 4.2: A hub and spoke system](image-url)
There are parallels with local bus services, where the spokes could be regarded as part of a low density network, possibly outlying rural areas, and the hubs as part of the high density urban network. An implication of the move towards such systems is that market concentration over the total network will increase, since the economies of density and scope mentioned above will only accrue to large firms. One example is that of Fife, an extensive rural area with a few large hub towns, where roughly ninety percent of buses are operated by Fife Scottish, the Stagecoach subsidiary. Stagecoach has argued that its near monopoly position can be justified by its provision of a comprehensive network of services, which could not be maintained if other operators entered the market (see Section 4.6.1).

Mergers have been a significant feature of the post deregulation bus market, and the network argument above provides one plausible explanation for this. The deregulated USA airline industry also experienced considerable merger activity: by 1988, the 8 largest airlines had gained effective control over 48 of the 50 largest regional carriers. A significant factor in this merger activity was the desire to take advantage of network effects (Fawcett and Farris, 1989). The importance of network effects should not be over-emphasised, however. Even in the context of the USA airline hub-and-spoke system, Viton concludes that:

"In the context of optimality, it was found that although the influence of networks of different extents was perceptible, it was not great."

(Viton, 1983, p. 609)

One reason may be that diseconomies of scope or density can arise through increased use of hubs: congestion has increased at USA hub airports during peak hours (Fawcett and Fawcett, 1988).

Another way of rationalising the merger activity is to view it as a method of internalisation within the market failure paradigm discussed in Chapter three.
Competition between many small operators will be Pareto-inefficient if the external benefits of connecting services (and through ticketing) are not taken into account. Merger activity will enable the resulting large firms to provide, via networks such as the hub and spoke system, connecting services that thereby internalise and remove the externalities.

The creation of fewer, larger firms by merger activity is sometimes regarded as anti-competitive, since larger firms may prevent entry by smaller firms. The issues raised by strategic entry deterrence, together with its practicability, are considered in Section 4.6. This section has shown that, even if economies of scale from a merger are not significant, there are likely to be benefits in terms of network effects. It may also be the case that the conditions for entry by smaller firms are enhanced. Merger activity, together with the development of the hub and spoke system in the USA airline industry, caused larger carriers to focus on high density routes and enabled smaller carriers to enter the market and offer specialised services (such as intrastate commuter services) on lower density routes. The result was an improved match between the supply of and demand for aircraft (Fawcett and Farris, 1989).

4.5.5 Economies of Experience

Another possible source of declining average costs is that of economies of experience, which are closely related to the cost savings from "learning-by-doing" (Button, 1989). Economies of experience are time related, whilst the learning curve relates costs to cumulative production. The existence of economies of experience may create an entry barrier that reduces the contestability of the industry:

"Firms already in a market have gained experience in both the problems of serving that market and the specifics of the demand conditions confronting suppliers ... New entrants require similar experience but know that this can only be done at a cost which cannot be recovered if market entry results in a financial loss."

(Button, 1989, pp. 91-2)
New entrants will have higher costs than the incumbent in any given time period, because they lack the same level of experience and thus the ability to achieve cost reductions arising from this experience. One example of an experience economy might be a more efficient organisation of the firm. If the amount of experience necessary to achieve such economies is finite, one would expect the cost differential to diminish over time, as the entrant approaches the incumbent's level of experience. It could be argued that the costs associated with the acquisition of experience are sunk costs, since they will not be recoverable if the firm exits from the industry. In this case the sunk cost nature of the experience economies will reinforce the entry barrier created by the cost differential.

Experience, or lack of it, will also be a factor in the consumption, as well as the production process. Incumbent firms may have advantages over entrants because bus services are "experience goods", in the sense that their quality cannot be ascertained prior to consumption (Nelson, 1970). Incumbents may have built up a high quality reputation with consumers which entrants cannot easily duplicate, thereby creating another entry barrier (Button, 1989). Evidence for experience economies in the bus industry is limited, and is considered in the next section in conjunction with the evidence on sunk costs.

To summarise the findings of Sections 4.5.1 to 4.5.5, there is no strong evidence to support the view that there are significant economies of scale in the bus industry, particularly if density economies are taken into account. There is evidence to suggest, however, that economies of scope and density, particularly density economies, provide perceptible cost advantages within a service network. Debate about the cost characteristics of the industry, together with public policy towards it, has sometimes been confused by a failure to distinguish between these cost economies. Although merger activity in the deregulated bus industry has increased significantly, this is not necessarily an anti-competitive development. It may be the result of a desire to increase efficiency via network effects rather than to deter entry.
4.5.6 The Extent of Innocent Sunk Costs in the UK Bus Industry

Sections 4.5.1 to 4.5.5 have examined the innocent entry barriers created by cost differentials arising from various cost economies. Essentially, these cost economies arise through a more efficient utilisation of fixed factors. But the existence of fixed costs does not in itself imply the existence of entry barriers. The critical consideration, as Section 4.4.3 emphasises, is whether these fixed costs are also sunk costs. If they are, then they will form an exit barrier and reduce the contestability of the industry.

The discussion of sunk costs has so far been concerned with those arising from innocent entry barriers. These "innocent" sunk costs do not appear to be a significant factor in the UK bus industry. There is an established second hand market for vehicles, as well as the possibility of using vehicles in other markets such as schools contracts and private hire. The multiproduct nature of the bus industry increases the transferability of vehicles and thereby reduces the incidence of sunk costs. Vehicle leasing is also well established, and enables operators to match supply and demand without a commitment to vehicle ownership.

Although the provision of "fixed" assets such as garage facilities may appear to give rise to sunk costs, a study of bus operations in eastern England concluded that there was a direct linear relationship between garage size and costs (Glaister and Mulley, 1983). In effect, fixed costs are relatively small, so that the possibility of sunk costs is reduced. This finding is consistent with the results of the Kessides (1990) study of USA manufacturing industry mentioned earlier in Section 4.4.3, where the fixity of an asset does not imply the existence of sunk costs. Indeed, the study concluded that fixed assets such as buildings were likely to incur fewer sunk costs than those associated with machinery and equipment.

Economies of experience may constitute a sunk cost, although it is difficult to quantify this effect. The costs of building up a reputation or goodwill may also be sunk, depending on whether these "intangibles" can be sold. Presumably, it would be more difficult to sell
a reputation than a vehicle, particularly if, as is likely, exit from the industry is caused by a poor trading performance. In view of the relatively low levels of advertising within the industry\textsuperscript{6}, however, the role of factors such as reputation and goodwill are not likely to be particularly significant.

One plausible explanation for these low levels of advertising is that bus travel tends to be an "experience" good. Experience goods are those that are purchased frequently e.g. soap, so consumers can experience their qualities without the need for high levels of advertising. "Search" goods are those which are purchased infrequently e.g. cars, so that advertising is necessary in order to provide information about the qualities of the product (Nelson, 1974). The fact that in practice heavy advertising occurs for many experience goods is related to market structure, rather than the nature of the product: the market for soap, for example, is oligopolistic, so that non-price competition tends to occur in the form of advertising.

It may be argued that the bus industry since deregulation is oligopolistic, so that one might expect higher levels of advertising than those observed. In the case of the bus industry, however, non-price competition has generally taken the form of higher service frequencies rather than increased levels of advertising. The reason for this relates to the temporal dimension of the industry: passengers will generally take the first bus to arrive, irrespective of the firm that supplies it\textsuperscript{7}. Expenditure on advertising existing services may or may not increase demand, but expenditure on increased services almost certainly will.

\textsuperscript{6} In a letter (10.5.94) to the author, Professor John Hibbs commented on the "extraordinary reluctance of the majority of companies to engage in serious marketing management" and identified this as the "big unanswered question" associated with the industry.

\textsuperscript{7} Professor John Hibbs, in evidence to the House of Commons Transport Committee, 1993. (Session 1992-93) makes the contradictory comment that:

"I know of cases where people who preferred the buses of one firm to another on the same route found out the times they ran, and went out to catch them."

(Fourth Report. Volume III. p. 373)

One suspects, however, that such people are in a minority.
Sunk costs may be significant where large scale infrastructure investments must be made by firms. This is not the case for the bus service industry, however, where the road network is provided by central government, and where non-discriminatory access to bus terminals is provided under the provisions of the 1985 Transport Act. This also implies that "first-mover" advantages are unlikely to be significant. These refer to the situation in which significant sunk costs have been incurred by the incumbent, usually through infrastructure or other large scale investments, but not by the potential entrant (see Section 4.4.2).

Sunk costs may also occur where operators invest in improved service provision e.g. new vehicles, as part of a quality partnership (or contract) arrangement (see Section 4.4.3), but much depends on the type of arrangement between the local authority and the operator. If it is for a relatively short duration, then the operator can exit and use the new vehicles on other routes. There is also the empirical difficulty of establishing whether the operator has invested specifically for the partnership arrangement, or whether the investment would have occurred anyway.

To conclude, the empirical evidence tends to suggest that the sunk costs arising from innocent (or non-strategic entry barriers) are not likely to be significant for the bus service industry. If these were the only sunk costs to consider, then the industry would be highly contestable. But sunk costs may also arise through strategic entry barriers. For example, a firm may install excess capacity in order to deter entry. If entry does occur, then incumbent firms retaliate with increased service frequencies. A pre-commitment of this type will create sunk costs. This, and other types of strategic entry deterrence, are discussed in the next section.

### 4.6 Strategic Entry Barriers

The last section emphasised the importance of sunk costs, rather than fixed costs, in determining the contestability of a market. The sunk costs discussed in that section arose as a result of factors such as capital indivisibilities and the development of knowledge.
and reputation, rather than as a deliberate attempt to deter entry. But strategic behaviour by incumbent firms is another important determinant of market contestability. Strategic behaviour could include the deliberate creation of sunk costs by incumbent firms wishing to deter entry. This section discusses the different types of strategic entry barrier, and identifies those that appear most relevant to the UK local bus service market.

4.6.1 Excess Capacity

Spence (1977) develops a model in which entry is deterred by the installation of excess capacity that is only used if actual entry occurs. In the event of actual entry, the incumbent firm will use its excess capacity to expand output to the point at which the entrant cannot make a profit. This pre-commitment by the incumbent firm to expand output may or may not be a credible threat, depending on whether or not the entrant perceives the strategy to be profitable for the incumbent. A large number of models of entry deterrence via excess capacity have been developed, of which Spence (op. cit.), Dixit (1980) and Schmalensee (1981) are seminal. In general, the studies mentioned tend to support the view that excess capacity as an entry deterrent strategy is unlikely, since other strategies may be more effective. This conclusion, however, is sensitive to the model specification. Given uncertain demand, for example, coupled with the minimum output constraint assumed in the Schmalensee model, the effectiveness of the excess capacity strategy is increased in comparison to the results of the Spence model (Perrakis and Warskett, 1986).

In the context of local bus service deregulation, it is not easy to test the validity of this theory since excess capacity could arise as a result of strategic entry deterrence or from an increase in competition. One of the most notable cases of excess capacity in the post-deregulation period was in Greater Glasgow, where the number of peak period buses nearly doubled in the month after deregulation (see Table 4.1):
### Table 4.1: Greater Glasgow peak period buses (Union Street south of St. Vincent Street)

<table>
<thead>
<tr>
<th>Operator</th>
<th>Before deregulation</th>
<th>After deregulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTE/Strathclyde Buses</td>
<td>143.5</td>
<td>PTE/Strathclyde Buses</td>
</tr>
<tr>
<td>SBG and Independents</td>
<td>4.0</td>
<td>SBG and Independents</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>147.5</strong></td>
<td><strong>Total:</strong></td>
</tr>
</tbody>
</table>


In this case, however, it is clear that the excess capacity resulted from increased competition rather than entry deterrence, since the capacity provided by the (highly) dominant incumbent operator prior to deregulation actually fell in the post-deregulation period. Excess capacity was produced by entrants, not by incumbents.

Even if excess capacity can be attributed to the incumbent firm, it would still be necessary, in order to demonstrate the validity of the excess capacity theory, to show that the excess capacity was installed by the incumbent prior to entry and was a genuine increase in capacity, rather than a diversion of resources from one area to another e.g. from off-peak to peak services. Given the heterogeneity of services typically provided by bus (and coach) operators, it could be difficult for this second condition to be satisfied in relation to excess capacity on part (as opposed to the whole) of a network.

In spite of the theoretical problems, there is evidence to support the view that excess capacity, in the form of excessive bus numbers on given routes, has been used as a strategy to punish entrants. In the year after deregulation, service provision by a dominant incumbent in some of the metropolitan areas, particularly Greater Manchester, was very high:

"Where competition has emerged, incumbent operators have found the resources to retaliate mainly by "route swamping"."

(Oxford Transport Studies Unit, 1987a, p. 27)
The view that a dominant incumbent is able to effectively respond to entry by this method also appears to be widely shared by operators:

"A commonly held belief was that larger companies could more readily respond to entry by reallocation of buses to the newly competitive routes. Spread over a large operation, these reallocations had little effect on general service levels whilst permitting formidable opposition to the entrant on his chosen routes."

(Price Waterhouse, 1990, p. 34)

Further support for the route swamping view comes from the experience of local bus service operation in the Trial Areas after the 1980 Transport Act. Savage (1985, p.39) notes that swamping, in the form of increasing services on a route such that it becomes unprofitable, occurred in nearly a quarter of the cases in which an incumbent reacted to an entrant's timings.

A more recent example of entry deterrence via excess capacity occurred in April 1995 when Yellow Buses, a Bournemouth municipal operator, replaced approximately 12 Fleetline double-deckers with 12 Dennis Dart midibuses. The double-deckers were not sold, but retained in order to discourage potential rivals:

"We are mothballing them for possible use should any competitor appear."\(^8\)

This statement provides clear evidence of the genuine, pre-entry excess capacity referred to earlier. Nevertheless, one might ask whether such a strategy is sustainable or even desirable from the incumbent's point of view. On the conservative assumption that each Fleetliner would have an average resale value of £20,000, the company is foregoing a capital sum of about a quarter of a million pounds. If the opportunity cost of this sum is measured in terms of the interest obtainable from it on a risk-free 5 year UK government

---

\(^8\) Quotation from Mr. E. Reid (Managing Director of Yellow Buses) in Coach and Bus Week. No. 162. 8.4.95.
bond (about 8% in April 1995), then the cost of this strategy to Yellow Buses is in the region of £20,000 per annum. There are also the costs associated with the storage and care maintenance of the vehicles. Given a total fleet size of only 105, including the Fleetliners, the total costs arising from the excess capacity would represent a significant proportion of annual operating costs. There is the paradoxical possibility that in following an excess capacity strategy the company may make itself less competitive and thus more prone to rival entry than would otherwise be the case.

It may be argued that the strategy is more likely to be successful where the incumbent operator is large, or has access to funds from a large parent company. In this case the costs of a local campaign will be relatively small in relation to total costs, and the campaign can thus be pursued for a longer time period. One possible example is that of Fife Scottish (owned by Stagecoach Holdings plc), which dominates the provision of bus services in Fife, a large and relatively rural area. Registration of new services by other operators was immediately matched by Fife Scottish, which timetabled the duplicate services a few minutes in front of their rivals'. In every case the entrant soon withdrew, one of the main reasons being that entrants perceived the battle to be against Stagecoach, not just Fife Scottish. The area Traffic Commissioner did not feel able to prevent the practice of duplication, but described it as a "grey area" in which Fife Scottish had taken full advantage of the deregulated environment.

Stagecoach responded to the allegations of predatory practices by issuing a full page newspaper statement, part of which is reproduced below:

"Stagecoach was born in 1980 as a result of the new deregulated environment and was itself a small company for many years experiencing the severest of competition from large state and privately owned companies. We survived the experience because our new routes and ideas were popular with passengers and our operations were significantly more efficient than that of our competitors. Most of the competitors referred to in this [TV]

---

9 This practice, together with allegations of predatory pricing on the part of Fife Scottish, were the subject of an investigative television programme ("Frontline", BBC Scotland, 27.4.95).
programme were neither innovative nor particularly efficient but simply creamed off traffic from our busiest routes at the best times using life expired vehicles. Stagecoach provides a comprehensive network of quality bus services at all times of the day and throughout the weekend but such benefits can only be sustained if we retain our market share on our profitable routes.\textsuperscript{10}

This statement is interesting for a number of reasons. First, it illustrates the earlier discussion of the Austrian view that competition is a process in which small firms grow large as a result of their ability to satisfy new patterns of demand at low cost. Whilst the public interest view of Chapter three would regard the resulting size of the firm as a problem because of possible market failure, the Austrian view of Chapter four would interpret the firm's growth as a natural outcome of the competitive process. As such, the size of the firm acts as a measure of its success within this competitive process. Secondly, the threat of entry does not disappear with the emergence of the large firm: having grown, the firm then itself becomes subject to predation through new entry. In the case above new entry by smaller firms was unsuccessful, but there is nothing inevitable about this outcome as demonstrated by the initial evolution of Stagecoach. The problem of "cream-skimming" by smaller firms using low quality vehicles has also been documented by South Yorkshire Transport\textsuperscript{11}, which accounts for approximately two-thirds of bus passenger journeys in South Yorkshire. Thirdly, the statement implies that cross-subsidy is still practised in the deregulated industry (at least by Stagecoach). If the industry is perfectly contestable, then this practice would not be possible. If, as is much more likely, the industry is imperfectly contestable, then cross-subsidy is still possible.

This leads to the question of why a commercial company should provide loss-making services on a sustained basis (assuming that this is, \textit{de facto}, the case). The answer would appear to be related to the network system argument discussed in Section 4.5.4. Whilst some routes may be considered as unprofitable on a stand-alone basis, a company large

\textsuperscript{10} Quote from Brian Souter, Executive Chairman of Stagecoach Holdings plc. in Scotland on Sunday. 30.4.95.

enough to obtain benefits from scope and density economies may be able to provide such services as part of a network system. The advantage of a network system is that revenue and cost interdependencies between services can be exploited. Some passengers may not use a profitable 8.00 hrs service if an unprofitable 20.00 hrs service is not available. It may be the case that if the second service is not provided, then the first becomes unprofitable. This situation is a variant of the "feeder" economies arising in a network system. As the number of services increases, there will be more opportunities to take advantage of service interdependencies: the number of possible links between four services is likely to be more than twice the number associated with two services.

The above argument applies to a service network on a given day. There may also be interdependencies arising from the provision of different services over time. Although the revenue from a weekday service is unlikely to be affected by the provision of a Sunday service, for example, a Monday service may still be affected by the provision of other weekday services. If a Friday service is not provided, the passenger may have to buy a car to make his journey. This may then lead to a decision to travel by car on all weekdays. Cost interdependencies through time are likely to be greater than for revenue. The provision of a Sunday service will always reduce average fixed costs, although higher variable (probably wage) costs may also have to be taken into account in this particular instance.

Whether one treats the reduction in unit costs as arising from economies of scope (i.e. provision of services differentiated in terms of time across a given network) or from economies of density (increased provision of the same service across a given network irrespective of the time period) is a matter of choice. The end result is the same: interdependencies across the network will enable the provision of some services which, on a stand-alone basis, would not be provided. Small firms cannot provide large networks, and cannot take advantage of service interdependencies to the same extent as large firms. Consequently, there will be a tendency to fewer, larger firms, particularly where the region concerned has a mix of high and low density routes.
It may be objected that firm expansion has little to do with serving the needs of the market in the ways outlined above, but is a device, via excess capacity strategies, to restrict competition. Excess capacity strategies are expensive, however, and there is no guarantee that they will be successful. Large firms may be able to finance a campaign in a given area for a longer period, but they also have a larger territory to defend. There appears to be some evidence from operators that "gap-filling" has occurred in the deregulated market in order to deter entry, but since this expansion has been on profitable routes (Mackie, Preston and Nash, 1995, p. 233) there seems no reason to presume that the strategy of filling profitable gaps in the network is not a legitimate part of the competitive process. But "gap-filling" has also occurred on unprofitable routes, where additional services are provided but then withdrawn as competition or threatened competition recedes.

Nonetheless, the current situation relating to service duplication is unsatisfactory, since there are no guidelines with which to prevent unscrupulous operators exploiting the "grey areas" referred to above. In an attempt to provide a simple guideline for these cases, the author has developed a "half-headway" rule (see Appendix two).

4.6.2 Raising Rivals' Costs

The excess capacity strategy involves an increase in the incumbent firm's costs and a decrease in short-run profit for the firm. Long run profit may be higher, depending on the success of the strategy. An alternative entry deterrence strategy is to attempt to raise entry costs. For example, in an industry where common wage rates are centrally negotiated and enforced, a capital-intensive incumbent firm could force higher average costs on labour-intensive entrants by agreeing to pay higher wages prior to entry (Williamson, 1968). One method by which a bus operator could raise rivals' costs is to provide integrated ticketing arrangements such as travelcards, which usually entitle the holder to unlimited transfer between buses, or even between bus and rail, for a fixed initial payment. Such schemes will be profitable if the revenue gained from generated patronage exceeds the
lost revenue from the existing patronage. They also have the effect of increasing customer loyalty.

Prior to deregulation, the West Midlands PTE operated a Travelcard scheme which accounted for some 40% of passenger journeys. Similar schemes existed in Tyne and Wear and South Yorkshire. The problem for a new entrant is that it would not be likely to have a sufficiently large service network to make its own "travelcard" relatively attractive, but without such a card it may be difficult to persuade passengers away from the incumbent (Oxford TSU, 1987b, p.29). In general, the smaller the service network, the higher the cost per passenger of providing travelcards. Administration and advertising costs, which tend to be fixed, will be spread over a smaller number of passengers.

Advertising may be used as an entry deterrence strategy (Cubbin, 1981, Schmalensee, 1982). If advertising by the incumbent has created a stock of goodwill or reputation which affects market demand, then an entrant will not be able to gain market share unless he has a higher level of advertising per unit of output than the incumbent. *Ceteris paribus*, the entrant's average costs will be higher than those of the incumbent. This argument does not seem to be significant in bus operations, where advertising, if it occurs at all, tends to be on a small-scale local rather than a national level. Although some operators have attempted to build up a brand image, there has traditionally been a limited use of advertising within the industry, both in the pre and post-deregulation periods (see Section 4.5.5). In evidence to the House of Commons Transport Committee (Session 1992-93), Professor Hibbs remarked in relation to the provision of customer information that:

"The industry's performance in this vital activity was appallingly bad before 1985 ... It has since become markedly worse ..."

(House of Commons Transport Committee, Fourth Report, Volume III, p. 369)
Where an incumbent has control over terminal facilities, an entrant may experience difficulty in gaining access to those facilities. Incumbent airlines can make it difficult to gain access to airport resources such as landing and take-off slots, boarding gates, check-in areas and lounge facilities. Access problems occurred after the 1980 Transport Act deregulated express services, when National Express refused access to some competitors (see Section 2.5.1.2). Even where access is allowed, terminal charges may be higher for the entrant, or the entrant may be assigned to parts of the terminal which are difficult to reach and manoeuvre in. If access is not allowed then, assuming the incumbent has already located in the least-cost site, the entrant will be forced to pay higher terminal costs if he decides to buy or rent alternative facilities.

Although the Buses White Paper states that "major bus stations should be operated .... on a commercial basis under arrangements which will provide for all operators to have equal opportunity of gaining access to them" \(\text{(op. cit., Ch. 5, para. 5.16)}\), there was evidence to suggest that, in the short term, this was not the case in certain areas\(^{12}\). Nevertheless, the number of complaints regarding bus station access is relatively small, and does not seem to have significantly increased in the longer term (see Table 4.2).

The development of quality partnerships (or contracts) in the deregulated market (see Section 4.4.3) could also provide an opportunity for strategic behaviour by large operators. They could put a case to the local authority for service standards that are too high for smaller operators to afford, with the (covert) intention of creating entry barriers to the market. Under the provisions of the Transport Bill (Dec. 1\(^{st}\), 1999), however, potential entrants will be able to object to the provisions of quality partnerships (and contracts). There is also the problem that raising entry costs in this manner will be expensive for the large operator, and this expense will increase with the duration of the partnership. The Transport Bill, for example, proposes a minimum partnership period of five years.

\(^{12}\) Useful references can be found in trade journals such as Buses, Bus Business, Bus Fayre and Coachmart.
4.6.3 Predatory Pricing

Predatory pricing can be defined as the use of price reductions by the dominant incumbent operator (or group of operators) so that an entrant cannot make a profit i.e. the post-entry price will be below the entrant's average cost. The incumbent sacrifices short-run profit in order to force the entrant out of the market: after exit, the incumbent raises price and earns a long run profit greater than that which would have been earned if the entrant had stayed in the market. Some definitions of predatory pricing emphasise that losses must be incurred by the incumbent firm. But the definition could also apply to any price reduction that reduces current profit in order to deter entry. On this criterion, even a post-entry profitable incumbent firm could be guilty of predatory pricing (Dodgson, Katsoulacos and Newton, 1991). As with all entry deterring strategies, it is the threat, rather than its execution, which is critical. Even if predatory pricing does not actually occur, the threat to use it may be sufficient to deter entry. If predatory pricing does occur in response to entry, it may serve a two-fold purpose: firstly, to force the exit of the current entrant(s); secondly, to build up the incumbent's credibility as a firm which is willing to carry out its threat, and thus deter future entry.

Models of predatory pricing can be classified into two broad groups: game-theoretic and non-game theoretic. The latter type of models (e.g. McGee, 1958,1980) tend to support the view that predatory pricing is unlikely to succeed, and thus unlikely to occur. These conclusions, however, depend on the assumptions of complete information and perfect capital markets. In addition, the essentially static nature of these models implies that an important element of strategic behaviour, namely the cumulative creation of a reputation for carrying out threats, is not considered.

Game-theoretic models have been widely used in the analysis of predatory pricing (e.g. Selten, 1978, Kreps and Wilson, 1982, Milgrom and Roberts, 1982a, 1982b, Benoit, 1984). In general, these models also show that the likelihood of predatory pricing depends on the nature of the assumptions made about information and capital markets. Given the existence of imperfect information, the likelihood of predatory behaviour
increases since incumbents (even weak ones) may have an incentive to build up a reputation for such behaviour (Krepps and Wilson, *op. cit.*).

An important feature of some game-theoretic models played over a finite number of time periods is the chain store paradox (Selten, *op. cit.*). In an n-period game, the incumbent will have no incentive to cut price in the nth period since there is no further period in which to raise price and offset the loss in profit incurred in period n. Entry in period n will not be opposed by the incumbent, and entry will occur as long as the post-entry price remains above the entrant's average cost. But if entry is never opposed in the final period, entry will also occur in period n-1, since the entrant knows that retaliation from the incumbent will not occur in period n. By a process of backward induction, entry will always be unopposed by the incumbent i.e. predatory pricing will not occur. But the logical unassailability of this result is at variance with actual business behaviour, where significant post-entry price reductions are observed\(^{13}\); hence the "paradox". Once again, this result hinges on the assumption of complete information: if incomplete information is introduced, the chain-store paradox is resolved (Kreps and Wilson, 1982, Milgrom and Roberts, 1982b).

A variant of the chain store paradox is where entrant and incumbent have complete information, but the entrant is subject to a financial constraint (Benoit, *op. cit.*). In this situation, the incumbent is always able to deter entry, so that a "reverse" result for the chain store paradox is obtained. If the assumption of complete information is relaxed, however, then the incumbent may not be able to deter entry even if the entrant is financially constrained. The financial constraint may involve the entrant paying a higher risk premium for borrowed funds than the incumbent. This is likely to occur when the incumbent is a long-established firm with a reputation for solvency, but the entrant has no such financial pedigree (Williamson, 1974).

---

\(^{13}\) The airline industry provides a well-known example: the collapse of Laker Airways in 1982, following alleged predatory pricing from other transatlantic carriers including British Airways, and the prompt increase in fares following Laker's exit.
The overall conclusion from these models is that the greater the degree of incomplete information, the more likely predatory pricing is to occur. It seems reasonable to assume that information in the initial stages of UK local bus service deregulation was particularly incomplete, since operators were in transition from a highly regulated system that had lasted for more than fifty years to a new environment in which they, not the Traffic Commissioners, could set fares and service levels. Entrants would not have prior experience of how incumbents would respond to entry, and this lack of information gives incumbents an opportunity to establish a reputation for predatory pricing. This strategy may work even where the incumbent is weak ((Kreps and Wilson, op. cit.). With the passage of time, however, one would expect that operators gain a greater knowledge of the deregulated market, particularly the likelihood and credibility of rivals' responses. Given an increase in such knowledge, the incidence of predatory pricing would reduce.

This conclusion is supported by a number of examples from the deregulated bus service market. In the Trial Areas, which were the first areas to experience local bus service deregulation, about 80 per cent of incumbents either matched or reduced their fares in response to entrants' fare reductions. In the Hereford trial area the National Bus Company ran a free bus service in response to competition. The National Bus Company also instructed its subsidiaries to adopt a policy of "fare matching". There were even suggestions that money from the rates was used to reduce fares on Cardiff City Transport (Savage, 1985, p.38). After the initial price wars, however, predatory pricing became the exception rather than the rule. This pattern was also evident in the post-1986 period, where price competition in any case has tended to be limited (Evans, 1990).

This is not to argue that predatory pricing has become non-existent in the deregulated industry. Mackie and Preston (1993) give examples of its occurrence, and also details of the OFT's approach to deciding whether pricing is predatory or not. The approach is firstly to identify whether predation is feasible: evidence for this could come from the incumbent's market share. Secondly, the relationship between price and costs is examined: if the incumbent prices below short run marginal or short run variable cost, this is considered to be clear evidence of predatory pricing. Thirdly, the motives and
intentions of the incumbent are scrutinised: in the case of Highland Scottish, for example, its Business Plan indicated a willingness to sustain losses and thus a possible intention to engage in predatory pricing. It should also be noted that since deregulation there have been a number of free bus services operated. Whatever the difficulties involved in defining and identifying predatory pricing, the provision of free services in response to entry is a pricing strategy that is clearly predatory.

Table 4.2 below shows a breakdown of complaints concerning anti-competitive behaviour made to the OFT. The official recording of such a complaint does not in itself imply that the complaint is investigated. Of the 541 complaints made to the OFT from 1987 to end 1994, only nine complaints were investigated under section 3 of the Competition Act - eight of these alleged predatory behaviour, and one concerned bus station access. Six of the investigations found evidence of anti-competitive practices, and two of these were referred to the MMC. During this period, two other references were made by the OFT to the MMC under the monopoly provisions of the Fair Trading Act (House of Commons Transport Committee First Report, 1995, Volume II, p. 98, and also Annex 1, pp. 102-3 of same).

<table>
<thead>
<tr>
<th>Year</th>
<th>Bus Stations (1)</th>
<th>Subsidised Services (2)</th>
<th>Predatory Pricing</th>
<th>Other (3)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>4</td>
<td>4</td>
<td>16</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>1987</td>
<td>9</td>
<td>7</td>
<td>23</td>
<td>18</td>
<td>57</td>
</tr>
<tr>
<td>1988</td>
<td>11</td>
<td>13</td>
<td>24</td>
<td>10</td>
<td>58</td>
</tr>
<tr>
<td>1989</td>
<td>7</td>
<td>15</td>
<td>19</td>
<td>12</td>
<td>53</td>
</tr>
<tr>
<td>1990</td>
<td>4</td>
<td>6</td>
<td>23</td>
<td>18</td>
<td>51</td>
</tr>
<tr>
<td>1991</td>
<td>7</td>
<td>4</td>
<td>39</td>
<td>18</td>
<td>68</td>
</tr>
<tr>
<td>1992</td>
<td>6</td>
<td>16</td>
<td>52</td>
<td>16</td>
<td>90</td>
</tr>
<tr>
<td>1993</td>
<td>10</td>
<td>7</td>
<td>51</td>
<td>40</td>
<td>108</td>
</tr>
<tr>
<td>1994</td>
<td>2</td>
<td>1</td>
<td>36</td>
<td>17</td>
<td>56</td>
</tr>
<tr>
<td>Total:</td>
<td>60</td>
<td>73</td>
<td>283</td>
<td>153</td>
<td>569</td>
</tr>
</tbody>
</table>

Source: 1986 data provided to the author by the OFT. 12.5.95. 1987-94 data from House of Commons Transport Committee First Report. 1995. Volume II. Table 1, p. 98.
Notes: 1. Complaints about bus station access and pricing.
2. Complaints concerning re-registration of commercial services by the incumbent on a route previously considered unprofitable, but where an entrant had won a tendered contract to operate subsidised services.
3. Other complaints including the running of duplicate services.

The largest number of complaints concerns predatory pricing. Although price competition has not been widespread in the deregulated industry, the figures show that the number of predatory pricing complaints has generally risen since 1986, and that it forms on average almost half the total complaints. But the figures themselves may understate the problem, since the threat of predatory pricing may also create a significant entry barrier. A possible solution to the problem of predatory pricing would be to construct a simple pricing rule that would be made explicit to operators and would unambiguously rule out such extreme cases as zero fares. Although the OFT uses information about the price-cost margin in cases of alleged price predation, there is no standard guideline made available to operators which would determine the minimum acceptable level of fares. If such information were made available, the OFT could undertake a more proactive, rather than its currently reactive role. The author has developed a pricing rule that might assist the OFT in this role (see Appendix two). A discussion of the roles which the OFT (and the Monopolies and Mergers Commission) could play in the deregulated system is also contained in Appendix two.

4.7 Summary and Conclusions

This chapter has developed alternatives to the market failure framework of Chapter three. The idea that market failure should be corrected by regulation is, from these alternative viewpoints, overly simplistic for two reasons. First, the public choice school emphasises that regulatory failure may occur, so that the cure may be worse than the disease: second, the Austrian School argues that markets per se do not fail, but that failure is essentially the result of entry barriers to the market and an inadequate property rights system.

Austrian School economists view (temporary) monopoly as a necessary reward for risk-

14 The author telephoned the OFT 28.02.00 to try to obtain complaints data for 1995 onwards. This data is not yet processed in the form of Table 4.2, and so is not readily available. It appears, however, that the
taking in a dynamic competitive process. Similarly, mergers are not necessarily anti-competitive, but may be a means of competing via increased cost efficiency. The crucial question is not market structure, but whether the market is contestable. In a perfectly contestable market, even a natural monopoly must set price and output at levels associated with perfect competition.

The re-emergence of local monopoly, together with increased market concentration via mergers in the deregulated bus industry, is thus not incompatible with a competitive market outcome. The formation of fewer but larger firms may also encourage the provision of more services via network effects. If the bus market is perfectly contestable, regulation is unnecessary.

In terms of actual entry to the deregulated market, the total number of licensed operators increased continuously between 1989 and 1994, from 6,415 in 1989/90 to 7,193 in 1993/94, an increase of 12%. This increase, in conjunction with the analysis of this chapter, is evidence that operators can enter the market with relative ease:

"UK local bus operation is a very easy business to enter with very low legislative, capital or market obstacles to entry ... It is far easier to become a bus operator than, for instance, a street corner grocer in most instances."


However, ease of market entry is one thing; coping with post-entry predatory behaviour by the incumbent is another. The Buses White Paper claims that the local bus service market is "highly contestable". In the case of non-strategic entry barriers the evidence provides strong support for this view. Sunk costs are low, and the multi-output characteristics of the industry increase the opportunities for cross-entry and thereby the overall number of complaints has fallen from 1995 onwards. Since complaints about predatory pricing tend to move in line with total complaints in Table 4.2, it is likely that they also have tended to fall.
The degree of contestability of specific markets. For strategic entry barriers the evidence is less clear: predatory behaviour has been a problem not only where it has actually occurred, but also in terms of the threat to use it. Since threatened use is virtually impossible to quantify, it is difficult to gauge the extent of the problem. Such problems are not uncommon in other markets, however, and could in principle be reduced by appropriate public policy, such as intervention by the OFT or some similar body.

