Measuring the Impact of Regional Projects: Developing an Evaluative Framework for Regional Economic Development

Thesis in Candidacy for the Degree of Doctor of Philosophy at the University of Glamorgan

Dafydd Williams, Part Time Student 99056666
dafydd.williams@wales.gsi.gov.uk
029 2089 8894

October 2007
Abstract

This thesis proposes solutions to assist agencies in optimising their contribution to regional development. It develops a technique of project appraisal that is both theoretically sound and of practical use. The objective is to help policy makers and practitioners to appraise, monitor and evaluate regional policy projects.

A review of theory and practice examines how various schools of economic thought assess the case for regional economic policy. The thesis argues that the context for evaluation has grown more complex and demanding, as regional policy moves away from simple job creation and tackles a greater variety of goals and objectives.

A taxonomy of evaluative techniques, both macro- and micro-economic, is presented. Particular attention is paid to cost-benefit analysis which, despite distributional, ethical and measurement difficulties, provides a consistent ranking of projects. The thesis describes how evaluators judge the true, ‘additional’ impact of intervention, and how this may be tracked by indicators and targets.

In the context of regional policy, project appraisal requires calculation of grant deadweight, displacement and substitution, multiplier effects and leakage from the target area. Current appraisal processes in the UK are compared with the more rigorous procedures of the Republic of Ireland.

The next stage in the thesis is the specification of an economic model, which assesses benefits and costs and produces two evaluative measures (Net Present Value and a Benefit-Cost Ratio) that can guide agencies faced with a budget constraint. This model is used to compare different project datasets, using data supplied by the former Welsh Development Agency and a sensitivity analysis is performed.

The thesis shows that outcomes are strongly affected by variations in project wages and in the shadow wage. It sets out implications for theory and makes a number of recommendations for policy and practice, including specific proposals for stakeholders in Wales.
Regional economic policy is big business. Governments invest considerable resources in its success and the number of regional development agencies has grown significantly in recent decades. Side by side with this expansion of policy, academic study of regional economic theory and policy has also grown in scope.

Policy makers and academics alike agree on the importance of evaluating the success or failure of both theory and policy. In the United Kingdom, detailed guidance is provided on how organisations should appraise the likelihood of favourable outcomes, monitor their progress and judge their effectiveness once complete.

Strangely, however, there is no evidence of agreement on the precise technique that should be used to achieve these goals. This means that there is relatively little opportunity to compare the effectiveness and efficiency of initiatives pursued in different regions, even where similar projects are being studied.

This lack of a common standard of measurement is all the more curious, given the emphasis placed by the European Commission on the need to evaluate the impact of its regional policy.

The thesis records the wide variety of ways in which UK ministries, agencies and devolved administrations interpret the evaluative advice provided by HM Treasury. It contrasts this with the more rigorous approach pursued in the Republic of Ireland.

This thesis explores the potential for developing a model that helps policy makers and practitioners to make the most of the resources at their disposal. The research question is defined as a search for the model that best helps agencies to optimise their contribution to regional development. It examines the concept of the region and its significance in economic theory and policy, considering the understanding of mainstream economic theory but also the critical contribution of spatial analysis and evolutionary economics. It then reviews how regional policy has evolved, and how in
turn techniques of appraisal and evaluation have been developed to assess their usefulness.

Building on this experience, the thesis then argues for development of a project appraisal technique based on cost-benefit analysis, which it suggests is best able to provide a method that is both theoretically sound and of practical use in appraising, monitoring and evaluating regional policy actions. It recognises that this device would be used as part of a wider evaluative process, and that projects with long-term or indirect outputs and outcomes pose particular measurement problems.

The thesis proposes a model intended principally for use in ex ante appraisal, although it is envisaged that the technique can also be adapted for ex post use. The model is designed to generate estimates of the Benefit/Cost Ratio produced for each project, while recognising that consideration of absolute levels of benefits and costs (expressed as Net Present Value) is necessary for use with agencies facing a budget constraint.

The thesis applies the model to a number of projects using data collected by the former Welsh Development Agency in its process of project appraisal. The projects are chosen for demonstration purposes rather than to arrive at conclusions between various economic initiatives. Benefits and costs entering the model are spread over time are converted to a net present value by applying a rate of discount of 3.5% in line with current UK Treasury guidelines.

The aim is to develop an instrument for use in deciding allocation of resources between competing projects. The model could therefore form a useful part of an evaluative framework. Use of the model to arrive at judgements on the relative merits of different initiatives is beyond the scope of the thesis but clearly this would be an application of great potential.

In the final chapter a number of conclusions are offered, together with recommendations for further development of policy and practice.
Acknowledgements

I gratefully acknowledge the generous help I have received during work on this thesis.

First and foremost, I thank my supervisor Professor Stephen Hill of the University of Glamorgan Business School for his never failing support and encouragement, including his great talent to inspire at times of crisis or occasional gloom. Professor David Pickernell kindly acted as deputy director of study at the University, and provided most helpful information.

I also thank Professor Mike Danson of the University of Strathclyde and Professor David Brooksbank of the University of Glamorgan for acting as assessors, and together with Professor Michael Sheehan making my viva interview an enjoyable as well as a challenging occasion.

Dr Max Munday of Cardiff Business School kindly reviewed my transfer document, and supplied valuable comments and suggestions which I have sought to implement. Kim Swales of the University of Strathclyde readily made time to discuss earlier drafts of a related paper co-authored with Steve Hill, and provided a valuable critique of our conclusions. As frequent references show, his work in many ways provided the launch pad for the core analysis of UK evaluative methods.

I also acknowledge with thanks the support of colleagues at the Welsh Assembly Government and the former Welsh Development Agency, especially Simon Jenkins, Gareth Hall, Gareth John and my former line manager Helen Butler, who made possible access to the time and resources required to undertake this study.

I thank all colleagues who willingly provided guidance and practical advice – especially Eleanor Knight and the WDA economics team James Price, James Owen and Ben Stokes. Justine Mabbs helped me navigate the workings of the Agency’s performance management system, Michelle Hillman afforded technical assistance while Rob Lasinki opened my eyes to the potential of Excel.

Despite all this help, errors and omissions are the sole responsibility of the author.
# Contents

Abstract i  
Preface ii  
Acknowledgements iv  
Contents v  
List of Tables xiv  
List of Figures xiv

## 1 Introduction

1.1 Background 1  
1.2 Research problem 2  
1.3 Need for research 3  
1.4 Research methodology 5  
1.5 Outline of thesis 6  
1.6 Definitions 12  
1.7 Delimitation of scope of thesis and key assumptions 13  
1.8 Conclusions 14

## 2 Context — the Region and Regional Economics

2.1 Introduction 15  
2.2 The region and regional economics 15  
2.3 Delimiting regions 17  
2.4 Regional economic theory and policy 19  
2.5 The theoretical base of regional science and policy 20  
2.6 Methodology of the thesis 22  
2.7 Counterfactual analysis 23  
2.8 Conclusions 25
3 'Spaceless' regional economic theories

3.1 Introduction 26
3.2 Taxonomy of Spaceless Regional Economic Models 27
3.3 Supply-based Models 28
3.3.1 Traditional Neoclassical Models 28
3.3.2 Modified Neoclassical Model 30
3.3.3 Endogenous growth model 32
3.3.4 Catch-up model and Convergence 33
3.3.5 Human Capital and Embodied Knowledge 37
3.3.6 Supply-based Models: Summary 39
3.4 Demand-based Models 40
3.5 Cumulative Causation models 42
3.6 Conclusions 46

4 Spatial Theories

4.1 Introduction 48
4.2 Development of spatial theory 49
4.3 Taxonomies of spatial models 51
4.4 Classical and neoclassical location theories 52
4.4.1 Land Use Models 53
4.4.2 Company Location Models 53
4.4.3 Classical and Neoclassical Location Theories: Summary 54
4.4.4 Central Place Theory 55
4.4.5 Market Area Models 56
4.5 Behavioural theories of location 59
4.6 Clustering and agglomeration 61
4.6.1 Industrial Districts 62
4.6.2 Growth Pole Model 62
4.6.3 Incubator Model 63
4.6.4 Product Cycle Model 63
4.6.5 Porter Cluster Model 63
4.6.6 New Industrial Areas Model 66
11.6 Selection of indicators
11.6.1 Employment as indicator
11.6.2 Other indicators: Gross Value Added, wages, turnover
11.7 Multi-faceted Programmes
11.8 Setting Decision Rules
11.9 Evaluative Framework
11.10 Appraisal Process
11.11 Determining Additionality
11.11.1 Net Grant Deadweight
11.11.2 Displacement
11.11.3 Substitution
11.11.4 Multiplier Effect
11.12 Technical Problems
11.12.1 Selection bias
11.12.2 Optimism bias
11.12.3 Double counting
11.12.4 Community Economic Development Programmes
11.13 Shadow Pricing
11.14 Conclusions

12 Criteria for Appraisal and Evaluation

12.1 Introduction
12.2 Taxonomy of Appraisal and Evaluation Criteria
12.3 Cost-effectiveness indicators
12.3.1 Gross cost per direct (gross) job (GCPGJ)
12.3.2 GCPGJ using ‘Job Years’ (not discounted)
12.2.3 GCPGJ using Permanent Job Equivalents
12.3.4 Discounted tests
12.3.5 Present value job years
12.3.6 Gross Cost per Net Job
12.3.7 Net Cost per Net Job
12.4 Cost-benefit analysis tests
12.5 A Cost-Benefit Critique: the Swales Model
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.5</td>
<td>General implications for policy and practice</td>
<td>311</td>
</tr>
<tr>
<td>17.6</td>
<td>Specific implications for stakeholders in Wales</td>
<td>314</td>
</tr>
<tr>
<td>17.7</td>
<td>Implications for further research</td>
<td>317</td>
</tr>
<tr>
<td>17.8</td>
<td>Final Conclusions</td>
<td>318</td>
</tr>
</tbody>
</table>
List of Tables

1. Structure of thesis 7
4.3 Approaches to Spatial Theory 52
4.8 Forces affecting industrial concentration or dispersion 73
6.7.1 Regional Micro-Policy Options 111
6.7.2 Regional Macro-Policy Options 112
9.3 Classification of Evaluative Methods 157
9.3.1 Shift-share approach 160
9.6 Cost-benefit Analysis Procedure 170
11.4.1 The Impacts Problem 199
11.4.2 The Impacts Problem (Modified) 200
11.5 Hierarchy of outcomes, outputs, and targets 202
11.9 Principles of an Evaluative Framework for Regional Development 209
11.11 Assessing Net Additionality 218
12.2 Classification of Appraisal and Evaluative Methods 224
15.3.1.1 @Wales Digital Media Initiative (Model Outturns) 291
15.3.1.2 @Wales Digital Media Initiative (WDA Value Added Predictions) 292
15.3.2.1 Accelerate Wales Clusters (Model Outturns) 294
15.3.2.2 Accelerate Wales Clusters (WDA Value Added Predictions) 295
15.3.3.1 Business Process Excellence (Model Outturns) 295
15.3.3.2 Business Process Excellence: WDA Value Added Predictions 296
15.4 Project Rank Comparison 298
16.2.4 Sensitivity analysis on parameters 303
15.A.1 Changing the Data (Appendix 15) 325
15.A.2 Projects Appraised by Cost-Benefit Model (Appendix 15) 325

List of Figures

4.6: Porter’s Competitive Diamond of Local Industrial Clustering 65
11.10 Department for Education and Skills Appraisal Checklist 211
13.2.1 Department of Work and Pensions two-tier cost benefit framework 246
PART 1 PROLOGUE

Chapter 1 Introduction

1.1 Background
1.2 Research problem
1.3 Need for research
1.4 Research methodology
1.5 Outline of thesis
1.6 Definitions
1.7 Delimitation of scope of thesis and key assumptions
1.8 Conclusions

1.1 Background

This thesis seeks to contribute to the understanding of appraisal and evaluation of regional economic policy. It investigates how the impact of regional development projects can best be measured, and whether such techniques can be applied world-wide to a variety of regions facing different economic challenges.

This topic has occupied growing attention in the wake of the increasing importance of regional economic development issues, the establishment of development agencies in many parts of the world and moves towards devolution in the United Kingdom. For example, it is estimated that regional policy will account for 36 per cent of the European Union’s budget agreed for 2007-2013, and by the end of the period regional spending will exceed that on the Common Agricultural Policy (Giordano, 2006: 3).

This thesis contributes to the state of knowledge about appraisal and evaluation by establishing a practical method for appraising and evaluating a range of economic development projects. Given the substantial spending and expectations involved in regional policy, this contribution is significant for practical as well as academic reasons.
As such, it represents a significant addition both to the underlying theory and the practice of project appraisal and evaluation.

1.2 Research problem

This thesis produces a simple and transparent model to help appraise and evaluate regional development projects. It fills a noticeable gap in the literature concerning the application of regional science to decision taking by development agencies and policy makers.

The research problem involves selection of a project appraisal technique that is both theoretically sound and of practical use, and assists in monitoring and evaluating regional policy:

What model best helps agencies to optimise their contribution to regional development?

This thesis investigates various systems currently used, especially those in the United Kingdom and the Republic of Ireland. It produces a model based on cost-benefit analysis to guide agencies in allocating scarce resources. It should be noted that the model assesses the impact of regional development projects rather than policies, and may be used for ex ante appraisal as well as ex post evaluation. The model is then used to study output data produced by the former Welsh Development Agency, now part of the devolved government administration in Wales. It concludes that there are significant gains to be made by introducing an effective international standard system of appraisal and evaluation that would enable comparison across a range of programmes and projects (17.6.2).
1.3 The need for research

The study is timely because of increasing interest in the theoretical and practical problems associated with evaluation of economic policies and the potential for the process of evaluation to add value to the regional policy cycle (Batterbury and Hill, 2003).

The need for research is indicated by:

0. The importance of regional economic development policy, indicated by the UK Government’s commitment of £2.3 billion for English RDAs in 2007-08 (HM Treasury, 2004: 174);
0. The lack of any systematic criticism of current appraisal and evaluative techniques in the United Kingdom since the contributions of Swales (1997a and 1997b) despite major changes during the last decade (HM Treasury, 2003);
0. The need for agreement on a clear and practical methodology of appraisal and evaluation that can assist practitioners and governments to compare and contrast different policies, programmes and policies, if possible across international boundaries;
0. The potential for securing beneficial outcomes for the design and delivery of regional economic development as a result of more precise and realistic measurement of public sector intervention.

A considerable volume of publications advise on methods of appraisal and evaluation. Guidelines have been produced both in the UK and Irish Republic, while regional development agencies have described their evaluative processes (Hill, 2001b). Yet there is relatively little critical analysis in the public domain of how appraisal and evaluation techniques are actually put into practice, nor of their comparative effectiveness, nor to what extent they influence agencies’ business planning. One particular gap is the lack of critical comparison of the different methodologies recommended for agencies in the United Kingdom (HM Treasury, 2003; DTI, 2003; English Partnerships, 2004) and the
Republic of Ireland (Forfás, 2003). Eurada, the association of European development agencies, has produced a paper on evaluation methods (Eurada, 2005) but this has not been followed up and no model has been developed to assess the added value of projects (Saublens, 2007).

A model that provides a standardised approach to assessing the effectiveness of individual projects has been developed on behalf of the English RDAs, this remains developmental and so has not been released to the public domain. Nor does it provide a means of comparing outputs (OffPAT, 2006: 13; Gibb, 2007). Scottish Enterprise operates a well-established evaluative framework, but no longer uses economic models to measure individual project impacts. The Welsh Development Agency, established in 1976, has developed a model that compares the outputs of some projects. Examination of current practices also raises a number of issues about their theoretical backdrop (Hill and Williams, 2005) (10.3.1).

While the thesis examines the theoretical framework relevant for regional economic policy, the emphasis is firmly on producing practical (if imperfect) guidelines for regional economic policy: the watchword is ‘better to be roughly right than precisely wrong’! The objective therefore is to develop a technique that not only measures but also helps maximise the impact of policies. The study draws on recent experience of regional economic policy practice within Wales, comparing this with techniques used elsewhere. The lessons are intended for general application.

The principal objectives of the research project are:

0. To develop a project appraisal model based on an explicit programme theory;
0. To assess the usefulness of the model as a measuring device and potential decision guide for international comparisons;
0. To decide whether it may be possible to use the model to derive preliminary judgements on the contribution of widely differing projects to prosperity.
1.4 Research methodology

The main focus of the research project is development and application of an appraisal technique that assists in making the best use of scarce public resources and is of practical use for practitioners and policy makers. Development of the research project requires an assessment of the state of knowledge about economic growth and development; of regional economic policy; and of existing techniques of appraisal and evaluation.

As argued in the preceding section, despite extensive research, there remains scope for further development and clarification of methodology. The thesis seeks to fill this lacuna by investigating the current state of knowledge about regional economic theory and policy; and by developing an appraisal technique that will assist agencies in deciding how best they can use their scarce resources. After considering the various techniques developed to measure the impact of policy, the thesis constructs a model based on cost-benefit analysis that provides a technique for comparing projects and programmes that fall within limits determined by availability of ‘hard’ output data.

This model is then applied to data of outputs available through the management information service (PerMIS) introduced by the Welsh Development Agency (WDA, 2005: 8), now incorporated within the Welsh Assembly Government as part of an enlarged Department for the Economy and Transport. It is emphasised that the thesis produces a model different to the one used by the WDA. The thesis model is used to generate values of the Benefit/Cost Ratio and absolute Net Present Value to produce rankings of projects and programmes, with the aim of helping assess the effectiveness of different regional development projects by providing a common standard of comparison.

As set out in greater detail in following chapter (2.7) the research approach follows reflects the positivist approach, the belief that the world is external and objective and can therefore be measured (Remenyi et al., 1998: 104). This does not however imply that all impacts of regional development projects can be accurately and consistently measured. Chapter 8 provides an account of the difficulties encountered in assessing the effects of
many projects, notably those involving long-term or indirect economic effects. The thesis follows a pragmatic (rather than a purist or dialectical) approach to the philosophical basis of evaluation (Patton, 1988; Smith, 1994), as detailed below (8.3.1).

1.5 Outline of thesis

The structure of the thesis is set out in Table 1. The thesis is organised in seven parts – entitled Prologue, Theoretical Framework, Policy Framework, Methodological Framework, Parameters for a New Approach to Evaluating Regional Policy Projects, Model Specification, and Analysis and Conclusions. These are subdivided into chapters, and within these chapters, sections and subsections. Cross-referencing is provided throughout the text by numbering sections and subsections (e.g. this section is 1.5). The numbering of figures, tables and appendices follow the numbering of the section to which they refer.

The rest of this section provides brief detail of the content of each part and chapter in turn. The following sections of Chapter 1 sets the scene by noting some key definitions, delimiting the scope of the thesis and providing a summary of conclusions.

Chapter 2 sets the context for the thesis by outlining the concept of the region and its significance in economic theory and policy. It looks at the way in which science develops and tests theory, and the implications for the various theoretical approaches to regional development. This involves a brief survey of the nature of theory and of evidence-based scientific enquiry, emphasising the practical imperative that a ‘good theory’ should both explain and predict. This treatment includes particular reference to the difficulty encountered in establishing the counterfactual, a familiar problem confronting practitioners and linked to the concepts of deadweight expenditure and displacement (Hill and Williams, 2005). This consideration underlines the importance of calculating net impacts rather than subsequent effects, i.e. what difference did the intervention make (or will it make) to the performance of individuals, communities and businesses (Hill, 2001b: 3). The focus of this discussion is on solutions – what insight
can theories bring to understanding the underlying causes of disparity between regions and whether they can be tackled.

Table 1: Structure of thesis

<table>
<thead>
<tr>
<th>Number</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front pages</td>
<td>Summary, Declaration, Acknowledgements, Contents, List of Tables, List of Figures, List of Appendices</td>
</tr>
<tr>
<td><strong>PART 1</strong></td>
<td><strong>PROLOGUE</strong></td>
</tr>
<tr>
<td>Chapter 1</td>
<td>Introduction</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>Context</td>
</tr>
<tr>
<td><strong>PART 2</strong></td>
<td><strong>THEORETICAL FRAMEWORK</strong></td>
</tr>
<tr>
<td>Chapter 3</td>
<td>'Spaceless' regional economic theories</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>Spatial economic theories</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Temporal dimension of economic theories</td>
</tr>
<tr>
<td><strong>PART 3</strong></td>
<td><strong>POLICY FRAMEWORK</strong></td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Regional Economic Policy</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>Regional Economic Theory and Policy Synthesis</td>
</tr>
<tr>
<td><strong>PART 4</strong></td>
<td><strong>METHODOLOGICAL FRAMEWORK</strong></td>
</tr>
<tr>
<td>Chapter 8</td>
<td>The Challenge of Evaluation</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>Techniques of Evaluation</td>
</tr>
<tr>
<td>Chapter 10</td>
<td>Cost Benefit Analysis in Policy Appraisal and Evaluation</td>
</tr>
<tr>
<td><strong>PART 5</strong></td>
<td><strong>PARAMETERS FOR A NEW APPROACH TO EVALUATING REGIONAL POLICY PROJECTS</strong></td>
</tr>
<tr>
<td>Chapter 11</td>
<td>Process of Appraisal and Evaluation</td>
</tr>
<tr>
<td>Chapter 12</td>
<td>Criteria for Appraisal and Evaluation</td>
</tr>
<tr>
<td>Chapter 13</td>
<td>Appraisal Methods in the UK and Republic of Ireland</td>
</tr>
<tr>
<td><strong>PART 6</strong></td>
<td><strong>MODEL SPECIFICATION</strong></td>
</tr>
<tr>
<td>Chapter 14</td>
<td>Defining the Model</td>
</tr>
<tr>
<td><strong>PART 7</strong></td>
<td><strong>ANALYSIS and CONCLUSIONS</strong></td>
</tr>
<tr>
<td>Chapter 15</td>
<td>Analysis of Data</td>
</tr>
<tr>
<td>Chapter 16</td>
<td>Sensitivity Analysis</td>
</tr>
<tr>
<td>Chapter 17</td>
<td>Conclusions</td>
</tr>
<tr>
<td>End pages</td>
<td>Appendices</td>
</tr>
</tbody>
</table>
| References | }
Part 2 of the thesis, the Theoretical Framework (Chapters 3-5), presents a review of literature dealing with regional economic theory. This presentation identifies aspects of theory and empirical evidence most relevant for determining regional economic policy, and lays the groundwork for development of the most appropriate method of appraisal. In line with the proposals of Emory and Cooper (1991: 122), it records important conclusions of other studies and particular methods or designs that could be duplicated or should be avoided.

The review of the theoretical framework includes a general survey of economic theory that seeks to account for growth and geographic variations in prosperity. It sets out the theoretical basis for regional policy, and in particular the justification provided by theory for publicly funded and directed intervention. Largely for convenience in dealing with the considerable volume of literature involved rather than any rigid analytical differences, the presentation of theory is divided into three parts (Chapters 3, 4 and 5).

Chapter 3 provides a summary of 'spaceless' theories of economic development and trade, a field dominated by the neoclassical school, but also encompassing models derived from Keynesian and other scholars. These theoretical models are intended to explain variations in productivity, specialisation and migration between different regions and nations. A taxonomy of spaceless theories is presented, distinguishing:

- Neo-classical (decreasing returns) models;
- Demand- or export-based (increasing returns) models;
- Endogenous growth or 'cumulative causation' models; and
- Evolutionary approaches.

The treatment identifies the approaches felt to be of the greatest relevance for regional economic policy makers, such as cumulative causation and cluster concepts.

Chapter 4 examines various attempts made since the nineteenth century to integrate geographic space within a framework of economic analysis, a treatment that includes
consideration of the importance of clusters for economic growth. Chapter 5 considers the role of time in economic theory, with particular consideration for criticisms made of orthodox analysis by advocates of ‘evolutionary economics’ and others. It is argued that integration of space and time, especially the influence of past events, reduces the likelihood of free markets achieving equilibrium in line with standard neoclassical assumptions.

One important theme of this framework concerns the relevance of theoretical models put forward to explain the causes of economic growth of regions and countries, and hence disparities between them. One major topic is whether regional economies converge. If, as some orthodox theorists contend, less prosperous regions will catch up with more prosperous areas, there is little point in spending resources on regional economic policy. Conversely, the case for intervention becomes more significant where spatial issues and path dependence reduce the likelihood of free markets ensuring convergence.

Part 3 turns to regional economic policy, with Chapter 6 tracing the main developments within the United Kingdom during the 20th and 21st centuries. This treatment is brief: the thesis makes no attempt to duplicate the ample literature on policy (Taylor and Wren, 1997; Armstrong and Taylor, 2000; MacKay, 2003; Fothergill, 2005) but rather aims to draw out a number of themes relevant for appraisal and evaluation, such as clustering, innovation and entrepreneurship. These are topics that have featured in contemporary economic policy in Wales, where government support for business start-ups is especially prominent, with 12.8 per cent of new starts accessing support compared with 8.4 per cent for the UK (Brooksbank and Jones-Evans, 2004: 6). Other measures to assist small business in Wales have included the simplification of application procedures for Regional Selective Assistance, reduction of the regulatory burden, and support for indigenous SMEs to gain more business from the public sector (Evans and Pickernell, 2002).

Chapter 7 briefly draws together the main themes to emerge from this review of theory and policy, focussing on the case for intervention and the need to weigh up its costs and benefits. It also provides a brief survey of the development of regional policy in the UK.
Part 4 surveys various methodologies applied by agencies in appraisal and evaluation. Chapter 8 sets out the framework for appraisal and evaluation, and considers the conflicting demands evaluators face from different agents.

Chapter 9 examines the main analytic techniques, both macro- and micro-economic, which are used to assess the effectiveness of regional economic policy. It proposes a taxonomy of evaluative methods, based on the approach of Vanhove (1999: 387-406), and notes the apparent consensus favouring the use of cost-benefit analysis (Taylor, 2002). In line with the emphasis placed on explicit recognition of underlying programme theory, this treatment precedes consideration of both the 'mechanics' of evaluation (in Chapter 11) and the various criteria employed (in Chapter 12).

This is followed in Chapter 10 by a more detailed critique of cost-benefit analysis, which highlights a number of reservations, especially in the way it approaches issues of time (and therefore sustainability and the interests of future generations). This theme is pursued in dealing with the approach adopted before 2003 by the UK Treasury to interregional displacement (the doctrine of '100 per cent crowding out' at a UK level) which provided some of the momentum for academic debate during the 1990s (Swales, 1997a); and in the apparent conflict in the treatment of 'tax deadweight' by government economists in the UK and Ireland (Hill and Williams, 2005).

Part 5 develops a fresh approach to the process of appraisal and evaluation, setting out the parameters for elaborating a cost-benefit based methodology. Chapter 11 examines the practical steps required to undertake appraisal and evaluation. This includes setting objectives, defining the precise techniques used and distinguishing truly 'additional' impacts of programme or policy studied; which in turn requires identification of targets, indicators and procedures to be used. This treatment considers the issue of focusing tests on defined objectives, such as growth of long term GVA per capita (Pickernell, 1999: 29).
Chapter 12 moves from process to underlying theory, examining the main types of criteria used or suggested to appraise, monitor and evaluate projects and policies. It begins by considering the simplest criterion, and proceeds by introducing progressively greater refinements – for example, by converting gross benefits and costs into net benefits and costs. This includes a number of methods based on cost benefit analysis, involving use of the concept of shadow pricing to take account of real resource cost to the economy.

Chapter 13 examines examples of evaluative practice in a number of UK departments and agencies, contrasting them with the more rigorous approach based on cost-benefit analysis used in the Republic of Ireland. It is noted that appraisal and evaluation methods based on monetized outputs encounter difficulties in assessing the value of long-term ‘capacity building activities’ such as urban and rural regeneration and community economic development. The thesis briefly considers methods adopted to deal with this problem. The aim of this survey is to draw out issues of relevance to the research question, not attempt a thorough study of the wide range of procedures discovered.

Part 6 deals with specification of the model and its sensitivity to assumptions and parameters. In Chapter 14, a model is developed for use by practitioners of regional economic policy, principally as an instrument of ex ante appraisal, although it is envisaged that the technique can also be adapted for ex post use. The model generates estimates of the Benefit/Cost Ratio produced for each project, while recognising that consideration of absolute levels of benefits and costs (expressed as Net Present Value) is necessary for use with agencies facing a budget constraint.

Part 7 is concerned with analysis of data and conclusions. Chapter 15 applies the model to a number of projects using data collected by the Welsh Development Agency in its appraisal process. The model produces ex ante estimates of the Benefit-Cost Ratio and Net Present Value. In the case of one project, for example, it is estimated that a net benefit to the economy of between 2.6 and 4 times the real resource cost involved, a
conclusion which however depends on the soundness of assumptions made, especially the accuracy of the raw output data estimated by project managers.

Chapter 16 investigates how the model responds to changes in its underlying assumptions and its various parameters. This is explored by a process of sensitivity analysis carried out on data of the project selected for use as a pilot, the @Wales Digital Media Initiative.

Chapter 17 sets out the conclusions arrived at by the thesis, together with recommendations for future actions and research.

1.6 Definitions

Definitions adopted by academic and government sources are not always uniform. This brief section therefore defines a limited number of the terms used in the body of the work – generally those that have involved contradictory treatment in the literature – where this assists understanding in advance of the area covered by the thesis. Most definitions appear throughout the text as required.

One issue that requires advance definition is the nature of evaluation, which is this thesis is interpreted in its broad academic sense as “the process of systematically and carefully assessing the value any action, policy or programme, and of capturing policy learning in a format that has intrinsic utility for subsequent policy formulation” (Batterbury and Hill, 2002:1). As discussed in 8.2, official literature tends to restrict the definition of evaluation to the measurement of past (ex post) events. Thus one UK ministry defines evaluation as the “retrospective analysis of a programme or policy to assess how successful it has been, and what lessons can be learnt for the future” (Department of Work and Pensions, 2005). Advance, ex ante evaluation is described as ‘appraisal’ – “the process of defining objectives, examining options and weighing up the costs benefits, risks and uncertainties of those options before a decision is made” (Department of Work and Pensions, 2005).
Project or option appraisal is defined as “the appraisal of multiple options chosen to achieve a specific objective” (ODPM, 2004b: 169). It is envisaged that project appraisal will be the principal use of the model developed in this thesis, although it may also be applied to monitoring and ex post evaluation.

1.7 Delimitation of scope of thesis and key assumptions

The research question defined in 1.2 above is framed in terms of seeking a model that guides agencies in optimising their contribution to regional development. This model should in principle be applicable internationally. Some reference is made to experience in Europe and North America. However, the focus of this thesis is on the various methods of appraisal and evaluation used in the United Kingdom and the Republic of Ireland; and the project data studied is that available in Wales.

The choice of research method was accommodated by the ready accessibility to data collected across the Welsh Development Agency (now incorporated within the Welsh Assembly Government as the Department for the Economy and Transport), and focuses on desk-based analysis. No use is made of survey, focus group or questionnaire techniques, although the thesis points to areas where its conclusions could be validated or further investigated by survey techniques. Further research could broaden this approach to methods and projects used in jurisdictions other than the UK and the Irish Republic.

The thesis also recognises the limitations that face such techniques, for example in seeking to measure projects and programmes that vary widely in their nature and in the time they take to bear fruit. This limitation means that projects appraised according to their contribution to ‘capacity building’ are not susceptible to measures of value added. As the thesis records (in 13.2.5), this limitation means that most of the Agency’s projects fell outside measurement by value added.
Other detailed assumptions used in framing the model which arise from the chosen methodology are set out in 14.6.

1.8 Conclusions

This introductory chapter has laid the foundations for this thesis by introducing the research problem and the research question, defined in 1.2.

The need for research into this issue was set out in 1.3. It rests on the significant scale of spending on regional policy and the surprising lack of comparison in published literature of the differing approaches to appraisal in the UK and Ireland.

A brief description of the chosen methodology, which follows a positivist approach and is based on cost-benefit analysis, appeared in 1.4. A full justification of this approach is set out in Chapter 9.

The chapter also presents an outline of the thesis (1.5) together with a brief description of the line pursued in each chapter.

In 1.6, a number of definitions were provided of concepts such as appraisal and evaluation that required explanation at the outset. However, definitions appear throughout the text, especially in Chapter 14 where the elements of the model are specified.

A number of key assumptions that underlie the overall approach of the thesis were set out in 1.7. Other assumptions relating to the model appear in Chapter 14.
Chapter 2  Context — the Region and Regional Economics

2.1  Introduction

This chapter lays out the groundwork for the theoretical foundation of the research project in Part 2. Sections 2.2 and 2.3 examine the concept of the region in economics, while 2.4-2.6 consider the nature of scientific inquiry, with particular focus on theories that guide development of policy. Section 2.7 specifies the methodology adopted for the thesis, which follows the positivist paradigm. Section 2.8 examines the nature of counterfactual analysis, an important aspect of economic theory and empirical enquiry.

A number of tentative preliminary conclusions are offered in Section 2.8.

2.2  The Region and Regional Economics

A survey of regional economic theory might be expected to begin by defining exactly what is meant by a ‘region’. Such definition is not as straightforward as might be supposed.

Richardson (1978: 19) described the ‘classical’ spatial method of conceptualising regions, a threefold division distinguishing:
Homogeneous regions, defined by unifying characteristics, where internal
differences and intra-regional interactions are considered unimportant;

Nodal (or polarised) regions, with little concern for uniformity but where the
cohesiveness is the result of flows, contacts and interdependences, usually
polarised towards one dominant centre or node;

Planning (or programming) regions, where unity derives from administrative
control or where the region is an ad hoc area over which specific programmes
apply.

This division, it should be emphasised, is more familiar among geographers than
economists. Richardson (1978: 20) points out that the concept of homogeneous
regions appeals to economists who use non-spatial techniques, for example regional
macroeconomic models that are ‘spaceless’. Such models require key economic
parameters (e.g. marginal propensity to consume and to import, and marginal capital-
output ratios) to have the same constant values over the region as a whole for them to
provide clear predictions. Similarly, the neoclassical approach (discussed in detail in
2.4) implies homogeneity and spacelessness:

"Perfect competition, a convenient if not inevitable assumption for the solution
of neoclassical models, implies perfect mobility of commodities and factors,
the free and instantaneous flow of information and zero transport costs. As a
result, to assume perfect competition we further need to assume the absence of
space and nodality characteristics.” Richardson (1978: 20-21)

Once space is admitted into economic analysis, a number of key assumptions in
orthodox economic theory break down. Imperfect competition becomes the norm,
increasing the tendency for phenomena such as clustering. This conflict between
simplified ‘spaceless’ economic models and geographic reality resurfaces throughout
the 20th century development of regional economics; and it constitutes a significant
element in debate about regional policy.

Spaceless economic models track the interactions of a small number of clearly
identifiable variables, and create a strong incentive to define regions homogeneously.
However, this approach does not satisfy those with an interest in how economic
activities are distributed geographically. The space economy is heterogeneous, with
population and industries concentrated as specific locations rather than being scattered evenly over space. Thus most regional and urban economists prefer the nodal region concept because it deals explicitly with what happens within regions and treats the spatial dimension as important (Richardson, 1978: 21).

2.3 Delimiting Regions

Despite the temptation to develop mechanistic criteria for determining regional boundaries, such an approach should be treated with great caution (Richardson, 1978: 17). Public opinion and politics are probably far more influential than the views of civil servants or economists, however elegant their models. For example, the creative division of Wales into two new NUTS 2 regions was driven by Welsh politicians and opinion formers concerned to secure Objective 1 status for the poorer parts of Wales during the 2000-2006 round of European Union funding (Morgan and Price, 1998).

These political realities have been widely recognised. Richardson (1978: 17) concludes:

"The delimitation of regions is not an easy task. Whether one uses economic, administrative, historical or other criteria there are no satisfactory methodologies. Compromise is inevitable, and disputes among economists and planners about whether particular boundaries are appropriate are very common. Indeed, defining regions precisely is such a nightmare that most regional economists prefer to shy away from the task, and are relieved when they are forced to work with administrative regions on the grounds that policy considerations require it or that data are not available for any other spatial units."

The inconclusive nature of the term is reflected in the four open-ended criteria suggested by Cooke et al. (1996) for defining a ‘region’: 1) no determinate size, 2) homogeneous in terms of specific criteria, 3) distinguished from bordering areas by a particular kind of association of related features, and 4) some kind of internal cohesion. Furthermore, the borders of regions are not fixed once and for all; they can change. In the same way, Armstrong and Taylor (2000: xi) adopt a pragmatic approach, noting that the term ‘regional economy’ is likely to vary according to the topic under discussion and the availability of data. ‘Regions’ may refer to areas as
large as states or provinces, or as small as local authority areas. A similar line is taken by McCann (2001: 2):

“A region is defined here as a spatial area which is larger than a single urban area, but which is different from the spatial definition of a single nation.”

This discussion suggests that institutional, historical and cultural factors make it unfruitful to search for a uniform or an optimal model of a region.

This wide definition means that some regions may be larger in terms of population or wealth than nation states. Indeed, the distinction between nations and regions is far from distinct. Some areas regarded as constituting single economic regions cut across national boundaries: for example, the economies of Detroit in the United States and Western Ontario in Canada are largely the same regional economy (McCann, 2001: 3). By contrast Catalunya, one of the most prosperous Spanish ‘regions’, is now formally recognised within Spain as a ‘nation’ following a controversial referendum in 2006, although this does not as yet affect international relations or defence. With seven million speakers, the Catalan language is more numerous than a number of languages that have full official status within the EU such as Estonian, Latvian, Lithuanian, Danish and Maltese (Pujol, 2005: 7).

While making the point that regional economies are far more open than the national economies within which they are located, Webber and White (2003: 773) maintain that following the liberalisation and harmonisation of trade barriers across EU member-states, the characteristics that used to make regions distinct from national economies have diminished. Vickerman (1992: 17) suggests that: “National states are becoming much less significant within a single market and we should look more fully at integration between regions”.

While offering no conclusion on this continuing and highly political ‘nation state vs region’ debate, this thesis follows the pragmatic consensus in using the term ‘region’ in an open-ended way. Trade and migration between regions follow a similar pattern to trade and migration between countries, although typically facing fewer obstacles. Globalisation, the influence of the World Trade Organisation, the rise of major trading
blocs such as the European Union and common currency areas make it more likely that the economic distinctions between many regions and nations will become still more blurred in future.

2.4 Regional Economic Theory and Policy

Having reviewed the nature of the region, this section seeks to answer two questions – what is regional economics, both theory and policy, and what is it for?

Regional economics analyses the spatial dispersion and coherence of economic activities (Nijkamp and Mills, 1987: 1) and centres on the question of “why economic agents (firms, workers and consumers) choose some regions over others as their preferred location for production work or consumption and why they often choose en masse to locate in some regions” (Brakman and Garretsen, 2003: 640). One important focus of theory is on what causes economic growth, and hence prosperity; and interest in its predictions is heightened by the opportunity of devising appropriate policies to assist this process.

However, regional economics has not evolved in a straight line. Although the roots of regional science can be traced as far back as nineteenth century economists and geographers, many of the concepts and models (such as ‘New Economic Geography’) are the work of recent decades, and reflect a revival of interest in a somewhat neglected field. It should be noted that regional science has been the focus of interest for both economics and geography, and that the neglect during much of the 20th century was on the part of economists. Brakman and Garretsen (2003: 637) maintain that “economists and geographers analyse similar issues, but do so in a very different framework and without taking too much notice of each others’ work”. Thus mainstream economists paid virtually no attention to the economics of location: “The analysis of regions was at best confined to the fringes of mainstream economics (e.g. regional science)” (Brakman and Garretsen, 2003: 639). Meanwhile attempts to develop location theory were taken forward by geographers.
On the other hand, the contributions of geographers have been criticised by economists. For example, Krugman (1993: 113) describes ‘what is called location theory’ as “a scatter of ideas, not really linked together – and in some cases location theorists seem to have failed to make linkages even where several different ideas are really different versions of the same insight.” However, some geographers respond that the New Economic Geography theory of economists (Krugman, 1991a, 1991b; Fujita at al., 1999) has done little more than rehash the work of classical geographers, and that in any event geography has moved on (Brakman and Garretsen, 2003: 639). Uncertainty and controversy continue to surround even fundamental questions of spatial behaviour and how it should be modelled: thus, for example, McCann and Sheppard (2003: 661) argue that modern clustering and agglomeration literature fails to address itself to basic problems posed by classical and neoclassical location models.

This thesis aims to summarise the contributions of both economics and geography to the regional science rather than attempt to resolve this issue; and the relevant insights of both sciences are reviewed in Part 2.

2.5 The Theoretical Base of Regional Science and Policy

Having described the concept of the region and the nature of regional economics, the thesis describes how theories relevant to regional growth and prosperity are developed. For the thesis, the main interest is how such theories may assist formulation of policy: hence, it is essential that theories are soundly based.

Economic theory is intended to assist the understanding of the way economies work, predict the outcomes of economic actions and so help guide policy. All economic development programmes are based on theory – whether or not that is explicitly acknowledged, or even realised. This underlying idea is often described as a programme theory, defined by Chen (1990: 43) as “a specification of what must be done to achieve the desired goals, what other important impacts may also be anticipated, and how these goals and impacts would be generated”.

20
The development and use of economic and other scientific theories has been strongly influenced by the work of the philosopher Karl Popper, whose approach displaced earlier views about the 'received view' of theories pronounced by Suppe (1974) and the 'hypothetico-deductive' model of scientific explanation (Hempel and Oppenheim, 1965).

Popper viewed any theory as consisting of a set of statements that satisfied a number of conditions, among them:

- Its set of 'axiom' statements must be independent and not contradict one another;
- No dependent statements should contradict any other dependent statements (Kenyon, 1984: 2).

Popper’s approach to the nature of scientific method is far from uncontested, as alternative ideas have been put forward in place of the strict use of falsification (Kuhn, 1970; Lakatos, 1978; Kenyon, 1984; Blaug, 1992). To summarise, the main point of contention is between the ‘logical positivism’ advanced by Popper – an ‘aggressive methodology’ in the sense that by its standards some of what is called ‘science’ can be dismissed as methodologically unsound – and the more defensive methodology of Kuhn, (1970) that tends to vindicate scientific practice (Blaug, 1992: 32). One difference in the approaches significant for regional economics is that logical positivism is inclined to strip out the influence of history from scientific methodology. By contrast, the Kuhn approach views science as advancing from one paradigm to another by means of occasional ‘scientific revolutions’ (Kuhn, 1971; Blaug, 1992: 36); or by new ‘scientific research programmes’ (families of theories) constantly superseding one another with theories of ever-increasing empirical content. This contrast in basic approach is paralleled by the alternative treatments offered by neoclassical theory and evolutionary economics, which also emphasises the role of history.

The philosophical discussion of proof and disproof may seem remote from the world of regional economic policy, and further discussion is beyond the ambit of this thesis: more detailed treatment of non-corroboration and falsification may be found in
Lakatos (1978) and Thornton (1997). However the nature of the scientific method is highly relevant to economic science. For example, doubtful theories may sometimes be defended because it is difficult to judge whether a hypothesis has been refuted, or rather that other circumstances may have interfered with observation. Scientific explanations can only be appraised in terms of the predictions they imply (Blaug, 1992: 23), but unlike traditional empiricists, Popper held that experience does not determine theory – theories cannot be argued or inferred from observation. Rather experience delimits theory: it shows which theories are false (Kenyon, 1984: 7) and also indicates how much predictive power they possess. Paradoxically, the more 'improbable' a theory, the better its scientific potential may prove (unless falsified), because probability and informative content vary inversely (Thornton, 1997: 7). This led Friedman (1953) to argue that the more unrealistic its assumptions, the more significant a theory would be, as it would explain much by little, an elaboration dubbed the 'F-twist' by Samuelson (Wong, 1973).

In the context of regional economics, this discussion is relevant to consideration of the various neoclassical and evolutionary explanations of regional convergence and divergence (3.3). This debate also resurfaces in the formulation of the model adopted by this thesis, which, as argued in Chapter 14, rests on theoretical ideas of how the economy works rather than empirical observations.

2.6 Methodology of the thesis

The methodology of this thesis follows a positivist paradigm, in that it is constructed on the basic premise that the world is objective and capable of measurement (Remenyi et al., 1998: 104). In pursuing a positivist line, however, alternative approaches also require consideration, in particular that of interpretivism as well as structuralist and postmodernist approaches (Audi, 1999). The rationale is that given limited resources for a large potential number of projects there is a need for an 'objective' allocation process in line with a clear programme theory.

It is also worth pointing up the important distinction that exists between the conceptual difficulties involved in constructing economic theories and the technical problems inherent in finding, measuring and analysing the data required to conduct
empirical tests. In the specific context of regional economics, this health warning has a special relevance. As shown below, especially in Chapter 4, theories applied to regional economics tended to overlook the influence of space and location. Yet all economic phenomena take place within geographical space, and many questions about the performance of particular local, urban and regional economies are in fact questions about the relationship between geography and the economy (McCann, 2001: 1). Spatial questions tended to be overlooked by economists and economic policymakers, a consequence partly of their education and partly of the post-War Bretton Woods world of relatively closed economies and currency convertibility restrictions. Globalisation makes these confining assumptions of spacelessness increasingly unacceptable in dealing with regional economic issues: As McCann (2001: 1) maintains:

"However, in the modern era of free trade areas, new information and communications technologies, currency convertibility and increased capital and labour mobility, many of these traditional assumptions can no longer be justified."

Much of the succeeding presentation in Part 2 describes the significant consequences of introducing the spatial dimension, which tends to restrict the influence of perfect competition (4.4).

2.7 Counterfactual Analysis

Empirical testing of theory or policy involves imagining what would be the outcome in its absence. While most mainstream economics texts are silent on the term, economic policy practitioners frequently talk of the ‘counterfactual’. Vanhove (1999: 7-8) defines the ‘counter-factual situation’ as what would have happened in the absence of a policy. In the context of regional economics, this is sometimes described as expected growth of income or employment based on ‘autonomous factors’ (Blake, 1973; Vanhove, 1999: 7-8). Various UK examinations of development programmes have concluded that the methodologies used were often weak because of the absence of any measure of the counterfactual in the form of appropriate control groups of firms outside the programmes (Pickernell, 1999: 18). In fact, it is a commonplace of regional policy evaluation, and of policy evaluation in general, that “the single most
significant obstacle to obtaining accurate results is the problem of the counterfactual” (Cheshire and Magrini, 2002: 209).

Counterfactual situations (or ‘counterfactuals’) are described as “alternative states of the world and/or alternative management decisions” (HM Treasury, 1997: 12). For example, the evaluation of the Regional Selective Assistance grant scheme in Assisted Areas of Great Britain between 1991 and 1995 made an estimate of an explicit ‘counterfactual scenario’ of what would have taken place without the grant (DTI et al., 2000: 32-34) involving eight counterfactual alternatives. The first explicit definition of causation in terms of counterfactuals may be traced back to the 18th century Scottish philosopher David Hume (Hume, 1748: VII; Menzies, 2001: 1). Full discussion of the term ‘counterfactual’ would be beyond the scope of this study: accounts are provided by Lyons (1967), Lewis (1973), Mackie (1974), Cowan and Foray (1999), Sledge (2000), and by Lewis (2000), who presents a revised counterfactual theory based on the central notion of influence (Menzies, 2001: 10).

However, one aspect of debate is worth noting for its possible explanation of unequal development between nations and regions. Cowan and Foray (1999: 3) maintain that counterfactual analysis takes on a different and more significant role for theories presented by evolutionary economics, in which the goal is to explain changes over time and where (unlike strict neoclassical theory) multiple equilibria may be possible (5.5). The presence of multiple equilibria raises the issue of selection, and for evolutionary economists selection mechanisms are inherently historical: “Historical processes can have the feature of path dependence, and this is commonly present in evolutionary models.” In other words, we are where we are because of our history, and while our history may not bind our future, it will certainly affect the way we get there. This theme of path dependence occupies a prominent role in the literature of regional economics, and is dealt with in greater detail in Chapter 5.

This clash between neoclassical and evolutionary economics, studied in greater detail below (5.4 and 5.5) poses particular problems for regional science and regional policy makers. It suggests empirical testing may not be able to identify what factors explain observed outcomes because (quite apart from any practical difficulties) the underlying programme theory is unclear. Yet governments demand answers, and economists are
obliged to proceed with models based on what seem to be the most likely theories. Part 2 therefore provides a survey of the main theories and empirical evidence.

2.8 Conclusions

This chapter briefly discussed the concept of the region, introduces the idea of space in economics and described how scientific enquiry seeks to find answers to the spatial distribution of economic activity. This treatment set the scene for considering the research question. It noted:

- Once space is admitted into economic analysis, imperfect competition becomes the norm, increasing the tendency for phenomena such as clustering;
- Institutional, historical and cultural factors make it unfruitful to search for a uniform or an optimal model of a region;
- Empirical observation shows which theories are false rather than which theories are true;
- Orthodox economic theory has tended to ignore the complications of space and therefore may overemphasise the influence of perfect competition and free markets;
- Counterfactual analysis has a particular significance for regional economics.
PART 2 THEORETICAL FRAMEWORK

Chapter 3 ‘Spaceless’ regional economic theories

3.1 Introduction

Part 2 reviews the theoretical basis for regional policy, and examines in particular what justification theory may provide for publicly funded and directed intervention. This is a necessary preparation for developing a sound methodology of appraisal and evaluation.

This requires an examination of regional growth theory, with special attention to explanations of persistent variations in growth and prosperity. This study of regional economies reflects understanding of the way the broader global economy works; borrowing from concepts evolved to explain economic growth, variations in output and employment, migration of factors and international trade. The focus is firmly on solutions – what insight theories can bring to understanding the causes of disparity between regions and whether they can be tackled.

Chapter 3 deals with the principal theories developed by economics to account for growth, trade and prosperity. Such models focus on economic relationships rather than presenting formulae for determining location decisions. Richardson (1978: 132) criticises previous formulations of both the supply- and demand-based models, maintaining that “Neither treats space seriously”. Space creates friction – an obstacle to interregional capital flows (and, he might have added, a more serious obstacle to migration of labour). As argued below (4.4), introduction of the spatial dimension implies a greater degree of imperfection of competition, removing key assumptions...
which neoclassical economics requires for equilibrium solutions. While it may be legitimate to set aside spatial issues in order to isolate economic factors for the purposes of analysis, it is not then possible to draw conclusions about geographic distribution of economic activity without their readmission.

It should be emphasised that this chapter considers only 'spaceless' models devised by economists - but these comprise a wide variety of different theories. As discussed below (3.3), there is no generally accepted consensus of why regional economies may grow at different rates. Explanations reflect the two main cultures in economics - formal neoclassical (orthodox) economics and heterodox political economy (Tabb, 1999). This thesis follows Martin and Sunley (2001: 150) in reserving the term 'mainstream' economics for neoclassical economics and its immediate offspring variants (such as endogenous growth theory).

3.2 Taxonomy of Spaceless Regional Economic Models

Orthodox neoclassical economics is defined as “that part of economics that subscribes to methodological individualism, mathematical models, rational maximising behaviour and the notion of general equilibrium” (Martin and Sunley, 2001: 150). Theoretical economic models developed to explain the causes of economic growth of regions and countries (and hence disparities between them) may be grouped for convenience within one of four broad categories (the first of which may be regarded as falling within the orthodox neoclassical approach):

- Supply-side models that explain growth as a consequence of factor endowment, including the neoclassical growth model, in revised forms of which technical progress is generally treated as a factor of production;
- Models that focus on demand for regional exports, and so put emphasis on the demand side;
- ‘Cumulative causation’ models incorporating both demand- and supply-side mechanisms, which describe the way in which growth, once started, tends to become self-propagating;
Evolutionary approaches, which adapt supply-side models and combine an explanation of the trial and error nature of innovation with the inherent differences in the economic environments of different regions (Chapter 5).

None of these approaches are mutually exclusive. The operation of all markets depends upon the interaction of both supply and demand; and the difference between them depends on the respective emphasis they place on each. Armstrong and Taylor (2000: 65) suggest it is tempting to interpret the demand-based export model as being the best explanation of growth in the medium term, while leaving explanations of long-term growth disparities to the neoclassical model: however, this is a temptation they resist, instead treating the supply- and demand-side explanations as two separate approaches. In the various developments of the cumulative causation model, however, both these approaches are combined; just as microeconomic price theory requires the interaction of both supply and demand schedules. Similarly, the evolutionary approach extends supply-side analysis over time; and this thesis suggests that cumulative causation can be interpreted to include the dynamic insight of evolutionary models (7.2). A number of other taxonomies have also been proposed, with Vanhove (1999) and Pop (2002) both classifying the cumulative causation model as a variant of export-based models.

3.3 Supply-based Models

3.3.1 Traditional Neoclassical Models

The first broad explanation, supply-based models, stress the role of factor endowment, and can be traced back to the English economist and Member of Parliament David Ricardo (1772-1823), who explained international trade by pointing to the advantages that nations possess in producing different commodities from varying levels of labour productivity (Backhouse, 2002). More recently, explanations have been based on the Heckscher-Ohlin theorem, the principle that a country will export those commodities that are intensive in the factor of production in which it is relatively best endowed (Bannock et al., 1992: 193). This ‘factors proportion theory’ maintains that regions specialise in production and export of goods in which they have a comparative
advantage, conferred on them by relatively abundant factors of production (Suranovic, 2006).

In the most traditional neoclassical models, the key explanatory variable is taken to be the distribution of capital and variations in the capital-labour ratio, while technology was treated as an unexplained residual (Borts and Stein, 1964). In an economy with no technical progress, output will be determined entirely by inputs of capital and labour, expressed in the general form of:

\[ Y = F(K, L) \]

where \( Y \) is real output, \( K \) is the stock of capital, and \( L \) is the labour force. One specific formulation of this relationship is the Cobb-Douglas production function (Armstrong and Taylor, 2000: 66). Assuming constant returns to scale, an unlikely scenario, the Cobb-Douglas production function may be expressed as:

\[ Y = AK^\alpha L^{1-\alpha} \]

where \( A \) and \( \alpha \) are parameters to be estimated (usually by regression analysis). The implication of \( \alpha \) is that the higher the contribution of capital to output, the lower is that of labour, and vice versa.

This production function can also be expressed in per capita terms (by dividing both sides of the equation by \( L \), so producing:

\[ y = AK^\alpha \]

where \( y = Y/L \) and \( k = K/L \).

Then output per worker or productivity (\( y \)) can only increase if capital per worker increases. With no technical progress, the model predicts that in long-run equilibrium, output, capital and labour will all grow at the same rate (Armstrong and Taylor, 2000: 68).
The neoclassical framework in its basic form therefore emphasises the importance of capital accumulation, seen as the reward of saving. This can be viewed as a constant average propensity to save (Solow, 1956) or, in more sophisticated rational expectations models, the result of the combination of a time preference rate and the intertemporal elasticity of substitution with respect to consumption (Lucas, 1988).

For regional economics, the importance of the neoclassical model is its prediction that factors of production will migrate to those regions where their relative price is highest and where they are therefore able to gain the greatest returns (Armstrong and Taylor, 2000: 72). This leads to the important consequent prediction that capital and labour should migrate in opposite directions, leading eventually to equal returns to both factors in all regions. In other words, different regions should experience convergence of output per head, a proposition that can be tested by empirical observation and made subject to the alternative trials of falsification and confirmation.

Yet in the real world, evidence of persistent regional disparities abound (Armstrong, 1995; Sala-i-Martin, 1996a; 1996b). Neoclassical models can explain different levels of development only as the result of (i) differences in saving behaviour, or (ii) as pure delays that will disappear over time (Castro and Jensen-Butler, 1999: 3). The first of these explanations implies immobility of factors of production, while the second cannot account for the persistence of development gaps. Explanations of such persistent gaps should involve a capital accumulation process where diminishing returns are absent (Castro and Jensen-Butler, 1999: 3), otherwise regional economies would eventually converge.

3.3.2 Modified Neoclassical Model

One significant development of this supply-side approach – which could provide a more adequate explanation of development gaps – was the theory of economic growth proposed by Solow (1956; 1957). Previously neoclassical theory tended to neglect technological change of the endogenous variety, claiming it was best treated as an exogenous input acquired as needed by purchase in the market; and only with Solow’s work was technology acknowledged as an important factor on productivity growth in
its own right (Cooke and Morgan, 1998: 195). Solow’s neoclassical growth model identified three sources of output growth – the capital stock, the labour force and technology. It marked a movement away from the idea that technological progress appears as ‘manna from heaven’ (Castro and Jensen-Butler, 1999: 3). Conceivably, therefore, different endowments of technology, if immobile, could account for spatial variations in growth and prosperity.

The modified neoclassical model assumes there is no impediment to the movement of factors of production and that there is perfect knowledge about factor prices in all regions, so that capital and labour are free to move to those regions that offer them the highest rates of return. Its emphasis on the dominance of technology may be significant for understanding the different pathways followed by regions and countries that otherwise may appear to have a similar potential for economic growth.

The dominant role in this model is played by ‘advances in knowledge’ and technology: thus Solow (1957) estimated that technology accounted for 87 per cent of per capita productivity growth (the remainder attributed to increases in capital inputs). In other words, the model suggests that technology is the ‘engine of growth’. Since then, further methodological and statistical refinements have reduced the unexplained “residual” of productivity growth that is equated to advances in knowledge and technology, but it remains the largest single source of long-run productivity and economic growth. It is estimated to account for more than one-third of total GDP growth in the USA since 1929 (Denison, 1985), and for between 34% and 63% of GDP growth in the OECD countries over the period 1947 to 1973 (Barro and Sala-i-Martin, 1992).

As noted, however, the alleged dominance of technology is derived from its definition as the residual factor in the model – accounting for whatever variation cannot be attributed to the inputs of labour and capital. This almost accidental dominance was felt to be unsatisfactory: Parkin et al. (2000: 815) observe that in the formulation presented by Solow (1956; 1957), economic growth does not influence the rate of technological change: “Rather economic growth is determined by chance.”
3.3.3 Endogenous growth model

While the Solow model treated technological progress as a separate variable, the cause of which is not identified, Romer (1986; 1990) suggests an endogenous growth model, in which the long-run growth rate of output per worker is determined by variables within the model, not an exogenous rate of technological progress as in the neoclassical growth model (Meyer, 2001). Endogenous growth analysis diverges from conventional neoclassical analysis by making technological change a function of economic incentives and behaviour (Freeman, 2001). It predicts that growth will be faster or slower depending on the costs and benefits of knowledge-creating and innovative activity, which policy can influence in various ways. What distinguishes endogenous growth theory is that “it attempts to provide an explanation of the causes of technological progress rather than treating it as a given” (ODPM, 2006: 10).

One important consequence involves returns to scale. According to Hulten (2000), “What is new in endogenous growth theory is the assumption that the marginal product of (generalized) capital is constant, rather than diminishing as in classical theories.” Generalized capital includes the result of investments in research and development (R&D). This conclusion is central for much of the controversy about the nature of regional economic development, because diminishing returns are a prerequisite for neoclassical convergence between regions (5.5). In the Romer model, technological progress is attached to the workforce. This produces a variant of the Cobb-Douglas production function, with the model predicting that long-run equilibrium for the economy requires output per worker, output per capital and technological progress all to grow at the same rate as each other. Technological progress is assumed to grow ‘endogenously’ due to the natural desire of entrepreneurs in ‘knowledge producing’ industries to make a profit (Armstrong and Taylor, 2000: 77). It predicts that in the long run global technological progress will therefore grow at the same rate as the population as a whole.

Romer argued that new ideas are ‘non-rival and non-excludable goods’, which have a cost of generation but which can then be spread and used with almost no additional costs. Such new ideas are public goods, and associated with increasing returns to scale because the long-run average cost of embodying new ideas into technical
progress is continually declining. Similarly, Karlsson and Stough (2002: 6) maintain that more knowledge-dependent manufacturing and service industries increase the prevalence of internal scale economies. The mechanism for their creation is by monopolies and oligopolies rather than by firms in perfect competition, who would have no incentive to create them, which, according to Castro and Jensen-Butler (1999: 4), is one of a number of weaknesses in Romer's model. Romer (1990) shows that increasing returns and imperfect competition, both previously somewhat heretical notions to neoclassical economists, play a central in technological change and growth of productivity (Cooke and Morgan, 1998: 196).

The neoclassical model implies that technological progress should take place instantaneously (Armstrong and Taylor, 2000: 80), and it could be argued that continued globalised competition and advances in communication makes rapid diffusion of technology more likely than in the past. Empirical studies suggest however that this is far from being the case (Guerrero and Sero, 1997; McCombie, 1982), and clearly none of the formulations described so far provide an adequate explanation of the persistence of major development gaps between regions and nations. The gaps in prosperity should be bridged over time by trade and migration, yet in the real world big differences remain evident. As Armstrong and Taylor (2000: 79) point out, the endogenous growth model is designed to explain the growth of the world economy as a whole, not that of individual countries and regions, for whom ability to keep up with latest technology may be more important than overall global progress.

3.3.4 Catch-up model and Convergence

A similar mismatch between neoclassical theory and evidence can be found in the model's prediction that regional per capita incomes converge rather than diverge. For example, one extension of neoclassical theory, the 'catch-up model', maintains that the further the technology of a region is behind the technology of the most advanced region, the faster is its rate of technological progress.
This can be portrayed algebraically as in equation 3.3.4:

\[ \Delta A_r = \lambda(A^* - A_r) \quad \lambda > 0 \]  

where \( A_r \) is the technology of the region in question, \( A^* \) is the technology of the most advanced region. The argument lying behind this assertion is straightforward. A region lagging technologically should be able to catch up rapidly by transfer of technology – by copying the processes used in more advanced regions, and by purchasing the latest capital equipment that embodies modern technology.

Even at a theoretical level, there are good reasons for doubting such straightforward convergence. For example, Groot and Schaik (1995) develop a two-country, two-sector endogenous growth model with a dual labour market based on efficiency wages. Growth is driven by research carried out in a high-tech tradable sector. The rate of growth of the follower country should be positively correlated with the size of the productivity gap between it and the leader country, in line with the catch-up model. However, differences in unemployment benefit systems may lead only to ‘relative convergence’, i.e. a steady state with the backward country lagging behind the leader country. The reason for this is that high social welfare benefits generate high unemployment and reduce the amount of labour employed for R&D purposes. Furthermore, it is shown that a shift in preferences towards non-tradeables can explain a global slowdown in economic growth.

Nor do empirical studies provide a convincing basis for the catch-up model. In a study of cross-sectional data for US states, McCombie (1982) found no evidence that each state’s technology gap had any effect on growth of labour productivity. Other studies have focused on spatial disparities in per capita income, which according to neoclassical theory should decline as capital flows to low wage regions and labour migrates to high wage regions (Sala-i-Martin, 1996a; Armstrong and Taylor, 2000: 81).
Two different measures of economic convergence have been distinguished:

- **beta convergence** (β-convergence), where poor regions grow faster than rich regions; and
- **sigma convergence** (σ-convergence), a more conventional measure of the fall in the dispersion of per capita income between regions over time.

Barro and Sala-i-Martin (1992), for example, obtained positive results for the neoclassical catch-up model predictions, with strong negative correlations for the growth of per capita income and the initial level of per capita income over the very long run for US states. However subsequent studies by Sala-i-Martin (1996a; 1996b), studying both β-convergence and σ-convergence, found that regional convergence is generally very slow in industrialised countries, while Armstrong (1995) found a similarly slow rate of convergence in the European Union. A cliometric study of β-convergence among developing countries in Africa finds that beta-convergence for growth is significant only for two subperiods, 1960-1970 and 1994-2001 (Charles and Ripoll-Bresson, 2002: 281).

Martin (1998) points out that there are different viewpoints as to what happens once a spatial distribution of economic cores and peripheries develops. The convergence school argues that spatial imbalance reduces naturally over time, while the contrary argument made by the divergence school suggests that such imbalances tend to reinforce themselves (Pickernell, 1999: 8). Persistent disparity in status and prosperity between core and peripheral regions is suggested by the theory of ‘internal colonialism’ (Hechter, 1969).

To demonstrate the complexity of the convergence process, Thirwall (2003: 270-272) reports that cross-section evidence on regional inequalities, and time-series evidence for different countries, suggests that the degree of inequality follows an inverted U-shape; that is, regional inequalities at first rise with the level of development and then decrease. Regional differences in a poor country tend to originate as a result of a favourable economic shock to one region or a set of regions – for example, growth of an export enclave or the development of industrial activities. Once established, the
differences will be accentuated by selective migration of labour from poor to richer regions and by the inflow of capital. ‘Spread effects’ from prosperous regions will be weak, owing to a general lack of political and economic integration. These factors will however become weaker over time, while external diseconomies of expansion in richer regions will appear.

Furthermore, the spread of technical knowledge between countries and regions appears far from being instantaneous (as suggested by the neoclassical model) and the pace of diffusion can change considerably over time (Suarez-Villa, 1999: 80-81; Armstrong and Taylor, 2000: 81). Economists have found evidence of persistence of disparities in wealth, stratification, and the formation of ‘convergence clubs’, with countries concentrating into groups of rich and poor (Galor, 1996). Veiga (1998:1) finds that there is empirical evidence of some real convergence among EU countries, mainly before the mid-1970s, but little evidence of real convergence at the regional level.

Most empirical studies have focused on convergence of income per capita, but the existence or otherwise of economic convergence between regions can also be tested by examining what happens to the prices of factors of production (Webber and White (2003: 773). The results are similarly inconclusive. Thus Tovias (1982) used data from the original six EEC member-states to establish whether factor-price convergence for labour took place: in fact, rapid divergence took place between 1968 and 1977, but over the whole period studied, 1957-1977, there was a long-run trend towards factor-price convergence for labour. Webber and White (2003: 773), in a study of regions in four EU states – Germany, Italy, Spain and the Netherlands – suggested that average regional factor prices for labour converged throughout the 1980s and into the 1990s, a time of economic integration, harmonisation and liberalisation within the European Union.

To sum up, empirical evidence on regional convergence is mixed. It fails to provide unequivocal support for the idea that left to themselves countries and regions will converge on similar levels of growth and prosperity. Armstrong and Taylor (2000: 85) conclude:
"The neoclassical model of growth therefore appears to predict correctly that there will be convergence of per capita incomes over the long run, but empirical studies suggest that the convergence process is likely to be painfully slow."

This conclusion is a major disappointment for those hoping for confirmation of the laissez-faire policy implications of orthodox economics and has led to a search for possible explanations.

3.3.5 Human Capital and Embodied Knowledge

One attempt to offer a credible rationale has involved the introduction of the concept of human capital (Lucas, 1988; Mankiw et al., 1990). Endowment of human capital (as distinct from labour) was seen as determining the ability to use technology available everywhere (Castro and Jensen-Butler, 1999: 4). As argued in this section, technology and human capital are crucial supply-side elements in determining regional prosperity. This significance is reflected in theoretical treatment of two ways in which regions acquire knowledge and technology. One way is through acquiring the latest capital investment, which embodies new production techniques. But more recently, attention has focused on regions’ supply of ‘human capital’, which is defined as the qualifications, skills and expertise that contribute to a worker’s productivity (Sloman, 1997: 263). A higher level of knowledge and skill enables better use of technology, so promoting competitiveness, and also helps a region to generate its own technological progress (Armstrong and Taylor, 2000: 86).

Castro and Jensen-Butler (1999), building on the model developed by Solow (1957), emphasise the importance of knowledge; both knowledge embodied in capital equipment or intellectual property and, more significantly, ‘disembodied knowledge’ embedded in the human capital of the workforce, and its capacity to transfer knowledge through effective networking. This modern variant of the neoclassical growth model seeks to account for the knowledge-based economy and may help explain the persistence of regional economic disparities – providing such knowledge is in some way immobile.
Such immobility could come about if a region’s capacity to absorb and create technological progress were determined by its institutional environment. Thus Rauch (1993) in a study of 237 US cities concluded that substantial external economies are gained from the geographical concentration of highly educated people. Knowledge-rich regions with effective institutional networks for generating and exchanging ideas and applications can maintain a continued lead over regions that are obliged to buy in exogenous technology embodied in capital goods. This advantage would in turn help explain evidence of increasing returns.

By contrast, regions poorly endowed with human capital are likely to depend on more routine activities that rely on technology embodied in capital equipment available world-wide. Such deprived regions will tend to rely for their economic growth on the spillover process described above (3.3.4), and on being cost-competitive in world markets (Armstrong and Taylor, 2000: 87).

Inward investment may in some instances provide a short cut to acquiring embodied knowledge as well as capital; but it is no substitute for the innovative networks of the ‘intelligent region’. Even though it can provide tutoring for local supplies and local training institutions, there are limits to the capacity of branch plants to provide a platform on which to build a more innovative, R&D-based economy (Cooke and Morgan, 2000: 149).

Armstrong and Taylor (2000: 87) present this modern variant of the neoclassical model algebraically as:

\[ Y/L = f(K/L, EXOG, ENDOG, HUMCAP) \]

where \( Y/L \) is the ratio of output to labour, \( K/L \) is the capital/labour ratio, \( EXOG \) is the technology embodied in the capital stock, and available to all regions, \( ENDOG \) is the technology generated endogenously within the region by its own skilled and innovative workforce, and \( HUMCAP \) is the ability of the region to absorb and use new technology, determined by its stock of human capital.
3.3.6 Supply-based Models: Summary

A number of significant points arise from this reworking of the ideas of Solow, Romer and earlier economists.

First, the model developed by Castro and Jensen-Butler (1999) provides a persuasive picture of the 21st century global economy. Regions rich in human capital and knowledge specialise in creative activities such as research and development, scientific research and high-level service activities. The key policy question, however, is to what degree can other regions replicate their success or catch up in other ways.

Secondly, the model, if soundly based, provides a possible explanation of the stubborn refusal of regions to converge economically in line with the neoclassical model’s prediction. Knowledge adjusted human capital enables regions to generate their own technical change: their favourable institutional environment equips them with a continuing advantage over regions dependent on ‘catch-up’ by buying in knowledge embodied in capital equipment from other areas.

Thirdly, the Castro and Jensen-Butler (1999) model provides a more realistic picture of the real world essentially by adding to the number of factors of production. In addition to capital and labour, three other separate variables show up on the right-hand side of the equation \((EXOG, ENDOG, HUMCAP)\). However a good portrayal is not necessarily the same as a good explanation. While capital and labour may be readily measured, concepts such as human capital may prove more elusive, although economists have sought to find good proxies e.g. the proportion of private sector investment devoted to research and development, or the number of patents (Dessyllas, 2005; Ejermo and Kander, 2006). Jones-Evans (2006: 8) describes the number of patents as arguably the nearest proxy to direct indicators of knowledge formation and knowledge creation, concluding: “Therefore the number of patents can be used to indicate how successful a region is in converting knowledge into potentially commercially valuable products and processes.”
The conclusion of this survey emphasises the difficulty encountered by neoclassical economists in adapting their equilibrium framework to explain persistent economic disparities between regions and countries. It is significant that in elaborating their models economists have been forced to look for factors outside economics (e.g. institutional environments) to explain this inconvenient disconnection between neoclassical theory and reality. The thesis now examines a number of other approaches to see if their adherents fare any better.

3.4 Demand-based Models

The second explanation of regional growth, the Export-Base Model, is a development of the Heckscher-Ohlin theorem (3.3.1). The theorem describes the way in which regions specialise in production and export of goods in which they have a comparative advantage, conferred on them by relatively abundant factors. There is therefore no fundamental contradiction to the supply-side explanation provided by the neoclassical model outlined above.

One reason for supposing the significance of this explanation is that export-based industries posses a potency that exceeds their scale. Porter (2003: 549) in a study of US clusters between 1990 and 2000 classifies industries into three groups – traded, local or resource-dependent. While traded industries account for about one-third of employment, they register much higher wages, far greater rates of innovation and exert a strong influence on local wages.

Export-oriented firms could be encouraged by such measures as tax relief, improved transport facilities, and high quality telecommunications. The danger is that such a strategy relies on external demand, and may lead to a skewed economy, vulnerable to vagaries in the global market. There is indeed considerable international evidence of countries with a sharp divide between the export sector and the ‘base sector’, which depends exclusively on the home economy (Encarta, 2007).

The key proposition is that initial development is linked to exploitation and then export of natural resources (Dixon and Thirwall, 1975). Kaldor (1970) argued that growth depended on scale economies and specialisation – and so applied to
manufacturing (rather than ‘land-based’ activities, such as mining of coal in 19th century South Wales). Regions specialising in export of manufactures will tend to import ‘land-based’ commodities.

Once specialisation is established, regions can benefit from outside demand for their goods and services, and the external demand for the region’s output will have a dominating effect on its growth (Armstrong and Taylor, 2000: 92). This provides a policy argument for giving priority to businesses that have an export market. This approach stresses the high multiplier effect of export-oriented firms; often with the underlying assumption that local non-export firms will then develop as a result of this increased activity (Blakely, 1994).

The model is typically expressed in the form of two equations:

- An export demand function; and
- An export supply function.

In the export demand function, the region’s output \( X^d \) will depend on the price of its exports, the income level of other regions \( Z \), and the price of substitute goods in external markets \( P_s \):

\[
X^d = f(P_x, Z, P_s)
\]

The export supply function includes all factors having a significant effect on production costs:

\[
X^e = f(P_x, W, P_k, R, C, T)
\]

where \( W \) is wage costs, \( P_k \) represent capital costs, \( R \) raw material costs, \( C \) intermediary input costs, and \( T \) the state of technology (Armstrong and Taylor, 2000: 94).

Growth of the export sector can then impact on the economy as a whole through:
1. A multiplier effect on regional income; and perhaps
2. An induced (accelerator) effect on investment; as well as

The scale of growth in the export industries will depend on the export multiplier, defined as the ratio of the total increase in a country’s national income to the increment in export revenue generating the increase (Bannock et al., 1992: 155).

This export-base model has been further developed to explain economic growth, through such models as the accelerator-multiplier. A positive shock, such as an increase in government expenditure, raises consumers’ incomes, which (through the multiplier) leads to an increase in output. This in turn (through the accelerator) raises investment, and this increased investment raises incomes, so repeating the process (Bannock, 1992: 9-10). Under some conditions, this may lead to unstable cyclical patterns of growth, as depicted by the cobweb theorem (Nerlove, 1958; Bannock, 1992: 72-73).

While discussion of the validity of such growth theories is beyond the scope of this thesis, it is worth noting the relevance of the acceleration principle to the subsequent development of cumulative causation models of regional economic growth. One particular feature of interest for regional economic science is under what circumstances growth can become self-sustaining.

### 3.5 Cumulative Causation models

A more general picture of regional growth is provided by models that (like variants of the neoclassical model) incorporate the idea of endogenous growth and develop an integrated process of cumulative causation. The concept of cumulative causation was familiar to economists such as Veblen (1898: 40).

In effect, such models simply draw together the supply- and demand-side factors facing regional economies. An explicit cumulative causation approach was first proposed by Kaldor (1970) and further developed by Dixon and Thirwall (1975) in a model that integrates neoclassical and post-Keynesian approaches to regional
economic growth. The Kaldor and Dixon-Thirwall models represent just one strand among a variety of approaches developed by regional economists to account for persistent gaps in regional prosperity: This formulation is considered in some detail because of the insight it gives into more complex models of regional economies.

Other explanations within the 'spaceless' family of theories have been offered by Myrdal (1957), McCrone (1969), Markusen (1983), Abraham and Van Rompuy (1992) and Rice and Venables (2003).

The Dixon-Thirwall model is set out in equations 3.1-3.4, with all the variables expressed as rates, not quantities). The four simultaneous equations propose the following relationships:

1. between output growth and productivity growth;
2. between prices and costs / production growth;
3. between exports and prices (of exports), prices (of competitor products), and growth of world income; and
4. between exports and output growth.

\[
q = \alpha + \lambda y_{-1} \quad (3.1)
\]
\[
p = w - q \quad (3.2)
\]
\[
x = - b_0 p + b_1 p f + b_2 z \quad (3.3)
\]
\[
y = \gamma x \quad (3.4)
\]

In the first equation, 3.1:
- \( q \) represents growth of productivity
- \( \alpha \) is autonomous growth of productivity
- \( y \) is growth of output
- \( \lambda \) is a constant (known as the Verdoorn coefficient, and discussed below)
- \( y_{-1} \) is growth of output lagged one period

This relationship maintains that productivity is partly determined by an autonomous factor (\( \alpha \)), and partly by growth of output in the previous period.
In equation 3.2:
\( p \) is price inflation in the region
\( w \) is cost inflation in the region
while \( q \), as previously, is growth of productivity.

Equation 3.2 states that the movement of prices in the region are determined by factor cost inflation and the growth of productivity: so if productivity increases less than factor cost inflation, regional prices will increase.

In equation 3.3,
\( x \) represents the growth of regional exports;
\( p \) is, as previously, price inflation in the region, while \( p_f \) is price inflation in the region’s main competitors;
both \( -b_0 \) and \( b_f \) are price elasticities of demand (of the three price elasticity coefficients, \( b_0 \) is negative, as the higher the price of the region’s exports, the lower will be the amount demanded);
\( b_f \) is the price elasticity of demand for the region’s main competitors; while \( z \) is the growth of world income, where the ‘world’ is understood as the region’s main export markets.

So the faster world income grows, the faster the region’s exports grow.

Finally, in equation 3.4, an export-base relationship (as discussed in 3.4 above),
\( y \) once more is output growth;
\( y \) is a coefficient measuring the responsiveness of the region’s output growth to its growth of exports; and
\( x \) is the region’s growth of exports.

The significance of the Dixon-Thirwall model is that it contains a feedback mechanism which makes the system self-perpetuating and which depends on an important connection between output growth and productivity growth, a relationship expressed by the so-called ‘Verdoorn coefficient’ (expressed as \( \lambda \) in the first equation,
3.1). This relationship takes account of the accelerator effect that output growth may have on investment. Providing the coefficient is positive, the relationship argues that greater output will promote greater productivity, and so lead on to greater growth, a process that is cumulative and self-perpetuating.

The model shows how an increase in world income \((z>0)\) leads to increased regional exports (equation 3.3) and therefore increased output growth (equation 3.4) (and immigration). This raises productivity (equation 3.1) which in turn reduces the region’s (relative) price inflation (equation 3.2): cost inflation is assumed to be constant. At this point second round effects begin, with improved competitiveness (3.2) creating a further increase in regional exports (via \(-b_o\), the price elasticity of demand for the region’s exports). In other words, it demonstrates a process of cumulative causation, in which second round effects are smaller than those in the first round, in the same way as the multiplier process (Dixon and Thirwall, 1975). Richardson (1976: 147n) notes that the Dixon-Thirwall model rivets attention on the export demand function, with regional growth treated as a function of demand for its exports. Myrdal (1957) allows for alternative growth triggers, such as a favourable endowment of resources. However, in a newly developed region this may be quickly translated into increased exports, so the difference may be unimportant.

The key feature of Kaldorian and Dixon-Thirwall models is that they incorporate a cumulative growth process (embodied in the Verdoorn coefficient) into what is essentially an export-led growth model. However the Verdoorn relationship is based on two different (and broad) concepts:

- Internal economies of scale, following the ideas of growth-pole theorists such as Perroux (1950), Myrdal (1957) and Hirschman (1958): here firms in an industry with scale economies gain a competitive advantage over their rivals;

- External economies of scale, in which the benefits are enjoyed by people other than the producer (Sloman, 1997:325).
The Dixon-Thirwall model has been criticised, for example Armstrong and Taylor (2000: 100-101), for a number of shortcomings:

- It does not identify the cause of specialisation in exports (e.g. factor endowment);
- It assumes export growth is the only source of regional output growth (Armstrong and Taylor, 2000: 100);
- The so-called ‘Verdoorn Law’ (that expansion of output improves productivity growth, or \( \lambda > 0 \), is a highly simplified relationship covering a number of variables (e.g. greater specialisation, or greater technical progress);
- It does not deal directly with the effect of growth on a region’s balance of payments.

It may also be argued that there is a danger of measuring output growth and productivity growth in a way that picks up the same quantities, and that the periods in question need therefore to be sufficiently distant from one another for the relationship reflected by the Verdoorn coefficient to be meaningful. Nevertheless the model has a strong intuitive appeal in accounting for disparities in regional growth.

The Dixon-Thirwall model also points clearly to an important implication inherent in the idea of cumulative causation: that economic growth, however generated in the first place, can become endogenous and self-sustaining. Under favourable circumstances – for example, a strong positive relationship between growth of output and growth of productivity – the various theories of ‘cumulative causation’ provide a plausible case for public intervention; and this argument is developed further below (7.2). In turn, productivity has been commonly attributed to ‘five key factors’ (skills, investment, innovation, enterprise and competition) as well as ‘place-based’ factors such as a quality environment (ODPM, 2006: 7-8), which may be susceptible to intervention.

3.6 Conclusions

This chapter examined the most familiar explanations of regional economic development, with the aim of providing a firm base for answering the research question (1.2). Although some scholars such as Blakely (1994) maintain that there is
no adequate set of theories that successfully explain variations in local and regional prosperity, even within the confines within of 'spaceless' theory this brief survey of theory and empirical investigation points to a number of possible conclusions for those involved in framing regional economic strategies.

First, although factor endowment in the way proposed by Ricardo may not explain 21st century economic growth, a general consensus has developed regarding the importance for economic growth of human capital in general and knowledge-based activity in particular, with obvious policy implications.

Secondly, models of cumulative causation stress the importance of getting ahead – however initial growth is generated, it is likely to accelerate. This in itself is a powerful argument for public intervention.

Thirdly, cumulative causation can, and does, have negative as well as positive effects. Taylor and Wren (1997: 837) maintain that unbalanced regional growth leads to the persistence and intensification of regional disparities. For example, selective migration of highly-skilled and better educated workers may damage the regions of origin, while destination regions may suffer because of increased pressure on available resources.

Finally, it is far from certain that free markets will even out geographic differences in prosperity. In fact Webber and White (2003: 773) acknowledge that few issues are more controversial than the 'so-called convergence hypothesis'; and the evidence for such convergence, crucial for arguments against government intervention, remains uncertain.

Thus even orthodox economics admits a point of entry for public policy. As argued in the following two chapters, consideration of the dimensions of space and time widen this chink in the armour of orthodoxy.
Chapter 4 Spatial Theories

4.1 Introduction

This chapter further develops the theoretical base required to attain the thesis objectives by introducing the role of space in regional economics. The models reviewed in the previous chapter operate by assuming a spaceless economy, with the emphasis placed on explaining how firms ‘solve’ their production problems – deciding on the output they produce and the factors inputs they use in line with their production function (3.1). Their goal is usually assumed to be profit maximization, whether under conditions of perfect or imperfect competition (Samuelson and Nordhaus, 2001).

This elimination of space may be justified by the need to concentrate on the interaction of key economic variables; but in the real world location counts. Isard (1956: 83-84) argues that if the concept of ‘time preference’ is allowed, so too should the concept of ‘space preference’. Weber (1911: 667-688) pointed out that classical theory ignored the transport costs involved in traversing space. At some point, therefore, the analysis must be extended to include the dimension of space and distance:

“To understand spatial patterns of economic activity it is necessary to understand how economic processes (and indeed social and political
processes) operate but equally, in order to understand how economies operate, an appreciation of the effects of space is required" (Healey and Ilbery, 1990: 4).

The models summarised in this section represent attempts made over the course of nearly two centuries by geographers and economists to explain the forces that influence location. Locational decisions by firms in turn effect the distribution of economic activities between and within regions and countries, and therefore play a major role in determining regional prosperity. This means they are important for policy makers, who wish to know what impact their actions are likely to have. For example, an evaluation of the Regional Selective Assistance grant system in Great Britain between 1991 and 1995 found that large firms are more likely to change their location in response to grant offers than small firms (DTI et al., 2000: 6).

4.2 Development of spatial theory

Yet interest in spatial issues, among economists at least, has fluctuated markedly. Not only have distinct traditions emerged within the disciplines of economics and geography, but several sharply conflicting explanations have been offered by economists (McCann and Sheppard, 2003; Ottaviano, 2003). In most areas of economics prior to the 1990s, spatial issues as a whole were regarded, at best, as minor issues, a neglect perhaps explained by the era of limited factor mobility, high tariff barriers and fixed exchange rates that followed the Second World War. Renewed interest in industrial location in recent years may be accounted for by increased economic integration, free trade, new information and communication techniques and greater availability of data as well as simple changes in academic fashion (McCann and Sheppard, 2003: 650).

Mills and Nijkamp (1987: 704) also point to more subtle linguistic barriers:

“Although most microeconomic analysis was written in English or quickly became available in English, virtually all early work on spatial analysis was in German. Therefore the use of spatial models to analyse urban economics had to await the availability of German literature to economists who were well versed in modern micro theory, but did not read German. The 1954 translation of Lösch’s classic was the key event in the education of English language scholars on spatial analysis.”
Similarly Arthur (1994: 3) relates how German-language treatment of agglomeration economies that emphasized "the role of chance in history and the evolutionary, path-dependent character of industrial location over time" tended to remain obscure and often untranslated.

Studies of spatial patterns within geography were traditionally ideographic in nature, stressing what was unique about the areas studied, and often making deterministic assumptions – for example, that physical factors such as soils and climate determined the distribution of agricultural activities (Healey and Ilbery, 1990: 20). This approach was subjected to growing criticism during the 1950s, as it failed to explain either the reasons for spatial development or provide the basis for developing general theories. Schaefer (1953) felt that the prevailing regional approach was too descriptive, and that geography needed rules and theories that apply to more than just specific regions. The resulting change in direction of economic geography involved a shift from ideographic to nomothetic studies (stressing the general) and towards emphasis on identifying the processes that account for regional patterns (Healey and Ilbery, 1990: 20).

In seeking clearer explanations for spatial development, geographers were able to turn to two sources of learning:

- Mainstream economic theory, which offered a microeconomic explanation of how firms decide what to produce and how much labour and capital to employ (but not where to locate);
- Earlier studies of location theory by such scholars as von Thünen (1826) and Weber (1909; 1929).

Neoclassical theory was used by both geographers and economists to model patterns of land use (Isard, 1956; Bunge, 1966; Harvey, 1969). These patterns were based on deductive reasoning (from general principles rather than observation); and hence most economic geography models were normative in nature, stipulating what patterns should emerge given a number of assumptions. 'Optimal' location or land-use
patterns were identified, where profits were maximised or costs minimised (Healey and Ilbery, 1990: 20).

Empirical evidence of the tendency of industries and services to gather together spatially has also led to the emergence of new theories, including cluster and New Economic Geography models. At the same time, dissatisfaction with many features of orthodox economics has prompted the development of models based on behavioural and evolutionary theories. Martin and Sunley (2001: 148) criticise the trend they observe among geographers such as Amin and Thrift (2000) to reject mainstream economics in favour of more heterodox socially and culturally based 'economic knowledges'. Table 4.3 maps the various approaches to spatial theory; the main part of this chapter summarises the different models developed; and the Conclusions section considers the implications for policy.

4.3 Taxonomies of Spatial Models

Several different taxonomies of spatial theory may be found in the literature. This thesis adapts the suggestions made by Healey and Ilbery (1990) and McCann (2001). This section first reviews the analysis developed to explain the locational decisions of individual firms. Then it examines the spatial distribution of activities across regions and urban areas. Theories of development in the very long term (Kondratieff, 1935; Schumpeter, 1937; Hall and Preston, 1988) are considered in the review of the influence of time (5.2.4). The literature also includes discussion of the philosophical nature of space and its relations to physical objects and flows between them (Jones, 2005), which is beyond the scope of this thesis.
Table 4.3: Approaches to Spatial Theory

<table>
<thead>
<tr>
<th>Type of Model</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical and Neoclassical Models</td>
<td>Land use models: Von Thünen (1826)</td>
</tr>
<tr>
<td></td>
<td>Location models: Laundhart (1885), Weber (1909), Moses (1958)</td>
</tr>
<tr>
<td></td>
<td>Central Place Theory: Christaller (1933)</td>
</tr>
<tr>
<td></td>
<td>Market Area models: Hotelling (1929); Palander (1935)</td>
</tr>
<tr>
<td>Behavioural Models</td>
<td>Simon (1957); Cyert and March (1963)</td>
</tr>
<tr>
<td>Cluster and Agglomeration Models</td>
<td>Industrial Districts (Marshall, 1920)</td>
</tr>
<tr>
<td></td>
<td>Growth Poles (Perroux, 1950)</td>
</tr>
<tr>
<td></td>
<td>Cluster models: Porter (1990)</td>
</tr>
<tr>
<td>New Economic Geography Models</td>
<td>Core-Periphery Models (Krugman, 1991b)</td>
</tr>
<tr>
<td></td>
<td>Dynamic Local Spillovers (Baldwin et al., 2004)</td>
</tr>
</tbody>
</table>

4.4 Classical and Neoclassical Location Theories

Conventional price theory deals with price and output decisions in a spaceless economy (Richardson, 1978: 38). The introduction of space into economic analysis means that prices for identical goods will vary among locations even in (neoclassical) equilibrium because of the existence of transport costs. Spatial price equilibrium therefore requires the simultaneous solution of both production and location problems. In the simplest example – a firm with one input (produced at M) and converting it in a single stage process for a single final market at C, a profit maximising firm will locate somewhere along the line MC (Richardson, 1978: 53). These models provide a number of explanations of how this simultaneous solution is achieved.

Classical and neoclassical location theories are derived from orthodox microeconomic analysis of price and output theory. McCann and Sheppard (2003) describe two distinct strands of classical location theory, a classification that corresponds with the assessment of Healey and Ilbery (1990: 4):

"There are two main kinds of spatial pattern of economic activity. One is concerned with the location of various establishments including factories, warehouses, offices and shops; while the other is concerned with land uses such as agriculture and forestry ... the distinction is useful because it emphasizes two different analytical approaches. In the first case the economic
activity is taken as given and the analysis is concerned with how its location pattern has evolved; in the second case the location is taken as fixed and the analysis concerns the changing land uses and activities that occupy the location.”

4.4.1 Land Use Models

The first strand, by von Thünen (1826) developed the land use tradition of Ricardo (1821) to differentiate land on the basis of its location rather than its agricultural use. This strand has evolved into modern urban locational concepts, although Isard (1956: 28) concludes that the science of economics suffered from the relative neglect of such methods. This approach is based on the concept of economic rent, defined as the difference between income and costs, which in turn is controlled by distance and the cost of transporting a firm’s output to the market (Healey and Ilbery, 1990: 21). Consequently the theory predicts that bulky and perishable products will be produced close to the market.

One feature of von Thünen’s treatment was that land was considered to be in inelastic supply. The ‘bid rent’ model, produced over a century after von Thünen by a number of scholars, including Alonso (1964), differed notably in that land, labour and capital were regarded as substitutable inputs: in other words, the coefficients describing their contribution to production were variable, not fixed. More complex approaches trace the behaviour of different income groups (McCann, 2001: 113-115). Urban economists have also debated serious critiques of the models, including issues such as land ownership and polycentricity, and whether it is possible to predict the optimal size of a city (McCann, 2001: 123-124).

4.4.2 Company Location Models

The second strand is represented by the early models of industrial location by Laundhart (1885) and Weber (1909; 1929), who sought to explain the decisions of firms, using the standard microeconomic assumptions, including profit maximisation (McCann, 2001: 7).
These theories were based on minimising the cost of transporting inputs and outputs, and may be characterised (using modern terminology) as a form of production function analysis (McCann and Sheppard, 2003: 651). Weber (1909) asserted that, given cost minimisation assumptions, a firm would locate in the place where its overall costs were minimised: given conditions of perfect competition, the firm would be a price taker; so maximum profits would be earned where total input plus output transport costs were lowest.

McCann (2001: 16) shows that (assuming equal prices of land and capital), an interregional equilibrium wage gradient may be constructed, showing the reduction in wages required to compensate the firm for locating at a sub-optimal position. The firm would be indifferent as between locations at any point along the wage gradient. McCann (2001: 16) observes: “Geography confers different competitive advantage on different locations, which can only be compensated for by variations in local factor prices.” Richardson (1978: 56) points out that economies of agglomeration (4.6) could also compensate for higher input costs. The existence of these different forms of competitive advantage imply that perfect competition becomes less likely once spatial variations are taken into account.

4.4.3 Classical and Neoclassical Location Theories: Summary

According to McCann and Sheppard (2003: 652), this twin-track classical approach of Von Thünen, Weber and Laundhardt provides the first ‘fundamental principle’ of location theory:

“In situations where both locational coordinates and production technology exhibit stable and identifiable fixed-coefficients relationships, if transport costs are known or can be calculated, then the fixed-coefficients assumptions embedded in all these models allows the equilibrium factor conditions to be determined. Alternatively, for any given set of factor prices, where production technology is governed by stable and identifiable fixed-coefficients relationships, the conditions under which firms will be mobile can also be determined.”

These two strands of classical location theory, represented by von Thünen and Weber, were not given a genuine ‘neoclassical twist’ until the middle of the twentieth century.
McCann and Sheppard (2003: 652). Von Thünen’s land use framework was extended to provide for substitution between land and all non-land inputs, by Alonso (1964), Muth (1969), Mills (1970) and Evans (1973). Similarly the work of Isard (1956) and Moses (1958) introduced neoclassical factor substitution into Weber’s locational framework, providing for new conclusions whose specific outcome was to depend on the firm’s production function.

McCann and Sheppard (2003: 653) conclude that the Weber-Moses framework provides the second ‘fundamental principle’ of location theory; that where production technology is governed by factor substitution, all location-optimum problems can be considered to be production problems, and vice versa. This principle makes it tempting to strive for some general equilibrium theory of location, but Richardson (1978: 75-76) concludes that the search for a general equilibrium model of location has proved unsuccessful, and notes that some economists believe that regional development is an inherently disequilibrating process.

4.4.4 Central Place Theory

A third strand of thought, central place theory, developed by Christaller (1933), provided a normative approach to modelling the distribution of settlements and services (Healey and Ilbery, 1990: 21). This seeks to explain world-wide evidence of a regular pattern of the numerical distribution and spatial organisation of cities (McCann, 2001: 71). Nations are generally dominated by one or two ‘primal’ cities, which produce most of the various outputs of the economy, while other more peripheral regions focus on smaller cities, more numerous, but each with a smaller range of outputs.

One explanation of this hierarchy was made by Christaller (1933) in his model of central places, which was based on inductive observation rather than deductive extrapolation from first principle (Parr, 2002). The model assumes homogenous land prices and transport costs in all directions, and suggests a hierarchy of different goods, markets and urban centres. Higher order goods tend to be produced in larger centres for larger market areas; and maximum market coverage is provided by a minimum number of production points. Christaller (1933) shows that a hierarchical urban
system can exist automatically with a variety of different sizes of spatial market areas (McCann, 2001: 73). Blakely (1994) describes how the basic concept of Central Place Theory, involving a 'hierarchy of places', in which each urban centre is supported by series of smaller supplying centres, may be applied to policy, as in the designation of one or two regional nodes by the Tennessee Valley Authority.

A model developed by Lösch (1944) refers to an ideal economic region, not the real world. Strict conditions of perfect competition would produce a hexagonal pattern of economic development. However, the process of clustering (4.6) implies concentration of the labour market, and (as recognised by Lösch) this disturbs the regularity of the hexagonal nets, which are based on an assumed evenly distributed population. The hexagons would become denser closer to population centres. Similarly major transportation routes alter the hexagonal pattern to one of irregular polygons, frequently approximating to rectangles (Isard, 1956: 272).

All three strands of classical and neoclassical location theory discussed so far are based on strict (and unrealistic) assumptions, including perfect knowledge, rational economic behaviour and maximization of profit. In each, the key variable is distance, working through transport costs. In the case of the von Thünen and Christaller models, the assumptions include a linear relationship between distance and transport costs, and a homogenous physical environment, often described as an isotropic surface (Healey and Ilbery, 1990: 21).

4.4.5 Market Area Models

The microeconomic approaches to location discussed are producer-oriented – they do not explain the nature either of market location or market price. Nor do they explain why many economic activities may be grouped together in space: instead the analysis has treated market location simply as a single spatial point. The agents underpinning agglomeration behaviour (Marshall, 1920; Hoover, 1948) have generally been treated as exogenous to the micro-location problem (McCann and Sheppard, 2003:653).
However markets themselves possess a spatial dimension, for two reasons:

- Market areas may differ over space because of variations in population density, distribution of income and differences in consumer taste;
- Geography may confer monopoly power on firms, which they are likely to exploit (McCann 2001: 27), another demonstration of how the likelihood of perfect competition breaks down with introduction of the spatial dimension.

A further strand of location theory – market area analysis, developed to deal with inter-firm oligopoly behaviour – does provide insight into the clustering process implied by the second of these points, monopoly power. With customers assumed to be distributed over space, the models of Hotelling (1929) and Palander (1935) provided a game theory method of dealing with the location problem.

The simplest model illustrating the problem of locational interdependence was the famous linear market duopoly outlined by Hotelling (1929) – two sellers of ice cream on a beach (Richardson, 1978: 59). In Hotelling’s locational game theory, each firm assumes competitors will not change their behaviour in response to its own actions (an environment known as ‘Cournot conjectures’). Two firms seeking to maximise market share under such conditions might shift their location until they end up at the same location. Here neither firm has any incentive to change location, because any change would involve a reduction in market share – a situation that, in game theory, is known as a ‘Nash equilibrium’.

McCann (2001: 32) shows that, given the assumptions of this model, these locational changes mean an overall loss of welfare for consumers. In this case, attempts to maximise profits have induced an agglomeration that is in fact wasteful. “Market forces result in a location pattern that is socially wasteful because transport costs are at a maximum whereas the minimum transport cost location (a social optimum) would require the firms to locate in dispersed fashion” (Richardson, 1978: 60). This outcome suggests a possible role for public intervention in the locational decisions of private firms, although it should be emphasised that that these models are based on highly specific assumptions, the relaxation of which can produce contrary results.
Thus Richardson (1978: 60) shows that different assumptions can produce circular rather than straight-line market areas.

Moreover, the stasis predicted by the Nash equilibrium breaks down with more than three firms, as the firms keep changing location (d’Aspremont et al., 1979); and if firms compete by price as well as by location, the equilibrium position becomes unstable: “Unless firms can agree to move away from each other, a price war becomes inevitable (the ‘Bertrand problem’).” (McCann, 2001: 33). Competitor firms will consequently locate next to each other only in situations where price competition is ruled out, either by agreement or by the use of non-price competition. In fact, in such situations of non-price competition, spatial clustering of competitor firms is a natural and familiar process – involving shops and showrooms for such products as clothing, electronic goods, cars, restaurants and furnishings. McCann (2001: 33) notes: “In these industries prices are used to indicate product quality, and to indicate the types of consumers for whom the good is intended. As such, prices in these industries tend to be fixed. Firms are unwilling to compete by lowering prices because this suggests that the product quality is falling, and this may actually have an adverse effect on sales.” Such ‘price placing’, familiar in the literature of marketing (Jobber, 1998: 294-315), often involves firms branding essentially similar products with differentiated packaging and appearance (the ‘Hotelling paradox’) (Hotelling, 1929; McCann, 2001: 34).

Building on the second principle of location theory (4.4.3), McCann and Sheppard (2003: 654) go as far as to claim: “Geographical proximity and industrial clustering are only possible in conditions in which price competition for identical products is not a major feature of the market behaviour.” This judgement is in line with the ‘spatial impossibility theorem’ proposed by Starrett (1978): if space is homogenous, there does not exist any competitive equilibrium with shipments between distant locations (Ottaviano, 2003: 666).
McCann and Sheppard (2003:654) conclude that spatial competition and market area analysis provide the third and fourth fundamental principles of location theory:

- The specific reasons for, and outcomes of industrial co-location or clustering, cannot be understood without a consideration of the cost interrelationships between the co-located firms, which provides the basis for their criticism of cluster and New Economic Geography models (4.8)

- Industrial co-location or clustering can only take place in markets characterized primarily by non-price competition, product heterogeneity, or the existence of transactions costs which militate against efficient price competition.

These results provide two important sets of conclusions:

- Firms engaging in non-price competition are encouraged to cluster spatially;
- Consumers close to a spatial cluster will make a welfare gain compared with those further away.

4.5 Behavioural Theories of Location

The spatial models discussed so far are all based on the assumption that ‘rational’ firms seek to maximise profits and have sufficient information to do so. Disillusion with deductive model-building exercises and assumptions about perfect competition led, in the 1970s, to the adoption of ‘behavioural’ approaches to the study of economic location and locational change (Healey and Ilbery, 1990: 23). Like neoclassical models, behavioural approaches seek generalizations, but these were based on what was thought actually occur at the individual level rather than an idealised (group) view of behaviour. Risk and uncertainty have an important role in such behavioural approaches (Healey and Ilbery, 1990: 25); and the introduction of uncertainty has a considerable impact on locational decisions. Richardson (1978: 61) argues that uncertainty reduces the willingness to risk capital and involves diminishing returns to management decisions; and that these factors will tend to reduce plant size. More important for the spatial context, uncertainty increases with
distance from market; small firms have to rely more on external services; and “both factors reinforce agglomeration in big cities”.

Simon (1952, 1959) argues that ‘bounded rationality’ – lack of information obstructing profit maximisation – is especially relevant for issues of location because of the inherent heterogeneity of land, property and local economic environments (McCann, 2001: 36). In larger firms, where different individuals might have a variety of objectives, the firm may seek to ‘satisfice’, that is to achieve a satisfactory (rather than optimum) performance over a range of measures (Simon, 1957; Cyert and March, 1963; McCann, 2001: 38). Relocation may involve a firm in substantial transactions costs; and given limited knowledge and bounded rationality, the firm may well reorganise factor allocation and activities rather than move its plant (McCann, 2001: 38-39).

None of these behavioural critiques indicate why a firm would have chosen a particular location in the first place, one of the apparent advantages of the classical and neoclassical theories, but McCann (2001: 39) argues that classical and neoclassical models should be interpreted in the light of behavioural critiques – bounded rationality, imperfect information, conflicting goals and relocation costs – all characteristic of the spatial economy. This is the approach followed by Alchian (1950), whose evolutionary treatment of an uncertain market environment, characterises uncertainty as leading to one of two stylised forms:

- An adoptive environment, in which no firm has any information advantage, and where the environment ‘adopts’ the firms best equipped to its needs;
- An adaptive environment, in which some firms (by virtue of their size) use information to their advantage, and where smaller firms take decisions that mimic or dovetail with those of the market leaders.

The real economy is argued to exist between these two stylised forms (McCann, 2001: 40). Alchian’s arguments suggest that while larger firms may choose their location in line with the models of Weber, Moses and Hotelling, smaller enterprise are likely to locate themselves close to major market leaders. Classical and neoclassical models of location therefore appear best suited to explain the spatial behaviour of large firms.
and firms with good information or low costs of relocation (McCann, 2001: 40). The behavioural critique of these models suggests that the ‘follow-my-leader’ actions will encourage smaller firms to cluster together in space close to large companies. Such clustering, however, appears certain to raise the price of local labour and property, so reducing profits and therefore the attractiveness of the area.

4.6 Clustering and Agglomeration

The previous section provided a number of theoretical reasons why individual firms might choose to cluster together, without however specifying which influence might prove dominant. Empirical observation of industrial locational behaviour reveals two distinct factors:

- Most industrial activities are clustered together in space;
- There also tends to be a distinct urban ‘hierarchy’, usually with a single largest city involved in almost all economic activities, followed by larger numbers of smaller clusters which increase in number as their size falls (McCann, 2001: 53).

One possible explanation of such clustering is that economies of scale are place-specific. If all firms encounter constant returns to scale, clustering will tend to increase the price of land and labour and so reduce profits. If, however, firms derive economies of scale precisely because of the presence of other firms in the area, this may more than compensate for the higher factor prices. As well as scale economies, firms may benefit from economies of ‘scope’ where the cost of producing two different products within one firm is lower than the total cost of producing them in different firms (Karlsson and Stough, 2002: 6). Essentially economies of scope reflect the result of several products using the same fixed resource within a firm, and provide a basic motive for product differentiation.
4.6.1 Industrial Districts

Such place-specific economies of scale and scope are generally known as economies of agglomeration or economies of localisation, and their first detailed description were provided by Marshall (1920). Armstrong and Taylor (2000: 104-105) distinguish between economies of localisation (which are gained by a single industrial sector) and economies of agglomeration (which are shared by a variety of industries). A number of other models of agglomeration have been proposed as alternatives to the classifications of Marshall and Hoover.

Agglomeration is more likely to take place in sectors with strong increasing returns, strong (monopolistic) market power and lower trade costs, and where customers and suppliers are mobile. This is because intense increasing returns and strong market power reduce the market crowding effect, while more mobile customers and suppliers amplify the market expansion effect. Lower trade costs reduce both market crowding and market expansion effects, but the former more than the latter (Ottaviano, 2003: 667).

4.6.2 Growth Pole Model

The growth-pole model developed by Perroux (1950) built on concepts originally set out by Schumpeter (1934) (McCann, 2001: 60). Richardson (1978: 164) explains that the model’s origins lie outside the field of regional economics:

“Blatantly borrowing from Schumpeter’s theory of innovations, Perroux argues that the pôle de croissance was confined to abstract space (i.e. economic space), rather than geographic space, which he regarded as ‘banal’”.

The growth pole was defined as a set of industries capable of generating dynamic growth in the economy, and strongly interrelated to each other by input-output linkages around a leading industry (the industrie motrice or propulsive industry) (Richardson, 1978: 164-165).
4.6.3 Incubator Model

Chinitz (1961; 1964) based his incubator model on observations of the industrial structure of Pittsburgh and New York. A similar hypothesis was proposed by Hoover and Vernon (1962), in a study of New York. The argument is that highly diversified industrial clusters provide superior ‘incubators’ for the growth of new firms because such diversification ensures the availability of business services (McCann, 2001: 61).