The effectiveness of OFT intervention in cases of alleged predatory behaviour could be increased by the following measures: first, specification of clear guidelines as to what constitutes predatory behaviour (see Sections 4.6.1 and 4.6.3); second, the introduction of financial penalties if the predatory pricing guideline is infringed; third, a more proactive approach, such as interim powers to order a return to the original fare in alleged predatory pricing cases that are under investigation. These measures would, in all likelihood, not have to be used very often. Operators would know the ground-rules and the penalties of infringing them.

The development of quality partnerships and contracts may act to reduce contestability, particularly for small operators. But much will depend on the exact type of partnership or contract arrangement, and at the time of writing there are a number of proposals in the Transport Bill (Dec. 1st, 1999) that may or may not pass into law. Again, however, there seems no reason why the OFT cannot be given appropriate powers to intervene in cases of alleged anti-competitive practices.

Given these changes, the influence of strategic entry barriers would be reduced and contestability increased accordingly. The industry would not be perfectly contestable, however, so entry barriers such as economies of scale could produce situations where some degree of monopoly power is exploited. Mergers could also create monopoly power. Given a situation of imperfect contestability, market structure becomes a factor that should be taken into account when determining whether a monopoly or merger is against the public interest. There is thus a continuing role for the MMC in the deregulated bus industry.
Overall, the evidence does not support the claim of the Buses White Paper that the local bus service market is highly contestable, but neither can the market be regarded as having low contestability. The weight of evidence tends to tip the balance somewhat towards high rather than low, so perhaps the fairest conclusion would be to regard it as "relatively" contestable. In conjunction with the public choice emphasis on the costs of regulation, the finding of relative contestability suggests that the introduction of competition through deregulation is an appropriate policy for local bus services in Britain. But there would still be a need for residual regulation, operating mainly through the Traffic Commissioners, the OFT and the MMC.

Before any firm conclusions can be reached on the deregulation issue, however, a number of other factors need to be considered. First, there is the actual experience of the deregulated local bus service market in Britain. Second, there is the empirical issue of how subsidy reduction affected the deregulated market. These factors are discussed in Chapters five and six respectively. Third, there is the issue of whether a system of competitive tendering for both profitable and unprofitable services might be preferable to the deregulated system. Competitive tendering is discussed in Chapter five (Sections 5.1.2 and 5.3.7).
CHAPTER 5

Deregulation of Local Bus Services

The 1985 Transport Act marked the demise of the highly regulated system introduced in the 1930 Road Traffic Act. Following the removal of quantity licensing in the Trial Areas and the long distance coach sector in 1980, the 1985 Act extended deregulation to the local bus service sector in Britain (excluding London). The legislation was designed to:

"... make major changes to arrangements for the bus industry, so that it is set free to give a better service to the passenger at less cost to the ratepayer and taxpayer."

(Buses White Paper, 1984, p. 3)

Chapters three and four discussed in some detail the general arguments for and against regulation, together with a number of specific examples from the local bus service industry. This chapter focuses these arguments on the deregulation of the industry, and deals with a number of issues within an analytical framework that distinguishes between the short, medium and longer term effects of deregulation. It is important to bear in mind that the deregulation of local bus services was accompanied not only by their privatisation, but also by a considerable reduction in subsidies to the industry. Section 5.1 presents the background to the deregulation of local bus services, and provides an evaluation of the main arguments discussed in Chapters three and four in the context of local bus service deregulation. Section 5.2 develops a chronological framework within which to analyse qualitatively the dynamics of the deregulation process. Section 5.3 discusses the short to medium term effects of the 1985 Transport Act, whilst the longer term effects are discussed in Section 5.4. Section 5.5 provides a summary and conclusions.
5.1 Local Bus Service Deregulation in Britain: the background

5.1.1 The Case for Deregulation

The case for deregulation of the industry is argued in a number of publications, notably the Buses White Paper (1984), Beesley and Glaister (1985) and Hibbs (1985). The main arguments concern the static and dynamic efficiency losses associated with the regulated system. This system resulted in the formation of protected local monopolies that created allocative and technological inefficiencies (see Chapter three). It also resulted in dynamic inefficiencies because of the lack of innovation. In addition, there were the costs associated with regulatory failure (see Chapter four).

The difficulties the industry was experiencing by the 1970s are illustrated by the following figures. Between 1972 and 1982 there was a 30% real increase in fares compared with a real increase of only 3% in the cost of running a car. Over the same period, subsidies to the industry (i.e. revenue support and concessionary fare schemes) increased in real terms by 1,112%. Even allowing for increases in unit costs caused by decreases in passenger density (see Section 4.5.2), the increase in subsidy is very large.

One plausible explanation is that lack of competition created technological or "X-inefficiency", so operators needed increased subsidies to cover the cost increases. Only 8% of pre-deregulation local bus service mileage in Britain was run by private operators, but the Buses White Paper (1984) argues that their costs were significantly lower than those of public sector operators. It cites the results of a Transport and Road Research Laboratory (TRRL) study based on two traffic areas, which concluded that private operators' costs were about 30% to 40% below those of public sector operators providing a similar type of service (Tunbridge and Jackson, 1980).

Further evidence is provided from a study of private urban bus operators in Australia where, after allowance is made for differences in the types of services, their average costs per vehicle mile were generally between half and two-thirds those of their public operator
counterparts (Wallis, 1980). These results are similar to those of Hensher (1988), who finds that unit driver costs for public sector operators in seven major Australian cities were over twice as much as those for private sector operators in Sydney. Operating costs per bus km. were 66% higher.

Within the British public operator sector there was considerable variation in operating costs: in 1979/80 cost per vehicle mile varied from 84p to £1.88 around an average of £1.22 (Higginson and White, 1982). These figures do not allow for special circumstances, so that it is unclear whether or not this range of unit costs indicates the existence of significant "X-inefficiency" on the part of (at least) some public sector operators. Using figures from four operators investigated by the Monopolies and Mergers Commission (1982), another study of British operators concluded that, after adjustments are made for differences in average operating speed and "peakiness", the differences between operators' costs per vehicle mile almost disappear (Mackie and Nash, 1982).

Cost reductions from competition were an especially important point for supporters of deregulation, since the cost reductions were needed to offset the fall in external and internal subsidy associated with deregulation:

"Government policy dictates that subsidy to the bus industry must fall substantially. If nothing else were to change then this would have serious consequences for the public. Fortunately, there is scope for substantial cost reduction. But this can only be achieved through competition ... deregulation will produce an instant response."

(Beesley and Glaister, 1985)

Critics of regulation also argued that lack of competition had made the industry unresponsive to changing patterns of consumer demand, especially in the provision of increased bus frequencies, different fare structures and different types of service such as taxis, minibuses and midibuses. This dynamic inefficiency restricts consumer choice and represents another source of welfare loss. A competitive market would stimulate
innovation and increase consumer choice (Hibbs, 1975, Walters 1982, Roth and Wynne, 1982).

Finally, the regulatory system may be subject to "regulatory capture", insofar as the regulators become excessively influenced by the views of those whom they are supposedly regulating, or by sectional interests within the regulated group. Additionally, the costs expended by producers in achieving and maintaining the regulated system may represent a considerable loss of welfare to society (see Section 4.1). Mulley (1983) argues that the provisions of the 1930 Road Traffic Act were excessively influenced by incumbent operators. A more recent example is where, shortly after the licensing changes in the 1980 Act, the Traffic Commissioners tended to uphold the objections of existing operators to new operators, even though "the supporting evidence was often meagre." (Buses White Paper, 1984, p. 73). This bias was subsequently reduced when the Traffic Commissioners began to take a wider interpretation of the term "public interest" (see Section 2.5.1.4).

5.1.2 The Case against Deregulation

The government did not issue a Green Paper prior to the publication of the Buses White Paper of 1984. Consequently, the White Paper was published without the benefit of a formal consultation process. The proposals in the White Paper generated considerable debate, and the proposals in the subsequent Transport Bill were given a generally critical response by the House of Commons Transport Committee (Bell and Cloke, 1990, p. 44). Although there was a broad agreement that change was necessary in the bus industry, and that the aims of the White Paper were appropriate to the type of change required, there was considerable disagreement as to whether the means contained in the White Paper would produce the desired end. The case against deregulation is succinctly articulated in the following comments:

"The analysis of the White Paper rests on four straightforward propositions: deregulation will produce a competitive market, competitive pressures will reduce costs, competition
The rejection of propositions one and two is based on the view that the UK local bus service market is not contestable. The evidence discussed in Chapter four, however, suggests that the market is "relatively contestable". The presence of competition, both potential and actual, is manifested most clearly by the large fall in bus operating costs in the two-year period after deregulation. Heseltine and Silcock (1990) estimate that there were average cost reductions per bus km of 30 per cent between 1985/86 and 1987/88 outside London. Mackie, Preston and Nash (1995, p. 238) concede that "From the Government's point of view, this is a spectacularly successful result, ...", although some part of the fall in costs may be due to factors not associated with competition. This decline in costs continued over the longer term: between 1985/86 and 1993/94 operating costs per bus kilometre (including depreciation) fell by 41 per cent in real terms outside London (Bus and Coach Statistics, 1996/97).

It is generally agreed that actual competition has not been widespread in the post-deregulation period, and that it has been less than the supporters of deregulation expected (Glaister, 1993). But the absence of widespread actual competition should not be used to infer that the deregulated market is not competitive. In conjunction with the large fall in unit operating costs, it strengthens the case for assuming that the UK local bus service market is contestable. If actual competition has not been widespread, then potential competition must be a significant force in exerting downward pressure on operating costs.

The rejection of proposition three is based on the view that market failure is significant in the UK local bus service industry. The conclusions of Chapter three are that the likeliest sources of market failure in the industry occur through wasteful competition and externalities, primarily unfair competition. Unfair competition arises because of the
failure to price car use at its marginal social cost. This situation results from government policy towards car users, and has existed in both the pre and post-deregulation periods. As such, it is not relevant to the arguments concerning the merits or demerits of the deregulation of bus services.

Wasteful competition may arise from the natural monopoly characteristics of the industry and the provision of sub-optimal fare and frequency combinations in an unregulated market (see Section 3.3.3). The problem with the wasteful competition argument is that it is based on a static view of competition. If the dynamics of the competitive process are taken into account, as in the approach of the Austrian school, the concept of wasteful competition becomes increasingly problematic. Competition that is considered wasteful from a static perspective may be interpreted as beneficial in a dynamic context (see Section 4.2). There is the further problem that a regulatory solution to wasteful competition may involve as much, if not more failure, than the market solution (see Section 4.1).

In terms of adverse side effects, it was argued that deregulation might result in lower safety standards, increased congestion, and loss of services (Gwilliam, Nash and Mackie, op. cit.). Safety standards have not declined in the post-deregulation period, however, and congestion has not generally caused serious problems (Mackie, Preston and Nash, op. cit.). It is more difficult to evaluate the loss of service argument, since data for vehicle kms will incorporate some element of schedule matching. Where schedule matching occurs there is likely to be a reduction in load factors, so that some buses may run with few, if any, passengers. In this case, it can be argued that the increase in vehicle kilometres does not represent a corresponding increase in service levels.

There is also the problem that increases in vehicle kilometres may arise through the use of smaller buses scheduled at increased frequencies. There has been a considerable degree of service innovation in the post-deregulation period, particularly with respect to the use of mini and midibuses (Mackie, Preston and Nash, op. cit.). It may be the case
that in terms of vehicle seating capacity, service quality actually fell. This issue is discussed in detail in Section 5.3.2.

Bus kilometres in Britain (outside London) increased by 24 per cent between 1985/86 and 1993/94 (Department of Transport, 1994). It is most unlikely that an increase of this size could be accounted for solely in terms of schedule matching. The increase has occurred mainly in the commercial service sector, with the subsidised sector remaining broadly constant. In terms of vehicle seating capacity, there was a slight increase between 1985/86 and 1988/89 for local bus services, with the rise in the non-metropolitan areas outweighing the decrease in the metropolitan areas and Scotland (see Section 5.3.2). These aggregate figures obviously conceal spatial and temporal variations in service quality. Where service cuts took place, it is important to bear in mind that they might also have occurred in the regulated system as part of a desire to reduce the rapidly increasing network subsidy. Overall, it is not unreasonable to infer that loss of services, whether measured in terms of vehicle kilometres or vehicle seating capacity, has not in general been a major problem in the post-deregulation period.

Gwilliam, Nash and Mackie (1985a) argue that deregulation policy should have encouraged competition "for the market", rather than competition "in the market" (or "on-the-road" competition). Their recommendation is to introduce competitive tendering for unprofitable and profitable services. Local authorities would invite operators to tender for the franchise to run a service. The highest bidder in the case of a profitable service, or the bidder requiring the lowest subsidy in the case of an unprofitable service, would be awarded the sole right to operate the service for a fixed time period, and the contract could specify the minimum service and the maximum fare. At the end of the allotted time period the service would be put out to tender again. In this way competition would be encouraged without the possibilities sometimes associated with "on-the-road" competition, such as schedule matching and "cream skimming".

Beesley and Glaister (1985) argue that competitive tendering for profitable services is undesirable for a number of reasons: first, collusion in the tendering process is likely to
occur between the larger incumbent operators; second, local authorities will in practice tend to implement competitive tendering in a gradual way over small parts of the network, so that the potential for collusion will increase: third, the tender contract will in practice only allow bidders to compete on price and possibly one other quality variable, so that competitive tendering will prevent bidders from competing in other ways e.g. season discounts, bus size, quality of vehicle; fourth, the difficulties of recontracting may enable the franchisee to "capture" the franchisor, perhaps by providing a lower service quality than originally specified. Large firms could also reduce competition for the market by block-bidding for route networks (Glaister and Beesley, 1991).

In a study of the competitive tendering system for London bus services, Glaister and Beesley (op. cit.) conclude that, for the period 1985 to 1988, public sector operators had systematically higher costs than private sector operators, indicating that the system was not in equilibrium. In a similar study, however, Kennedy (1995) found no cost differences between public and private sector operators for the period 1985 to 1993, indicating that the tendering system had moved to a competitive equilibrium. The study also found cost savings from the competitive tendering process of 20 per cent, a result consistent with that of Domberger et al. (1986) for refuse collection. This does not, however, preclude the possibility that cost savings would have been greater if on-the-road competition had been allowed for London's profitable routes. Domberger et al. have also been criticised on the grounds that their cost equation is mis-specified, that they have measured the gains from privatisation without including the losses, and that service quality declined following the tendering process (Ganley and Grahl, 1988).

The arguments for competitive tendering in the commercial network, like those for regulation generally, are based on the static efficiency losses resulting from market failure and the belief that regulatory failure, particularly regulatory capture, is not significant. The discussion of Chapter four emphasises that this traditional or public interest approach is not sufficient to evaluate the benefits and costs of the regulation-deregulation debate. By omitting the public choice and Austrian school approaches, it biases the outcome of the debate towards regulation. When these other approaches are
taken into account, the arguments in favour of deregulation and "on-the-road" competition are stronger.

More recently, another argument against deregulation has been emphasised. This is the view that deregulation creates a loss of public control of bus services, so that local authorities cannot use public transport policy as an instrument to achieve social and environmental aims (Mackie, Preston and Nash, op. cit., Royal Commission on Environmental Pollution (RCEP), 1994). The view of the RCEP depends on the argument that local authorities can integrate fares and services in an optimal way. The Transport Act 1985, having removed the powers of local authorities to provide and integrate services, means that:

"It is not therefore possible for them [local authorities] to ensure the most cost-effective implementation of the measures which they adopt in pursuit of land-use and transportation strategies."

(Royal Commission on Environmental Pollution, 1994)

The arguments concerning integration have been discussed in Sections 3.3.2 and 4.5.4. Their overall conclusion is that integration can be provided by market forces, although the degree of integration is likely to be less than that in a regulated system. There is also no particular reason why a local authority can provide transport services more cost-effectively than the market, particularly if the costs of regulatory failure are taken into account. As far as land use is concerned, the biggest single problem associated with transport is the underpricing of car use (see Section 3.3.3.5). The first best solution to this problem is to charge car users more, rather than allow local authorities to subsidise public transport. Indeed, the use of public transport subsidy to achieve land-use aims appears to be a questionable practice:
"... it would appear that in general transport subsidy policy is a possible instrument of long-term land-use strategy, but is likely to be a weak and relatively expensive means of achieving any objective."

(Gwilliam, 1987, p.13)

To conclude, the theoretical case against deregulation, as argued by Gwilliam, Nash and Mackie (1985) is based firmly on the public interest model. This model is useful for identifying static welfare losses, but does not take into account the dynamic benefits of competition, the costs of regulatory failure or the importance of potential competition. As such, it provides an incomplete framework for the evaluation of deregulation. The large fall in unit operating costs provides firm evidence for the emergence of a competitive industry, in contrast to the predictions of Gwilliam, Nash and Mackie (op. cit.).

5.1.3 Privatisation

Privatisation formed a cornerstone of the supply-side policies of the 1979 Conservative government. In the words of Mr. Nicholas Ridley (then Financial Secretary to the Treasury):

"It must be right to press ahead with the transfer of ownership from state to private enterprise of as many public sector businesses as possible ... The introduction of competition must be linked to a transfer of ownership to private citizens and away from the State. Real public ownership - that is ownership by people - must be and is our ultimate goal."

(Treasury Economic Progress Report, May, 1982)

Privatisation is based on the belief that ownership affects economic performance. It is a wider term than denationalisation and can take a number of forms, ranging from the sale of council houses, the introduction of competitive tendering, and the sale of shares (either wholly or partly) in public corporations. Proponents, using the experiences of the post-
war UK nationalised industries, argue that state-owned enterprises were producer oriented and unresponsive to the needs of consumers. In particular, they were a bastion for restrictive Trades Union practices. They were also "X" inefficient because of the lack of capital market discipline, since (Labour) governments tended to write off losses rather than suffer the political unpopularity associated with rationalising "lame duck" industries (Polanyi and Polanyi, 1974). It was argued that the transfer of assets from the public to the private sector would increase consumer sovereignty and allocative efficiency, decrease "X" inefficiency, and also provide management and employees with an opportunity to buy shares in their companies. The alternative view, that nationalised industries performed better than the private (manufacturing) sector, is supported by Pryke (1971), although this conclusion was subsequently modified (Pryke, 1981).

The National Bus Company was privatised in 1986. It was not privatised en bloc, but split up into free-standing subsidiaries which were gradually sold off to the private sector. This fragmentation began with the management buy-out of Devon General in August 1986 and was completed in April 1988 with the sale of London Country Buses (North East) to Parkdale. The decision not to privatise en bloc was because:

"The Company's size and financial strength might deter other operators from competing with it particularly in view of the strong position from which it will start."

(Buses White Paper, p.17-18)

This view is supported by a number of considerations, both general and specific. British Telecom, which enjoyed a monopoly of the UK telecommunications market, was privatised en bloc. Ten years after privatisation, British Telecom still retained at least 85 per cent of the overall market. Its success was due in part to price-cutting, innovation and improved service, and also to "sheer muscle". In the specific case of the bus market, it is supported by the experience of the Trial Areas, where NBC used its size to deter entry:

---

1 The Times, 11.1.95, p.29.
"NBC had a policy of fighting new entrants in the Courts together with a policy of subsequently reducing fares to match competitors' prices and refusing access to NBC-controlled bus stations as a means of running them off the road."

(Mulley and Wright, 1986, p. 18)

It should be noted that, even after the privatisation of individual NBC subsidiaries, the average size of a subsidiary's fleet was about 500 buses. This is in sharp contrast to the private operator sector prior to deregulation, where 93 per cent of the 5,848 private operators owned less than fifteen buses (Transport Select Committee, 1985, p.172). The size of ex-NBC subsidiaries relative to private operators may be a reason why actual competition after deregulation has been less than anticipated. On the other hand, the form of privatisation did encourage competition between ex-NBC subsidiaries. Ideally, the government could have encouraged buy-outs to take place at sub-subsidiary level, although this would have complicated the privatisation process.

5.1.4 Subsidy

This section reviews the arguments for and against subsidy in the context of the bus industry, and examines the legislative changes introduced by the 1979 Conservative government to control expenditure on subsidy. Subsidy can be divided into two types: external and internal (i.e. cross-subsidy). The former is provided by local and central government to operators. The latter is provided by the operator when part of the revenue from profitable services is used to subsidise the running of unprofitable services.

5.1.4.1 External Subsidy

The case for subsidy depends on the existence of market failure. If markets fail to maximise social welfare, then there may be a case for the provision of subsidy to offset the market failure, provided that the benefits of the subsidy outweigh any undesirable side effects (Gwilliam, 1987). Market failure is here taken to include equity, as well as
economic efficiency considerations. It should be noted that it is not the desire to promote equity which creates a market failure, but rather the violation of the Pareto-efficiency conditions through the non-lump sum transfers which are often used to achieve this aim.

Chapter three concluded that the main sources of market failure in local bus services are likely to be wasteful competition and externalities. Although a lack of equity was not considered as a form of market failure, it obviously provides an important justification for subsidy in terms of social policy. Low income groups make significantly more use of buses than other income groups, and subsidising bus travel would be one method of income redistribution.

Ruling out subsidy as an appropriate form of intervention in the case of wasteful competition, the economic efficiency argument in favour of subsidy to local bus services depends mainly on the existence of externalities. Although economies of scale are often given as a justification for subsidy under marginal cost pricing, the weight of evidence suggests that there are no significant scale economies in bus operations (see Section 4.5.1).

There are two forms of externality that justify the provision of subsidy. First, the existence of unpriced congestion and pollution costs from car use implies that subsidy to public transport could be used as a second best solution to the externality problem (see Section 3.3.3.5). In general, buses use road space twenty times more efficiently than cars, and energy four times as efficiently. It is also the case that bus passenger deaths per passenger kilometre are one eighth those for car occupants (House of Commons Transport Committee, 1985, Second Report, Vol. II, op. cit., p. 2). The theoretical problems of calculating the benefits of a subsidy to reduce congestion are considered by Jackson (1975) and Shilony (1981). Second, the existence of unpriced benefits accruing from the "Mohring effect", whereby existing bus users gain from the provision of increased service frequencies (Mohring, 1972).
In a study of the effectiveness of urban public transport subsidy, Glaister (1987) concludes that, outside London, subsidies to reduce traffic congestion have limited effect. On the other hand, the "Mohring effect" was more important than expected, although this was estimated in the context of a reduction in both fares and service frequencies. In a study of major Australian cities, however, Dodgson (1987) concludes that, once the extra costs of operating more vehicle kilometres are taken into account, there are disbenefits from subsidising increased service frequencies. Both studies find that subsidy is generally more effective if it is used to decrease fares rather than increase services. This is not to argue that subsidy is justified, however, since neither study takes into account the problems of subsidy "leakage" or the appropriate shadow price of public funds.

Subsidies can also be provided for unremunerative services deemed "socially necessary". Hibbs (1971) has questioned this policy, arguing that "If people are not prepared to pay for it then there is at least a prima facie case for saying it is not required." But this is a dubious argument, since lack of demand for a service may arise because of an inability rather than an unwillingness to pay. In effect, this is the equity argument for subsidy. Generally speaking, it is the desire to promote equity, rather than reduce market failure, which has been the dominant influence in determining local authorities' policies towards public transport subsidy. The method of promoting greater equity has been the provision of concessionary fares for eligible groups, mainly the elderly, visually impaired, disabled and school children. As noted in Mackie and Preston (1996, p. 179), concessionary fares are a form of price discrimination and may also promote greater efficiency by reducing fares towards marginal costs.

The two main components of local authority subsidy to bus operators are revenue support and reimbursements towards the cost of concessionary fare schemes. Other sources of non-fare income include grants towards capital expenditure and fuel duty rebates (the latter being paid direct from central government). Between 1972 and 1982 expenditure on revenue support and concessionary fares increased from £22 million to £725 million, a real increase of 1,112%2. Prior to the 1985 Transport Act external subsidy was paid on a

---

2 Figures are based on Tables 17 and 18 of the Buses White Paper (1984).
"blanket" basis i.e. to a service network rather than to individual services within that network. The amount of subsidy paid was determined by negotiation between the local authority and the operator. In 1982/83 fares and other receipts, revenue support and concessionary fare scheme payments formed 62%, 28% and 10% respectively of bus operators' income (House of Commons Transport Committee, 1985, Second Report, Vol. II, p. 177).

Table 5.1 below shows the considerable differences in the regional distribution of these subsidies. These differences reflect the tendency of Labour controlled metropolitan counties to use transport subsidies as a means of achieving equity objectives. The low fares policy of the now defunct Greater London Council is a well known example. The abolition of the metropolitan county councils, together with tighter controls on public sector spending, would be expected to reduce the regional subsidy differential.

<table>
<thead>
<tr>
<th></th>
<th>Revenue Support</th>
<th>Concessionary Fares</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLC</td>
<td>29.7</td>
<td>7.3</td>
</tr>
<tr>
<td>Metropolitan Counties</td>
<td>19.9</td>
<td>7.2</td>
</tr>
<tr>
<td>English and Welsh non-metropolitan counties and districts</td>
<td>3.0</td>
<td>1.6</td>
</tr>
<tr>
<td>Scottish councils</td>
<td>7.4</td>
<td>7.9</td>
</tr>
<tr>
<td>Average</td>
<td>10.2</td>
<td>4.0</td>
</tr>
</tbody>
</table>


It is interesting to consider the subsidy figures in an international context: for example, the subsidy per capita in 1980 for New Zealand, Britain and Sweden was £2.3, £8.0 and £28.8 respectively (Webster, 1986). Taking into account both population and gross
domestic product, Britain has a higher level of subsidy than most other European countries covered in the above survey. The conflict between local authorities' desire to provide services and central government's desire to control public spending was illustrated in 1984/85, when local authorities' budgeted spending on public transport subsidy was £8.1 million, or 34% above the government's planned spending.

Given this background, the "blanket" subsidy system described above was criticised in two ways. Firstly, it did not encourage incumbent bus operators to improve their efficiency: on the contrary, efficiency improvements might have tempted the local authority to reduce the subsidy in a subsequent round of negotiations. It is possible that the provision of a blanket subsidy may have encouraged operators to become less efficient. Bly and Oldfield (1985), analysing subsidy data from over a hundred cities in selected countries, estimated that, on average, between 40 and 60 per cent of any increase in subsidy was associated with higher average costs, particularly wage increases. This process of subsidy "leakage" is corroborated by studies using American data (Anderson, 1983, Pucher et al., 1983, Windle, 1988). Although it could be argued that the direction of causation was from higher average costs to increases in subsidy, empirical tests by Bly and Oldfield (op. cit.) do not support this view4. Nevertheless, there are problems with the methodologies used in these studies, and the importance of subsidy leakage, if it does exist in Britain, may be overstated (Gwilliam, Nash and Mackie, 1985a).

Secondly, the system meant that subsidy was paid to a service network, so that losses on individual services were not made transparent to outside parties. Thus it was generally not possible for an outside party to ascertain the underlying structure of external subsidy across services. In the extreme case, a given network loss could arise from a large loss on a single service, or from small losses on a large number of services. It might be of interest to rate and taxpayers to know the destination and degree of subsidy to which they contribute, but under the blanket system this information was not available.

---

3 The figures refer mainly to stage carriage bus services, but also include subsidies to cover urban rail services of the Underground or Metro types (but not mainline rail services).
5.1.4.2 Internal Subsidy

The Traffic Commissioners administered the 1930 Road Traffic Act on the basis that operators were expected to provide unremunerative (but socially necessary) services in return for monopoly power on the profitable routes. Thus cross-subsidy was an integral part of this system. Ponsonby has defined internal or cross-subsidisation as a policy which:

"... involves deliberately and knowingly providing services or parts of services at a loss so that, if the undertaking as a whole is to pay its way, more than normal profits (i.e. monopoly) must be made from its other services."

(Ponsonby, 1969)

The practice of cross-subsidy, whilst enabling a network of profitable and unprofitable services to be maintained, is open to objections which can be summarised in terms of efficiency, equity and accountability. Firstly, in the absence of both demand and cost interdependencies there will be a welfare loss because price will not equal marginal cost on either the profitable or unprofitable service. The subsidised services may also become subject to the problem of "X" inefficiency. Secondly, the direction of cross-subsidy may be from low to high income groups (Posner, 1975). In this respect the White Paper cites the example of low income car-owning groups who make frequent use of urban bus services and high income car-owning groups who make infrequent use of (subsidised) rural services. Finally, the amount and direction of cross-subsidy is determined by bus operators rather than elected representatives accountable to their electorate.

Cross-subsidy occurs both spatially and temporally, although in practice its measurement is problematic because services may be related on the demand or cost side or both. Faulhaber (1975) recommends appropriate tests for the determination of cross-subsidy.

---

4 The study concluded, on the basis of introducing time lags into the regression analysis, that "the overall tendency was for change in subsidy to precede the changes in cost and output per employee, rather than vice versa". (Bly and Oldfield. 1985, p. 19).
Glaister and Mulley (1983) argue that, contrary to the usual view, peak bus users are generally subsidised by off-peak users because peak capacity can only be used productively for a relatively short time period. To give a rough indication of the importance of cross-subsidy, the NBC estimated that deregulation might force it to withdraw approximately £122 million of cross-subsidies. This is about twice the amount it received from external subsidies (Institute for Transport Studies 1984). The Institute study indicates that inter-urban routes support rural and, to a lesser extent, intra-urban routes. Temporally, the direction of subsidy depends on the method of cost allocation. Using a "preferred" allocated cost system, the study concludes that the general temporal direction of cross-subsidy was from weekday inter-peaks and Saturdays to weekday peaks and Sundays.

Cross-subsidy arises not only because of the desire to maintain a service network of profitable and unprofitable services, but also because of the practice of adopting uniform fare scales, usually on a pence per mile basis. This practice was favoured by the Traffic Commissioners because of its intrinsic simplicity for administrative purposes, and also as a means of establishing "proper" competition between operators. Given that different services will tend to have a different cost per mile, and given a break-even constraint, such a practice will inevitably lead to cross-subsidy between those services.

The introduction of competition reduces the ability of operators to cross-subsidise because the entry of new operators into profitable services decreases the profits available to cross-subsidise. In a situation of perfect competition, or perfect contestability, cross-subsidy would be completely eliminated.

5.1.4.3 Legislation and Subsidy Control 1983 - 88

Given the large increases in public expenditure on subsidies to public transport throughout the 1970's, together with a desire to reduce public spending generally, the

---

5 As with all estimates of cross-subsidy, this figure should be treated with some caution. In this particular case, it is also open to the charge of special pleading.
1979 Conservative government introduced a number of legislative changes designed to restrict the spending power of local authorities, particularly the metropolitan authorities. Between 1979 and 1989, for example, there were around twenty Acts passed that affected local government finance. This section describes the main features of those Acts that had most influence on subsidy to the bus industry. The issue of subsidy and the direction of causality with bus fares is an important topic analysed in more detail in Chapter six.

The 1983 Local Government Act introduced "Protected Expenditure Levels" set by the relevant Secretary of State. If Passenger Transport Authorities exceeded the level set by the Transport Secretary, they could be subject to legal challenge. In the same year the Transport Act required the PTAs to submit three-year financial plans to the Department of Transport, in order that proposed PTA spending could be more closely scrutinised. It also introduced the possibility of tendering for services on a limited scale. The 1984 Rates Act introduced "rate capping", whereby limits were imposed on the local taxes which local authorities could set to pay for services such as public transport. In 1985 the Local Government Act abolished the metropolitan counties and established the PTAs as joint-boards composed of elected district council members from within the metropolitan area. These new boards were also subject to rate capping. Finally, the 1988 Local Government Finance Act replaced the rating system with the Poll Tax (later the Community Charge). Business rates were replaced by a national non-domestic rate, so that local authority spending in excess of central government limits had to be financed by increases in the community charge.

The effect of these legislative changes was to place a much tighter financial constraint on local authorities. This was particularly evident in the PTAs such as South Yorkshire, where the changes forced the abandonment of low fare policies.

5.2 Effects of the 1985 Transport Act: a framework for analysis

The 1985 Transport Act deregulated local bus services and introduced a process of privatisation to the industry. The details of the Act are described in Appendix one.
order to analyse the effects of the 1985 Act, it is convenient to distinguish between the short, medium and long term effects. This section provides an explanation of these concepts in the context of the bus industry, and a justification for the preferred chronological structure. Most of the fundamental changes in the industry took place in the short and medium term, and these changes are discussed in Section 5.3. Discussion of the longer term effects are dealt with in Section 5.4. The key issue of subsidy reduction and the long term effects of deregulation are dealt with in Chapter six, where an econometric model is specified and estimated.

The short to medium term effects of the 1985 Act are defined as those effects occurring within the two and a half year period after deregulation on October 26 1986. The attempt to define the difference between the economic concepts of short, medium and long run periods in terms of chronological time will always depend, to a greater or lesser extent, on the nature of the market involved and the judgement of the analyst. Financial markets adjust relatively quickly, and the short run may be a matter of days or even hours. Employment contracts are relatively flexible, redundant computers and company cars can be resold fairly easily, and interest rates can be adjusted at very short notice.

The bus industry, on the other hand, faces a number of constraints on its ability to adjust. The bus industry is labour intensive (Monopolies and Mergers Commission, 1982), and for this reason alone is likely to adjust more slowly than less labour intensive industries. Trade union influence, although reduced by the effects of deregulation, still exists (Pickup et al., 1991). There is also the point that regulation of local services lasted for fifty years, engendering a set of values amongst many operators which deregulation would take time to modify. Given these considerations, it seems reasonable to assume that full adjustment to the effects of the 1985 Transport Act will take years, rather than months. Since the 1985 Act took effect on 26 October 1986 ("Deregulation Day"), the following sections often refer to the pre and post deregulation periods for convenience. It should be remembered, however, that the 1985 Act introduced deregulation of local

---

services and privatisation. Other legislation also introduced subsidy reduction for the industry (see Section 5.1.4).

The adjustment process is specified as follows. The short term effects are those occurring over the financial year periods 1985/86 - 1986/87. Data for financial years are used, since this is the method by which they are presented in the Bus and Coach Statistics, and Transport Statistics, publications from 1985/86 onwards. The change from calendar to financial year data causes two problems. Firstly, strict comparisons of time series data up to and including 1984 with those from 1985/86 onwards are made more difficult because calendar year periods do not coincide completely with financial year periods. Secondly, and more importantly for this section, the period 1986/87 covers both pre and post deregulation periods (7 and 5 months respectively). In comparing 1985/86 with 1986/87 data, more weight will be given to the pre-deregulation period so that the short term effects of the 1985 Act will tend to be understated.

A more equally weighted comparison could be obtained by comparing the 1986 and 1987 calendar year periods, but this is not possible using the published data. Another possibility is to compare the pre and post-deregulation periods within the 1986/87 period, namely the period 1 April 1986 to 25 October 1986 with that for 26 October to 31 March 1987. Bus and Coach Statistics, however, only gives this information for the vehicle mileage variable. Using this information, there is a fall in total local service vehicle mileage of 22.4%: using the data for 1985/86 to 1986/87, however, there is an increase of 4%. This shows clearly the sensitivity of the results to the choice of comparison periods. Another possibility is to use quarterly data to compare the pre and post deregulation periods within the 1986/87 period, but quarterly data is less reliable than annual data and is not readily available for most of the variables required for a comprehensive comparison. Given these difficulties, it seems appropriate to retain the 1985/86 - 1986/87 comparison, but to bear in mind that it will tend to underestimate the short term effects.

---

8 Data is collected by the Department of Transport from a sample of operators who are obliged to provide annual, but not quarterly, returns. Consequently, quarterly data tends to be less reliable.
Even so, some of the changes that occurred in this period are fairly dramatic, as Table 5.2 in Section 5.3.1 demonstrates.