4.6.4 Product Cycle Model

A product cycle model, describing city clusters (Vernon, 1960) and international investment flows (Vernon, 1966) is frequently used to demonstrate the qualitative aspect of spatial investment patterns. Vernon argues that firms choose their locations to suit the particular stage in the life-cycle of their product. Thus research and development associated with a new product is likely to take place in high-level industrial clusters such as dominant central cities, because of the benefits of informal information spillovers and the availability of highly skilled employees. At later stages in its life-cycle, the product becomes ‘mature’, and the production process can move to other lower cost and lower skilled areas. This suggestion is of interest to regional policy practitioners:

"The product-cycle argument therefore implies that more geographically peripheral areas, which tend to exhibit lower labour costs and lower labour skills, will also tend to have plants producing more mature, less novel, and more standardised products." (McCann, 2001: 62)

4.6.5 Porter Cluster Model

One of the most striking features of regional economies is the presence of geographic concentrations of linked industries or ‘clusters’, defined by Porter (2003: 562) as “a geographically proximate of interconnected companies, suppliers, service providers and associated institutions on a particular field, linked by externalities of various types.”
A well-known model, elaborated by Porter (1990, 1998a, 1998b, 2003) has been developed as an alternative to the standard markets and hierarchies dichotomy associated with Williamson (1975). This formulation of the cluster concept has proved the most influential, not least because of its ‘applied’ and ‘prescriptive’ character (Lagendijk, 1998), although Porter’s own approach has been ‘wholly nonformal’ (Martin and Sunley (2001: 157n). This popularity has led to a global proliferation of policies to promote clustering, promoted by a range of international bodies and drawing on the same theoretical sources (Raines, 2002: 21).

Using case studies from different countries, Porter developed the core concept of the ‘diamond of national advantage’. Examining what issues provided the foundations of national competitiveness, Porter (1990: 77) argued:

“The answer lies in the four broad attributes of a nation, attributes that individually and as a system constitute the diamond of national advantage, the playing field that each nation establishes and operates for its industries.”

These attributes are:

- Factor conditions, such as skilled labour or infrastructure, necessary to compete in a given industry;
- Demand conditions, the nature of home-market demand for the industry’s products or services;
- Related and supporting industries, particularly supplier or related industries that are internationally competitive; and
- Firm strategy, structure and rivalry, the conditions governing how companies are created, organized and managed.
Each of these factors is essential for international competitive success, and in addition the diamond has a clear systemic character, so that individual factors cannot be considered in isolation from each other (Ache, 2002: 9). Porter (2003: 549) finds that the mix of clusters differs markedly across US regions, with major implications for regional prosperity: “The performance of regional economies is strongly influenced by the strength of local clusters and the vitality and plurality of innovation.”

Clustering is argued to assist firms to maximise the transfer of technology and information flows, and is of particular importance for small firms who rely mainly on external sources of knowledge. A cluster “allows each member to benefit as if it had greater scale or as if it had joined with others formally – without requirements to sacrifice its flexibility” (Porter, 1998a: 80). The key feature of the argument is that proximity means that competing firms can observe each other’s activities and spur each other on, so increasing the overall competitiveness of the cluster (McCann, 2001: 62). Better access to staff with relevant skills, closer supply-chain links, local institutional support and enhanced contact with sources of information all lead to economies of scale and scope, while competitive pressures and a highly demanding
customer base force increased productivity. Relationships and contacts mean that firms within the cluster perceive new technological opportunities and changing buyer needs earlier than their more isolated counterparts (Porter, 2000: 261-262).

Clusters are argued to have a particular significance because of different patterns of externalities connecting constituent industries such as common technologies, skills, knowledge and purchased inputs:

"Software, for example, is connected with other IT industries in terms of technology and demand, but also linked with medical devices because software is embedded in many types of devices and software development is crucial to medical device product development.” (Porter, 2003: 562)

This complexity of clustering has important implications for policy, as it may not be sufficient to encourage the development of one particular industry without the presence of other linked industries and related skill bases (6.7.3).

4.6.6 New Industrial Areas Model

An alternative formulation to that of Porter is the new industrial areas model, based mainly on the work of Scott (1988) in the field of urban planning and geography. A number of industrial clusters have emerged as key centres of innovation; and the model suggests that spatial clusters of small firms are more innovative than large firms (Saxenian, 1994). Examples include the electronics cluster in Silicon Valley, California (Saxenian, 1994), the electronics and biotechnology cluster around Cambridge, and the manufacturing region of Emilia-Romagna in Italy Scott (1988). The relationship between company size, innovation and clustering has been of particular interest among policy planners, the main reason being the hope that such developments may be replicated elsewhere (Castells and Hall, 1994). However, scholars disagree about the conclusions (Arita and McCann, 2000), and it is unclear whether the empirical observations represent something genuinely new, or a simply new forms of an old phenomenon (McCann, 2001: 63).
Criticism of Cluster Concepts

Criticism of cluster concepts often focus on the lack of clear theoretical foundations for clusters and related models. Martin and Sunley (2002: 7-9) argue that the popularity of Porter's particular formulation of clustering is based on its "business-and policy-friendly" style, its stress on competitiveness and its vagueness and non-testable nature. The definition offered by Porter (1998a) seems "intentionally opaque and fuzzy" (Martin and Sunley, 2002: 11).

Discussion of both the theoretical background and empirical observation of clustering are complicated by the differing forms clusters may take. In response to this shortcoming, Gordon and McCann (2000) distinguish three main cluster models:

- **Pure agglomeration**, in which firms act atomistically, there are no barriers to cluster membership, and the value of being in the cluster is reflected in local property rentals;

- **Industrial complexes**, characterised by long-term, stable relationships between cluster firms, where entry costs are high, and rental payments not relevant;

- **Social network model**, whose mutual trust and loyalty, with only partially open membership, and the value of membership is only partly reflected in local rents (Granovetter, 1991).

In practice, elements of all three types may be present in the same region or area, making analysis difficult (McCann, 2001: 64). Even the spatial nature of clustering is questioned: thus a study of innovation processes among firms in the Netherlands concluded that networking was more relevant than proximity -- that "a far friend appears to be better than a good neighbour" (De Bruijn and van Oort, 2005: 12). The fuzziness of cluster concepts makes the theory difficult to specify or subject to empirical testing (Martin and Sunley, 2002: 21).
4.7 Urban Hierarchy Theory

This thesis studies what causes some regions to prosper more than others, rather than how spatial distribution of economic activities takes place within urban areas. Therefore it touches only briefly on the considerable literature dealing with the ranking of cities of different sizes within countries and regions, and with the patterns of settlement and economic activities within urban areas. Urban economics has developed as a separate discipline since it evolved from neoclassical economics in the early twentieth century (Balchin et al., 2000: xviii).

A city that captures the most ‘national market area activities’, and at the same time engages in all other activities common to smaller market areas tends to be the largest in size (Isard, 1956: 58; Richardson, 1978: 330). This empirical observation has led to a formulation of a ‘rank-size rule’, that the product of a city’s size and its rank is a constant equal to the size of the primate city, described as that ‘most incredible law’ (Christaller, 1966: 59; Zipf, 1949).

Urban economics have also constructed models to explain why people and firms should pay more to occupy land at certain locations. Empirical observations show that higher land prices and rentals are paid at city centres, and that larger areas of land are used by households and firms further out from city centres (McCann, 2001: 93-94). Detailed description of these urban models, already mentioned in the context of Central Place Theory (4.4.4) is beyond the scope of this thesis.

4.8 New Economic Geography (NEG) models

One distinctive approach to explaining economic location is offered by various new economic geography (NEG) models. New Economic Geography assumes that the comparative advantage of regions are created by dynamic interaction between geographic market potentials and rational firms (Karlsson and Stough, 2002: 1). The mechanisms proposed show clear similarities to cluster theory, and like cluster models they have generated considerable controversy.
NEG models grew out a set of new trade theories developed in the late 1970s and 1980s (Ethier, 1982; 1979; 1980). New trade theories sought to explain intra-industry trade and the predominance of trade flows between larger, developed countries, neither of which can be explained by the Heckscher-Ohlin theorem (3.4). Places differ in terms of their relative abundance of natural resources, proximity to natural means of communication and climatic conditions. This was the explanation proposed by international trade theory, but that seemed “an inadequate explanation of the dramatic differences in economic development that one observes even between areas that are not very different in terms of those exogenous properties” (Ottaviano 2003: 66).

New trade theories focus on industries characterised by increasing returns to scale and by imperfect competition, which, as argued above (3.3.3) may be associated with non-convergence of regional economies. They highlight the competitive advantages enjoyed by regions or countries with the biggest home markets – or the best access to wider markets (Armstrong and Taylor (2000: 135). Without this ‘home market effect’, it is virtually impossible to explain the geographic concentration of firms and regional specialisation (Karlsson and Stough, 2002: 3).

The home market effect benefits industries with economies of scale and an array of differentiated products; and would be expected to produce some combination of higher local wages and increased net exports (Venables, 1998: 3). New economic geography models extend this analysis spatially to explain the geographic clustering of industries: regions given a head start find their market size advantage enhanced by a process of cumulative causation (Armstrong and Taylor (2000: 135), with distance and transport costs playing a key role (Thirwall, 2000: 282). They explain the geographic pattern of economic development between countries, and between regions and countries, in terms of a contest between centripetal forces which lead to industrial concentration and centrifugal forces which lead to industrial dispersal (Krugman, 1991; 1998; Thirwall, 2000: 282).

The NEG models developed by Krugman (1993) and Fujita and Krugman (1995) “simulate Christaller-type general equilibrium results within a monopolistic competition framework” (McCann, 2001: 86-87). Two main variants may be
distinguished – the ‘Core-Periphery’ Model (Krugman, 1991b) and the Dynamic Local Spillover Model (Baldwin et al., 2004; ODPM, 2006: 34). To deal with the challenge of increasing returns and externalities, it is common to use a spatial version of the Dixit-Stiglitz model of monopolistic competition that allows for multiple locations and transport costs between different locations (Dixit and Stiglitz, 1997; Karlsson and Stough, 2002: 2). This model may be elaborated to take account of spatial interaction costs – which include not only transportation costs but also ‘geographical transaction costs’, which for many products is non-linear (Johansson and Karlsson, 2000).

Full exploitation of increasing returns presuppose a market large enough for firms to develop economies of scale, suggesting that such firms will choose to locate in regions with large market potential, and equally that some large regional markets will evolve because firms with economies of scale locate there. So consideration of spatial issues implies a mechanism that generates regional growth and concentration in a self-reinforcing way:

“In this way, a cumulative relationship is established that is driven by the interaction between internal economies of scale, demand growth and geographical interaction costs.” (Karlsson and Stough, 2002: 5).

This suggestion provides a new dimension for the spatial working out of the process of cumulative causation described above (3.5); and (a most important consideration for policy makers) implies that regions with similar production processes may specialise in quite different ways. This argument is developed by evolutionary economists into the treatment of path dependence and ‘lock-in’ (5.5). Once the analysis is widened to take account of the effect of time – or history – this factor points to a significant opportunity for governments and agencies to influence the future course of development.

Different variants of NEG stress the role of different elements of cumulative causation. Thus in the footloose labour models pioneered by Krugman (1989, 1991a, 1991b; Armstrong and Taylor, 2000: 135), once manufacturing firms begin to congregate in a region, mobile labour is drawn in. This further stimulates the home market (‘expenditure shifting’), encouraging more firms to move in. Footloose labour
will tend to check upward pressure on wages; and the region may also benefit from lower prices of goods produced locally (the ‘cost-shifting effect’).

An alternative version is the vertically linked industries model developed by Krugman and Venables (1995) and Venables (1996). In this version, the cumulative causation process is driven not by migrant labour but the growth of intermediate suppliers who form a cluster of firms. This provides the growing region with input-output linkages which help reduce costs and reinforce its competitive advantages (Armstrong and Taylor, 2000: 136).

The patterns of clustered industries produced by such mechanisms had already been observed by Marshall (1890), who stressed the role of localized technological and pecuniary externalities in explaining spatial variation in economic activities. Market prices may not reflect the contribution of these externalities. In the case of technological externalities, they may not be priced at all, but transmitted by sheer proximity: California’s Silicon Valley is cited frequently as an example of such transmission (Saxenian, 1994). Pecuniary externalities require market transactions; but the pricing may be distorted by market power (Ottaviano 2003: 666).

Ottaviano (2003: 666) maintains that local pecuniary externalities are at the core of New Economic Geography models, although modelling their microeconomic parameters has proved difficult. Pecuniary externalities are likely to play a particularly important role in sectors where there are significant trade costs (of transport as well as surmounting administrative and cultural barriers), increasing returns to scale and monopolistic competition. Such localised increasing returns to scale, which may take the form of highly competitive and rapidly growing industrial clusters, provide a major source of comparative advantage for regions (Karlsson and Stough, 2002: 1).
In such sectors, the entry of a new firm will lead to two different (and countervailing) effects (Ottaviano 2003: 666):

- Market expansion effect – a new firm increasing demand for ‘upstream’ inputs;
- Market crowding effect – where the firm creates additional supply downstream, exerting competitive pressure on other companies.

This is analogous to the contest already described in this subsection in which centripetal forces promoting industrial concentration battle it out with centrifugal forces that encourage dispersal. Agglomeration takes place where the impact of market expansion effect dominates the impact of the market crowding effect (Venables, 1966): in other words, the influence of centripetal force overcomes that of centrifugal force. Assume that there are three vertically linked activities – intermediate production, final production, and consumption. If a new firm begins producing intermediates, it will increase labour demand and intermediate supply; so wages rise while intermediate prices fall. Other intermediate producers will suffer a market crowding effect. But final producers will benefit from falling production costs and from greater demand by richer workers (and possibly a growth of population). The entry of new final producers will feed back into stronger demand for the intermediate producers (the market expansion effect). When market expansion effect dominates market crowding, both final and intermediate firms end up being agglomerated in the same place (Ottaviano, 2003: 667).

Krugman (1998) distinguishes a number of significant forces, both centripetal and centrifugal, that affect industrial concentration, and these are set out in Table 4.8.
Table 4.8: Main categories of force affecting industrial concentration or dispersion

<table>
<thead>
<tr>
<th>Centripetal forces</th>
<th>Centrifugal forces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market-size effects</td>
<td>Immobile factors of production</td>
</tr>
<tr>
<td>‘Thick’ labour markets</td>
<td>Land rents</td>
</tr>
<tr>
<td>Pure external economies</td>
<td>Pure external diseconomies</td>
</tr>
</tbody>
</table>

Source: Krugman (1998: Table 1); Armstrong and Taylor (2000: 137)

Centripetal forces include:
- Market-size effects – the advantages of backward and forward linkages;
- ‘Thick’ labour markets – defined as “the advantages that can arise when an industrial cluster leads to the development of a large and appropriately skilled local labour force into which all the firms can tap” (Armstrong and Taylor, 2000: 137);
- Pure external economies – the web of additional external economies produced by exchange of technological expertise.

Centrifugal forces include:
- Immobile factors, including not only land and natural resources but also labour where barriers to mobility are high;
- Land rents, which deter clustering where increased economic activity leads to their being rapidly bid up;
- Pure external diseconomies, including spillovers such as congestion and pollution.

It should be noted that these explanations involve considerable academic controversy. Thus while the processes are similar to those described by Marshall (1890) and Ohlin (1933), Ottaviano (2003: 667) claims that New Economic Geography models transform them into “a general equilibrium model with solid microeconomic foundations”, although this striking claim is disputed by McCann and Sheppard (2003).

Armstrong and Taylor (2000: 137-139) maintain there is empirical evidence to suppose they successfully account for trade and investment between advanced economies which is not explained by the Heckscher-Ohlin theorem.
4.9 Conclusions

Consideration of space introduces a new dimension into economic relationships. This chapter considered the implications of this spatial element for the aims of the thesis and the issue posed by its research question (1.2).

Classical and neoclassical location theories sought to lay out the conditions for solving firms’ location problems, as well as their production problems, in line with standard assumptions of perfect competition. Inconveniently, however, the real world seemed unwilling to comply: notably, these theories failed to explain the strong evidence of firms’ concentrating in certain favoured locations and regions of similar development trading with each other (4.8).

In response, a range of agglomeration and cluster models grew up during the late 20th century, recasting the earlier contributions of Marshall (1920). The formulation proposed by Porter (1990) proved popular among policy makers; while New Economic Geography (NEG) models emphasised ‘tipping points’ at which market forces encouraging concentration might overcome those promoting dispersion.

According to its critics, however, cluster ‘theory’ has grown without solid microeconomic foundations. It may describe what happens, but does not explain why. Unlike cluster ‘theories’, in the view of scholars such as Ottaviano (2003), NEG models do rest on the microeconomic foundations provided by new trade theory; but have proved less attractive among policy makers.

One important theme to emerge is the association between both clustering and NEG models with non-price competition. These representations of the spatial economy involve imperfect competition, implying restriction on the ability of free markets to produce convergence of prosperity. Such restriction provides a more plausible argument for government intervention.
Chapter 5 The Temporal Dimension

5.1 Introduction

This chapter examines how consideration of time affects understanding of the way regional economies work. As with spatial issues, this has implications for regional economic theory and therefore for the way in which the thesis seeks to answer the research question (1.2). This chapter sets out the way in which orthodox economic theory deals with the concept of time, including the question of how to balance present and future needs, an issue crucial for economic policy. It then examines how the process of economic development could be affected both by conjecture about events in the future and the influence of events in the past. One major issue of controversy surrounds the concept of ‘path dependence’, the existence of which is contested on both theoretical and empirical grounds. This examination exposes regional economics to the insight of evolutionary theory; and raises the question of whether the gain in realism afforded by evolutionary models is offset by a loss of predictive power.

5.2 Treatment of Time in Economics

Time is treated in a number of distinct ways by orthodox economic theory. Rather like money, it can be thought of alternatively as a store of value in its own right; or else as a medium applied to utility of present or future consumption (Blyton et al., 1989: 10). It is also perceived as a calibration of change, allowing changes in economic variables to be measured (Hill, 1989b: 57).
The principle of consumer choice suggests that people will make the best use of time when they equalise the marginal utility of the last minute spent on each activity (Samuelson and Nordhaus, 2001: 89-90). Individuals choose the combination of work and leisure (mapped as indifference curves) that best meets their preferences; and competition in the product and factor markets ensures employers offer alternatives that meet these preferences (Hill, 1989b: 58-59).

Time-related costs and efficiency of working time arrangements are prominent in managerial practice, notably the ‘just-in-time’ systems of production scheduling and inventory control. These aim to reduce stocks of raw material, components and finished goods and so minimise any periods of unproductive time in which expensive capital is tied up; such as the North American auto plant that has car seats delivered every twenty minutes (Blyton et al., 1989: 133).

5.2.1 Time preference

Classical and neoclassical theory accords a prominent role to time in the theory of capital and interest. Consumers are thought to possess a rate of time preference – the amount of money sufficient to persuade them to save rather than to spend (Bannock et al. 1992: 422-423). The rate of time preference will be affected by the level of current consumption, the expected level of future consumption, and consumer taste (for example, the belief among some consumers that they will enjoy spending money more when they are young).

The market rate of interest expresses the amount a consumer will be compensated for saving, and rational consumers should save enough for their rate of time preference to equal the rate of interest. Fisher (1906; 1907) argued that the key factors determining the rate of interest were based on two fundamental pillars – impatience, revealed by ‘time discounting’, and investment opportunity, reflected in “the marginal rate of return over cost” (Samuelson and Nordhaus, 2001: 277). This aspect of according a lower value to costs and benefits in the future is a key element in measuring the impact of economic policy, captured by the concept of Net Present Value, which is defined as the present value of an investment’s future net cash flows minus the initial investment. This concept can be used to appraise in advance (ex ante) whether or not
to proceed with a development project. If Net Present Value is positive, the investment should be made (unless an even better investment exists), otherwise it should not (investorwords.com, 2004); although imposition of a budget constraint may rule out some projects with a positive value. Estimation of Net Present Value occupies an important role in the evaluation of regional policy (9.4).

5.2.2 Time as a unit of measure

Many of the references to time in economics are not concerned with its allocation but use time as a proxy for change. Time is employed as a convenient unit for spacing out measurements of economic indicators such as those relating to output, prices and wages (Sharp, 1981: 8).

This use of the concept of time occupies a particular role in comparative static equilibrium analysis, the method favoured by neoclassical economics. This methodology analyses markets or economics in terms of their different equilibrium positions, without reference to the process by which adjustment between equilibria is achieved (Bannock et al., 1992: 79). It typically assesses the characteristics of the equilibrium state, and discovers the position of a new equilibrium when some variable is changed, without tracing the path or speed of adjustment.

In developing a theory of production, neoclassical economics generally makes do with two conceptualised time periods. In the ‘short run’, firms can adjust production by changing variable factors such as materials and labour, but cannot change fixed factors such as capital. The long run is sufficiently long for all factors to be adjusted (Samuelson and Nordhaus, 2001: 113). However Marshall (1890: Book V, Chapter XV: 3) distinguished four time periods, noting that the main problem for formulating a general equilibrium theory of demand and supply was posed by variations in the area of space, and the period of time over which the market in question extended. Marshall maintained that the influence of time was more fundamental than that of space.

Marshall (1890: V.XV.4) defined a “very short period” market, such as that of “a provincial corn-exchange on market-day”. Here the “higgling and bargaining” would
probably oscillate around a mean position, which “would have some sort of a right to be called the equilibrium price”: but the action of dealers in offering one price or refusing another would depend little, if at all, on calculations with regard to cost of production: “They would look chiefly at present demand on the one hand, and on the other at the stocks of the commodity already available.” Marshall’s short and long periods were comparable to those described above, but he also distinguished a very long period, in which technology, population trends, habits and customs are not taken as given, but allowed to vary (Wikipedia, 2004).

5.2.3 Trade cycles

The mainstream treatments of time discussed so far have not dealt with the uncertainty attached to future events or with the possible influence of the past in determining subsequent events. Such a connection between time and the process of change is made in orthodox theory’s treatment of cyclical movements in the economy. The trade cycle is defined as a regular fluctuation in the level of national income, typically of the order of five years, although often on a generally upward cycle (Bannock et al. 1992: 425). Samuelson and Nordhaus (2001: 478) define ‘business cycles’ as “economywide fluctuations in total national output, income and employment, usually lasting for a period of 2 to 10 years, marked by a widespread expansion or contraction in most sectors of the economy”.

The best known ‘internal theories’ of the cause of trade cycles, developed by Samuelson, Hicks, Goodwin, Phillips and Kalecki in the 1940s and 1950s, attributed cycles to the combination of multiplier effects with the accelerator theory of investment (3.4). In certain cases, investment can be related to output one period back, and negatively related to output two periods back, causing an oscillating pattern of income (Bannock et al. 1992: 425).

While detailed treatment of trade cycles is beyond the scope of this thesis, the cause and effect nature of the cyclical mechanism over time points to the significance of actual events in determining later events in the sequence, such as government spending designed to ‘pump prime’ the economy during a period of recession. Classical and neoclassical theory implies that the free market economy will tend to
move back towards competitive equilibrium. This view is in sharp contrast to the views of dissenting economists, whose alternative presentations are described below (5.4).

5.2.4 Long waves

A number of theories have also been adjusted to account for apparent cycles over the very long periods identified by Marshall (1890). Such theories may be grouped into three broad categories, which vary according to the emphasis they place on history, the internal structure of the economy, and international relations (Harris, 1988). These are:

- Long-wave theories (discussed in this section);
- World-system theories concerned with the evolution of the global capitalist system (Harris, 1988); and
- Regulationist theories, which deal with changes in systems of production, such as the move from ‘Fordist’ mass production to more flexible regime of accumulation (Blackburn et al., 1985).

Long-wave theories suggest that economic change occurs in cycles, and is associated with major shifts taking place every fifty years or so in prevailing technologies (Healey and Ilbery, 1990: 14-16). Using commodity prices as an indicator, Kondratieff (1935) identified three such economic waves – the industrial revolution in the first half of the nineteenth century in Britain (subsequently linked with iron and textiles); the Victorian boom and recession of 1847-1894 (railways, steel); and the imperialist expansion of the early twentieth century in Britain, Germany and the United States (electric power, chemicals and the motor industry). Later writers identified a fourth ‘Kondratieff cycle’ from 1940-1990 (petrochemicals, electronics and aerospace), and discussed the possibility of a fifth, based on microelectronics and information technologies (Healey and Ilbery, 1990: 15).

An explanation of these long waves was offered by Schumpeter (1939), who argued that at certain points in history, a bunching of innovations produces rapid economic
expansion and waves of 'creative destruction'. Both Schumpeter (1939) and Mensch
(1979) argued that long waves reflect an internal regulating system within capitalist
economies. As each wave of innovation works its way through the economy, the
market for a particular range of goods becomes saturated, prices fall, and the resulting
recessionary period stimulates a search for new inventions and commercial
application of previous inventions not marketed (hence the bunching process) (Healey
and Ilbery, 1990: 16).

This portrayal of innovation is relevant for the way particular regions gain a
competitive advantage. For neoclassical theory, such an advantage would be
temporary, a lead that would be eroded as others caught up:

“Every successful innovation creates a temporary pool of monopoly. We can
identify innovational profits (sometimes called Schumpeterian profits) as the
temporary excess return to innovators or entrepreneurs. For a short time,
innovational profits are earned. These profit earnings are temporary and are
soon competed away by rivals and imitators. But just as one source of
innovational profits is disappearing, another is born. An economy will
generate innovational profits as long as it produces new products and
processes.” (Samuelson and Nordhaus, 2001: 275)

This neoclassical adaption of Schumpeterian innovation has clear implications for
regional economics, because it suggests a mechanism that might produce regional
convergence. ‘Rivals and imitators’ located in other regions would join in the process
of competing away the gains of innovators in lead regions, enabling their regions to
catch up. If the convergence process were assured, then the need for a specifically
regional policy would arguably be reduced. As shown above however (3.3.4) it is not
possible to demonstrate empirically an unambiguous process of regional convergence,
while the theoretical basis for supposing convergence is also strongly disputed (5.4).
5.3 Criticisms of neoclassical treatment of time

In the view of critics of orthodox theory, however, neoclassical economics had never been concerned with processes of change. Thus Joan Robinson repudiated neoclassical economics (along with her own previous work on imperfect competition) for failing to pay sufficient attention to problems of time and uncertainty (Backhouse, 2002: 315). Keen (2001: 79) maintains that one of the fundamental weaknesses of conventional economics is the absence of time in its analysis, and states: “Economic theory in general ignores processes which take time to occur; and instead assumes that everything occurs in equilibrium” (Keen, 2001: 166). Neoclassical growth theory is described as simply “old-style stationary state analysis in which an element of compound growth is introduced by adding factor-augmenting technical change and exogenous increases in labor supply to an otherwise static general equilibrium model of the economy” despite the fact that “no economy has ever been observed in steady-state growth” (Blaug, 1992: 238). These criticisms echo the views of Veblen (1898) who asked why is economics not an evolutionary science:

“The problem as he and others saw it was that, following Marshall and the marginalists, economics had been colonised by the mechanical metaphor of static equilibrium, exogenous impulses for change, and a desiccated notion of humanity as ‘rational economic man’” (Cooke and Morgan, 1998: 196-197).

According to evolutionary critics, this ‘colonisation’ rendered orthodox economic largely irrelevant:

“Neoclassical economic analysis is based on the concept of equilibrium and the attendant definitions of the economic problem as one of optimal substitution (best allocation) under conditions of known resource scarcity. This is certainly an economic problem, but it is not the main economic problem faced by modern globally connected economies in which competition is mostly about introducing new options for consumers in the face of ongoing uncertainty, and not simply about beating down existing suppliers facing a known opportunity set.” (Potts, 2003: 61)

In similar vein, Nelson and Winter (1982: 8) contend that “reliance on equilibrium analysis, even in its more flexible form, leaves the discipline largely blind to the phenomena associated with historical change”. If there is no correspondence between the steady-state path and the actual historical experience of economic development,
“it is not easy to see how growth theory can be expected to throw light on the causes of unbalanced growth” (Blaug, 1992: 238). If the economic world is in continuing flux, as evolutionary theory suggests, “the normative properties associated with competitive equilibrium become meaningless, just as that equilibrium is meaningless as a description of behaviour” (Nelson and Winter, 1982: 356).

Accordingly, evolutionary economics has produced its own theories to deal with situations of change, open systems and innovation processes (Reschke, 2003: 1). These revived ideas produced many decades before (Martin and Sunley, 2001: 151), with origins that can be traced back at least as far as Veblen (1898), who criticised classical economists and other scientists for seeking to systematise knowledge as a natural law, which “is felt to exercise some sort of a coercive surveillance over the sequence of events, and to give a spiritual stability and consistence to the causal relation at any given juncture” (Veblen, 1898: 4). Following the tradition of Schumpeter, evolutionary economics views the economy as an evolving system. It places a strong emphasis on dynamics, changing structures (including technologies, institutions, beliefs and behaviour) and disequilibrium processes such as innovation, selection and imitation (Economist.com, 2004). It regards the survival of firms as an evolutionary process:

“Because of limited information and bounded rationality, firms cannot deliberately choose strategies which guarantee survival. Actual survival is due to a combination of a corporate philosophy which leads to the choice of strategies appropriate to a firm’s economic environment, and sheer chance.” (Black, 1997: 157)

This requires a theory of transformation and change, not stasis and equilibrium (Cooke and Morgan, 1998: 197). Evolutionary economics “must be a theory of a process of cultural growth as determined by the economic interest, a theory of a cumulative sequence of economic institutions stated in terms of the process itself.” (Veblen, 1919: 77; Hodgson, 1996). A theory of economic change was needed that “does not merely rely on external factors propelling the economic system from one equilibrium to another” Schumpeter (1937: 166). Schumpeter took from Marx the idea that capitalist evolution is driven by technological competition between firms. Marx had asserted that competitive pressure forced firms to increase productivity by
introducing new and more efficient machinery. Those firms that succeeded would improve their competitive position and be rewarded by above average profits, while those who failed would be unprofitable and driven out of the market (Marx, 1887; 1954; Fagerberg, 2002: 7).

According to this line of reasoning, the whole process of economic evolution was thus closely associated with innovation, and the pressures unleashed were very different from traditional text-book competition: rather “it commands a decisive cost or quality advantage and ... strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives.” (Schumpeter, 1943: 84; Fagerburg, 2002: 7). While Marx limited his analysis to mechanisation – process innovation – Schumpeter included other elements such as new products, new types or qualities of raw materials or intermediary products, new markets and new ways to organise business, and developed a theory of how innovations take place, stressing the role of the entrepreneur (Schumpeter 1934, 1943; Fagerburg, 2002: 8-9). Important innovations did not occur at random but tended to cluster in certain time-periods and sectors of the economy; and this was likely to give rise (or contribute) to the discontinuous pattern of growth known as ‘long waves’ (Schumpeter, 1939; Fagerburg, 2002: 20).

Innovations alone did not provide the force for economic development. Such impulse was also provided by new and rebuilt plants, new firms and new people (Schumpeter 1964: 68-71). Diffusion of innovation takes place by what Schumpeter described as the “swarming” process – the entry of a large number of imitators. This process might be delayed by a decade or more until profitability is clearly demonstrated or basic organisational or related social changes occur. Once swarming occurs, it has powerful multiplier effects in generating additional demand, which in turn may induce a further wave of process and applications innovations. “It is this combination of related and induced innovations which gives rise to expansionary effects in the economy as whole” (Freeman et al., 1982: 65).

However, the ‘wave’ analogy, which requires a strong regularity, has not been proven. The empirical basis for assuming that ‘long waves’ in the GDP of the world economy actually exist, is weak (von Tunzelmann, 1995). Fagerberg (2002: 23) concludes that
the strong mechanistic flavour of long wave literature is arguably quite alien to the evolutionary approach.

Yet although academic interest in long waves has faded, applied researchers have taken up Schumpeter’s emphasis on the cumulative and path-dependent nature of innovation (Dosi, 1988). Fagerberg (2002: 24-27) provides a survey of the different accounts of innovation systems. The variety of approaches suggests that this remains a somewhat formless area of study that is still evolving. However, one frequent theme is the importance of close or ‘dense’ and enduring relationships between firms, customers and suppliers (Lundvall, 1992). For example, Etzkowitz and Leydesdorff (2000) emphasised the role played by the interaction between universities, firms and governments in regional and local knowledge agglomeration (the so-called ‘triple-helix’ approach), while Braczyk et al. (1997) studied the regional basis that innovation systems often display.

A parallel approach is provided by insights on the importance of institutions for economic change. Commons (1934) put forward the concept of diverse pathways to economic development being institutionally determined, while North (1993) emphasised the role of trust, co-operation and institutions. This approach equates institutions with rules and behavioural norms, including the way these rules and norms are enforced; the key point being that these minimise transactions costs, which accounted for 45 per cent of US GDP in 1970 (North, 1993; Cooke and Morgan, 1998: 197).

A number of themes emerge in the way evolutionary theory treats economic development: thus (Cooke and Morgan, 1998: 199-200) note that evolutionary economics lays stress on the role of:

- Processes of change that are determinate (rather than exogenous), cumulative and pursue diverse economic pathways;
- Disequilibrium situations in markets and institutions, with changes tending to be cumulative and path dependent (5.5);
- Because of imperfect information, tending to deal with radical changes by opting for agglomeration;
• Imitation and innovation processes, both radical and incremental;
• Institutions – in the sense of rules, norms and routines;
• Social processes of interactive learning;
• Non-mechanical and non-biological theorising, with the aim of avoiding the use of analogy.

Attempts to develop formal models of economic evolution date back to the 1970s and a series of papers later summarised by Nelson and Winter (1982), which emphasised technological competition as the driving force of capitalist development. The firms in Nelson and Winter’s models compete by reinvesting their profits in new and more productive technology. Those that succeed are rewarded by higher profits and faster growth, while those who fail risk being eliminated altogether.

5.4 Dealing with the Future – Uncertainty and Bounded Rationality

Like other critics of neoclassical theory such as Simon (1959, 1965) and Cyert and March (1963), Nelson and Winter (1982) dispense with unrealistic assumptions about human behaviour, such as perfect knowledge. People, they argue, are simply not able to calculate the consequences of all possible actions. What they actually do is to practice a less demanding type of decision making called ‘bounded’ (or ‘procedural’) rationality. One notable example of this is ‘satisficing’ behaviour – sticking to a familiar behavioural rule as long as it leads to a satisfactory outcome.

Actors may follow all kinds of rules, but only those that lead to a successful outcome, i.e. are profitable, will in the long run survive the competitive struggle (Alchian, 1950). One neoclassical response to this assault on orthodoxy was to justify the existence of an unconscious ‘rational behaviour’ by results:

“Unless behaviour of businessmen in some way or other approximated behaviour consistent with the maximisation of returns, it seems unlikely that they would remain in business for long” (Friedman, 1953: 22).

Winter (1964; 1971) argued however that evolution and selection took time. To explain how one type of behaviour becomes dominant, economic theory must allow
for different types of behaviour to persist – or reproduce – over time, on the lines of the role of genetic inheritance in evolutionary biology. Well-specified routines among firms (such as hiring and firing, ordering new inventory, or stepping up production of products in high demand) play the same role as genes; while more unpredictable events are accommodated by recognising that there are stochastic elements both in decision taking and in the outcomes of those decisions (Nelson and Winter, 1982: 14). Which type of behaviour, if any, eventually become dominant would depend on the characteristics of the different types of behaviour and the relevant market conditions – the selection environment. Different market conditions could easily lead to different types of behaviour being selected (Winter 1971: 244-245).

Orthodox theory’s microeconomic assumptions of equilibrium and maximisation mean that it fails to deal with economic change (Nelson and Winter, 1982: 400). For example, it is unable to explain the effect of increased energy prices on economic actors: in reality, firms will try to do things they had previously not seriously considered, and will differ in the extent to which they can cope (Nelson and Winter, 1982: 403). Instead a dynamic, evolutionary theory of economic change was proposed, recognising the diversity of firms, and based on three concepts:

- Routines, playing a role comparable to that of genes in biology, which cut across the orthodox ideas of capabilities (techniques available) and choice (the maximisation part of orthodox theory of the firm);
- ‘Search’, activities that lead to change in routines (and may be portrayed as playing the role of ‘mutation’ in biology); and
- The ‘selection environment’, defined as the ensemble of considerations which affects the well-being of an organisation and determines whether it contracts or expands (Nelson and Winter, 1982: 400-401).

This theoretical debate points to the possibility of one crucial issue that underlies much current debate in regional economics. Uncertainty – such as, will a war break out? – is fundamentally different from the risk associated with probability, such as the outcome of the roll of a dice (Keen, 2001: 200). Not only are outcomes uncertain, but there may be a number of different possible outcomes with very different
characteristics – multiple equilibria. Initial conditions may play an important role in determining outcomes. This is a very different scenario to the prediction of neoclassical theory which suggests that, regardless of initial conditions firms, industries and regions will be guided to a determinate equilibrium position or equilibrium pathway; in the words of Liebowitz and Margolis (1995), “the neoclassical model of relentlessly rational behavior leading to efficient, and therefore predictable, outcomes” (Puffert, 2000: 4n).

The neoclassical approach reached the peak of its ambition in the general equilibrium theory proposed by Walras (1877) which took the form of system of simultaneous equations that given sufficient information could solve supply and demand in goods and factor markets by a process described as *tatonnement* ('groping' or adjustment in accordance with successive prices cried out by an imaginary auctioneer). The implication is that progress towards general equilibrium would also produce the convergence of regional economies, in line with the prediction of the neoclassical model that factors of production will migrate to those regions where their relative price is highest and where they are therefore able to gain the greatest returns (Armstrong and Taylor, 2000: 72) (3.3.1).

Potts (2003) maintains that uncertainty and bounded rationality break down the implicit neoclassical assumption of integral space. Economic systems have multiple agents with different mental maps of the world. For example, consumers cannot possess a prior preference map of the many products of a supermarket; and of the few products the brain can account for, the selection process cannot be regarded as optimal because it is local and time dependent: all elements cannot be connected. Uncertainty cannot be brushed aside by assuming a complete set of forward contingent markets; rather cognitive imagination matters (Schmid, 2003: 1). The argument is that increasing returns and path dependency destroy general equilibrium possibilities – and with them the inevitability of regional convergence.
5.5 Dealing with the Past – Path Dependence

Moreover, the vision of general equilibrium is ill-equipped to deal with the sort of shocks to the economy envisaged by Schumpeter. This can be demonstrated by instances where ‘market success’ was cumulative or self-reinforcing:

“There was typically more than one long-run equilibrium outcome. The one arrived at was not predictable in advance; it tended to get locked in; it was not necessarily the most efficient; and its ‘selection’ tended to be subject to historical events. If the problem was symmetrical in formulation, the outcome was typically asymmetrical.” (Arthur, 1994: 2)

The likelihood of multiple equilibria is often associated in the literature with the alleged occurrence of increasing returns, whose existence was treated with scepticism among economists until the 1980s (Fagerberg, 2002: 30). Arthur (1994: 1-2) recorded that, although examples with increasing returns and non-convexities were mentioned from time to time, “in the main they were treated like the pathological specimens in labelled jars that used to be paraded to medical students – anomalies, freaks, malformations that were rare, but that nevertheless could serve as object lessons against interference in the natural workings of the economy.”

Increasing returns disturb the previous orthodox acceptance that all properly specified economic problems should show a unique equilibrium solution, an event that demanded sufficient convexity, i.e. diminishing returns at the margin (Arthur, 1994: 1). Increasing returns by contrast introduce the possibility of multiple equilibria and an indeterminate outcome. Arthur (1994) maintains that increasing returns may well result from research and development investment, learning by doing, or ‘externalities’, discussed in 4.6.5. In such a situation, small differences in initial conditions may determine long-run outcomes.

By contrast, the orthodox understanding of how markets and businesses operate, passed down more than a century previously ago by a handful of European economists – was said to be based squarely upon the assumption of diminishing returns:
"...products or companies that get ahead in a market eventually run into limitations, so that a predictable equilibrium of prices and market shares is reached. The theory was in rough measure valid for the bulk-processing, smokestack economy of Marshall's day. And it still thrives in today's economics textbooks." (Arthur, 1996: 2)

But modern economies have become divided into two interrelated, intertwined parts, each of them dominated either by diminishing or by increasing returns (Arthur, 1996: 1). Increasing returns generate instability rather than equilibrium:

“If a product or a company or a technology – one of many competing in a market – gets ahead by chance or clever strategy, increasing returns can magnify this advantage, and the product or company or technology can go on to lock in the market.” (Arthur, 1996: 1).

If true, this depiction is of great importance for regional development and policy.

Moreover, initial advantages with increasing returns may provide firms or regions with a lead that others cannot overhaul – even if the technology at their disposal is superior to that of the early market leader. In other words, with increasing returns, there is no guarantee that the solution ‘selected’ by market forces will in any sense be ‘optimal’ (Fagerberg, 2002: 30). One example frequently given in support of this contention is the QWERTY keyboard (named for the first letters appearing on the top row of keys), initially introduced to slow down typing to avoid jamming of the type keys on mechanical typewriters (David, 1985; Arthur, 1994). The QWERTY standard prospered long after this problem ceased to exist, and this despite the existence of allegedly more efficient keyboards – arguably a typical example of how network externalities and associated ‘switching cost’ may lead to selection of an inferior technology (Fagerberg, 2002: 30).

It was David (1985)’s study of the QWERTY keyboard that first gained widespread attention among economists for the concept of path dependence. Path dependence is described as a process of economic allocation where the history of the process has lasting effects on subsequent allocations: “A path-dependent process does not converge to a unique, globally stable equilibrium allocation or even, in stochastic contexts, to a unique limiting distribution of allocations” (Puffert, 1999: 1). As a positive definition, David (2000) suggests that a path dependent stochastic process is
one whose asymptotic distribution evolves as a consequence (or function of) the process’s own history.

Crucially, “instead of converging to a determinate, predictable, unique equilibrium, such processes have multiple potential equilibria, and which one is selected depends on the specific history of the process” (Puffert, 2000: 1). Positive feedbacks among agents’ choices lend persistence and, indeed, increasing impact to particular early choices and other events; and so path dependence means that apparently insignificant events and choices can have huge consequences for the development of a market or an economy (economist.com, 2004).

Under increasing returns, many outcomes are possible. Insignificant circumstances become magnified by positive feedbacks to ‘tip’ the system into the actual outcome ‘selected’: “The small events of history become important” (Arthur, 1989: 127). Such ‘contingent events’ are defined as events not necessitated by systematic a priori factors (Puffert, 2000: 4-5). Outcomes may also be affected by the sequence of events – they are ‘non-ergodic’ (Liebowitz and Margolis (1995: 13).

Significantly, it is argued that there is a strong tendency towards ‘path dependence’ in the patterns of specialisation and trade between regions (Karlsson and Stough, 2002: 5). As argued in the previous discussion of spatial issues (4.8), the New Economic Geography approach allows regions with similar production processes to specialise in different ways, and so regional specialisation becomes to some extent a matter of historical accident. The focus of a particular industry – with the exception of industries dependent on natural resources – is to a large extent indeterminate and history-dependent. Once a pattern of specialisation is established, for whatever reason, such a pattern gets ‘locked in’ by the cumulative gains from international trade. In short, ‘history matters’ (Karlsson and Stough, 2002: 5).

The ideas of path dependence and lock-in provide a major battleground for contemporary economics. Puffert (1999: 5) notes that advocates of path dependence tend to argue that demand-side increasing returns produce positive feedbacks in market share, so that contingent events, including differences among corporate strategies, decide which alternative becomes the standard. By contrast, “opponents of
path dependence argue that the purposeful behaviour of supplying firms rules out path-dependent explanations and that competitions are won simply by offering the best products” (Puffert, 2000: 4). Liebowitz and Margolis (1995) set out a neoclassical response, disputing both the theoretical nature of path dependence and much of the claimed empirical evidence; and suggesting that the inferior outcomes predicted by models of path dependence are often not true equilibrium positions. This interpretation is contested by the evolutionary school (Puffert, 2000: 4-6).

While David (1985) regarded path dependence as a dynamic feature of an allocation process, whether or not the resulting allocation is efficient, it was widely interpreted by others primarily as a source of market failure, through lock-in to a suboptimal technology (Puffert, 1999: 4). (Market failure is defined in 6.4.1). This interpretation was adopted by some supporters of the concept, as well as such neoclassical critics as Liebowitz and Margolis (1995), who argue that competition in such areas as videocassette recording (between the rival VHS and Betamax systems) have actually been won by firms offering the best products rather than as a result of path dependence.

Uncertainty about the future, as well as the past, plays a role in the idea of path dependence. Lack of foresight and bounded rationality greatly limit pursuit of a rationally chosen equilibrium. If agents have perfect knowledge of future events, including technological possibilities, “future markets clear at time zero, leaving no deciding role for the dynamics of the process as such, and thus there is no path dependence” (Puffert, 2000: 5). The development path itself becomes an object of rational choice for optimising agents. Increasing returns may still produce multiple potential Pareto-optimal equilibria, but selection among them takes place through some process of formation of rational (and subsequently fulfilled) expectations (Puffert, 2000: 6). But in the real world, this perfect foresight does not exist – new technologies and market opportunities are progressively revealed, rather than being clear in advance.

Evidence of path dependence has been investigated in a range of products and technologies, such as nuclear power (Cowan, 1990); electric power (alternating- and direct-current systems) (David, 1990); and regional standard railway track gauges.
(Puffert, 1991). It has also been applied to regional competitiveness, as in the establishment of Silicon Valley and other concentrations of industrial activity (Krugman, 1994). These locations are not the result of systematic advantages but rather of accidental origins reinforced by agglomeration economies and other factors that lead new firms to locate in the vicinity of similar established firms (Puffert, 1999: 6). Krugman (1994b) also discusses how path dependence applies to patterns of international trade (Arthur, 1994: 5). Puffert (1999: 6) concludes that while historians had long believed that individual events make a difference in shaping the world, “economists have been hindered from such a view as a result of mechanistic models featuring unique equilibria and no decisive role for dynamics”.

The concept of path dependence injects particular relevance into the new growth theory advanced by the economic geography models of Krugman (1991b) discussed above (4.8). This analysis begins with international trade, in particular the growing phenomenon of inter-industry trade specialisation among developed economies, which results from imperfect competition and increasing returns to scale. Cooke and Morgan (1998: 200-201) maintain that, unlike recent neoclassical economists, Krugman’s treatment of imperfect competition takes the spatial dimension of economic theory seriously, focusing on regional disequilibrium, and using the concepts of externalities and cumulative causation to explain industrial location and path dependence. However, from an evolutionary standpoint, Cooke and Morgan (1998: 201) find a number of unconvincing aspects which economic geography models share with neoclassicism more generally, such as treating imperfect knowledge and competition as if perfect in order to enable the model to function.

5.6 Conclusions

The introduction of a time dimension into economics requires consideration of future uncertainty and the influence of past events. While disputed by neoclassical critics, advocates of such concepts as bounded rationality and path dependence raise further doubts about the ability of economic systems to produce optimum equilibrium positions. This in turn has implications for the appraisal and evaluation of regional economic projects and the way in which this thesis approaches its research question (1.2).
For regional economics, the most interesting feature of the alleged existence of path dependence is the possibility it raises of regions stranded at different local equilibrium positions. Evolutionary economists contend that this is more likely with increasing returns (5.5), which are associated with innovation and the knowledge-driven economy.

Puffert (2000: 2) speaks of some empirical cases that “result not in a single, ‘global’ de facto standard but rather in multiple local or subnetwork standards”: some cases of standardisation or ‘lock-in’ appear permanent, while others have given way to new standards, sometimes showing a tendency to converge to an optimal technique. Whatever the conclusions drawn, it must be conceded that this description tallies with the mixed evidence reviewed above (3.3.4), which shows cases of regional convergence coexisting with divergence.

Clearly from a policy point of view, it is in a region’s interests to avoid a situation where it becomes locked into an outmoded technology or stuck in a suboptimal position. It may therefore be worth an additional tax burden and risking distortion if (admittedly, a big if) governments can somehow manage contingent events so that their nation or region progresses to a more favourable development trajectory. Successful management of such intervention requires an effective evaluative system.
6 Regional Economic Policy

6.1 Introduction

This chapter examines the basis of regional economic policy, relating it to its theoretical base in the concept of social welfare. This prepares the way for considering appraisal and evaluation methodologies, in line with the issues raised by the research question (1.2).

Consideration of regional economic policies is required to understand how policy instruments are intended to operate, and (especially relevant for the research question) because there may be different ways of evaluating different policies. It is important because of the magnitude and prevalence of regional economic policy. For example, while there have been notable changes in economic thought and government policy during the twentieth century, regional economic policy in the United Kingdom has proved to be surprisingly tenacious, remaining on the statute books for over 70 years (Armstrong and Taylor, 2000: 205). As indicated in the Introduction (1.1 and 1.3), a considerable amount of public money is devoted to regional economic policy within both the UK and the European Union.
The existence and the implementation of regional economic policy are therefore matters of considerable significance and controversy, both academic and political. It is important to ensure clarity about the objectives of regional policy instruments, and how their performance should be measured.

This review examines the theoretical base for policy, offering a revised taxonomy of regional economic policies, integrating measures that seek to develop entrepreneurship, knowledge and clustering with more traditional approaches concerned mainly with subsidizing factors of production.

It examines the link between economic policy and social welfare, and considers two broad frameworks within which formal public policies are delivered, the ‘market based’ and ‘market failure’ frameworks. It then examines the main approaches to regional policy, relating each back to the theoretical models discussed, and suggesting that policies may be classified in a number of ways – micro or macro, ‘carrot’ or ‘stick’, or according to the factors of production or networks they subsidise. This treatment does not repeat the detailed summaries of the history of UK regional policy in the literature (Moore et al., 1986; Taylor and Wren, 1997; Armstrong and Taylor, 2000; and MacKay, 2003). Its aim is rather to highlight those areas of policy which are relevant to the theme of appraisal and evaluation of projects and programmes. The section therefore concludes with a brief review of the development of UK regional policy.

6.2 Theoretical Base for Policy

Public policy interventions have two major purposes which, by precedence and public acceptance, have come to be recognized by society as legitimate (Bjornstad, 2004). These two purposes are to improve efficiency and to improve equity. A policy is said to be efficient if it maximizes the total net benefits (benefits less costs) available to society, independent of who receives the net benefits. Equity, on the other hand, is not concerned with the ‘size of the pie’, but on how the pie is distributed among the members of society (Bjornstad, 2004: 2 section 1). Intervention may in turn be delivered by policies that help
markets work more efficiently or which counteract market failure. This chapter will argue that this distinction is frequently misplaced because it ignores political and institutional realities. It also suggests that intervention justified by appeals to market failure is often more properly attributable to social equity.

Intervention may seek to promote overall 'people prosperity', or else be spatially targeted to enhance 'place prosperity' (Winnick, 1966). Targeting 'people prosperity' implies policies such as assisting the unemployed, the poor or underprivileged minorities in every region without distinction, while 'place prosperity' would support poorer regions, inner city ghettos etc. Although policies aimed at securing spatial equity may therefore promote 'people prosperity', there may also be conflict, for example where subsidies to assist migration to areas of greater job opportunity (and so increase 'people prosperity') may depress 'place prosperity' in the region from which migration takes place (Richardson, 1976: 221-223). In a similar approach, Schofield (1989a: 87) classifies policies broadly as either labour supply or labour demand, characterized respectively as 'taking workers to the work' (roughly, an aspect of 'people prosperity') and 'taking work to the workers', which is clearly a case of place prosperity and includes regional economic policy.

6.2.1 Efficiency

Economic efficiency is a measure of the net contribution of an activity or project to overall social welfare. Bjornstad (2004: 3 section 1) offers the example of an objective that increases social welfare – such as flood protection – being pursued by more than one policy, such as:

- inexpensive educational programmes informing people of their rights, such as flood insurance;
- moving communities to less flood prone areas; or
- increasing taxes to build a dam to protect property and other aspects of quality of life by preventing flooding.
While each of these measures pursues the same overall social welfare objective, they may each have different levels of effectiveness. Thus dam building might “disrupt local ecosystems, increase downstream flooding and can damage or destroy other valuable resources, as well as costing significant amounts to construct” (Bjornstad, 2004: 3 section 1). This raises the issue of how to assess the value of actions that raise the welfare of some people while making others worse off.

Economists have tended to focus on two particular criteria for determining what is the best outcome in such circumstances – (1) a Pareto improvement, and (2) a potential Pareto improvement.

6.2.1.1 Pareto improvement

A project is said to constitute a Pareto improvement if it improves the quality of life for some people, but does not make anybody worse off. Clearly, society should pursue all attainable Pareto improvements, because they help some people, but do not hurt anybody. However, in a complex, modern society, every project or policy is likely to place some segment of society at a disadvantage, if only because of the tax revenue required to finance it. Rigid application of the concept of a Pareto improvement for screening potential projects or policies would make it unlikely that any policies or projects would meet the criterion and be implemented (Bjornstad, 2004: 3 section 1).

6.2.1.2 Potential Pareto improvement

Economists have therefore suggested a less stringent criterion for determining if a project or policy improves the welfare of society as a whole. This criterion, the potential Pareto improvement (also known as the Hicks-Kaldor criterion) lies behind conventional cost-benefit analysis (Sloman, 1997: 349):
“A policy or project is said to constitute a potential Pareto improvement if those who benefit as a result of the project or policy gain by more than the losses of those who were made worse off as a result of the project or policy. This type of arrangement of costs and benefits is called a potential Pareto improvement, because those who gain could compensate the losers for their losses, and still be better off. In fact, if the winners did compensate the losers, the potential Pareto improvement would become an actual Pareto improvement.” (Bjornstad, 2004: 3 section 1)

6.2.2 Equity

Unlike efficiency, which reflects aggregate gains, measures of equity seek to determine if costs and benefits are systematically reallocated in ways that discriminate against citizens least able to protect themselves, or in favour of citizens who already enjoy advantage:

“Thus, some potential Pareto improvements may be deemed undesirable no matter how large the difference between gains and losses. For example, many would argue that a policy which benefited Bill Gates by 10 billion dollars, but increased the taxes of middle class tax payers by one billion dollars would be undesirable, even though it generated net benefits to society as a whole of nine billion dollars.” (Bjornstad, 2004: 2 section 4)

The concept of equity may be also be broadened to take into account a sense of spatial equity. This could include global equity (such as the ‘Make Poverty History’ campaign during 2005). Equally it may apply to the efforts of the European Union to promote cohesion between Member States or to regional policy, whether across the EU or within Member States.

6.3 Intervention – For and Against

The previous subsection set out the conditions under which a policy may increase social welfare, whether by acting on efficiency or equity. It remains to be demonstrated, however, whether or not public intervention in the shape of economic policy actually improves on the performance of free markets left to their own devices. As the review of economic theory has shown, there is dispute between economists about the need for government intervention; with orthodox theory implying that markets are best left to
themselves. Such is the case in the debate between post-Keynesian and monetarist schools on the wisdom of government demand management (Sloman, 1997: 486).

In the field of regional economics, the neoclassical models discussed above (in Chapter 3) offer two key concepts – equilibrium and mobility; which suggest that in the absence of restrictions, capital should flow from high-wage to low-wage cost areas, while labour will be attracted in the opposite direction towards areas offering greater rewards. This argument forms part of a wider approach towards the economy as a whole, relying on the concept of ‘market clearing’, but without the qualifications attached to the concept by its originator, Adam Smith (MacKay, 2003). On the contrary, according to MacKay (2003: 308), describing ‘counter-revolutionary’ opponents of Keynesianism:

"The equilibrium assumption, which lay at the heart of the Counter Revolution belief, emphasised that all productive resources released in the process of structural change should find alternative use. This assumption was always heroic, it becomes increasingly difficult to reconcile with reality when unemployment (open and concealed) climbs and persists."

The notion of such a market-clearing equilibrium, especially in the highly elaborate form proposed by Walras (1874) (see 5.4 above) may appear so remote from reality as not to be seriously countenanced. Yet this simple yet powerful view of the world provided the impetus for the Radical Right’s ‘Counter Revolution’ that helped overthrow Keynesianism and lent intellectual force to attempts to ‘roll back’ the State under Margaret Thatcher in the UK and Ronald Reagan in the US. Part of that process of rolling back, in the UK at least, was the running down of regional policy which, as Moore et al. (1986) show, began under the Labour government in the 1970s and continued under the Thatcher administration.

6.4 Policy and Neoclassical Regional Convergence

According to the neoclassical models – and confining analysis to only two factors of production, labour and capital – if markets operate effectively, regional disparities should be automatically eliminated by three mechanisms (Armstrong and Taylor, 2000: 209):
Wages would fall in areas of high unemployment, and rise in areas of low unemployment, tending to even out variations in employment;

Workers would migrate from low wage to high wage areas; and

Firms and mobile capital would move from high wage to low wage areas (in the opposite direction to labour).

The implication of these spaceless models is that in a perfectly competitive world, with no barriers to labour and capital mobility, all regions would tend to enjoy equal wages and full employment. This outcome would obviate the need for regional economic policies, which would serve only to distort the workings of the market and so produce a lower level of social welfare. While such models can be used as a basis for arguing for removal of barriers which harm deprived areas, they are more usually employed as a case for allowing capital to move freely and not spending resources to ‘save lame ducks’ (Blakely, 1994).

Even with the strict conditions imposed by neoclassical analysis, however, it is easy to see why markets may not work flexibly in the real world. There are formidable barriers to migration of both labour and capital. Many unemployed (or low paid) workers will not wish to leave their homes for more expensive or uncertain accommodation elsewhere. Wages are likely not to be flexible, because of minimum wage levels, the existence of benefit systems, and industry-wide collective agreements that transcend geographic areas (Armstrong and Taylor, 2000: 210).

The UK provides ready evidence both of immobility of labour and inflexible wages. In 1973-76, at the end of a 30-year active period of regional policy, average earnings for men in full-time employment in Wales were actually 1 per cent higher than in the West Midlands. However, because of a male work gap between Wales and the West Midlands of 5 jobs per 100 men of working age (and presumably a lower level of economic activity among women), GDP per head of population in Wales was nine percentage points lower than in the West Midlands (MacKay, 2003: 307). This evidence of a divide in prosperity
provided a continued rationale for regional intervention to promote greater equality of employment opportunities.

Although considered more mobile than labour, capital too may be reluctant to migrate: for example, firms may not wish to face the disruption of a move and the loss of locally established input-output linkages (Armstrong and Taylor, 2000: 210). Moreover, there is evidence that owners of capital, such as 'business angels', prefer to take up investment opportunities located nearby (Bayoumi and Rose, 1993).

Even the increased importance attached to knowledge does not guarantee a more even spread of wealth. Sotarauta and Hukkinen (2002: 6) contend that knowledge creation leads to polarised regional development: “Socialised processes of knowledge production suggest that regional development is not leading towards a process of convergence and that the free flow of information and knowledge does not stimulate economic convergence.” This is often explained by the contrasting institutional environments of different regions: technical progress is thus determined by a collective learning process in which many individuals interact and exchange ideas and information (Armstrong and Taylor, 2000: 86). In such an environment, information passes rapidly from one agent to another, ensuring a wide variety of new ideas (Rauch, 1993). Interest in the concept of tacit knowledge, embedded in certain social environments (Granovetter, 1991) has generated policies aimed at promoting innovation and learning, and seeking to contribute directly to economic growth.

The implication of this reasoning for policy is that different measures may be needed to encourage innovation and enterprise; and also that it may not be easy to transplant processes that require possession of embedded knowledge.
6.5 Rationales for Intervention

The case for intervention is usually based on one of two rationales – that of market failure or of clearly defined government distributional objective based on equity considerations (HM Treasury, 2003: 11; Graham, 2006). However, as argued in this section, arguments based on alleged market failure may often prove flimsy.

6.5.1 Market failure

The market failure approach provides the most familiar (if dubious) framework for putting forward a case for public sector support for businesses. Market failure has been defined as “A situation in which economic efficiency has not been achieved through imperfections in the market mechanism” (Bannock et al., 1992: 274); and in the UK, regional policy has been seen very much in this context, with market failure referring to a situation where “the market has not and cannot of itself be expected to deliver an efficient outcome; the intervention that is contemplated will seek to redress this” (HM Treasury, 2003: 11).

Market failure may manifest itself either in the inability of the system to produce goods which are wanted or by a maldistribution of resources which could be improved in such a way that some consumers would be better off and none worse off; in other words resource allocation is not ‘Pareto optimum’. Thus there may be a ‘failure’ to produce a certain aggregate level of output; or a ‘failure’ to achieve a particular desirable pattern of distribution. Economic theory predicts that markets will usually fail in some sense except under conditions of perfect competition, and so the term is reserved for cases where it is believed that a serious maldistribution of resources has occurred. In these cases, it is often argued that government intervention is justified.

Four varieties of market failure are proposed: (a) public goods, (b) externalities, both positive and negative, (c) imperfect information and (d) market power (HM Treasury, 2003: Annex 1; Graham, 2006). Intervention is justified by any one of these forms of
market failure or by promotion of social equity by such means of local or regional regeneration. Examples of market failure in the field of education are provided by DfES (2004: 17), while in the context of regional development the UK Government acknowledges that market failures need to be addressed at the appropriate spatial levels (DTI, 2006: 3n).

The relatively tight definition of market failure discussed above does not extend to the sense in which it was frequently employed outside academic circles in discussion of government economic policy. For example, Bannock et al. (1992: 27) notes that the concept of market failure has been widely used to justify government intervention:

"The government has the responsibility to provide goods that the market will not supply in sufficient quantities, and to use its power to tax in order to correct the market mechanism. In the 1960s such beliefs fitted in well with the belief that the government also had to intervene at the macroeconomic level to ensure full employment."

In this case market failure apparently signifies a 'failure' by the market to produce the prosperity expected of it, especially in comparison with other more favoured areas. It is this much looser sense that the term is applied to the failure of the economy of Wales and other nations and regions to keep pace with more prosperous areas of the United Kingdom in the late 1990s. Yet it is quite possible to argue that the market is simply doing what markets do, and that policies designed to second-guess markets are misplaced. Such arguments have been put forward, especially by the 'Radical Right'. Maintaining that firms are the best judges of where they should locate, they favour policies that would make markets more perfect, such as locally negotiated wage agreements (instead of state-wide agreements bargained by trade unions), control of local authority taxes, and reduction of unemployment and welfare benefits (Sloman, 1997: 672; MacKay, 2003: 305).
6.5.2 Social Equity

In practice, however, political realities mean that governments in the developed world feel obliged in any event to provide a minimum floor for living standards for people in all regions. Such action has little to do with market failure as described above (6.5.1). Instead, it reflects the alternative rationale for intervention, social equity. This involves, at least in the short run, a redistribution of income, from richer to poorer people, and consequently a spatial redistribution from richer to poorer areas. Given this imperative, it can be argued that it is therefore better to spend money to help people to become productive and self-supporting, rather than spend it on providing them with handouts. The same rationale can be applied to the case for regional policy – that it is better to invest resources in enabling disadvantaged regions to be self-supporting than spend the same money on permanent subsidies. Policies aimed at decentralising government activities, such as that proposed for the civil service (Lyons, 2004) or the Welsh Assembly Government's office relocation programme (Welsh Assembly Government, 2004), implicitly recognise the substantial (and inequitable) advantages capital city areas derive from the presence of government.

A number of the theoretical models discussed above lend weight to this line of reasoning. For example, the cumulative causation models of economic development discussed above (3.5) suggest that successful intervention at a critical stage in the regional economic development path can play a crucial role in setting a region on a positive growth trajectory. In contrast, a negative cumulative process may unfold in regions suffering economic decline. For example, workers who lose jobs may become ‘detached or discouraged’, losing contact with the labour market (MacKay, 2003: 304). More prominently, the popularity of cluster theory among policy makers has led to the proliferation of policies designed to promote the growth of clusters, especially of the fast growth, knowledge-led variety (Martin and Sunley, 2002).

In the same way, the various evolutionary and complexity models reviewed above (5.4 and 5.5) imply a distinct role for public intervention, rather than leaving economic
development to the mercy of free markets. These heterodox models suggest that the
economy as a whole may not necessarily move to an equilibrium position; instead, the
course of development is affected by initial conditions and path dependence (5.5).
Systems with unstable equilibria may not break down, but instead display complex
cyclical behaviour, including movement around a ‘strange attractor’ point (Keen, 2001:
308). If the real economy behaves more like these heterodox models than the predictions
of neoclassical theory, the case for intervention and regional policy becomes more
convincing.

This viewpoint provides the justification for the alternative ‘interventionist’ policy argued
by the political centre and left, which in broad terms favours moving ‘jobs to people’
rather than ‘people to jobs’. This emphasis on the labour market seems fair in the context
of 20th century UK regional policy, the case for which was usually related to disparities in
regional unemployment (Armstrong and Taylor, 2000: 205-208). While such policies
have generally aimed to influence the labour market, the methods employed may involve
enhancing the supply of all factors of production, and also other aspects of regional
policy such as supporting networks.

Whether regional underemployment of labour, or lack of highly paid occupations, is an
example of market failure however, is open to challenge. For example, it may be argued
that the economic inactivity in the South Wales valleys that followed the rundown of the
coal industry cannot be ascribed to any of the types of market failure cited by HM
Treasury (2003: Annex 1) (6.4.1). Rather, it took place because markets did not deliver
alternative private sector employment, while social benefits enabled the local population
to maintain a level of subsistence without moving home to other areas as neoclassical
trade theory would suggest.

So the true rationale for regional policy in this instance derives from social equity, not
market failure; and this ‘equity’ consideration is arguably more frequently used as a
rationale for public intervention, whether or not this is explicitly recognised. In one
example of such recognition, after reviewing various market imperfections, a review of
the former Welsh Development Agency adds:

"Governments also act to correct perceived unfairness in the distribution of
incomes that result from market processes, and to ease adjustments that occur as a
result of market processes" (National Assembly for Wales, 2001: 4.15).

Naturally, theoretical justification is not the same as demonstrating that real world
regional policies actually work. There are a large variety of such policies, and arguably
the task of evaluating their effectiveness is still at an early stage. After briefly examining
the different forms of policy, the thesis examines practical means of evaluation devised
so far and the conclusions on the effectiveness of regional policies.

6.6 Policy Objectives

Once the case for government intervention is established, the precise objectives of such
action needs to be clearly specified (English Partnerships, 2004: 1). Clarity in defining
the benefits expected assists the process of appraisal and evaluation, but is by no means
easy to secure.

An idea of the wide range of policies is provided by Diamond and Spence (1983), who
identify a number of regional policy objectives:

• Increase use of existing physical resources within the region;
• More effective allocation of resources between uses within a region;
• Optimum growth;
• Efficient re-allocation of factors between regions to maximise national income and
growth;
• Equalise growth rates between regions;
• Equalise per capita incomes between regions;
• Reduce demand differences between regions;
• Minimise location specific costs to maximise net social benefits; and
• Preserve regional and cultural identity.
Because political pressure for regional policy has arisen mainly where there are strong spatial disparities in employment, ad hoc policy programmes were devised in an attempt to solve specific problems. As a result, a patchwork of objectives and policies grew up, with diffuse objectives and inconsistent policy tools. Without clear policy objectives, evaluation of the effectiveness of regional policy is extremely difficult (Pickernell, 1999: 9).

6.6.1 Employment

Traditionally regional economic policy has sought to promote greater employment, and has been defined as the promotion of growth and employment in lagging areas (Regional Studies Association, 2001). Swales (1997a: 864) suggests that “regional policy is primarily adopted in order to counter labour market failure over space”.

However, Armstrong and Taylor (2000) indicate that policy objectives and resources have changed considerably over the course of the last seven decades. Thus between 1945 and the 1980s there was a trend for UK regional policy to focus increasingly on economic rather than social goals. Ideology played a clear role, especially after 1945, and by the late 1980s, there was a particular stress on ‘enterprise’. While promotion of employment continues as a priority, the latter half of the 1990s have witnessed ‘mission creep’, with the appearance of multiple objectives – social, political, environmental as well as economic (Armstrong and Taylor, 2000: 220). Yet in practice, generation of employment continues to have primacy over enterprise. For example, only four of 18 sampled projects under the EU Objective 1 Programme in Wales (Priorities 1 and 2) had set targets for new business creation (Old Bell 3, 2006: 3).

6.6.2 Income (GVA or GDP) per head

If regional variations in unemployment are largely evened out, the focus may switch to equalising wages and incomes (Forfás, 2003). One notable change has been the adoption
of Gross Value Added or Gross Domestic Product per head as a policy target alongside job creation and maintenance (H.M. Treasury, 2003). However a number of objections have been made to the pursuit of GVA or GDP as a goal:

- Some projects may be intended to promote ‘capacity building’ assisting communities to raise their own ability to take responsibility for decisions and for actions as well as generate wealth for themselves (e.g. the Coed Darcy regeneration scheme on the site of a former BP oil refinery near Swansea). The ‘pay-off’ for such projects may take many years to accomplish, and may even then not be measurable in purely money terms.

- In a similar way, some projects may promote the general level of research and development (R&D) activity in an area rather than focus on immediate outputs. While the goal may well be to raise local GVA, this may well take place indirectly through creation of ‘externalities’ and take a considerable time to accomplish.

- Pursuit of GDP may also ignore side effects such as damage to the environment, whether within or beyond the target area or population group it is intended to assist. This has led to attempts to provide for sustainability within the framework of neoclassical growth theory (10.3.1) (Toman et al., 1995: 140) and the proposed use of a number of measures designed to assess the sustainability of economic policies (Midmore, 2001).

Layard (2005) even argues that happiness, rather than economic growth, should be made the objective of economic policy; suggesting that once basic needs for food, clothing and shelter are more or less universally met, higher gross domestic product does not seem to make societies happier.
6.6.3 Export-oriented policy

The survey of theory pointed to the potential of focusing on an export-based sector to drive the economy as a whole (3.4). Export-oriented firms could be encouraged by such measures as tax relief, improved transport facilities, and high quality telecommunications. The danger is that such a strategy relies on external demand, and may lead to a skewed economy, vulnerable to vagaries in the global market. There is indeed considerable international evidence of countries with a sharp divide between the export sector and the ‘base sector’, which depends exclusively on the home economy. For example, the growth of the oil industry in Nigeria lead to a five-fold appreciation in the exchange rate and a sharp fall in the relative profitability of domestically produced goods (World Bank, 1996: 14).

6.7 Regional policy instruments

El-Agraa (1997: 1,504) defines industrial policy as “any state measure designed primarily to affect the allocation of resources between economic activities”, and regional policies fall within this broad field (Wren, 2001: 850). Armstrong and Taylor (2000: 222-223) suggest that regional policy instruments could pursue the following goals:

- Influence the locational decisions of households (micro-policy instruments);
- Change aggregate regional income and expenditure by measures such as fiscal, monetary and exchange-rate changes which may have regionally discriminating effects (macro-policy instruments); and
- Co-ordinate the impacts of micro- and macro-policy instruments and those of different layers of government.
### Table 6.7.1 Regional Micro-Policy Options

<table>
<thead>
<tr>
<th>Reallocate Labour</th>
<th>Reallocate labour in situ (e.g. occupational training, education intermediate labour market schemes)</th>
<th>Spatial reallocation of labour</th>
<th>Reallocate Capital</th>
<th>Taxes and Subsidies</th>
<th>Inputs</th>
<th>Outputs (e.g. export rebate and price subsidies)</th>
<th>Technology (e.g. subsidies for R&amp;D, subsidies for dissemination of technological information)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Migration policies (e.g. subsidies to cover financial and psychic costs of migration)</td>
<td>Mobility policies (e.g. information for potential migrants, housing help for migrants, policies to ease house sale and purchase)</td>
<td>Improve efficiency of labour markets (e.g. encourage local collective bargaining)</td>
<td>Taxes or subsidies on capital, land and buildings (e.g. building grants, capital grants, interest rate relief, tax allowances on investment, local tax and rent relief, taxes on firms located in prosperous areas)</td>
<td>Subsidies or taxes on labour (e.g. wage subsidies, subsidies for key worker migration)</td>
<td>Subsidies or taxes on other inputs (e.g. freight or energy subsidies)</td>
<td>Policies to improve efficiency of capital markets (e.g. loan agreements, provision of venture capital, exchange guarantees, credit unions, local exchange and trade schemes [LETS])</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Measures to support Higher and Further Education (e.g. University ‘Third Mission’ activities)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Support for networking activities (e.g. trade networks and associations)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on Armstrong and Taylor (2000: 233-234)
Table 6.7.2 Regional Macro-Policy Options

<table>
<thead>
<tr>
<th>Devolve trade, fiscal and monetary policy powers to regions</th>
<th>Regionally discriminating tax and expenditure policies</th>
<th>Regionally discriminating monetary policies (e.g. easier credit facilities for assisted areas)</th>
<th>Regionally discriminating tariff or trade controls (e.g. import quotas on goods produced in assisted areas)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retain central control of trade, fiscal and monetary policies</td>
<td>Automatic stabilisers (e.g. income tax, social security)</td>
<td>Discretionary actions (e.g. government contract preference for tenders from assisted areas; targeted spending programmes for health, roads, schools and colleges etc.)</td>
<td></td>
</tr>
</tbody>
</table>

Based on Armstrong and Taylor (2000: 233-234)
The breadth of this policy array underlines the reality that ‘national’ policies may well have spatial impacts, perhaps not intended. An undifferentiated monetary policy may for example produce different results in different areas, because of regional differences, as indicated by de Lucio and Izquierdo (2002) in their analysis of Spanish regions.

Macroeconomic adjustments required because of inflation depend primarily on the non-accelerating inflation rate of unemployment (NAIRU) in regions in which full employment prevails (ODPM, 2006: 17). This means that expansionary policy may ‘over-heat’ fully employed parts of the economy, resulting in departures from target inflation and a reversal of macroeconomic policy – a ‘national policy failure’ that “makes targeted (spatial) policy intervention necessary to correct for existing market failures or other imbalances in the inter-urban system” (ODPM, 2006: 20).

In practice, regional policy has been dominated by micro-policy instruments designed to bring about a reallocation of capital and labour to areas that would not otherwise have been chosen, whether that is achieved by movement of factors from outside the region or by promoting the development of indigenous resources (Armstrong and Taylor, 2000: 234). This thesis adapts the taxonomy they propose to offer a classification of regional policy instruments (Tables 6.7.1 and 6.7.2). This builds on from the taxonomies offered above in specifying measures to promote knowledge and networking, increasingly regarded as playing an important role in the process of clustering and agglomeration.

Taylor and Wren (1997: 840) also suggest measures to control development in overheated areas – such as location controls or congestion taxes – and (curiously in this context) measures to reduce job search and relocation costs of migrant workers from less well-off areas, which in view of the process of cumulative causation discussed above is likely to worsen rather than improve regional disparities.

Policies thus may be either of the ‘carrot’ or ‘stick’ variety, the ‘stick’ actively discouraging development in non-assisted areas. In the UK, use of the ‘stick’ has declined markedly since the 1970s, as demonstrated notably by abolition of the UK’s Industrial and Office Development Certificate controls in 1981 (Armstrong and Taylor,
Establishment of regional development agencies throughout England, including more prosperous areas, by the Blair government might be regarded as final abandonment of judicious use of the regional stick. This is the view of Wren (2001: 856), who regards the setting up of the English RDAs, including that for London, as agents that formulate and implement national policy at a regional level, adding that the main regional instrument, Regional Selective Assistance (RSA) no longer funds the interregional transfer of firms where there is no net increase in UK activity. There may be room for scepticism on this view of the use of RSA, as RDAs are permitted to promote expansions involving firms moving from one part of the UK to another. This is a topic of some controversy, as in the case of the engineering company Fenner, which announced plans to open a factory in the Rhondda in 1996, closing two plants in England (UK Business Park, 2006). However, it is difficult to resist the conclusion that:

“What is termed regional policy therefore has little or nothing to do with redistribution, and increasingly it is the implementation at the regional level at national level of national competitiveness policy measures.” (Wren, 2001: 856)

6.7.1 Policy Instrument Mechanisms: Factor Subsidy

Regional economic policies have generally sought to counter perceived market failure or promote social equity by subsidising one or more factors of production in order to improve the region’s competitiveness. Labour, capital and land have all been subsidized in this way. UK regional policy during the 20th century was dominated by capital subsidies, except for the period between 1967 and 1977 when labour subsidies (in the form of Regional Employment Premium and Selective Employment Premium) played a significant role (Wren, 1996). It should be noted that the aim of these capital subsidies was to help the employment of underused labour in the assisted areas rather than encourage use of capital for its own sake. All these options have been pursued as part of UK regional and employment policy, with various governments subsidizing both capital and labour to varying degrees at different times (Armstrong and Taylor, 2000: 242-243).
Sloman (1997: 673) suggests a range of options open to such intervention, involving both ‘carrot’ and ‘stick’:

- Subsidies and tax concessions in depressed areas;
- Provision of facilities in depressed areas e.g. infrastructure, sites and premises, or decentralisation of Government departments; and
- Direct restriction of expansion in prosperous areas, by differential taxes or by regulation (this form of policy clearly using the stick rather than the carrot, and raising revenue rather than costing the Exchequer money).

A more detailed although “simple crude” taxonomy is proposed by Richardson (1976: 243), who notes that there are several ways of classifying the variety of regional policy instruments available:

(i) subsidies to firms;
(ii) government expenditure policies including
    - short-run discriminatory expenditures (transfer payments, stabilisation expenditures, fiscal equalisation policies)
    - long-run infrastructure expenditures
    - regional discrimination as an element in the spending decisions of government agencies and nationalised industries;
(iii) direct controls;
(iv) measures to promote factor mobility; and
(v) implicit spatial policies, i.e. the regional implications of State policies (monetary and fiscal, agricultural, sectoral).

These divisions include market-based measures (such as promotion of factor mobility) and measures that fall under the market failure approach. Behind these broad headings lie further subsets of policies. For example, the first category (subsidies to firms) may be further subdivided in a variety of ways; such as promoting initial investment or early stage growth (perhaps by a selective policy) or by subsidising capital or labour within the
region (Richardson, 1976: 243-244). It should be noted that this classification, suggested during the 1970s, does not include policies to promote trust, networks, clusters and institutional thickness, all high on the agendas of policy practitioners in the 21st century.

One detailed taxonomy of regional policies within the framework of the ‘market-failure approach’ is suggested by Taylor and Wren (1997: 840), who suggest a number of policies that can be deployed to increase the productive potential of less well-off areas:

- Encourage indigenous development, through new firm formation and the growth of existing firms (by provision of business support, industrial sites and premises, financial support and loan guarantees);
- Encourage inward investment from other regions or from abroad (through provision of investment and other incentives);
- Improve physical infrastructure and environment to raise competitiveness and make areas more attractive to potential investors;
- Improve the skill level of the local workforce by investing in human capital programmes, including both education and training;
- Encourage unemployed people back into the workforce (by offering subsidies targeted at these individuals).

Detailed discussion of the range of policies available and of the most appropriate form of classification is beyond the scope of this thesis. To begin with, many policies may assist regions economically – for example measures to develop higher and further education, transport infrastructure, agricultural production and location of research and defence establishments. Many of these are not considered as falling within the ambit of regional policy but may nevertheless exert a powerful influence on how regional economies develop. For example, the low level of publicly funded science research in Wales (Davies, 2004: 301-302) has been cited as arguably entailing a negative regional policy. The remarkable economic performance of the Republic of Ireland in the late 20th century is widely ascribed to its policy decision to invest in human resources, as well as a tighter fiscal regime, shrewd use of instruments such as corporation tax and its influence within
the European Union, attributable to its political independence (Bradley and Hamilton, 1999). The EU’s Common Agricultural Policy has also frequently been described as benefiting relatively wealthy producers of grain in areas such as eastern England (Jenkins, 2005) and penalizing producers in northern and western parts of the UK, thus in effect operating as a negative regional policy.

A number of these options can be aimed at adjusting the industrial mix of the less well off areas to make them less vulnerable to economic change. Such strategies may not be easy: Cooke and Morgan (2000: 78) maintain that

"There are not large numbers of cases of trajectory switching to a totally new path-dependence in the recent history of regional development".

6.7.2 Selection of Factors Subsidised

Whatever its goals, traditional regional policy usually took the form of subsidising certain factors of production in designated areas. This approach raises the question of which factors of production should be subsidised, especially in the light of recent theories that seek to explain the importance of knowledge and human capital in generating economic growth, such as the model developed by Castro and Jensen-Butler (1999) discussed above (3.3.4). This earlier theoretical treatment showed that such models emphasized the advantage enjoyed by knowledge-rich regions with effective institutional networks for generating and exchanging ideas. It noted the possible role played by inward investment as a short cut to acquiring embodied knowledge as well as capital, but also stressed the innovative networks of the ‘intelligent region’ (Cooke and Morgan, 1998: 149) (3.3.5).

Such treatment provides a starting point for assessing the contribution regional economic policy can make to building a knowledge-rich economy. If credible as an explanation of the real world, then as well as subsidising capital and labour, public support services may be used to enhance the region’s endogenous capacity to generate knowledge and its wealth of human capital; and equally techniques of evaluation should seek to monitor and measure the impact of their contribution to those factors.
The model however leaves unspecified one element that has gained growing attention from policy-makers – entrepreneurship. Is entrepreneurship to be treated as an undifferentiated component of human capital, or should it be regarded as a variable in its own right? This is a question of some import, if only because of the attention (and the cash) devoted to a variety of programmes intended to raise levels of entrepreneurship: the Welsh Development Agency set aside a budget of £15.4 million during 2002-2003 for encouraging participation and enterprise (WDA 2002a: 2.5).

Academics differ in their views of how exactly entrepreneurs should be defined. At one end of the spectrum are the ‘Silicon Valley’ or ‘Route 128 (Boston)’ group, who see entrepreneurship as something related to founding and growing high technology businesses – creating new wealth by implementing new concepts. At the other end of the spectrum are those who see any novel activity that creates organizational change and economic value as entrepreneurship. However researchers agree that ‘entrepreneurship’ is different from ‘small business’: most small businesses have limited growth potential, and are primarily focused on getting a fair return for the small number of people they employ (Birley, 1997). It is more difficult to explain the wide variation in ‘entrepreneurial culture’ found by many observers (Brooksbank and Jones-Evans, 2000). Can entrepreneurship be regarded as a factor of production, and measures taken to increase its availability? Certainly this is the approach followed by development agencies in both Scotland and Wales (Scottish Enterprise, 2002; Entrepreneurship Action Plan for Wales Steering Group, 2000).

6.7.3 Subsidy of Networks and Clusters

So far this discussion has viewed economic policy as assisting less advantaged regions by subsidising individual factors of production. This approach, as shown above (3.3.5), can accommodate the role of human capital, embodied and disembodied knowledge, and entrepreneurship by treating them as factors of production. However it appears unsuited to the task of developing the networks that many writers believe play an increasingly
important role in regional prosperity. Evidence of the importance of clusters and networking indicates that factors of production, however elaborated, cannot be treated in isolation. This is argued to be especially significant because of globalisation of trade and economic activity (Castells, 1996; Asheim and Dunford, 1997).

As a result of the perceived importance of developing networks and clusters, policy-making has increasingly focused on supporting the growth of such clusters rather than competitiveness of individual firms (Sotarauta and Hukkinen, 2002: 4). The model developed by Porter (1990) suggested that the critical competitive factors are based not only on such factors as labour costs, resource availability and the general macro-economic environment, but also on qualitative factors of the local environment that are intensified through clustering. While individual firms may come and go, clusters as a whole may continue to create employment and generate wealth. Global competitiveness appears paradoxically to depend to a greater extent on local economic, political and social institutions, processes and networks, and it is argued that regional agglomeration is growing in importance as a mode of economic co-ordination (Cooke and Morgan 1998; Sotarauta and Hukkinen, 2002: 5). The principal reason for this argument is that regional agglomeration provides the best context for an innovation-based economy (Asheim and Isaksen, 1997); and this has led to an increasing interest in national and regional innovation systems (Lundvall, 1992; Braczyk et al., 1998). This interest has been reflected in Wales by the launch in 1994 of the Wales Regional Technology Plan, jointly funded by the WDA and the European Union (Cardiff University, 1994), succeeded by the Welsh Assembly Government’s Wales for Innovation action plan (Welsh Assembly Government, 2002).

Sotarauta and Hukkinen (2002: 1-2) argue that globalisation tendencies are one of the most important sources of overall change, and that the basic puzzle for policy makers is how to create locations for opportunities in the “borderless world”. In the global economy, borders are fuzzier than ever before, so that regional success is determined by competencies and skills at learning and developing in a continuous process: “Activities and processes are increasingly organised in networks and participation in the networks
and network dynamics are critical sources of power.” A developed information technology enables an ever-accelerating interconnection of flows, diminishing the time-space ratio. Castells (1989; 1996) maintains that flows of capital, information and technology are more relevant now than a space of places: such flows pass through certain nodes and hubs, but if a city-region becomes less attractive or some other area becomes more attractive, the routes of flows may change quite rapidly. Therefore regions need to choose which flows they want to attract and what type of development strategies they need to implement (Kostiainen, 1999). In the network society information and innovations move faster than ever from one place to another (Thrift 1995).

However, an overwhelming majority of people live in places and so they perceive their space as place-based. These places do not vanish: networks are linked up to specific places (Castells, 1996: 413). In this world of rapid flows, to be competitive is to be attractive: international money, nomadic by nature, chases opportunities in the world economy that might produce profit; increasing competition for capital as well as expertise, information and technology (Sotarauta and Hukkinen, 2002: 3). Global competition is based on comparative and competitive advantage, and therefore on specialisation, with place-specific advantages created on the basis of innovation, skilled labour and different institutional and cultural environments. Florida (2002) attributes the economic success of cities to their willingness to accept creative classes, a conclusion disputed by Markusen (2006), who argues that occupations with distinctive spatial and political proclivities are bunched together purely on the basis of educational attainment, and with little demonstrable relationship to creativity. Frontier Economics (2004) identify a number of place-based factors influencing regional disparities such as peer group, aspirational, neighbourhood and intergenerational effects. There is thus a sound policy rationale for improving the quality of place by community investment in urban areas, well designed green spaces, river frontages and local housing estate renewal that can catalyse large scale urban regeneration (OPDM, 2006).

These factors offer subtle distinctions in the operating environment for various actors (Maskell and Malmberg 1999), a feature described by the concept of ‘competitiveness’.
Sotarauta and Linnamaa (1997: 62) go so far as to define regional competitiveness largely in terms of connection to networks and quality of life, viewing competitiveness as the region's ability to:

- Connect the urban area and its actors to the best possible networks as tightly as possible;
- Maintain and develop the quality of life of local residents;
- Attract new, competitive companies to the area; and
- Assist existing companies to maintain and develop their competitiveness.

During the 1990s Porter’s influence led to a number of supranational organisations such as the OECD and the EU to promote policies based on agglomeration or clustering. In Europe regional policies in several regions are geared to clustering (Konstadakopulos, 1998: 11). Cluster policies at regional level have been adopted by some of the German Länder (i.e. Nordrhein-Westphalen and Baden-Württemberg), some the autonomous communities in Spain (i.e. Catalonia and the Basque Country), by Northern Ireland in the UK and by many states in the United States (Lagendijk, 1998: 319). Lagendijk shows that cluster policies and initiatives in the UK emerged out of sectoral policies that were evolved from the ‘industrial district’ model and have developed in the context of a more comprehensive policy in Scotland, Wales and Northern Ireland rather than in England (Lagendijk, 1998: 328-330).

The complexity of cluster structure discussed above (4.6) has clear implications for regional policy. In the UK, cluster policy has been taken forward by the Department of Trade and Industry, which used location quotient analysis to identify regional conurbations of industry and suggest potential cluster policy actions (DTI, 2001); although cluster policy was developed in Scotland several years before it was enthusiastically embraced by the DTI (Adams and Robinson, 2005: 5).

Based on an analysis of clusters in the US, Porter (2003: 571) highlights the need for regional economic development policies to be particularly attuned to traded clusters,
because these not only support higher wages but also appear to drive local employment and especially local wages. He finds some evidence that growing clusters in the US are becoming more concentrated, while those in decline become more dispersed (Porter, 2003: 568). If this is a general feature of the growth on advancing industries, it would underline the significance of the cluster concept for regional policy.

The nature of the interrelationship between different industries found by Porter (2003: 562) in the United States suggests that policies intended to promote a specific industry may not prove successful unless the region assisted also possesses or secures linked industries and related skill bases. Rather than focusing solely on developing ‘high-tech’ clusters (designated as aerospace engines, aerospace vehicles, analytical instruments, biopharmaceuticals, computer equipment, IT and medical devices), regions therefore need to upgrade all the clusters that are present:

"Regions should focus on upgrading the productivity of all clusters in which they have a meaningful position, rather than attempting to migrate to more ‘desirable’ clusters." (Porter, 2003: 571)

This approach is reflected in attempts to extend cluster policy beyond the limits of a ‘natural cluster’. Thus the regional development agency Advantage West Midlands has grouped together seven distinct mature manufacturing sectors in a High Value Added Consumer Products Cluster: the reason for this artificial clustering is that the different activities are all design driven, market led and customer facing (Berkeley et al., 2005: 1). Such policy efforts do not readily fit the traditional cluster concepts of horizontal and vertical integration (Berkeley et al., 2005: 9).
6.7.4 Critiques of cluster policy

A number of authors question both the cluster concept and its widespread adoption as a policy for regional development (Adams et al., 2003; Berkeley et al., 2005: 7). Martin and Sunley (2002) suggest that "the rush to employ 'cluster ideas' has run ahead of many fundamental conceptual, theoretical and empirical questions". Palazuelos (2005) concludes that clusters are far from being a panacea, and that, apart from the risks of excessive congestion, "few local economies possess the appropriate characteristics for cluster-creation to be successful as a local economic development strategy".

Interest in the networking associated with clusters has led policy makers to take seriously the significance of institutions and established practices (Maskell 1996; Cooke and Morgan, 1998; Dosi and Orsenigo, 1988: 19). One effect of this growth of interest has been to emphasise the positive role in promoting development that can be played by regional levels of government, a feature evident during the course of the 1997 referendum on devolution for Wales (Richard Commission, 2004). However, even in promoting disembodied knowledge, there is no consensus that the regional milieu is more important than the State environment (Asheim and Dunford, 1997; Amin and Tomaney, 1995). Sotarauta and Hukkinen (2002: 8) agree with this critique, especially from the viewpoint of small countries like the Nordic nations; and suggest that the question is how development processes can be global, national and local at the same time.

6.7.5 Regional competitiveness

There has been considerable discussion of the concept of regional competitiveness, sometimes cited as a policy goal. Competitiveness is defined at the national level as "the degree to which [a country] can, under free and fair market conditions, produce goods and services which meet the test of international markets, while simultaneously maintaining and expanding the real incomes of its people over the longer run" (OECD, 1992: 237). This much-quoted definition has been adopted by the UK government (Wren, 2001: 58n). However, competitiveness is neither well defined nor easily
measured according to Krugman (1997), who describes it as a kind of ‘ineffable essence’ that can neither be defined nor measured.

Competitiveness is not a formal economic concept, and has developed as part of the policy-making process in an ill-defined way (Wren, 2001: 848). Fagerberg (1996) suggests that competitiveness is of its nature a relative concept, involving comparison of performance; that it can be applied at different levels, including the firm, industry or national economy; and that at the country level it can relate both to the well-being of citizens and to trade performance.

The nature of competitiveness between firms is well understood. The terms on which firms compete across national boundaries are determined both by comparative advantage and the level of industrial ‘competitiveness’. The government defines the latter as “the ability to produce the right goods and services of the right quality, at the right price, at the right time” (DTI, 1994: 9). However, it is not necessarily the case that this idea of competitiveness at firm or industry level can be interpreted in the same way for the nation or region (Dunford et al., 2001). Thus Krugman (1994b) contests the notion of regional or state competitiveness, since while firms may rightly be seen as competing in a ‘win-lose’ fashion, the same metaphor does not carry over to nations which can potentially all gain from trade:

“Coke and Pepsi are almost purely rivals: only a negligible fraction of Coca-Cola’s sales go to Pepsi workers, only a negligible fraction of the goods Coca-Cola workers buy are Pepsi products. So if Pepsi is successful, it tends to be at Coke’s expense. But the major industrial countries, while they sell products that compete with each other, are also each other’s main export markets and each other’s main suppliers of useful imports. If the European economy does well, it need not be at U.S. expense; indeed, if anything a successful European economy is likely to help the U.S. economy by providing it with larger markets and selling it goods of superior quality at lower prices.” (Krugman, 1994)
International trade, then, is not a zero-sum game, and in Krugman (1994)'s view, the idea of state competitiveness can be a 'dangerous obsession' because it may serve to promote protectionism. In the short run, the factors which affect competitiveness are those leading to adjustments in the real exchange rate, such as a reduction in nominal wages but "in the long run the real exchange rate is in equilibrium, and competitiveness is virtually synonymous with trend productivity growth" (Wren, 2001: 848).

Similarly, Turner (2001) contests the assumptions that because companies are locked in ever more global competition, so too must countries: that assumption is almost entirely wrong, and "the fact that it is wrong has profound consequences for our ability to think clearly about relative economic performance and about policy choices in economic and social life". Martin and Tyler (2003) agree that nations and regions do not compete in the same way as firms; but regions do compete for investment, for labour (especially skilled and creative labour) and entrepreneurs, and for technology.

6.8 Conclusions

This chapter has related regional economic policy to its theoretical foundation in the concept of social welfare, preparing the way for studying appraisal and evaluation methodologies.

Traditionally the case for economic policy is advanced on the grounds that it improves social welfare whether by improving efficiency or equity. Advocacy of policy implies that market systems are failing to produce an optimum social welfare position; and in the context of regional policy are unable to lead with sufficient speed to spatial convergence of economic activity. Policy is usually argued either (1) to counter market failure or (2) to promote greater social equity; and the case for intervention is stronger where free markets fail to secure optimum spatial distribution of social welfare.

However, a wide range of regional economic policy objectives may be distinguished, with an even greater number of associated policy instruments (6.7.1). Policies may be
developed to pursue goals linked to various theories of regional economic growth, and operated by subsidising various factors of production or the links that promote networking and clustering.

One problem facing policy-makers is that the microeconomic basis of cluster 'theory' is unclear, so that it is difficult to identify which policy instruments will be most effective. This uncertainty makes sound evaluative procedures still more necessary. The following chapter considers the implications of theory and policy for devising a sound and robust system of appraisal and evaluation.
Chapter 7 Regional Economic Theory and Policy Synthesis

7.1 Introduction

This chapter seeks to draw out a number of arguments presented in the earlier treatment of regional economic development. The aim is to present any conclusions that may be relevant for policy makers and practitioners.

The key question is whether the growth or decline of regions should be left to the operation of free markets. Does either theory or empirical evidence present a clear case for or against specifically regional economic policies? If government intervention passes this test, does theory or empirical evidence suggest any specific way in which that intervention should be managed? And are there any particular concepts from theory that are of use to policy makers and policy practitioners? The chapter concludes with a brief review of the impact of regional economic policy in the United Kingdom.

7.2 Lessons of Theory

What lessons, if any, can be drawn from the survey of regional economic theory (Chapters 3, 4 and 5)? Regional economics reflects the same divisions that appear within economics as a whole. Orthodox economists tend to argue the neoclassical school’s case for reliance on markets to even out widely differing living-standards between regions and nations, attributing persistent differences to various rigidities that prevent adjustments through trade and investment. Remove these rigidities, the argument goes, and regions will move irresistibly towards general equilibrium, which in turn implies regional
convergence – with less developed regions catching up with their more prosperous counterparts (3.3.4). Intervention by government to promote regional development is unnecessary.

Even within the realm of orthodoxy, however, and before admitting empirical evidence, breaches appear in the wall erected against policy. Once the dimension of space is admitted to the analysis, the presumption of perfect competition that drives the doctrine of convergence breaks down. Distance implies imperfect competition (4.4.2, 4.4.5). Firms located in certain areas may then enjoy indefinite advantages over firms in other areas. Movement towards equilibrium would continue but need not imply equilibrium of prices or factor incomes over space.

Such adjustment problems increase as strict neoclassical assumptions are progressively relaxed, for example by admission of modifications in neoclassical growth theory and by Keynesian export-based models, themselves scarcely meriting proscription as heterodox. Models of endogenous growth show how regions that gain a head start, especially by developing export sectors, may secure self-sustaining growth. Evidence of such ‘cumulative causation’ (3.5) provides a strong case for interventionist policy, even within the confines of spaceless theory.

Combined with the insight of spatial theories, and their demolition of perfect competition as a driver of spatial equilibrium, the rationale for public policy appears more reasonable. A further widening of the breach in the neoclassical defences against intervention is made by the arguments of new economic geography and clustering. New economic geography extends and elaborates the nature of imperfect competition implied by spatial theories. Agglomeration and cluster ‘theory’ arguments show regions can develop a feature highly disturbing for orthodox theory – increasing returns (5.5). Both new economic geography and clustering are consistent with the idea of cumulative causation, and together constitute a powerful argument for persistence of regional disequilibrium, and hence for a countervailing regional economic policy.
Arguments in favour of public policy therefore hardly need the support of thoroughly heterodox theories. Evolutionary theories, however, provide further ammunition, especially by stressing the possibility of path dependence (5.5) and the likelihood of multiple equilibria and increasing returns. Finally, while hardly as yet respectable, heterodox theories that dispute the nature of overall equilibrium (5.3) throw another spanner into the works by suggesting there is no inherently equilibrating process in the economy as a whole.

The conclusion therefore is that the science of regional economics leaves the door wide open for regional economic policy. Certainly there is no clear disproof of the need for policy. Some near orthodox theories, such as export-based models, clustering, disembodied knowledge and cumulative causation imply plausible roles for government intervention. This is very different, of course, from knowing that policy works, which is what the process of evaluation is all about.

7.3 UK Regional Economic Policy: An Assessment

As discussed in the theoretical review (3.3.1), one important conclusion of orthodox economics was that region’s economic activity and prosperity would tend to converge over time. Most neoclassical economists believed that the capitalist system had strong self-correcting and self-regulating tendencies (Cornwall, 1993: 86). Such a belief has major implications for government policy: if it were the case that, left to themselves, market forces would even out prosperity, there would be little argument for spending scarce resources on policies aimed at adjusting the spatial balance of economic activity.

From the end of the Second World War to the mid-1970s, there was indeed a trend towards greater regional equality in the distribution of income and wealth across the countries of Britain. This ‘Golden Age’, as the period between 1950 and 1973 is styled by MacKay (2003: 304), was accompanied by growth and relatively full employment, in which overall as well as regional inequalities of prosperity and standards of living were ameliorated.
For most Western economies, this era of exceptional stability came to an abrupt end with the sharp depression of the early 1970s, triggered by the Yom Kippur War and the oil price hike of 1973-74 (Bank for International Settlements, 1991). So too did the trade-off assumed to exist between employment and price inflation proposed by the ‘Neoclassical-Keynesian Synthesis’ school of economists and traced by the ‘Phillips curve’ (Phillips, 1958). From the mid-1970s, high inflation and high unemployment marched hand in hand, a phenomenon dubbed ‘stagflation’ (DeLong, 2006).

Not surprisingly, therefore, the end of the post-War period witnessed an attack on its prevailing economic ideology, Keynesianism, the advocacy of government monetary and fiscal programmes designed to increase employment and stimulate business activity (Phelps, 1967). In the UK and the USA at least, a neo-liberal ‘Counter Revolution’ (MacKay, 2003: 305) identified government as the problem, and saw the welfare state as an obstacle to progress. Welfare benefits were presumed to add to ‘natural unemployment’ by reducing the real returns to work. The Counter Revolution’s prescription was that government economic policy should retreat from welfare and direct intervention and concentrate on managing the supply of money to ensure a sound currency (Friedman, 1968). This was reinforced during the 1980s by a move in the UK Government’s approach away from directly interventionist policies in favour of providing an overall sound business environment for the private sector (Wren, 2001:849).

That the call to cut spending on welfare failed to materialise in the UK is not in dispute: the benefits system became considerably more expensive during the Thatcher years. The UK’s social security spending increased in relative terms from 8.2 percent of GDP in 1973-1974 to 11.4 percent in 1995-1996 – more than double in real terms (Glennerster and Hills, 1998).

What is open to debate is the reason for failure. Or, put another way, could it have succeeded given different circumstances? This brings economists face to face with the ‘What If?’ question, the counterfactual, whose nature is discussed in detail in section 2.8.
Friedman (1953: 11) lamented "the denial to economics of the dramatic and direct evidence of the 'crucial' experiment". Some economists of a similar philosophy felt that the monetarist Counter Revolution was not pursued with sufficient conviction or effectiveness (Pepper and Oliver, 2001). Others believed that the experiment had been too 'crucial' by far. For example, greater social security spending was in part a result of increases in levels of poverty and unemployment, themselves arguably caused by the monetarist experiment (Glennerster and Hills, 1998).

The case argued by adherents of the Counter Revolution rests on a belief of how the economy works, in which a central role is played by the concept of market-clearing and its idealised target, general equilibrium, whose nature was examined in detail in the theoretical review (3.3.4):

"The equilibrium assumption, which lay at the heart of the Counter Revolution belief, emphasised that all productive resources released in the process of structural change should find alternative use." (MacKay, 2003: 308)

Thus Friedman (1977: 47), calling for a cut by one-sixth in government spending, argued that this would render people available for productive rather than unproductive employment, stressing (a considerable act of faith in real world market-clearing): "Most of these people would be absorbed in a fairly brief period". This assumption, always heroic, became increasingly difficult to reconcile with reality when unemployment (open and concealed) climbed and persisted (MacKay, 2003: 308).

This highly political debate is relevant for this thesis in that regional policy in the UK formed part of overall government intervention. Regional policy was in part driven by wider policy issues. For example, determined that that Britain should join the Common Market in January 1973 with business booming, British Prime Minister Edward Heath persuaded his Chancellor, Tony Barber, to introduce regional investment grants and to provide favourable terms for investment in plant and machinery (Daily Telegraph, 2005: 25).

131
MacKay (2003: 305) distinguishes “two periods of active or strong regional policy, 1945–51 and 1960–75, and two periods of passive or weak regional policy, 1952–59 and the years after 1975.” Furthermore, he maintains that the clear and considerable contrast between these four periods, two strong and two weak, provides ‘a good guide as to the impacts of economic regional policy’; to use Friedman’s phraseology, a ‘crucial’ experiment.

The evidence showed that annual plant movement into the regional policy UK, Development Areas was approximately three times higher in the strong policy years MacKay (2003: 305). This would tend to confirm the conclusion of Moore and Rhodes (1973: 99) that “the results . . . which are remarkably consistent indicate a regional policy effect which is encouragingly large.” Other research has agreed that there is a consistent, substantial employment effect which built over time (Ashcroft and Taylor, 1977; MacKay and Thomson, 1979; MacKay, 1982; and Moore et al., 1986).

Thus there has been no lack of scholars ready to attribute strong positive effects for regional policy in the UK. What is especially significant about the approach of MacKay (2003), however, is the way in which a counterfactual comparison is provided by examining strong and weak policy periods within the same period of growth and relatively full employment. Given that two strong and one weak period of regional policy all lie in the post-War ‘Golden Age’, there may be grounds for claiming that the problem of establishing the counterfactual has been met.

In attempting to judge the impact of regional economic policy, however, it is necessary to take a broad view. It is not sufficient to measure what direct effects may be attributed to policy, and compare these benefits with the cost. Wider impacts on the economy must also be considered; these may be beneficial (for example, through the multiplier effect, or promoting externalities) or negative, as in the case of displacement of existing activity. Even this consideration may not be sufficient: it is also necessary to enquire what would be the consequence of not proceeding with the policy – the counterfactual. This includes
considering whether the results would have taken place without the policy, in which case money spent would be regarded as ‘deadweight’ expenditure.

Wider considerations demand that policy makers consider the benefits and expense of regional economic policy alongside other policies. For example, if an effective regional policy is abandoned, how much will society have to spend on welfare benefits for those made unemployed as a result?

On the other hand, regional economic policy, as generally understood, is not the only method of promoting prosperity. To begin with, scarce resources could be spent on other economic policies not spatially targeted, such as encouragement of innovation or enterprise. More widely, government could choose to invest more in human resources – in staff and equipment in schools, colleges and universities, or in training at work, in the belief that these may have a greater effect than regional policy per se. Just such a wide policy framework was enumerated by the UK Government in its summary of the ten main influences on competitiveness (DTI, 1994; Wren, 2001: 848-849).

The point is that real-world evaluation has to take account of the wide range of real-world variables and counterfactual situations. Those undertaking the task of appraisal and evaluation therefore need to approach their role with some humility. It is unlikely to be within their power to determine which specific policies would make a real difference for any economy in any situation. What they can hope to do, however, is arrive at a judgement that holds good at least within certain limits – for example, between projects of a roughly similar character or within similar economic environments. This thesis shows how one particular model of appraisal and evaluation can help achieve that limited ambition.
7.4 Conclusions

The case for regional economic policy rests on its claimed ability to raise social welfare, itself a complex calculation involving questions of distributive as well as aggregate measures of prosperity. The preceding analysis concludes that free markets may not always deliver a spatially optimum pattern of welfare, implying a potential case for advocating government intervention.

However policy implies cost as well as benefits. Evaluation of regional economic policy as a whole, as well as individual programmes and projects, must therefore weigh up and compare these cost and benefits. Part 4 of this thesis examines how this process may be undertaken.
8.1 Introduction

Part 4 sets out the methodology required to develop a model for use in practical appraisal and evaluation of regional economic development programmes and projects. The term 'methodology' is defined by Blaug (1992: xxv) as "the investigation of the concepts, theories, and basic principles of reasoning of a subject".

This chapter examines the overall methodological framework for appraisal and evaluation, relating it to the previous survey of theory and policy in Parts 2 and 3. Section 8.2 reviews the overall context for evaluation while 8.3 discusses its relation with economic theory. Section 8.4 discusses the debate about evaluation and learning and the significance of the institutional context. Section 8.5 presents a brief conclusion.