The medium term effects are defined as those occurring over the periods 1986/87 - 1988/89. This chronology is designed to capture the "settling down" process after the initial changes caused by the 1985 Act. In comparing the changes between 1986/87 and 1987/88 there will be a tendency to overstate the effects of the Act, for reasons analogous to those given in the previous paragraph. In fact, some of the changes during this period are greater than those in the short term period. Nevertheless, by the period 1987/88 - 1988/89, the rate of change generally becomes much lower than in previous periods. Thus the short to medium term effects of the 1985 Act occur over the period 1985/86 - 1988/89 (see column 5 of Table 5.2).

It follows that the long term effects of the 1985 Act fall within the period from 1988/89 onwards. There are two advantages associated with this chronology. Firstly, the long term is often defined in economic theory, in so far as the theory is concerned with the dynamics of the adjustment process, as the period during which the industry reaches, or approaches, an "equilibrium". Equilibrium in this sense implies a steady-state or stable situation. This corresponds fairly well with the pattern of the data observed for the bus industry. The settling down process begins in the medium term, and continues into the long term\(^9\). Secondly, the long term can also be defined as the period during which fundamental changes take place in the structure of the industry. In the context of the bus industry, such a change began to occur sometime after 1989 when companies started to merge with or take over other companies:

"Since 1989, the industry has begun a further phase of adjustment. A process of concentration has begun based on the policy that market entry is most profitable by company merger and takeover rather than competition on-the-road or through tenders. This process is, in some areas, producing local monopolies and is likely to continue with

\(^9\) Long term changes are discussed in more detail in the next chapter.
the expansion of holding companies and with small operators retaining a significant though limited role."

(Pickup et al., 1991, p. 65)

This is a significant change since the move towards greater concentration, particularly amongst the larger companies, represents the beginning of a process that reverses the trend towards smaller, more independent companies initiated by the 1985 Act.

5.3 Short to Medium Term Effects of the 1985 Transport Act

The purpose of this section is to provide a discussion of the short to medium term effects of the 1985 Act on the local service bus industry, using the analytical framework developed in the previous section. Following deregulation there was a considerable amount of research on these effects, and more detailed accounts can be found in a large number of sources including Balcombe et al. (1987), Hopkin and Perrett (1987), Meadowcroft and Pickup (1987), Cahm and Guiver (1988), Rickard et al. (1988), White (1988), Perrett et al. (1989) and Pickup et al. (1991). The Transport Research Laboratory (TRL) published forty-six reports between 1983 and 1992 on deregulation and competition in the bus and coach industries. Table 5.2 below summarises some of the major trends following the 1985 Act in terms of a selection of indicator variables: column 2 captures the short term changes, columns 3 and 4 the medium term changes, and column 5 the changes occurring over the short to medium term. The changes are for the English metropolitan areas and the rest of Britain (excluding London and Strathclyde).

5.3.1 Overall Changes 1985/86 - 88/89

It is clear from Table 5.2 overleaf that changes in the (English) metropolitan areas were generally significantly different to those in the rest of Britain excluding London. This is

---

particularly evident for fare increases, where the 30.5% fare increase in the metropolitan areas contrasts with a negligible increase in the non-metropolitan areas.

Table 5.2: Short to medium term effects of the 1985 Transport Act

<table>
<thead>
<tr>
<th>Indicator/Variable</th>
<th>85/86-86/87</th>
<th>86/87-87/88</th>
<th>87/88-88/89</th>
<th>85/86-88/89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met (a) Non-Met</td>
<td>% Change</td>
<td>% Change</td>
<td>% Change</td>
<td>% Change</td>
</tr>
<tr>
<td>Passenger Journeys</td>
<td>-12.5</td>
<td>-4.3</td>
<td>-2.2</td>
<td>-18.1</td>
</tr>
<tr>
<td>Passenger Receipts (b)</td>
<td>+0.5</td>
<td>+0.3</td>
<td>-2.2</td>
<td>+1.4</td>
</tr>
<tr>
<td>Fares (b)</td>
<td>+24.8</td>
<td>+4.1</td>
<td>+0.3</td>
<td>+30.5</td>
</tr>
<tr>
<td>Vehicle Kilometres</td>
<td>-3.0</td>
<td>+10.6</td>
<td>+2.8</td>
<td>+10.3</td>
</tr>
<tr>
<td>Vehicle Stock (c)</td>
<td>-5.5</td>
<td>-19.2</td>
<td>-9.5</td>
<td>-34.5</td>
</tr>
<tr>
<td>Operating costs (b) (d)</td>
<td>-10.4</td>
<td>-19.2</td>
<td>-9.5</td>
<td>-34.5</td>
</tr>
<tr>
<td>Operating costs (b) (e)</td>
<td>-13.0</td>
<td>-10.7</td>
<td>-7.0</td>
<td>-27.8</td>
</tr>
<tr>
<td>Staff (c) Employed</td>
<td>-15.6</td>
<td>-12.8</td>
<td>-1.1</td>
<td>-27.2</td>
</tr>
<tr>
<td>Revenue Support (b)</td>
<td>-19.9</td>
<td>-34.2</td>
<td>-16.9</td>
<td>-56.2</td>
</tr>
</tbody>
</table>


Notes:
(a) "Met" refers to data for local services operated within the English metropolitan areas. Although this classification excludes the Strathclyde metropolitan area, it is chosen because it enables direct comparisons to be made between most of the indicator variables in Table 5.2. "All" refers to data for local services operated in Britain, excluding London. "Non-met" is equal to "All" minus "Met".
(b) Figures are inflation adjusted: 1985=100 for fares and passenger receipts. 1990/91=100 for other nominal variables.
(c) Figures are for bus and coach services i.e. they refer to the provision of local and non-local services. They are also based on metropolitan PTCs (rather than English metropolitan areas) and London Buses Ltd. (rather than London). Figures in rounded brackets in column 5 represent the estimates for staff employed in the provision of local services only. They are derived by using the proportion of local to total vehicle kilometres given in Table 1.3 of Bus and Coach Statistics (op. cit.) for the appropriate years as a weighting factor.
(d) Operating costs per vehicle kilometre, excluding depreciation.
(e) Total operating costs, excluding depreciation; data derived by grossing up from Tables 1.1 and 4.1 of Bus and Coach Statistics (op. cit.).
This difference was caused by the ending of low fares policies in the English metropolitan areas and the consequent sharp increase in fares, especially in South Yorkshire and Merseyside, although fares had started to rise before deregulation both as a result of rate capping and a desire to partially pre-empt the anticipated fare increases. The increase in fares may be understated because of a failure to adjust fare index weightings in a period of rapid change (Mackie, Preston and Nash, 1995). There was also a considerable amount of change in the Strathclyde metropolitan area where, in addition to fare increases, the PTC initially experienced strong competition from Scottish Bus Group subsidiaries and Stagecoach.

There was a substantial fall in operating costs per vehicle km in both metropolitan and non-metropolitan areas. The size of this fall matches the predictions made in the White Paper, and supports the view that there was considerable X-inefficiency in the regulated industry. The source of these cost reductions was mainly from increased labour productivity and decreased unit labour costs, although there may also have been a pure privatisation effect (Heseltine and Silcock, 1990).

The biggest change was the reduction in revenue support in both metropolitan and non-metropolitan areas. Revenue support is significantly greater than concessionary fare payments in these areas, and also for public sector operators. For private sector operators revenue support is less important than concessionary fare payments. Although the data is for areas rather than operators, from the limited evidence available there would appear to be a not unreasonable correlation between the two. For the metropolitan areas, about 80% of revenue support paid by the metropolitan counties in 1982/83 went to their Passenger Transport Executives (House of Commons Transport Committee, 1985, Second Report, op. cit., Table 2.8, p. 177).

The relationship between reductions in revenue support (subsidy) and fare increases can be interpreted in two ways: firstly, subsidy reduction can act as a cause of the increase in fares; secondly, it could be caused by the increase in fares (together with the reduction in operating costs). On the first interpretation, subsidy reduction is an exogenously
determined change to which the industry must react by increasing fares. On the second it is endogenously determined as a result of the ability of operators to force up fares and push down costs in the post deregulation period. The issue of causality between subsidy reduction and fare increases is discussed in the context of an econometric model in Chapter six. Whatever the exact status of the causality relationship it will be associated, \textit{ceteris paribus}, with fare increases and a decrease in passenger journeys. Table 5.2 shows that this is what happened, although the changes in the non-metropolitan areas were comparatively modest given the two-and-a-half year period over which they occurred. The change in passenger journeys must also be considered in the light of the long term decline occurring before deregulation: some of the change shown in Table 5.2 will be attributable to this long term decline, rather than to the effects of the 1985 Act \textit{per se}. This issue is also examined within the econometric analysis of Chapter six.

The exception to these broad movements was the independent operator sector (excluding NBC or ex-NBC operators), which experienced an increase in local service passenger journeys from 173 million to 339 million over the period 1985/86 to 1987/88. This is a very large increase of 96%, albeit from a very small base. In 1985/86 this sector accounted for approximately 4% of local service passenger journeys outside London. A similar picture emerges for non-local services, although the increase was smaller at 14.5%. Unfortunately it is difficult to track the performance of these operators after 1987/88, since official statistics do not distinguish between them and ex-NBC operators after this date.

Given the continued overall fall in demand, as measured by passenger journeys, it is likely that the industry would respond by contracting, or at least not increasing, its labour and capital stock. This observation is supported in the case of the metropolitan areas, where local service work forms a large proportion of the total local and non-local services. The evidence for the non-metropolitan areas is more difficult to assess, since the data in Table 5.2 does not distinguish between the vehicles and staff used for local and non-local services. The increase in vehicle stock may have occurred because of the upward trend in non-local service work since deregulation, particularly for the
independent operators for whom this type of work is extremely important. It may also have occurred because of the change to midi and minibuses over the period: a minibus will still count as one vehicle in the data above, even though its seating capacity will be much less than that of a double decker. It is possible for capital stock to increase on one measure (total vehicle stock) but to decrease on another (total vehicle seating capacity). The issue of capital stock change is discussed in the next section.

5.3.2 Vehicle Seating Capacity and Trends in Vehicle Size

In an attempt to test the hypothesis that vehicle seating capacity may have fallen in the non-metropolitan areas over the period, information on vehicle size as well as vehicle stock is taken from Table 6.3 of Bus and Coach Statistics 1990/91. This shows that the total public service vehicle stock in Britain outside London rose from 62,700 in 1985/86 to 66,900 in 1988/89, an increase of 6.7%. Within this overall total, the stock of small to medium size single decker vehicles rose from 9,600 in 1985/86 to 17,000 in 1988/89, an increase of 77.1%. By far the biggest user of these vehicles is the independent operator sector (which is now defined to include NBC or ex-NBC operators). Although neither the metropolitan nor the municipal PTEs used any small single deckers over the period, their use of medium single deckers increased from 200 to 1,700. This is a very large increase, albeit from a very small base.

The corresponding figures for large single and double decker vehicles are 53,000 and 49,400, a decrease of 6.8%. The use of large single deckers is concentrated overwhelmingly in the independent operator sector, but there was no significant change in their stock over the period. The biggest users of double deckers are the metropolitan PTCs and the independent operator sector (particularly the NBC and ex-NBC operators), with both sectors having fleets in the 7 to 8 thousand range in 1985/86. Over the period there was a significant difference between the sectors in terms of their double decker stock: in the independent operator sector it increased by 8%, but fell by 20% for the metropolitan PTCs.
Inspection of the figures for changes in the size distribution of vehicle stock can only reveal part of the picture, however, since they do not give any precise measure of the change in vehicle seating capacity. Assuming maximum seating capacities of 16, 35, 50 and 77 for small, medium, large single and double deckers respectively, total vehicle seating capacity rose over the period from 3,415,600 seats to 3,459,700, a fairly negligible increase of 1.3%. The distribution of this change across operator type and vehicle size is shown in Tables 5.3(i) and 5.3(ii) overleaf. Estimates for local and non-local service seating capacity are derived by allocating total seating capacity according to the number of local and non-local vehicle kilometres run by each operator type given in Table 1.3 of Bus and Coach Statistics 1990/91. Thus the change in total local service seating capacity is from 2,125,200 seats to 2,197,600 seats, an increase of 3.4%. For non-local services there is a decrease of 2.2%.

Table 5.3(i): Vehicle seating capacity (000s) 1985/86

<table>
<thead>
<tr>
<th>Vehicle Size (a)</th>
<th>Metropolitan PTCs</th>
<th>Municipal PTCs</th>
<th>SBG</th>
<th>Independent</th>
<th>Total (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>104.0</td>
<td>104.0</td>
</tr>
<tr>
<td>Medium</td>
<td>3.5</td>
<td>3.5</td>
<td>0.0</td>
<td>101.5</td>
<td>108.5</td>
</tr>
<tr>
<td>Large</td>
<td>45.0</td>
<td>60.0</td>
<td>120.0</td>
<td>1415.0</td>
<td>1640.0</td>
</tr>
<tr>
<td>Double</td>
<td>623.7</td>
<td>292.6</td>
<td>77.0</td>
<td>569.8</td>
<td>1563.1</td>
</tr>
<tr>
<td>Total:</td>
<td>672.2</td>
<td>356.1</td>
<td>197.0</td>
<td>2190.3</td>
<td>3415.6</td>
</tr>
<tr>
<td>Local</td>
<td>648.8</td>
<td>336.3</td>
<td>158.5</td>
<td>981.6</td>
<td>2125.2</td>
</tr>
<tr>
<td>Non-local</td>
<td>23.4</td>
<td>19.8</td>
<td>38.5</td>
<td>1208.7</td>
<td>1290.4</td>
</tr>
</tbody>
</table>


Notes:
(a) The terms "small", "medium" and "large" refer to the different sizes of single decker public service vehicles and are based on the distribution of seating capacities given in Bus and Coach Statistics i.e. up to 16 seats, 17-35 seats and 36 plus seats respectively; "double" refers to a double decker. Note that the figures in column 2 are for metropolitan PTCs, rather than English metropolitan areas.
(b) Totals refer to total estimated public service vehicle seating capacity provided in Britain, excluding London Buses Ltd.
Table 5.3(ii): Vehicle seating capacity (000s) 1988/89

<table>
<thead>
<tr>
<th>Vehicle Size</th>
<th>Metropolitan PTCs</th>
<th>Municipal PTCs</th>
<th>SBG</th>
<th>Independent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>121.6</td>
<td>121.6</td>
</tr>
<tr>
<td>Medium</td>
<td>38.5</td>
<td>17.5</td>
<td>10.5</td>
<td>255.5</td>
<td>322.0</td>
</tr>
<tr>
<td>Large</td>
<td>35.0</td>
<td>65.0</td>
<td>80.0</td>
<td>1350.0</td>
<td>1530.0</td>
</tr>
<tr>
<td>Double</td>
<td>500.5</td>
<td>284.9</td>
<td>84.7</td>
<td>616.0</td>
<td>1486.1</td>
</tr>
<tr>
<td>Total</td>
<td>574.0</td>
<td>367.4</td>
<td>175.2</td>
<td>2343.1</td>
<td>3459.7</td>
</tr>
<tr>
<td>Local</td>
<td>555.1</td>
<td>338.2</td>
<td>143.0</td>
<td>1161.3</td>
<td>2197.6</td>
</tr>
<tr>
<td>Non-local</td>
<td>18.9</td>
<td>29.2</td>
<td>32.2</td>
<td>1181.8</td>
<td>1262.1</td>
</tr>
</tbody>
</table>

Source and notes: as for Table 5.3(i) above.

Taking Tables 5.2 and 5.3 together, a number of comments can now be made on vehicle stock and seating capacity. Table 5.2 shows that total vehicle stock for all services (i.e. local and non-local) rose by 6.7%, but Table 5.3 shows that total vehicle seating capacity rose by only 1.3%. The inference is that, for the bus and coach industry as a whole, there was a slight trend away from the use of larger towards smaller vehicles. The next step is to determine whether the trend applies to both local and non-local services, and the strength of the trend in each sector. For local services only, the trend can be observed by noting that the 16.5% increase in vehicle kilometres given in Table 5.2 is matched by the much smaller increase in seating capacity of 3.4% implied in Table 5.3. On the basis of these figures, it would appear that there was a significant trend towards the use of smaller vehicles for local services. Given that the overall trend for the bus and coach industry is not as significant, the conclusion is that for non-local services the trend is less pronounced, or possibly non-existent.

Having established that the trend towards smaller vehicles is more likely to have occurred in the local service sector, further inspection of Table 5.3 shows how this trend is distributed between operator types. In order to measure the degree of vehicle size "concentration" (VSC) across operator types, the VSC is defined in terms of vehicle seating capacity as

\[
\text{VSC} = \frac{[(\text{Large} + \text{Double}) \times 100]}{\text{Total}}
\]
In 1985/86 the VSCs for the metropolitan, municipal, SBG and independent operator sectors were 99.5%, 99%, 100% and 90.6% respectively. By 1988/89 they had fallen in all sectors to 93.2%, 95.2%, 94% and 83.9%. This is not to say that all these decreases were fully accounted for by local services, since some part of the decrease may have occurred in non-local services, but given the discussion in the preceding paragraphs it is likely that most of the decrease occurred in local services.

Although total local service seating capacity increased by a modest 3.4%, the metropolitan PTCs had a decrease of 14.4% whilst the independent operators had an increase of 18.3%. This is consistent with the contraction of the metropolitan local service industry noted earlier in relation to Table 4.2, since the metropolitan PTCs account for most of the metropolitan area local services. Excluding the metropolitan PTCs, the increase in local service seating capacity rises from 3.4% to 11.3%. The only operator type to experience an increase in both local and non-local seating capacity was the municipal operators, where the 47.4% increase in non-local seating capacity was in sharp contrast to the rest of the non-local sector.

To conclude, the implications of the preceding analysis are that, for the local service sector in the short to medium term after the 1985 Transport Act, there was a clear trend towards decreasing vehicle size for all types of operator. Overall, there was a slight increase in the number of local service vehicle seats provided. In other words, reductions in the number of large single or double deckers were broadly offset (in terms of seating capacity) by increases in the number of medium to small single deckers. But there was a difference between the metropolitan and non-metropolitan areas: local service vehicle seating capacity decreased for the metropolitan PTCs, but increased elsewhere. The exception to this was the Scottish Bus Group, where local service vehicle seating capacity fell by 9.8%, although this operator forms a relatively small proportion of total seating capacity. Thus the hypothesis that the capital stock used for local services, as measured by local service vehicle seating capacity, fell in the non-metropolitan areas after the 1985 Act is not supported by the evidence. As far as non-local services are concerned, vehicle
seating capacity fell by 19.2% in the metropolitan PTCs, and by the much smaller amount of 1.9% for other operators combined.

5.3.3 Profit Margins

The second biggest change in Table 5.2 is for operating costs per vehicle kilometre. This is partly because of the large increase in vehicle kilometres travelled. The change in operating costs per passenger journey is much less because of the fall in passenger journeys. Although the change in total operating costs is not given explicitly in Bus and Coach Statistics, it is possible to derive an estimate of this figure by using the formula:

\[
\frac{\text{(Operating costs per vehicle kilometre} \times \text{vehicle kilometres)} + \text{(Operating costs per passenger journey} \times \text{passenger journeys})}{2}.
\]

In theory, grossing up from either vehicle kilometres or passenger journeys should give the same figure for total operating costs, but in practice there is a small difference. Hence the final figure is derived from the average of the upper and lower values. The results are given in Table 5.4, along with figures for total revenue and operating profit:

Table 5.4: Costs, revenues and operating profit 1985/86-1988/89

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Met (a)</td>
<td>Non-Met</td>
<td>Met</td>
<td>Non-Met</td>
</tr>
<tr>
<td>Total Operating</td>
<td>831</td>
<td>1286</td>
<td>725</td>
<td>1211</td>
</tr>
<tr>
<td>Costs (£mill)(b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Revenue</td>
<td>875</td>
<td>1361</td>
<td>820</td>
<td>1295</td>
</tr>
<tr>
<td>(£mill)(c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating Profit</td>
<td>44</td>
<td>75</td>
<td>95</td>
<td>84</td>
</tr>
<tr>
<td>(£mill)(d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating Profit</td>
<td>5.0</td>
<td>5.5</td>
<td>11.6</td>
<td>6.5</td>
</tr>
<tr>
<td>Margin % (e)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Notes:
(a) "Met" refers to data for local services operated within the English metropolitan areas. "Non-met" refers to data for local services operated in Britain, excluding London and the English metropolitan areas. Figures are inflation adjusted (1990/91=100).
(b) Costs exclude depreciation and are net of fuel duty rebate.
(c) Total revenue consists of passenger receipts, concessionary fares and revenue support.
(d) Operating profit = Total revenue - Total operating costs.
(e) Operating profit margin % = (Operating profit*100)/Total revenue. Figures in brackets are the profit margins with depreciation included. Data for operating costs including depreciation are not available in Bus and Coach Statistics (op. cit.) prior to 1987/88.

Once again, there is a marked difference in performance between the metropolitan and non-metropolitan areas for most years. Between 1985/86 and 1988/89 total operating costs in the metropolitan areas fell by 27.1% in real terms, whilst total revenue decreased by 17.8%. In the non-metropolitan areas, the corresponding figures were 9.6% and 11.5%. The large reduction in operating costs enabled the metropolitan areas to generate a much higher profit margin than in the non-metropolitan areas. In the latter areas the profit margin after the inclusion of depreciation costs is negative for the years of 1987/88 and 1988/89. Taking the two areas together, the overall profit margin (including depreciation costs) for these years is 0.7% and 1.9%. Since these figures represent pre-tax profits, it is not clear whether even the metropolitan sector made sufficient post-tax profits to cover long run costs.

The figure of 1.9% overall pre-tax profit margin for 1988/89 contrasts with the 3.2% figure (excluding coach operators) implied in the Bus and Coach Industry Monitor (1991) and the figure of 5% given in the Price Waterhouse report (TRRL, 1990, p. 16). All three figures are based on very large samples, and there is no reason to assume that one figure is more reliable than the others. The financial year 1988/89 is the only one covered by all three sources. Different accounting practices between operators should have a similar effect on all three figures. The accounts of SBG subsidiaries, for example, are prepared on a current cost replacement basis, which will tend to reduce their profit margins below that of other operators. The Price Waterhouse report (PWR) figure excludes depreciation costs, but whether their inclusion would reduce it nearer to 1.9% or 3.2% is difficult to determine.
A feature of the Bus and Coach Industry Monitor (BCIM) results for 1988/89 is that there is no major difference between the profit margins of the metropolitan PTCs and other operators. The figure for the metropolitan PTCs (excluding London Buses Ltd. from the BCIM results) is 3.8%, compared with the overall figure of 3.2%. This contrasts with the conclusions based on the Bus and Coach Statistics (BCS) figures in Table 5.2, although these relate mainly to metropolitan and non-metropolitan areas rather than companies. The PWR does not distinguish between metropolitan and non-metropolitan areas or companies, so that no comparison is possible from this source.

To conclude, comparison with the profit margins achieved in a range of other industries shows that the bus and coach industry has one of the lowest profit margins in the sample over the period 1986/87 to 1987/88 (PWR, op. cit., p. 24). Although there is some evidence of a recovery in 1988/89, there seems little doubt that profit margins in the local service industry as a whole were too low in the short to medium term after the 1985 Transport Act to ensure long run financial viability. Government statistics imply a significantly different profit margin performance between the metropolitan and non-metropolitan areas, but this difference is not evident from the BCIM figures.

5.3.4 Return on Capital Employed

The performance of the bus and coach industry is even worse if the return on capital employed is used as a performance indicator (PWR, op. cit., p. 24). From 1985/86 to 1988/89 the return on capital employed was by far the lowest amongst a range of other industries, and well below the rate of return from risk-free government securities. Strictly speaking, it is the rate of return on capital employed, rather than the profit margin, which determines long run industry viability. A persistently low return implies that the industry will not attract sufficient funds to ensure long term survival. There are problems in calculating the rate of return on capital employed, however, because of significant variations in company accounting practices in relation to "returns" and "capital employed".
The PWR uses a linear multiple regression model to estimate the influences on rates of return in the bus industry for 1987/88. The model and estimation results are as follows:

\[ P = -84.0 - 0.086S + 139.0C - 250.0M - 1.8SB + 0.01PD + 51.1PR - 44.9G \]

\[ (52.7) (0.36) (64.8) (87.6) (52.2) (0.86) (15.3) (15.5) \]

where:
- \( P \) = profits per seat (before interest and tax)
- \( S \) = total seats
- \( C \) = market concentration
- \( M \) = proportion of minibuses in bus fleet
- \( SB \) = subsidised miles run
- \( PD \) = population density
- \( PR \) = private ownership
- \( G \) = group ownership

Figures in brackets are standard errors. \( R^2 = 0.3 \)
Sample size = 109
\( F_{(7,101)} = 6.18 \)

The dependent variable \( P \) is a proxy for return on capital employed, with total seats acting as the measure of capital employed. Four of the seven explanatory variables are statistically significant: market concentration, proportion of minibuses, private ownership and group ownership. Surprisingly, the minibus variable has a negative influence on rates of return whilst population density, although it has the expected positive coefficient, is statistically insignificant. The \( F \) value of 6.18 is higher than the critical \( F \) value for 7 and 101 degrees of freedom, so that the null hypothesis that all the true regression coefficients are simultaneously zero can be rejected.

The \( R^2 \) value is low, however, and implies that the model explains only 30% of the variation in rates of return. In addition, there is a simple, but potentially serious, problem in the specification of the model. The dependent variable is defined as profits divided by total seats, and this is regressed on a number of explanatory variables including total seats. The point of dividing by total seats is to remove the effect of changes in capital stock on profits, in the same way that using income per capita removes the effect of changing population on incomes. Having done this, it makes little sense to include total seats as an explanatory variable in the model. If total seats is used as an explanatory variable, then the appropriate dependent variable is simply profits, not profits divided by...
total seats. The consequence of including total seats on both left and right-hand sides of the equation will be a higher $R^2$, and lower standard errors, than would otherwise be the case. In other words, criteria such as goodness of fit and parameter significance will be overstated in the model above.

5.3.5 Labour Productivity

There was a large decrease in the number of staff employed in the metropolitan PTCs over the period 1985/86 to 1988/89, particularly in the short run. The reduction was most pronounced in the area of maintenance staff where the number employed fell from 10,300 to 5,900, a decrease of 42.7%. For the non-metropolitan operators outside London the number of staff employed stayed roughly constant over the period: broadly speaking, the municipal and SBG operators shed 2,100 jobs, whilst the NBC and independent operators gained 2,800. This latter total consisted of an increase of 6,900 platform staff (i.e., drivers, conductors and other on-vehicle staff) and a decrease of 4,100 in maintenance and other staff. The decrease in maintenance staff across both metropolitan and non-metropolitan operators reflects in part the move towards the contracting out of maintenance work.

Since the data in BCS does not distinguish between local and non-local service staff, an estimate of the number of staff employed in local service work is derived by using the appropriate split between local and non-local vehicle kilometres to allocate staff between the two service categories. The resulting adjusted figure for metropolitan operators between 1985/86 and 1988/89 is a decrease of 27.1%, compared to the unadjusted figure of 27.2%. The corresponding figures for non-metropolitan operators are +7.5% and +0.6%. These (adjusted) changes in staff employed of -27.1% and +7.5% are broadly consistent with the changes in total vehicle seating capacity of -14.4% and +11.3% estimated in Table 5.3. If labour productivity is measured as staff employed per vehicle seat, then it is clear that it rose significantly in the metropolitan PTCs and less significantly amongst the non-metropolitan operators.
5.3.6 Commercial and Subsidised Services

The 1985 Transport Act requires operators to register the local bus services they wish to provide. In general, these services will be commercial services, although there may be instances where unprofitable services have been registered or re-registered in order to deter entry. Given the importance of cross-subsidy in the regulated industry, and the decrease in operators' ability to cross-subsidise in the deregulated industry, it was widely expected that registered vehicle mileage would be significantly less than the vehicle mileage run before deregulation.

This expectation was confirmed. Using the pre-deregulation NBC (which provided about one-third of Britain's bus services) as an example, the pattern of its registrations before 28th February 1986 can be seen in Table 5.5:

Table 5.5: Registration of NBC services (000s bus miles)

<table>
<thead>
<tr>
<th>Service</th>
<th>Original Mileage</th>
<th>Registered Mileage</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional bus</td>
<td>448,876</td>
<td>312,758</td>
<td>-30.0</td>
</tr>
<tr>
<td>Minibus</td>
<td>23,630</td>
<td>79,658</td>
<td>+237.0</td>
</tr>
<tr>
<td>Urban</td>
<td>293,144</td>
<td>167,242</td>
<td>-8.5</td>
</tr>
<tr>
<td>Rural</td>
<td>162,667</td>
<td>105,127</td>
<td>-35.4</td>
</tr>
<tr>
<td>Main day (Mon-Sat)</td>
<td>321,812</td>
<td>293,433</td>
<td>-8.8</td>
</tr>
<tr>
<td>Early morning (Mon-Sat)</td>
<td>29,756</td>
<td>18,670</td>
<td>-37.2</td>
</tr>
<tr>
<td>Evening (Mon-Sat)</td>
<td>50,406</td>
<td>28,599</td>
<td>-43.3</td>
</tr>
<tr>
<td>Sunday</td>
<td>21,906</td>
<td>10,356</td>
<td>-52.8</td>
</tr>
<tr>
<td>NBC Total</td>
<td>472,506</td>
<td>392,416</td>
<td>-16.9</td>
</tr>
</tbody>
</table>

Source: "Bus Special", No. 6, June 1986. TGWU Passenger Services Group.
Note: Figures not available for Southern Vectis.

The only case where registered mileage exceeded the original mileage was for minibus operations, and the extent of the increase indicates the potential for innovation in this area. In fact, the NBC had already been experimenting with minibus services on urban routes from 1984, introducing services in housing estates and a "hail stop" system (Birks et al, 1990). Elsewhere, registered rural and off-peak vehicle mileage (especially
Sundays) was significantly below the pre-deregulation mileage. In total, NBC registered mileage was 17% lower than its pre-deregulation mileage.

Local authorities can subsidise the provision of non-registered services via the competitive tendering system. It appears that, in the short term, local authorities chose to maintain these services at "broadly" their pre-deregulation level (Balcombe et al., 1987, p. 10).

Frequency reductions did occur in certain areas, particularly on off-peak and Sunday services, because local authorities decided that the required subsidy did not produce sufficient benefit. In Greater Manchester, frequency reductions were initially made which amounted to a loss of 20% of bus mileage and an estimated subsidy saving of some £10 million (although the PTE subsequently replaced some lost services). Whilst this was an extreme case it serves to focus attention on the distributional effects of the 1985 Transport Act. Although the Government made extra money available to maintain vital rural services, cuts that do occur may seriously harm the quality of life for those who lack access to a car. Guiver and Hoyle (1987) provide an account, on the basis of the "Buswatch" survey, of the social and economic problems caused by the loss of services.

Nonetheless, the proportion of subsidised to total vehicle mileage remained at around 16% in the short to medium term. Since total vehicle mileage followed an upward trend over this period, it follows that the subsidised network did not in aggregate decrease. The figure of 16% subsidised total vehicle mileage is highly consistent with that implied for the NBC in Table 5.5 above.

5.3.7 Competitive Tendering

The system of competitive tendering set up by the 1985 Transport Act is generally perceived to have worked successfully in the post-deregulation period (Glaister, 1993, Nash, 1993). Balcombe et al. (1987) estimate that subsidy (net of costs such as company restructuring and additional administration) fell in the short run from about £95m to
£73m in non-metropolitan counties in England (at 1985-86 prices), from £24m to £18m in Scotland, and from £9m to £6m in Wales. This estimate may be subject to downward revision in the medium to longer term if operators misjudged market conditions. On the other hand, restructuring costs will be a one-off cost so that the estimated figure may be subject to upward revision for the medium to longer term. In general, it seems reasonable to assume that the competitive tendering system reduced subsidy by approximately 20 to 30%. The average number of tenders per contract was estimated to be between two and three, although there were a significant number of single tender cases, especially in rural areas. Lack of competition may cause the subsidy requirement to increase in these areas.

Some operators were apparently willing to cross-subsidise their tendered network from their commercial services, either to deter new entrants or gain a foothold in a rival's market (Gwilliam, 1989). One example occurred in Tayside, where tenders were invited for the evening and Sunday operation of two services in Dundee. For the first service, the tenders were: Strathtay Scottish (a member of the Scottish Bus Group) £6,500, Greyhound Luxury Coaches £36,140, and Tayside Public Transport Co. Ltd. £37,960: for the second service, the respective tenders were £26,260, £61,672 and £73,645. This prompted an official complaint to the Scottish Office from the Roads and Transport Convener of Tayside region about the "predatory policy" operated by the Scottish Bus Group. The contract was eventually awarded to Strathtay Scottish.

It is sometimes argued that the tendered market is more contestable than the commercial market (Gwilliam, 1989, Mackie, Preston and Nash, 1985). The main reason given is that reaction times in the tendered market are longer than those in the commercial market, so that strategic entry deterrence cannot be implemented quickly. If a tender is lost, it will typically be three years time before the losing operator can compete for the tender again. There are two comments in order here: first, competition can still occur in the form of registering (or reregistering) services which compete with those operated by the tender winner; second, and more important, the ability of the tender winner to exit the market is constrained to a much greater extent than in the commercial market, where only 42 days notice of service withdrawal is normally required. In practice, an operator wishing to exit
from the tendered market may be allowed to exit before the contract expiry date, although this is by no means a certainty. For these reasons, it is not clear that the tendered market is more contestable than the commercial market.

5.4 The Longer Term Effects of the 1985 Transport Act

This section reviews some of the longer term effects of deregulation and is based on information given in the Transport Committee First Report (TCFR) in November 1995.

5.4.1 Changes in Market Structure

Since deregulation there have been two major phases of consolidation in the local bus service industry. The first, roughly from 1988 to 1991, concerned the weakest of the ex NBC companies. The second, from 1993 to 1995, concerned the ex PTE and STG companies as well as other ex NBC companies. The main reason for this consolidation is the force of competition in the deregulated market, where the search for cost savings created considerable structural adjustments in the market. Table 5.6 shows the changes in market share associated with different types of operator since 1989:

By 1995 the industry was dominated by the "big three" i.e. Stagecoach, First Bus (Badgerline/GRT) and British Bus. A second tier of smaller, but still substantial groups, had also developed i.e. West Midlands Travel, Go-Ahead Group, Cowie Group and MTL Trust Holdings. The rest of the industry is made up of smaller groups (two or three companies each) and single units. These single units are owned either by local authorities, company managers or company employees. The independent sector (defined as operators owning less than 50 vehicles each) has seen a small, but steady growth in market share despite the consolidation amongst the larger operators.
Table 5.6: Changes in market share (%) by turnover in the bus industry 1989-95

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First tier groups:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stagecoach Holdings</td>
<td>3.9</td>
<td>3.7</td>
<td>4.9</td>
<td>4.9</td>
<td>6.9</td>
<td>13.4</td>
<td>13.4</td>
</tr>
<tr>
<td>Badgerline</td>
<td>3.1</td>
<td>4.3</td>
<td>5.1</td>
<td>5.0</td>
<td>5.0</td>
<td>9.3</td>
<td>9.3</td>
</tr>
<tr>
<td>GRT</td>
<td>0.6</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
<td>1.8</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>British Bus</td>
<td>2.8</td>
<td>3.5</td>
<td>3.4</td>
<td>3.4</td>
<td>3.9</td>
<td>7.9</td>
<td>9.8</td>
</tr>
<tr>
<td><strong>First tier total:</strong></td>
<td>10.4</td>
<td>12.6</td>
<td>14.6</td>
<td>14.5</td>
<td>17.6</td>
<td>34.0</td>
<td>35.9</td>
</tr>
<tr>
<td><strong>Second tier total:</strong></td>
<td>2.3</td>
<td>2.3</td>
<td>8.3</td>
<td>8.2</td>
<td>12.5</td>
<td>22.0</td>
<td>22.1</td>
</tr>
<tr>
<td><strong>First and second tier:</strong></td>
<td>12.8</td>
<td>14.9</td>
<td>22.9</td>
<td>22.6</td>
<td>30.1</td>
<td>56.0</td>
<td>58.0</td>
</tr>
<tr>
<td><strong>Smaller groups</strong></td>
<td>8.1</td>
<td>8.9</td>
<td>9.1</td>
<td>9.2</td>
<td>7.1</td>
<td>5.5</td>
<td>4.3</td>
</tr>
<tr>
<td><strong>Employee owned</strong></td>
<td>1.7</td>
<td>1.7</td>
<td>3.8</td>
<td>4.2</td>
<td>6.6</td>
<td>3.5</td>
<td>3.7</td>
</tr>
<tr>
<td><strong>Management owned</strong></td>
<td>13.8</td>
<td>16.6</td>
<td>17.2</td>
<td>16.8</td>
<td>17.8</td>
<td>13.0</td>
<td>11.9</td>
</tr>
<tr>
<td><strong>Publicly owned</strong></td>
<td>30.4</td>
<td>25.4</td>
<td>18.8</td>
<td>18.1</td>
<td>9.2</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Scottish Bus Group</td>
<td>5.7</td>
<td>4.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>London Transport</td>
<td>15.5</td>
<td>15.5</td>
<td>15.4</td>
<td>15.2</td>
<td>15.2</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Independent Operators</td>
<td>12.0</td>
<td>12.2</td>
<td>12.9</td>
<td>13.9</td>
<td>14.1</td>
<td>14.6</td>
<td>14.6</td>
</tr>
</tbody>
</table>


Note: Where data for 1995 was not available, it is based on 1994 data.