8.2 Review of appraisal and evaluation methodologies

8.2.1 Nature of Evaluation

The term 'evaluation' is used to refer to the forecasting of results of actions in advance (ex ante), to monitoring their progress (ex itinere) against agreed benchmarks and to assessing their consequences after the event (ex post). Much academic literature interprets 'evaluation' in this broad sense (Armstrong and Taylor, 2000: 364). However, the practice among UK practitioners and government officials is to
use the term ‘appraisal’ to refer to ex ante estimation, that is a forward-looking assessment of the likely future effects of new policies of proposals.

As discussed above (1.6), appraisal concerns the estimation and presentation of the costs and benefits of each potentially worthwhile way of fulfilling defined objectives, and taking account of associated risks (English Partnerships, 2004: 1). By contrast, ‘evaluation’ tends to be restricted in the official lexicon to the ex post “retrospective analysis of a policy, programme or project at its completion, conclusion or revision” (HM Treasury, 2003: 1). Ex ante analysis provides answers to ‘what if?’ questions, whereas ex post evaluation looks backwards at the effects of policies already introduced (European Environment Agency, 2004). Ex post modelling identifies the relative contributions of different causal variables (Jensen-Butler et al., 2002: 329).

Recent academic discussion emphasises the need to build evaluation into all these stages, and to use it as part of policy formulation and as an instrument of performance assessment (Batterbury and Hill, 2002: 4), although there is no explicit endorsement of this approach in official UK Government guidelines such as HM Treasury (2003).

This thesis follows the academic convention of using the term ‘evaluation’ in the broader sense of describing the process as a whole, except where circumscribed by the term ‘ex post’.

8.2.2 Context for evaluation

Gordon (2002: 154) distinguishes three broad contexts for evaluation – those of project, performance and process. The research question adopted for this thesis is concerned with project evaluation – arguably the most significant of the three in that it enables comparison of different policies and selection of the most promising.

8.2.2.1 Project evaluation

Project evaluation, whether ex ante or ex post, is essentially economic with an answer that is either binary (yes / no) or at least ordinal (allowing ranking of the project alongside its competitors or substitutes). Conceptually, such project evaluations
belong to the family of cost-benefit or cost-effectiveness studies: according to Gordon (2002: 154), "One key variation within this family is in terms of whether an absolute standard of worthiness is to be set, commensurate across all projects, or just one permitting relative evaluation among those with related objectives." A second source of variation is whether objectives depend on the preferences of the policy-maker or the preference functions of those affected. Gordon (2002: 154-155) stresses that counterfactuals play a central role in this form of evaluation; with opportunity costs to be compared with expected benefits in ex ante studies, and attribution of outcomes to the project in ex post analysis.

8.2.2.2 Performance evaluation

Performance evaluation takes it for granted that the projects or programmes themselves are worthwhile, and examines how well they are being delivered. An important requirement is to develop performance indicators that are regularly available, objective and consistent, a requirement developed in practical guidance for evaluators (11.5). Often, there may be a gap between what is measured and the real wishes of policy-makers or consumers (Gordon, 2002: 155).

8.2.2.3 Process evaluation

Process evaluation analyses the way in which polices are delivered, with emphasis on the potential for improvement, and implying co-operation between practitioners and analysts. Tensions may well exist between this method, which is related to the 'learning by evaluation' approach, and pursuit of performance (or 'instrumental') evaluation.

8.3 Evaluation and Theory

Practical evaluation rests on a concept of how the economy works – in other words, some sort of theory. For example, governments and agencies assist small businesses to make use of broadband technology because they believe it will help them to compete more effectively. This belief constitutes a 'programme theory' (2.5), whether or not that is explicitly acknowledged. Systematic evaluation requires a theory or
‘intervention logic’ linking the activity to its desired outcomes. Practitioners are advised to assess the factors in the additionality ‘logic chain’ that link the intervention activity with its net outputs and outcomes (DTI, 2006: 16; 102).

According to this approach, theory plays a central role in all aspects of the evaluation process:

“Theory directs us to look for certain relationships and not others. It tells us that certain variables are fundamental, others can be ignored. It supplies a causal narrative linking the policy instruments with the target variables. These causal narratives allow us to understand how policy works and to explain its operation to others. Where the theory that underlies evaluation is not stated, it remains implicit.” (McVittie and Swales, 2003: 17).

This programme theory is often explicitly recognised in the process of evaluation by constructing a model, in which the policy being assessed may be presented as a dependent variable. Such formal presentation requires clear definition of variables, especially if it is proposed to use the model to make comparisons between different programmes. It is necessary to avoid uncertainty of the sort caused by use of a ‘fuzzy concept’, defined by Markusen (2003: 702) as “one which posits an entity, phenomenon or process which possess two or more alternative meanings and thus cannot be reliably identified by different readers or scholars”.

As discussed previously (6.5), the case for intervention is usually based on one of two rationales which in their turn rest on an underlying programme theory: (1) market failure or (2) clearly defined government distributional objectives based on equity considerations (HM Treasury, 2003: 11). Clarity in the specification of the objectives of regional policy helps identify the associated programme theory. So too does clear statement of the key target variables needed to observe to decide success in meeting the stated objectives. Therefore, a statement of objectives should form an inherent part of the evaluation process; assessed as part of prior (ex ante) appraisal and subject also to reassessment during the progress of the project (Batterbury and Hill, 2003).

Without clear objectives, it is difficult to construct targets by which success can be measured; but as argued above (6.6), such clarity is not always easy to secure. Making the programme theory explicit has the advantage of enabling the theory to be
tested against reality. For those who argue for the use of evaluation as a learning device, it also permits feedbacks from outcomes to theory, and helps assess the value of pursuing the stated objectives. So evaluation can play an active role in the process of policy formulation for regional and local development, and arguably, this contribution constitutes its ultimate objective (Batterbury and Hill, 2003: 1-2). The success of learning through evaluation may well depend heavily on the institutional context, as Figueiredo (2004) discusses in the case of Portugal, with its strong municipalities and weak planning regions.

Evaluation may have both backward and forward looking purposes:

- Summative evaluation describes what outputs and outcomes were generated;
- Formative evaluation explains how, why, and under what conditions a policy intervention worked, or failed to work (DTI, 2006: 36)

Formative evaluations may therefore be important to understand what for determining what actions succeed, and learning from experience

**8.3.1 Philosophical Paradigms**

While all appraisal or evaluation should relate to a programme theory, a variety of techniques may be employed. Some (e.g. cost benefit analysis, multicriteria analysis) have been developed specifically for the purpose of evaluation, while others are borrowed from the main scientific disciplines. Thus the questionnaire survey is taken from sociology, regression analysis was created by statisticians, the input-output model comes from economics, the expert panel draws largely upon legal practices, while the geographic information system and ethnographic observation emerged from their corresponding sciences (Williams, 2000: 2).

One point to note is that techniques may be derived from disparate philosophical paradigms constructed around different assumption about the nature of social phenomena, and that these assumptions may in certain cases be incompatible (Williams, 2000: 5). For this reason, combining of techniques within a mixed-method
evaluation is the subject of an ongoing debate between evaluation theoreticians and practitioners. Smith (1994) distinguishes three broad approaches to this debate (purist, dialectical and pragmatic).

This thesis follows the pragmatic approach (Patton, 1988). As argued below (8.4.1), practitioners grapple with the problem of measuring the impacts of programmes whose nature, scale and time-scale vary widely. The focus of the thesis is fixed on developing a model suitable for the task of project appraisal and evaluation, and in particular for comparing the relative contribution of different policies and projects. All the appraisal and evaluative methods presented below fall into this category of project evaluation.

8.3.2 Alternative models

While recommended evaluative practice in the UK and the Republic of Ireland has moved increasingly in the direction of expressing as many outputs as possible in money terms (6.6.2), some observers have questioned the use of such value for money studies. Instead they suggest a ‘theory of change’ evaluation model of regional development policy in which evaluators, practitioners and social actors work together to chart out a desired trajectory of change and the key steps necessary to bring it about (Diez et al., 2002). This argument bears similarity to the ‘learning through evaluation’ approach.

Connell and Kubisch (1998) define a theory of change as “a systematic and cumulative study of links between activities, outcomes, and contexts of the initiative.” In a theory of change approach, all stakeholders – community participants, funders, and evaluators – work together to develop the basic ‘theory’ or reasons why the activities may lead to outcomes. The evaluation determines the degree to which the trajectory of change unfolds as projected, while it also links the strategies to the final effects produced (Milligan et al., 1998: 55). This approach has already begun to permeate official UK evaluative advice, e.g. DTI (2006: 18).

Recent developments in the understanding of regional economic theory also have implications for the process of evaluation. Karlsson and Stough (2002: 14) maintain
that the central message of the new economic geography (4.8) is that the dynamic interaction between geographic market potentials and rational firms itself “creates the comparative advantages of regions in the form of localised increasing returns to scale”. This in turn requires regional policy evaluation to be based on a new theoretical framework, with dynamic models capable of dealing with increasing returns and interrelationships between regions. For example, pursuit of a cluster policy (6.7.3) may require evaluators to assess the contribution projects make to the development of a local skill base and other externalities, as well as their direct economic returns (Forfás, 2003: 37-41).

Evaluative techniques should therefore reassess the goals, methods and demarcation of regional policies; and develop a long time perspective, new types of data and new methodological perspectives. This makes it still more important to be clear about the motives for evaluation: for example does it seek to determine whether some absolute level of return is required (e.g. cost per job against an arbitrary threshold) or are schemes compared with each other to allocate scarce resources more effectively (e.g. costs per job for different schemes)? In the latter case, decision-taking may be more difficult because of different objectives between schemes (Pickernell, 1999: 18).

8.3.3 Evaluating long-term activities

Evaluation in such fields as research and innovation may take the form of critical external examination by representatives of the political system rather than an insider peer review process (Kuhlmann, 2003: 144). Such an efficiency review, questioning whether ‘return on investment’ of taxpayers’ money is justified, runs the risk of overlooking possible long-term, desirable effects of creative science and innovation activity. In other words, narrowly focused evaluation may miss the potential of beneficial externalities, which may justify going ahead with investment, even though the impacts on assisted businesses themselves may not seem adequate.

A further complication is that regional and local development policies are implemented in a complex economic milieu where many observed data are interdependent and interrelated. This means it is not enough simply to employ policy instruments that have succeeded elsewhere: an understanding of the socio-cultural and
economic dynamics within particular localities is required to optimise policy impact and delivery (Batterbury 2002; Tsipouri 2002; Batterbury and Hill, 2003: 2). Context, mechanism and outcomes all need to be examined in evaluating policies and programmes: thus Pawson and Tilley (1997: 81) conclude that evaluations should be “mechanism- and context-driven” rather than programme-led. This suggests that specific policy instruments should be analysed against a variety of contexts, enabling the degree of ‘context-specificity’ to be ascertained. This notion sits well with the idea that evaluation needs to operate “not as a performance assessment tool but as an integrated component of a learning-focused policy” (Batterbury and Hill, 2002: 8). On the other hand, it may well complicate the process for those at the sharp end.

8.3.4 Meta-evaluation

The lessons learned may be extended by techniques of meta-evaluation, defined as “the process of delineating, obtaining, and applying descriptive information and judgmental information about an evaluation’s utility, feasibility, propriety, and accuracy and its systematic nature, competence, integrity/honesty, respectfulness, and social responsibility to guide the evaluation and publicly report its strengths and weaknesses” (Stufflebeam, 2001). Meta-evaluation examines different policies and evaluations in order to determine policy elements that work in different contexts and environments (Batterbury and Hill, 2002: 4). Policy makers need to be able to identify which approaches can be generalised and which depend closely on their context, a dilemma that meta-analysis may help resolve: “In collecting and, importantly, analysing evidence evaluation proffers a solution to this dilemma, enabling intelligent evidence-based policy to function better.” (Batterbury and Hill, 2003: 5).

Without contesting this assertion, it is apparent that relating the success or failure of policies to their contexts itself implies a theory, in which the context determines how well a certain policy works. It may be argued that such contextual explanation – expanding the ground theory is obliged to cover – makes it more difficult to apply meaningful tests. At some point, policy-makers or practitioners are required to decide for themselves how wide the context should be.
Meta-evaluation has also been described as a relatively weak form of analysis because, its results may be derived using “different and not necessarily consistent methods – which may not lend themselves to ‘adding up’ without adjustments being made to compensate for these differences” (DTI, 2006: 47).

8.4 Evaluation in practice: issues

8.4.1 Multiple objectives

A particular problem arises where practitioners are set multiple objectives – for example, specifying targets for particular geographic areas, genders or socioeconomic or ethnic groups. The emergence of more complex multiple objectives for regional policy, including social, political and environmental goals, has made the disentangling of its economic effects a more difficult task (Armstrong and Taylor, 2000: 262).

Policy-makers and practitioners are faced with the reality that there is no direct and universally accepted measure of social welfare (6.2.1). Before any process of evaluation can take place, the real objectives of the policy have to be specified; and such specification is not always apparent.

Even within the confines of regional development, the idea of ‘well-being’ can take a number of forms – growth in some quantitative factor such as jobs or GDP, changes in some quality such as ‘capacity’ or ‘empowerment’, or redistribution in wealth or income (McGovern and Hill, 2002). Increased awareness of social exclusion has been accompanied by the emergence in the 1990s of theories of local and regional development based on the concept of social capital, the view that strong community bonds and local networks can significantly improve an area’s economic prospects (Armstrong and Taylor, 2000: 254). Linked to this idea of social capital is the idea of capacity building, not only developing the ability of individual people to rejoin the mainstream economy but also raising the institutional capacity of the community as a whole (Armstrong and Taylor, 2000: 256). Emphasis on measures to enhance capacity building has featured strongly in the growing range of policy instruments favoured by the European Community, such as Community Economic Development...
(Commission of the European Communities, 1996). Thus Eurada (2005: 1) stress that “the main role of RDAs is to change or improve the framework conditions in which the socio-economic fabric of a region is operating”. Such a wide field of activity encourages regional development bodies to operate essentially as “networked organizations, achieving their objectives by working in partnership with other public and private actors” (Danson and Whittam, 1998: 3.4).

While UK regional policy has been linked to employment and (to a lesser extent until the 1990s) income, there have also been various allied objectives, such as land reclamation, urban regeneration, factory creation and improvement, and encouraging clustering in growth industries (Pickernell, 1999: 5). McVittie and Swales (2004: 16) maintain that English RDAs have been set unclear and imprecise policy objectives. This confronts evaluators with a conceptual problem: how can a straightforward increase in GDP be compared with greater equality or empowerment? There have also been challenges to the orthodox use of money as the common comparator, with emphasis on developing substitutes or complements to GDP (Midmore, 2001) and work on developing an Index of Sustainable Well-being in Wales (Hill, 2002: 7).

Recent iterations of UK evaluation have set a stronger emphasis on consideration of capacity building. Thus DTI (2004: Annex C) recognises that agencies are “increasingly active in non-project activities that do not produce traditional outputs, and so capturing this ‘catalytic’ and ‘influencing’ role is essential to full monitoring of RDA performance.” Agencies’ contribution to capacity building is captured within the concept of ‘strategic added value’, defined as the influence an RDA can exert on partners and key stakeholders to achieve strategic objectives over and above the contribution it can make through its project and programme spend (DTI, 2006: 143n).

However, a major difficulty is introduced by the very different time-spans associated with different projects, including capacity building. Thus it may be possible to measure the impacts of programmes to encourage, for example, inward investment or marketing within a period of 5–10 years. Scottish Enterprise estimated their promotion of clusters would occupy a 7–10 year time-span (Richmond, 2001: 10). But the beneficial outcomes of major regeneration projects may take decades to become apparent. A similarly long time-frame may be expected in the case of some
of the basic awareness raising measures set for implementation under the Entrepreneurship Action Plan for Wales, such as the Dynamo project, which sought to provide school children with entrepreneurial role models (WDA, 2002d). In principle ex ante evaluation can and should compare the discounted benefits of widely differing programmes; although it is difficult to imagine policy-makers taking seriously ex post evaluations conducted after a lapse of many years. McVittie and Swales (2004: 16) cite the hypothetical example of Regional Development Agencies failing to achieve long-term regional growth targets, and ask who would be accountable – the then serving relevant minister(s) and RDA chief executives, or those who had held these offices in the past?:

“This problem is increased because it is difficult to monitor the RDAs’ actions or to decide whether the actions are appropriate ex ante. Of course, the lack of agreement of the determinants of regional development in general will make identifying the counterfactual – what would have occurred had there been no regional policy – uncertain, so that even ex post evaluations with clear targets will face difficulties.”

According to DTI (2006: 8), (ex post) evaluation has to supplement monitoring because it requires assessment of effects on third parties (e.g. multiplier and displacement effects) and unintended effects.

8.4.2 Capacity building measures

This difficulty becomes apparent when attempting to compare the benefits of conventional economic policies with measures designed to promote ‘capacity building’, which is generally defined as development work that strengthens activities and the ability of community organisations and projects to build their structures, systems, people, networks and competencies (PACT, 2004). In the context of the European Commission’s LEADER I programme, capacity building refers to the development of an individual’s or a group’s ability to identify opportunities, mobilise resources, and develop and implement plans, while a parallel term, animation, has been used to describe various processes adopted to improve the use of latent or under-utilised resources in local areas (O’Reilly, 2004: 9).
Capacity building may cover a spectrum from purely technical inputs – such as training staff or new computer systems, through strengthening organisations and the institutional environment, to developing people’s understanding of their own needs, rights and entitlements:

“To be meaningful in the long term, capacity building is best done as part of a process, rather than as an end in itself; carried out in partnership, and not as a condition of funding or as something which is imposed” (Green and Battock, 2004: 2).

Measuring the impact of capacity building activities raises both conceptual and practical issues. First of all, many capacity building programmes are long-term in nature, and in deciding how current resources are to be committed, the value of their future benefits have to be compared with those of short-term initiatives. An arbitrary measuring rod is provided by the traditional cost-benefit approach of applying a discount rate to future outcomes in order to obtain a measure of Net Present Value (NPV), which is discussed in greater detail below (9.4).

More difficult for practitioners is the problem of how to isolate the effect of regeneration (and other) activities from the many other influences that affect the local economy. The case for regeneration usually takes the form of arguing that environmental and amenity improvements attract businesses and individual people to establish themselves or remain in the area. In theory, such benefits will be reflected in the level of local income and should therefore be measurable. In practice, it is not easy to disentangle the precise effect of such regeneration programmes, whose cumulative effects are spread over many years, from changes arising from other influences.

One conceptual solution to this problem might be the use of a counterfactual measure. This could take the form of identifying another area similar in population, sparsity, income and so on, but which does not benefit from the specific regeneration programme analysed. Then a number of indicators could be selected for the comparator areas, such as employment, income per head or other measures of economic and social welfare. In practice, however, it may well prove difficult to find a truly comparable area that provides a genuinely counterfactual measure, as
discussed in the context of area based urban regeneration in England by Lawless (2007: 5).

In part the problem of comparing conventional policies with capacity building measures is a further manifestation of the conceptual ‘long-run versus short-run’ dilemma discussed above (6.6.2). However it also makes plain the very real difficulty inherent in judging between different objectives – even though capacity building may well in the long run lead to higher levels of income. It may well be worth while investing in projects that raise local people’s skills and self-confidence, which in turn will ultimately produce a higher standard of living. It is often difficult however to measure exactly how great a benefit has been derived. Ideally practitioners should have at their disposal a ‘common currency’ measure against which every programme could be measured and compared. After all, development agencies are obliged by their business planning processes to decide at a particular point of time how much of their annual budgets should be allocated to long-term regeneration strategies rather than to shorter run measures promoting such activities as innovation, inward investment, property development or business support.

This complexity of regional policy makes analysis and comparison of different policy packages more difficult, especially where objectives vary. This is particularly relevant considering that ‘official’ (as opposed to academic) analysis in the UK has concentrated on ‘cost per job’ measures. Although such analyses have attempted to move from gross costs per job to more sophisticated net measurements, taking account of additionality, displacement and multiplier effects defined below (11.11), it may be argued that they are not the most appropriate for present activities (Pickernell, 1999: 5).

In principle at least, it is possible to conceptualise use of a formula such as Net Present Value (defined in 5.2.1) that enables comparison of like with like across different time spans, although fixing the rate of discount would always involve subjective judgement. However, in reality it is far from clear how comparisons can be made between such outcomes as jobs created by increasing the area of industrial floorspace and participation in community activities, even assuming a large measure of ex ante knowledge. So in the real world, practitioners may be obliged to give up
seeking the 'common currency' solution and confine direct comparisons to activities that have a broadly similar outcome – such as increased GVA.

This is precisely the position that confronts development agencies which are required by government to undertake a wide variety of activities, not all of which can be linked to readily measurable value-adding outputs of the sort described below (11.6). The Welsh Development Agency responded to the difficulty of measuring the impacts of long-term projects by developing a number of benchmark measures to enable the tracking of progress. In essence, these capacity building indicators record movement either toward or away from desired targets. The approach seeks to establish benchmarks for a three-tier range of indicators, using movement from the baseline measures to assess progress (Williams, 2005: 17-18).

The non-comparability of short- and long-term projects need not vitiate attempts to create standard impact assessment models, as many of the agencies’ programmes will produce measurable, comparable outputs. It does however complicate the task of assessing the true contribution made by agencies to overall welfare, and evaluating their own value as institutions.

8.4.3 Measurement, Learning, Objectives

Evaluation in the broader sense has been defined as “the process of systematically and carefully assessing the value any action, policy or programme, and of capturing policy learning in a format that has intrinsic utility for subsequent policy formulation” (Batterbury and Hill, 2002:1). This definition suggests two distinct purposes, (1) measurement and (2) learning, which are pursued in order to improve policy. Most evaluations have focused on the first half of this definition, assessing the value of initiatives; and, as discussed below, there can be conflict between the two distinct objectives of measurement and learning.

To some extent, the learning and policy-framing element is now generally recognised and accepted. For example HM Treasury (1997: 12), in its official guidance on appraisal and evaluation in central government in the UK stated:
“Evaluation examines the outturn of a project, programme or policy. When carried out it adds value by providing lessons from experience to help future project management or development of a specific policy. It may also contribute to the quality of wider policy debate.”

It was significant that this Treasury guidance for practitioners appeared to favour a distinct pecking order – measurement accorded the first place, improvement of project management or policy development next, and consideration of wider policy issues coming in a distant third. Later formulations of Treasury guidelines appear to have downgraded the significance of the learning element, although commending use of the ROAMEF cycle approach (Rationale, Objectives, Appraisal, Monitoring, Evaluation, Feedback) (HM Treasury, 2003: 2).

While the guidelines specify that the main purpose of (ex post) evaluation is to ensure that lessons are widely learned, communicated and applied when assessing new methods (HM Treasury, 2003: 5), there was no advocacy of the ‘learning by evaluation’ approach per se. Nevertheless, the learning element features in the work of practitioners: for example, one aim of the mid term evaluation of the EU Objective 1 programme in Wales was to help prepare for ex post evaluations, which were not expected to be carried out until after final programme closure in 2008 and would therefore be too late to influence the initial shape of any post-2006 programming (Old Bell 3, 2006: 8).

Later formulations of UK practice appear to upgrade learning significantly. Thus DTI (2006: 8) advocates the use of learning for “development of future projects and programmes through benchmarking, good practice guides and case studies which, over time, will help all RDAs deploy their resources more effectively”. It argues that its evaluative framework will generate learning opportunities as various interventions by different agencies are evaluated on a consistent basis, allowing comparison of what approaches work best in particular circumstances.

The increased emphasis on evaluation, learning and development of evidence-based policies are arguably part of the same process as the emergence of the knowledge-based economy, often described as developing the ‘learning region’ (Florida, 1995; Morgan, 1997; Hill, 2002; Batterbury and Hill, 2003: 2). Evaluation may seek to
develop this learning process by helping to formulate new policy objectives as well as playing its more obvious role of assessing which mechanisms fulfil defined objectives best. Batterbury and Hill (2003) describe the various stages at which evaluation can promote policy formulation and the problems that this can involve.

One problem implicit in this discussion is the apparent trade-off between practical measurement on one hand and learning through evaluation on the other. Practitioners (and, more particularly, their paymasters) need measuring devices, however imperfect, that enable them to measure the contribution of economic policies and justify their existence to the taxpayer. Quite often, this process is regarded by practitioners as a threat – evaluation has for some had a negative connotation, seen as “an enforced activity which has to be done to satisfy the needs of others, usually funding organisations and most notably Government” (Scottish Enterprise, 2003: 2).

While evaluation as a form of learning may help avoid that perception of threat, it also runs the risk of allowing those in the know to move the goalposts. The more numerous the objectives, the longer the time required to fulfil those objectives, and the more indirect and nebulous the benefits, the greater this risk becomes and the more difficult it is to deliver the tangible measurements governments demand. For example, Diez (2002: 301) observes that the European Community finds itself caught between a desire to develop new evaluative methods, and a reality in which the European Parliament and other institutions wish to know what gains are being made by spending money on EU Structural Funds. This is not the only pitfall: Weiss (1998: 25) notes that evaluation may provide an excuse for inaction at a policy or programme level, while Palumbro (1987) describes how it can be caught up in the political context, such as the emergence of new policy priorities, findings that are politically insensitive, or failure to ensure the findings politicians seek as an excuse for unpopular decisions. Such problems may limit the scope for effective use of evaluation, whether for measurement or learning, but do not negate the need for useful techniques to appraise, monitor and judge the results of important and costly policies.

There is however a more serious conceptual dimension to the ‘measurement vs learning’ debate. Thus McVittie and Swales (2003: 17-21) stress that the construction of economic models linking policy instruments and target variables is theory-
dependent, and influenced by availability of data, in a circular process which is itself affected by past activities and past theories. They argue that various techniques advocated to "cut through this particular Gordian knot" are 'theory-lite'. For example, evaluation using matched samples of companies, one of which receives programme support while the other does not, itself implies a prior theory and faces difficulty in identifying such samples in the normal operation of policy, requiring setting up the exercise experimentally (Taylor, 2002). Similarly, use of shift-share analysis, removing national effects operating through the region's industrial structure to leave only a region-specific residual (Fothergill and Gudgin, 1979), requires a true theory to identify the national effect correctly (McVittie and Swales, 2003: 20).

One 'solution' to this problem is to abandon such 'intellectual rumination' and apply a trial and error, iterative evaluation process, in which policies are used as long as they produce positive change, and adjusted when they do not. The existence of devolved administrations could enable simultaneous iterative comparison of a number of policies by a process of benchmark competition (McVittie and Swales, 2003: 20). This approach however brings with it a number of disadvantages:

- To facilitate frequent adjustment, policy instruments would need to be short-term in nature;
- Government credibility would be damaged because it could not commit itself to future policy;
- Use of iterative evaluation processes might lead to local rather than global optimum positions, a pitfall reminiscent of evolutionary school's criticism of neoclassical macroeconomics (5.3) (Nelson and Winter, 1982; Ormerod, 1994; Potts, 2003).
8.4.4 Evaluators and their social and institutional environment

Naturally, those who carry out evaluation are affected by the social environment in which they find themselves. Describing evaluation of the European Social Fund Objective 3 programme in France, Barbier (1999: 375-376) describes how evaluators, far from being neutral providers of the methods and technical tools of their craft, inevitably play a role in “the tradeoffs between complex and often conflicting requirements from stakeholders”. This requires them to possess strong professional ethics, especially given the long time-period involved. Leca (1997: 13) distinguishes between ‘political assessment’ – the prerogative of elected representatives (and, periodically, of the people electing them) – and ‘evaluation of policies’, which is nevertheless an integral part of the implementation process of those policies. Evaluators need to create an ‘area of autonomy’ for evaluation, separate from the rest of the political and managerial systems, and drawing on ethical frameworks as well as technical skills (Leca, 1997: 15). But the evaluator’s role is to promote “authentic co-operation between independent actors acting strategically”, who may have differing, and more often than not conflicting, interests: this means that “the evaluator’s stance cannot remain purely technical/scientific, but is also political/normative” (Barbier, 1999: 376).

Evaluation of regional policies may proceed from social, economic or purely financial perspectives, the latter from the point of view of particular firms and individuals relocated under the policy or the government responsible for funding the policy (Schofield, 1989a: 194). It may be limited to a regional perspective, or may also consider ‘national’ impact. These distinctions are relevant to later discussion of deciding how to assess both the costs and benefits of development programmes.

The UK Labour government’s devolution and ‘New Localism’ approach has raised issues for delivery and evaluation of regional economic policy (Balls, 2002). A new institutional arrangement described by McVittie (2005: 2) as ‘constrained discretion’ has emerged with the delegation of policy delivery within England to local Regional Development Agencies (RDAs), while overall control rests with a set of Whitehall Departments. This has been accompanied by abandonment of previous insistence that
regional policy has zero-sum ‘national’ impact because of crowding out (McGregor and Swales, 2003:17) (12.5).

McVittie and Swales (2003: 1) voice the fear that “because of standard asymmetric information problems, such policies will be extremely difficult to evaluate, which has implications for both their efficiency and democratic accountability”. Fothergill (2005: 666-667) voices the concern that regional policy is being driven into uncharted territory, in which divergence in absolute levels of prosperity would be more acceptable. On the other hand, the variety of regional initiatives operated by the UK Government, the devolved administrations and their agencies could be argued to provide new opportunities for making comparisons and for promoting best practice (HM Treasury et al., 2003: 4; McVittie and Swales, 2004: 12).

Subsidiarity and devolution of power have opened up space for local and regional involvement in dialogue and policy formation and have helped foster an evaluation culture (Williams, 2000). The European Union has contributed to raising standards and increasing the comparability of regional policy, although its centralised control of evaluation design is argued to have constrained innovative approaches to evaluation (Batterbury and Hill, 2003: 4-5). Evaluation is cited as one of six principles underlying the reform of the 1994-1999 European Structural Funds (Vanhove, 1999: 489-490).

The need to evaluate socio-economic programmes financed by the European Union has been a major factor behind rapid growth of evaluative practice in Europe during the 1990s; and this expansion has been accompanied by the MEANS Programme (Methods for Evaluating Structural Policies) launched by the European Commission in 1995 (Williams, 2000: 1).

8.5 Conclusions

To sum up, the context for evaluation continues to become more complex and demanding. This reflects the increasingly diverse objectives set for practitioners as regional policy moves away from simple job creation and tackles a greater variety of goals. The basic institutional architecture has also evolved with the growth of
influence of the European Union and what in some ways can be seen in part as the complementary development of devolved administrations and regional development agencies.

Along with new opportunities, recent theoretical, policy and institutional developments have placed new demands on the process of appraisal and evaluation. On the one hand, governments increasingly call for evidence of value for money at a time when policy goals are becoming more complex and their achievement more difficult to record. On the other hand, evaluators are being urged by academics and others to ensure the potentially threatening process they manage contributes to the sharing of insights and learning from experience. This dilemma presents a major challenge to agencies driven to demonstrate evidence of success in meeting the various objectives they are set.
Chapter 9  Techniques of Evaluation

9.1  Introduction

This chapter moves from consideration of the overall methodological framework for appraisal and evaluation to examine the various evaluative techniques used. Section 9.2 reviews the concepts of effectiveness and efficiency, while 9.3 presents a taxonomy that distinguishes between measurement of impacts on economies as a whole – macro-economic evaluation – and measurement of impacts on groups of companies, micro-economic evaluation. It also discusses the main features of these techniques. Chapter 12 examines the criteria used. The remaining sections look in greater detail at the nature of cost-benefit analysis, which is argued to offer the best means of effecting appraisal and evaluation.

9.2  Effectiveness and Efficiency

Following the treatment of macro-economic evaluation by Vanhove (1999: 388), this thesis further suggests that the purpose is the same in both the macro and micro cases, to estimate (a) the effectiveness and (b) the efficiency of any policy or technique. Consideration of effectiveness implies that the policy has clearly defined objectives, but this is far from being universally the case (Armstrong and Taylor, 2000: 261; Vanhove, 1999: 388; McVittie and Swales, 2003; Wren, 2005: 257). Given clear objectives, however, it is desirable that benefits and costs are assessed together (as in cost-benefit analysis) so that a gain in benefit is offset by the cost incurred, and an overall net measure of value obtained.
For both the macro- and micro-economic strands, consideration of efficiency also implies that some calculation is being made of the benefits obtained for the costs incurred, but that the focus of the evaluation is on the cost incurred for a certain agreed benefit (as in cost-effectiveness studies) rather than on comparison of different benefits and costs. Thus Wren (2005: 258) states that cost-effectiveness analysis is the most appropriate technique of resource allocation “if the policy has a single overriding objective and the resources are constrained”. This would be the case, for example, if the type of job created by a project is not an issue (Wren, 2005: 258).

With diffuse objectives, however, the deficiencies of cost-effectiveness analysis become obvious; for example, it would be absurd if appraisers compared a project creating employment in high performance engineering with one providing work in a low-wage call centre purely on the basis of project cost-per-job. Thus Swales (1997a) argues for use of cost-benefit analysis where jobs created are heterogeneous (Wren, 2005: 258). The need for a more discriminating system becomes greater as an economy approaches full employment, and where job-creating projects inevitably involve displacement and ‘crowding out’ of jobs in non-assisted occupations (Forfás, 2003: 12).

9.3 Macro- and micro-economic evaluation

The principal macro- and micro-economic models used during the course of evaluation of regional policies and projects are set out in Table 9.3 and discussed briefly in the following sections. This classification develops the approach proposed by Vanhove (1999: 387-406) and does not rule out the potential for one model informing the other: thus values of target variables and parameters obtained by system-wide methods such as input-output analysis may be used in micro-economic evaluation of individual projects. At the boundaries, moreover, one methodology may merge into another: thus one strand of cost benefit analysis in the field of housing and urban renewal concentrates on the minimisation of cost subject to provision of a given standard of accommodation (Schofield, 1989a: 108). This approach may be considered analogous to adoption of a simple ‘cost per job’ evaluation of regional economic policy where jobs are assumed to be homogenous.
Among macro-economic models, the techniques of multiple regression (especially those using time-series data) are of particular interest. Application of models to evaluation implies that the policy studied may be treated as a dependent variable and its impact identified. In principle, the techniques listed may be used both for appraisal and evaluation; and carried out either singly or in combination with each other. The task set by the research problem involves selection of a method of project appraisal that helps agencies to optimise their contribution to regional development. Among the techniques considered above, this thesis identifies an appraisal model that is both theoretically sound and of practical use, and assists in monitoring and evaluating regional projects.

Table 9.3  Classification of Evaluative Methods

<table>
<thead>
<tr>
<th>Evaluative Approach</th>
<th>Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macro-economic evaluation</td>
<td>Modified shift-share analysis</td>
</tr>
<tr>
<td></td>
<td>Multiple regression approach</td>
</tr>
<tr>
<td></td>
<td>Time-series</td>
</tr>
<tr>
<td></td>
<td>Cross-sectional</td>
</tr>
<tr>
<td></td>
<td>Cost-benefit approach (macro)</td>
</tr>
<tr>
<td></td>
<td>Input-output</td>
</tr>
<tr>
<td></td>
<td>Computable general equilibrium (CGE)</td>
</tr>
<tr>
<td>Micro-economic evaluation</td>
<td>Survey techniques</td>
</tr>
<tr>
<td></td>
<td>Cost-benefit approach (micro)</td>
</tr>
<tr>
<td></td>
<td>Cost effectiveness approach (micro)</td>
</tr>
</tbody>
</table>

Based on Vanhove (1999: 387-406)

Evaluation may aim at assessing the impact of regional policies as a whole, for example by comparing the performance of assisted regions with those not receiving assistance, which provide a counterfactual control group. Such a broad-based approach for example could assess the effectiveness of a development agency. In such cases, the evaluative process would need to take account of system-wide effects – for example, additional benefits secured through the multiplier effect or the savings to the exchequer resulting from additional tax revenue.

Such a macro modelling approach requires a systematic and theoretically rigorous treatment of regional capacity and labour market constraints: thus Gillespie et al. (2001: 157)
demonstrate the important role of regional wage setting and migration in determining employment adjustment over time. While official guidelines have declared that evaluation should take into account the displacement of other activity in both product and factor markets (HM Treasury, 1995: 26), ex post evaluations of Regional Selective Assistance based on (microeconomic) sample surveys of companies involved have confined themselves to product market effects alone (Gillespie et al., 2001: 126 and 126n). These official evaluations have adopted input-output/Keynesian multiplier techniques that assume excess capacity and rigid wages, rather than initial capacity constraints, a bargained real wage, price flexibility and migration (Gillespie et al., 2001:127).

Alternatively, evaluation may be focused on the contribution made by individual projects or programmes, which are defined as groups of related projects (HM Treasury, 2003: 1). Assistance schemes can be individually evaluated against their stated aims and against the value they provide in meeting those aims as far as the taxpayer is concerned (Pickernell, 1999: 18). Among the reasons for the latter course, policy makers may wish to compare the relative success of public funds devoted to different programmes – for example, the performance of an innovation programme in comparison with financial support for all start-ups. Even in such cases, it may be necessary to consider system-wide effects in coming to a judgement; especially if one programme is subject to greater regional leakage and therefore a lower regional multiplier effect. Moreover, as suggested in Table 9.3, the same technique may be used in a macro or micro context.

### 9.3.1 Shift-share approach

Shift-share (or ‘components of change’) analysis is a traditional and well-known method, which was applied to empirical analysis as early as the 1960s and underwent numerous extensions and improvements in the 1970s (Richardson, 1978; Fritz and Streicher, 2004: 2). Shift-share analysis separates the structural growth determinants from those related to the competitiveness of a country or region (Fritz and Streicher, 2004: 2). It enables the standardisation of data about changes in regional employment or income to take account of (1) the particular economic structure of regions and (2) different national growth rates of sectors (Vanhove, 1999: 117).
Shift-share analysis is the best-known example in regional economics of decomposition, a form of ex post analysis that requires an underlying theoretical model, with data for both dependent and independent variables corresponding to different points in time (Jensen-Butler et al., 2002: 337). Schofield (1989b) used a shift-share model of employment change to assess regional economic policies pursued by the Canadian government in Quebec and the Atlantic Provinces. Two main variations of the technique have been developed, one of them adjusting a region's sectoral structure to the national pattern, the other adapting regional sectoral growth rates to national sectoral changes (Vanhove, 1999: 117).

Traditional shift-share analysis separated out various components of changes (for example, in regional employment growth):

1. National 'share' component (N), which shows employment change expected from the overall national employment growth, providing the norm from which regional deviation can be measured;

2. Net regional 'shift' component, which shows any deviation from the national share of employment growth, and which can be further subdivided into two parts:
   a) A proportionality shift component, also called the 'structural' or 'industrial mix' component (S), which for each sector shows the employment change expected from the national growth rate of that sector after allowing for the overall growth rate (in other words, the portion of regional growth that is due to an abundance of either quick or slow growing sectors);
   b) The regional component (R), the difference between the actual employment change observed and the sum of the other two components.

Calculating the regional component (R) indicates all influences on the total regional employment that are not accounted for by its sectoral structure, and this enables tentative conclusions to be drawn about a region's competitive advantage. A positive regional component may indicate a beneficial degree of competitiveness, although the fundamental factors explaining from this cannot be derived from the analysis (Vanhove, 1999: 119).
At one time, the shift-share technique was asserted to be ‘atheoretical’ (Fothergill and Gudgin, 1979). This claim, however, is disputed by McVittie and Swales (2003: 19), who stress the role of programme theory:

“The problem with this claim is that shift-share will only accurately identify the impacts of national effects working at the level of the region if it accurately replicates these effects. In order to know that we need the true theory of how these effects work.”

Vanhove (1999: 120-121) notes a number of limitations of shift-share analysis. First, it does not attach any special importance to sectors that may have a strong regional multiplier effect, in contrast for example, to input-output analysis. Secondly it does not take into account differences in labour productivity – a negative differential employment component may reflect above average labour productivity growth rather than slow growth of output. Stilwell (1970: 452-453) concludes that although shift-share analysis integrates well with some regional growth theories, it is not a theory of regional growth, and, as in the case of multiple regression, any conclusions drawn rest upon the underlying programme theory. Vanhove (1999: 120) also records sensitivity of results to the level of data disaggregation used, whether by sectoral breakdown, spatial boundaries or time period. In general terms, it is necessary to ask whether any regional growth deficiency can be attributed to its structure or to other factors. If regional economic structure is indeed the problem, then measures to alter regional distribution of economic activity may be effective. If not, more general improvement in infrastructure or other instruments may be required (Vanhove, 1999: 121-122).
In the context of regional policy evaluation, a modified form of shift-share analysis was used by Moore and Rhodes (1973) to analyse the effects of British regional policy between 1950 and 1971, an exercise described as "simplistic but simple" (Cheshire and Magrini, 2002: 210). Shift-share analysis was used to produce the expected share of employment for all Development Areas in the United Kingdom, eliminating the influence of regional differences in industry-mix. The technique employed then divided the study into two distinct periods -- 'policy off' and 'policy on', a counterfactual approach (Taylor, 200: 175). A linear trend was constructed for the 'policy off' period; this was extrapolated forward into the 'policy on' period to produce expected values of employment for each year. The difference between these values and the actual values observed was taken as an estimate of the difference made by regional policy (Vanhove, 1999: 389-399; Armstrong and Taylor (2000: 368). Shift-share methods are currently used in forecasting employment, for example for Eastern England, with and without imposition of long-term supply-side constraints (ODPM, 2004c: 230-231).

The main criticism of this process is that only regional policy is treated as an explanatory factor, and that other possible autonomous factors (e.g. labour reserve) are not considered. Moreover, no allowance is made for underlying cyclical movements in aggregate demand, while the values observed of the regional shift component (R) are reduced to the level of sectoral disaggregated increases (Vanhove, 1999: 400).

Taylor (2000: 175) observes that dissatisfaction with this trend-based approach led researchers to construct economic models designed to explain the time-path of employment rather than relying simply on extrapolation from policy off into policy on periods. In fact, according to Armstrong and Taylor (2000: 368), this proved the most significant result of the shift-share approach. These time-path studies generally used regression analysis to estimate the effect of a number of potential explanatory variables, including that of regional policy.

9.3.2 Multiple regression approach

In the multiple regression approach, actual growth of income or employment is compared with the growth that would have happened anyway in the absence of the policy studied (i.e.
Multiple regression can make allowance for influences other than the policy that affect the dependent variable, for lags in policy effect and for significance testing. For example, Vanhove (1962) sought to explain employment growth in the 78 economic areas of the Netherlands between 1950 and 1960 by setting up a basic equation using as independent variables (termed 'autonomous factors') the labour reserve (A), the wage level (L), the degree of industrialisation at the beginning of the period (I) and the share of growing industries (S), defined as chemicals and metal production. The following basic equation was then used to estimate $Y^*$, the calculated percentage growth of employment or industrial employment over 1950-1960:

$$Y^* = \alpha_0 + \alpha_1 A + \alpha_2 L + \alpha_3 I + \alpha_4 S \quad (9.3)$$

The formula was then used to compute growth of employment in each of the assisted regions, and the value of $Y^*$ calculated by the equation compared with the actual growth of employment ($Y$). The difference between the two values ($Y^* - Y$) was said to be explained by the impact of Dutch regional policy, with observed employment in the assisted regions about 50 per cent higher than the $Y^*$ values calculated by the equation (Vanhove, 1999: 396).

### 9.3.2.1 Regression time-series data

Many multiple regression studies analyse time-series data, in which phenomena such as real gross domestic product, interest rates, or money supply are observed at specific points of time (Koop, 2000: 5). Such approaches are felt to provide a valuable basis for judging the effectiveness of policy. Thus Begg et al (1998: 414) conclude:

"In some cases the use of aggregative time-series data has been very successfully analysed through relatively simple econometric models whereby the net additional employment has been considered as a dependent variable in regression models where the independent variables are measures of the strength of policy and other non-policy variables that might be expected to influence the dependent variable concerned (see for example Moore et al., 1986)."

Such analysis however needs to be treated with some caution, even assuming that the data collected is accurate. Its conclusions essentially depend upon the underlying 'programme
theory', which in the case cited above maintains that any residual not accounted for by the dependent variables chosen (the labour reserve, wages, prior industrialisation and the share of growing industries) can be explained by regional policy.

This treatment is reminiscent of early neoclassical attempts at growth theory by Solow (1957), which attached a high explanatory value to "advances in knowledge and technology", claiming it accounted for 87% of per capita productivity growth. In a sceptical view of such reasoning, The Economist describes regression analysis as "number-crunching to discover the relationship between different economic variables" and advises:

"The findings of this statistical technique should always be taken with a pinch of salt. How big a pinch can vary considerably and is indicated by the degree of statistical significance and R squared" (Economist.com, 2004)

However as part of a body of evidence-based study, in which programme theories are tested from different standpoints, regression analysis models can play a valuable role: as discussed below, the conclusions of Vanhove (1962) discussed in the previous subsection were supported by a further study (Netherlands Economic Institute, 1972). A number of other models using multiple regression analysis have also contributed to understanding the workings of regional economies, for example Moore et al (1986). Hill and Munday (1992) analyse how regional distribution of foreign direct investment within the UK is determined by a number of independent variables – preferential government assistance, male earnings relative to that of the UK, share of spending on trunk roads (a proxy for access to markets) and the region’s share of total UK employment. The results provide evidence that both financial incentives and access to markets are important determinants of the regional location of foreign direct investment projects (Hill and Munday, 1992: 543).

9.3.2.2 Cross-sectional models

An alternative to the time-series method of modelling the interregional movement of industry is provided by the use of cross-sectional data, in which typically the ordering of the data does not matter (Koop, 2000: 6). This cross-sectional approach was used in studies by Twomey and Taylor (1985) and Taylor and Twomey (1988) to study the movement of industry between pairs of regions in Great Britain.
Using cross-sectional data, Twomey and Taylor (1988) constructed a model that studied the movement of industry between each pair of regions during 1960-1977, employing three non-policy and two policy variables:

- Size of region of origin;
- Geographic position of destination region in relation to region of origin, or distance between them;
- Availability of female labour (for manufacturing industry);
- Investment incentives (policy variable); and
- Location controls (policy variable).

Using multiple regression techniques to obtain the effect of each variable on movement of industry between each pair of regions, the influences of the policies were estimated by simulating a ‘policy off’ situation. The conclusion was that about 42 per cent of moves by manufacturing industry in the UK between 1960 and 1977 could be explained by location controls (29 per cent) or investment incentives (13 per cent) (Taylor, 2002: 181). Again, however, policy instruments may work interactively, so that estimates of the impact of single instruments may be misleading.

Use of this cross-sectional approach has not been followed up. Taylor (2002: 181) suggests that this reflects a changed policy environment in the UK. Encouraging firms to move to Assisted Areas was thought too costly for the benefits gained; regional policy has focused more on stimulating indigenous growth; and foreign direct investment is regarded as benefiting the economy as a whole. Moreover the interregional movement of industry database has not been maintained (Taylor, 2002: 181).

9.3.2.3 Cost-benefit approach

This section introduces the basic cost-benefit approach to economic evaluation, which forms the basis of the research project developed in this thesis. It examines the steps required to set up a cost-benefit based evaluation. It includes treatment of problems that often arise in the case of large-scale projects but which may be overlooked with smaller
scale projects undertaken by regional development agencies. The arguments in this section apply equally to macro- and micro-evaluation. The rationale for considering cost-benefit analysis as an appropriate mechanism for evaluating regional economic policy together with the practical methods and problems involved in implementation are discussed in greater detail in the following chapter.

Every evaluation has to take account of the benefits of a policy, and make some comparison of that benefit with the costs incurred; in other words perform an analysis of the costs and benefits in its widest sense. For example, Swales (1997a: 74) maintains that the national efficiency requirement set by the UK Government for ex ante project appraisal (12.5) is essentially a cost-benefit hurdle. However ‘cost-benefit analysis’ refers to a more specific and tightly defined methodology, whose underlying theory can be traced back to welfare economics of the nineteenth century, but which was developed independently as a practical device in connection with the Flood Control Act of 1936 in the United States (Pearce, 1983: 14). Cost-benefit analysis is defined as “A method of reaching economic decisions by comparing the costs of doing something with its benefits” (Economist.com, 2004), although this seemingly simple proposition involves complex calculations of the true value of future costs and benefits to society as a whole.

9.4 The cost-benefit approach in practice

The essence of the cost-benefit approach is comparison of likely costs with a stream of future net benefits discounted back to the present using an appropriate discount rate (Swales 1997a: 75). Use of the discount rate means that costs and benefits at different times can be converted to ‘present values’ so that they can be compared (HM Treasury, 2003). For example, at a ten percent discount rate, the present value of one pound ten years from now is 37 pence and the present value of one pound fifty years into the future is 0.67 pence. In general, the longer the time frame, and the higher the discount rate, the smaller will be the impact of any given year on total net benefits (Bjornstad, 2004: 5).

This calculation produces a single figure, reflecting the present value of the surplus generated by investing in the project rather than at the discount rate. This surplus is called net present value (NPV). The net present value of a project is defined as the difference
between discounted future benefits and initial costs (plus discounted costs if they are incurred during the life of the project) (Swales, 1997a: 75 and 75n). The NPV formula seeks to apply a rate of discount that makes future costs and benefits comparable to their current equivalents, with the discount rate reflecting society’s preference for consumption today rather than consumption tomorrow (Armstrong and Taylor, 2000: 388).

Net present value is regarded by the UK Government as the primary criterion for deciding whether public action can be justified (HM Treasury, 2003: 26-27). The UK Treasury’s current recommended rate of discount is set at 3.5 per cent; although “for projects with very long-term impacts, over thirty years, a declining schedule of discount rates should be used rather than the standard discount rate” (HM Treasury, 2003: 26-27). It should be noted that cost-benefit analysis is said to have has an in-built tendency for to “discriminate against the future” (Pearce, 1983: 53). Moreover, as discussed below (10.3.1), consideration of the underlying theoretical foundations is open to question.

Net present value is associated with an internal rate of return (IRR), defined as the discount rate \( r \) which makes the net present value of the stream of future benefits equal to the net present value of the stream of costs (Vanhove, 1999). Often used in capital budgeting, the IRR is in effect the interest rate that makes net present value of all cash flow equal zero: for example, this could be viewed as the return that a company would earn if they expanded or invested in themselves, rather than investing that money abroad. A variant occasionally encountered is provided by the Modified Internal Rate of Return (MIRR), also known as ‘financial management rate of return’, which is the Internal Rate of Return (IRR) adjusted for negative cashflows. This is accomplished by using a ‘safe rate’ to provide for future negative cashflows. The MIRR better reflects the profitability of a project: the IRR assumes the cash flows from the project are reinvested at the IRR, whereas the Modified IRR assumes that all cash flows are reinvested at the firm’s cost of capital (Investopedia, 2004).

Vanhove (1999: 390) describes net present value and its associated IRR as “the centrepiece of the cost-benefit approach”. As Bjornstad (2004: 14) explains, to determine whether or not a particular project is acceptable, the calculated IRR must be compared to the ‘hurdle rate’, which is defined as the minimum acceptable rate of return that should reflect the time
value of money, risk, etc. However, using the internal rate of return as a decision tool may generate results that are inconsistent with the use of net present value. This inconsistency arises because of the effect of the timing of the cost-benefit flows on the IRR calculation. Any project that has relatively large positive cost-benefit flows early in its life will generate a relatively large IRR (Bjornstad, 2004: 14).

Bjornstad (2004: 14) further maintains that the use of IRR as a measure for choosing between projects is inappropriate when capital rationing exists. This problem arises because of the assumption that the cost-benefit flows are reinvested at the internal rate of return rather than at the cost of capital as in NPV. What this implies for the decision maker is that the ranking of projects will depend as much on their relative size and the timing of their cost-benefit flows as it will on the actual cost-benefit flows, where the actual flows should be the only determinant of acceptance or rejection. Bjornstad (2004: 15) concludes:

"The inconsistency implies that the usefulness of the IRR method is limited. Further, difficulty arises when calculating the IRR of a project that has negative cost-benefit flows after the first period. Due to the mathematics of the calculations, it is possible under these circumstances to calculate multiple IRR's that equate the net present value of costs with the net present value of benefits. This is clearly an undesirable situation."