Overall, even allowing for the emergence of the big three, the structure of the industry remains reasonably diverse in terms of market share. It also appears to be the case that ownership in geographical terms is reasonably diverse (Transport Committee, First Report, op. cit., Vol. III, p. 271).

### 5.4.2 Profitability

Pre-tax profit in the bus industry fell and then improved between 1988/89 and 1993/94. There are wide variations in profitability within the industry according to the type of ownership, but the major groups have maintained a significantly better profit performance overall. Table 5.7 summarises this profit performance according to type of ownership:
Table 5.7: Pre-tax profit (as % of turnover) by operator type 1988-94

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Major groups</td>
<td>6.08</td>
<td>9.62</td>
<td>6.53</td>
<td>9.50</td>
<td>9.10</td>
<td>10.19</td>
</tr>
<tr>
<td>Second tier groups</td>
<td>1.55</td>
<td>-4.02</td>
<td>2.08</td>
<td>3.19</td>
<td>8.51</td>
<td>11.79</td>
</tr>
<tr>
<td>Small groups</td>
<td>4.88</td>
<td>5.92</td>
<td>5.44</td>
<td>5.23</td>
<td>6.57</td>
<td>5.73</td>
</tr>
<tr>
<td>Management owned</td>
<td>5.24</td>
<td>5.34</td>
<td>4.85</td>
<td>5.61</td>
<td>6.53</td>
<td>7.03</td>
</tr>
<tr>
<td>Employee owned</td>
<td>4.16</td>
<td>7.25</td>
<td>5.41</td>
<td>2.28</td>
<td>3.40</td>
<td>2.89</td>
</tr>
<tr>
<td>Publicly owned</td>
<td>4.60</td>
<td>3.19</td>
<td>2.07</td>
<td>1.94</td>
<td>3.05</td>
<td>4.81</td>
</tr>
<tr>
<td>Scottish Bus Group</td>
<td>2.54</td>
<td>0.53</td>
<td>-0.34</td>
<td>6.37</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>All operators</td>
<td>4.4</td>
<td>3.3</td>
<td>2.8</td>
<td>4.0</td>
<td>5.3</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Source: Bus Industry Monitor (TAS Partnership) for House of Commons Transport Committee First Report, 1995, Vol. III, adapted from Figure D and Table 4, p. 274.

Note: No data available for independent operators i.e. those owning less than 50 vehicles each.

The fall in profitability from 1988 to 1991 was undoubtedly a factor in encouraging further consolidation in the industry. Although profitability increased after 1991, it seems likely that only the larger companies made sufficient profit to provide adequately for long-term needs such as investment. Nonetheless, the pre-tax profit figure of 6.8% for 1993/94 represents the industry's best performance since deregulation.

5.4.3 Costs and Employment

Operating costs have fallen significantly since deregulation. Between 1986 and 1994 operating costs per vehicle kilometre (excluding depreciation) fell by 41% in England. Part of these cost savings were from reductions in the number of people employed, particularly in management and administrative staff, and part were from reductions in the real wage paid to drivers. Table 5.8 shows the changes in average earnings of bus (and coach) drivers from 1975 to 1994:

185
Table 5.8: Average earnings of bus and coach drivers (£ per hour at April 1994 prices)

<table>
<thead>
<tr>
<th></th>
<th>Bus and Coach drivers (1)</th>
<th>All Manual Workers (1)</th>
<th>Bus and Coach as % of Manual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>5.51</td>
<td>5.38</td>
<td>102</td>
</tr>
<tr>
<td>1983</td>
<td>5.58</td>
<td>5.59</td>
<td>100</td>
</tr>
<tr>
<td>1984</td>
<td>5.48</td>
<td>5.61</td>
<td>98</td>
</tr>
<tr>
<td>1985</td>
<td>5.48</td>
<td>5.60</td>
<td>98</td>
</tr>
<tr>
<td>1986</td>
<td>5.70</td>
<td>5.80</td>
<td>98</td>
</tr>
<tr>
<td>1987</td>
<td>5.34</td>
<td>5.91</td>
<td>90</td>
</tr>
<tr>
<td>1988</td>
<td>5.44</td>
<td>6.08</td>
<td>89</td>
</tr>
<tr>
<td>1989</td>
<td>5.27</td>
<td>6.07</td>
<td>87</td>
</tr>
<tr>
<td>1990</td>
<td>5.18</td>
<td>6.09</td>
<td>85</td>
</tr>
<tr>
<td>1991</td>
<td>5.08</td>
<td>6.18</td>
<td>82</td>
</tr>
<tr>
<td>1992</td>
<td>5.17</td>
<td>6.29</td>
<td>82</td>
</tr>
<tr>
<td>1993</td>
<td>5.09</td>
<td>6.37</td>
<td>80</td>
</tr>
<tr>
<td>1994</td>
<td>5.03</td>
<td>6.31</td>
<td>80</td>
</tr>
</tbody>
</table>


Note: Males on adult rates since 1984 and full-time men over 21 in earlier years.

It is clear from Table 5.8 that the pre-deregulation real wage for bus and coach drivers was generally somewhat below that of all manual workers, whereas in the post-deregulation period it fell by a substantial margin. Table 5.9 shows how labour costs per employee varied according to sector, together with the changes in number of employees over the period 1989/90 to 1993/94:

Table 5.9 shows that there are significant variations in labour costs per employee between different sectors. Particularly striking are the figures for the PTA/Ex-PTA sector, where there is a large differential between this sector and the independent sector. This is an important factor in explaining why the PTA sector, particularly the ex-PTAs, has been so vulnerable to new entrants to the market.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Independents</td>
<td>11,483</td>
<td>12,929</td>
<td>12,282</td>
<td>11,138</td>
<td>11,236</td>
</tr>
<tr>
<td>London Buses</td>
<td>14,701</td>
<td>16,603</td>
<td>17,256</td>
<td>17,239</td>
<td>16,601</td>
</tr>
<tr>
<td>Local Authority owned</td>
<td>11,820</td>
<td>12,470</td>
<td>13,319</td>
<td>14,025</td>
<td>14,236</td>
</tr>
<tr>
<td>Ex-National Bus</td>
<td>10,736</td>
<td>11,356</td>
<td>12,119</td>
<td>12,613</td>
<td>12,276</td>
</tr>
<tr>
<td>PTA/Ex-PTA</td>
<td>13,867</td>
<td>13,211</td>
<td>14,156</td>
<td>14,952</td>
<td>15,235</td>
</tr>
<tr>
<td>Ex-Scottish Transport</td>
<td>9,944</td>
<td>11,880</td>
<td>12,222</td>
<td>12,619</td>
<td>12,988</td>
</tr>
<tr>
<td>All Companies</td>
<td>11,750</td>
<td>12,718</td>
<td>13,562</td>
<td>13,441</td>
<td>13,389</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Annual changes (%) in:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Labour costs</td>
<td>-</td>
<td>+8.2</td>
<td>+6.6</td>
<td>-0.9</td>
<td>-0.4</td>
</tr>
<tr>
<td>All employees</td>
<td>-</td>
<td>-1.7</td>
<td>-1.2</td>
<td>-2.9</td>
<td>-0.8</td>
</tr>
<tr>
<td>Management and admin staff</td>
<td>-</td>
<td>-6.6</td>
<td>-6.2</td>
<td>-6.1</td>
<td>-3.9</td>
</tr>
</tbody>
</table>


Although wages have fallen, this does not necessarily imply that net pay has also fallen. Stagecoach claims that average net pay for its employees has actually increased through the introduction of profit-related pay schemes and share participation schemes. The purpose of these arrangements is to encourage employees to create a long term career in the industry and thereby reduce costly labour turnover. It is uncertain, however, whether these arrangements are typical of the industry as a whole.

### 5.4.4 Vehicle Seating Capacity and Trends in Vehicle Size

Given the increase in product differentiation in terms of vehicle types since deregulation, measuring the supply of the industry by vehicle kilometres is not an ideal indicator of local service bus supply. Vehicle size ranges from the minibus, with a capacity of 16 seats or less, to the double decker with a capacity of 80 seats and above. Although vehicle kilometres have increased substantially since deregulation, a better indication of supply is the number of seat miles which operators have delivered. Vehicle kilometres are not disaggregated by the Department of Transport by type of vehicle, so it is not possible to provide a direct figure for seat miles from official statistics. A survey of local bus
operators (49+ vehicles) shows that seat miles have in fact fallen by around 5% since 1989/90. The changes are summarised in Table 5.10:

Table 5.10: Bus supply by bus and seat miles

<table>
<thead>
<tr>
<th></th>
<th>Bus miles (mill)</th>
<th>Average bus size</th>
<th>Seat miles (mill)</th>
<th>(% change on previous year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989/90</td>
<td>1516</td>
<td>58</td>
<td>87928</td>
<td>-</td>
</tr>
<tr>
<td>1990/91</td>
<td>1519</td>
<td>57</td>
<td>86585</td>
<td>-1.5</td>
</tr>
<tr>
<td>1991/92</td>
<td>1539</td>
<td>57</td>
<td>87710</td>
<td>+1.3</td>
</tr>
<tr>
<td>1992/93</td>
<td>1562</td>
<td>54</td>
<td>84340</td>
<td>-3.8</td>
</tr>
<tr>
<td>1993/94</td>
<td>1574</td>
<td>53</td>
<td>83441</td>
<td>-1.1</td>
</tr>
<tr>
<td>Overall change</td>
<td>58</td>
<td>-5</td>
<td>-4487</td>
<td>-5.1</td>
</tr>
</tbody>
</table>


The change to smaller vehicles, which was noted for the short to medium term in Section 5.3.2, has continued in the long term. This has outweighed the increase in vehicle miles, so that seat miles have declined. Since the number of passenger journeys for all operators fell by 13.5% over the period, however, there is still evidence of over-capacity in the industry in this period.

5.4.5 Investment in the Bus Industry

There was a severe shortage of investment in the British bus industry from 1985 to 1992, when the age profile of the bus fleet worsened considerably and pre-tax profit as a percentage of turnover never exceeded 4.5%. Investment began to recover from the early 1990s onwards. These trends were reversed for London. Capital investment (i.e. new vehicles and other new equipment) is shown in Table 5.11:
Table 5.11: Capital investment in the bus industry (£m) 1989-94

<table>
<thead>
<tr>
<th></th>
<th>Britain exc. London (£m)</th>
<th>London Buses (£m)</th>
<th>Total (£m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989/90</td>
<td>155.34</td>
<td>24.76</td>
<td>180.10</td>
</tr>
<tr>
<td>1990/91</td>
<td>127.49</td>
<td>26.01</td>
<td>153.50</td>
</tr>
<tr>
<td>1991/92</td>
<td>116.82</td>
<td>37.08</td>
<td>153.90</td>
</tr>
<tr>
<td>1992/93</td>
<td>122.28</td>
<td>45.82</td>
<td>168.10</td>
</tr>
<tr>
<td>1993/94</td>
<td>168.70</td>
<td>20.66</td>
<td>190.60</td>
</tr>
</tbody>
</table>


The improvement in profitability (see Table 5.7) provided a considerable stimulus to the recovery in investment. In general, the larger groups invested more per unit of turnover than the rest of the industry. During 1993/94, for example, the larger groups invested £12 per £1000 turnover in new vehicles and equipment compared to an average of £9 per £1000 turnover for the rest of the industry.

Investment in new buses by vehicle type is shown in Table 5.12 below. The figures in Table 5.12 show the decline of the double decker from the late 1980s onwards and, somewhat surprisingly, the supermini and the minibus. In the case of these latter two vehicle types, the reduction in investment may arise as a result of operators having achieved their desired levels of stocks of these vehicles, so that investment is no longer to expand vehicle stocks and replace existing vehicles, but mainly replacement investment. The midibus appears to be the vehicle which operators have favoured most since their introduction in the early 1990s. Apart from the recession years of 1990/91 and 1991/92, the big three operators had the highest capital investment as a proportion of turnover.
Table 5.12: Investment in new buses (number) 1988-94

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Double deck</td>
<td>529</td>
<td>797</td>
<td>543</td>
<td>317</td>
<td>223</td>
<td>275</td>
<td>178</td>
</tr>
<tr>
<td>Single deck</td>
<td>342</td>
<td>606</td>
<td>607</td>
<td>390</td>
<td>322</td>
<td>502</td>
<td>487</td>
</tr>
<tr>
<td>Super low floor</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>90</td>
</tr>
<tr>
<td>Midibus</td>
<td>0</td>
<td>0</td>
<td>211</td>
<td>277</td>
<td>546</td>
<td>529</td>
<td>830</td>
</tr>
<tr>
<td>Supermini</td>
<td>1711</td>
<td>1341</td>
<td>1194</td>
<td>922</td>
<td>589</td>
<td>990</td>
<td>941</td>
</tr>
<tr>
<td>Minibus</td>
<td>975</td>
<td>575</td>
<td>377</td>
<td>271</td>
<td>296</td>
<td>205</td>
<td>294</td>
</tr>
<tr>
<td>Total</td>
<td>3557</td>
<td>3319</td>
<td>2932</td>
<td>2177</td>
<td>1976</td>
<td>2509</td>
<td>2820</td>
</tr>
</tbody>
</table>


Note: Typical number of seats for each vehicle type is 74, 50, 47, 36, 27 and 19 respectively.

5.4.6 Market Entry and Contestability

The 1985 Transport Act requires operators of local bus services to obtain an Operators Licence (or "O" Licence). This licence is the main qualitative entry constraint to the bus service industry, and was issued subject to three conditions. First, applicants must be of good repute; second, applicants must have appropriate financial standing i.e. capital and reserves of the lower of 3000 ECUs (approx. £2,450) per vehicle or 150 ECUs (approx. £122) per passenger seat; third, applicants must be "professionally competent".

The capitalisation requirements are very modest and may have encouraged too high an entry level from operators who have little chance of long term viability. The capitalisation of the larger groups is much greater than the minimum figures above e.g. Stagecoach had a capitalisation of around £12,500 per bus (Transport Committee, First Report, Vol. III, p. 286). The total number of licensed operators increased continuously from 6,415 in 1989/90 to 7,193 in 1993/94, an increase of 12%. Only around 1% of new applicants for "O" Licences were refused, and only about 0.5% of existing "O" Licences were suspended or revoked. Thus the "O" Licence did not appear to be a significant barrier to market entry.

New entrants to the market are generally attracted to local authority tendered operations, where their income is usually guaranteed, and to high frequency urban routes where
"cream-skimming" can take place. In a survey of 76 new service registrations in four different Traffic Areas, these two types of entry formed 47% and 43% respectively of the total entries. In general, new entrants did not offer new routes or higher quality services (Transport Committee, 1995, First Report, Vol. III, p. 287-8).

The continuous increase in "O" Licensed operators over the period 1989-94 supports the view that the local bus service market is contestable. The attractiveness of entry for new and small businesses in particular is enhanced by a very fast cash flow from customer to producer. Service frequency has increased on many existing routes, although new entrants do not generally appear to develop new routes. In fact, the low capitalisation requirements may be attracting a number of low quality operators with irregular time-keeping and low quality vehicles and facilities. A 1994 survey of customer perceptions found that 69% of bus passengers had a more frequent service compared to ten years previously, although 86% thought that bus services could be improved. The most important non-frequency service characteristic desired by passengers was good time-keeping (Transport Committee, 1995, First Report, Vol. III, p. 288). On the other hand, vehicle maintenance standards do not appear to be falling. There was a 40% increase in vehicle inspections between 1989/90 and 1993/94, and the proportion of immediate and delayed bans on vehicles to total vehicles inspected decreased over this period.

5.4.7 Subsidy

The term subsidy here refers to public transport support from the government to the local bus service industry. It excludes concessionary fare reimbursements, fuel duty rebates and rural and new bus grants. Between 1984/85 and 1994/95, total subsidy fell in real terms by 61% in Britain (excluding London) and by 86% in London. This overall fall is somewhat misleading, however, since London has the highest subsidy per passenger in comparison to other areas (see Table 5.13):
The figures in Table 5.13 show that London bus services are heavily subsidised in comparison to other areas. These subsidies are directed into the Travelcard scheme, where 70% of the subsidy is spent on peak services, and 30% on off-peak. This allocation of subsidy is reversed outside London. Thus the low rate of decline in passenger journeys within London relative to areas outside London has been achieved at a high cost in subsidy. This is an important point, since the regulated London local bus service system, where contracts are awarded by a tendering system, is often cited as an improvement on the deregulated system outside London. Given the differences in subsidy per passenger, however, the comparison is biased in favour of London.

The problem of subsidy levels, and changes in them, is a general problem for the analysis of passenger journey trends. This problem is analysed in the next chapter.

5.4.8 Bus Travel Demand Factors

The TAS model of bus travel demand factors (TCFR, Vol. III, p. 285) estimates that an overall reduction of 7.63% passengers might have been expected over the period 1987-93. The actual loss is 17.32% over the period, a significantly higher amount than the model estimate. The variables used, together with their relative influences on bus travel demand, are shown in Table 5.14:

---

Table 5.13: Subsidy per passenger (pence, current prices) by area

<table>
<thead>
<tr>
<th></th>
<th>London</th>
<th>Mets</th>
<th>English shires</th>
<th>Scotland</th>
<th>Britain exc. London</th>
</tr>
</thead>
<tbody>
<tr>
<td>1987-88</td>
<td>8.95</td>
<td>7.39</td>
<td>4.58</td>
<td>3.40</td>
<td>5.61</td>
</tr>
<tr>
<td>1988-89</td>
<td>9.57</td>
<td>6.72</td>
<td>5.26</td>
<td>3.25</td>
<td>5.52</td>
</tr>
<tr>
<td>1989-90</td>
<td>7.15</td>
<td>6.13</td>
<td>5.76</td>
<td>1.96</td>
<td>5.46</td>
</tr>
<tr>
<td>1990-91</td>
<td>9.59</td>
<td>6.33</td>
<td>6.51</td>
<td>3.42</td>
<td>5.72</td>
</tr>
<tr>
<td>1991-92</td>
<td>13.83</td>
<td>7.37</td>
<td>7.35</td>
<td>3.50</td>
<td>6.74</td>
</tr>
<tr>
<td>1992-93</td>
<td>14.79</td>
<td>7.88</td>
<td>7.11</td>
<td>3.77</td>
<td>6.92</td>
</tr>
<tr>
<td>Increase (%)</td>
<td>65.3</td>
<td>6.6</td>
<td>55.2</td>
<td>10.9</td>
<td>23.4</td>
</tr>
</tbody>
</table>

Table 5.14: Effects of bus travel demand factors 1987-93

<table>
<thead>
<tr>
<th>Factor</th>
<th>Influence</th>
<th>Effect(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car ownership and use</td>
<td>Each 1% growth in car use reduces bus travel by 0.4%</td>
<td>-7.50</td>
</tr>
<tr>
<td>Economic activity</td>
<td>Each 1% growth in GDP increases demand for bus travel by 0.15%</td>
<td>+0.99</td>
</tr>
<tr>
<td>Employment</td>
<td>Each 1% change in number of persons employed (full time equivalents) increases/decreases demand for bus travel by 0.36%</td>
<td>-0.27</td>
</tr>
<tr>
<td>Population</td>
<td>Each 1% change in resident population increases/decreases demand for bus travel by 0.3%</td>
<td>+0.36</td>
</tr>
<tr>
<td>Service reliability</td>
<td>Each 1% of scheduled bus mileage not operated reduces demand by 0.15% and each 1% of mileage operated seriously late (&gt;10 minutes) reduces demand by 0.03%</td>
<td>-0.19 (est.)</td>
</tr>
<tr>
<td>Real fare changes</td>
<td>A price elasticity of demand of -0.3% is assumed</td>
<td>-3.06</td>
</tr>
<tr>
<td>Service level changes</td>
<td>A service elasticity of demand of +0.2 is assumed</td>
<td>+2.04</td>
</tr>
<tr>
<td>All factors: Actual:</td>
<td>Overall effect Reduction of passengers in period</td>
<td>-7.63</td>
</tr>
</tbody>
</table>


This large difference is possibly accounted for by model misspecification. One of the objectives of the next chapter is to develop a model that provides a better fit to the data. In particular, the model in Chapter six estimates a long-run own-price elasticity of around unity for bus travel.

Different (implicit) price elasticities have been estimated by Stagecoach, who argue that estimates (prior to this study) tend to be out-of-date and based on the experience of fare increases rather than decreases. Stagecoach experience of fare decreases is shown in Table 5.15:
Table 5.15: Stagecoach fare decreases

<table>
<thead>
<tr>
<th>Discount to fares</th>
<th>Increase in passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>-50%</td>
<td>+100%</td>
</tr>
<tr>
<td>-33%</td>
<td>+50%</td>
</tr>
<tr>
<td>-20%</td>
<td>+25%</td>
</tr>
<tr>
<td>-10%</td>
<td>+4%</td>
</tr>
</tbody>
</table>


Although the first two of these price elasticities seem very high by historical standards, the third value is broadly in line with the long-run results obtained from the econometric model estimated in Chapter six.

5.5 Summary and Conclusions

This chapter has focused on the short to medium term effects of local bus service deregulation. The issue of the long run welfare effects of deregulation is discussed in Chapter six, where an econometric model is specified and estimated.

The experience of the metropolitan areas in the short to medium term since deregulation was significantly different to that of the non-metropolitan areas. This is primarily because of the large reduction in revenue support for the metropolitan areas, which resulted in large fare increases and a large fall in passenger journeys. Only the independent operator sector (excluding NBC and ex-NBC operators) managed to achieve an increase in passenger journeys.

Overall, operating costs per vehicle kilometre fell significantly after deregulation. The extent of the fall, together with other evidence, provides support for the view that the deregulated local service bus industry is relatively contestable and thus competitive. The fall in operating costs was not matched by fare reductions because of the loss of revenue support. If revenue support had not been reduced so dramatically in the metropolitan areas, fare increases would not have been so large. Indeed, it is possible that fares might have fallen, and this issue is the focus of the next chapter. In spite of fears that the service
network would reduce after deregulation, neither commercial nor non-commercial vehicle mileage fell in the short to medium term. Some of this mileage may have been wasteful duplication but, in a competitive industry, this is not a sustainable strategy (see Section 4.6). Overall, the competitive tendering system for non-commercial services has worked well and delivered significant cost savings.

Profitability in the industry was very low, and fundamental adjustment in the long run was inevitable. This adjustment has taken the form of mergers. Increased market concentration was accompanied by an increase in profitability and investment, particularly for the large operators. There was a trend towards the use of smaller vehicles. The reduction in vehicle size tended to outweigh the increase in vehicle kilometres so that, overall, seat miles supplied fell, although this fall was not large.
CHAPTER 6

Subsidy and Local Bus Service Deregulation in Britain: a re-evaluation*

Bus vehicle kilometres increased after local bus service deregulation in October 1986, but passenger journeys fell and bus fares increased in real terms. The inability to reverse the long-run decline in passenger journeys and the increase in bus fares is often cited as evidence of the failure of deregulation to promote greater competition in the industry. This evaluation is not clear-cut, however, since government macroeconomic policy caused significant reductions in subsidy to the bus industry concurrent with deregulation. It can be argued that it is the reduction in subsidy, rather than the lack of competition, which caused fares to increase. If this is the case, then the evaluation of deregulation should allow for the effects of subsidy reduction. This chapter specifies and tests an econometric model in which the role of subsidy reduction is explicitly incorporated in a price–markup equation. The model can be used to generate forecasts of bus fares and passenger journeys after allowing for subsidy replacement. These forecasts are compared with those for the continuation of the regulated system. A cost-benefit analysis calculates that the net present value of the welfare gains from deregulation per se in Britain excluding London for 1986-97 is £90 million. If the external effects of bus for car substitution are included, this figure rises to £117 million or £9.7 million per annum. A similar analysis is also conducted for the metropolitan and non-metropolitan areas of Britain.

This chapter is structured as follows. An introduction to the background of the methodological problem is given in Section 6.1. Section 6.2 discusses the data, methodology and model specifications. Section 6.3 describes the estimation methods and

---

*A modified version of this chapter has been accepted for publication in the Journal of Transport Economics and Policy. The author would like to thank Professor David Starkie (Consultant), John Preston (Director of the Transport Studies Unit, University of Oxford) and two anonymous referees for their comments, as well as Mick Clary and Paul O’Hara (Department of the Environment, Transport and the Regions) for their help in providing extra data for local bus service travel.*
results, and Section 6.4 uses these results to forecast what would have happened to subsidy, bus fares and passenger journeys if deregulation had not been accompanied by subsidy reduction. Section 6.5 uses univariate time series modelling to generate forecasts for these variables on the assumption that the regulated system was allowed to continue. The welfare changes associated with the transition from the regulated system to a deregulated system (with subsidy replacement) are identified and calculated in Section 6.6. Section 6.7 concludes.

6.1 Introduction

Considerable academic debate preceded the deregulation of local bus services in Britain (excluding London) in October 1986. The arguments for and against deregulation have been discussed in the preceding chapters. Other reviews of the deregulation debate and developments since 1986 are contained, inter alia, in Mackie et al. (1995), Fairhurst and Edwards (1996), Mackie and Preston (1996), Glaister (1997) and White (1997). It is important to bear in mind that when we refer to “deregulation” there are three distinct changes that occurred: deregulation per se, which removed the quantity controls established by the Road Traffic Act of 1930; privatisation, which split up the nationalised undertakings and sold them, along with the local authority undertakings, to the private sector; and subsidy reduction, which followed from the desire of the Conservative government to reduce the public sector borrowing requirement (PSBR).

Glaister (1997, p. 146) comments that deregulation was “primarily a result of a determination to reduce overall government expenditure – both central and local authority”. It is difficult to argue against this view, given the Conservative government’s emphasis on PSBR reductions and the continuous large overruns of local authority bus subsidies relative to planned expenditures in the five annual Public Expenditure White Papers between March 1980 and March 1984. From 1972 to 1982 revenue support in Britain increased from £10 million to £529 million at current prices: even after allowing for inflation, this represented a hefty increase of 1300 per cent. Over the same period local bus service passenger journeys fell by 30 per cent. The spiralling costs of revenue
support made the local bus service sector a prime target for the privatisation policies of the incoming 1979 Conservative government and, after the 1980 Transport Act deregulated the long distance (or express coach) market, the publication of the Buses White Paper (Department of Transport, 1984) signalled the end of the regulated local service market.

The implication of these changes from the perspective of econometric modelling is that subsidy reduction can be regarded as an exogenous or “driver” variable in any model specification for evaluating deregulation. In particular, causality should run primarily from subsidy reduction to changes in bus fares, thereby avoiding any simultaneity problem that might have existed between the two variables. It also provides a benchmark for the proper specification of the model: if the model is properly specified, causality tests should confirm the direction of causality from subsidy reduction to bus fares. Failure to confirm this would indicate model misspecification.

Although privatisation is not a necessary adjunct to deregulation, it is an integral part of the proposals in the Buses White Paper. In this sense both are part of the same policy package, and henceforth this chapter uses the term “deregulation” to include privatisation. It is the subsidy reduction that causes particular problems for the evaluation of deregulation. It is conceivable that if subsidy reduction had not occurred then bus fares might have fallen and passenger journeys increased. Combining these hypothetical outcomes with the cost reductions that actually occurred might give a more favourable picture of the effects of deregulation.

Mackie and Preston (1996, p. 56) observe that “What is clear is that any review of deregulation comes up against the intractable problem of disentangling the effects of deregulation itself from those of the cuts in revenue support”. The aim of this chapter is to show that the problem is difficult rather than intractable. By specifying an appropriate econometric model it is possible to identify the influence of subsidy reduction on bus fares and, by estimating what subsidy would have been in the absence of reductions, to estimate the demand for bus travel in the “counter-factual” situation.
6.2 Data, Methodology and Model Specification

6.2.1 Data

The data set consists of forty-five annual observations for the years 1953 to 1997 on passenger journeys by local bus, bus fares, motoring costs, personal disposable income, vehicle kilometres, fleet structure, operating costs, subsidies, consumer prices and population. Since London bus services are not deregulated, the data is for Britain excluding London. This exclusion creates a (small) number of problems in obtaining the data for some variables, as does the fact that transport data tends to be collected for Britain rather than the UK whilst income and price data are collected for the UK rather than Britain. Details of the data and variable definitions are provided in Appendix four.

The effects of these differences in data collection should not be exaggerated, however. For example, inflation in Northern Ireland has been broadly similar to that in the UK, so that the UK consumer price index will be a close proxy for that in Britain. The author collected data on house prices in the South-east of England (mainly London) and the UK and used these two series as a measure of differential demand pressure in order to determine whether consumer prices might have risen faster in London than in the rest of the UK. In fact, between 1969 and 1995 house prices rose in current terms by 1259 per cent and 1224 per cent in the two areas respectively, a very similar increase. Thus changes in the UK consumer price index seem a reasonable approximation to changes in the consumer price index for Britain excluding London.

The data on subsidy to the bus industry requires particular comment. Subsidy reduces operating costs and, ceteris paribus, allows the operator to set a lower fare for bus travel. Subsidy began in 1969, when revenue support (now called public transport support) was paid to bus operators, as well as reimbursement for participation in concessionary fare
The effect of the definitional change from revenue support (RS) to public transport support (PTS) can be seen in the figures for Britain from 1977-79 inclusive, where the former is 16 per cent higher than the latter in all three years. The RS data for 1969-77 is thus adjusted upwards by the same percentage and spliced into the PTS data for 1977-97. No data is available for PTS in or outside London prior to 1977, but inspection of the data after this period shows that PTS for Britain is on average one-and-a-half times greater than that for PTS outside London for 1977-86, with a small margin of fluctuation in each of these years. This ratio is used to derive the data for PTS outside London for 1969-76. The same procedure is used to derive the data for concessionary fares (CF) outside London for 1969-76.

Operators of most local bus services also receive a fuel duty rebate from central government. This rebate is not included in the operating cost calculations, where diesel prices are measured net of fuel duty. A rural bus grant was also paid to operators between 1 April 1986 and 31 March 1990 to support rural bus services in the transitional period following deregulation. This grant is a very small amount in both absolute and relative terms, however, and is not included in the calculations.

6.2.2 Methodology

Economic theory suggests that local bus service deregulation should reduce fares and increase demand if there is an increase in competition. But between 1985/86 (the financial year immediately preceding deregulation) and 1996/97 bus fares increased in real terms by an average of 24 per cent and passenger journeys fell by 31 per cent. Although it is possible to argue that the fare increase was caused by a lack of

---

1 Concessionary fare payments are actually a subsidy to the passenger, rather than the bus operator. For convenience, however, they are included with public transport support under the general heading of subsidy to the bus industry.

2 Unless stated otherwise, data refer to Britain excluding London. It should be noted that the real fare increase in the English metropolitan areas was 58% between 1985/86 and 1996/97, much higher than the national average. This was not a steady increase since fares increased by 23% between 1985/86 and 1986/87 alone, mainly because low fares policies in some areas prior to deregulation had kept fare changes well below inflation. Thus the subsequent analysis deals not only with Britain excluding London but also the different experiences of the metropolitan and non-metropolitan areas.
competition, it is hard to see how the deregulated system created a reduction in competition sufficient to produce such a fare increase. The most plausible explanation for the increase in fares is the reduction in subsidy, which was 59 per cent in real terms over the same period. Given that subsidy forms about one-quarter of passenger and concessionary fare revenue, it does not seem unreasonable to hypothesise that, if subsidy reduction had not occurred, the increase in bus fares might have been significantly lower. Taken in conjunction with the decrease in operating costs per vehicle kilometre of one-third in real terms between 1985/86 and 1989/90, it is conceivable that bus fares might actually have fallen below those that would have occurred in the regulated system.

Figure 6.1 shows the actual and possible outcomes. The actual price/quantity combination after deregulation and subsidy reduction is given by point A. Point B shows the price/quantity combination which might be observed if the regulated system had continued in operation. Point C shows the price/quantity combination which might be observed if subsidy reduction had not occurred i.e. deregulation with subsidy replacement. Given the demand curve ABC, the triangle BCD can then be interpreted as the welfare gain from deregulation with subsidy replacement (see Section 6.6).

Figure 6.1: Bus fares with subsidy reduction, subsidy replacement and regulation
6.2.3 Model Specification and Estimation Methodology

In order to find points A, B and C it is necessary to estimate the demand function for bus travel for 1953-97 after allowing for changes in bus fares and other explanatory variables. The demand curve ABC can be interpreted as the long-run demand curve for bus travel, showing price/quantity combinations after adjusting for shifts in the short run demand curves. The bus fare is assumed to be cost-determined. The initial model specification is as follows:

\[ D_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 P_{b_t} + \alpha_3 P_{c_t} + \alpha_4 V_t + \alpha_5 F_t + \alpha_6 DU_t + u_t \]  
(1)

\[ P_{b_t} = \beta_0 + \beta_1 P_{s_t} + \beta_2 C_{F_t} + \beta_3 O_{C_t} + \beta_4 F_t + \beta_5 DU_t + u_2 \]  
(2)

where

- \( D \) = demand for bus travel (passenger journeys per person)
- \( Y \) = personal disposable income per person (real)
- \( P_{b} \) = bus fare index (1953 = 100, real)
- \( P_{c} \) = motoring cost index (1953 = 100, real)
- \( V \) = service frequency (vehicle kilometres per person)
- \( F \) = bus fleet structure (ratio of double-deckers to total bus fleet)
- \( O_{C} \) = bus operating costs (pence per vehicle kilometre, real)
- \( P_{S} \) = public transport support (pence per vehicle kilometre, real)
- \( C_{F} \) = concessionary fare revenue (pence per vehicle kilometre, real)
- \( DU \) = bus deregulation dummy variable (\( DU = 0 \), 1953-86; \( DU = 1 \), 1987-97)
- \( u_i \) = a normally and independently distributed error term with zero mean and constant variance \( \sigma^2 \) for \( i = 1,2 \).

Equation (1) is the demand function for local bus service travel, which is hypothesised to depend on own-price, cross-price and income variables. Bus travel is also influenced by service frequency and other variables such as comfort. In the absence of reliable and consistent data on factors such as comfort, the model specification relies on vehicle kilometres travelled to proxy service frequency. A difficulty with this measure is that it may overstate service frequency if network size is increasing. This effect is not tested for. Another difficulty with using vehicle kilometres as a measure of output is that it does not take account of changes in the structure of the bus fleet. Since deregulation there has been a significant degree of substitution of single-decker buses for double-deckers, so that the
number of vehicle seat kilometres has not increased at the same rate as vehicle kilometres. Using vehicle kilometres as a measure of supply may thus overstate the quality of service being offered and thus the number of passenger journeys. On the other hand, the trend towards smaller vehicles may have increased passenger journeys if, for example, a single-decker such as a minibus can access housing estates that a double-decker cannot. Thus the effects of changes in fleet structure on passenger journeys are somewhat ambiguous. An attempt is made to determine the significance of this effect by incorporating the variable $F$ into the model specification.

Equation (2) is a price-markup equation. Following the seminal paper by Hall and Hitch (1939), a large amount of empirical evidence has accrued to suggest that many firms set prices on the basis of a markup on costs, or cost-plus pricing. More recently Hall et al. (1996) in a major survey of the pricing policies of 654 UK companies confirmed this view, finding that over 47 per cent of them regarded cost-plus pricing as “important”. The $PS$ and $CF$ variables are included as (negative) costs, whilst $F$ is included in order to control for changes to the structure of the bus fleet.

The advantage of the price-markup specification is that it is general enough to cover bus operator pricing policies before and after deregulation. In the pre-deregulation period covered by this study (1953-1985), the bus industry was largely a nationalised undertaking. Profit maximisation was not considered an appropriate objective, although the industry was generally expected to behave commercially i.e. total revenue should cover total cost subject to achieving a target rate of return on capital employed. Nash (1978) discusses possible management objectives given this overall requirement.

Setting price as a markup on costs is one way of achieving this target. But such pricing is not incompatible with the more market oriented environment resulting from deregulation, since the profit markup can be adjusted in line with changes in market conditions (Hawkins, 1973). Indeed, some authors argue that cost-plus pricing is a rule of thumb for setting a profit-maximising price. Thus Equation (2) can be regarded as a reasonably
flexible specification which encompasses price setting behaviour in both the pre- and post-deregulation periods.

Both equations contain a dummy variable to allow for the effects of bus deregulation which are not explicitly modelled. These effects include instability, loss of information and lack of integration in the demand equation and changes in the markup in the price-markup equation. Other dummy variables are included in at least one of the equations, although they are not reported above in order to keep the specifications concise. These dummies account for events such as the oil price shocks of 1973/74, low fares policies (and their termination) in the metropolitan areas, and the initial period of non-privatisation in Scotland. In general, these other dummy variables are found to be insignificant.

6.3 Model Estimation

6.3.1 Estimation Methodologies

Choosing an appropriate estimation methodology is not straightforward. Time series analysis has developed an extensive literature following the realisation that regressions containing trended variables may be subject to the problem of spurious relationships. Within this literature there has been a growing emphasis on system rather than single-equation estimation, and on the specification of the short-run dynamics (via the error correction mechanism) and the long-run equilibrium properties of the system (via cointegration relationships). System estimation has a number of advantages over single-equation estimation. It makes full use of the data, it can allow and test for endogeneity/exogeneity amongst the variables without implicitly imposing coefficient restrictions, and it can be used to determine the dynamics of the system.

For example, service frequency can be regarded as a supply variable and there is a possibility of simultaneity between $D$ and $V$: passenger journeys are partly determined by service frequencies, whilst service frequencies may be adjusted by operators in response
to changes in passenger journeys. There are other potential feedback mechanisms between the variables in equations (1) and (2), such as simultaneity between $D$ and $Pb$. It also seems reasonable to assume that decisions to use bus services are subject to time-lags, since travel-to-work patterns in particular are often well-established and difficult to vary at short notice. Similarly there is likely to be some degree of inertia in changing service frequencies, since such changes involve alterations to timetables, fleet schedules, driver hours and so on. This implies that appropriate specification of variable lag orders is an important part of the modelling process. For these, and other reasons, system estimation of the above model has significant advantages.

But system estimation requires a relatively large number of observations. Although the number of observations (= 45) in the data set is large in comparison to that of most time series transport studies, it is small relative to the typical data set used for system estimation. Using the now widely established Johansen maximum likelihood (JML) estimation method in small samples, for example, is likely to lead to tests of low power and a tendency to over-reject the null hypothesis of no cointegration (Reimers, 1992, Cheung and Lai, 1993). Additionally, it may be very difficult to identify some of the multiple cointegration relationships often detected by the JML method, particularly where economic theory has nothing specific to say about the coefficient values in the cointegrating vector. The approach taken in this paper is to use both single and system equation methods to determine the properties of the data, and to determine the appropriate model specification(s) by means of the model diagnostics and the plausibility of the coefficient estimates.

6.3.2 Model Estimation

The starting point for estimation is the general to specific approach advocated by Hendry and others (cf. Hendry and Richard, 1982), together with the now standard procedures for testing for cointegrating relationships. These are: first, to determine whether the variables contain stochastic and/or deterministic trends; second, to use the JML approach to determine the number of cointegrating relationships; third, to estimate the implied vector...
error correction model (VECM); fourth, to determine which variables are endogenous or (weakly) exogenous and thereby obtain a parsimonious model specification. Given the remarks in the preceding sub-section, steps two to four are also conducted within a single equation framework using the autoregressive distributed lag (ARDL) estimation method.

In terms of capturing all the possible interactions between the variables in equations (1) and (2), it would be preferable to include them all in a single vector autoregression (VAR). Given the sample size, however, the loss of degrees of freedom in response to increased VAR lag orders is prohibitive. Accordingly, the demand and price-markup equations are modelled within separate VAR systems. The equations are estimated in both linear and log-linear form, although the latter generally gives better estimation results than the linear form and tends to reduce problems of heteroscedasticity. The log-linear form also has the convenient property that the coefficients on the explanatory variables can be interpreted as constant elasticities with respect to the dependent variable. Consequently, the remainder of this paper refers to estimation procedures conducted with the variables in logarithm form.

6.3.3 Unit Root Tests

The variables from equations (1) and (2) are tested for unit roots using the augmented Dickey-Fuller (ADF) equation with a time trend:

\[ \Delta y_t = \alpha_0 + (1 - \rho)\alpha_1 T - \rho y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + u_t, \]

where \( y_t \) is the time series of interest, \( T \) is a linear deterministic time trend, \( p \) is the order of augmentation of the test, and \( u_t \) is a white noise error term. The \( p \)th order ADF test statistic is given by the t-statistic of \( \rho \). There are four possibilities: first, \( y_t \) is stationary with no time trend \( (\alpha_1 = 0, \rho = 0) \); second, \( y_t \) is stationary with a time trend i.e. trend stationary \( (\alpha_1 = 0, \rho \neq 0) \); third, \( y_t \) is non-stationary with no time trend \( (\alpha_1 = 0, \rho = 0) \);
fourth, $y_t$ is non-stationary with a time trend ($\alpha_t \neq 0, \rho = 0$). The lag order of the ADF test, up to a maximum of four, is chosen with reference to three model selection criteria: the Akaike information criterion (AIC), the Schwarz Bayesian criterion (SBC) and the Hannan-Quinn criterion (HQC). The lag order is determined with reference to the value chosen by at least two of the criteria. Where the criteria all give different values, the lowest lag order is selected. The ADF test statistic is compared with the 95 per cent critical values given in MacKinnon (1991, Table 1).

The results from the ADF test should be treated with some caution since it lacks power in finite samples where $\rho$ is close to zero. There is also a constraint on the maximum size of the lag order of the ADF test, since there is a size-power trade-off. Rather than increasing the size of the lag-order to eliminate serial correlation in the regression residuals, the Phillips-Perron (PP) (1988) test can be used to carry out a non-parametric correction to the ADF statistic. Thus the PP test is particularly useful when it is difficult to eliminate serial correlation from the ADF regression residuals. Another problem is the presence of a structural break in the $y_t$ series. Perron (1989) shows that there is a tendency to over-reject the null of non-stationarity when the series contains a structural break, and provides adjusted critical values for the ADF test in this situation. The Perron break (PB) test used in this paper is the innovational outlier test (Perron, op. cit., p. 1380) which allows for changes in both the intercept and the slope of the ADF regression:

$$
\Delta y_t = \alpha_0 + (1 - \rho) \alpha_1 T - \rho y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \beta_1 D1 + \beta_2 D2 + \beta_3 D3 + u_t
$$

where $D1 = 1$ if $T \cdot TB$ (zero otherwise) and $TB$ is the time of the structural break, $D2 = T$ if $T \cdot TB$ (zero otherwise) and $D3 = 1$ if $T = TB - 1$ (zero otherwise). This general specification nests the "unit root with a break" null hypothesis ($\rho = \alpha_t = \beta_2 = 0$) with the "trend-stationary around a breaking trend" alternative ($\rho, \alpha_1, \beta_1, \beta_2 \neq 0$ and $\beta_3 = 0$). The t-statistic of $\rho$ is compared with the critical values given in Perron (op. cit., Table VI.B,
If the unit root null is rejected, then hypothesis tests on the remaining coefficients can be carried out using the conventional t and F tests.

The series are first tested for stochastic and/or deterministic trends using the ADF and PP tests. In a number of cases the tests give the contradictory result that the series are non-stationary with and without a deterministic trend. This is an indication that the series may contain a structural break. The PB test (along with the PP equivalent) is then used to test the unit root null hypothesis against the trend-stationary alternative in the presence of a structural break. Although the date of the structural break can be determined using a recursive, rolling or sequential approach (Banerjee, Lumsdaine and Stock, 1992), the finite-sample properties of these tests are uncertain. The determination of the break date is also a purely statistical process, with no reference to actual economic shocks. The approach of this paper is to determine the break date visually, or if no break date is visually apparent, in the light of the economy's history. The series may be subject to one or more possible breaks (oil price shocks and bus deregulation, for example). If this is the case, each break is tested individually.

The results from these test procedures can be summarised as follows. All the series are non-stationary in levels form, either with or without a deterministic trend. The non-stationarity result is consistent with a visual inspection of the series plots, where all the plots display evidence of trending. The first differences of the series are all stationary. In the case of \( \ln D, \ln OC \) and \( \ln F \) the first differences are stationary with a deterministic trend. For the other series, the first differences are stationary without a deterministic trend. All the series are thus integrated of order one, or \( I(1) \). The presence of deterministic trends in some of the first differenced series has important implications for the cointegration analysis in subsequent sections.
6.3.4 Model Estimation: the demand equation

6.3.4.1 System Estimation: Johansen maximum likelihood

The variables ln \( D \), ln \( Y \), ln \( P_b \), ln \( P_c \), ln \( V \), ln \( F \) from equation (1), plus the dummy variable and an intercept, are first estimated in terms of an unrestricted VAR model to determine the appropriate lag order. Setting the initial lag order to five, the maximum allowable given the number of variables and observations, eliminates serial correlation from each of the VAR equations. The lag order chosen by the SBC and AIC is 5, which is too high for practical estimation purposes given the small sample size. The next lag order chosen by the SBC is 1. Using this lag order for the unrestricted VAR gives no serial correlation in five of the six equations. The equation for ln \( F \) shows serial correlation at the 5 per cent level, but not at the ten per cent level. Given the lack of serial correlation, this VAR can be used as the basis for cointegration testing.

Using the JML test for cointegration, the maximal eigenvalue and trace tests find two cointegrating vectors, as does the SBC. In view of the deterministic trend found in ln \( D \), a time trend (= \( T \)) is included in the cointegrating regressions. The first vector appears to be the demand function given by equation (1), with ln \( Y \), ln \( P_b \), ln \( V \) and \( T \) significant at the 5 per cent level, but ln \( P_c \) and ln \( F \) insignificant.\(^3\) Various sets of restrictions are imposed to identify both vectors, but identification of the second vector is problematic. In addition, the underlying VECMs for the various restrictions give poor results in terms of \( \bar{R}^2 \) and coefficient significance.

One reason for the poor performance may be that the VECM’s are not conditioned on the (weakly) exogenous variables. Inspection of the VECM for each variable shows that neither error correction term is significant in the VECMs for ln \( Y \) and ln \( P_c \), so that they can be regarded as the weakly exogenous variables in the system. Conditioning on these variables, however, produces poor results in terms of \( \bar{R}^2 \) and coefficient significance.

\(^3\) The dummy variable is not normally included as part of the cointegrating regression(s) in the JML approach, but as part of the VECM. The justification is that dummy variables capture the short-run “shocks” to the system, but this is a somewhat restrictive use of dummy variables. In the present case, it is...
variables (i.e. adding them to the VECMs for the endogenous variables) does not noticeably improve the performance of the endogenous variable VECMs, however.

6.3.4.2 Single Equation Estimation

In view of the inconclusive results from the JML cointegration tests, another possibility is to use single equation estimation in the form of the ARDL estimation method. This is preferred to the Engle-Granger two-stage (EG2S) method, since the latter may be subject to small-sample bias (Wickens and Breusch, 1988). The appropriate lag order can be determined by testing down from the general to the specific model specification or, more conveniently, using the AIC, SBC and HQC. The latter method gives the following ARDL(1,0,0,0,1) model:

\[
\begin{align*}
\ln D_t &= 1.414 + 0.630 \ln D_{t-1} + 0.226 \ln Y_t - 0.380 \ln P_h_t + 0.166 \ln c_t \\
&+ 0.526 \ln V_t - 0.417 \ln V_{t-1} - 0.051 D/7 - 0.009 T \\
&+ 0.526 \ln V_t - 0.417 \ln V_{t-1} - 0.051 D/7 - 0.009 T
\end{align*}
\]

\(t\)-statistics: (1.567)(7.069)(2.224)(-4.226)(2.489)

\(R^2 = 0.999 \chi^2_{sc}(1) = 0.126 \chi^2_{fl}(1) = 0.044 \chi^2_{nn}(2) = 8.700 \chi^2_{he}(1) = 2.296
\]

where the \(t\)-statistics are in parentheses, and the \(\chi^2\)-statistics (with degrees of freedom in parentheses) are Lagrange multiplier tests for serial correlation, functional form, normality of residuals and heteroscedasticity respectively. This specification performs well. All the coefficients (apart from the intercept) are significant at the 5 per cent level, and they have the signs generally expected by economic theory. In the case of the income coefficient, the positive sign implies that the increase in demand for public transport

likely that deregulation has long-run as well as short-run effects, so it is included in the cointegrating regression(s).

\(^4\) The maximum lag order for the computations is set at 2 and the final lag order is determined by the SBC, which gives the most parsimonious model specification. Given that there are six variables in the ARDL model, higher lag orders cause very large increases in computing time. Further tests show that \(\ln F\) can be
induced by economic growth outweighs the inferior good aspect whereby increasing incomes cause car use to rise and bus use to fall. The $R^2$ is close to unity and there are no problems with serial correlation, functional form and heteroscedasticity. Although there is evidence of residual non-normality at the 5 per cent level, this is not the case at the 10 per cent level. The coefficients on $\ln Y$, $\ln P$, $\ln Pc$ and $\ln V$ are the short-run elasticities with respect to $\ln D_b$, and have plausible magnitudes. The dummy variable captures the effects of deregulation and subsidy withdrawal. Its negative coefficient implies that the latter effect outweighs the former. The negative coefficient on the time trend captures the decline in bus use caused by the increase in car ownership from the 1950s onwards.

Furthermore, the F-test procedure (Pesaran and Pesaran, 1997, p.304) shows that there is a significant cointegrating relationship between the five $I(1)$ variables in equation (3), and that it is only significant in the equation for which $\ln D$ is the dependent variable. This is an important result, since it implies that, apart from $\ln D$, the other $I(1)$ variables are weakly exogenous. Thus it is valid to include the current values of these exogenous (differenced) variables in the corresponding error correction model (ECM). Based on the underlying ARDL (1,0,0,0,1) model, the estimated ECM is:

$$\Delta \ln D_t = 1.414 + 0.226 \Delta \ln Y_t - 0.380 \Delta Pb_t + 0.166 \Delta Pc_t + 0.526 \Delta V_t - 0.009 \Delta T - 0.051 \Delta DU - 0.370 (CR)_{-1}$$

where $CR$ is the cointegration relationship. All the variables in the regression are now $I(0)$, so there should be no problem with spurious relationships. The regression results are the same as those for equation (3), apart from the introduction of the lagged

omitted from the model specification: its coefficient is not significant either individually or jointly, and its omission does not affect the performance of the model.

It is perhaps more precise to use the terminology of Pesaran and Pesaran (1997) and refer to them as the "long-run forcing variables".
cointegration relationship. The coefficient on the lagged CR is negative and highly significant, and implies that passenger journeys adjust in a given time period by 37 per cent towards their long run equilibrium value. The long run equilibrium relationship embodied in the cointegration relationship is estimated as:

\[
\ln D_t = 3.825 + 0.611 \ln Y_t - 1.028 \ln P_b_t + 0.450 \ln P_c_t + 0.300 \ln V_t - 0.025 T - 0.137 DU
\]

\[(5)\]

All the coefficients have the expected sign and are significant at the 5 per cent level, apart from the coefficient on \( \ln V \) which is significant at the ten per cent level. This seems another good result given that there are seven regressors specified but only 45 observations over which to estimate the respective coefficient values. The coefficients on \( \ln Y \), \( \ln P_b \), \( \ln P_c \) and \( \ln V \) represent the long-run constant elasticities of demand for bus travel with respect to income, bus fares, motoring costs and service frequency respectively. These long-run values should be compared with their short-run counterparts in equations (3) or (4). They are determined by the speed of adjustment coefficient (= 0.370) according to the formula: long-run elasticity = short-run elasticity \* (1/0.307), where the short-run service frequency elasticity is given by 0.526-0.417= 0.109. The short-run bus fare elasticity (= -0.38) is fairly close to the "accepted wisdom" value of -0.33, whilst the long-run elasticity of around (negative) unity is also in line with other studies (Button, 1993, p. 41, Goodwin, 1992).

6.3.5 Model Estimation: the price-markup equation

Estimation of the price-markup equation is more problematic than that of the demand equation for the following reasons. First, public transport support and concessionary fare payments do not begin until 1969, so that the number of observations available for
estimation falls from forty-five to twenty-nine\(^6\). This makes estimation of the five variable model in equation (2) a difficult proposition. Second, the period 1969-1984 is distinctly different from that of 1985-1997. In the first period, bus fares, public transport support, concessionary fares and operating costs rise together. From 1985 onwards, bus fares continue to increase but public transport support and operating costs start to fall. This difference is compatible with the expected effects of both deregulation and the reduction in public transport support in the second period, but from an econometric viewpoint it is difficult to satisfactorily estimate a model which explains the data over the period 1969-1997, even after the inclusion of dummy variables allowing for these effects. The non-linearity of public transport support and operating costs is taken into account by including quadratic terms in some of the model specifications, but this also fails to give satisfactory results. Finally, the estimation difficulties are compounded by a multicollinearity problem, since operating costs and public transport support move fairly closely in line with each other over the period 1969-1997.

In order to resolve these problems, the following modelling strategy is adopted. Public transport support and concessionary fare payments are combined into a single subsidy variable (= \(SUB\)), and the relationship between \(Pb\), \(SUB\) and \(OC\) is estimated for the period 1985-97. Since there is a multicollinearity problem between \(SUB\) and \(OC\), the latter variable is dropped from the model specification. The final specification is a simple log-linear bivariate model which, using OLS with a second order autoregressive error specification to remove serial correlation, yields the following results:

\[
\ln Pb = 5.406 - 0.252 \ln SUB \\
(117.279) (-4.334)
\]

\[R^2 = 0.878, \quad DW = 1.922\]

\(^6\) A possible way to overcome the shortage of observations is to assign zero values to public transport support and concessionary fares from 1953 to 1968, and to use the whole of the sample from 1953 to 1997 for estimation of the price-markup equation. The results of this strategy are discussed in Appendix five.
The estimated equation has a high $R^2$ and the $SUB$ coefficient is significant and has the expected sign. It also has a plausible value, given that subsidy per vehicle kilometre is around a third of operating costs per vehicle kilometre for the 1980s and 1990s. If price is determined mainly by subsidy and operating costs, then a bus fare elasticity of 25 per cent with respect to subsidy is not unreasonable. The estimated residuals all lie within the two standard error band, and most of them fall within just one standard error band. Inspection of the autocorrelation function for the estimated residuals confirms they are stationary, although this conclusion must be tempered with caution given the small sample size.

The small sample size also implies that it is not advisable to test for causality between the variables in equation (6). It is possible to make a qualitative evaluation, however, since the estimation results for equation (6) for the period 1969-1984 give a positive $SUB$ coefficient. This is consistent with the view that in the pre-deregulation period it is bus fares and public transport support which rise together, caused by the shift from bus to car use and the resulting loss of revenue for bus operators. The negative $SUB$ coefficient in the 1985-97 period implies that it is the decrease in public transport support that is causing the increase in bus fares. The issue of causality is discussed further in Appendix five, where Granger causality tests strongly support the hypothesis of uni-directional causality from subsidy to bus fares.

Equation (6) provides a baseline estimate for the whole of Britain excluding London. The next step is to disaggregate the data between the English metropolitan areas and the rest of Britain excluding London (or the "non-metropolitan" areas), since there are significant differences in the amount of subsidy withdrawal and fare increases between these two areas. The same modelling strategy is adopted as for equation (6), and yields the following results:

Metropolitan areas: $\ln Pb = 5.548 - 0.519 \ln SUB + 0.150 DU$

$$\begin{align*}
(147.757) & \quad (-16.813) & \quad (4.883)
\end{align*}$$

Non-metropolitan areas: $\ln Pb = 4.718 - 0.159 \ln SUB$

$$\begin{align*}
(234.091) & \quad (-3.555)
\end{align*}$$
Both equations have a very high $R^2$ equal to 0.962 and 0.886 respectively. The coefficients are significant at the 5 per cent level and have the expected sign. Equation (7) is estimated by OLS, with the dummy variable ($DU = 1$ for 1986, zero elsewhere) included to account for a single residual lying just outside the two standard error band. Equation (8) is estimated by OLS with a third order moving average error specification to remove serial correlation. All the residuals in (7) and (8) fall within the two standard error band, and the corresponding ACFs indicate stationarity of the residuals. Given the large reduction in public transport support in the metropolitan relative to the non-metropolitan areas, the higher (absolute) value of the metropolitan $SUB$ coefficient compared to the non-metropolitan $SUB$ coefficient is what one would expect. These coefficient values are also broadly consistent with the value in equation (6), given the relative weightings between the two areas. 

6.3.6 Model Estimation: a summary

Cointegration single-equation modelling for the demand equation yields a set of coefficient estimates that are plausible and statistically significant for both the short-run disequilibrium and long-run equilibrium processes. Similar procedures for the price-markup equation are not possible given the smaller number of observations and the nature of the data. Even in the latter case, however, it is possible to derive a plausible relationship between subsidy and bus fares for the deregulation period. Taken overall, these results suggest that it is sensible to proceed to the next stage of the analysis.

6.4 Replacing the "Lost" Subsidy.

This section provides an answer to the crucial question "what would have happened to bus fares and patronage if subsidy had not been reduced?" To do this requires some estimate of the level of subsidy that would have occurred in the absence of subsidy.

---

7 Although the price-markup equation (2) is disaggregated between the metropolitan and non-metropolitan areas, the demand equation (1) is not. The justification for this is that consumer behaviour with respect to bus travel is assumed to be unaffected by geographical area e.g. the own-price elasticity with respect to bus travel will be the same in the metropolitan and non-metropolitan areas.
reduction. One possibility is to use estimates made by the author. This runs into the legitimate objection that the author may be tempted to use an estimate which gives "good" results. Clearly, it would be better to have an independent estimate to preclude this possibility. Fortunately, such an estimate exists. In a response to the Transport Select Committee in 1995, Peter Bottomley stated that if pre-deregulation increases in subsidy (i.e. public transport support) had continued then the subsidy bill would have risen to £2 billion by that date (Hibbs and Bradley, 1997, p. 3)\(^8\). Inspection of the data shows that subsidy reduction began in 1985, so that it is necessary to estimate an equation which provides a good fit to the subsidy data from 1969 to 1984 and generates a forecast of £2 billion by 1995.

After some experimentation, the specification which provides the best match to these criteria is given by maximum likelihood estimation of an AR(4) model with \(PS\) as the dependent variable, combined with an AR(3) error specification. The model has an \(\bar{R}^2 = 0.953\) for 1969 to 1984, and generates a forecast of £1.97 billion for 1995 which is very close to that made for the Transport Select Committee. This forecast compares with the actual subsidy figure of £388 million for 1995, so that the forecast for Britain for subsidy in the absence of subsidy reduction (\(PSGBF\)) is around five times the actual figure for subsidy (\(PSGB\)). Since the estimate is for Britain, rather than Britain excluding London, it is assumed that the forecast values for public transport subsidy in London from 1985 onwards (\(PSLTF\)) increase in the same ratio as that for Britain i.e. \(PSLTF = (PSGBF/PSGB) \times PSL\). Thus the forecast values for public transport support in Britain excluding London (\(PSF\)) are derived as \(PSF = PSLTF - PSL\). The forecast values for public transport support are then added to the actual concessionary fare values to generate a forecast for total subsidy to the bus industry (\(SUBF\)).

The \(SUBF\) variable contains actual subsidy data for 1969 to 1984, and forecast values for 1985 to 1997. This variable is substituted for \(SUB\) in equation (6) to generate forecasts of \(Pb\) from 1985-97 (\(= PbF\)). The \(PbF\) variable is then substituted for \(Pb\) in equation (5) to

---

\(^8\) From other information given by Hibbs and Bradley (op. cit.) it is clear that the subsidy forecast is for public transport support in Britain in current terms.
generate forecasts of $D$ from 1985-97 ($= DF$). Thus the variables $SUBF$, $PbF$ and $DF$ represent the values which would have occurred in the deregulation period if subsidy had not been reduced. The actual and forecast values are shown in Table 6.1. Note that $PbF_R$ and $DF_R$ refer to forecast values for the continuation of the regulated local bus service system, and their derivation is explained in the next section.

Table 6.1: Deregulation (with and without subsidy reduction) and regulation: actual and forecast values for Britain excluding London

<table>
<thead>
<tr>
<th>Subsidy</th>
<th>Bus Fares</th>
<th>Passenger Journeys</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SUB$</td>
<td>$SUBF$</td>
<td>$Pb$</td>
</tr>
<tr>
<td>1985</td>
<td>3.41</td>
<td>4.18</td>
</tr>
<tr>
<td>1986</td>
<td>2.92</td>
<td>4.46</td>
</tr>
<tr>
<td>1987</td>
<td>2.29</td>
<td>4.21</td>
</tr>
<tr>
<td>1988</td>
<td>2.15</td>
<td>4.18</td>
</tr>
<tr>
<td>1989</td>
<td>2.03</td>
<td>4.42</td>
</tr>
<tr>
<td>1990</td>
<td>1.97</td>
<td>4.15</td>
</tr>
<tr>
<td>1991</td>
<td>1.92</td>
<td>3.93</td>
</tr>
<tr>
<td>1992</td>
<td>1.85</td>
<td>4.02</td>
</tr>
<tr>
<td>1993</td>
<td>1.74</td>
<td>4.09</td>
</tr>
<tr>
<td>1994</td>
<td>1.70</td>
<td>4.21</td>
</tr>
<tr>
<td>1995</td>
<td>1.68</td>
<td>4.42</td>
</tr>
<tr>
<td>1996</td>
<td>1.64</td>
<td>4.56</td>
</tr>
<tr>
<td>1997</td>
<td>1.60</td>
<td>4.86</td>
</tr>
</tbody>
</table>

Notes: Data is presented in calendar rather than financial years for convenience. Thus 1985 data is for the financial year 1985/86. The forecasts begin in 1985 because this is when subsidy reduction starts.

$SUB = $ public transport support plus concessionary fare payments, pence per vehicle km., real.
$Pb = $ local bus service fare index, 1953 = 100, real.
$D = $ local bus service passenger journeys per person.
$SUBF = $ forecast of $SUB$ with deregulation and subsidy replacement
$PbF = $ forecast of $Pb$ with deregulation and subsidy replacement (similarly for $DF$)
$PbF_R = $ forecast of $PbF$ with continuing regulation (similarly for $DF_R$)

Table 6.1 shows that by 1997 forecast subsidy is around 200 per cent higher than actual subsidy. Given the model parameters in equations (5) and (6), this implies that forecast bus fares are around 30 per cent lower, and forecast passenger journeys are around 30 per cent higher, than their respective 1997 actual values. In fact, forecast passenger journeys stabilise at around 80 passenger journeys per person, in contrast to the continuing decline in actual passenger journeys.
These aggregate figures do not distinguish between the metropolitan and the non-metropolitan areas. Equations (7) and (8) are used to forecast bus fares in the two areas respectively, and these values are then used in equation (5) to generate forecasts of passenger journeys in both areas. The results are shown in Tables 6.2 and 6.3. Forecast passenger journeys for the metropolitan areas decline sharply at first, although from a higher level than the actual values, and then begin to increase and stabilise at around 160 passenger journeys per person. The sharp decline is consistent with the “instability” effects identified by White (1990) and Fairhurst and Edwards (1996) in the metropolitan areas for the period 1985/86 to 1988/89. By 1997, the forecast value for metropolitan area passenger journeys is 47 per cent higher than the actual value for that year. In the non-metropolitan areas, the forecast values for passenger journeys decline and then stabilise at a value of around 57 passenger journeys per person which, for 1997, is 25 per cent higher than the actual value.

These results imply that, if subsidy reduction had not occurred, deregulation would have halted the long-term decline in bus patronage in both the metropolitan and non-metropolitan areas. It would not, however, have reversed the trend, since the 1997 forecast values for passenger journeys are still below the actual values for the first few years of deregulation. Nonetheless, the results do appear to vindicate the White Paper’s claim that deregulation would end the “spiral of decline”.

---

9 The values obtained for the metropolitan areas from equation (5) must be weighted by the proportion of total metropolitan population to total British population excluding London, and by analogy for the non-metropolitan areas in Britain excluding London. These weighted values are the ones shown in Tables 6.2 and 6.3. Given that the population weights are around 0.24 and 0.76 respectively, it can be checked that the sum of the forecast passenger journey values in Tables 2 and 3 are close to those given for Britain excluding London in Table 6.1.
Table 6.2: Deregulation (with and without subsidy reduction): actual and forecast values for the English metropolitan areas

<table>
<thead>
<tr>
<th>Year</th>
<th>Subsidy</th>
<th>Bus Fares</th>
<th>Passenger Journeys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUB</td>
<td>SUBF</td>
<td>Pb</td>
</tr>
<tr>
<td>1985</td>
<td>6.21</td>
<td>7.72</td>
<td>100</td>
</tr>
<tr>
<td>1986</td>
<td>5.56</td>
<td>8.68</td>
<td>122.38</td>
</tr>
<tr>
<td>1987</td>
<td>4.14</td>
<td>7.74</td>
<td>127.61</td>
</tr>
<tr>
<td>1988</td>
<td>3.74</td>
<td>7.23</td>
<td>129.38</td>
</tr>
<tr>
<td>1989</td>
<td>3.46</td>
<td>7.20</td>
<td>129.96</td>
</tr>
<tr>
<td>1990</td>
<td>3.33</td>
<td>6.69</td>
<td>131.91</td>
</tr>
<tr>
<td>1991</td>
<td>3.21</td>
<td>6.23</td>
<td>139.22</td>
</tr>
<tr>
<td>1992</td>
<td>3.05</td>
<td>6.33</td>
<td>143.41</td>
</tr>
<tr>
<td>1993</td>
<td>2.86</td>
<td>6.37</td>
<td>148.04</td>
</tr>
<tr>
<td>1994</td>
<td>2.75</td>
<td>6.43</td>
<td>149.99</td>
</tr>
<tr>
<td>1995</td>
<td>2.76</td>
<td>6.83</td>
<td>152.70</td>
</tr>
<tr>
<td>1996</td>
<td>2.73</td>
<td>7.34</td>
<td>157.05</td>
</tr>
<tr>
<td>1997</td>
<td>2.56</td>
<td>7.26</td>
<td>161.38</td>
</tr>
</tbody>
</table>

Notes: See Table 6.1 for variable definitions. There are no bus fare indices for the metropolitan areas prior to 1975. The metropolitan area bus fare index is rebased to 1985 = 100.

Table 6.3: Deregulation (with and without subsidy reduction): actual and forecast values for the non-metropolitan areas

<table>
<thead>
<tr>
<th>Year</th>
<th>Subsidy</th>
<th>Bus Fares</th>
<th>Passenger Journeys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUB</td>
<td>SUBF</td>
<td>Pb</td>
</tr>
<tr>
<td>1985</td>
<td>2.11</td>
<td>2.53</td>
<td>100</td>
</tr>
<tr>
<td>1986</td>
<td>1.81</td>
<td>2.69</td>
<td>101.30</td>
</tr>
<tr>
<td>1987</td>
<td>1.49</td>
<td>2.70</td>
<td>101.92</td>
</tr>
<tr>
<td>1988</td>
<td>1.48</td>
<td>2.87</td>
<td>101.86</td>
</tr>
<tr>
<td>1989</td>
<td>1.41</td>
<td>3.20</td>
<td>101.66</td>
</tr>
<tr>
<td>1990</td>
<td>1.37</td>
<td>3.04</td>
<td>101.86</td>
</tr>
<tr>
<td>1991</td>
<td>1.36</td>
<td>2.90</td>
<td>103.61</td>
</tr>
<tr>
<td>1992</td>
<td>1.31</td>
<td>2.98</td>
<td>104.83</td>
</tr>
<tr>
<td>1993</td>
<td>1.24</td>
<td>3.07</td>
<td>106.48</td>
</tr>
<tr>
<td>1994</td>
<td>1.22</td>
<td>3.19</td>
<td>108.77</td>
</tr>
<tr>
<td>1995</td>
<td>1.21</td>
<td>3.36</td>
<td>109.51</td>
</tr>
<tr>
<td>1996</td>
<td>1.17</td>
<td>3.45</td>
<td>112.61</td>
</tr>
<tr>
<td>1997</td>
<td>1.16</td>
<td>3.79</td>
<td>115.48</td>
</tr>
</tbody>
</table>

Notes: See Table 6.1 for variable definitions. There are no bus fare indices for the non-metropolitan areas prior to 1975. The non-metropolitan bus fare index is rebased to 1985 = 100.
6.5 Forecasts for Regulated Local Bus Services

So far the forecasts have focused on how deregulated local bus services would have responded if subsidy had not been reduced. It is also of interest to make some estimate of what would have happened to passenger journeys and bus fares in the absence of both subsidy reduction and deregulation or, equivalently, the continuation of the regulated system. One strategy is to build a structural model analogous to equation (5), but this would imply having to forecast vehicle kilometres as well as bus fares and subsidy levels in the regulated system. Another strategy is to note that from 1953 to 1981 there is a continuous and steady decline in the number of passenger journeys per person for Britain excluding London. If one assumes that the future would have been similar to the past, then there is no need for structural modelling since forecasts for passenger journeys could be made from their previous values. This can be achieved either by identifying the trend for passenger journeys within a deterministic framework or by allowing for stochastic variation in the series and modelling it in terms of the Box-Jenkins ARIMA methodology (Box and Jenkins, 1970).