9.5 Selection of inputs and input values

Cost-benefit analysis requires identification of input variables, which may be divided into (1) Parameters, (2) Costs and (3) Benefits; and various classifications of costs and benefits are offered by a number of scholars, including Bjornstad (2004: 6), Armstrong and Taylor (2000: 389) and Vanhove (1999: 389). Because the values chosen for these variables will significantly influence the final values calculated, decision-makers must satisfy themselves that the values chosen are reasonable (Bjornstad, 2004: 6). This is especially important in ex ante evaluation, where values cannot be predicted with certainty:

"For these reasons, it is important that a cost-benefit analysis does not present a single number as the sole estimate of net present values. Rather sensitivity and scenario analysis should be conducted to illustrate how the results change with different analytical choices and with variation in the uncertain levels of key costs and benefits." (Bjornstad, 2004: 6)
Such a sensitivity analysis would assess how net benefit would vary given different outcomes – for example, different levels of pollution that might be difficult to predict in advance. Where there are several uncertain outcomes, the typical approach is to carry out three cost-benefit calculations – the most optimistic, the most pessimistic and the most likely (Sloman, 1997: 346-347). The use of sensitivity analysis is fundamental for ex ante appraisal in order to test the vulnerability of options to unavoidable future uncertainties (HM Treasury, 2003: 32). In cases where there many variables with significant uncertainties, appraisals may use Monte Carlo analysis, which is defined as “a risk modelling technique that presents both the range, as well as the expected value, of the collective impact of various risks” (HM Treasury, 2003: 33: Annex 4). A simple form of sensitivity analysis is applied to the model proposed by this thesis (Chapter 14).

9.6 Cost-benefit analysis procedure

The procedure required for a cost-benefit analysis at first sight appears straightforward, but each of the stages involves a number of difficulties (Sloman, 1997: 343). These steps are detailed in Table 9.6, which also refers briefly to complications that may arise.

**Direct (private) monetary costs** are relatively easy to identify, and would include such items as construction, operating and maintenance costs.

**Non-monetary costs** may involve such items as pollution, spoiling the landscape, noise and inconvenience to residents. While in the case of some projects, such as a tunnel, these costs may be confined to the construction phase, other projects (e.g. a new airport) may lead to considerable ‘externalities’ resulting from their operation, such as aircraft noise (Sloman, 1997: 343). These non-monetary externalities are usually the most difficult to identify.

**Direct (private) monetary benefits** also relatively easy to identify, and consist of revenues received from the users of the project (e.g. money taken at a toll bridge) (Sloman, 1997: 344).
**Private non-monetary benefits** comprise benefits consumers derive over and above what they actually pay, an element defined by economists as ‘consumer surplus’. This can be illustrated by the area between the demand curve (showing what people are willing to pay) and the price charged (Sloman, 1997: 344). Samuelson and Nordhaus (2001: 98) note that generally an economist would recommend that a free road should be built if its total consumer surplus exceeds its costs.

**External benefits** are enjoyed by non-users of the project – for example, road users who benefit from less congestion when a new underground railway is opened (Sloman, 1997: 344). Such gains may be both monetary and non-monetary.

**Measuring costs and benefits**

While conceptualising different costs and benefits is relatively easy, measuring them may prove much more difficult.

**Measuring direct (private) monetary costs and benefits**

As recorded in Table 9.6, even quantification of direct (private) monetary costs and benefits poses problems – such as estimating future prices and deciding whether to base prices on actual market prices even if these are known to be distorted by monopoly power. Since such distortions exist throughout the economy, it is usual to use actual market prices unless there is a distortion applying only to the specific project (Sloman, 1997: 344).

In the case of non-monetary costs and benefits, indirect means have to be used to attempt estimates. One method of measuring non-monetary benefits – reflected by the idea of consumer surplus – requires estimation of the demand curve and the area under it. A further problem is noted by Sloman (1997: 344), who cites the example of the Channel Tunnel, where consumer surplus gained by the project may replace a smaller consumer surplus from a competing service (e.g. cross-channel ferries), in which case the non-monetary benefit will be made up of (i) only the additional consumer surplus of those who switch modes of transport and (ii) the full consumer surplus of people who otherwise would not have crossed the channel.
### Table 9.6: Cost-benefit Analysis Procedure

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Elements</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify costs and benefits</td>
<td>Direct (private) monetary costs</td>
<td>Easy to identify, but market prices may be distorted by e.g. monopoly power</td>
</tr>
<tr>
<td></td>
<td>External costs (monetary and non-monetary)</td>
<td>Difficulty of identifying externalities, especially those that are non-monetary</td>
</tr>
<tr>
<td></td>
<td>Direct (private) monetary benefits</td>
<td>Relatively easy to identify</td>
</tr>
<tr>
<td></td>
<td>Private non-monetary benefits</td>
<td>Requires estimates, e.g. by survey</td>
</tr>
<tr>
<td>Assign monetary values to each cost and benefit</td>
<td>Direct (private) monetary costs and benefits</td>
<td>Uncertainty about future prices</td>
</tr>
<tr>
<td></td>
<td>Non-monetary costs and benefits</td>
<td>Difficulty of measuring consumer and producer surplus</td>
</tr>
<tr>
<td></td>
<td>Monetary externalities</td>
<td>Requires estimates, e.g. by survey</td>
</tr>
<tr>
<td></td>
<td>Non-monetary externalities</td>
<td>Difficulty of assigning monetary values to externalities</td>
</tr>
<tr>
<td>Assess risk and uncertainty</td>
<td>Risk (calculable) and uncertainty (incalculable)</td>
<td>Use of sensitivity analysis to deal with uncertainty</td>
</tr>
<tr>
<td>Discount future costs and benefits</td>
<td>Based on current rate of interest</td>
<td>Issue of ‘discounting against the future’</td>
</tr>
</tbody>
</table>

**Source:** Based on Sloman (1997: 343)

An alternative method is to focus on specific non-monetary benefits such as the hours of travelling time saved by a new motorway. An estimate would have first to be made, and a monetary value attached by using the average hourly wage of e.g. business people and...
lorry drivers to estimate the value of the time saved. Leisure time saved is usually measured at a fraction of the average hourly wage, an admittedly arbitrary approach (Sloman, 1997: 345).

**Monetary externalities** (e.g. the cost of losses suffered by taxis and bus companies following construction of a new underground line, or the benefit earned by a service station on a new motorway) would normally be counted at face value (Sloman, 1997: 345).

**Non-monetary externalities** are likely to prove the most difficult to calculate. The general principle is to try and find out how much people would be prepared to pay to obtain the benefits or avoid the costs. According to Sloman (1997: 345), this can be attempted in two ways:

- Ask people (by questionnaires) how much they need to be compensated (although there are practical problems e.g. ignorance of the scheme, or dishonesty;

- Make inferences from behaviour to assess how people have reacted to similar projects elsewhere, e.g. has the noise of a new airport been reflected in a dip of local housing prices, or has a new bypass resulted in reduced road accidents?

Such calculations raise difficult issues – for example, how much value should be set on a human life?

**Assess risk and uncertainty**

The third step involves taking account of the likelihood of a cost or a benefit occurring, which means assess risk and uncertainty. Risk is relatively straightforward to estimate – the value of a cost or a benefit is simply multiplied by the probability of its occurrence (Sloman, 1997: 345). In the case of uncertainty, however, all that is known is that a certain outcome might occur – the likelihood is unknown. One approach, as discussed above (9.5) is the use of sensitivity analysis, a method for analysing uncertainty by changing input variables and observing the sensitivity of the result (Bjornstad, 2004: 6), where “a range of
possible values of uncertain costs and benefits are given to see whether the project's desirability is sensitive to these different values” (Sloman, 1997: 346).

Discounting

While it is straightforward to apply an agreed interest rate to cost and benefit streams, discounting in fact raises complex theoretical issues (10.3.1).

9.7 Conclusions

This chapter has suggested a taxonomy of evaluative techniques, both macro- and micro-economic (9.3). These include cost-benefit analysis, whose main features are introduced in 9.4. A brief presentation of the steps required to carry out a cost-benefit analysis (9.5 and 9.6) indicates the complexity of the issues raised. A full discussion of the advantages and disadvantages of cost-benefit analysis is presented in Chapter 10.
10 Cost Benefit Analysis in Policy Appraisal and Evaluation

10.1 Introduction

This chapter examines cost-benefit analysis in further detail, with special regard for its advantages and disadvantages as an instrument of appraisal and evaluation.

10.2 Advantages of cost-benefit analysis

One clear advantage of cost-benefit analysis is that it offers a means of expressing economic variables in quantifiable and often monetized terms. The UK Government defines cost-benefit analysis as an “analysis which quantifies in monetary terms as many of the costs and benefits of a proposal as feasible, including items for which the market does not provide a satisfactory measure of economic value” (HM Treasury, 2003: 4). This emphasis on the monetary nature of the method reflects the view of Pearce (1983: 3), who regards cost-benefit analysis as a procedure for:

- measuring the gains and losses to individuals, using money as the measuring rod of those gains and losses; and
- aggregating the money valuations of the gains and losses of individuals and expressing them as net social gains and losses (sic).

The UK Government recommends use of cost-benefit analysis rather than an alternative approach, cost-effectiveness analysis, which is defined as “analysis that compares the costs of alternative ways of producing the same or similar outputs” (HM Treasury, 2003: 4). This recommendation reflects an academic consensus that cost-benefit analysis provides
the best way of evaluating the impacts of regional economic policy. Thus Vanhove (1999: 389) maintains:

"The cost-benefit approach is the most appropriate method of analysing the efficiency of regional economic policy with respect to a specific assisted region or the economy as a whole."

Cost benefit analysis performs the useful function of forcing a careful consideration of the trade-offs involved in determining whether to proceed with a development or not (Auty, 2005: 9). Armstrong and Taylor (2000: 388-389) agree that cost-benefit analysis provides the most comprehensive approach to ex ante evaluation:

"The cost-benefit approach attempts to quantify the full range of effects and to translate them into measurable costs and benefits so that the net benefits of policy action to society as a whole can be estimated."

Similarly, Swales (1997a, 1997b) argues that cost-benefit analysis provides the most comprehensive approach to ex ante evaluation of regional policy, noting that it seeks to measure the overall social benefits and costs of any action and attaches a lower value to benefits and costs in the future. Adler and Posner (1999) maintain that cost-benefit analysis, suitably constrained, is consistent with a broad array of appealing normative commitments, and it is superior to alternative methods of project evaluation: "It is a reasonable means to the end of maximising overall welfare when preferences are undistorted or can be reconstructed."

This widespread academic support for the use of cost-benefit analysis for evaluation is based on both theoretical and empirical grounds. For example, Swales (1997a: 84) argues that cost-benefit analysis "is a technique firmly based in welfare theory, is flexible to use and is specifically designed to deal with situations where market failure exists and where costs and benefits are distributed over time".

The practical advantages of the cost-benefit approach are summed up by Bjornstad, (2004). As society has evolved, it has become more complex, seemingly at an increasing pace, and the social benefits of governmental activities are ever more questioned. Decision-makers therefore need a framework that structures information in a way which makes this
complexity more tractable while still taking into account its implications. Cost-benefit analysis has the potential to advance this process significantly.

In simple terms, cost-benefit analysis imposes an accounting framework that prescribes classes of benefits and costs to consider, techniques of measuring them, and approaches for their aggregation. Key parameters, such as the discount rate, are highlighted, together with recognition of and methods to deal with inherent uncertainties. While originating as an analogy to private studies of investment, and thus to calculate a ‘go-no-go’ decision, the technique is flexible and can be used to choose among a range of alternatives, to compare projects of differing lengths, and to identify instances where costs and benefits place identifiable groups at special advantage or disadvantage.

In the context of the United Kingdom, Swales (1997a: 79-80) advances pragmatic arguments for adopting a cost-benefit approach:

“Cost-benefit is a practical evaluation technique that is firmly rooted in sound economic principles. It is not esoteric or obscure. It is routinely taught in undergraduate economics degrees and its strengths and weaknesses are therefore very well known. Moreover, as a framework generally recommended by the Treasury for the evaluation of public expenditure projects, it is familiar to administrators in the civil service.”

Finally, Samuelson and Nordhaus (2001: 374) maintain that cost-benefit analysis enables efficiency to be set by balancing the marginal costs of an action against the marginal benefits of that action. This is a notable advantage. Hill (1989b: 330) suggests that the technique of marginalism has an overwhelming logic: “Whenever a decision is made, extra benefits must be compared to extra costs.”

10.2.1 Cost-benefit analysis and discounting

A further important characteristic of cost-benefit analysis – the idea of working out the value today of outputs that will only be delivered in the future – is noted by Isaacs and Martin (1993: 185). Defining cost-benefit analysis as ‘a method of deciding whether or not a particular project should be undertaken, by comparing the relevant economic costs and the potential benefits’ they stress the uses to which this technique may be put:
“It can be used for private investment projects, calculating outlays and returns, and estimating the net present value of a project: if this is positive the project would be profitable. Cost-benefit analysis is also frequently used by governments in an attempt to evaluate all the social costs and benefits of a project (e.g. road building), which is much more problematic, involving such considerations as externalities, public goods, macroeconomic consequences, etc.”

Thus cost-benefit analysis can be used for both macro- and micro-economic evaluations; and the comparison of likely costs with future net benefits discounted back to the present using an appropriate discount rate to produce a single figure is described as ‘the essence of the cost-benefit approach’ by Swales (1997a: 75). The concept of net present value was examined in detail above (9.4); and the following subsection examines disadvantages associated with discounting against the future.

However, as Pearce (1983: 12-13) concludes, the definitions of both ‘benefits’ and ‘costs’ have to be widely drawn. A benefit may accrue in money form, or in some other sense of pleasure or happiness: “in the language of economists, we require only that it be some welfare or utility gain”. Thus the standard cost-benefit procedure may be considered as an example of a Social Welfare Function, a mechanism for ranking different economic states, which is required to judge whether changes to the economy generated by application of policy have brought about a net benefit (McVittie and Swales, 2003: 3; 2004: 3). Similarly, a ‘cost’ need not be defined in terms of a money flow, but is anything that imparts a loss of utility or welfare: in the context of cost-benefit analysis, ‘cost’ always implies opportunity cost, and this condition raises issues for the model developed in this thesis (10.3.2).

In the context of environmental decision making, Bjornstad (2004: 1) maintains that cost-benefit analysis has a particular role to play in dealing with the highly complex nature of modern society. Unexpected outcomes may mean that well-intended decisions can produce losses as well as gains in social welfare:

“Decision makers therefore have a great need for a framework which structures information in a way which makes the complexity more tractable, but still takes into account the implications of the complexity. Cost-benefit analysis is an analytical tool which has the potential to significantly advance this process.”
Cost-benefit analysis provides a means of comparing complex projects, even when benefits and costs occur during different time periods, and enables the systematic comparison of the value of outcomes with the value of resources needed to achieve them. This makes it particularly appropriate to deal with issues of public policy, considering classes of benefits and costs that are more far-reaching than a business decision focusing only on net profits for the enterprise (Bjornstad, 2004: 1).

Many such public decisions require considering the widest possible impacts of policy, both monetary and non-monetary effects – for example, on such untraded amenities as peace and quiet, for which the housing market could provide a surrogate, with house prices as the indicator, or goods and services for which no real or surrogate market exists. An example of the latter is the value (or ‘willingness to pay’) that people attach to conserving a cherished species such as the blue whale, which would not be reflected in the commercial market for whale meat or oil. Such ‘willingness to pay’ either has to be ignored, or constructed by hypothesis; and there have been a number of attempts to establish such hypothetical or experimental markets (Pearce, 1983: 10-11). HM Treasury (2003: 21-23) sets out the main techniques recommended to elicit non-market values.

10.2.2 Measuring efficiency and equity

As discussed above (6.2), public policy has tended to focus on two major purposes, improvement of efficiency and improvement of equity. Cost-benefit analysis has traditionally been focused on efficiency – on providing policy makers with an indication of the magnitude of net benefits associated with a particular project or policy. However, it may also be constructed to adjust for various goals, for example to provide a different weighting than that of the existing distribution of income (Pearce, 1983: 4). Thus Bjornstad (2004: 1) maintains that the cost-benefit analyst should track the distribution of costs and benefits among the various segments of society:

"In an ideal world, the analyst would attempt to determine how benefits are distributed by age, sex, income, race, geographic location, and time. At a minimum, the analysis should attempt to ascertain, to the degree possible, if imbalances between benefits and costs are present for those segments of the
population which are most vulnerable.”

Cost-benefit analysis, when properly conducted, will identify potential ‘Pareto improvements’ and measure the magnitude of the difference between gains and losses based on the Hicks-Kaldor criterion (6.2.1.2). But although it can, to some degree, identify gross imbalances in the distribution of benefits and costs, and even provide special weighting for certain issues of equity, such as unbalanced impacts on readily identifiable and disadvantaged social groups, it cannot measure the multi-dimensional aspects of overall project desirability that may include such factors as sustainability, altruism, ethics, public participation in the decision process and other social values. Thus, cost-benefit analysis informs the decision making process, but it does not by itself make decisions (Bjornstad (2004: 4).

Pearce (1983: 3) agrees that a seemingly ‘rational’ social decision where the benefits to society exceed costs may not necessarily be morally correct: cost-benefit analysis “makes no claim to produce morally correct decisions”. Instead it is an ingredient of decision-making, not a method that supplants political judgement. In fact it is possible to distinguish at least two different objections to exaggerating the utility of cost-benefit analysis in decision-making (Pearce, 1983: 4):

- Applying market-place philosophy to non-market situations;
- Implying market-place values are equivalent to moral correctness.

10.2.3 Shadow Pricing

One significant theme to emerge in the conduct of cost-benefit analysis is the need to take into account overall costs and benefits to the economy as a whole, a task usually accomplished by the device of using shadow rather than market prices (Swales, 1997a). While a commercial enterprise may appraise an investment activity in line with its own financial profitability alone, a social cost-benefit analysis seeks to discover how much society as a whole gains or loses. Thus it “claims the right to modify the prices used in the commercial accounts because market forces are not operating as predicted by neo-classical economics and/or the distribution of wealth is not considered just” (Cameron, 2005: 1). By contrast, it is argued that market prices fail to represent “the General Equilibrium and/or
Social Justice valuations that would optimise efficiency and equity for the whole society”, a feature commonly described as “market failure”, which implies a misallocation of resources, because they are not priced in line with their marginal costs and benefits (Helios, 2004: 7).

The concept of shadow pricing derives from mathematical programming as the marginal value imputed to an input or output at the optimum (Schofield 1989a: 53):

“More generally, as interpreted in cost-benefit analysis, it refers to the value associated with a change of social welfare following use or loss of the marginal unit of an input or output. Shadow prices thus reflect true economic or social value at the margin.”

The shadow price is defined as the price attributed to a good or service by an economist in the absence of an explicit market price (Isaacs and Martin, 1993). In the context of a maximisation problem with a constraint, the shadow price on the constraint is the amount that the objective function of the maximisation would increase by if the constraint were relaxed by one unit, and is equivalent to the value of a Lagrangian multiplier (Econterms, 2004).

In a perfectly competitive economy, the shadow price would be equal to the market price, but in general, the two may differ. This difference is frequently attributed to the phenomenon of market failure, although as argued above (6.5), it could often be more correctly viewed as a ‘social equity failure’, that is the failure to deliver social and spatial objectives that would increase overall social welfare.

It is common, therefore, to allow for the estimated shadow price when carrying out a cost-benefit analysis for Government projects, since few believe that real-world economies are perfectly competitive. However, Isaacs and Martin (1993: 769) maintain that there is “no generally accepted method for calculating shadow prices and they are usually considered unreliable.”

In social cost-benefit analysis, instances of market failure may justify modifying observed prices to so-called “shadow prices”. A shadow price may be higher or lower than observed
prices depending on the specific nature of the market failure (Cameron, 2004: 1). Shadow prices are used where costs and benefits need to be considered from the standpoint of society as a whole, so that among other issues the opportunity costs of inputs are considered (Helios, 2004: 5). Taxes and subsidies should be omitted from such calculations except in circumstances where a pollution tax, for example, internalises external costs that would otherwise be ignored.

One example of the use of shadow pricing is in the labour market. In practice, the usual justification for providing grants has been to reduce the level of involuntary unemployment. Jobs created by such projects do not, however, reduce involuntary unemployment one-for-one. Instead, some workers may move from existing jobs, so that the activity entails an opportunity cost for the economy. Use of cost-benefit analysis, and in particular the device of the shadow wage, also permits explicit consideration of the spatial dimension, by attaching a higher net value to benefits in the assisted region. It should be added that a comprehensive cost-benefit treatment would also require use of macroeconomic model of full regional effects to obtain a more accurate measure of benefits and costs for the economy as a whole; and also that it may be difficult to set a money value on wider social impacts. While this approach has not to date been adopted in the United Kingdom, it is significant that such a methodology has been approved for use by development agencies in the Irish Republic (Barry et al., 2002; Forfás, 2003).

10.2.4 Shadow Wage

Shadow pricing may be applied to the markets for all factors of production, but the shadow price for labour – the shadow wage – has a particular significance for regional economics. Bannock et al. (1992: 389) define the shadow price of a factor as the opportunity cost to society of engaging in some economic activity, a concept applied to situations where the actual price cannot be charged, or where the actual price charged does not reflect the real sacrifice made when some activity is pursued.

Its significance arises because regional policy has tended traditionally to focus on adjustments to employment patterns, and therefore the shadow wage features more prominently in regional economics than shadow prices of other factors of production.
Moreover, as Cameron (2004: 2) points out, “labour comes with people attached and therefore has enormous complications”. For example, cost-benefit studies of rural areas have generally argued that wage rates have strong civil society elements that maintain them above the general marginal productivity of labour, which would be the free market price.

In assessing the shadow wage, therefore, such considerations have usually modified the observed price downwards. In the context of regional policy, the shadow wage is used to estimate the opportunity costs of creating jobs in the assisted area. However, “this is rarely done to zero as people’s time is usually seen as having some opportunity cost, perhaps in the form of domestic tasks that frees up others for wage work” (Cameron (2004: 2). The shadow wage is used to estimate the opportunity costs of creating jobs in the assisted area, and so the higher it is, the lower is the value added by the project (Taylor, 2002: 194).

The shadow wage determines what proportion of the wage bill should be treated as a benefit in an appraisal model (Forfás, 2003: 32). Setting a lower shadow wage rate in a cost-benefit appraisal means that labour intensive projects are more likely to be selected. This is because the difference between actual wages paid and the shadow wage rate is used to estimate net benefit; so the lower the shadow wage rate, the higher the overall benefit recorded (Swales, 1997b: 863; Taylor, 2002: 194). This plays an important role in the basic decision rule developed in this thesis.

Since labour is actually paid the going wage rate, not the shadow wage rate, shadow pricing means that the actual wage bill will be higher than that in the cost-benefit exercise. In other words, the real resource cost to the economy will be lower than the actual financial cost. Setting a low shadow wage therefore implies a readiness to support a more expensive regional policy (Cameron, 2004: 2).

Shadow wages may have to take account of large variations in the types of labour, depending on skills, regions within countries, and even individual jobs, especially in developing countries. It is thus often necessary to use a set of shadow wage rates, one for each skill, location, economic sector, and even season, rather than a single rate for the whole country (Asian Development Bank, 1997: Appendix 12). Such variation may also be required for economic development policies that have the specific aim of raising
employment or income levels in certain disadvantaged regions. International evidence suggests that the shadow wage should be set at a minimum of 50 per cent (Hegarty, 2002) – that is, the net benefit of a project is only 50 per cent of the gross benefit while, as shown, a minimum shadow wage rate of 80 per cent is used in Ireland.

Variation in shadow pricing enables the appraisal system to take account of changing economic circumstances, and offers the opportunity of taking into account a range of desired social objectives that are not being realised because of perceived market failure. However, it is felt that such changes should not be applied to short-term cyclical variations (Forfás, 2003: 56). Irish agencies are recommended to reduce the shadow wage only if unemployment exceeds 6 per cent for four consecutive quarters.

The practical difficulties involved in setting shadow prices may well account for the call in official UK Government guidelines for costs and benefits normally to be “based on market prices as they usually reflect the best alternative uses that the goods and services could be put to (the opportunity cost)”, although adjusting these for tax differences between different options (HM Treasury, 2003: 19). This guideline appears to rule out the general use of shadow prices in normal appraisal and evaluation. Yet exclusion of shadow pricing from ‘social’ evaluation implies that the market’s price system perfectly reflects the best interests of society, and as already argued (10.2.3) this is far from being the case. The use of market rather than shadow wage in estimating the value of jobs generated by a project would exaggerate its true contribution. Using a shadow pricing system therefore assists construction of a decision rule to determine whether aid should be given to projects which the market would not deem viable, but which are found to be worth supporting when all costs and benefits are considered (Pickernell, 1999: 18-19).

10.3 Disadvantages of cost-benefit analysis

This review of the limitations of cost-benefit analysis begins with outright opposition. Self (1970: 8) attributes the plausibility of cost-benefit analysis to the use of a common monetary standard deriving from exchange situations:
“Outside such situation, common values cannot be presumed, and symbol and reality become easily confused ... To call these judgements £s is to engage in a confidence trick – to exploit the ordinary man’s respect for the yardstick of money in what are actually non-monetary situations.”

A still more trenchant and uncompromising criticism is made by Schumacher (1973: 41-42), who feels that cost-benefit analysis is used by economists to press non-economic values into the framework of the economic calculus:

“This is generally thought to be an enlightened and progressive development, as it is at least an attempt to take account of costs and benefits which might otherwise be disregarded altogether. In fact, however, it is a procedure by which the higher is reduced to the lower and the priceless is given a price. It can therefore never serve to clarify the situation and lead to an enlightened decision. All it can do lead to self-deception or the deception of others; for to undertake to measure the immeasurable is absurd and constitutes but an elaborate method of moving from preconceived notions to foregone conclusions ... what is worse, and destructive of civilisation, is the pretence that everything has a price or, in other words, that money is the highest of all values.”

Without abandoning the cost-benefit approach, these criticisms should be borne in mind when attempting to apply the method to assess the value of policies whose outcomes may be in part non-monetary, such as enhancing the natural or built environment, or raising a community’s self-confidence or capacity to make decisions for itself. Descriptions of cost-benefit as a method of reaching economic decisions by comparing the costs of doing something with its benefits makes the method sound simple and commonsense. In practice, however, it can easily become complicated and is much abused (Economist.com, 16 September, 2004). Careful selection of assumptions can be made to support, or oppose, almost anything, particularly if the decision contemplated involves some cost or benefit for which there is no market price or which, because of externalities, is not fully reflected in the market price. Typical examples would be a project to build a hydroelectric dam in an area of outstanding natural beauty or a law to require factories to limit emissions of gases that may cause ill-health (Economist.com, 16 September, 2004).

10.3.1 Discounting against the Future – the Sustainability Debate

As Pearce (1983: 53) points out, the seemingly neutral mechanical process of cost-benefit
analysis has a basic problem, namely “that the effect of discounting is to discriminate against the future”. One illustration of this discrimination is the imaginary example of a nuclear waste storage facility where an accident costing £10 billion in today’s prices is certain to occur in 500 years’ time. Discounting this at the rate of 5 per cent would produce a ‘present value’ cost of the £10 billion accident of just 25 pence. Thus fixing the rate of discount, far from being a neutral calculation, is in fact laden with value judgement.

According to Bjornstad (2004: 5), discounting the future in this way is one of the weaknesses of cost-benefit analysis because it calculates its results from the present generation’s perspective, raising concerns about ‘intertemporal equity issues’, that is, fairness to future generations:

“In fact, costs that occur far into the future may be given little weight in traditional cost-benefit analysis. Sustainability has developed as an additional consideration for public policy decision making precisely because of the concern that the process of discounting may steer us towards policies that overly emphasise short term gain.”

Some economists have attempted to provide for sustainability within the framework of neoclassical growth theory (Toman et al., 1995: 140). Early work in this field “implicitly modelled sustainable development as non-declining consumption over time and was concerned with intergenerational efficiency rather than equity” (Hartwick, 1977). Theoretically such approaches often proceeded by “making the sum of discounted net benefits across generations as large as possible”: one conventional approach is to maximise present values using some intergenerational social welfare function (Toman et al., 1995: 140-141). Later work, such as the Commons-Perrings model of sustainable development, stress the need for the ecosystem to be resilient to the demands of any one generation (Commons and Perrings, 1992), a major adjustment of traditional emphasis on utility as the dominant element in social welfare.

In the field of policy, agencies are increasingly urged to take into account the sustainable development context in which their impacts are evaluated, particularly by considering the consistency of their programmes with sustainable development principles and targets (DTI, 2006: xvi). Thus agencies are advised that ‘wider effects on sustainable development’
require evaluation. Such consequences are described as “the effect an intervention has on sustainable development through its consequences for economic, social and environmental development – the ‘triple bottom line’” (DTI, 2006: 17). English RDAs’ responsibility is regarded as delivering the ‘weak’ version of sustainability: the ‘strong’ strong interpretation of sustainable development “identifies the need to preserve a critical stock of natural resources or environmental capital as the primary goal – failure in this respect being seen to compromise the achievement of all other goals over the medium term and into the next generation” (DTI, 2006: 58).

While the contributions to GVA of two of the ‘four pillars of sustainable development’ (economic and human capital) may be readily counted, the other two (environmental and social capital) pose greater difficulty and hinder evaluators from assessing the benefits or otherwise of trade-offs between them (DTI, 2006: 60). UK policy makers have sought to overcome this problem by developing a set of framework indicators, 68 in total (DTI, 2006: 61). The complexity of this process raises obvious questions for policy makers seeking simple measures of project impacts. DTI (2006: 64-66) presents suggestions of how evaluators should accommodate sustainable development issues within their approach.

There is little doubt that concern for climate change and sustainability raises fundamental questions about economic appraisal that are difficult to reduce within the format of a cost-benefit framework. Nevertheless, Stern (2006) argues that concern for the future can be justified on economic grounds: “Climate change presents a unique challenge for economics: it is the greatest and widest-ranging market failure ever seen” (Stern, 2006: 1).

Not everyone agrees with this interpretation. Thus Nordhaus (2006) claims that the ‘apocalyptic conclusions’ of the Stern Review are overstated because of its assumption of a near-zero discount rate (0.1 percent per year) and log utility (so marginal utility does not decline much as technological progress causes consumption to rise). Therefore, “the Review’s unambiguous conclusions about the need for extreme immediate action will not survive the substitution of discounting assumptions that are consistent with today’s market place” (Nordhaus, 2006: 21). This seems to confirm the impression that, as with consideration of efficiency, “consideration of sustainability provides the decision maker with additional information, but does not by itself make the decision” (Bjornstad, 2004: 5).
Decisions about balancing the needs of present and future generations involve profound value judgements, and ultimately are beyond economics.

10.3.2 Social Discounting

While a private firm would probably choose the market rate of interest as its discount rate, in the case of cost-benefit analysis, it is argued that the government should use a social rate of discount, which reflects society's preference for present over future benefits. Since the government has a responsibility to future generations and not just to the present generation, it can be argued that a relatively low social rate of discount should be chosen (Sloman, 1997: 348). This issue resurfaces in specifying the appropriate rates of discount for assessing future costs and benefits in the model developed for the research project (14.3.4).

In 9.4, the net present value of a project was defined as the difference between discounted future benefits and discounted costs. While this appears straightforward enough, the underpinning theoretical foundations turn out to be anything but firm, reflecting once more the problems neoclassical economics has in dealing with time.

In order to compute the present (or terminal) value, an appropriate rate of discount (or interest) is required, described as the minimum rate of return (MRR) that a public investment is required to earn if it is to be worthwhile (Schofield, 1989a: 87). Unless a project can earn this rate of return, the resources would be better employed in an alternative use. In principle, there is little difficulty in identifying the MRR as the opportunity cost of the funds used on the project. However, apart from problems of measurement, the MRR rests on two underlying concepts, the social time preference rate (or STP), and the social opportunity cost rate (SOC).

Displacement of private consumption is measured by the social time preference rate (STP), defined as “the rate of return required by society to induce it to sacrifice present consumption for the promise of future consumption as generated through investment” (Schofield 1989a: 88). This displacement of private consumption may be regarded as pure time preference/opportunity cost (Hill and Williams, 2005).
The displacement of private sector investment is measured by the social opportunity cost rate (SOC), defined as "the marginal pre-tax rate of return in the private sector, the rate of return which the funds would have yielded for society, abstracting for externalities, had they not been commandeered for the public sector" (Schofield 1989a: 88). This displacement of private sector investment may be regarded as pure cost of distortion. It is generally assumed that such 'commandeering' involves a loss of social welfare, although the basis for this reflects a strong value judgement (Hill and Williams, 2005).

The shadow pricing method involves a measure based on the 'shadow cost adjustment factor', the ratio of SOC/STP (Schofield 1989a: 90), the rationale of which is provided in the following paragraph. Thus (assuming private investment and consumption given up because of taxation are foregone in perpetuity) if SOC = 6% and STP = 3%, then the present value of the capital outlay of £1 on a project = £1(0.06/0.03) = £2.

So this ratio reflects an underlying presumption that the two rates differ – in simple terms that the 'cost of distortion' discount rate is greater than 'pure time preference/opportunity cost' discount rate. Accepting this proposition, this process in itself implies that future resources have to be discounted by the degree to which the 'cost of distortion' (SOC) is greater than the 'time preference' rate (STP). Both rates attempt to represent opportunity cost, so the shadow cost adjustment factor represents the degree to which there is an additional opportunity cost caused by the public sector taking over private resources (Hill and Williams, 2005).

Adjusted costs and benefit are then both further discounted (by STP). The net present value (NPV) of the project is defined as:

\[
NPV = \sum \left( \frac{B_t - (SOC/STP)K_t}{(1 + STP)^t} \right)
\]

where \(B_t\) is the net annual benefit and \(K_t\) represents annual capital cost flows.

The argument that the 'shadow cost adjustment factor' is the ratio of
Rate of cost of distortion / Rate of pure time preference means that the higher the cost of distortion in relation to the rate of pure time preference, the higher the shadow price.

Hill and Williams (2005) suggest that this raises the question of double counting: having already allowed for additional opportunity cost once – by grossing up the capital outlay by $K_t$ by the ratio (SOC/STP) – the excess of benefits over cost is itself reduced by the rate of discount. If its theoretical roots are taken seriously, the process of shadow pricing of cost is based on opportunity cost, with the presumption that the opportunity cost of public sector investment will exceed that of the private sector. But having grossed up capital outlay, the resulting benefit / cost outcome is discounted once more.

While resolution of this controversy is beyond the scope of this thesis, the debate ought to caution evaluators from accepting at face value the tenets of cost benefit analysis.

10.3.3 Technical Difficulties

Given the positive attributes of the cost-benefit approach, Bjornstad (2004) concludes it is little wonder that pressures to apply the technique to important decisions are growing. However, this comes at a price. The technique is difficult, specialised and costly to apply. Adapting it to issues of moderate size and to agencies with modest resources, as are found at the sub-national level, provides a significant challenge. It has usually been applied to appraise large infrastructure projects rather than evaluate programmes involving a large number of projects (Wren, 2005: 258n).

Despite the favourable academic consensus (10.2), cost-benefit analysis has not been widely used as an evaluative method (Vanhove, 1999: 390), although Blake (1973), Moore and Rhodes (1976) and Armstrong and Taylor (1993) are cited as examples of those who have followed a cost-benefit approach. Selection of the cost-benefit approach, in itself, is a choice; and in business decisions, other simpler models are commonly applied (Bjornstad, 2004: 6). This reflects the reality that in practice economic cost benefit studies can be technically complex and controversial because they call for subjective decisions that can radically affect the outcome (Tietenberg, 2003; Auty, 2005: 9). These subjective decisions
include the assignment of costs to non-market items, the appropriate rate at which the costs and benefits should be discounted, the relative value accorded to different groups within society and the time period and geographical boundary of a project's costs and benefits.

Swales (1997a: 80) attributes the reluctance of policy-makers in the UK to use cost-benefit analysis when evaluating government subsidies to the complexity of the analysis and the difficulty of gathering data: "civil servants do not have the time to undertake a full cost-benefit analysis on each project seeking support". Schofield (1989a: 194-195) concurs that use of the cost-benefit approach to regional policy is more tentative than in other fields, but cites a number of examples where has been applied to evaluation of labour mobility and local labour demand policies in the United States and Canada as well as advance factory construction in the UK.

Cost-benefit analysis was used to a limited extent to analyse the UK's distribution of industry policy in the 1960s and 1970s, but in each case a number of factors on both sides on the cost-benefit account were omitted (Schofield, 1976: 181). In one assessment by NEDC (1963), which claimed a high rate of return for industry policy, Schofield (1976: 182) maintains that the analysis compared benefits accruing to the economy as a whole with costs incurred by only one sector (the government). However, an appraisal model based on cost-benefit methodology has been recommended for use by development agencies in the Republic of Ireland since the late 1970s (Honohan, 1998: 1-2); and a review of economic appraisal in the Republic of Ireland concluded that cost-benefit procedure remained the most appropriate mechanism for assessing value for money for enterprise development projects (Forfás, 2003: ii) (13.4).

10.4 Temporal Issues - the Crowding Out Debate

Policy-makers in the UK preferred to use an approach, which (among other shortcomings) ignores the positive macro-economic effects of regional policy (Swales, 1997b: 861). This is in line with the political outlook commonly encountered during the Thatcher years, which according to Taylor and Wren (1997) regarded the economic case for regional policy as weak and unproven: "Regional policy was viewed as a zero-sum game in which one region could benefit only at the expense of other regions." For example, the UK
Treasury at one stage assumed that jobs created in one part of the UK simply displace jobs elsewhere in the UK (unless created by foreign direct investment) and set down the following stipulation for evaluators:

"Because of crowding-out at the macroeconomic level, effects on employment should not be included as benefits of projects in an efficiency test." (HM Treasury, 1991: 70)

This position was advanced despite evidence to the contrary provided by the Computable General Equilibrium model for Scotland developed by Strathclyde University (McGregor, 2001). Thus strict application of the Treasury rule could have meant rejection of projects even though they have a positive present net value for the UK economy as a whole (Armstrong and Taylor, 2000: 392).

Even during the 1990s, the 100 per cent crowding-out rule was not universally applied within the UK Government: thus the Department of Trade and Industry argued that increasing the competitiveness of the assisted areas raises national productivity and has beneficial effects on the economy as a whole (DTI, 1995; Taylor, 2002: 173n). McGregor and Swales (2003:17) concluded that bureaucratic expedience and politics rather than economics accounted for the crowding-out 'rule':

"There is, in our judgement, no rationale for this view in economic theory, although we recognise the administrative convenience in terms of project evaluation and, in particular, avoidance of the need to conduct detailed investigations of demand effects by region. The view also neatly eliminates the multi-regional lobbying activity that might otherwise arise."

This position has changed considerably. Official UK guidance, as reflected in English Partnerships (2004: 49), notes that crowding-out (regarded as the macroeconomic effect arising from the government’s budget constraint) leads to the displacement of all outputs with the exception of those arising through supply side improvements (ODPM, 2006: 6). Such improvements could take the form of increased employment of factors of production (notably labour) or greater productivity (ODPM, 2006: 6). So overall national economic performance could be enhanced if policies could lift the productivity and employment rate levels of relatively poorly performing cities and regions of the UK towards those of better
performing areas, rather than merely securing a spatial redistribution of economic activity (ODPM, 2006: 7). It is also suggested that social capital (8.4.1) is an asset that could be enhanced without compromising activity elsewhere (ODPM, 2006: 9); and that additional employment secured by regeneration policy will contribute a net addition to national output providing productivity in other regions is unaffected (ODPM, 2006: 17).

Identification of a national impact requires not only a visible supply-side effect, but also that it be greater than alternative uses of the same call on the public finances (OPDM, 2006: 6).

Pearce (1983: 8-9) sets out a philosophical defence of cost-benefit analysis by specifying its normative base in the form of four propositions:

- Cost-benefit analysis requires two (and only two) basic value judgements, that (a) people's preferences count, but careful qualification is needed of whose preferences are to count, and (b) how preferences are to be weighted should also be specified;
- It is legitimate to adjust either of these value judgements as there are no rules for choosing between ultimate value judgements;
- Cost-benefit analysis is a normative procedure, reflecting the nature of economics itself;
- Any value judgements chosen should be made clear.

While stating these propositions helps to clarify the issues involved, it does not in itself resolve the difficulties faced in seeking to recognise the value of the environment and the needs of future generations, the issue of social capital or cultural issues such as the Welsh language. These problems are very real and constitute practical questions for governments and agencies. One contemporary example is the debate about the complex impact of marina developments on Welsh-speaking communities (Newidiem, 2002; Brooksbank et al., 2005). Pearce (1983: 12) concedes that cost-benefit analysis cannot itself take account of the problems raised by such issues, and to the extent that it may fail to identify certain values, he agrees with Schumacher (1973)'s critique (10.3). Practitioners of regional policy have therefore to be constantly aware of the inherent limitations of this and other purely economic measurement techniques.
10.5 Prescribing boundaries for estimating costs and benefits

One problem with no readily apparent solution is specifying who counts in calculating costs and benefits. For example, the use of national boundaries may be inappropriate in examining a policy of sulphur dioxide emissions, where pollution may be 'exported' to another country: "There are in fact no clear rules for setting the 'boundaries' for a cost-benefit analysis." (Pearce, 1983: 13). This view echoes earlier discussion of the indeterminate view of what constitutes a 'region' (2.2). Such boundary issues may be particularly relevant in evaluating regional policies, especially in determining whether or not to take into account displacement effects in other regions.

A wide-ranging evaluation of a business support project, for example, might involve estimating policy effects that reach far beyond firms receiving assistance, and the direct impacts on new jobs and increased output of those firms. They include for example the effects of spillover into other regions as a consequence of interregional trade flows as well as to firms not receiving assistance (Armstrong and Taylor, 2000: 388-389).

As noted, the complications of cost-benefit analysis discouraged its adoption as a widespread technique of evaluation in the UK (10.3.2). A review of appraisal systems in Ireland (Forfás, 2003) discussed the possibilities of various alternatives to cost-benefit analysis:

- A one-dimensional evaluation criterion, such as profitability or number of jobs created, considered too narrow an approach owing to its failure to take account of several outcomes and broader impacts that are important from a policy perspective;
- Non-aggregative procedures, which suffer from the limitation of not presenting overall assessment of impact; and
- Multi-aggregative procedures, including decision (utility) analysis, planning balance sheets and goal achievement matrices, which may be more flexible than cost-benefit analysis but are open to the risks of subjectivity and double counting.
These alternatives are dismissed in favour of cost-benefit analysis:

"Cost-benefit analysis is generally justified on the grounds that it is the only methodology that provides a consistent ranking of projects' measured benefits and costs. It reduces the overall effects to a monetary amount and aggregates. It appears to be superior to rival methodologies in enabling agencies to evaluate projects according to the extent that they contribute to overall economic well being and meet defined economic criteria. It allows agencies to take into account all relevant influences on economic well being, unlike more simple (one-dimensional) decision procedures and it enables agencies to weigh the advantages and disadvantages of projects in a clear, systematic and relatively objective way." (Forfás, 2003: 22).

These advantages explain why cost-benefit analysis is regarded as an appropriate method for the evaluation and ranking of projects. While the process may well be too burdensome except in the case of major projects, evaluative procedures based on the technique may be developed to provide a workable method of appraising and evaluating projects.

10.6 Conclusions

In developing a practical model for evaluating regional policy it is necessary to ensure theoretical soundness as well as practical utility. A system based on cost-benefit analysis, for the reasons studied in this section, provides one means of achieving this goal. It offers a systematic way of counting of costs and benefits, the technique of discounting to allow for the effect of time and a comprehensive measure of net value that can provide practitioners and decision makers with a useful measuring rod.

The survey presented in this chapter points to a strong consensus in favour of basing appraisal and evaluation on cost-benefit analysis. However, there are evident limitations. First, not everything of value to society can be measured in terms of money. Secondly, policies designed to assist specific regions may have effects outside their target area. Thirdly, as noted, there are considerable technical difficulties to be surmounted. Practical evaluation of many individual projects requires a common system to be based on cost-benefit analysis, not repetition of a full cost-benefit exercise on each and every project. Finally, the growing prominence of sustainability highlights the value judgements involved in balancing the wishes of people today with the interests of future generations.
While cost-benefit analysis provides the best available means of constructing a theoretically sound and practicable evaluative method, the limits to its effectiveness need always to be borne in mind. These limitations become particularly evident in comparing projects whose effects are indirect or of long-term duration. The method has also been criticised for its treatment of issues such as do with the environment (extinction of species etc.) and health (value of life) (Forfás, 2003: 22). The formulation of net present value by Schofield (1989a) has been interpreted as involving ‘double discounting’ if both shadow pricing and a rate of discount are applied to future benefits (Hill and Williams, 2005) (10.3.2). Many distributional and ethical issues are unresolved, while measurement problems are pervasive and, as Little and Mirless (1994) note, there are many implementation problems, especially in developing countries. Nevertheless, the use of cost-benefit analysis has been justified because “it provides a consistent ranking of measured benefits and costs of projects, and no better alternative is available.” (Forfás, 2003: 22).

Disadvantages there surely are; but these can be regarded as limitations on scope rather than fundamental weaknesses. Paying due account to non-money quantities, the needs of future generations or other jurisdictions may limit reliance of the technique but does not rule out its use within defined parameters where like is compared with like. This restriction, however, means it is necessary to regard cost-benefit analysis (and the model based on it in this thesis) as a specific technique to be used within an overall framework that pays due regard to the many issues that confront agencies.
Part 5
Parameters for a New Approach to Evaluating Regional Policy Projects

Chapter 11: Processes of Appraisal and Evaluation

11.1 Introduction
11.2 Setting Objectives
11.3 Defining Techniques
11.4 Assessing Impacts
11.5 Outputs and Outcomes
11.6 Selection of Indicators
11.7 Multi-faceted Programmes
11.8 Setting Decision Rules
11.9 Evaluative Framework
11.10 Appraisal Process
11.11 Determining Additionality
11.12 Technical Problems
11.13 Shadow Pricing
11.14 Conclusions

11.1 Introduction

Chapter 10 set out the rationale for using a cost-benefit model for appraisal, monitoring and evaluation of regional economic development projects and programmes. This chapter considers the practical steps required for appraisal and evaluation recommended in the UK. This sets the scene for Chapter 12, which looks in greater depth at theoretical issues involved.
11.2 Setting Objectives

The process of impact assessment should begin by asking the question 'impact upon what?' – in other words, fixing the purpose or purposes of the intervention. Underlying this treatment is the need to have agreed, clear objectives, so that the system adopted assesses what difference has been made by the policy or project studied. Thus Pickernell (1999: 29) argues the case for the testing of regional policies to focus on defined objectives, such as growth of long term GDP per capita. UK policy makers identify three broad intervention categories, through which agencies can evaluate their impact – business development and competitiveness, regeneration (through physical infrastructure) and human resource and community development (DTI, 2006: xi).

Typically, however, agencies will be set a wide range of different objectives, some of which may conflict with each other. For example, it may be easier to achieve a headline target of raising Gross Value Added in Wales (at least in the short run) by concentrating effort on areas already favoured in terms of communications, infrastructure and skills. Such a concentration might however make it more difficult to achieve other objectives (e.g. raising employment activity, promoting equality among different social groups, or spreading prosperity more evenly). Politicians may prefer simply to set out desirable targets and leave it to agencies to resolve any contradictions and face criticism, a process which has been argued may prove more complex in Wales following absorption of the WDA by the Welsh Assembly Government in April 2006 (Morgan, 2004).

Apart from such conflict, some objectives may be difficult to measure within an acceptable period of time. Thus McVittie (2005: 8) concludes that the use of growth in real GVA per head as a policy target is highly questionable, since GVA cannot be measured with sufficient accuracy to identify improvements in growth rates with precision over reasonable time scales.

Overall objectives, as discussed later in this chapter, are often translated into targets, which are intended to focus activity and provide appropriate incentives, especially where policy delivery has been delegated to external agencies (McVittie, 2005: 13-14). Achievement of targets, however, does not automatically mean that costs outweigh benefits: even if targets are set
optimally, meeting the target is not a good indicator of whether the project has been effective in cost benefit terms.

11.3 Defining techniques

Once objectives are specified as precisely as possible, evaluators need to define the techniques best able to measure the direct consequences of actions taken to achieve them. Storey (2000: 6) maintains that evaluation of the estimated or eventual outcomes depends on measuring the difference between (i) Observed results, and (ii) What would have happened in the absence of the project (the ‘counterfactual’), discussed in detail above (2.7).

However, policies or projects may vary widely in nature, scope and time-scale; and this has clear consequences for the method of evaluation adopted. Of particular reference to this thesis is the apparent clash of cultures between the traditional economic opportunity cost approach to evaluation and the more recent social capital paradigm that characterises community regeneration and community development (Armstrong, 2001). This rift plays a prominent role in the differences that appear between the treatment of more readily measurable ‘value-adding’ programmes and those capacity-building initiatives whose effects are less tangible and take longer to appear (8.4.1).

A project’s additional impact may be assessed by using two alternative approaches:

- Top-down (the outcome indicator approach), by estimating changes in overall indicators (e.g. employment, total population or number of dwellings); and/or
- Bottom-up, appraising the expected impact of individual actions or projects, through considering their likely outputs and outcomes (English Partnerships, 2004: 4).

The choice of technique may depend upon scale, with most projects appraised using the bottom-up approach. In the case of smaller projects, it will often not be feasible to undertake any detailed assessment, and appraisal may be informed by a qualitative analysis, as well as programme and policy level assessments (English Partnerships, 2004: 1-2). Standard projects of a type that have been delivered before will typically require less original analysis and can more easily draw on
appraisals and evaluations of previous projects. Similarly projects that are novel, contentious, repercussive or involve a high level of risk will require more in-depth analysis, as will larger projects that involve a significant amount of public expenditure (English Partnerships, 2004: 1).

11.4 Assessing Impacts

The objectives of projects are often presented as ‘outcomes’ and ‘outputs’. Outcomes are defined as the intended results of the project in terms of its key or ultimate objectives, such as sustainable increases in economic activity; while outputs are the things immediately produced by a project, such as land reclaimed or number of trainees (English Partnerships, 2004: 1n). The efficiency of a project may be judged by analysing the relationship between project financial inputs, and net outputs in terms of activities, results and impacts; while its effectiveness is gauged by comparing outputs with what was planned (Munday and Williams, 2006: 2). UK practitioners are advised that the objectives of projects are more attuned to outputs, whereas for programmes the main objectives are outcomes (DTI, 2006: 14).

The objectives adopted often rely on a rationale of market failure, with the case for policy actions depending on absence of information, knowledge, competition or affordable premises (Hill, 2002: 8). As argued above however (6.5), arguments for regional policy in fact often imply achievement of social equity rather than dealing with market failure.

In the case of intervention by a government or a regional development agency, the real benefit is likely to be reflected indirectly, in the greater success of businesses, communities or individual people assisted, a factor that considerably complicates calculation of the true impact. Evaluation therefore needs to measure the true benefit by gauging the difference made by that intervention, and that requires full consideration of the ‘counterfactual’ – what would have happened in the absence of the intervention. Hill (2002a: 3) sets out one approach to resolving the practical problems posed by the indirect and counterfactual nature of the ‘impacts problem’, by drawing a clear distinction between the three layers of evaluation detailed in Table 11.4.1.
Table 11.4.1: The Impacts Problem

<table>
<thead>
<tr>
<th>Level of Evaluation</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>What we do</td>
</tr>
<tr>
<td>Results</td>
<td>Effects on direct recipients</td>
</tr>
<tr>
<td>Impacts</td>
<td>What difference will (or did) the intervention make</td>
</tr>
</tbody>
</table>

Source: Hill (2002a: 3)

Take the case of an RDA encouraging, say, the take-up of broadband technology by small businesses by staging a series of seminars. In this instance, running the seminars would constitute the first level ‘activity’. Evaluation would readily measure the second level ‘results’ of that intervention, often termed ‘intermediate outputs’ (Begg at al., 1998: 421) – in this example, perhaps a record of those attending the seminars and receiving the information on offer. The most significant measurement, however – and by far the most difficult to attempt – is that of the difference the seminars actually made in terms of a change in some defined impact measure, such as trade, jobs or GDP, which constitutes the real objective of taking the action. This brings the evaluator face to face with the problems of how to make complex, counterfactual estimates:

“In an ideal world, trade would be measured before and after the seminars, with nothing else changing in between. In practice the world keeps on turning, making the analysis of cause and effect difficult.” (Hill, 2002a: 3)

In the real world, the impact of specific policy measures may be overwhelmed by changes in the macroeconomic environment (such as a collapse in the value of the currency) or external shocks (for example, an oil-price hike).

In practice, evaluation may well stop short at the first stage of the process. For example, an examination of best practices and lesson learned from community-based ICT programmes across Canada concludes:

“Most community based networks manage projects with a focus on lower level project outputs and outcomes in terms of service provision, and do not manage with a focus on higher level social and economic development results. Most of the organisations
contacted track network use, membership data, and event- and project-specific outcomes. They recognise the value of this data to demonstrate their relevance to funding agencies and clients/members. There are few examples of client/member involvement in developing performance measurement approaches or defining the indicators that will be measured.” (Industry Canada, 2002)

However, it is clear that estimation of the third stage – the impacts – may involve two distinct processes of estimation:

- Change in the chosen indicator revealed by recipients
- Overall net change in the chosen indicator, allowing for (positive) multiplier effects, and (negative) displacement and deadweight effects (11.11.1).

This would suggest elaborating the evaluation process to the four-stage model suggested in Table 11.4.2.