The estimation period is from 1953 to 1981, with the four observations from 1982 to 1985 used for comparing forecasting performance. Since \( \ln D \) is first-difference stationary, the ARIMA models are specified with \( \Delta \ln D \) as the dependent variable. Sixteen ARMA models are specified for \( \Delta \ln D \) both with and without a deterministic trend, ranging from ARMA (0,0) to ARMA (3,3). Model selection is based on the AIC and SBC, together with Theil's inequality coefficient ("U") to evaluate forecasting performance. The best overall performance is given by an ARMA(0,0) without a trend. Inspection of the ACF indicates residual stationarity, and this result is confirmed by an ADF test on the residuals. This ARMA specification is then re-estimated for 1953 to 1985, and forecasts generated for 1986 to 1997 for \( D \). These forecasts for passenger journeys under the assumption of the continuation of the regulated system are shown as \( DF_R \) in Table 6.1. The results are quite similar to those for deregulation with subsidy reduction, declining gradually from 90 to 60 passenger journeys per person over the period 1986 to 1997.
Given the good performance of the zero order ARMA model, it is also possible that a simple regression of $\ln D$ on a time trend is an appropriate characterisation of the data generating process. This is indeed the case and, following the forecasting procedure in the preceding paragraph, passenger journeys per person decline from 81 to 51 over the same period. The statistical advantage of the ARMA specification, however, is that it gives more weight to the turning point associated with the interruption in the decline of passenger journeys in the 1982-85 period, which is why its forecasts are higher than those of the simple time trend regression for 1986-97. Another advantage of the ARMA specification results is that they imply a rather simple and intuitively appealing policy outcome: if the actual passenger journeys associated with deregulation and subsidy reduction are slightly below those forecast for the regulated system, then overall the positive effects on bus patronage from deregulation must be slightly outweighed by the negative effects of subsidy reduction\(^\text{10}\). In a nutshell, deregulation and subsidy reduction have tended in practice to cancel each other out as far as bus patronage is concerned.

Apart from a "spike" in 1976, there is a fairly continuous and steady increase in bus fares over the period 1953-85 so that $\ln Pb$ can be forecast in a similar way to $\ln D$. A total of thirty-two ARMA models are evaluated according to the AIC, SBC and Theil's U criterion, together with a check for residual stationarity. The selected model for $\Delta\ln Pb$ is an ARMA(1,0) specification without a time trend. The forecasts for bus fares under the assumption of the continuation of the regulated system are shown as $Pb_{FR}$ in Table 6.1. These forecasts are slightly lower than the actual values ($= Pb$), a result consistent with the forecasts for regulated passenger journeys ($= DF_{FR}$) being slightly higher than actual passenger journeys ($= D$).

\(^\text{10}\) This conclusion is reinforced by the negative sign and relatively small coefficient value of the dummy variable in equations (4) to (5), where $DU$ captures the effects of both deregulation and subsidy withdrawal.
6.6 A Welfare Evaluation

6.6.1 Methodology

Table 6.1 contains forecasts for the continuation of the regulated system \((PhF_R, DFR_R)\) and for the deregulated system with subsidy replacement \((PhF, DF)\). Referring back to Figure 6.1, these price/quantity combinations are associated with points B and C respectively. The movement from B to C occurs because of the change from the regulated system, characterised by territorial monopolies, to a deregulated system where there is greater competition and a reduction in bus fares (after subsidy replacement). The increase in competition is not through actual competition, which has been fairly limited in the deregulated market, but through potential competition. Chapter four concluded that the local bus service market is relatively contestable, so that the increase in industry concentration after deregulation does not imply an absence of competitive pricing.

Given certain assumptions, the standard measure of the change in welfare is given by the area BCD. The first assumption is that the long-run average cost curve is horizontal between Pb3 and C. This assumption is reasonable, given the generally accepted view that there are no significant scale economies in bus operation. Second, increased competition is likely to have resulted in a reduction in the X-inefficiency associated with the regulated bus industry. White (1990, p. 322) estimates that the increase in physical productivity was up to 20 per cent in the metropolitan areas for 1985/86 to 1988/89. This efficiency gain is permanent and applies to all years in the deregulation period. The welfare gain from reduced X-inefficiency, however, is offset by the losses associated with instability effects, reduced integration and reduced passenger information about bus services. If these gains and losses are assumed to cancel each other out, then area BCD can be used as a measure of the annual change in welfare over this period.

These conclusions are reached within the framework of the public interest theory of regulation discussed in Chapter three. The public choice approach discussed in Chapter four draws attention to the possibility that the costs of monopoly may be higher than the
deadweight loss of area BCD because of "Tullock-Posner" costs. These are the costs associated with producers expending resources to achieve and maintain regulation. If these costs are significant, then area BCD may understate the welfare gain from deregulation.

6.6.2 Britain excluding London

Area BCD is calculated for each year over the period 1986 to 1997 for Britain excluding London using the bus fare and passenger journey data in Table 1. The bus fare indices are converted to actual fares paid by using the mean bus fare figure of 32.7 pence for 1985/86 given by Mackie and Preston (1996, p. 124) in their welfare analysis. The resulting annual welfare gains represent the "internalities" associated with the movement from B to C in Figure 6.1. These gains are then discounted using a discount rate equal to the average real rate of return on UK Treasury Bills over the period (= 4.67 per cent). The net present value (NPV) for 1986 of these welfare gains from deregulation with subsidy replacement is £92.087 million, or an average of £7.674 million per annum over the period.

The movement from B to C is also associated with a shift from car to bus travel because of the fall in the price of bus relative to car use. The reduction in car use creates various cost changes that are not generally taken account of in the private decision to substitute bus for car travel. In this sense, these cost changes can be regarded as the externalities associated with the movement from B to C. Romilly (1999) considers the unpriced cost changes arising from changes in traffic congestion, air and noise pollution, global warming, fuel savings, road accidents and road damage caused by the shift from car to bus travel in urban areas. The annual overall cost reductions range from £639.5 to £1234.1 (1995 prices) for each car no longer on the road because its occupants have

11 For the non-metropolitan areas, the DF values are below those for the regulated system for 1987 and 1988, even though the PbF values are also below the regulated values. This implies either an upward sloping demand curve over the period or a shift inward of the demand curve. In both cases measurement of the welfare change is problematic. The approach taken is to set the welfare gain (or loss) equal to zero for these years.
transferred to bus travel (Romilly, *op. cit.*, p. 123). A possible external benefit not considered in Romilly (*op. cit.*) is the user benefit arising from increased bus service frequencies, since increasing service frequencies in response to increased demand will reduce waiting times for existing bus users (Mohring, 1972). To the extent that these benefits are significant, the above procedure will understate the welfare gains from bus for car substitution.

The reduction in car travel is estimated by using the results of cointegration methodology applied to car ownership and use (Romilly, Song and Liu, 1998). The cointegrating equation (9) is derived from an ARDL specification, although EG2S estimation produces similar figures for the reduction in the number of cars owned:

\[
\ln CO = -1.580 + 2.393 \ln Y - 1.300 \ln P_c + 1.926 \ln P_b - 0.0727T
\]

\[
(-0.553) (3.998) (-4.409) (3.743) (-4.050)
\]

where \(CO\) = car ownership per capita. The total reduction in cars owned is obtained by substituting \(P_bF\) and \(P_bF_R\) in turn for \(P_b\) in equation (9), and multiplying the results by the population of Britain excluding London. The total reduction in cars owned is then multiplied by the conservative external cost saving estimate of £639.5 per car per annum (rebased from 1995 to 1986) to give the total reduction in external costs for each year. These annual cost savings are discounted to a NPV for 1986 using the discount rate of 4.67 per cent. The resulting figure is £25.463 million. Combining the internal and external welfare gains from deregulation with subsidy replacement yields a NPV equal to £117.6 million, or an average of £9.796 million per annum over the period (see Table 6.4).

6.6.3 Metropolitan and Non-metropolitan Areas

The bus fare index in the regulated system is more difficult to forecast for the metropolitan areas than for Britain excluding London. It is more volatile, and there are fewer observations on which to base a forecast since bus fare data for the metropolitan areas is only available from 1975 onwards. It is possible to derive an index by noting that
the forecasts for the regulated system in Britain excluding London are similar to (in fact, slightly below) those which are actually observed for the deregulated system with subsidy reduction. It is assumed that this relationship holds for the metropolitan areas, so that an index can be derived by taking the actual metropolitan fare index for 1986 and multiplying it by the ratio of the forecasts for the regulated system fare indices to the actual fare indices for Britain excluding London.

A similar procedure is used to forecast passenger journeys per person for the regulated system in the metropolitan areas. The regulated and deregulated fare indices for the metropolitan areas are converted to actual fares by using the figures given in White (1990, Table 3) and Mackie and Preston (1996, p. 124). The reduction in cars owned for the metropolitan areas is derived from the figure for Britain excluding London multiplied by the ratio of metropolitan area population to that of Britain excluding London. Using the welfare methodology above, the NPV for the internal and external adjusted welfare gains is £47.515 million and £5.753 million respectively. Corresponding estimates of bus fares and passenger journeys are also derived for the non-metropolitan area, giving a NPV for the internal and external adjusted welfare gains in this area of £44.572 million and £19.710 million respectively. The overall results are summarised in Table 6.4:

\[ \text{NPV for the internal and external adjusted welfare gains is}\]

\[ \text{£47.515 million and £5.753 million respectively.}\]

\[ \text{The unadjusted internal welfare gain for the metropolitan areas is}\]

\[ £123.025\text{ million, but this must be weighted by the ratio of metropolitan to non-metropolitan area passenger receipts, since the actual fare for Britain excluding London is a weighted (by passenger revenue) sum of the constituent areas. The passenger revenue ratio is fairly constant at around 35 per cent, giving an adjusted internal welfare gain for the metropolitan areas of}\]

\[ £123.025\times0.35=£43.059\text{ million. This figure is then adjusted slightly upwards by an aggregation factor (see next footnote).}\]

\[ \text{The sum of the passenger revenue weighted internal welfare gains (equal to £83.45 thousand, consisting of £43.059 and £40.392 million for the metropolitan and non-metropolitan areas respectively) falls slightly short of the total for Britain excluding London (equal to £92.087 million). To maintain consistency between the aggregate and disaggregate totals, the latter are scaled up by a factor equal to 92.087/83.451.}\]
Table 6.4: Welfare gains from deregulation (with subsidy replacement) (£m)

<table>
<thead>
<tr>
<th></th>
<th>Britain excluding London</th>
<th>Metropolitan areas</th>
<th>Non-metropolitan areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal welfare gains</td>
<td>92,087</td>
<td>47,515</td>
<td>44,572</td>
</tr>
<tr>
<td>External welfare gains</td>
<td>25,463</td>
<td>5,753</td>
<td>19,710</td>
</tr>
<tr>
<td>Total welfare gains</td>
<td>117,550</td>
<td>53,268</td>
<td>64,282</td>
</tr>
</tbody>
</table>

Notes: The figures represent the welfare gains derived from comparing the forecast values for the regulated bus system with those for the deregulated system with subsidy replacement for the period 1986-97. The figures are net present values derived by discounting the welfare changes from 1986-1997 using a discount rate of 4.67 per cent.

The preceding analysis assumes that subsidy replacement is associated with an overall increase in the forecast level of subsidy from 1985 to 1997 (see values for $SUBF$ in Table 6.1). Another possibility is to assume that subsidy is frozen at its 1984 level for the period 1985 to 1997. This is similar to one of the scenarios considered by Mackie and Preston (1996, p. 123, p. 127), although their analysis does not account for externalities or the split between the metropolitan and non-metropolitan areas. Actual subsidy in 1984 was 3.54 pence per vehicle kilometre, and the bus fare index was 166.69. These values are used for the whole of the period 1985 to 1997 to generate forecasts of passenger journeys for Britain excluding London using the methodology developed above. Passenger journeys per person fall from 91 in 1985 to 81 in 1987, and stabilise at around 75 by 1997. The associated internal and external welfare gains are £63.526 million and £13.752 million respectively, giving a total welfare gain of £77.277 million or an average of £6.440 million per annum over the period 1986 to 1997. Thus the total welfare gains from the “frozen” subsidy are two-thirds those of the “replacement” subsidy case.

White (1990, Table 9) estimates the welfare changes from deregulation, excluding externalities, for the period 1985/86 to 1988/89. These are a welfare gain of £10-£40 million for the metropolitan areas, and a welfare loss of £63 million for the English shire counties and Scotland. The estimates do not allow for the effects of subsidy replacement, however. Mackie and Preston (1996, Table 6.17, Table 9.8) provide a number of estimates for Britain excluding London which do allow for the effects of subsidy replacement.

14 The author is grateful to Professor David Starkie for this suggestion.
reduction. Taking their scenario that corresponds most closely to the “frozen” subsidy case above, there is an estimated welfare gain of £7 million per annum for the period 1985/86 to 1987/88 (excluding externalities). This is not dissimilar to the author’s “frozen” subsidy internal welfare gain estimate of £63.526 million, or £5.294 million per annum for the period 1986 to 1997.

Comparison between these studies and the present paper, however, may not be appropriate. The welfare change estimates in the above studies are based on fewer years and tend to capture the short to medium-term effects of deregulation. This contrasts with the present paper, where the welfare changes are estimated over the twelve-year period 1986-97 and are thus associated with the long-run deregulation effects. The above studies estimate the welfare changes on the basis of elasticity values which are taken from external sources and may not be representative of the true values for the time period under consideration. This paper uses the long-run elasticity values estimated directly from the data set. Although the Mackie and Preston (1996) study allows for the effects of subsidy reduction, the method used is to determine the increase in fares necessary to make up for the lost revenue from subsidy reduction. This is different to the methodology of the present paper, which forecasts the reduction in fares after subsidy replacement by explicitly specifying and estimating the relationship between costs and fares. Additionally, the Mackie and Preston (op. cit.) study does not distinguish between welfare changes in the metropolitan and non-metropolitan areas, and consequently may be subject to aggregation bias.

6.7 Conclusions

Using an econometric model for 1953-97 in which the effects of subsidy reduction and replacement are determined with reference to coefficient values derived directly from the data, this chapter concludes that the discounted “internal” welfare gains from local bus service deregulation with subsidy replacement in Britain excluding London are in the region of £92.1 million, or £7.7 million per annum for the period 1986-97. If the change in externalities from bus for car substitution is also taken into account, this figure rises to
£117.6 million or £9.8 million per annum. If the forecast subsidy is assumed not to increase but to remain at its 1984 level, then the total welfare gain is reduced by one-third to £77.3 million, or £6.4 million per annum.

These values are derived on the basis of fairly cautious assumptions concerning the discount rate, external costs and the projected behaviour of the regulated system for the period 1986-97. If the welfare losses from instability effects, lack of integration and information are assumed not to outweigh the gains from increased efficiency, then these estimates can be regarded as being at the lower end of the possible values for welfare gains. These gains should not be over- emphasised, however, since they are small relative to total passenger receipts for the industry, which were around £1,340 million in 1985/86 for Britain excluding London.

Metropolitan and non-metropolitan areas gain from deregulation with subsidy replacement, although the gain per person is greater in the metropolitan areas. The total net present values of the internal welfare gains are roughly similar between the areas (£47.5 and £44.6 million respectively, or £4.0 and £3.7 million per annum). Adjusting for population size, the total net present values of the internal welfare gains for the metropolitan and non-metropolitan areas become about £4.2 and £1.2 per person respectively.

Based on the results of univariate time series estimation, the actual outcome of deregulation with subsidy reduction in terms of passenger journeys per person and bus fares is very similar to the forecasts made for the regulated system for 1986-97. In other words, the positive effects of deregulation per se on fares and passenger journeys are broadly cancelled out by the negative effects of subsidy reduction. It is possible, however, that deregulation with subsidy replacement and a system of competitive tendering would yield greater benefits than those calculated above. This possibility is not explored in this chapter, however, and represents an avenue for further research.
CHAPTER 7

Substitution of Bus for Car Travel:

external benefits and costs*

Previous chapters have tended to focus on the local bus service market itself, rather than the relationship between bus and car travel. This chapter addresses the issue of the externalities arising from bus and car travel, and the monetary changes arising from the substitution of bus for car travel following deregulation with subsidy replacement. The role of externalities as a source of market failure was discussed in Chapter three, although no attempt was made to quantify their impacts. In principle, externalities in the road passenger transport sector can be internalised through first best pricing policies such as increased taxation of motorists or, in the absence of first best pricing, through second best policies such as the subsidisation of bus travel. In practice, fiscal policies in the UK and elsewhere have not tended to allow for the internalisation of externalities, and there is a perception that the bus (and rail) sector suffers from unfair competition from underpriced car use (see Section 3.3.3.5). It is important, therefore, that any welfare evaluation of deregulation with subsidy replacement should include at least some of these externalities.

The aim of this chapter is to estimate a monetary value for some of the external benefits and costs associated with bus for car substitution. These estimates are used as part of the social welfare evaluation of local bus service deregulation in Chapter six.

Increasing car use is associated with a number of problems including traffic congestion, air and noise pollution, road accidents, loss of countryside from new roads, community severance and global warming. A switch from car to bus use is generally thought to increase social welfare because of a reduction in these costs. This view is generally correct, although there are some exceptions.

The focus of this chapter is on the air pollution and global warming costs associated with bus for car substitution. This is because more reliable data has recently become available with respect to emissions of petrol and diesel engines, and also because existing research concentrates on the overall costs of transport emissions rather than the specific costs associated with bus for car substitution. The effects of substituting bus for car travel in urban areas are simulated by specifying a spreadsheet model incorporating two types of car (petrol and diesel engine) and three types of bus (mini-, midi- and large bus). Six types of exhaust emission are considered for each vehicle type for the years 1992, 1995 and 1999: carbon monoxide, volatile organic compounds, nitrogen oxides, sulphur dioxide, (small) particulate matter and carbon dioxide. The chapter provides a synthesis of monetary estimates of these exhaust emission and other costs.

The other costs considered are traffic congestion, fuel consumption, noise pollution, road accidents and road damage. These costs are discussed within the context of a sensitivity analysis that allows for changes in parameters such as load factors, emission factors and the individual exhaust emission cost estimates.

Section 7.1 examines the environmental and health effects of petrol and diesel fuel exhaust emissions, while Section 7.2 develops a framework for the analysis of these emissions. Section 7.3 quantifies the effects on air pollution and global warming of substitution of bus for car travel. Section 5.4 reviews the literature on the monetary costs of air pollution and global warming and estimates the change in these costs resulting from a substitution of bus for car travel. A sensitivity analysis of these results is conducted in Section 7.5. Other cost changes are discussed in Section 7.6, and Section 7.7 concludes.
7.1 Environmental and Health Effects of Petrol and Diesel Exhaust Fumes

There are two main types of vehicle engine: petrol and diesel. A mixture of fuel and air is drawn into the cylinders of a petrol engine, and is then compressed and ignited by a spark. No spark is needed in a diesel engine, since fuel is injected into air already made very hot by compression. The more efficient combustion process of the diesel engine means that its fuel consumption is at least 25% less by volume and 15% less by mass than a petrol engine of the same power (Royal Commission on Environmental Pollution (RCEP), 1994). The average rate of fuel consumption for petrol and diesel cars has been estimated at 28.8 and 41.4 miles per gallon respectively (National Travel Survey, 1989/91, Table 5.3.8). Buses are generally powered by diesel engines, although some operators have introduced engines powered by alternative fuels. Cars are mainly petrol engined, although there has been an increasing trend towards the use of diesel engines because of their higher fuel efficiency and, to a lesser extent, their perceived environmental benefits. In 1986, only 4% of new private car registrations were for diesel engines: by 1994 this figure was 23%.

There are different health and environmental costs associated with the exhaust emissions from petrol and diesel engines. These emissions include: carbon dioxide (CO₂), carbon monoxide (CO), nitrogen oxides (NOₓ), particulate matter (PM), volatile organic compounds (VOC), sulphur dioxide (SO₂) and lead (Pb). A switch from car to bus travel implies that less petrol and more diesel fuel is consumed, with corresponding changes in health and environmental costs. A brief description of these costs is given below.

Carbon dioxide is the main "greenhouse gas", probably accounting for more than half the total man-made greenhouse effect. Although it has a relatively harmless direct effect on human health, its contribution to global warming is likely to cause, along with other greenhouse gases including methane, significant long term environmental effects (Intergovernmental Panel on Climate Change, 1990; Cline, 1992). Carbon monoxide is produced as a result of incomplete combustion of the carbon contained in the fuel. Diesels are more efficient in their combustion process than petrol engines, and consequently produce significantly less CO emissions than equivalent petrol engines.
Although CO soon oxidises to CO\(_2\) when released into the atmosphere, it is a highly poisonous gas that interferes with the absorption of oxygen by red blood cells. It can affect the central nervous and cardiovascular systems, and is fatal in large doses. Even at the relatively low concentrations typically found in city centres, it can cause sore throats and drowsiness.

VOCs are a complex group of carbon compounds of sufficient volatility to exist as vapour in the atmosphere. The main constituent of VOC is hydrocarbons (HC), which are formed mainly from the incomplete combustion of petrol. Methane is the lightest hydrocarbon. Other organic compounds include benzene, formaldehyde and polyaromatic hydrocarbons (PAH). HC may be emitted as a gas or in solid form as particulate matter. Diesel engines emit fewer hydrocarbons than equivalent petrol engines because of their better fuel combustion, although incomplete combustion of diesel fuel creates a wide variety of polyaromatic hydrocarbons (PAH) which may have important synergistic adverse effects on health. PAH can become absorbed onto particulate matter and carried deep into the lungs. HC contain several carcinogenic organic compounds, including benzene which is known to cause leukaemia. They also react with NO\(_x\) to form ozone (O\(_3\)) at ground level. Ground-level ozone is the main component of photochemical smog.

Particulate matter consists of small solid or liquid particles suspended in the atmosphere. They are conglomerates of carbon (soot) and other pollutants including sulphur and hydrocarbons. Vehicle use is generally the main source of PM in industrialised countries, arising from exhaust gases, wear and tear of tyres and brakes, and from road dust caused by vehicle use. High concentrations of carbon particles become visible as black smoke, a problem particularly associated with the diesel engine. The small particles can penetrate deep into the lungs, aggravating respiratory problems. Those particles less than 10 \(\mu\)m in diameter (PM\(_{10}\)) penetrate the lung fairly efficiently and are the most harmful to health. Diesel particulate emissions may also have potential carcinogenic effects. A study of US railroad workers suggests that the possibility of developing cancer may be 42% higher in individuals exposed to diesel exhaust than in those who are not (OECD, 1988). The link between PAH and cancer is
disputed, however (Grace, 1994). Diesel engines emit higher levels of PM than equivalent petrol engines.

Nitrogen oxides are created when oxygen combines with nitrogen in the air and, to a lesser extent, nitrogen in the fuel. The amounts of nitrogen oxides in exhaust gases increase as combustion temperature increases, so that they are highest at high speeds and during rapid acceleration. Some of the nitrogen is oxidised into nitric oxide (NO) and nitrogen dioxide (NO\textsubscript{2}) which together constitute NO\textsubscript{X}. More than half the NO\textsubscript{X} emissions in the UK are from road transport. NO\textsubscript{X} has significant direct adverse effects on human health in the form of respiratory problems, even when the exposure time is relatively short. It also reacts with other pollutants in the atmosphere to create indirect effects including acid rain and smog. NO\textsubscript{2} combines with water vapour in the form of rain, mist, fog or dew to produce a weak nitric acid. In conjunction with weak sulphuric acid from sulphate emissions, their depositions help to create acidity in lakes and stunt the growth of vegetation. In the presence of strong sunlight, NO\textsubscript{X} combines with other pollutants such as CO and HC to form photochemical smog. It is NO\textsubscript{2} which helps to give smog its brown colour.

Sulphur occurs naturally in crude oil, and becomes more concentrated in diesel fuel. Sulphur dioxide (SO\textsubscript{2}) emissions from the transport sector are a relatively small part of total SO\textsubscript{2} emissions, and are produced by the combustion of fuel containing sulphur. They oxidise to form sulphur trioxide (SO\textsubscript{3}) which combines with water vapour (H\textsubscript{2}O) to form sulphuric acid (H\textsubscript{2}SO\textsubscript{4}). Sulphur compounds constrict air passages and can thus cause particular problems for asthmatics. In 1952, sulphur emissions from domestic coal burning caused a smog in London which killed around 4000 people in two weeks (Parliamentary Office of Science and Technology, 1994, p. 1). Sulphur dioxide, although colourless, is an irritant with a particularly strong and unpleasant smell. As noted earlier, sulphur dioxide also reacts with nitrogen oxides to form acid rain. Both SO\textsubscript{2} and NO\textsubscript{X} emissions also contribute to secondary particulate formation: this indirect effect appears to be more important for human health than the direct effects (Small and Kazimi, 1995, p. 8). Ostro (1994, p. 23, p. 51) concludes that SO\textsubscript{2} emissions do have significant direct effects on human health, although these effects per unit of emission are much smaller than those from particulate matter and...
(tropospheric) ozone. On the other hand, Ball et al. (1991) and Kageson (1993) consider SO₂ emissions to be relatively unimportant.

Lead is used in petrol to enhance its performance, although it cannot be used in conjunction with a three-way catalytic converter (TWC). Diesel fuel is lead-free. Lead is known to have a number of adverse effects on health, including damage to the kidney and liver, the reproductive system, blood formation, bone marrow and the brain. Children are particularly affected by lead poisoning, and a number of studies suggest that it causes behavioural problems, lower IQs, and a reduced ability to concentrate in children (OECD, op. cit.). There is also a possibility that lead may have carcinogenic effects.

Substitution of unleaded for leaded petrol has occurred because of differential tax treatment, but unleaded petrol contains aromatics, particularly benzene, which are used instead of lead to maintain octane values. If a car using unleaded fuel does not have a catalytic converter, which is generally the case for older vehicles, then lead pollution is simply replaced by pollution from aromatics. Although exhaust gases produce only trace quantities of benzene, concentrations are much higher at petrol pumps because of evaporation. These concentrations enter the vehicle interior and remain for some time after refuelling. It has been estimated that "the levels [of benzene] inside cars are higher than those outside, even in the most polluted areas" (House of Commons Transport Committee, 1994, Vol. 1, p. xxix).

The amount of pollutants emitted by a given vehicle depends not only on the type of engine and the conditions under which it is driven, but also on the age of the engine and the extent to which it has been maintained properly. Newer engines comply with more stringent emissions standards. As a new engine becomes older, however, it will emit more pollutants unless it is maintained properly. Differences in vehicle age and degree of maintenance can cause significant differences in the distribution of emissions across a vehicle stock: for example, around 10 percent of goods vehicles account for some 90 percent of particulate emissions within this vehicle category. The following section describes the analytical framework used to deal with some of these issues.
7.2 Substitution of Bus for Car Travel: an analytical framework for the determination of exhaust emission changes

Increasing car ownership and use in Britain and elsewhere has contributed to a number of well-publicised problems, including traffic congestion, air and noise pollution, road accidents, community severance and loss of countryside because of road-building. These problems are acknowledged in the recent White Paper on transport (Department of the Environment, Transport and the Regions (DETR), 1998), which sets out proposals to constrain car use and promote alternative forms of transport, particularly the bus.

A switch from car to bus use is generally thought to reduce many, if not all, of these problems. But there are difficulties in trying to present a detailed case for encouraging road users to transfer from their car to a bus. First, in spite of an extensive literature on the monetary valuation of transport related costs such as air pollution, the empirical results tend to be at a highly aggregate level: it is usual to see monetary estimates for the environmental costs of all road passenger and freight travel, for example, but not for sub-categories such as bus and car travel. This makes it difficult to estimate the specific monetary effects of substitution between these modes, and even more so if one wants the substitution to be disaggregated across different types of bus and car. Second, the estimated aggregate values are often very different from each other, having been derived from different methodologies or local conditions. Third, there is some uncertainty over the emission factors for individual exhaust emissions. In short, there is a need to develop a consistent set of values that can be used in a cost-benefit analysis of bus for car substitution.

This chapter identifies six exhaust emissions and discusses their monetary costs using research results primarily from the UK and the USA. In order to place the exhaust emission costs in context, it also considers the monetary valuations of fuel savings, road congestion, noise pollution, road accidents and road damage. To check the robustness of the results to different scenarios, a sensitivity analysis is conducted within a spreadsheet simulation model called “CarBus”. An example of such a simulation is given in Appendix three.
Given the different fuel and vehicle technologies associated with bus and car travel, substitution between these modes will cause changes in the amounts of exhaust emissions. This is true not only at the aggregate level, but also at a more disaggregate level. The air pollution and climate change effects of a road user switching his or her journey from a petrol engine car with a three-way catalyst (TWC) to a large bus are different to those where the switch is from a diesel engine car to a minibus. Vehicle distribution is modelled in terms of two types of car (petrol and diesel) and three types of bus (mini-, midi- and large bus), and this distribution is allowed to change over time in response to changes in consumer preferences and exhaust emission regulations.

Substitution of bus for car travel is assumed to occur in urban areas, where around 75 per cent of bus travel by distance takes place. The focus on urban areas seems appropriate since this is where car users are most likely to switch to buses: bus service frequencies are higher than in rural areas, people are likely to be nearer a bus stop, and journey to work distances are shorter (so that the inconvenience of bus relative to car travel is less important). Buses are assumed to be powered by diesel engines although a small, but increasing, number of operators have introduced engines powered by alternative fuels including liquid petroleum gas (LPG) and compressed natural gas (CNG). Cars are mainly petrol engined, but the analytical framework allows for the increasing use of diesel engined cars.

Exhaust emissions are disaggregated into six constituents: carbon monoxide (CO), volatile organic compounds (VOCs), nitrogen oxides (NO$_X$), small particulate matter (PM$_{10}$), sulphur dioxide (SO$_2$) and carbon dioxide (CO$_2$). The analysis does not include lead exhaust emissions, which now form around 14% of their 1975 levels in the UK and are likely to fall further. A detailed description of these emissions and their effects can be found in, *inter alia*, RCEP (1994) and Maddison et al. (1996, ch. 4). Past and future trends in road transport emissions, as well as strategies for achieving reductions in air pollution, are set out in the United Kingdom National Air Quality Strategy (Department of the Environment (DOE), 1996).

The calculations are based on the scenario of a reduction in the number of cars by 100, a figure chosen for expositional simplicity. Using the 1994/96 National Travel Survey
(DETR, 1997b) estimate of 1.6 for the average car occupancy rate, there will be a reduction of 160 in the number of car users. It is assumed that all these car users substitute bus for car travel. A realistic expectation of load factors for both small and large buses is about 45 per cent (Glaister, 1985a, p.69). Given these load factors and the actual distribution of cars and buses, it is then possible to estimate the increase in the number and type of buses required to accommodate the 160 former car users, as well as the decrease in the number of petrol and diesel engine cars. In 1992, for example, the number of petrol cars without a TWC, the number with a TWC, and the number of diesel cars fall by 92, 3.5 and 4.5 respectively, whilst the corresponding increases for mini-, midi- and large buses are 3.73, 1.93 and 2.67 respectively. Full details of these calculations are contained in Appendix three. The load factors assumed for car and bus travel are critical parameters in determining the outcome of this modal substitution, and the effects of changing load factors and other parameters are examined in the sensitivity analysis of Section 7.5.

Research in the areas of exhaust emissions and their monetary costs is continuously evolving. The conclusions reached in this chapter are necessarily tentative, not only because of the uncertainties involved in the estimation of the main parameters but also because some important issues are not considered within the CarBus model. This is primarily because of the lack of data, particularly for greater levels of disaggregation, and there is a need for research into these issues. For example,

- to what extent are the impacts of exhaust emissions on mortality and morbidity rates non-linear rather than linear, as is generally assumed?
- to what extent do mortality and morbidity rates differ with age? How are these rates affected by road dust particles rather than those emitted directly from vehicle exhausts?

1 This figure is an average for urban peak period bus travel, and may not be the same as the marginal value for substitution of bus for car travel: hence the need for a sensitivity analysis of load factor assumptions. The paper assumes that, for marginal changes in the off-peak, the resulting load factor will be the same as that in the peak. The rationale is that spare capacity in the off-peak can be used to increase bus load factors up to the peak value for marginal passenger increases.

2 A small proportion of petrol cars in 1992 conformed to Stage I emission limits, which were effective from 1.7.92. This proportion was around 3.5 per cent of the total car stock in 1992. The proportion of diesel cars was about 4.5 per cent.
how are load factors and other parameters affected by journey purpose, type of road user and the geographical area under consideration?

how does the method of reducing car travel affect the consumer surpluses (or “internalities”) of bus and car users? Who are the people likely to transfer from car to bus in response to specific government policies?

how far will the degree of competitiveness in the local bus service market affect bus load factors in urban and rural areas?

where substitution occurs on a significant scale, to what extent will the reduced congestion and increased traffic speeds affect emission factors?

7.3 Bus and Car Urban Exhaust Emissions

The results in this section are derived from studies commissioned by the DETR, including those by Gover et al. (1994) and Salway et al. (1997), as well as from recent unpublished data revisions. Table 7.1 shows urban emission factors for cars and buses based on three sets of emission regulations: Pre-stage I, Stage I and Stage II (or Euro I and Euro II). Broadly speaking, Pre-stage I vehicles are those registered before 1993: in particular, petrol cars were not required to have TWCs. Stage I vehicles are those registered from 1993 to 1995 inclusive: the Stage I regulations required all new petrol cars to be fitted with TWCs, as well as tightening emission limits for all diesel engine vehicles. In general, there are substantial reductions in the emission factors of CO, VOCs, NOX and PM10 between the Pre-stage I and Stage I sets of regulations. Further reductions are achieved under Stage II from the beginning of 1996. The limit values for these emissions are agreed by European Union member states and applied throughout these states. The exceptions are for SO2 and CO2, where emissions are determined mainly by fuel consumption.

3 The author would like to thank Tim Murrells of AEA Technology and Chris Mcgaincy of DETR for their assistance in providing the revised data. This data will eventually be published in the tenth National Atmospheric Emissions Inventory produced by the National Environmental Technology Centre as part of the Air Quality Research Programme of the DETR.

4 The exact dates are dependent on whether the vehicles are diesel or petrol engined, the type of diesel engine, and whether or not the vehicle is a new model.
Table 7.1: Urban exhaust emissions (gm per km) from buses and cars in Britain

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>VOCs</th>
<th>NOx</th>
<th>SO2</th>
<th>PM10</th>
<th>CO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petrol car:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-I</td>
<td>13.599</td>
<td>2.318</td>
<td>1.788</td>
<td>0.029</td>
<td>0.036</td>
<td>174.2</td>
</tr>
<tr>
<td>I</td>
<td>3.134</td>
<td>0.296</td>
<td>0.382</td>
<td>0.031</td>
<td>0.010</td>
<td>188.5</td>
</tr>
<tr>
<td>II</td>
<td>2.195</td>
<td>0.130</td>
<td>0.169</td>
<td>0.031</td>
<td>0.010</td>
<td>188.5</td>
</tr>
<tr>
<td>Diesel car:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-I</td>
<td>0.781</td>
<td>0.187</td>
<td>0.725</td>
<td>0.111</td>
<td>0.207</td>
<td>193.6</td>
</tr>
<tr>
<td>I</td>
<td>0.472</td>
<td>0.079</td>
<td>0.523</td>
<td>0.111</td>
<td>0.057</td>
<td>193.6</td>
</tr>
<tr>
<td>II</td>
<td>0.331</td>
<td>0.055</td>
<td>0.377</td>
<td>0.111</td>
<td>0.025</td>
<td>193.6</td>
</tr>
<tr>
<td>Minibus:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-I</td>
<td>1.201</td>
<td>0.315</td>
<td>2.236</td>
<td>0.151</td>
<td>0.348</td>
<td>263.3</td>
</tr>
<tr>
<td>I</td>
<td>0.252</td>
<td>0.103</td>
<td>1.952</td>
<td>0.151</td>
<td>0.141</td>
<td>263.3</td>
</tr>
<tr>
<td>II</td>
<td>0.201</td>
<td>0.096</td>
<td>1.394</td>
<td>0.151</td>
<td>0.087</td>
<td>263.3</td>
</tr>
<tr>
<td>Midibus:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-I</td>
<td>2.360</td>
<td>0.804</td>
<td>6.354</td>
<td>0.253</td>
<td>1.652</td>
<td>441.0</td>
</tr>
<tr>
<td>I</td>
<td>0.503</td>
<td>0.263</td>
<td>5.562</td>
<td>0.253</td>
<td>0.668</td>
<td>441.0</td>
</tr>
<tr>
<td>II</td>
<td>0.402</td>
<td>0.245</td>
<td>3.973</td>
<td>0.253</td>
<td>0.412</td>
<td>441.0</td>
</tr>
<tr>
<td>Large Bus:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-I</td>
<td>15.565</td>
<td>4.574</td>
<td>15.094</td>
<td>0.594</td>
<td>1.391</td>
<td>1037.7</td>
</tr>
<tr>
<td>I</td>
<td>3.169</td>
<td>1.496</td>
<td>13.205</td>
<td>0.594</td>
<td>0.563</td>
<td>1037.7</td>
</tr>
<tr>
<td>II</td>
<td>2.535</td>
<td>1.396</td>
<td>9.432</td>
<td>0.594</td>
<td>0.347</td>
<td>1037.7</td>
</tr>
</tbody>
</table>

Source: Gover et al. (1994), Salway et al. (1997) and unpublished DETR data (see footnote 4).

Notes:
Pre-I = Pre-stage I emissions, I = Stage 1 emissions, II = Stage II emissions.
Emissions for petrol and diesel cars include hot and cold running driving conditions.
Bus seating capacities: minibus – up to 16 seats; midibus – 17-35 seats; large bus – over 36 seats
VOCs do not include methane emissions.

The emission factors are dependent on a number of parameters, including vehicle type, engine fuel, vehicle speed and the emission regulations in force when the vehicle is registered. Test results for exhaust emissions show considerable variability, even for the same model of vehicle. No two vehicles are ever identical, and driving conditions cannot be exactly replicated over different tests. Thus the emission factors in Table 7.1 are not the legislative limits set for each Stage, but are the mean values from tests covering a sample of different vehicle types within each Stage. The emission factors are based on urban driving conditions and include cold starting conditions for petrol and diesel cars. These cold start emissions are estimated to form 33.8% of total journey distance by car (Eggleston et al., 1991). Following Gover et al. (1994) and Salway et al. (1997) there is no allowance for bus cold starts. Data is not available for
bus cold start emissions, and in any event these are likely to form a very small proportion of total journey distances by bus.

Two approaches are used to measure the exhaust emission effects of substituting bus for car travel. First, three years are chosen (1992, 1995 and 1999) such that each of these years represents the last year in which a particular set of emission regulations are in force (Pre-stage I, Stage I and Stage II respectively). Other choices can be made, but this particular choice of years allows as much time as possible for a given set of regulations to influence vehicle distributions before the introduction of a new set. The corresponding vehicle distributions and distances travelled, as well as a number of other parameters, are derived for each of these years. The most significant change in vehicle distribution over these years is in the number of cars fitted with TWCs, which rises from a very small proportion of the total car stock in 1992 to nearly one half in 1999. The data for 1999 is estimated using the forecasts of car and bus use made in the National Road Traffic Forecasts (Great Britain), (DETR, 1997a).

The second approach is based on the observation that emission regulations tend to change every three or four years, but vehicle life expectancy is generally much longer than this period. Given the relatively long life expectancy of a large bus, for example, the vehicle distribution in 1999 still contains a significant number of Pre-stage I as well as Stage I and Stage II large buses. As Table 7.1 shows, emissions from Pre-stage I buses are significantly higher than those from Stage I and Stage II, with the exception of SO$_2$ and CO$_2$. Any estimate of exhaust emissions in 1999 is affected by this residual bloc of high emission vehicles, even though Stage II regulations are in force. Rather than focus on the actual emissions of a technologically disparate vehicle stock in a given year, therefore, one might wish to know what the exhaust emission effects would be from substituting bus for car travel if all vehicles in a given year conformed to the most recent set of emission regulations for that year. Using this second approach, the effects on exhaust emissions of changes in fuel and vehicle engine technology over time are evaluated using a criterion which is not affected by differences in vehicle life expectancy.
Table 7.2 shows the effects on exhaust emissions of car to bus transfer given these two approaches. For the years 1992, 1995 and 1999 the changes in emissions are calculated firstly on the basis of the emission characteristics of the actual vehicle stock for these years (the “actual emissions” case), and secondly assuming that all vehicles conform to Pre-Stage I, Stage I and Stage II emission limits for 1992, 1995 and 1999 respectively (the “emission limits” case). In general, the results depend not only on the absolute changes in emission regulations, but also on the rate at which emission regulations are tightened for cars relative to buses over time, the changes in vehicle distribution (particularly the absolute and relative increase in the number of cars with TWCs), and changes in distances travelled.

**Table 7.2: Changes in urban exhaust emissions (kg) from a reduction of 100 cars**

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>VOCs</th>
<th>NOx</th>
<th>SO2</th>
<th>PM10</th>
<th>CO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual emissions case:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>-8345</td>
<td>-1214</td>
<td>507.8</td>
<td>50.9</td>
<td>196.1</td>
<td>-4382</td>
</tr>
<tr>
<td>1995</td>
<td>-6985</td>
<td>-960</td>
<td>705.9</td>
<td>49.3</td>
<td>174.9</td>
<td>-5726</td>
</tr>
<tr>
<td>1999</td>
<td>-3879</td>
<td>-412</td>
<td>979.3</td>
<td>45.1</td>
<td>125.8</td>
<td>-6141</td>
</tr>
<tr>
<td>Emission limits case:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>-8634</td>
<td>-1270</td>
<td>469.1</td>
<td>50.9</td>
<td>195.3</td>
<td>-3988</td>
</tr>
<tr>
<td>1995</td>
<td>-1986</td>
<td>-62.3</td>
<td>1342</td>
<td>48.2</td>
<td>87.1</td>
<td>-13991</td>
</tr>
<tr>
<td>1999</td>
<td>-1213</td>
<td>57.2</td>
<td>1040</td>
<td>44.5</td>
<td>52.1</td>
<td>-10255</td>
</tr>
</tbody>
</table>

Source: Table 7.1 and assumptions in text.

Notes:
- Actual emissions case based on the emission characteristics of the actual vehicle stocks for 1992, 1995 and 1999 respectively.
- Emission limits case based on the assumption that all cars and buses in 1992, 1995 and 1999 conform to Pre-stage I, Stage I and Stage II emission regulations respectively.
- Results based on a car occupancy rate of 1.6 persons and a bus load factor of 45 per cent.

For the actual emissions case, the upper half of Table 7.2 shows that substitution of bus for car travel decreases the emissions of CO, VOCs and CO2, but increases the emissions of NOx, SO2, and PM10. This is because buses have high emission factors for NOx, SOx and PM10 relative to cars. The reductions in CO and VOCs become smaller over time because of the tightening of emission regulations, which also accounts for the increases in PM10 emissions becoming smaller. As previously noted, these limits do
not apply to SO\textsubscript{2} and CO\textsubscript{2} emissions. The reductions in CO\textsubscript{2} become larger over time because of the increase in the ratio of diesel to petrol cars, since diesel cars have higher CO\textsubscript{2} emissions than petrol cars. On the other hand, the increases in SO\textsubscript{2} emissions become smaller because of the increase in the ratio of mini- and midibuses to large buses. Finally, the increases in NO\textsubscript{X} emissions are becoming larger over time because the reductions in NO\textsubscript{X} emission limits for cars are significantly higher than for buses. In 1995, for example, people are transferring from petrol cars, where NO\textsubscript{X} emissions have become relatively low, to buses where these emissions are still relatively high.

Similar remarks apply to the emission limits case in the lower half of Table 7.2 between 1992 and 1995, although the changes are generally much greater. This is to be expected, given that all buses and (most) cars in 1992 are assumed to conform to Pre-stage I limits, and all buses and cars in 1995 to conform to Stage I limits. In 1999, however, the change in VOC emissions becomes positive rather than negative because bus VOC emission limits fall only slightly between 1995 and 1999, whilst the limit for petrol cars falls by over 50 per cent. This reduces the scope for VOC emission reductions from petrol cars, so that overall there is a rise in VOC emissions for 1999. The reduction in CO\textsubscript{2} emissions becomes smaller in 1999 because of a (slight) fall in the average distance travelled by car and a (slight) increase for buses. Even though the change in average distance travelled by car is small, the total reduction in distance travelled by car is large because of the high proportion of cars to buses in the vehicle stock, so that the reduction in CO\textsubscript{2} emissions becomes significantly smaller. Finally, NO\textsubscript{X} emission changes in 1999 are still positive but smaller than for 1995. This is partly caused by the change in average distances for cars and buses, but also by the relatively large decrease in bus emission factors for NO\textsubscript{X} between Stage I and Stage II regulations.

To summarise, substitution of bus for car travel has a diverse impact on the total amount of exhaust emissions. Although CO, CO\textsubscript{2} and VOC emissions are generally reduced, the total emissions of NO\textsubscript{X}, SO\textsubscript{2}, and PM\textsubscript{10} are increased. The next section discusses ways of placing a monetary valuation on these emissions, and determines whether or not there is a net increase or decrease in the total monetary costs of exhaust emissions resulting from the substitution of bus for car travel. It is important, however,
to place these results in the context of the assumptions made. Accordingly, Section 5.5 conducts a sensitivity analysis for changes in parameters such as load factors, emission factors and monetary costs.

7.4 The Monetary Costs of Bus and Car Exhaust Emissions

There are a number of ways in which the monetary costs of exhaust emissions from road passenger vehicles can be estimated. They can be estimated directly by tracing the links between emission sources and their effects on health, and placing a value on these effects. Alternatively, indirect estimates can be made through techniques such as hedonic pricing, where emission costs are inferred from observed price differentials in markets such as housing, or the revealed preferences of policy-makers, where the inference is based on the costs of meeting emission regulations. In practice, data limitations mean that studies of emission costs tend to use a variety of estimation methods e.g. Small (1977), Krupnik and Portney (1991), Hall et al. (1992), and Small and Kazimi (1995).

In spite of the extensive literature on the monetary costs of air pollution and climate change, many empirical results are not appropriate for the purpose of this chapter because they are insufficiently disaggregated. The study by Tinch (1995), for example, uses PM$_{10}$ as a proxy for all the health effects of urban air pollution. Although this avoids the problem of double-counting the health effects of individual pollutants, sole use of PM$_{10}$ as the measure of health effects would bias the results of this chapter in favour of petrol engine cars since PM$_{10}$ emissions from road passenger vehicles are mainly from buses and diesel engine cars.

This chapter uses the work reported by Pearce (1993), Small and Kazimi (op. cit.) and Maddison et al. (1996) to calibrate the simulation model CarBus. These studies are chosen because they are at a relatively disaggregated level and deal with similar exhaust emissions. They contain, however, differences in approach and empirical results that require some discussion. Pearce (1993) gives estimates for the 1991 unit and overall monetary damage costs of air pollution from UK road transport, where these estimates are revisions of those in Pearce et al. (1992). Further revised figures
(updated to 1993) given in Maddison et al. show significant differences in the damage costs. The estimates are shown in Table 7.3:

Table 7.3: Emission costs from UK road transport 1991 and 1993

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Emissions (tons) 1991</th>
<th>Unit Value (£/ton) 1991</th>
<th>Damage Costs (£m) 1991</th>
<th>Damage Costs (£m) 1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂</td>
<td>30,000,000</td>
<td>13.33</td>
<td>400.0</td>
<td>110.1</td>
</tr>
<tr>
<td>CH₄</td>
<td>10,000</td>
<td>70.00</td>
<td>0.7</td>
<td>N/A</td>
</tr>
<tr>
<td>SO₅</td>
<td>58,000</td>
<td>220.69</td>
<td>12.8</td>
<td>3,760.0</td>
</tr>
<tr>
<td>NO₂</td>
<td>1,400,000</td>
<td>190.00</td>
<td>266.0</td>
<td>6,150.0</td>
</tr>
<tr>
<td>CO</td>
<td>6,000,000</td>
<td>10.43</td>
<td>62.6</td>
<td>N/A</td>
</tr>
<tr>
<td>VOCs</td>
<td>970,000</td>
<td>N/A</td>
<td>N/A</td>
<td>2,870.0</td>
</tr>
<tr>
<td>PM</td>
<td>208,000</td>
<td>9,778.84</td>
<td>2,034.0</td>
<td>5,550.0</td>
</tr>
<tr>
<td>Pb</td>
<td>1,800</td>
<td>N/A</td>
<td>N/A</td>
<td>280.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>2,776.0</td>
<td>18,720.1</td>
</tr>
</tbody>
</table>

Source: Data for 1991 from Table 10.4 in Pearce (1993), data for 1993 from Box 3.7 and Box 4.10 in Maddison et al. (1996).

Notes:
N/A = data not available.
Monetary values are at current prices. Metric tons are assumed here and elsewhere in the paper.
Damage cost figures for 1991 and 1993 are not strictly comparable.
Individual values may not sum to the corresponding total because of rounding error.

Small and Kazimi (SK) (op. cit.) use Los Angeles data on road vehicle exhaust emissions. Their focus is on the effects of particulate matter and ozone formation, which SK assume to be the forms of air pollution most damaging to human health. This assumption is supported by the results of Ostro (1994, p. 51). Although no estimates are made in the SK study for the direct health effects of ambient SO₂ and NO₂ emissions, estimates are made for their indirect effects via their role in particulate formation. The emission cost estimates are shown in Table 7.4.

Using the findings of a number of previous studies, SK derive the monetary costs of mortality from particulates, and morbidity from particulates and ozone, based on a value of life of $4.87 million (US dollars, 1992 prices) or £2.77 million. This value is the geometric mean of the range of values given in Fisher et al. (1989) and is very close to the figure used in Hall et al. (1992). It is, however, considerably higher than the 1994 Department of Transport (DoT) “official” figure of £744,000 used in road
accident and safety project valuations. In relation to this latter value, Tinch (1995, p.124) notes that "There is substantial evidence that this value is rather low ...", and Pearce *et al.* (1992, p. 4-5) comment that "A valuation of £1.5 million would seem fairly conservative". The value of £1.5 million is used by Pearce and Crowards (1996) in their study of the links between human health and particulate matter.

<table>
<thead>
<tr>
<th></th>
<th>VOCs</th>
<th>NOX</th>
<th>SOX</th>
<th>PM10</th>
<th>CO2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality from Particulates</td>
<td>0.96</td>
<td>4.80</td>
<td>59.55</td>
<td>55.23</td>
<td>-</td>
<td>120.53</td>
</tr>
<tr>
<td>Morbidity from Particulates</td>
<td>0.05</td>
<td>0.24</td>
<td>2.95</td>
<td>2.73</td>
<td>-</td>
<td>5.97</td>
</tr>
<tr>
<td>Morbidity from Ozone</td>
<td>0.65</td>
<td>1.03</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>1.68</td>
</tr>
<tr>
<td>Climate Change</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.035</td>
<td>0.035</td>
</tr>
<tr>
<td>Total</td>
<td>1.66</td>
<td>6.06</td>
<td>62.44</td>
<td>57.95</td>
<td>0.035</td>
<td>128.2</td>
</tr>
</tbody>
</table>

Source: adapted from Small and Kazimi, *op. cit.*, Table 5, p. 22.

Notes:
- Estimates based on a $4.87 million value of life, the geometric average of the high and low particulate mortality coefficients, the geometric average of two ozone morbidity figures with costs allocated equally to NOX and VOCs, and the particulate morbidity costs used by Hall *et al.* (1992) and Krupnik and Portney (1991).
- CO2 estimate based on Manne and Richels (1992) and information in SK (*op. cit.*, p. 28, footnote 22).
- Individual estimates may not sum to the corresponding total because of rounding error.
- Conversion from US$ to UK£ on the basis of £1 = $1.76 in 1992.

SK argue that there is little quantitative information on the health costs of CO, and what little information there is tends to suggest that they are less important than either particulates or ozone in costing motor vehicle emissions (SK, *op. cit.*, p.16). For these reasons, they do not include the costs of CO emissions. The effects of including different values for the monetary costs of CO and other emissions are explored in the sensitivity analysis of Section 7.5, although the results of this analysis suggest that changes in CO costs do not have a significant impact on the overall cost changes.

Of the emissions accounted for in Table 7.4, by far the biggest single source of exhaust emission costs per unit of emission is mortality from particulates. This result is consistent with the finding of Bown (1994) who claims that PM10, mainly from car exhausts, is responsible for up to 10,000 premature deaths a year in England and
Wales, a figure also used in RCEP (1994, p. 31). The mortality costs are split roughly equally between $SO_\text{x}$ and $PM_{10}$.

$CO_2$ emissions have indirect effects on health and the environment via the induced change in global temperatures. A detailed account of the methodological problems involved in the quantification of $CO_2$ emissions costs, and the results of empirical studies, is contained in Cline (1992). Although estimates of the monetary costs of $CO_2$ emissions are subject to considerable uncertainty, three early studies (Nordhaus and Yohe, 1983, Edmonds and Barnes, 1990, Manne and Richels, 1992) provide monetary valuations which are of approximately the same order of magnitude i.e. between $100 and $300 per ton of carbon at 1990 prices (Cline, op. cit., p. 188). Manne and Richels (op. cit., p. 59) estimate a cost of about $208 per ton of carbon (1990 prices), based on the costs of reducing $CO_2$ emissions to sustainable levels i.e. stabilisation at 1990 levels up to the year 2000, followed by a gradual 20 per cent reduction over the next decade, and then a permanent continuation of these reduced emission levels. Updated to 1992 prices, the cost of $CO_2$ emissions becomes about $225 per ton of carbon, equivalent to about £130 per ton. More recent research by Fankhauser and Tol (1996) suggest much lower climate change costs, somewhere in the region of $20 to $28 per ton of carbon, although these estimates are highly sensitive to changes in the discount rate.

Small (1977, pp. 123-4) suggests that Los Angeles is likely to have six times as many days as twelve other USA cities outside California on which climatic conditions combine to form a "trap" for air pollutants, so that the estimates for Los Angeles may not be representative of other cities. In order to apply these exhaust emission cost estimates to the UK, it is assumed that the SK figures for Los Angeles exhaust emission costs represent the worst case and thereby an upper limit to UK exhaust

---

5 Pearce and Crowards (1996) argue that this figure is likely to be an overestimate, although their central estimate of 6,668 premature deaths still implies total mortality costs of £7 billion (1994 prices) from anthropogenic PM emissions.

6 One ton of carbon is equivalent to 3.67 tons of carbon dioxide. Thus the figure of £130 per ton of carbon is equivalent to £35.4 per ton of carbon dioxide, which is the figure implicitly given in the SK study and the one used in Table 7.5 of this chapter.

7 These lower climate change cost estimates are of a similar order of magnitude to those used in Maddison et al. (1996).
emission costs. UK costs will thus lie somewhere between the best case i.e. zero costs, and the worst case.

In the absence of any firm evidence to indicate where to draw the line between the best and worst case costs, it is initially assumed that UK exhaust emission costs are one-half those of the SK figures for Los Angeles (SK*0.5), which is roughly equivalent to assuming a value of life of $2.35 million dollars, or £1.34 million. This value of life is still about twice that of the official UK figure, but is of a similar order of magnitude to the preferred value of £1.5 million used in Pearce and Crowards (op. cit.). Another set of estimates is then generated by halving again the SK figures (SK*0.25), so that the SK*0.25 value of life is roughly the same as the official UK figure. These scaling down factors are applied to all exhaust emissions except CO$_2$, because global warming occurs in the stratosphere and is unlikely to be affected by climate conditions in the troposphere. UK costs from CO$_2$ emissions are thus assumed to be the same as those for Los Angeles. The no-change assumption for CO$_2$ is later relaxed for the sensitivity analysis of Section 7.5.

Table 7.5 summarises the unit costs (£000s per ton of exhaust emission) of bus and car transport using the different cost assumptions associated with the SK, SK*0.5, SK*0.25, Pearce (1993) and Maddison et al. (1996) studies. For ease of reference, the unit cost values for the SK, Pearce and Maddison cases are the same as those in the original studies. The net change column shows the change in monetary costs (£000s, adjusted for inflation up to 1995) from a reduction of 100 cars, and is based on the actual emissions case for 1995 from Table 7.2. Finally, the total cost column shows the total exhaust emission costs (£billion, adjusted for inflation up to 1995) arising from urban bus and car travel, also based on the actual emissions case for 1995 from Table 7.2.
Table 7.5: Bus and car urban exhaust emission costs 1995: unit, net change and total costs.

<table>
<thead>
<tr>
<th>Study</th>
<th>CO</th>
<th>VOCs</th>
<th>NO\textsubscript{X}</th>
<th>SO\textsubscript{2}</th>
<th>PM\textsubscript{10}</th>
<th>CO\textsubscript{2}</th>
<th>Net change (£000)</th>
<th>Total Cost (£bn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK*0.5</td>
<td>0</td>
<td>0.83</td>
<td>3.03</td>
<td>31.22</td>
<td>28.98</td>
<td>0.035</td>
<td>8.34</td>
<td>2.81</td>
</tr>
<tr>
<td>SK*0.25</td>
<td>0</td>
<td>0.42</td>
<td>1.52</td>
<td>15.61</td>
<td>14.49</td>
<td>0.035</td>
<td>4.06</td>
<td>2.00</td>
</tr>
<tr>
<td>Pearce</td>
<td>0.01</td>
<td>N/A</td>
<td>0.19</td>
<td>0.22</td>
<td>9.78</td>
<td>0.013</td>
<td>1.84</td>
<td>0.61</td>
</tr>
<tr>
<td>Maddison</td>
<td>N/A</td>
<td>2.96</td>
<td>4.39</td>
<td>64.83</td>
<td>26.68</td>
<td>0.004</td>
<td>8.72</td>
<td>3.04</td>
</tr>
</tbody>
</table>

Source: Tables 7.2, 7.3 and 7.4, and assumptions in text.

Notes:
Figures for unit exhaust emissions are for cost per ton of emission (£000s).
Figures for Pearce (1993) unit costs are 1991 current values. Figures for Maddison et al. (1996) unit costs are 1993 current values. Figures for SK, SK*0.5 and SK*0.25 unit costs are 1992 current values. The Pearce VOC and Maddison CO unit costs are set to zero in the calculations. Since the main constituent of PM is PM\textsubscript{10}, the unit costs of PM emissions given in the Pearce and Maddison et al. studies is used to proxy the corresponding PM\textsubscript{10} costs in Table 7.5. SO\textsubscript{2} unit cost values are proxied by the SO\textsubscript{X} values given in the SK, Pearce and Maddison et al. studies.

Figures for the net change and total cost columns are based on the unit cost values updated to 1995 values, actual emission values based on 1995 vehicle distributions and distance travelled, a reduction of 100 cars, a car occupancy rate of 1.6 persons, and a bus load factor of 45 per cent.

The net change column in Table 7.5 shows that urban exhaust emission monetary costs increase as a result of switching from car to bus travel. The reason is the relatively high monetary cost values given to SO\textsubscript{2} and PM\textsubscript{10} emissions. Table 7.2 shows an increase in these emissions (as well as those of NO\textsubscript{X}), and the net change column in Table 7.5 implies that the increase in their monetary costs outweighs the decrease in the monetary costs of CO, VOCs and CO\textsubscript{2}. PM\textsubscript{10} exhaust emission costs are always the largest single cost change in the net change column, whilst around 90 per cent of the total urban exhaust emission costs are from petrol and diesel cars. Similar results for net change and total costs are obtained when the same calculations are performed for the actual emissions case for the 1992 and 1999 vehicle distributions and distances travelled.

It should be noted that the similarity in the net change and total cost figures for the SK*0.5 and Maddison et al. estimates does not imply similarity in the cost changes for the individual exhaust emission categories. SK*0.5 gives cost changes (£000s) of \(-0.86, 2.30, 1.66, 5.45\) and \(-0.22\) for CO, VOCs, NO\textsubscript{X}, SO\textsubscript{2}, PM\textsubscript{10} and CO\textsubscript{2} emissions respectively, whilst the corresponding figures for Maddison et al. are \(-3.06, 3.34, \ldots\)
3.44, 5.02 and −0.02. Thus the SK estimates tend to give a much greater weighting to the monetary costs of PM$_{10}$ relative to the CO, VOCs, NO$_X$ and SO$_2$ emissions.

Apart from the SK*0.5 and Maddison et al. results, however, there are considerable differences in the net change and total cost estimates in Table 7.5. Which, if any, of these estimates are comparable to other estimates made for UK car and bus exhaust emission costs? One point of comparison is with the total cost estimates given in RCEP (op. cit., Table 7.2) for both road freight and passenger transport air pollution and climate change costs. RCEP estimates that these costs range from £3.54 bn to £8.26 bn. Assuming that road freight transport accounts for roughly 25 per cent of these costs (a figure consistent with RCEP, Table 7.3), then the total air pollution and climate change costs from road passenger transport range from £2.66 bn to £6.20 bn for 1994/95. Given UK GDP at factor cost was £579.14 bn in 1994, these costs form about 0.5 per cent to 1.1 per cent of UK GDP. The total (urban) costs given in Table 7.5 imply that the SK*0.5 and Maddison et al. urban exhaust emission costs are around 0.5 per cent of UK GDP for 1995. If the non-urban costs are assumed to be around one-half to one-third of the urban costs (a figure based on the non-urban distances travelled by cars and buses), then the SK*0.5 and Maddison et al. national estimates would fall well within the 0.5 to 1.1 per cent range derived from RCEP. The SK*0.25 national estimates would fall just inside the lower part of this range if the ratio of non-urban to urban costs is assumed to be one-half.

A corresponding analysis is performed for the emission limits case given in Table 7.2, although the detailed results are not reported here. The net change values are generally similar to those for the actual emissions values given in Table 7.5, although total costs are lower for 1995 and 1999, as one would expect. In terms of the individual net changes for 1995 and 1999, however, the largest single net cost change occurs for NO$_X$ rather than PM$_{10}$ emissions because of the large increase in the former relative to the latter (see Table 7.2). But overall the results are comparable: in particular, the net change in urban exhaust emission costs is always positive. The next section discusses the robustness of this result in relation to changes in some of the assumptions made earlier.
7.5 Substitution of Bus for Car Travel: a sensitivity analysis

This section explores the effects of changing a number of the parameter values associated with the results in Section 7.4. The sensitivity analysis is based on the actual emissions case for 1995 although other cases, including the effects of Stage III and Stage IV regulations, are considered. Following the discussion in Section 7.4, the sensitivity analysis is confined to the SK*0.25, SK*0.5 and Maddison et al. estimates. The full CarBus spreadsheet model contains over sixty parameters, so that the number of possible permutations in these parameter values is extremely large, even when the changes are restricted within reasonable bounds. The results reported below are some of the more realistic and relevant simulations performed by the author.

7.5.1 Changes in Load Factors

It can be argued that the assumed bus load factor of 45 per cent is too conservative, since any transfer of car users to bus could be accommodated by allowing bus load factors to increase rather than using extra buses to maintain the existing load factor. Of particular interest is the bus load factor at which the change in exhaust emission costs begins to decrease (or break-even bus load factor), ceteris paribus. Using the SK*0.25, SK*0.5 and Maddison et al. unit cost estimates from Table 5.5, these load factors are 66, 76 and 75 per cent respectively. These load factors are much higher than the historically observed urban peak value of 45 per cent. The emission limit case for 1995 implies even higher load factors of 70, 89 and 150 per cent respectively. This latter value occurs because the NO\(_x\) and SO\(_2\) emission factors do not fall very much (if at all) for buses between Pre-stage I and Stage I regulations (see Table 7.1). Since the Maddison et al. estimates give very high unit costs for NO\(_x\) and SO\(_2\) emissions, their total monetary costs are relatively large so that a high (in fact, impossible) bus load factor is required to achieve the break-even point.

It is also of interest to determine the corresponding break even car occupancy rate. For a bus load factor of 45 per cent, the respective values are 1.1, 0.9 and 1.0 persons per car, which are well below the values typically assumed for European driving
conditions\textsuperscript{8}. In the emission limit case for 1995, the values are 1.0, 0.8, and 0.5 persons respectively.

### 7.5.2 Changes in the Monetary Costs of Individual Exhaust Emissions

Another possibility is to consider whether changes in the costs per ton of the exhaust emissions cause any significant change in the net change costs associated with a reduction of 100 cars. One possibility is to change the cost per ton of CO\textsubscript{2} in the SK estimates, since these are considerably higher than those used in the Maddison \textit{et al.} estimates. But changing the SK*0.25 cost per ton of emission figures for CO\textsubscript{2} from £35 to £4 (the figure used in the Maddison \textit{et al.} estimates) causes the net change in emissions costs to increase from £4,060 to £4,250. Thus a decrease of almost 90 per cent in CO\textsubscript{2} unit emission costs results in a relatively small increase of 4.7 per cent for the net change figure. For SK*0.5, the corresponding increase is also small, rising from £8,340 to £8,530. It is also possible that the Maddison \textit{et al.} CO\textsubscript{2} unit cost estimate is too low. Substituting the SK unit cost estimate for CO\textsubscript{2} into the Maddison \textit{et al.} estimates gives a small decrease in the net change costs from £8,720 to £8,520.

CO emissions are generally given a value of zero in Table 7.5. The exception is the Pearce study, which estimates the damage costs of CO emissions at £10 per ton. Substituting this value into the SK*0.25, SK*0.5 and Maddison \textit{et al.} estimates gives a very small decrease in the net change costs of 2.0, 0.8 and 0.9 per cent respectively.

The net change figures in Table 7.5 are for the actual emissions case and are dominated by the change in the cost of PM\textsubscript{10} emissions, particularly for the SK estimates. Suppose it is assumed (somewhat arbitrarily) that there is a margin of error of 50 per cent around the estimates given in Table 7.5 for the unit costs of PM\textsubscript{10} emissions. Halving the unit costs of PM\textsubscript{10} emissions causes the net change costs to fall by around 30 per cent for the SK*0.25, SK*0.5 and Maddison \textit{et al.} estimates, a fairly substantial decrease. For the emission limit case, however, the biggest single cost change is that

\textsuperscript{8} For example, the National Travel Survey 1994/96 gives an average car occupancy rate of 1.6 persons, whilst de Borgcr \textit{et al.} use a figure of 1.7 for urban peak and urban off-peak driving conditions in Belgium.
for NO\textsubscript{x} emissions. Halving the unit costs of NO\textsubscript{x} emissions in the emission limit case causes net change costs to fall by around 27 per cent for the SK*0.25, SK*0.5 and Maddison \textit{et al.} estimates, another substantial decrease. In comparison, halving other exhaust emission unit costs has less of an impact on net change costs. In all cases the changes are less than 20 per cent, and in most cases the changes are well below this figure.

7.5.3 Looking Ahead: Stage III and Stage IV regulations

From the year 2000 more stringent exhaust emission regulations for road freight and passenger transport are introduced in Europe to continue the process initiated by Stages I and II: Euro III starts in 2000/1, and Euro IV in 2005/6. The reductions in exhaust emissions for Euro IV relative to Euro II are substantial: it is envisaged that most car and light goods vehicle (LGV) exhaust emissions will fall by around 55 to 70 per cent, with somewhat lower reductions for buses and heavy goods vehicles (HGVs). As far as PM\text{10} emissions are concerned, however, there are no reductions for petrol cars and petrol LGVs whilst those for diesel vehicles fall significantly (about 70 per cent for diesel cars and diesel LGVs, and 50 per cent for buses and HGVs)\textsuperscript{9}.

To determine the difference such reductions would make to the results obtained previously, these PM\text{10} emission reductions for diesel cars and buses are substituted into the 1999 actual emissions case. This causes the net change in urban exhaust emission costs to fall by 29, 25 and 18 per cent in the SK*0.25, SK*0.5 and Maddison \textit{et al.} results respectively, another substantial decrease.

To summarise, the general conclusion from this sensitivity analysis\textsuperscript{10} is that the positive net change in urban exhaust emission costs resulting from the substitution of bus for

\textsuperscript{9} The PM\text{10} exhaust emission reductions for diesel vehicles occur because of the introduction of ultra-low sulphur diesel fuel and new technologies such as the Continuously Regenerating Trap (CRT) which can remove up to 90 per cent of particulate material from diesel exhaust emissions.

\textsuperscript{10} The effects of bus for car substitution were also analysed in terms of a model based on national averages for emission factors, load factors and so on. Although this model is based on earlier work by the author and contains unrevised data for a number of emission factors as well as a different analytical framework, the results are qualitatively similar to those of the urban model. The net change in exhaust emission costs remains positive, although the net change and total cost values are higher because of the inclusion of relatively low rural bus load factors in the simulations.
car travel is relatively insensitive to reasonable changes in nearly all the parameters. An important exception is for PM$_{10}$ emissions where changes in monetary costs or emission factors do have a relatively large impact on the net change figure for urban exhaust emission costs. A policy of developing improved fuel and bus technologies in order to satisfy (or go beyond) Euro III and IV regulations for particulate reductions will significantly reduce the net change figure. Policies relying on increasing bus load factors appear less likely to succeed in reducing total exhaust emission costs, since the load factors required are very high by historical standards, although higher load factors will generate other important cost reductions (see next section).

7.6 Substitution of Bus for Car Travel: other cost changes

This section deals with a number of other cost changes arising from the substitution of bus for car travel, namely fuel consumption, traffic congestion, noise pollution, road accidents, and road damage. A number of these cost changes are traditionally regarded as externalities arising from market failure, but Nash (1997) shows that the concept of an externality is much more complex than is generally thought. Nonetheless, the concept is a useful starting point for the subsequent analysis. The range of costs considered is by no means comprehensive, but is determined by two factors: first, data availability, and second the likelihood that they will not be internalised through market or other transactions. Monetary estimates of these cost changes are numerous but, as in the case of air pollution, often highly aggregated and sometimes unsuitable for the type of analysis conducted in this chapter. The following estimates are based on a literature review by the author and are customised for a welfare analysis using the CarBus model. Like most, if not all estimates in this area, they are subject to considerable uncertainty.

For expositional convenience, the following description of the derivation of the estimated values is mainly for 1992, but the CarBus model also calculates values for 1995 and 1999. The estimated values are calculated in terms of vehicle kilometres travelled. Once the overall changes in bus and car vehicle kilometres are known, the overall change in the external costs can be calculated. Estimation of the changes in vehicle kilometres for each of the six vehicle categories is not a straightforward
process: details are given in Appendix three. An example of the CarBus simulation used to derive the values in Table 7.6 is also given in Appendix three.

7.6.1 Fuel Cost Savings

Oil is a non-renewable resource. The more we use today, the less there is available for tomorrow. This may not be of great concern to current generations, but future generations may feel very differently. As we run out of new reserves to exploit, increasing scarcity will increase oil prices and encourage the use of substitute forms of energy, both renewable and non-renewable. Unless renewable energy sources can compensate for the loss of non-renewable resources, future generations will suffer. The problem is that the market mechanism is an inherently defective process for determining prices that adequately reflect the utility of future generations (Dasgupta and Heal, 1979).

This market failure is a form of intergenerational externality. Current generations use resources at prices that do not fully incorporate the costs imposed on future generations. The difficulty lies in establishing a monetary value for this externality. The approach used in this paper is to use the current (pre-tax) price of fuel as a measure of this discounted value. Some may object that this is too high a value, on the grounds that the scarcity problem may be exaggerated, whilst others that it is too low. But in the absence of any knowledge concerning the likely deposits of undiscovered oil or the utility functions of future generations, this value seems as reasonable as any other.