**Table 11.4.2: The Impacts Problem (Modified)**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Level of Evaluation</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Activity</td>
<td>What we do</td>
</tr>
<tr>
<td>2.</td>
<td>Results</td>
<td>Effects on direct recipients</td>
</tr>
<tr>
<td>3.</td>
<td>Impacts (specific)</td>
<td>What difference will (or did) the intervention make to recipients</td>
</tr>
<tr>
<td>4.</td>
<td>Impacts (overall)</td>
<td>What difference will (or did) the intervention make, allowing for (positive) multiplier effects, and (negative) displacement and deadweight effects</td>
</tr>
</tbody>
</table>

Based on Hill (2002a: 3)

While the estimates required by Stage 3 could be achieved by direct interrogation of recipients of the intervention, Stage 4 presents a more challenging problem, involving a move “beyond the counting and costing of activities to the careful assessment of the consequences of these activities, with multiple activities interacting within an uncertain and complex social, economic and cultural environment, leading to continual reformulation of policy” (Hill, 2002a: 10). Such (Stage 4) allowance for deadweight, displacement and multiplier effects is made in the updated
While the framework suggested above allows for a wide range of effects, policies may also exert a long-term impact beyond the Stage 4 elements – for example, if a policy promoting links between business and higher education succeeds in raising local levels of education and skills. The difficulties in estimating the contribution of such ‘externalities’ are recognised in the appraisal model recommended in the Irish Republic (Forfás, 2003: 43; Barry et al., 2002: 46), and the model proposed in this thesis may also be adapted to favour high-value projects by varying the level set for the shadow wage (14.3.1.7).

The more sophisticated measures required for net job estimates also raise a number of practical problems – who should carry them out, when they should be made (pre- or post-grant offer) and how they should be calculated. This suggests the need for an independent system, possibly where past performance measures could generate guidance for future estimates (Pickernell, 1999: 18).

11.5 Outputs and Outcomes

Official UK guidelines distinguish between objectives, outcomes and outputs, as well as targets (HM Treasury, 2003: 13). They emphasise the need for clear specification of policy objectives so that it is plain what proposals are intended to be achieved, and maintain that the ultimate objectives are ‘outcomes’ which outputs help achieve:

“There is usually a hierarchy of outcomes, outputs, and targets that should be clearly set out in an appraisal. Outcomes are the eventual benefits to society that proposals are intended to achieve. Often, objectives will be expressed in terms of the outcomes that are desired. But outcomes sometimes cannot be directly measured, in which case it will often be appropriate to specify outputs as intermediate steps along the way. Outputs are the results of activities that can be clearly stated or measured and which relate in some way to the outcomes desired.” (HM Treasury, 2003: 13)
Targets can be set to check progress in terms of producing outputs, delivering outcomes, and meeting objectives. According to the UK guidelines, such targets should be ‘SMART’ (Specific, Measurable, Achievable, Relevant and Time-bound) (HM Treasury, 2003: 13; DfES, 2004: 5). An example of the hierarchy of outcomes, outputs, and targets, closely based on official UK guidelines, is given in Table 11.5. It should be recalled that agencies are likely to be set wider long-term objectives, so that summing the net outputs of projects will provide only “one rather narrow window on impact” (DTI, 2006: 103).

Table 11.5 Hierarchy of outcomes, outputs, and targets

<table>
<thead>
<tr>
<th>Overall policy objective</th>
<th>Outcomes</th>
<th>Outputs</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘To address the major skills deficiency in the UK by increasing training to be reflected in the numbers of people holding vocational qualifications’.</td>
<td>A socially optimal level of training</td>
<td>Human capital as a share of GDP</td>
<td>The number of training places that will be provided by a certain date</td>
</tr>
<tr>
<td></td>
<td>Higher productivity for both trainees and co-workers</td>
<td>Proportion of workforce with vocational training</td>
<td>Reduction in the percentage drop-out rate by a certain date</td>
</tr>
</tbody>
</table>


11.6 Selection of indicators

Thorough evaluation of regional development policies requires measurement of net impacts at both micro- and macro- levels, using surveys at firm level as well as models of the economy as a whole. This process requires the selection of indicators that will reflect the chosen objectives. In practice, it may not be possible to observe the selected indicator directly, usually because of a lack of dependable statistical data within the chosen time-period. This may require selection of a proxy indicator, capable of readier measurement. Thus Hill (2000a: 5) suggests that the activities of the WDA may be evaluated by an evaluation framework on four layers:
Objective: Increase prosperity

Measured by: Increased GDP

Indicators: Wage levels, wage bills, profits, value-added per job

Proxyed by: Turnover, turnover per job.

This approach can be applied to individual programmes or groups of programmes, as well as to the activities of the agency as a whole, with the goal of constructing an impact assessment framework capable of measuring the real cost-benefit impacts; i.e. it should enable comparison of the real value of different programmes. Since most appraisals compare costs and benefits which occur in different time periods, this requires conversion to “present values” or discounting, which is accomplished by multiplying costs and benefits by a discount factor (DiES, 2004: 8), the nature of which was discussed in 10.2.1.

The European Commission (2000) recommend that indicators are separated into:

- Output indicators: i.e. those indicators relating to activity and measured in either physical or financial units;
- Result indicators relating to the effects of the programme (e.g. changes in numbers of jobs created or businesses assisted); and
- Impact indicators relating to the net effects of the programme (Munday and Williams, 2006: 3).

Eurada, the association of around 150 development agencies in Europe, suggests “it would be judicious to evaluate the effectiveness of public sector intervention in support of the economy in terms of the tax revenue it generates” (Eurada, 2005: 1). Such an analysis could be performed at the level of both strategic programmes and individual measures.

As discussed above (11.4), the evaluation process requires comparison of at least two different sets of impacts – one resulting from the activities of the project under investigation; the other that would come about in the absence of the project, the ‘counterfactual’ (2.7). The extent of the impacts is revealed by ‘indicators’, and the choice of the indicators depends on the objectives of the policy. Practitioners may choose either a large number of indicators (in which case, it is
probably easier to justify the policy) or a minimal number (meaning zero weight for those left out) (Hill, 2001b: 1). In Northern Ireland, for example, the IDA’s chosen indicators included turnover, profit, employment, sales, and sales to the rest of the UK (IDB, 2001: 10). Wren and Storey (2002: 335) chose as indicators for evaluation of a marketing project (i) sales turnover, (ii) employment, and (iii) survival of the business, a critical issue in small firm policy.

This choice of multiple indicators reflects the existence of more than one presumed policy objective – for example, policy makers could seek both increased employment and a rise in the number of small businesses. Gambardella et al. (2001) estimate the determinants of labour productivity in European NUTS regions between during 1989 and 1996 by comparing three potential explanations of regional advantages: technological capabilities (proxied by regional patents), agglomeration economies (proxied by employment density), and openness (proxied by the number of air passengers embarked and disembarked in the region); and found that openness had a positive effect for labour productivity as well.

Some indicators may well be chosen for convenience, such as the fact they become available sooner than others. For example, Hill (2002: 11) notes that while aggregate regional data on employment, wages and value-added are normally available only after some time lag, and therefore prove poor indicators in a responsive learning framework, “business turnover or sales revenue may be a useful proxy, with a significant relationship to these other indicators”. So even raw data on business turnover, preferably combined with information on employment and wages (to assess productivity) can provide a useful and flexible indicator for likely impacts on prosperity.

From the standpoint of simplicity, other things being equal, the fewer the number of indicators, the better. Thus the UK Government guidelines published in 2005 reduced the number of performance indicators for local authorities in England to 94, “less than half the number originally used in 2000/01” (ODPM, 2005: 4). Indicators for English RDAs are set out by DTI (2006: 141-143).
11.6.1 Employment as indicator

This discussion has a number of implications for the conduct of evaluative techniques and the choice of indicators involved. As discussed above (6.6.1), the traditional focus of UK regional policy has been employment. ‘Jobs’ have the advantage of being relatively easy to measure; but they are not homogenous and vary widely in quality and in wage. It may be possible to use a single indicator of ‘jobs’ to measure unskilled employment in assembly plants of the sort attracted by regional policy in the 1980s, but which are now increasingly attracted to lower cost locations. Such a measure is of less use in assessing the contribution to the economy of fast growth, high technology and high value enterprises. Adopting the concept of ‘job years’ as an indicator may help to get over some of the difficulty: jobs likely to last a decade or so may well involve greater skill and be more competitive.

Within the broad category of jobs, a distinction needs to be drawn between jobs that are created and jobs that are safeguarded. The two concepts have been a recognised goal of UK policy for decades. However, concerns are frequently heard about the validity of jobs safeguarded as an indicator: the ‘heroic assumption’ underlying the concept of jobs safeguarded is that firms’ jobs would otherwise vanish if not supported by public policy. If the assumption is erroneous, this could mean an overestimate of the total number of jobs safeguarded. A particular issue is raised by the tendency to add up the total jobs safeguarded over a long period: there is nothing to prevent the same job being safeguarded several times during its lifetime. Again, the use of ‘job years’ would be a more accurate reflection of the real output gained. The concept of jobs safeguarded has been treated as a legitimate objective of the UK’s Regional Selective Assistance grant system: jobs created have been treated as Category A, while jobs safeguarded were listed as Category B and subject to less favourable conditions (Wren, 2003: 7). Generally, it is assumed that jobs have a lifetime equal to that of the value of the asset provided by the project, and that this lifetime is on average about ten years (Wren, 2003: 16 and 16n). The Regional Selective Assistance programme is thought to bring forward through time about a quarter of assisted projects (Wren, 2003: 16).
11.6.2 Other indicators: Gross Value Added, wages, turnover

Arriving at a more accurate estimate of the contribution jobs attributed to a project make to the economy could be attempted by qualifying the number of jobs (or job years) by the amount of value they generate. One way of doing this would be by recording their economic sector, and using the mean or median measure of Gross Value Added generated. Clearly, the more detailed the sectors, the more accurate the qualification can be. Information on the contribution to Gross Value Added (GVA) is provided for 74 sectors by input-output tables for Wales (Bryan et al. 2004: 6).

One alternative to GVA as an indicator of benefit may be provided by wages. Wages account for a significant proportion of value added – about 63 per cent in the year 2000 according to the input-output tables for Wales (Bryan et al. 2004: 15). Qualifying the number of jobs by wages may give a better guide than qualification by industrial sector, although jobs in manufacturing are likely to be more capital intensive and hence costlier in real economic terms. Alternatively the wage bill itself could be used as an indicator. In both cases, however, it would be necessary to take account of profit, another major component of GVA.

A further alternative to GVA is provided by company turnover. Information on turnover is more easily obtained from companies than information on their value added. Information on the ratio of value added to turnover for each industrial sector is found in the input-output tables for Wales: for the Welsh economy as a whole, the estimated ratio in the year 2000 was 22.6 per cent (Bryan et al. 2004: 45).

11.7 Multi-faceted Programmes

One practical problem confronting evaluators is how to disentangle the different elements of support that may be contained within a single programme. For example, the WDA subsidiary Finance Wales provided a wide range of funding support measures, all of which were accompanied by management advice. These management support programmes are said to be “a
vital element of our work, which ensures our funding is put to the best possible use, reducing the level of risk both for the businesses assisted and for Finance Wales itself" (WDA, 2001a).

This in no way reduces the need for evaluation, but it points to the possible complexity of seeking to compare the relative effectiveness of, say, ‘marketing’ support compared with ‘financial’ support, where both elements are combined in a single measure.

11.8 Setting Decision Rules

Regional development agencies and other agents who support policies or projects are usually faced with competing demands for a limited resource of money and staff time. The process of deciding which initiatives to favour can be likened to the theoretical device of maximising a targeted benefit subject to the limit imposed by a fixed budget or limit on capacity (Hill, 1989a: 15-18; 39-40). This can then be used to construct a decision rule, sometimes described as an ‘allocation rule’ (or set of decision rules).

A decision rule is defined either as (1) a function that maps from the current state to the agent’s decision or choice or (2) a mapping from the expressed preferences of each of a group of agents to a group decision (Econterms, 2006a). The first is more relevant to decision theory and dynamic optimisation; the second is relevant to game theory. Intuitively, both forms could play a part in practical decision taking by RDAs: the first could guide decisions aimed to raise prosperity throughout the target territory (such as the Business Eye advisory service accessible throughout Wales) while the second would be more applicable to policies that support e.g. low income groups. One example of a decision rule is the suggested guideline developed by Swales (1997a) to estimate the maximum level of support that should be paid to sustain employment over a ten-year period (Pickernell, 1999: 20) (12.5).

One problem in applying such a system is that essentially it provides a short term measure of the fiscal neutrality of an individual scheme (i.e. do its benefits outweigh costs?), rather than comparing schemes in terms of their long term effects on the economy (Pickernell, 1999: 20).
In the real world, the behaviour of regional development agencies is likely to be much more complex than the simple maximization of net benefit subject to a budget constraint. For example, in the case of the Welsh Development Agency prior to merger with the devolved Welsh Assembly Government administration in April 2006, a number of specific targets and instructions would be issued by the Assembly Government Minister in the form of an annual remit letter. These constituted the basis of the Agency's Business Plan and formal reporting on progress in the form of the WDA's Annual Report and six-monthly reports (WDA, 2004a). In addition, its system of internal reporting included a management information system (PerMis) tracking the performance of individual projects and programmes (Williams, 2005).

The Agency's Board and management team would wish to meet targets on all formal objectives, although some of these might have greater psychological importance than others. During the 1990s, for example, foreign direct investment projects had a particularly high profile, whereas a decade later emphasis was laid on indigenous business and encouraging more people to set up their own companies (House of Commons, 2005: clause 34). The Agency would seek to control the level of risk, both financial and political, for example by avoiding unduly risky or controversial property schemes. It would also be aware that achieving results well over target might result in subsequent targets being raised. It would certainly seek to keep within the budget laid down, although some budget items e.g. management running cost, might well be regarded as being especially sensitive.

Given this highly complex process of decision taking and monitoring, 'mechanical' appraisal and evaluation devices are likely to play only a supporting role in affecting allocation of resources. Nevertheless, agencies operating an extensive mix of programmes need some way of mapping their impacts, and this suggests the need not only for a comprehensive business plan but also for a system that tracks the progress of a large number of indicators.

11.9 Evaluative Framework

One method of keeping track of performance on a broad front is by adopting a framework of impact assessment to measure and guide the impact of current activities and plan future actions.
Such an evaluative framework would enable practitioners to assess the impact of a particular programme as it unfolds. Hill (2001a: 2) suggests that development agencies require a management information system to assess the consequences of their activities, both direct and indirect. This would be in addition to the normal practice of evaluation within the public sector as an instrument of ensuring proper accounting for expenditure of resources, and suggests the need for a clearly understood evaluative framework that encourages policy change to reflect evidence. In the case of regional development policies, Hill (2002) suggests that such a framework should rest on eight key principles, as set out in Table 11.9.

### Table 11.9 Principles of an Evaluative Framework for Regional Development

<table>
<thead>
<tr>
<th>Principles of Evaluative Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Evaluative process to be open, transparent and simple</td>
</tr>
<tr>
<td>2. Assessment should look forward as well as back</td>
</tr>
<tr>
<td>3. Process to distinguish between activities, results and impacts or outcomes</td>
</tr>
<tr>
<td>4. Impacts to be assessed in terms of objectives</td>
</tr>
<tr>
<td>5. Framework to allow comparison of different actions and policies</td>
</tr>
<tr>
<td>6. Framework to compare performance with, and without, intervention (or institution)</td>
</tr>
<tr>
<td>7. Framework to adopt the minimum indicator set that is compatible with measuring performance against objectives</td>
</tr>
<tr>
<td>8. Framework to drive actions in pursuit of priorities</td>
</tr>
</tbody>
</table>

Source: Hill (2002:10)

These eight principles restate a number of conclusions already discussed:

- Principles 1 and 5 are necessary for the evaluative process to produce tangible measures useful for comparison.

- Principle 2 asserts the need for ex ante evaluation.

- Principle 3 involves distinguishing between gross and net effects, and involves the need to estimate the magnitude of variables such as the regional multiplier, net grant deadweight, displacement and substitution.

- Principle 4 reiterates the case for clear objectives, capable of measurement by target variables or proxies; but it also points up the fact that 'learning' may well lead to reassessment of the programme theory or the objectives or both.
• Principal 5 underlines the major potential to learn from evaluation and switch resources to more beneficial actions.

• Principle 6 emphasises the role of the counterfactual in evaluation.

• Principle 7 recognises the practical problems of obtaining data and the need for techniques to be transparent and simple to use.

• Principle 8 makes the case for the roles of learning by evaluation within a policy that is inevitably determined politically.

One study of such an approach by Scottish Enterprise is described by Jackson (1998) and outlined below (13.2.6).

The importance of sustainable development enhances the need for a framework approach to evaluation to accommodate the complex issues involved (10.3.1), as emphasised by DTI (2006: 66) in its elaboration of the Sustainable Quality Management process.

11.10 Appraisal Process

A number of specific steps are recommended as part of the appraisal and evaluation process, although UK Treasury guidelines, for example, stress that appraisals are often iterated a number of times before their proposals are implemented in full, and the various stages may be repeated and taken out of sequence:

“In particular, as options are developed, it will usually be important to review more than once the impact of risks, uncertainties and inherent biases. This helps to avoid spurious accuracy, and to provide a reasonable understanding of whether, in the light of changing circumstances, the proposal is likely to remain good value for money. As the stages of an assessment progress, data must be refined to become more specific and accurate.” (HM Treasury, 2003: 4).
The steps required are:

- Justify the action – to establish (a) that there is a clearly identified need and (b) that any proposed intervention is likely to be worth the cost;
- Set objective – by specifying the goals of the intervention in order to identify the full range of options that may be available to deliver them and creating associated targets;
- Appraise options – (a) creating a wide range of options to set parameters for a solution, and select a shortlist that includes a ‘do minimum’ option to act as a check against more interventionist action and (b) establishing a Base Case to provide the best estimate of costs and benefits, which can be adjusted by considering different scenarios or sensitivity to changes in key variables.

Figure 11.10 illustrates how one UK government department has incorporated this process into its own appraisal procedure (DfES, 2004).

**Figure 11.10 Department for Education and Skills Appraisal Checklist**

1. Complete or Review Policy Rationale and Objectives Statement
2. Identify and Value Costs and Benefits of Each Policy Option
3. Complete NPV and IRR Calculations
4. Make Adjustments for Distribution of Costs and Benefits
5. Analyse Potential impact of key risks and uncertainties
6. Present the Appraisal and Select the Best Option
11.11 Determining Additionality

As concluded above (11.4), accurate appraisal involves forecasting the real difference that will be made by the project – that is distinguishing between its gross and net effects. This difference is described as ‘additionality’. Taylor (2002: 191) suggests that additionality may be regarded as the probability that the project takes place as a direct consequence of the assistance provided, so that greater non-additionality reduces benefits.

The precise definitions in the literature of some of these effects (especially additionality and deadweight) vary (Hill and Williams, 2005). This issue is discussed in this subsection and clear specifications provided as part of the outline of concepts used in the model (14.3). In the context of regeneration, ODPM (2006: 5) states:

“An impact arising from an intervention is additional if it would not have occurred in the absence of the intervention. It is the extent to which activity which takes place at all, on a larger scale, earlier or within a specific designated area or target group as a result of the intervention.”

However, the concept of additionality requires clarification, not least because it may be used as an operational element in some evaluative models, such as that proposed by Swales (1997b: 859). The estimation of additionality depends on a number of other concepts, including grant deadweight, displacement and the multiplier. A detailed treatment of the argument for consistent definition of additionality, deadweight and displacement is provided by Hill and Williams (2005). This thesis follows the proposal made by Begg et al. (1998: 413), who define ‘gross additionality’ as the output of the initiative that would not otherwise have been there in the absence of the policy:

“Gross additionality is then translated into net additionality to allow for displacement effects which occur within local economic systems. The net additional activity is then usually grossed up to reflect local income and linkage multiplier effects. The result is an estimate of the net additional output which has arisen as a result of the policy operating in the local economy concerned.”
Additionality may also be referred to as a 'supply side' or 'structural' impact, which operates by altering the productive capacity of the economy, either because of a change in the size of the workforce or in the productivity of the workforce (ODPM, 2006: 1).

11.11.1 Net Grant Deadweight

This thesis follows the suggestion of Hill and Williams (2005: 3) that 'grant deadweight' reflects the extent to which companies or projects assisted would in any event have developed without grant aid, or with lower grant aid. Grant deadweight should be clearly distinguished from 'tax deadweight', the cost of distortions imposed on the economy by higher taxes required to fund regional policy, qualified by the additional tax revenue resulting from industrial projects (Forfás, 2003: 51), discussed in greater detail below (13.5).

Empirical studies distinguish between full (or ‘pure’) grant deadweight (where the grant makes no difference to the number or timing of jobs created) and partial deadweight, where the project takes place sooner or on an increased scale (Armstrong and Taylor, 2000: 251; Lenihan, 2002: 15). Thus a project intended to create 200 jobs and costing £10m in public money would have an additionality rating of 50 per cent if 100 of those jobs would have been created without any public money being provided. Honahan (1998: 31) outlines a simple model of the relationship between grants and employment linked to a Nash bargaining framework.

The main problem involved in estimating the grant deadweight effect of assisting a project is calculating what would have happened if it had not been assisted – the counterfactual (Forfás, 2003: 27-29), especially given an extensive stream of policy initiatives funded by various agencies (Munday and Williams, 2006: 5). Three approaches are generally followed to identify deadweight levels (Forfás, 2003: 27-29):

- Interview techniques, which suffer from the drawback that executives have no particular incentive to reveal the truth;
- Econometric modelling, which require substantial data and a good deal of variation in the level of assistance provided to pin down likely deadweight effects of various levels and types of assistance; and
Control/comparison groups, which compare differences over time in the performance of assisted firms in one area with unassisted firms in another, regarded as the most convincing.

The evidence reviewed shows a considerable variation in deadweight estimates. A comparison of small to medium-sized firms in Northern Ireland compared with (non-assisted) firms in the North of England suggested a grant deadweight level of about 50 per cent (Hart and Gudgin, 1999), comparable to the 48 per cent for UK Regional Selective Assistance (Pickernell, 1999: 16) but considerably lower than the 80 per cent suggested by Honahan (1998) or Lenihan (1999). A further study by Lenihan et al. (2004) estimated deadweight of 24.4 per cent for “firms with a net increase / static employment and partial deadweight only”. English Partnerships (2004: 12-14) cite evidence of grant deadweight for regeneration activity ranging from 25 per cent (offices), 35 per cent (industrial property), 30 per cent (recreational) and 40 per cent (retailing).

However DETR (2000) show that survey-based estimates vary widely according to who is interviewed. Thus programme and project managers responsible for business support estimate deadweight at 15 per cent, while beneficiaries set it at 36 per cent. This variation appears to confirm the judgement of (Forfás, 2003) above.

The revision of the grant deadweight parameter (0) in the appraisal system for the Republic of Ireland retained the 80 per cent level for Greater Dublin, but made variations for region and according to whether the assisted firms were start-ups or expansions (with slightly higher values for the latter) (Forfás, 2003: 54).

Further elaborations of the term ‘deadweight’ include project-level deadweight (the extent to which funding available to projects would have been available even in the absence of the programme); and secondly, deadweight at the level of the individual final SME recipient: in other words, whether the assistance received itself stimulated the creation of the business or job or whether it would have occurred in the absence of the project’s support (Munday and Williams, 2006: 18).
11.11.2  Displacement

Displacement may be defined as "The proportion of project outputs/outcomes accounted for by reduced outputs/outcomes elsewhere in the target area" (English Partnerships, 2004: 19). Armstrong and Taylor (2000: 251) describe displacement as taking place where a subsidy results in new jobs in the recipient firm, but at the cost of jobs in other firms in the assisted area. Displacement may arise where private sector business support (e.g. by banks) is replaced by publicly funded support, and may take place in product and factor markets (English Partnerships, 2004: 7). However, displacement is difficult to measure. Munday and Williams (2006: 5-6) conclude:

“Ultimately, it is only ever possible to gain an approximation of elements of displacement, and this largely in the context of firms as opposed to other agents.”

It is quite possible that an economic policy designed to promote employment and output in one region might result in displacement in another, so that definition of the target area or group is crucial for estimating the extent of displacement (Hill and Williams, 2005: 4).

The likelihood of displacement within assisted areas plays a significant role in appraisal. Thus EU Structural Funds tend to exclude retail as a sector eligible for business development support on the basis that displacement effects would be high, although a number of mid-term evaluations of the 2000-06 programmes identified the need for intervention in the retail sector (ECOTEC, 2003: 29). Lenihan et al. (2004: 26) estimate a 4.4% level of displacement for high tech firms in Ireland, but most industrial sectors might be expected to have higher levels of displacement, reflecting greater local competition in goods and factor markets.

The appraisal model used in the Irish Republic uses a single measure for deadweight and displacement, described (confusingly) as ‘grant deadweight’ (Forfás, 2003: 33). Similarly, a single estimate for both project deadweight and displacement is made by the Welsh Development Agency (Williams, 2005: 10). This model does not follow this procedure. While Hill and Williams (2005) suggest that both ‘grant deadweight’ and displacement may be viewed as
aspects of the same theoretical welfare loss caused by monopoly, taxes, tariffs or other distortions (Samuelson and Nordhaus, 2001: 761), their amalgamation in this way risks confusion.

Deadweight and displacement are more usually treated as distinct concepts (Armstrong and Taylor, 2000: 251; English Partnerships, 2004: 39; ODPM, 2005), and have been estimated empirically by surveys that involving different groups of firms: these studies provide some insight into the way the two elements operate and the magnitudes that should be attributed to them (Lenihan, 1999; 2002). Deadweight and displacement are also specified independently in the model proposed by Swales (1997b) and by most UK studies. For the avoidance of doubt, therefore, grant deadweight not including displacement is described throughout this thesis as ‘net grant deadweight’.

11.11.3 Substitution

The UK Government also distinguishes ‘substitution’ impacts, which may be regarded as internal displacement – “the situation in which a firm substitutes one activity for a similar activity (such as recruiting a different job applicant) to take advantage of government assistance” (HM Treasury, 2003: 105). English Partnerships (2004: 21) suggest that substitution is a very specific form of non-additionality that has in the past been largely subsumed within the displacement effect and as a result not considered sufficiently.

11.11.4 Multiplier Effect

The multiplier is defined as the ratio of the final change in national income resulting from an autonomous change in expenditure, to that change in national income (Livesey, 1982). In the context of regional economic development, the multiplier effect takes the form of further economic activity (jobs, expenditure or income) associated with additional local income and local supplier purchases (English Partnerships, 2004: 22).

The impact of a regional project will be positively correlated with the regional multiplier effect. The appraisal process therefore adjusts impacts upward to take account of this multiplier effect,
which will depend in turn on the regional propensity to consume locally produced goods and the rate of income tax (Armstrong and Taylor, 2000: 9). The precise magnitude of a locality’s multiplier will depend on its size (smaller regions are more likely to import goods and services), its industry mix (specialised industries will import more) and location in relation to other labour markets and retail centres. For these reasons, multipliers are project-specific as well as region-specific (Armstrong and Taylor, 2000: 9-12).

The multiplier effect is a process that includes (i) backward linkages to intermediate suppliers and (ii) consumption multiplier effects. An increase in final demand for a particular commodity will tend to produce an increase in the output of that commodity, as producers react to meet the increased demand; this is known as the direct impact. This will lead to an increase in demand on their suppliers and further down the supply chain (the indirect impact). As a result of these direct and indirect impacts, the level of household income throughout the economy will rise as a result of increased employment, and a proportion of this increased income will be re-spent on final goods and services: this is known as the induced effect (Scottish Executive, 2005).

A distinction is usually made between different types of multiplier effect: Type I multipliers sum together direct and indirect impacts, while Type II multipliers also include induced income effects, (Scottish Executive, 2005). Type III multipliers, a variant of Type II, adjusts for non-linear relationship between growth of income and consumption spending (Dumas, 2003: 3-4).

The extent of economic linkages, and therefore the multiplier effects, will be greater if a larger spatial area is being considered. As a result, when appraising a project the spatial level and target group within that area against which the project is being assessed must be clearly specified. They should relate directly to the identified need (the rationale for intervention) and be wide enough to take into account spill over or unintended effects on other groups, areas or markets. It is quite common in project appraisals to consider effects at different spatial levels. The four most often used geographical levels are: site, local/sub-regional, regional and national (English Partnerships, 2004: 8).
A number of studies have estimated the size of local income and supply linkage multipliers of regeneration initiatives (English Partnerships, 2004: 23). The magnitude of the multiplier effect in assessing the impact of policy instruments such as RSA has also been calculated by studying local purchasing patterns of assisted firms, and then estimating the additional employment needed to create this given output, sometimes with analysis developed from reference to the UK-wide Input-Output table framework (Munday and Williams, 2006: 6-7).

**Table 11.11 Assessing Net Additionality**

<table>
<thead>
<tr>
<th>Gross Attributable Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MINUS Net Grant Deadweight</td>
</tr>
<tr>
<td>MINUS Leakage from Target Area</td>
</tr>
<tr>
<td>↓ Gross Direct Outputs</td>
</tr>
<tr>
<td>MINUS Displacement</td>
</tr>
<tr>
<td>MINUS Substitution</td>
</tr>
<tr>
<td>↓ Net Direct Outputs</td>
</tr>
<tr>
<td>PLUS Multiplier Effect</td>
</tr>
<tr>
<td>↓ Net Outputs</td>
</tr>
<tr>
<td>PLUS Crowding-In Effects</td>
</tr>
<tr>
<td>MINUS Crowding-Out Effects</td>
</tr>
<tr>
<td>PLUS / MINUS Wider Effects on Sustainable Development</td>
</tr>
<tr>
<td>↓ Total Net Outcomes and Outputs</td>
</tr>
<tr>
<td>Based on DTI (2006: 75-81)</td>
</tr>
</tbody>
</table>

Having precisely defined its component elements, net additionality may now be fully defined as in Table 11.11. This scheme provides a conceptual framework for the model developed in this thesis (Chapter 14), although crowding in and out, and wider effects on sustainable development need to be assessed outside the model using an evaluative framework. An alternative method of arriving at net additionality is provided by English Partnerships (2004: 25).
11.12 Technical Problems

Once the decision is taken as to what is to be measured, across what period of time, and using what indicators to track the chosen target variables, evaluative techniques have to wrestle with a number of more mundane problems involved in measuring a complex and changing economy.

11.12.1 Selection bias

Any comparative evaluation has to avoid falling prey to ‘selection bias’ – failing to compare like with like. A familiar pitfall is that firms applying to take part in an agency-supported programme may well be more ambitious and better aware of opportunities available than the generality of firms in the region: if so, it would be misleading to ascribe their subsequent success to the impact of the programme being evaluated. Storey (2000: 189), Wren and Storey (2002: 335), Bondonio, (2002) and Schmitt et al. (2002) present econometric techniques for dealing with this problem; but again their true significance is bound to rest on a presumed underlying programme theory.

11.12.2 Optimism bias

Practitioners are counselled to make allowance for ‘optimism bias’, which is defined as “a demonstrated, systematic, tendency for project appraisers to be overly optimistic” (HM Treasury, 2003: 29). The recommendation by HM Treasury (2003: 5) is that practitioners in the UK should “adjust for risk and optimism to provide the Base Case”, providing the best estimate of the costs and benefits of each option considered, and consider the impacts of changes in key variables and of different future scenarios on the Base Case.

This changed previous Green Book guidance, in which Optimism Bias was taken into account in a generalised way through a percentage premium embodied in the test discount rate. This ‘unbundling’ by HM Treasury of Optimism Bias from the discount rate accounts for its reduction from 6 per cent to 3.5 per cent (Defra, 2003: 2).
However, for the reasons set out below (14.3.2.4), no adjustment is made within the model developed in this thesis.

11.12.3 Double counting

A familiar problem for evaluators is that different project managers may claim credit for the same results. Munday and Williams (2006: 25) find high levels of double counting (of SME recipients) and that projects were "not always careful to ensure strong attributional links between interventions and claimed jobs".

11.12.4 Community Economic Development Programmes

Estimating the gross and net impact of community economic development programmes is notoriously difficult. Community based projects are likely to produce greater displacement because of the smaller scale of projects and the intensity of local trading. At the same time multiplier effects are expected to be larger because community-based firms and activity may purchase more locally than larger organisations and firms (Munday and Williams, 2006: 8).

11.13 Shadow Pricing

One conspicuous absentee from this list of terms favoured by UK official literature is any reference to shadow pricing (10.2.3). Yet as argued in that section, shadow pricing is a necessary component of any cost-benefit system that seeks to measure real resource costs and benefits to the economy and is therefore an integral part of the model specified in Chapter 13.

11.14 Conclusions

Chapter 11 examined concepts involved in project appraisal and evaluation. This treatment involved a clear definition of objectives and identification of the ultimate outcomes sought by intervention. The effectiveness of policy can only be truly judged by the extent to which its
impact is additional to what would have transpired without intervention, a state of affairs conceptualised as the counterfactual.

The process also requires selection of indicators to measure progress on desired objectives, which may be represented by targets. Agencies’ behaviour may be portrayed as maximising target objectives subject to a budget constraint, although real world behaviour is likely to pose much greater complexity. On these foundations, an evaluative framework may be constructed.

Estimation of the ‘additionality’ of interventions requires calculation of grant deadweight, displacement and substitution, multiplier effects and leakage from the target area.

This summary laid out the ‘mechanics’, while Chapter 12 looks in greater detail at the theory that underpins evaluative tests and Chapter 13 considers how agencies in the UK and the Republic of Ireland have put them into practice.
Chapter 10 concluded that cost-benefit analysis presents the most promising basis on which to construct a system of appraisal and evaluation. This chapter examines various tests used in evaluative methods, both for ex ante appraisal and ex post evaluation. These are listed in section 9.2, which suggests a revised taxonomy of criteria used in the process of appraisal and evaluation. These are described under two broad headings – cost-effectiveness (9.3) and cost-benefit (9.4). Section 9.5 moves to a critical analysis, presenting the critique of UK evaluative methods developed by Swales (1997a; 1997b). Section 9.6 describes the alternative approach suggested, as well as the cost-benefit economic model used in the Republic of Ireland which together form the departure point for the model proposed in this thesis.

12.2 Taxonomy of Appraisal and Evaluation Criteria

This section examines criteria developed in the UK to appraise, monitor and evaluate projects and policies. It sets out a taxonomy by considering the simplest available criterion, and proceeds by introducing progressively greater refinements. The objective of this part of the exercise is to demonstrate the different inherent characteristics of the
criteria used, because their usefulness depends upon their theoretical and empirical 'construct validity', which measures the correspondence between a theoretical concept and a specific measuring device or procedure (Cronbach and Meehl, 1955).

This treatment includes consideration of the critique by Swales (1997a; 1997b) of the appraisal recommendations of the UK Treasury in the 1990s, now superseded; because the issues revealed are relevant for the model developed in Chapter 13. This chapter is concerned with theoretical evaluative techniques: more 'mechanical' processes were discussed in Chapter 11.

All methods of evaluation must involve weighing up estimated or actual benefits of regional economic policy against its costs. As argued above (9.3), the various approaches may be allocated to one or more of three broad categories:

- (Micro) cost-effectiveness tests, in which defined objectives (such as jobs created) are taken as given, and comparisons then made of the costs of achieving them;
- (Micro) cost-benefit tests, in which a more systematic effort is made to identify the benefits achieved, and produce a 'net' measure of the surplus of social benefit over cost (the model in this thesis belongs to this category); and
- Macro-evaluations, in which system-wide impacts are taken into account (including cost-benefit methods).

Within these categories, a large number of variations can be distinguished, depending on whether or not future costs and benefits are discounted, whether or not gross measures are refined to produce more meaningful net values, and whether supply-side influences (such as labour market effects, migration or inflation) are treated as dependent variables, or ignored. Further, minor variations involve the precise units chosen for evaluation, such as Present Value Job Years or Permanent Job Equivalents, which are described in 12.2.2 and 12.2.3 below. In practice, attention in the UK has focused primarily on just two outcomes: the total jobs created and the average cost (to the government) of each job created (Taylor, 2002: 190). The methods are presented in Table 12.2, identifying the various tests, their most significant characteristics and examples of their application. The
possible tests are listed in ascending order of complexity (i.e. the simplest first); with examples of their use or suggested use (not all tests have been implemented). It should be emphasised that these models simply offer differing ways of estimating the impacts of projects and programmes: they are different ways of obtaining estimates rather than separate concepts – thus costs and benefits may be assessed using Input-Output or Computable General Equilibrium models etc.

Table 12.2: Classification of Appraisal and Evaluative Methods

<table>
<thead>
<tr>
<th>Type</th>
<th>Method</th>
<th>Features</th>
<th>Example of use or recommendation of method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost effectiveness</td>
<td>Gross cost per direct (gross) job (GCPGJ)</td>
<td>Not discounted</td>
<td>Wren (1989)</td>
</tr>
<tr>
<td></td>
<td>Gross cost per direct (gross) job per year (GCPGJY)</td>
<td>Not discounted</td>
<td>Wren (2003)</td>
</tr>
<tr>
<td></td>
<td>Gross cost per direct (gross) job, using Permanent Job Equivalents (GCPGJPJE)</td>
<td>Not discounted</td>
<td>King (1990)</td>
</tr>
<tr>
<td></td>
<td>Gross cost per direct (gross) job (GCPGJ) (D)</td>
<td>Discounted</td>
<td>King (1990)</td>
</tr>
<tr>
<td></td>
<td>Gross cost per direct (gross) job per year (GCPGJY)</td>
<td>Discounted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gross cost per direct (gross) job, using Permanent Job Equivalents (GCPGJPJE)</td>
<td>Discounted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gross Cost per Net Job (GCPNJ)</td>
<td>Not discounted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gross Cost per Net Job (GCPNJ) (D)</td>
<td>Discounted</td>
<td>HM Treasury (1997) guidelines</td>
</tr>
<tr>
<td></td>
<td>Net Cost per Net Job (NCPNJ)</td>
<td>Not discounted</td>
<td></td>
</tr>
<tr>
<td>Cost-benefit analysis</td>
<td>Modified value-added</td>
<td>Discounted</td>
<td>WDA (Williams, 2005)</td>
</tr>
<tr>
<td></td>
<td>Net present value</td>
<td>Discounted</td>
<td>Swales (1997a, 1997b)</td>
</tr>
<tr>
<td></td>
<td>Net present value, with labour movement</td>
<td>Discounted</td>
<td>Forfás (2003)</td>
</tr>
<tr>
<td>Large-scale regional models</td>
<td>Input-output tables</td>
<td></td>
<td>Hill and Roberts (1996)</td>
</tr>
<tr>
<td></td>
<td>Computable general equilibrium models</td>
<td>Reflects migration and labour market adjustment</td>
<td>Gillespie et al. (2000) (9.3)</td>
</tr>
</tbody>
</table>
12.3 Cost-effectiveness indicators

A common feature of many of the simpler cost-effectiveness tests is that they tend to treat the benefit measured (usually jobs) as a single, undifferentiated item, reducing the decision rule to one of comparing the different costs of using a number of methods to secure a similar benefit. Cost-effectiveness tests may be undiscounted or discounted.

12.3.1 Gross cost per direct (gross) job (GCPGJ)

The simplest conceptual approach is to measure the gross cost of securing a gross benefit - with the gross benefit usually expressed in the form of direct (gross) jobs. This takes no account of how long the jobs created or safeguarded are likely to last, nor does it in all cases apply a rate of discount to future benefits or future costs.

Most evaluations of regional policy acknowledge the impact of time, both in estimating the value of future employment and the level of future costs. However, Wren (2003: 17) noted that “the discounting of future jobs is highly unusual in evaluation work”. There is a practical argument in favour of using undiscounted tests, where this will make possible like-for-like comparison. In practice, scope has been found for its use in comparisons of similar projects, where wider effects may be ignored: thus estimates of GCPGJ have been used to compare cost per job of different types of financial assistance and the effectiveness of regional financial incentives for different types of firm (Taylor, 2002: 192 and 192n).

A relatively crude, undiscounted GCPGJ measure was used in the evaluation of the employment effects of the principal regional policy instruments of the UK Government and local authorities in north-east England between 1978 and 1984 by Wren (1989). Gross cost is defined as average annual expenditure on public schemes in 1981 prices; the measure does not take account of tax flowbacks; while the definition of additional
employment is specified in terms of the number of jobs created, not the duration of employment (Wren, 1989: 499; 508).

Another example of use of gross cost per gross job is that of King (1990: 20), for the first stage of an evaluation using econometric techniques to explain the 'cost-effectiveness' of Regional Selective Assistance (RSA) projects in the UK's Assisted Areas. This evaluation took the number of (gross) jobs achieved by each project at the date of the final RSA payment (King, 1990: 37). GCPGJ was used in the ex ante appraisal of individual offers of RSA, and the predictions were subsequently compared with later outturns. It should be noted that in this instance the estimates of cost used in these calculations were discounted. The reason may be significant; it was because the DTI needed to satisfy European Community conditions regarding state assistance to industry: "Both assistance and costs were discounted at a broadly commercial rate." (King, 1990: 16).

One argument in favour of using GCPGJ in this first stage of evaluation is its relative simplicity:

"In principle, these gross measures are not adequate indicators of the cost-effectiveness of the scheme as a stimulus to employment in the Assisted Areas. But they are available for monitoring purposes while RSA projects are in progress, in a way that the more comprehensive net cost per net job are not." (King, 1990: 80) (12.3.1)

King offers no justification for this conclusion, and the rationale should not pass without challenge: if well established estimates of such variables as deadweight, displacement, multiplier effect and tax flowback are available, it may be relatively straightforward to apply them in order to produce net figures. It is not clear whether King is referring to the absence of measures of these variables, or to difficulties in estimating administration and compliance costs. A number of studies have sought to make such estimates, e.g. Lenihan, (2002); Lenihan and Hart (2004). Later consideration of actual practice however supports the conclusion that, on the case of grant deadweight at least, it is difficult to set quantitative values to the variables for the purpose of appraisal and evaluation (11.11.1).
One clear disadvantage of the GCPGJ measure is that it takes "no account at all of when these jobs are created and of how long they will last." This is frankly acknowledged to be unsatisfactory: "A more satisfactory measure must therefore take account of the timing and duration of jobs achieved – so that projects which generate jobs for different periods of time can be compared on a consistent basis." (King, 1990: 50) The GCPGJ measure also treats new 'Jobs Created' as being of the same value as 'Jobs Safeguarded', and pays no consideration of the value added generated by the jobs.

These objections are derived from a perceived lack of construct validity of the GCPGJ method rather than from any empirical conclusions derived from the tests themselves. Policy makers need to have some idea of the length of time jobs are likely to last in order to assess their true benefit. They also should require some estimate of the value they generate in terms of wages or profit: employment in a knowledge-driven, high skill sector such as bioscience or optronics is surely to be preferred to that in low-cost manufacturing. This observation points to the main defect of cost-effectiveness tests – that they fail to distinguish between high- and low-value jobs. It is because of this theoretical weakness in the concept of GCPGJ that it is not used in developing a robust appraisal model; but the criticism is equally valid in the case of a number of other tests.

12.3.2 Gross cost per direct (gross) job (GCPGJY), using ‘Job Years’ (not discounted)

One difficulty facing practitioners is that the expected duration of jobs created by projects may vary widely. Projects generating long-term employment are clearly preferable to those producing only short-term jobs. Estimating ‘job life’ may be difficult, but King (1990: 81) suggests the "rather crude" working assumption that project jobs will last for the same length of time as the capital assets of the project. The outcomes can then be expressed as ‘job years’, which provide a more realistic measure of the impact made by the project rather than the number of jobs at various points in time.
This refinement makes a significant improvement on the previous test by taking into account the expected (or actual) duration of jobs. However, the GCPGJY figure in this formulation remains undiscounted: a job in five years’ time is treated as having the same value as a job now. Nor does this refinement distinguish between high- and low-value employment.

12.2.3 Gross cost per direct (gross) job, using Permanent Job Equivalents (GCPGJPJE)

An alternative common basis of measurement can be provided by the conversion of Present Value Job Years into an annuity, described as Permanent Job Equivalents (PJE), which represent the number of ‘permanent jobs’ which would be equivalent (at the same rate of discount) to the temporary jobs that were actually created (King, 1990: 50).

Conceptually, this is identical to the previous test, but instead of job years, it employs an alternative formulation:

“At a given discount rate, these two methods of reducing jobs at different times to a common basis are equivalent to one another: they simply scale the numbers in a different way.” (King, 1990: 50)

Wren (2003: 17) estimates that “a job that lasts for 10 years is equal to 7.4 Present Value Job Years, and a job that lasts forever is equal to 16.7 PJYs or 1 PJE”. This conversion rate clearly depends on the discount rate; and the values produced by these calculations has been questioned (NERA, 2003: 36).

A number of further questions arise concerning the use of Permanent Job Equivalents. To begin with, two methods are suggested in the literature for calculation of PJE. Armstrong and Taylor (2000: 388) simply apply the rate of discount to the gross total; whereas King (1990: 50) applies it to the discounted total, involving the seeming paradox of double use of the discount rate, a conceptual issue that resurfaces in objections to methods of calculating real resource cost (Hill and Williams, 2005).
A more significant issue concerns the nature of a Permanent Job Equivalent, and in particular why it should depend upon the discount rate. The implicit justification appears to be that jobs are presumed to last the same length as the capital assets they are linked to, as in King (1990), so that the price of capital is relevant. But sharp, short-term variations in interest rates would compromise decisions involving long-term projects. Resolution of this issue is beyond the scope of this thesis, but the confusion revealed in the treatment of what appears at first sight a mundane concept is disturbing. A similar fuzziness in elaboration of the theoretical base of shadow pricing is examined by Hill and Williams (2005) (10.3.2). This thesis avoids use of the concept of PJEs.

12.3.4 Discounted tests

The first major variation of evaluative methods takes account of the temporal dimension, in particular the generally accepted preference of society for consumption now rather than in the future. This involves applying a rate of discount to reduce either future costs or future benefits, and has been criticised for lack of reality (8.4.1 and 10.3.1). This approach produces a set of tests parallel to the undiscounted tests already discussed, using discounted cost-effectiveness measures of gross jobs, job years and permanent job equivalents. Again, there are a number of ways in which the discount rate can be applied:

- Direct use of the discount rate to produce a measure of Net Present Value (defined in 5.2.1) of future jobs or profit, which always assumes that returns are invested at the available interest rate (it should be noted that this process takes account of time-preference but does not differentiate between high- and low-value jobs);
- Use of the internal rate of return (IRR), the discount rate equating the present value of a project’s cash inflows to the present value of its costs, or outflows, or the modified internal rate of return (MIRR), the average annual rate of return that will be earned on an investment if the cash flows are reinvested at a specified rate of return;
- Use of an average present value calculator, which converts the flow of Assisted Area employment over time into present value job years (PVJYs) (Swales, 1997b: 859).
12.3.5 Present value job years

In their evaluations of the UK's Regional Selective Assistance grants, both King (1990) and PA Cambridge Economic Consultants Ltd. (1993) express the additional employment generated in terms of present value job years, using a discount rate of 5-6 per cent. However, HM Treasury (1995) concedes: "There is no general agreement on whether it is appropriate to discount employment back to a base year, or at what rate". Swales (1997b: 861) maintains that this problem reflects the fact that projects delivering different time profiles of jobs are not providing homogenous output; and hence a cost-effectiveness ratio based on current cost cannot be constructed. This difficulty arose because the Treasury failed to advocate an operational method to measure the discounted net present value of the jobs created, despite the general commitment made in their guidelines (HM Treasury, 1991: 70).

King (1990: 50n) argues that unless there are persuasive reasons for doing otherwise, it seems appropriate to discount both costs and benefits at the same rate. However, two caveats should be made about the use of discount rates in assessing the present 'value' of future benefits or future costs as part of ex ante appraisal:

- While there is a strong case to be made on grounds of opportunity cost for using prevailing interest rates to discount future costs, choice of a particular level of discount rate for application to future benefits may be highly subjective, involving a judgement of the relative needs of present and future generations;
- Interest rates may be subject to frequent change, and techniques of evaluation and appraisal need to be constructed in a way that permits easy adaptation of the model to such alterations (such changes would be more important where the timing of revenue and costs differ markedly).

Variations could be made to the rate of discount applied to future benefits, for example to reflect the need to encourage sustainability. However, as argued below (13.4), the device of shadow pricing offers a preferable means of reflecting future needs.
The model developed in this thesis uses the first of the methods discussed above (12.3.4), namely direct use of the discount rate to produce a measure of Net Present Value and a Benefit-Cost ratio. Direct use of the discount rate is familiar and readily applied to individual annual outputs, while the average present value calculator coefficient used by Swales (1997b) measures the present value of an average flow over a number of years.

12.3.6 Gross Cost per Net Job (GCPNJ)

A further variation is introduced by modifying the benefits side, in this case the figure for direct jobs, job-years or permanent job equivalents. This variation seeks to make the estimate of benefits more realistic by taking account of deadweight, displacement, substitution and multiplier effects, which were described in 11.11 (Taylor, 2002: 191).

Two GCPNJ approaches may be identified, one undiscounted, the other using the discount rate. In practice, however, for the reasons given above, only the discounted measure is used. The GCPNJ measure is significant in that it has been recommended for use in its discounted form by the UK Treasury.

12.3.7 Net Cost per Net Job (NCPNJ)

A further logical variation involves refinement of gross cost as well as gross benefit figures. Gross costs to the exchequer are likely to differ from net costs (Moore and Rhodes, 1976; Taylor, 2002: 192). This difference arises because of a number of considerations that determine net cost figures. These include the additional benefit of tax clawback and reduced benefit payments to people in new jobs created by projects (Armstrong and Taylor (2000: 386-387). This method has been used in the two major evaluations of the UK Government’s Regional Selective Assistance policy (King, 1990; PACEC, 1993; Armstrong and Taylor (2000: 386-387). Thus King (1990) adjusts the gross cost figures by subtracting the ‘tax flowback’ ($f$) that results from taxation of Regional Selective Assistance receipts. The importance of these refinements to both costs and benefits is underlined by the work of King (1990), who found little correlation
between gross costs per direct job created and the net costs per net job created (Taylor, 2002: 193).

This NCPNJ process differs from the GCPNJ in that it subtracts the ‘tax flowback’ (£) that results from taxation of RSA receipts. It should be noted that this definition does not allow for the full net change in the government’s budgetary position as a result of the subsidy, which should also involve subtracting:

a) additional tax revenue (from new incomes generated, as well as RSA receipts); and
b) reduced unemployment benefits (Swales, 1997b: 1).

Arguably, therefore, net cost per net job is a misnomer for this calculation, for the same reasons advanced in criticism of the UK Treasury’s former ‘100 per cent crowding out’ rule (10.4).

12.4 Cost-benefit analysis tests

A further broad set of tests involve various forms of cost-benefit analysis, which as argued above, provides the most suitable method of evaluating regional policy (10.6). Typically, the results of cost-benefit analyses may be expressed as a measure of net present value or a benefit-cost ratio. For example, the appraisal system favoured in the Irish Republic produces a present value benefit/cost ratio (13.4) (Forfás, 2003: 53). This thesis suggests that both measures are necessary for agencies that seek to optimise their impact under conditions of budget constraint.

The key difference between cost-benefit analysis and the cost-effectiveness tests reviewed so far lies in the systematic measurement of the benefit side by cost-benefit analysis. Instead of treating the benefit measured as an arbitrary and homogenous item (e.g. ‘jobs’ or ‘job years’), cost-benefit analysis seeks to attach a money value to the benefit, which can be compared using the same scale as cost. The limitations to the cost-benefit method have already been examined in Chapter 10, and constitute a barrier to its universal application to project appraisal and evaluation, as discussed for example in the
Despite these restrictions, however, there remains a valid case for a cost-benefit based system of appraisal and evaluation of at least a considerable range of projects and programmes. This point is made convincingly by Forfás (2003: 30), which argues that the task of project appraisal becomes more rather than less important as the economy approaches full employment. Most projects that create jobs are likely to yield net benefits when unemployment is high, but at full employment any new project will almost inevitably lead to the decline of any existing one. Selectivity between projects therefore becomes more important and an effective evaluation framework should allow for differences in their quality, longevity and wealth-creating ability (Pickernell, 1999: 5).

Moreover, at full employment, more complex issues such as evaluation of externalities come to the fore. Such externalities can be project-specific, and more difficult to measure than the gap between the market and shadow wage levels:

"Thus cost-benefit analysis, while more important than ever when the economy is at full employment, also becomes far more difficult." (Forfás, 2003: 30)

### 12.5 A Cost-Benefit Critique: the Swales Model

Some allowance is made for system-wide effects in the cost-benefit method of evaluation proposed by Swales (1997a and 1997b). This approach was developed in response to the recommendations made at that time by the Treasury, but now superseded, as detailed in later in this section. The rationale and construction of this method are analysed in some detail as the arguments involved are highly relevant to the development of a robust method of appraisal and evaluation.