Given the six vehicle categories in Table 7.1, the urban fuel consumption (in litres per 100 kilometres, starting with petrol car without TWC) is 8.3, 7.8, 6.2, 10.92, 18.94 and 44.26 respectively (Gover et al., op. cit., Tables 5.1 and B.12). In 1992 about 50% of motorists (excluding diesel cars) normally bought unleaded petrol (Transport Statistics Great Britain (TSGB), 1995, Table 2.2). It is assumed that half the petrol cars without TWC run on leaded, and the other half on unleaded petrol. It is also assumed that all petrol cars with a TWC run on unleaded petrol only. Excluding duty and VAT, the prices of 4 star leaded petrol, unleaded petrol and diesel fuel were 50.6, 46.1 and 44.6 pence per litre respectively in 1992 (TSGB, 1995, Table 2.4). These
parameter values can then be incorporated in the CarBus spreadsheet to find the value of fuel saved as a result of bus for car substitution.

7.6.2 Reductions in Congestion Costs

The amount of congestion caused by a vehicle can be measured in terms of passenger car units (PCUs). These values are used by the DETR in conjunction with other measures of road use to allocate the capital costs of the road network between different vehicle classes. The rationale for this procedure is that congestion causes new roads to be built, or existing ones to be upgraded, thus incurring capital expenditure. The DETR assigns PCU values of 1, 1.2 and 1.4 to cars, minibuses and midibuses respectively. PCU values of 1.6 and 1.7 are given to 36 to 60 seat buses and over 60 seat buses respectively. Since this study does not distinguish between these latter two types of large bus, a weighted PCU value is calculated for a single large bus category on the basis that there are approximately 1.5 times as many 36-60 seat buses as there are over 60 seat buses (BCSGB, various). This gives a PCU value of 1.64 for the large bus category. An overall weighted average bus PCU value can be derived by using vehicle stock as a weight, where the weights are taken from Bus and Coach Statistics (BCSGB) (1992/93). This gives a PCU value of 1.54.

These PCU values provide a very approximate guide to the amount of congestion caused by different types of vehicle. Newbery (1990) estimates that the average congestion cost across all roads in 1990 was 3.4 pence per PCUkm, although there is considerable variation around this average: urban central roads at peak periods, for example, have an average congestion cost of 36.37 pence per PCUkm. Inflation adjusting this national average figure to 1992 figures gives a value of 3.7 pence per PCUkm.

In order to derive a value for urban areas, the six main categories of urban road type congestion costs given in Newbery (1990) are weighted according to their categorisation as either peak or off-peak values, where the weightings are based on those given in Glaister (1985a). The resulting weighted average congestion cost for
urban areas is 15.04 pence per PCUkm for 1990. Inflation adjusting this figure to 1992 values gives a value of 16.36 pence per PCUkm.

Assuming that the occupants of 100 cars all transfer to bus travel, and using the figures on vehicle kilometres travelled in 1992, there is a reduction of 1,690,591 PCUkm for car travel, and an increase of 768,182 PCUkm for bus travel, yielding an overall reduction of 922,409 PCUkm. Using the value of 16.36 pence per PCUkm, the CarBus spreadsheet can then compute the reduction in urban congestion costs.

7.6.3 Reductions in Noise Costs

Tinch (1995) provides a review of monetary estimates of noise costs from different types of motor vehicle. The preferred figures used in the Tinch study (op. cit.) are £5.5 and £10.9 per 1000 kilometres for cars and buses respectively. These figures are based on a weighting of 3:2:1 for the relative noise nuisance from heavy goods vehicles, buses and cars (Peirson, Skinner and Vickerman, 1994). This weighting appears to be based on the noise from large buses. Since no weightings are available for the three main types of bus, it is necessary to make an estimate of their relative noise nuisance value (NNV). Given a car NNV equal to one, the suggested values for a large bus, midibus and minibus are 2, 1.6 and 1.2 respectively. These figures are based broadly on the relative vehicle weights within the bus categories on the assumption that vehicle noise is roughly proportional to vehicle weight. The problem with this procedure is that the average gross vehicle weight of a large bus is about nine times greater than that of a car, yet the Tinch estimates assume that large buses have a NNV only twice that of a car. Nonetheless, given the lack of any reliable data for NNVs amongst these vehicle categories, the values above may not be too unreasonable to serve as a starting point for the subsequent analysis.

Assuming the Tinch (op. cit.) values are for 1994, and inflation adjusting them to 1992, a set of noise costs is derived using the NNV weightings above. These costs are £5.3, £10.5, £8.5 and £6.4 per 1000 kilometres for cars, large buses, midibuses and minibuses respectively.
7.6.4 Reductions in Road Casualty Costs

Although road casualty costs can be partly internalised through the payment of insurance premiums, there is reason to believe that these payments do not fully account for all the effects of road casualties (Jansson, 1997). In the case of a fatal accident, for example, compensation is paid to all the affected parties except the deceased. Whilst the compensation may provide a monetary measure of the deceased's concern for his or her dependants, it does not generally provide a measure of the value of the deceased's life.

There is a large literature on the monetary valuation of a human life. Mishan (1988) provides a review of the techniques commonly used, and also a highly critical appraisal of the work in this area (op. cit., p.352). A survey of studies for the UK, US and Sweden, using mainly stated preference techniques (Jones-Lee, 1990), concludes that the distribution of the value of life in 1989 has mean and median values of $3.4 million and $1.1 million respectively.

The personal injuries resulting from road accidents are classified by the Department of Transport into three broad groups: fatal, serious and slight. Accidents concerning vehicle damage only are not included in the Department of Transport's Road Accidents Great Britain: the Casualty Report. It is estimated that the total cost-benefit value of road accidents in 1995 was £13,280 million, of which £9,500 million is attributable to personal injury accidents, with damage-only accidents accounting for the remainder (Road Accidents Great Britain 1995, 1996, p.22). One accident may give rise to several casualties. The average cost of a casualty (at 1994 prices) in each of these groups is estimated to be £784,090, £89,380 and £6,920 respectively (Road Accidents Great Britain 1994, 1995), although these valuations appear to be substantial underestimates (see Section 7.4). These costs include lost output and pain, grief and suffering caused by the accident, but exclude the direct accident costs such as police work and property damage. In effect, they represent a measure of the external costs associated with road accidents.
The procedure is to calculate the average car and bus casualty costs per thousand kilometres, and use these values to estimate the reduction in total casualty costs resulting from a substitution of bus for car travel. Casualty figures are given in TSGB (1995, Table 4.14). One problem is that bus and coach casualty data is not given separately. The solution adopted is to allocate casualties on the basis of vehicle kilometres travelled, although this procedure may be somewhat unfair to bus travel since it is likely that bus travel has a lower accident rate than coach travel. Another problem is that some accidents occur on motorways, where local bus services do not generally operate, so that inclusion of motorway casualty costs will bias the casualty cost savings of car to bus transfer in favour of bus travel. Nonetheless, road accidents on motorways in 1992 formed only 3 per cent of all accidents (TSGB, 1995, Table 4.15), so that the bias is likely to be small.

Adjusting the casualty costs above to 1992 prices, the total casualty costs associated with car accidents in 1992 were £4.206 bill., and £73.058 mill. for bus travel. The average casualty cost per thousand kilometres for car and bus travel is £12.44 and £27.30 respectively. No differentiation is made between petrol and diesel cars in respect of these casualty costs, since it is unlikely that the casualty cost per petrol car is significantly different to that of diesel cars.

7.6.5 Increases in Road Track Costs

The UK road network is provided by the government, which sets a price for the use of this network based on the costs of provision. These costs (sometimes referred to as road track costs) are divided into current and capital costs. Current costs include maintenance, policing and administration, whilst capital costs cover the building and improvement of roads. These costs are allocated to different vehicle categories on the basis of various measures of road use and wear such as vehicle kilometres, vehicle weight, standard axles and PCUs. Details are given in the annual publication "The Allocation of Road Track Costs" (Department of Transport).

This section shows that road track costs increase as a result of substituting bus for car travel. Strictly speaking, these cost increases are not externalities since the government
could in principle charge bus users for them. It could also reduce the amount paid by car (and commercial vehicle) users. In practice, however, it is not obvious that the government would follow a policy of increased bus use charges since it may conflict with more fundamental environmental objectives such as encouraging a shift from private to public transport. It follows that, for practical purposes, the evaluation of bus for car substitution should treat these costs as external and include them in the same way as the other externalities discussed earlier. Given the size of these costs (see Table 7.6 below), however, it should be noted that this procedure is substantially disadvantageous to bus use on the baseline assumption of a 45 percent load factor.

A critical determinant of road track costs is road-wear. The standard axle parameter is a measure of the relative road-wear caused by each vehicle. Road-wear increases exponentially with respect to vehicle weight for a given number of axles. A ten ton two axle vehicle will cause more than twice the road wear associated with a five ton two axle vehicle. Thus a car is given a standard axle value of 0.00002 whilst a 60+ seat bus has a value of 0.306, over 15,000 times greater than that of the car (Allocation of Road Track Costs, op. cit). These values include an allowance for passenger loadings.

The procedure is to calculate the average road cost per vehicle kilometre for each vehicle category i.e. cars, minibuses, midibuses and large buses. Individual parameter values for the three bus categories are those for 1993/94, since no individual values are given prior to this year. Taking the total road costs allocated in 1993/94 to each of the vehicle categories, and dividing by the number of vehicle kilometres travelled, the following figures (adjusted to 1992/93 prices) are obtained for cost per vehicle kilometres: £0.009, £0.015, £0.021, and £0.080 for cars, minibuses, midibuses and large buses respectively. Thus a large bus creates nearly nine times the road costs per vehicle kilometres associated with a car, the reason being the relatively high levels of road wear caused by the large bus.

7.6.6 Overall Cost Changes: a synthesis and simulation

The calculations have so far been made on assumptions which incorporate an average bus load factor of 45%. It may be argued that at the margin many of the 160 car users
who transfer to bus could be accommodated by existing bus services, given the observed spare capacity which exists in much of the local bus service network in Britain. Table 7.6 summarises the results of a simulation based on the actual emissions case for 1995 using the SK*0.5 unit cost estimates, and takes into account a range of bus load factors (from 35 to 55 per cent). As in the preceding sections, the simulation is based on urban conditions.

Table 7.6: Overall urban cost changes at different bus load factors (£000s)

<table>
<thead>
<tr>
<th>Type of cost</th>
<th>35%</th>
<th>40%</th>
<th>45%</th>
<th>50%</th>
<th>55%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhaust emissions</td>
<td>14.29</td>
<td>10.95</td>
<td>8.34</td>
<td>6.26</td>
<td>4.55</td>
</tr>
<tr>
<td>Congestion</td>
<td>-74.85</td>
<td>-83.96</td>
<td>-91.04</td>
<td>-96.71</td>
<td>-101.35</td>
</tr>
<tr>
<td>Noise</td>
<td>-2.00</td>
<td>-2.32</td>
<td>-2.57</td>
<td>-2.78</td>
<td>-2.94</td>
</tr>
<tr>
<td>Accidents</td>
<td>-2.68</td>
<td>-3.69</td>
<td>-4.48</td>
<td>-5.10</td>
<td>-5.62</td>
</tr>
<tr>
<td>Road damage</td>
<td>5.43</td>
<td>3.78</td>
<td>2.50</td>
<td>1.47</td>
<td>0.63</td>
</tr>
<tr>
<td>Overall cost</td>
<td>-63.95</td>
<td>-84.39</td>
<td>-100.29</td>
<td>-113.00</td>
<td>-123.41</td>
</tr>
</tbody>
</table>

Source: Assumptions in text and (for exhaust emissions) Tables 7.2 and 7.5.

Notes:
Results based on 1995 actual emissions case, SK*0.5 unit cost exhaust emission estimates, a car occupancy rate of 1.6 persons, and bus load factors ranging from 35 to 55 per cent. All values are inflation adjusted to 1995 values.
Fuel consumption figures based on data in Gover et al. (1994) and valued at the retail price of fuel excluding fuel duty and value added tax.
Congestion costs based on measures of congestion caused by different vehicle types in “The Allocation of Road Track Costs” (Department of Transport (DOT), 1993), and valued using cost estimates from Newbery (1990).
Accident costs based on data in “Road Accidents Great Britain: the Casualty Report” (DOT, 1995) and “Highways Economics Note No. 1” (DOT, 1996).
Road damage costs based on data in The Allocation of Road Track Costs (op. cit.).

When all the above costs are taken into account, substitution of bus for car travel in urban areas decreases the overall costs. The decrease is dominated by the fall in congestion costs. Overall costs only start to increase at bus load factors below 25 per cent, which is lower than the historically observed bus load factors of around 45 to 30 per cent for peak and off-peak bus travel respectively. It could be argued that this simulation is unduly favourable to bus travel, since it is based on urban areas where bus load factors and congestion costs are higher than on a national basis (i.e. including
rural as well as urban travel). Using national data\textsuperscript{11}, the overall cost change figures are substantially lower than those above, but still positive for bus load factors greater than 33 per cent. As noted in Section 7.5.1, exhaust emission costs do not start to decrease until a bus load factor of 76 per cent is reached.

7.7 Conclusions

Substitution of bus for car travel in urban areas yields significant monetary gains when a number of costs (exhaust emissions, fuel consumption, traffic congestion, noise pollution, road accidents and road damage) are considered. The most important gain is the reduction in congestion costs, although fuel savings also become substantial at higher load factors. For load factors in the range of 35 to 55 per cent, the monetary cost reductions vary from £63,950 to £123,410 per annum (1995 prices) for every 100 cars ceasing to travel. Within these totals, however, exhaust emission and road damage costs increase. The increase in exhaust emission costs occurs because buses have relatively high emission factors for those emissions that have the highest health costs, particularly PM\textsubscript{10}. Because of this, policies which reduce emission rates (for PM\textsubscript{10} in particular) through the introduction of new fuel and bus technologies are likely to be the most effective way of reducing the exhaust emission cost changes arising from substitution of bus for car travel.

Changes in exhaust emission regulations over time do not tend to have a significant qualitative effect on these results, since emission limits are progressively tightened for both buses and cars. An exception is for PM\textsubscript{10}, where the introduction of Euro III and Euro IV regulations from 2000 onwards will give further PM\textsubscript{10} reductions for buses (and diesel cars) and a substantial decrease in the exhaust emission costs associated with substituting bus for car travel.

These conclusions are based on an evaluation framework that does not take into account a number of costs associated with car use. Road building causes a loss of countryside, community severance and a pattern of land-use (e.g. out-of-town}

\textsuperscript{11} See previous footnote.
shopping centres) which encourages car-dependency and discriminates against those without access to a car. These costs are not included because of a lack of information concerning their monetary magnitude. This is not to imply, however, that they are unimportant. If quantifiable, inclusion of these costs would strengthen the case in favour of substituting bus for car travel. Other areas for research are identified in Section 7.2.

Finally, the estimates used in this chapter are subject to considerable uncertainty and should not be considered definitive. Hopefully they will stimulate further research in an area which has received limited quantitative attention at the disaggregate level. They also have two important policy implications for any transport strategy that has environmental objectives. First, improvements in fuel and bus vehicle technologies designed to reduce particulate emissions will yield important reductions in exhaust emissions costs. Second, transport policies aimed at increasing bus load factors yield substantial overall cost savings, even if the increase in the load factor is modest.
CHAPTER 8

Conclusions

8.1 Opening Remarks

This thesis began with a description of the lightly regulated local bus service market that existed before the introduction of the 1930 Road Traffic Act, continued with a discussion of the highly regulated system which evolved after 1930, and then moved on to discuss the deregulation of the market. Viewed from this historical perspective, one might be forgiven for thinking of a wheel that has turned full circle. In some ways, this view is not unreasonable. Reading Garcke's (1923) account of the behaviour of different types of bus operator in Section 2.1, it is hard not to be struck by the similarities with some of the facets, actual or alleged, of the deregulated market after 1986.

In other ways, however, much has changed. Perhaps the biggest change in the context of this thesis has been in the development of economic theories used to understand both the outcome and process of competition, as well as the outcome and process of regulation. Accordingly, this thesis devotes two chapters (three and four) to a discussion of these theoretical perspectives and their application to the local bus service market. The evaluation of deregulation is further complicated by the considerable reduction in external subsidy to bus operators that took place between 1985/86 and 1996/97: over this period, external subsidy fell by nearly 60 per cent in real terms. Some commentators have argued that it was a lack of competition after deregulation that caused bus fares to increase, but one could also argue that bus fares increased because of the subsidy reduction rather than a lack of competition. Finally, the transport sector is one that generates a number of externalities. Much work has been done on the monetary valuation of these externalities, but less appears to be published on the externalities generated by substitution between transport modes such as bus and car.
To recapitulate, the aims of this thesis have been to:

- Develop the theoretical perspectives necessary for the evaluation of local bus service deregulation in Britain (excluding London)
- Establish the outcome of deregulation without subsidy reduction
- Estimate the externalities associated with substitution of bus for car travel.

This final chapter structures the overall conclusions of the thesis around these aims. Section 8.2 provides conclusions relating to the choice of theoretical perspective and the implications for the evaluation of deregulation. Section 8.3 comments on the results of the econometric modelling of external subsidy reduction and replacement, and draws conclusions about the changes in social welfare that might have occurred if subsidy had not been reduced. Section 8.4 provides conclusions about the monetary valuation of externalities arising from bus for car substitution. Finally, Section 8.5 gives an overview of the chapter and concluding remarks.

8.2 Theoretical Perspectives

The Public Interest Approach

The highly regulated local bus service system inaugurated by the 1930 Road Traffic Act was a response to problems over safety, "unfair" competition and lack of co-ordination between road, rail and tram interests. Unfair competition was equated with unregulated competition. These problems can be regarded as instances of market failure which, in the public interest approach, can be remedied by regulators acting to increase social welfare. Drawing also on the experience of the post-1986 deregulated market it is possible to identify other instances of market failure that characterise local bus services to varying degrees. These are monopoly, externalities, wasteful competition, unstable competition, unfair competition, loss of integration, public goods and common goods. Unfair competition is here interpreted more narrowly to mean the underpricing of car use.
Regulation on safety grounds is a relatively uncontentious issue. The question is whether there are other instances of market failure that might justify some form of regulation. Chapter three discussed in detail the instances of market failure, both from a theoretical and empirical perspective, and concluded that deregulation *per se* has not resulted in significant instances of market failure. Even if it were the case that significant market failure exists, however, this may still not be sufficient to justify regulation.

For example, the prevalence of merger activity in the deregulated local bus service market is often cited by opponents of deregulation as an indicator of a lack of competition and thus a source of market failure. Using the public interest approach, it can be argued that some form of regulation is preferable to the deregulation outcome. Mackie and Preston (1996) favour a quasi-regulated system in the form of competitive tendering for both profitable and unprofitable services. In other words, competition “in the market” should be replaced by competition “for the market”.

But this emphasis on market structure and market failure is not sufficient to encompass the complex *historical* development of the local bus service market. In particular, two key points about this historical development should be emphasised. First, deregulation was preceded by a regulated system that lasted for over fifty years. Any evaluation of the deregulated system has to take into account the defects of the regulated system that it replaced. This regulatory failure may not be specific to the local bus service market, but may be inherent in any system of regulation. Second, deregulation has been a process characterised by merger activity. To focus on an outcome of deregulation i.e. increased concentration, may run the risk of paying insufficient attention to the process that generated the outcome. Merger activity may be the result of an ongoing process of competition, rather than an impediment to it. These two features of the historical development of the local bus service market necessitate analysis in terms of other theoretical frameworks, namely the public choice, Austrian School and contestability approaches.
The Public Choice Approach

A key question is not whether market failure exists, as stressed by the public interest approach, but whether its costs outweigh those associated with regulation. The public choice approach emphasises the process of decision making and the concepts of regulatory capture and regulatory failure. In particular, it stresses that the welfare costs arising from the monopolies created by regulation may be much higher than those associated with the static allocative inefficiencies identified in Chapter three. It follows that the removal of the protected local monopolies that developed in the regulated local bus service market may give substantial welfare gains.

The costs of regulatory capture and failure are difficult to quantify, but there is evidence to suggest that the regulated local bus service market suffered from a number of deficiencies, particularly “X-inefficiency” and a lack of innovation. Between 1953 and 1980 bus fares increased by 86 per cent in real terms to offset the decline in bus use caused by increasing car ownership, but costs managed to increase faster. External subsidy became necessary, and from 1969 to 1980 total subsidies and grants to the industry jumped from £24 million to £570 million, an increase of over 500 per cent in real terms. If one uses changes in bus fleet structure as a measure of innovation, then the structure of the bus fleet used by operators changed fairly slowly before deregulation, but rather more quickly afterwards. Using the proportion of double deckers to total bus fleet (1953 = 100) as a proxy for changes in bus fleet structure, the index fell by 24% from 1953 to 1986. After deregulation, the index fell by 37% in the much shorter period from 1986 to 1997. The trend towards smaller buses also enabled operators to provide new services in areas such as housing estates that were previously inaccessible to larger buses.

Although “off the road” competition in the form of competitive tendering for unprofitable and profitable services is sometimes advocated in preference to “on the road” competition for local bus services, this solution is also a form of regulation that may be susceptible to the regulatory costs identified by the public choice school. It is unclear how high these costs would be, and the debate about regulatory costs in a competitive tendering...
framework is one that would benefit from further research. It is argued by Mackie et al. (1995) and Mackie and Preston (1996) that competitive tendering has worked well in London and that “what is good for London is good for the country” (1995, p. 249; 1996, p. 206). This seems an over-optimistic view, however, given the very different nature of the London local bus service market to the rest of Britain.

*The Austrian School and Contestability Approaches*

Another fundamental critique of the market failure and public interest paradigm is that of the Austrian School approach. The Austrian School argues that it is not markets *per se* that fail, but rather that failure is the result of entry barriers to the market and an inadequate property rights system. Competition is regarded as a dynamic process in which a Darwinian struggle for survival takes place. Temporary monopoly is a necessary reward for risk-taking in this competitive process. Similarly, mergers are not necessarily anti-competitive, but may be a means of competing via increased cost efficiencies. The key indicator of market competitiveness is not market structure, but market contestability. In a perfectly contestable market, even a natural monopoly must set price and output at levels associated with perfect competition.

An important implication of this view is that the re-emergence of local monopoly, together with increased market concentration via mergers in the deregulated local bus service market, is not incompatible with a competitive market outcome. The actual trend towards fewer but larger firms may also encourage the provision of more services via network effects. Thus a lack of actual competition, which was a feature of the deregulated market, is not necessarily a good indicator of market competitiveness. The key determinant of competitiveness is not actual competition (or the lack of it) but *potential* competition.

Thus the critical question is the contestability of the local bus service market. If it is contestable, then market structure is of secondary importance. If it is not contestable, then market structure is of crucial importance. The issue of contestability is dealt with in some
detail in Chapter four. The Buses White Paper regards the deregulated local bus service market as "highly contestable". In the case of innocent or non-strategic entry barriers such as economies of scale and scope, this is probably a fair assessment. Sunk costs tend to be low, and the multi-output nature of the industry (firms operating both local and long-distance services, for example) increases the possibility of cross-market entry and the degree of contestability.

For strategic entry barriers, particularly those created by predatory pricing, the evidence is less conclusive. In the period 1986 to 1994 the largest number of complaints to the Office of Fair Trading (OFT) about anti-competitive practices in the deregulated market concerned actual predatory pricing. This may understate the problem, since there may also have been threats to use predatory pricing which were not reported to the OFT. Also, market entry is not made easier by the requirement to give 42 days notice of a new service to the Traffic Commissioners: this gives the incumbent operator a reasonable period of time in which to prepare any entry-deterrence response.

The general conclusion from the theoretical literature on predatory pricing, however, is that it is likely to occur under conditions of incomplete information. As time passes and information becomes less incomplete, one would expect the incidence of predatory pricing to reduce. This appears to be the case both in the post-1986 period and the Trial Areas introduced by the 1980 Transport Act. Bus operators engaging in predatory pricing also run the risk of investigation by the OFT. Nonetheless, predatory pricing must be considered an obstacle to market entry, particularly where attempted entry is by smaller operators. The development of quality partnerships and contracts in the deregulated market may also create strategic entry barriers for smaller operators, although much depends on the specific form of the partnership or contract between the local authority and the operator(s). The possibility of strategic behaviour by incumbent operators implies a continuing role for the OFT (or a similar body) in preventing anti-competitive practices in the provision of local bus services.
To summarise the discussion on entry barriers and contestability, the local bus service market is not highly contestable as the Buses White Paper claimed. This is because of the existence of strategic entry barriers, primarily predatory pricing. It would be wrong, however, to leap to the opposite conclusion, namely that the local bus service market has low contestability. Strategic entry-deterrence is probably not sustainable in the long-run, particularly where attempted entry is by large operators with a “long purse”. For these reasons, the market could best be described as relatively contestable. One key piece of evidence that supports this view is the performance of the industry in terms of reducing operating costs. In the decade following deregulation, operating costs per vehicle kilometre fell by forty per cent in real terms. Although some of these cost reductions can be attributed to structural and other changes not related to productivity increases, there is little doubt that productivity related costs “…fell by a large amount” (Mackie et al, 1995, p. 239).

Given this relative contestability, deregulation per se should result in a fall in bus fares and a net welfare gain. The econometric analysis of Chapter six confirms this view, although the welfare gains are not large. Since the industry is not perfectly contestable, however, even non-strategic entry barriers such as economies of scale could produce situations where some degree of monopoly power is exploited. Mergers could also create monopoly power. Given a situation of relative contestability, market structure becomes a factor that should be taken into account when determining whether a monopoly or merger is against the public interest. There is thus a continuing role for the Monopolies and Mergers Commission (MMC) in the deregulated bus industry.

8.3 Subsidy and Deregulation

The single most important issue that confounds any economic evaluation of local bus service deregulation is the large reduction in public transport support to bus operators that accompanied deregulation. If subsidy had not been reduced, it is conceivable that bus fares might have fallen, rather than increased, in the post-1986 period. The results of the econometric analysis in Chapter six show that, with deregulation and subsidy
replacement, bus fares in Britain excluding London would have fallen slightly (see Table 6.1). Given this fall in bus fares, it is possible to calculate the welfare gain from deregulation per se. On the basis of a fairly cautious set of assumptions, the welfare gain (including externality effects from bus for car substitution) is £9.8 million per annum in real terms for the period 1986-97. If the forecast subsidy is assumed not to increase but to remain “frozen” at its 1984 level, then the total welfare gain is reduced to £6.4 million per annum. These gains are small in relation to the total passenger receipts of the industry, but they are gains nonetheless. The annual welfare gains are likely to be higher than £6.4 to £9.8 million, because the calculations in Chapter six do not include the benefits arising from the reduction in Tulloch-Posner costs associated with the demise of the regulated system. The magnitude of these costs is another area in need of research, similar to that mentioned in the previous section.

It is important that the evaluation of deregulation takes into account the very different experiences of the metropolitan and non-metropolitan areas. Further econometric analysis shows that both areas experience a similar welfare gain (excluding external effects) from deregulation with increasing subsidy replacement, with gains of £4.0 million and £3.7 million per annum for the metropolitan and non-metropolitan areas respectively. There is a clear difference between the areas in terms of gain per person, however, with figures of around £4.2 and £1.2 per person for the metropolitan and non-metropolitan areas respectively.

The analysis also forecasts what would have happened to bus fares and passenger journeys per person if the regulated system had been allowed to continue. The forecasts are based on the Box-Jenkins ARIMA methodology rather than the econometric structural modelling used earlier. The results imply that bus fares and passenger journeys in the continuing regulated system would be similar to those that actually occurred with deregulation and subsidy reduction. In other words, the positive effects of deregulation per se are broadly cancelled out by the negative effects of subsidy reduction.
8.4 External Costs and Benefits of Deregulation

It is a commonly held belief that promoting bus use at the expense of car use has desirable economic consequences. This belief is generally valid, although there are some exceptions. The exceptions identified in Chapter seven include a number of vehicle exhaust emissions that tend to increase rather than decrease as people transfer from car to bus.

The previous section summarised the welfare gains from deregulation *per se*. Most of these welfare gains arise because bus fares fall and demand increases when the “lost” external subsidy is replaced. These gains can be characterised as the “internalities” associated with deregulation *per se*. Part of the increase in demand for bus travel occurs because some people substitute bus for car use in response to the change in the relative price of bus to car use. This substitution generates changes in a number of externalities, the main change being the reduction in traffic congestion costs. Although there is an extensive literature on the monetary valuation of transport related externalities, it tends to focus on the overall costs of these externalities rather than the changes that occur as people transfer from one transport mode to another. The effect of modal transfer on total external costs and benefits is not straightforward to establish since, in the case of transfer from car to bus travel, it depends on parameters such as car occupancy rates and bus load factors. In the case of air pollution the situation is more complex, because petrol and diesel engines have different exhaust emission characteristics which are subject to legislative limits that change over time.

The analysis of Chapter seven considers the externalities associated with exhaust emissions, fuel consumption, traffic congestion, noise pollution, road accidents and road damage. It finds that for bus load factors in the range of 35 to 55 per cent, the total monetary cost reductions arising from changes in these externalities in urban areas vary from £639.5 to £1,234.1 per annum (1995 prices) for every car ceasing to travel. The single largest cost reduction is for congestion, although fuel savings also become substantial at higher load factors. Within the 35 to 55 per cent bus load factor range,
however, exhaust emission and road damage costs increase. The increase in exhaust emission costs may surprise some people, but the result is robust to a number of sensitivity tests. The reason for the increase is that bus engines emit relatively large amounts of particulate matter, and this emission has particularly high health costs. The policy implication is that bus for car substitution should be encouraged, but particulate emissions from bus engines reduced by improved engine design and the use of new low-sulphur fuels.

The results above do not take into account a number of costs associated with car use. These include loss of countryside from road building, community severance and the development of land-use patterns that encourage further car dependency. These costs cannot be included because of a lack of research on them. It is possible that they are substantial and, if so, would encourage further the case for promoting bus travel.

Finally, a health warning is in order. Much of the research in this area (if it exists at all) is subject to considerable uncertainty, and the author’s is no exception. Hopefully, the analysis of Chapter seven provides a basis on which further research can build, particularly for the monetary valuation of the benefits and costs of transferring from car to bus travel.

8.5 Overall Conclusions

The local bus service market is a complex one, not only because of the special characteristics of local bus service provision, but also because of its historical development. This complexity means that the welfare gains and losses from local bus service deregulation must be carefully analysed from a number of theoretical perspectives, rather than a single one. Each perspective has its uses and limitations.

The public interest approach focuses on market outcomes, and uses the neoclassical framework of equilibrium and comparative statics to determine the extent of the welfare loss arising from market failures such as monopoly. This approach has conceptual clarity.
and empirical applicability, and is the basic foundation of the quantitative welfare evaluation of deregulation in Chapter six. But the approach does not consider market processes, and may give misleading conclusions if used in isolation. The strength of the public choice and Austrian School approaches is that they focus on regulation and competition respectively as processes evolving over time. Public choice theory emphasises that regulatory capture and failure may be just as much a problem as market failure, whilst the Austrian school highlights the importance of (temporary) monopoly and mergers as a legitimate part of the competitive process. On the other hand, the emphasis on market process as opposed to outcome means that these approaches are not as amenable to the type of quantitative welfare analysis used in Chapter six.

Contestability theory is a more recent development, although its emphasis on entry and exit barriers places it in the Austrian School tradition. In particular, it predicts that even natural monopolies will price competitively if they are in a contestable market. It is the entry and exit barriers to the market, rather than the structure of the market itself, which determines whether or not the monopolist sets competitive prices.

The evidence surveyed in Chapter four suggests that local bus services comprise a market that is relatively contestable. The degree of contestability is not high or low, but somewhere between these extremes, probably nearer the “high” end of the contestability spectrum rather than the low. This conclusion is reinforced by the large fall in bus operating costs per vehicle kilometre after deregulation, in spite of the absence of widespread actual competition.

The welfare evaluation of local bus service deregulation is made more difficult by the concurrent reduction in external subsidy to bus operators. If this lost subsidy is replaced, then econometric analysis shows that there are welfare gains from deregulation per se. The welfare gains arise because bus fares fall and demand for bus travel increases in comparison to the actual outcome associated with deregulation and subsidy reduction. These welfare gains accrue to both metropolitan and non-metropolitan areas, although the gains are small in relation to the total passenger receipts of the industry.
It is likely that these welfare gain estimates err on the side of caution, since they are based on the public interest approach and do not include the gains from the reduction in a number of costs associated with the demise of the regulated system. These are the regulatory costs emphasised by the public choice school. The public interest approach also views monopoly as an instance of market failure, so that the increase in industry concentration after deregulation would be regarded as a source of welfare loss. The Austrian School, on the other hand, would regard the merger activity as a legitimate consequence of the process of competition. Indeed, temporary monopoly for the Austrian School could be regarded as a welfare gain. These remarks emphasise the importance of setting the deregulation debate explicitly within the context of these different analytical frameworks, rather than solely within the public interest approach.

It is possible, however, that competition “for the market” (i.e. a competitive tendering system for both unprofitable and profitable services) might have produced greater welfare gains than the deregulated system introduced by the 1985 Transport Act. This possibility is not explored in detail in this thesis.

The welfare analysis takes into account a number of externalities associated with the substitution of bus for car travel. This substitution causes a large fall in traffic congestion costs, and also falls in the costs of fuel consumption, noise pollution and road accidents. On the other hand, exhaust emission and road damage costs increase. The main culprit for the increase in exhaust emission costs is particulate matter, which has high health costs and is emitted in relatively large quantities by bus diesel engines. Overall, however, external costs decrease as a result of bus for car substitution. Policy makers should seek to continue to encourage people to travel by bus instead of car, but ensure that particulate emissions are reduced to much lower levels.

Finally, the discussion of the previous sections has drawn attention to the need for further research in a number of areas. First, the regulated local bus service system in Britain was subject to regulatory failure. At least one study has been made of the monetary costs of regulatory failure in other transport related markets, but no estimates have so far been
made for the local bus service market in Britain. Given the importance attached by the public choice and Austrian School paradigms to the problems of regulation, this area of research is worth pursuing in the context of a system that lasted for over fifty years. Second, and on a related theme, there is a need for research into the costs and benefits of competitive tendering systems for profitable services in the context of the deregulated local bus service market. The avoidance of wasteful competition by encouraging “off the road competition” through competitive tendering is an appealing idea, and could result in welfare gains greater than those achieved from deregulation with “on the road” competition. But this is a speculative view, and a proper evaluation should take into account the monetary costs, as well as benefits, of a competitive tendering system. Third, there is considerable scope for refining and developing existing research work on the externalities associated with bus for car substitution. Given the emphasis by policy makers on the need to reduce externalities such as traffic congestion, air pollution and global warming, there is a surprising lack of detailed research on the monetary values that should be used in the evaluation of bus for car substitution.

To summarise, the main conclusions of this thesis are that:

- Local bus service deregulation cannot be fully evaluated in terms of the public interest approach alone
- The deregulated local bus service market is relatively contestable, although there is still a need for supervision by the OFT and the MMC
- Deregulation resulted in a welfare gain in both metropolitan and non-metropolitan areas, although the gain is relatively small
- Broadly speaking, the positive effects of deregulation per se have been cancelled out by the negative effects of subsidy reduction
- Policies to promote bus for car substitution will result in an overall increase in welfare
- Further research is desirable in three areas: the costs of regulating bus services, the costs and benefits of competitive tendering, and the monetary valuation of externalities.
REFERENCES


Bleasdale, C. 1983. *The effects of deregulation on Inter-City rail services*. PTRC Summer Meeting, Volume L.


Department of Transport. (various) *Allocation of Road Track Costs*. London: HMSO.


Department of Transport. (various) *Road Accidents Great Britain*. London: HMSO.


282


286


Intergovernmental Panel on Climate Change (IPCC), 1990. Bracknell: Meteorological Office (UK).


Liebenstein, H. 1980. Microeconomics and X-Efficiency Theory: if there is no crisis, there ought to be. *Public Interest* (Special Issue) 97-110.


Transport and Road Research Laboratory, 1980. *Factors affecting the Demand for Bus Services*. Crowthorne: TRRL.