Swales (1997a; 1997b) analysed the ex ante methods proposed at that time by the Treasury in its Green Book (HM Treasury, 1991), those of the interdepartmental Evaluation Group on Regional and Urban Programmes (EGRUP) (HM Treasury, 1995) and two government-funded ex post evaluations of Regional Selective Assistance (King,
A number of different criticisms were made of the ex ante and ex post evaluations; for example, the Treasury did not incorporate multiplier impacts in its ex ante assessment of benefits (HM Treasury, 1991: 70-71) for reasons which are unclear (Swales, 1997a: 80). As discussed below, later Treasury guidelines have modified this position.

The Treasury’s 1991 Green Book guidelines set out four criteria for determining ex ante eligibility for a project support:

1) Project additionality – the requirement that the project would not go ahead without assistance (or only on a smaller scale or at a later date).
2) Project viability – that both the recipient firm and the project should be viable once subsidies are taken into account.
3) A national efficiency requirement, essentially a cost-benefit hurdle that “if all inputs and outputs could be correctly valued and discounted at an appropriate rate the project should have a positive NPV [net present value]” (HM Treasury, 1991: 70); King (1990: 68) explains that this test is designed to ensure that each project was of some benefit to the economy of the UK as well as that of the Assisted Areas.
4) A cost-effectiveness condition reflecting the exchequer cost-per-job limit (Swales, 1997a: 74).

The additionality and national efficiency criteria (1. and 3. above) together constitute a coherent cost-benefit appraisal scheme for a selective subsidy, while the other conditions (viability and cost-effectiveness) yield an additional cost-per-job condition. However, Swales (1997a: 75-76) maintains that as then formulated, the Treasury approach was flawed. For example, on the cost side no account was taken of the costs faced by the private sector in complying with the rules concerning the subsidy, nor was any direction given on how the cost-per-job ceiling should be fixed. However, more significantly, the failure to specify how total benefit is calculated makes it impossible to compare the respective costs and benefits of different programmes. This means there was no uniform trade-off between the costs and benefit measures:
"A project with a high NPV will be rejected if it does not generate enough jobs, and similarly a project that generates jobs at a low exchequer cost will be rejected if it does not yield a positive NPV." (Swales, 1997a: 76).

The failure to assess the real benefit of the RSA programme was also inherent in the way the way the Treasury treated the contribution of job creation in the Assisted Areas to UK employment as a whole. The Treasury assumption (now superseded) was that jobs created in one part of the UK simply displace jobs elsewhere in the UK (unless created by foreign direct investment):

"Because of crowding-out at the macroeconomic level, effects on employment should not be included as benefits of projects in an efficiency test." (HM Treasury, 1991: 70)

Thus two distinct criticisms of the previous Treasury approach are made, one micro and one macro. The micro-economic shortcoming of the cost-effectiveness technique recommended has already been described above (9.2), namely that (unlike cost-benefit analysis) job benefits are specified as a homogenous item, without regard to the levels of wages and profits generated by different projects.

The macroeconomic objection concerns the assumption underlying ‘100 per cent crowding out’, namely that regional financial incentives do not lead to positive macroeconomic effects, because jobs are simply ‘diverted’ and not ‘created’ (Taylor, 2002: 194). Therefore, the Treasury view has been that it is appropriate to use the gross costs of a project rather than the net cost, with the object of assessing whether or not the taxpayer is getting ‘value for money’. Thus, in the Treasury’s view at the time, a new job in Wales (while desirable as a policy goal) would lead to the loss of a job elsewhere in the UK. Policy-induced employment increases in geographical areas receiving regional assistance would be matched by corresponding reductions in other areas, so that regional policy had a purely spatial redistributive (or equity) role, which in turn was enshrined in evaluation practices (Ferguson et al, 2004: 4). Therefore, it would not be reasonable to reduce the cost of the programme to take account of the reduction of unemployment.
benefit or the additional tax collected as a result of the new job in Wales, because that gain would be wiped out by a comparable loss elsewhere.

The problem is that this ‘value for money’ indicator fails to provide policy makers with a clear cut decision rule. Its usefulness is confined to comparing value for money across projects, comparing the cost per job of different types of financial assistance, and for different types of firms (Taylor, 2002: 194). It may be argued that this is in itself a considerable contribution to appraisal and decision taking; but Taylor (2002: 194) stresses the fact that it is therefore possible for a project to be rejected even though it may have a positive net value for the economy.

Furthermore, it is argued that “key variables, such as the level of displacement, the multiplier value and even the degree of additionality itself are determined by system-wide forces”, while the ex post evaluation process carried out by civil servants takes place at the level of individual projects (Swales 1997b: 861). This meant that the Government’s evaluations said nothing about the way the UK economy operates across space.

Strict application of the previous Treasury rule could mean rejection of projects even though they have a positive present net value for the UK economy as a whole (Armstrong and Taylor, 2000: 392). In effect, this stipulation prevents consideration of the full benefit of a regional policy aimed at the spatial redistribution of employment, which “implies that, at the margin, it values jobs in some parts of the country more highly than jobs in others” (Swales, 1997a: 75-76). Even if regional policy is merely redistributive across the UK, this higher value should be taken into account as part of evaluation (Pickernell, 1999: 19).

The Treasury assumption that jobs created in one region of the UK cause 100 per cent displacement in another ruled out consideration of the likely overall saving to the exchequer of reduced unemployment benefit or gain from additional income tax. To estimate such effects, a multi-regional, macroeconomic model would be needed, with identification of effects on the different nations and regions of the UK. The presumption
that this process would be too difficult or costly is contested by McGregor (2001), using evidence provided by the Computable General Equilibrium model for Scotland developed by Strathclyde University. It is claimed that such a spatially disaggregated computable general equilibrium model of the UK, distinguishing the English, Scottish, Northern Irish and Welsh economies, could be produced for a third of the cost of the King (1990) report (Swales, 1997b: 861). Ferguson et al. (2004: 16-17) call for development of an interregional model of the UK economy in which the impact of policies on all regions and on the UK economy can be assessed.

While the analysis presented by Swales (1997a; 1997b) identifies a number of problems, the key criticism concerns the way in which the benefit of the subsidy – the value of jobs generated – is measured. The Treasury approach at the time appeared to be ‘a job is a job’, regardless of the wages it generates (whereas estimating the value added in the form of additional employment, profit and rent would provide a more meaningful measure). The policy implication is therefore clear: the various criteria should be amended, and combined to produce a single, consistent operational method of appraising individual RSA applications ex ante and evaluating the ex post RSA programme as a whole (Swales, 1997a: 76). Two steps were identified to enable this process to take place:

- Most significantly, on the benefit side, valuing additional (regional) employment directly – or, equivalently, costing labour at a shadow wage set below the market wage;
- On the cost side, incorporating the costs of private compliance, public administration and tax-raising incurred by the policy (Swales, 1997a: 76-77), estimated to impose a cost of around 20 per cent of the value of the subsidy (Swales, 1997b: 860).

In place of the now superseded Treasury approach, Swales (1997b: 863) advocated valuing the benefit as “the increase in employment multiplied by the difference between the market wage (w) and the ‘shadow’ wage (w∗).” The methodology involved is set out in detail in the next section (12.6). This approach goes a long way towards differentiating between high- and low-value projects, even though it does not take specific account of additional profits generated. It provides a decision rule that has the great advantage of
being embedded in cost-benefit analysis, which argues that a project is worthwhile provided the social benefits outweigh the social costs (Taylor, 2002: 194). The resultant equation generates the maximum proportion of grant aid which should be given for each job so that over a 10-year time horizon the effect of giving the grant is fiscally neutral (Pickernell, 1999: 19).

It should be noted that in recent years the previous Treasury approach has been superseded by a ‘new regional policy’ focus on each region ‘realising its full potential’ and seeking to grow through supply-side-reform-induced productivity growth. Part of this new departure involves dumping the previous crowding out formula: “Indeed, here not only is there apparently no adherence to the old notion of a 100% crowding out, but the impact of any individual region’s actions on any other region appears to be ignored entirely.” (McGregor and Swales, 2003:17)

Thus, in effect, the Treasury appears to have taken up a polar opposite standpoint in which no account at all needs to be taken of possible displacement effects in regions other than those assisted by regional policy. National efficiency gains are acknowledged, though how they are to be identified and measured is not spelt out (Ferguson et al, 2004: 4).

12.6 The Modified Cost-Benefit Model

A relatively straightforward decision rule based on cost-benefit analysis has been devised by Swales (1997a; 1997b), and this rule could be adapted to assist the evaluation of the relative impact of different programmes. Swales (1997b) studied two Government-funded ex post evaluations – King (1990) and PA Cambridge Economic Consultants Ltd. (1993) – of Regional Selective Assistance (RSA), a form of discretionary assistance to projects in UK Assisted Areas. Both these evaluations used a common index of policy efficiency, ‘net cost per job’.
The net present value of a project is defined by Swales (1997b) as:

\[
\text{(Value added by project) - (Cost of administration, application) - Subsidy}
\]

This equation is developed to produce a ratio of cost per job / wage to calculate the maximum value consistent with a net present value that is positive and therefore worthwhile. This approach suggests that provision of publicly funded regional policy can be justified providing that its net present value is positive. The theoretical basis of this process is straightforward: it will be worth increasing public investment in regional policy instruments as long as the marginal product in terms of net present value (a measure of utility to society as a whole) is positive.

However this approach may be further developed in a direction that serves the goal of this thesis: the formula for value added also suggests a method of judging the relative contribution of different policies or programmes to economic development. Assume that two different projects A and B are compared to assess their relative contribution to a defined policy objective such as employment in the context of a budget constraint. The net present value of each programme for each year can be calculated using the formula proposed by Swales (1997b: 862), namely:

\[
NPV = J(a(1-d)(1+m)(w-w^*)\sum - C([1-\Delta] + K)
\]

where \( J \) is the gross number of jobs generated over the period of the project; \( a \), the degree of additionality, is the proportion of employment in the assisted project which would not have occurred without the subsidy where \((0 < a < 1)\); \( d \) is the displacement of other jobs in the locality, \( m \) is the additional jobs created in the locality through supply chain and household income multiplier effects; \( w \) is the average market wage locally for the type of workers employed by the project; \( w^* \) is the market wage; \( w^* \) is the shadow wage, reflecting the opportunity cost of employing workers in the locality; \( \sum \) converts extra jobs generated over time into present value job years; \( C \) is the cost to the exchequer of the subsidy; \( \Delta \) is the proportion of subsidy expenditure which is deadweight; and \( K \) is
the proportionate relationship between the subsidy payment and the administrative, compliance and tax-raising costs. Armstrong and Taylor (2000: 392) present a simplified formula for calculating the value added by a project to the assisted area.

It will be worth supporting any project providing its net present value is positive. Given a budget constraint on the total funding available for support measures, it will be worth supporting programme A rather than B as long as the marginal increase in its net present value produced by each additional pound invested in programme A exceeds that for programme B. The factor \( (w - w^*) \) produces a rough comparison of overall benefit rather than a simple jobs measure; although calculation of impact on GDP requires a similar assessment of the rewards of all factors of production. In the original formulation, while net present value is discounted, net cost is treated as an undiscounted item, as though incurred at the outset of the project (Swales, 1997b: 863). In the research project, however, allowance is made for discounting costs incurred during the life of the project.

Agencies such as the WDA in practice also measure the contribution made by projects to profit (Williams, 2005).

12.7 Large-scale regional models

Large-scale regional models have also been used to measure the potential impact of policy. In particular, input-output analysis, which can provide detailed estimates of the impacts of exogenous shocks on the regional economy, enables economists to assess the impact of various shocks to the economy. For example, Welsh input-output tables for 1994 detail trade flows between 67 separately defined sectors of the Welsh economy (Hill and Roberts, 1996). These, among other applications, have been used to generate multiplier impacts of the Welsh forestry industry (Munday et al., 1999). The Scottish Input-output tables have been updated annually to measure impacts such as the effect of inward investment on Scotland’s income and employment (Alexander and Martin, 1997; Taylor, 2002: 196).
Economists differ in their views on the relevance of large-scale regional models to the evaluation of regional economic policy. Some have called for adoption of large-scale computable general equilibrium models, partly to identify the effect of regional policies on either non-target regions and/or the economy as a whole (Ferguson et al., 2004: 3). This has to date been resisted by policy makers and official evaluators, with others feeling that computable general equilibrium models are too opaque to be ‘sellable’ to politicians. There is no doubt however that the information gained from such models can be used to provide information on variables used in smaller scale project appraisal and evaluation: for example, information from input-output tables of the Welsh economy are used to generate GVA figures per full-time equivalent worker for all Unitary Authorities in Wales which provide targets for use in project appraisal (Williams, 2005: 8).

This thesis aims to develop a practical method of appraising and evaluating development projects which, individually, are unlikely to create impacts that could be measured by macroeconomic methods. Detailed consideration of large-scale regional models is therefore beyond the scope of this study; but the insights they provide can be of great value.

12.8 Conclusions

This thesis seeks to construct a model that enables us to predict, monitor and evaluate the outcomes of regional economic policy. This requires a combination of clear, well-founded theory and a robust and practicable system of assessment. Even a correct theory is of limited value if we are unable to adapt it to carry out observations of the real world. Even less useful is accurate measurement if the basic theory on which they rest is false.

The earlier part of the thesis reviewed what lessons from economic theory can help ensure that evaluation rests on a firm theoretical base. This chapter examined a range of tests that have been used or suggested in the evaluation of regional economic policies, and has argued that such tests should be grounded in cost-benefit analysis. It has noted the developments in methods of evaluation used in the United Kingdom and more
especially the Republic of Ireland – such as the discounting of future costs and benefits in line with society’s perceived time-preferences, the treatment of risk and the refinement of gross measures of costs and benefits to produce more meaningful net figures.

This review of the practical application of appraisal and evaluation involved consideration of a wide variety of tests that have been used to assess the contribution of regional economic policies. Often there is a trade-off between simplicity and accuracy – between simple tests that are relatively easy to undertake and more refined tests that involve greater complexity and perhaps more time and more expense, and which policy makers may find difficult to appreciate.

As suggested in 12.3, tests may be classified in a number of ways. One clear distinction is between micro- and macro-evaluations: this thesis is oriented to development of a micro-evaluative technique for use at project and programme level. Within the micro family, an important distinction may be drawn between cost-effectiveness tests, which tend to take the benefit as a given, and cost-benefit tests that seek to estimate the value of the benefit as well as the cost to society as a whole.

For the reasons set out previously (10.6), this thesis favours the use of cost-benefit tests. These make possible comparison of real benefit and real resource cost to the economy by means of shadow pricing. Among the advantages of cost-benefit is its potential for generating a consistent scale with which to rank projects. Further refinement could be introduced by attaching a higher value to Gross Value Added or employment generated in regions facing particular economic and social deprivation; and by scoring according to sustainability criteria. Chapter 14 constructs a model based on these foundations.
Chapter 13  Appraisal methods in UK and Ireland

13.1 Introduction

This chapter considers how the principles of appraisal and evaluation described so far are implemented by a range of institutions in the UK and Ireland. This examination shows that the guidance of the UK Treasury in its various Green Books (HM Treasury, 1997; 2003) is tailored by various departments and public bodies at a UK level, in England and the devolved administrations to suit their own requirements (13.2). This practice is in contrast with the appraisal system used in the Republic of Ireland (13.4). This treatment does not attempt to provide a full overview of the methods used, but rather to draw out a number of themes relevant to development of the model.

13.2 UK Appraisal Methods

The purpose, process and methods of policy, programme and project evaluation are set out in a number of UK Government publications – generically in the Treasury’s Green Book and the Magenta Book (Cabinet Office, 2003) and in a number of derivative treatments, such as those for regeneration (ODPM, 2004a; 2004b and 2006) and small businesses (Small Business Service, 2005) in England. A recent overview is provided by DTI (2006). Detailed advice on project management, such as Prince2 (Projects in a Controlled Environment) are provided by OffPAT (2005).
While agencies' own procedures duly reflect the Treasury's guidelines, it would be naïve to suppose they are applied to the letter on all but a handful of projects. Practitioners regard Green Book guidelines as being more appropriate for major capital projects, wider strategies or new regulations or deregulation, while projects costing less than £1 million are generally regarded as falling below Treasury limits (Beattie, 2006). As the Green Book itself comments, the effort applied at each step should be proportionate to the funds involved, outcomes at stake, and the time available:

“In the early steps of identifying and appraising options, summary data only is normally be required. But later on, before significant funds are committed, the confidence required must increase.” (HM Treasury, 2003: 4).

Generally, the resources allocated to assessing the additional impact of a project should be proportionate to the nature and scale of the project, as reflected in the guidance of English Partnerships (2004: 1-2):

“Projects that are novel, contentious, repercussive or involve a high level of risk will require more in-depth analysis, as will larger projects – such as those that involve a significant amount of public expenditure. Standard projects of a type that have been delivered before will typically require less original analysis and can more easily draw on appraisals and evaluations of previous projects. For smaller projects, it will often not be feasible to undertake any detailed assessment of additionality. In these cases, project appraisal should be informed by a qualitative assessment and statement about project additionality, as well as programme and policy level assessments.”

Moreover, in practice agencies are unlikely seriously to appraise unrealistic options, and may well cover themselves by formal appraisal of two or three, one of which should be the 'do nothing' option (Beattie, 2006). For average projects (less than £5 million) the real choice is between the project and 'do nothing'. The Green Book guidelines themselves have not been used in accredited training procedures. Instead, government departments and agencies have adopted features from them in their own procedures, such as the Single Programme Appraisal Guidance (SPAG) in England (DTI, 2003). Pilot projects may be used to generate estimated outputs where there is no hard evidence (DfES, 2004: 7).
The main steps recommended for appraisal have already been set out in Chapter 11 with further discussion of criteria used in Chapter 12. The conduct of appraisal and evaluation in the UK reflect the guidelines laid down by the UK Treasury in the ‘Green Book’ (HM Treasury, 2003; McVittie, 2005:10), which are regarded as binding on government departments and executive agencies in England and have also been tailored for use by the devolved administrations and used widely by local government and regional development agencies, especially for proposals that require funding from central government (Defra, 2003: 6). Guidelines by Forfás (2003) govern the process in the Republic of Ireland.

The UK ‘Green Book’ guidelines were revised in 1997 and again in 2003 to reflect changes in procurement, reduced interest rates and concern to improve delivery of public services (Defra, 2003: 6). Among the changes made by the 2003 Green Book were:

- increased emphasis on valuing the benefits of interventions;
- greater transparency for risk, including “unbundling” optimism bias from the discount rate;
- methods to better take into account long-term and distributional impacts of actions (Defra, 2003: 6).

### 13.2.1 Department of Work and Pensions

The Department of Work and Pensions uses cost benefit analysis to provide a systematic comparison of its Active Labour Market Programmes (DWP, 2005: 4). However, such analysis is one of a range of methods employed, depending on the nature of the project studied (Allaker, 2006).

A description of its appraisal methods available in the public domain states that proposed reductions in the department’s headcount and financial resources are driving a change in policy focus away from a ‘what works?’ towards a ‘what is cost effective?’ approach (DWP, 2005: 4). The recommended ‘cost benefit’ method is published (DWP, 2005) while detailed analyses using the framework are sent to ministers but are not themselves
in the public domain (Allaker, 2006). However, an appendix presents a copy of an excel spreadsheet showing how calculations could be made (DWP, 2005: 13; 28).

This framework is intended for use as an appraisal tool to allow the systematic review of different policy options (DWP, 2005: 6). Not every piece of information should be loaded on to the framework: rather it should act a reference point to other information sources such as research reports (DWP, 2005: 7). As evidence becomes available, it should move away from being an appraisal tool to providing evidence for evaluation (DWP, 2005: 6).

The perceived advantages of this system include the systematic recording of information on cost effectiveness and to assist in carrying out appraisal within ROAMEF statements (8.4.3) to help determine which polices offer the greatest net benefits (DWP, 2005: 4-5). The framework is organised in two tiers as shown in Figure 13.2.1 to reflect “a general shift in policy away from focusing on individual programmes to looking at the bigger picture of client groups” (DWP, 2005: 8). Client groups are described as “defined groups of people with similar characteristics, barriers to work or claiming similar benefits e.g. young people, lone parents, disabled people” (DWP, 2005: 8n). Difficulties in aggregating data collected at project level mean that analysis needs to take place at the broader client group level as well as for individual programmes.

Figure 13.2.1 Department of Work and Pensions two-tier cost benefit framework

Source: DWP, 2005: 7
In its description of the framework, the Department distinguishes between the fiscal and economic effectiveness of programmes studied. The Net Fiscal Effectiveness to the Exchequer of the programme measures “the difference in the costs of a particular programme (the operational costs) and the benefits that flow back to the Exchequer in terms of increased tax revenue and benefit payments saved” (DWP, 2005: 20). By contrast, Net Economic Effectiveness measures the effects of the programme on the economy more widely; it sets the benefits of increased output in the economy (produced by more people flowing into work) against the net costs of running the programme including the cost of raising taxes to pay for it.

On the benefits side, the framework seeks to score for profit benefit as well as additional wages generated by the project. This is achieved by applying a ‘product market corrector figure’, a variable used in the economic effectiveness calculations to estimate the value of output produced by additional workers (DWP, 2005: 24). The rationale for this process is that in an economy with perfectly competitive product markets, workers would be paid a wage equal to their marginal revenue product. However, in reality product markets are likely to be imperfectly competitive, which implies that not all the additional output generated by the net increase in output goes to workers in increased wages. This additional output should however be scored as a benefit.

The value of workers’ output is divided between wages and firm’s profits, and the total value of output is estimated by multiplying the worker’s wage by the product market corrector to account for the profit share of output. The figure for 2003/4 was set at 1.26, but it is noted that this figure changes over time and should be checked with the Treasury’s model (DWP, 2005: 24).

Moving from the fiscal to the wider economic measure also requires an estimate of how impacts are effected by grant deadweight, displacement and the multiplier effect. The Department’s framework makes allowance for deduction of grant deadweight from gross wages in calculating benefits (DWP, 2005: 20). It also states that substitution and
displacement effects must also be built into the evaluation of programmes (DWP, 2005: 14). However, it makes no adjustment for the multiplier effect.

On the cost side, the wider economic measure takes account of the deadweight cost of taxation – this is intended to reflect the fact that raising taxation to fund programmes impacts on work and output incentives in the wider economy that distort labour supply decisions (DWP, 2005: 21; 24). It is noted that estimates of the deadweight cost of tax lie between 0.2 and 0.3, and recommended that a mid point of 0.25 is used (DWP, 2005: 24). Fiscal cost of the programme is multiplied by this tax deadweight factor to arrive at an economic cost figure, which is subtracted from overall benefit, and divided by the total number entering work to produce the ‘net cost per person’ (although the figure could in theory be positive).

Table 13.2.1 Net Benefit/Cost per Person

<table>
<thead>
<tr>
<th>Net Economic Cost in £s:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Additional Flow into Work x Gross Wages x Product market corrector x Job Duration)</td>
</tr>
<tr>
<td>Minus</td>
</tr>
<tr>
<td>(Deadweight Cost of Taxation x Net Fiscal Cost)</td>
</tr>
<tr>
<td>All Divided by Additional Flow into Work (to derive net cost per person)</td>
</tr>
</tbody>
</table>

Source: (DWP, 2005: 22)

While there is no explicit reference to discounting, it is clear from the sample spreadsheet that the Treasury’s 3.5 per cent discount rate should be applied to future costs and benefits. In line with current UK recommendations, the guidelines allow for best, worst and base (or ‘central’) case scenarios (DWP, 2005: 28) and recommend sensitivity analysis, especially in cases where acquiring robust data is problematic (DWP, 2005: 6)

In summary, the Department of Work and Pensions follow the overall Treasury guidelines. The most notable feature of their treatment is the recognition of tax
deadweight as an item used to increase estimates of cost, which appears out of line with the recommendations of the Treasury’s Green Book (HM Treasury, 2003).

As argued in 10.2.4, the absence of an explicit system of shadow pricing to reflect real resource cost and benefit to the economy means that these, along with other UK guidelines, fall short of providing a full cost-benefit framework, in contrast to the process favoured in the Republic of Ireland (13.4).

13.2.2 Department for Education and Skills (DfES)

The Department for Education and Skills has published its evaluation and appraisal guidance to provide Ministers with “the best advice we can on the relative effectiveness and pros and cons of different policy options” (DfES, 2004: 1.) As with the Department of Work and Pensions, the recommended process is considered within the ROAMEF policy making cycle described in 8.4.3 (DfES, 2004: 2). The rationale for intervention is ascribed to market failure (causing insufficient investment in education and skills) or unfairness and unequal access, reflecting the general approach suggested by HM Treasury (2003: 11) (13.4). An explicit step-by-step appraisal process is advocated.

The approach is set clearly within a cost-benefit framework, with examples provided of costs and benefits most frequently encountered (DfES, 2004: 7). This requires identification of additional costs and benefits – ‘additional’ referring to an impact over and above that which would have happened under the ‘do nothing’ option. Clear guidance is given to discount future costs and benefits at the 3.5 per cent rate recommended by the Treasury (DfES, 2004: 8). Appraisals should also consider the distributional effects on different groups (DfES, 2004: 9).

The DfES guidelines recommend a ‘robust judgement of risks’, adjustment for optimism bias and the use of sensitivity analysis to show how changes in particular assumptions made in the appraisal affect costs and benefits, and therefore the NPV and rate of return (DfES, 2004: 11). They also advise consideration of whether it is necessary to carry out a
regulatory impact assessment and screening for environmental impacts. The appraisal results should be presented in the format shown in Table 13.2.2, with the suggestion that it may be sensible to bring together this information in a matrix so that NPVs, risks, unvalued benefits and other impacts can be easily compared between options.

Table 13.2.2 Presenting the Results of an Appraisal

- Policy background, making reference to the policy’s rationale and objectives
- The options considered to meet the objectives, including the “do nothing” case
- The categories of anticipated costs and benefits for each option
- References to available evidence justifying the anticipated costs and benefits
- Overall Net Present Values (NPVs) and rates of return for each option
- NPV or rate of return estimates after any adjustments for distribution
- Sensitivity analysis of the effects on NPV and rates of return of changing key assumptions
- Any additional costs and benefits not included in the NPV calculation
- Practicality and acceptability of each option

Source: DfES (2004: 11)

The Department recommends that ex post evaluation should follow the same key steps as appraisal, together with a summary of:
- How effective the activity was in achieving its objectives, and why;
- Whether the activity has been more beneficial for some groups than others, and why;
- Suggestions for how the policy could be improved to make it more effective;
- The cost effectiveness of the activity – if practicable, a brief discussion of cost-effectiveness relative to other ways of achieving the objectives is useful;
- Why the outturn may have differed from that foreseen in the appraisal;
- What the results imply for future management or policy decisions.

In summary, the Department for Education and Skills provides clear guidance for practitioners on the conduct of appraisal and evaluation which, in general, closely follow UK Treasury recommendations. Unlike the Department of Work and Pensions (13.2.1),
no provision is made to gross up fiscal cost by an allowance for tax deadweight. Again, the absence of shadow pricing means that the guidelines fall short of providing a full cost-benefit framework.

13.2.3 Department for Environment, Food and Rural Affairs (Defra)

Defra, the UK Department for Environment, Food and Rural Affairs recommends ‘appraisal-led design’ for the process of option development for managing flood prevention work (Ministry of Agriculture, Fisheries and Food, 1999: 2).

Economic evaluations are ex-post assessments of the Department's policies, usually undertaken after the policy has been in operation for sufficient time to have some impact or when it has terminated. Evaluations provide evidence of the rationale for the Department's policies, the effectiveness and efficiency with which they deliver their objectives and of any unforeseen side effects. Evaluations are therefore used to inform policy reviews as well as assisting in the formulation of new policies and the policy process generally.

Defra describe evaluation as a policy assessment that is carried out ex post either when the policy has been in operation for sufficient time to have had some impact or, in some cases, after the policy has ended (Defra, 2006: 3). In some cases the project specification may also include specifically an ex ante appraisal of alternative policy options or mechanisms in addition to the ex post evaluation (Defra, 2006: 4).

Defra stipulate that evaluations should include an assessment of the extent to which the policy objectives are met (effectiveness), and what the successes and difficulties have been; it does not however specify use of the technique of cost-benefit analysis.
13.2.4 English Regional Development Agencies

In England, the regional development agencies established after 1997 are required to base their project appraisal techniques on the DTI Single Programme Appraisal Guidance (or SPAG) (DTI, 2003). Each RDA has designed their own appraisal forms and guidance to deliver the requirements of this framework (Gilbert, 2007). The agencies are supported by OffPAT (Office of Project Advice and Training) as a secretariat for the development and sharing of best practice advice and guidance on project appraisal and related processes. OffPAT has developed a practical approach that presents the process in ten key steps (Beattie, 2006). The agencies are represented on a practitioners’ group and a steering group and the groups have produced a number of Appraisal Advice Notes. However, no generalised model has been developed to compare project outputs (Gibb, 2007).

Specific guidance on estimating the additionality of regeneration projects is provided by English Partnerships (2004) and ODPM (2004b). These too are based on the Treasury Green Book.

While not binding in the same way on the devolved administrations in Scotland, Wales and Northern Ireland, agencies in these countries have also adapted the same approach. For example, the Welsh Assembly Government has required public bodies in Wales to ensure that their own guidance on appraisal and evaluation is consistent with the principles set out in the Green Book (Welsh Assembly Government, 2003). Representative from the UK nations take part in the OffPAT groups’ work (Gilbert, 2007).

13.2.5 Appraisal in Wales

The appraisal methodology introduced by the Welsh Development Agency during 2004 employs a number of techniques familiar in cost-benefit analysis: however, it does not generate a single measure of net present value. Instead, it derives a figure for gross value
added based on estimated employment and profit attributable to the project, expressing this as a ratio against cost for comparison with regional targets (Williams, 2005). It should be noted that:

- Costs are defined as those associated with the project, not including overall administrative costs;
- Although allowance is made for multiplier effects, full system-wide effects are not measured (for example, there is no measurement of opportunity cost to the economy by means of a shadow wage, although an analogous process is reflected in the setting of regional value-added targets to assist project appraisal, with lower targets for projects which are located in areas of lower income).

Only a minority of projects are assessed by the value added criterion, with others judged by contribution to capacity building (8.4.1), and the rank orders generated from only one part of business planning (Williams, 2005). Further information is provided on this method in comparing the output measures it produces with those generated by the model proposed in this thesis (15.3.1).

13.2.6 Scottish Enterprise

Scottish Enterprise present an example of the practical use of an impact assessment framework (11.6 and 11.9). Jackson (1998) carried out an analysis of the output measurement framework introduced by Scottish Enterprise for its network of 13 local enterprise companies (LECs). The purpose of this is to provide monitoring, performance review, quality assurance and strategic planning (Jackson, 1998: 562-563). This approach helps compare different schemes, rather than simply assessing whether a single scheme passes the ‘efficiency test’ of generating a positive net value (Pickernell, 1999: 20). Each LEC activity must be allocated to one of 23 activity categories, broken down into activity measures and output measures:

“Activity measures quantify levels of network impact and delivery, whereas output measures quantify the direct effect of these efforts on the targeted area of assistance.” (Jackson, 1998: 560)
For example, activity category D1 is defined as ‘Individual export companies’ assistance’, and one of the activity measure associated with this category is ‘Number of companies assisted’. One of the associated output measures is ‘Rise in export sales’ – the real gain sought by engaging in the activity.

Jackson’s study of the early results of this method of evaluation found a wide range of impacts – with benefit-cost ratios ranging from over 16:1 to 0.15:1 – prompting the conclusion that the measurement framework offered significant potential for improving cost-effectiveness, providing robust estimates of net impacts were obtained on a disaggregated basis and used to reallocate resources to attain strategic goals (Jackson, 1998: 563).

However, two separate sets of problems were identified, technical and strategic. Technical problems included the difficulty of securing firm-level estimates of the preferred multiplicand, value-added (as opposed to jobs or productivity). A second technical problem was that the grossing up of local value-added estimates by Type I or II multipliers to calculate impacts was sensitive to leakage from the region. This was likely to lead to selection bias, as the system would favour labour-intensive projects even if they had limited local linkages (Jackson, 1998: 564).

One strategic problem related to the wide range of outcomes observed, namely that “some of these programmes have a welfare rather than a developmental basis, and should not be compared with other forms of development assistance” (Jackson, 1998: 565). This problem surfaces conspicuously in the practical difficulties agencies encounter in comparing initiatives with discernible impacts on jobs and profits and those with less direct or longer term effects (8.4).

A second strategic problem recorded is that the framework process is based on partial equilibrium analysis, assessing the effect of operating existing economic structures at an
increased level (more of the same). But local development is often concerned with changing structures:

"The correct comparison ..... is not between the deadweight unavoidably incorporated in non-discretionary forms of development assistance and the potential for much greater benefit from perfectly targeted support. It is between the costs of deadweight in the former and the management and information costs of attempting to target in the latter." (Jackson, 1998: 565)

Scottish Enterprise no longer use economic models to calculate and compare the value added by individual projects, as the agency doubted the robustness of such models to generate comparisons at a small area level (Richmond, 2007). Its appraisal process is based on best evidence (e.g. economic intelligence, past evaluation) and estimates of job and/or turnover impacts for a project. The agency uses broad rules to assess GVA (based on ABI data on the ratio of turnover to GVA or GVA per employee). For ex post evaluations, where possible the agency obtains information on profits, employee costs and depreciation from assisted companies to assess GVA, or uses the same approach as for appraisal.

13.2.7 Northern Ireland

A guide to project appraisal is provided by the Northern Ireland Department of Finance and Personnel (DFPNI, 2003). Again, this follows the broad approach of the UK’s Green Book (HM Treasury, 2003), and advocates assessment of cost effectiveness by the use of benchmarking, that is comparing the ratios for a specific project with those for other similar projects (DFPNI, 2003 :10).

13.3 European Structural Fund Evaluation

Guidance on evaluation of socio-economic development in the European Union is provided by CEC (2003) and follows the MEANS (methods for evaluating structural policies) collection. This has a specific focus on evaluation within European Structural Funds, but is not confined to the evaluation of these interventions. Evaluation of EU
structural funds has provided the most comprehensive approach to qualitative assessment (DTI, 2006: 67). Eurada (2005) provides a critique from the point of view of development agencies.

While the European Commission has pressed the case for evaluation and suggested what indicators may be used (11.6) there is no evidence of advocacy of any standard method of assessing impacts at project level in a way that could assist comparison between projects. Nor has any system been proposed by Eurada, the association of regional development agencies (Saublens, 2007).

<table>
<thead>
<tr>
<th>Key Task</th>
<th>Evaluative approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continued relevance of programme strategy and need for any changes in delivery</td>
<td>Reflects ROAMEF framework</td>
</tr>
<tr>
<td>Integration of crosscutting themes (equal opportunities, information society, environmental sustainability)</td>
<td>Attempts to marry multiple objectives</td>
</tr>
<tr>
<td>Progress towards quantified targets</td>
<td>Benchmarking approach</td>
</tr>
<tr>
<td>Progress towards programme impacts</td>
<td>Benchmarking approach</td>
</tr>
<tr>
<td>Progress against performance reserve indicators</td>
<td>Benchmarking approach</td>
</tr>
<tr>
<td>Effectiveness of Programme processes.</td>
<td>Process evaluation</td>
</tr>
<tr>
<td>Further benefits, including community added value</td>
<td>Attempts to assess long-term capacity building measures</td>
</tr>
</tbody>
</table>

The mid-term evaluation of EU Objective 1 programmes in Wales (Williams, 2004; Munday and Williams, 2006) is an example of the combination of a number of analytical methods. The methodology summarized in Table 13.3 seeks to measure the seven ‘key tasks’ set out in the table by using a three-fold methodology:
• Desk-based analysis of documents and data;
• Detailed examination of a sample of approved projects; and
• A 'process evaluation' (8.2.2.3), involving a survey of all members of local and regional partnerships followed by face-to-face interviews with a range of participants (Williams, 2004: 22).

The first column of Table 13.3 indicates the key tasks, while the second column notes the nature of the evaluative approach already discussed (11.9).

While the various conclusions are not within the remit of this thesis, it should be noted that this evaluative approach measures success against targets rather than carrying out a systematic weighing up of costs and benefits. The final mid-term evaluation report nevertheless makes arrives at judgements between the effectiveness of widely differing programmes without adopting a cost-benefit approach (Old Bell 3, 2005). Again, no use is made of shadow pricing.

An updated evaluation of 18 projects selected randomly and balanced in terms of project size and sponsoring organisation provides evidence of the relationship between the gross and net outcomes from the largest EU structural funding programme in the UK, the Objective 1 West Wales and the Valleys programme. The focus is on job-related outcomes (Munday and Williams, 2006: 2; 10-11).

13.4 Appraisal in the Republic of Ireland

While not used for project appraisal in the UK, a cost-benefit appraisal system has been operated in the Republic of Ireland since the 1990s (Honohan, 1998; Barry et al., 2002; Forfás, 2003). The economic model is used consistently by both Enterprise Ireland and IDA Ireland for the ex ante evaluation of industrial projects. The cost-benefit ratio derived from the model is used as one of the criteria underpinning the investment decision, together with a technical and commercial assessment (Compernolle, 2007).
In the appraisal model used in the Republic of Ireland, the opportunity cost of labour is reflected in the level set for the ‘shadow wage’ (10.2.4). In the original Irish appraisal model introduced in 1996, the shadow wage was set at 80 per cent of the market wage nationwide to reflect the net economic gain associated with employment generated. This implied that out of every £100 of a project’s domestic wage bill, £80 is required to elicit the supply of labour, which is an opportunity cost, while the remaining £20 is the value attached to the reduction in involuntary unemployment, which is a net economic gain. By contrast, in a fully employed economy the market wage is the same as the shadow wage – no net economic gain is attributed to generating employment per se (Forfás, 2003: 24).

Achievement of near full employment was felt to have made this 80 per cent level inappropriate. Therefore a reference or benchmark shadow wage \((v)\) set a value of 100 per cent for Greater Dublin, but with regional variations elsewhere – 90 per cent for the Border, Midlands and West, and 95 per cent for the Rest of Ireland. A further refinement proposed was that lower levels of the shadow wage should be set for projects with higher wage levels than the regional average, and these variations were set at different levels for the three regions (Forfás, 2003: 37-43).

“We believe that these values of the shadow wage strike a balance between the need to take into account the benefits of attracting relatively high-wage firms to relatively depressed regions of the country, on the one hand, and the full-employment/increasing congestion situation that prevails in the Greater Dublin region. The same project would be deemed to have a higher benefit if it located in a low wage region, while within a region the higher a project’s average wage, the greater the benefit attributed to it.” (Forfás, 2003: 42).

Regional economic policy has focused principally on the need to tackle unemployment in assisted areas, and the approaches recommended by Swales (1997a and 1997b) and Forfás (2003) both use the difference between market wages and the shadow wage to measure the benefit (to the State as a whole rather than a particular region within it). However, agency support of projects with high wages, skill and R&D can be justified even with full employment because replacement of other jobs by these higher value
activities represents a net gain. This argument leads to the advocacy of a lower shadow wage being set for higher wage projects (Forfás, 2003: 18). This approach suggests a mechanism that could be used to encourage projects that are particularly attractive from the point of view of promoting agreed policy objectives. Different values could be set for the shadow wage to provide for certain defined goals, such as sustainability, promotion of the Welsh and Irish languages or social inclusion.

13.4.1 Migration

Regional wage setting and migration can play an important role in determining employment adjustment over time (Gillespie et al., 2001: 126). Agencies in the UK have not integrated issues of migration into their appraisal systems, despite the policy issues involved, such as impact on the Welsh language, house prices, provision of health and social services and so on. The Irish appraisal system, however, places considerable emphasis on the treatment of migration both into and out of the Republic. The 1996 model had been based on the view that Irish unemployment was insensitive to the rate of job creation, because the openness of the labour market maintained an equilibrium rate of unemployment that equalised the expected value of wages in Ireland with those obtainable by emigrants such that

\[ W_{ir} = (1-u) W_{uk} \]  \hspace{1cm} 13.4.1

where \( W_{ir} \) and \( W_{uk} \) are Irish and UK wages respectively, and \( u \) is the Irish unemployment rate (Forfás, 2003: 24-25).

The Irish appraisal system defines the relevant criterion as "the welfare of the existing indigenous population", treating the increased income of immigrants as a private gain and not including it within the criterion (Forfás, 2003: 26). This approach is frankly admitted to "reflect a relatively conservative perspective on this issue" in the absence of a clear national consensus to the contrary (Forfás, 2003: 26). The method adopted is to treat a fraction \( (\lambda) \) of the opportunity cost part of the wage bill as a net addition to GDP.
reflecting net immigration or higher participation, even though that part of the wage bill does not count as a benefit. The value chosen for $\lambda$ in Greater Dublin was 50 per cent, with higher values outside: previously $\lambda$ had been set at 100 per cent throughout the Republic (Forfás, 2003: 44).

13.5 Tax deadweight: Treatment in Ireland and UK

The Irish system provides an explicit treatment of the cost of distortions imposed on the economy by higher taxes required to fund regional policy, qualified by the additional tax revenue resulting from industrial projects. This burden is described as ‘tax deadweight’; and should not be confused with the more familiar concept of grant deadweight, which was defined above (11.11.1).

In the Republic of Ireland, adjustment for tax distortion has been promoted as orthodoxy since the 1990’s. The rationale is that the act of raising taxes itself distorts the economy by altering input or output prices, thereby changing the allocation of resources. The effect of this is that grants are regarded as costing the economy some multiple of the level of grant. Allowance is made for this cost of distortion by a premium (known as the shadow price of public funds) applied to all government revenues and expenditures. This premium was estimated in 1998 to be 1.5, down from over 2 at a time of high taxation in the Republic in the mid-1980s (Honohan 1998: 3).

This level was justified on the grounds of the high marginal tax rate prevailing in Ireland at the time (ESRI, 1996; Honohan, 1998). In the Swales (1997b) model this is described as “the costs of the inefficient use of other resources implied by the requirements of the subsidy” ($\Delta$), in the expression $1 - \Delta$ (Swales, 1997b: 863). The reduction in Irish tax rates during the 1990s and academic reconsideration of the true burden of tax deadweight (Ng, 2000) led to a further reduction in the tax deadweight parameter $\phi$ to 25 per cent, treating the marginal cost of using €1 of public funds as €1.25 (Forfás, 2003: 36). Both tax-financed costs and taxes included as benefits are
grossed up by 25 per cent to take account of the effect of tax deadweight (Forfás, 2003: 51).

As a “rule of thumb”, deadweight costs of taxation are regarded as “roughly proportional to the square of the tax rate” (Honohan 1998: 25n). This rule is derived from standard theory regarding loss of welfare following imposition of a tax, and rests on the assumptions usually made about an individual’s utility indifference curve and budget line (Creedy, 2003: 5-12). This is expressed as a movement along an individual’s indifference curve tracing out part of a Hicksian demand curve, which shows changes in demand as price for a good changes, while utility is constant (Creedy, 2003: 8). In the United States, it has been estimated that these ‘tax deadweight’ costs of the Federal government raising an additional dollar equal 39 cents, i.e. a multiplier of 1.39 (Jorgenson and Kun-Young, 1991). Lee (2000) argues that such deadweight costs go largely unperceived.

The case for adjusting for tax-distortion is argued strenuously by Honohan (1998: 20-25) in terms of the Government as a discriminating monopolist, raising prices (by taxation) for a low-elasticity group (e.g. the non-traded sector) and lowering them for a high-elasticity group (such as “mobile or footloose investors”). The adjustment applies to both costs and benefits: for example, given the presumed additional burden of distortion, taxes clawed back as a result of grant-aided activities should be valued at a premium reflecting the marginal cost burden avoided (Honohan, 1998: 25). In seeking to identify real resource cost, the system proposed by Forfás (2003: 35-36) allows for the burden of this distortion on the cost side by imposing an extra shadow price of public funds (φ) on cost, so that total cost becomes (1+φ) times the grant. However, this calculation of tax deadweight is based on the theoretical “rule of thumb” described above; and it seems curious that little effort has been made to subject it to empirical testing.

No general allowance has been recommended for the cost of public funds in the UK since revision of the 1997 guidelines – compare HM Treasury (1997: 81) and HM Treasury (2003: 97), which bases the public sector discount rate solely on social time preference.
comparisons – although it can be argued that the concept of tax deadweight is implicit in the Treasury’s use of ‘cost per job’ as a measure and the idea of a government budget constraint. In the absence of an explicit cost of public funds, to single out the budget for regional economic development for such a charge would be inconsistent (Hill and Williams, 2005: 6).

This change of practice has involved some political controversy. It has been claimed by Conservative politicians that the UK Treasury is now almost unique amongst finance ministries in assuming that taxes have no economic cost:

"In line with standard theories of public finance used around the world, the Green Book guidance inherited by the current government made explicit reference to the costs of taxation. In addition to the opportunity cost from government expenditure, it stated there was ‘a further cost of the distortions caused by taxation.’ (The Green Book: Appraisal and Evaluation in Central Government, HM Treasury (1997), p.81). But when the Green Book came to be revised last year, after political intervention, these appraisals of public spending now assume that taxation has no extra economic cost, the public sector discount rate being solely a foundation of social time preference comparisons (see Annex 6 of the 2003 version of the Green Book).” (Conservative Party, 2003: 4).

It is worth noting however that despite acknowledgement of the “further cost of the distortions caused by taxation, at the margin, to fund increments in public expenditure”, no explicit price of public funds was set by the 1997 Green Book, while the rates of discount set (6 per cent in most cases, or 8 per cent for ‘appraising efficiency’) (HM Treasury, 1997: 61 and 81) is much lower than the shadow price of public funds set in the Republic of Ireland, with the most recent reduction in 2003 explained by explicit allowance for optimism bias (11.12.1).

While a cost benefit appraisal or evaluation should include the costs of administration, compliance and raising taxes as well as those caused by distortionary impacts on the economy, this is not the present practice advocated by Green Book in the UK (HM Treasury, 2003). Exceptions can be found: thus the Department for Work and Pensions advocate adjustment of benefits to allow for the burden of tax deadweight (DWP, 2005:
24). It is argued that such costs as applied to evaluation of development projects should only be counted if applied to other appraisals in such fields as health and education (Hill and Williams, 2005).

13.6 Conclusions

This chapter compared appraisal processes in the UK with those of the Republic of Ireland. It examined how the general guidelines of HM Treasury (2003) (13.2) are implemented in practice by ministries and agencies.

This study finds a broad similarity of processes use, for example the ROAMEF evaluative cycle, discounting and allowance for grant deadweight and displacement in calculating additionality. One surprising feature is the use by the Department of Work and Pensions of shadow pricing to correct for 'tax deadweight' distortions.

English RDAs established since 1997 have developed their own appraisal systems based on the SPAG guidelines (DTI, 2003) and the advice on regeneration of English Partnerships (2004), which in turn are based on the Green Book (HM Treasury, 2003). No generalised model has been developed for English RDAs to compare project outputs (Gibb, 2007). In the devolved administrations, where development agencies have been active since the 1970s, a variety of different appraisal systems have evolved, but these too reflect the same overall approach.

In the Republic of Ireland, however, an appraisal system more rigorously based on cost benefit analysis has been developed, with allowance for migration effects, a regionally differentiated shadow wage adjustment and allowance for tax distortion through imposing a shadow price of public funds.

Finally, it should be noted that neither the Irish or Welsh appraisal systems place reliance entirely on cost-benefit techniques. They therefore offer practitioners a decision guide rather than a decision rule. Thus the system adopted by the Welsh Development Agency

263
uses estimates of discounted value-added as one part of a complex process of business planning (Williams, 2005). Similarly, the use of cost-benefit analysis in Ireland is "one part of a wider project evaluation system used by each of the development agencies, which includes an assessment of management, financial track record, target markets and so on". (Forfás, 2003: 23).
14.1 Introduction

This chapter describes the specific nature of the model and the variables presented. It takes into account the guidelines set out for agencies in the United Kingdom and the appraisal model proposed for use in the Republic of Ireland, although it departs from them in several notable aspects noted in this chapter.

Section 14.2 sets out the rationale for adopting the model proposed, while 14.3 describes its main characteristics, listing benefits and costs with their associated parameters and describing the way in which these benefits and costs enter the model. Section 14.4 then describes how the model as a whole is specified. In Section 14.5, the sources of data are specified and the set for parameters. Limitations to the methodology proposed and the key assumptions on which it is based are set out in 14.6, while ethical considerations are presented in 14.7.
14.2 Justification of selected appraisal and evaluation methodology

14.2.1 Aim of model

The aim of the model (as argued in 1.3) is to design a clear, robust and simple instrument, which will be useful to two groups of agents:

- Practitioners of economic development who need to know the likely impact of their activities, and use this information to shape their business planning;
- Government and other agencies who receive reports on the objectives and achievements of publicly funded intervention, and require information to guide the process of resource allocation.

The appraisal and evaluation model is designed to measure what difference is made to the economy as a whole by engaging in particular projects or programmes (Hill and Williams, 2005: 3). This process therefore needs to go beyond identifying project outputs, and take into account wider effects (both positive and negative) on the local economy (Armstrong and Taylor, 2000: 384-386).

The primary use envisaged for the model is as an instrument of appraisal, enabling comparisons to be made between the likely consequences of deciding between competing uses of public money. As noted by Honohan (1998: 34), “the purpose of a formal appraisal system is to rank projects, and to identify a cut-off point in the ranking”. Providing the appraisal system provides a consistent ranking, the absolute values generated for projects are less crucial.

In practice such comparisons may involve different projects that aim to deliver an agreed policy, rather than making judgements about widely differing policies (such as health spending versus subsidies for inward investment). As noted, it is difficult to apply such a model where project outcomes are long-term or indirect in their impacts (8.4.1). In this context, it is important to stress that its contribution is one of a range of processes used in business planning, often reflecting specific demands by government for particular strategies to be pursued (Williams, 2005). Such a process may take the form of a wider appraisal system, embodied in a ‘Quality Ranking
Matrix' that takes into account additional intangible, non-quantifiable benefits associated with a project, as part of a multi-objective approach to take full account of the value of support provided by the development agencies (Forfás, 2003: 23).

The thesis also investigates whether the model has the potential to be used in the monitoring and ex post evaluation of programmes and projects. Clearly, it would be advantageous to students of economic science and to policy makers if a way could be found to make international comparisons of the value of specific programmes and projects. Moreover, despite the limitations noted, the information gathered might be used as part of a wider and more complex process of macroeconomic evaluation.

14.2.2 Key factors

The accuracy of predictions and measurements made using the model depend on four broad groups of factors:

1) The theoretical validity of indicators selected as dependent variables (such as the choice of industrial sector);
2) The accuracy with which they are represented by the data used (e.g. wages);
3) The relation between the gross effects of the policy and the true ‘net’ impacts and the accuracy with which that is represented in the model;
4) As an ex ante appraisal device, on how closely project boards or managers are able to estimate in advance raw outputs generated (in this case, jobs and profit).

In seeking to estimate item 3), the model looks beyond the immediate impacts scored by projects on variables such as jobs or profits:

- An estimate is made of the wider effects on the economy through the multiplier effect, which tends to increase impacts;
- A correction is made for loss of welfare caused by net grant deadweight (as defined in (11.11.1), displacement and substitution, which tend to decrease impacts.)
The nature of these concepts, and a number of conceptual and empirical problems that arise, are discussed in the following sections. No provision is made for leakage (see 14.3.1.6).

Even allowing for ‘optimism bias’ (11.12.1), estimates of raw outputs under (4) are bound to be subject to wide margins of error which will be reflected in estimates of net benefit.

14.2.3 Determining the level of complexity

In designing a model for use in the real world, it should be recognised that there is an inevitable trade-off between complexity and practicality. It is important that the model can be used without undue dependence on data which may be difficult, expensive or time-consuming to collect. It is also an advantage to avoid complexities that would make it difficult for decision takers to comprehend how the model operates. The model developed in this thesis therefore seeks to combine adherence to fundamental principles of evaluation with a process that remains as clear and transparent as possible.

For these reasons, the model avoids complications that might be introduced by allowing for such macroeconomic effects as migration, the possible distortionary (‘tax deadweight’) effects of taxation or government borrowing and impact on balance of payments through exports or imports. It also opts for a simple treatment of cost, preferring to take the discounted identified project cost as the basis for calculation rather than make ‘real resource cost’ adjustments whose theoretical base is the subject of academic controversy (Hill and Williams, 2005).

Ceteris paribus, the model is specified in terms familiar to evaluators and practitioners in the UK. However, as recorded by Hill and Williams (2005), a number of terms used by practitioners, such as deadweight, displacement and additionality, require more precise definition if they are to used in the construction of quantitative models. Section 11.1 above set out a set of clear definitions.
14.2.4 Territorial boundaries of model

One important element in framing the model is the definition of its territorial boundaries, because the definition of benefits and costs varies with the jurisdictional viewpoint (Schofield, 1989a: 4). The model is intended for use in managing costs and benefits applying to a specific territory, using Wales as an example. This distinction is significant in determining the level of the shadow wage applied.

Territorial definition is also important in treating the profits accruing to firms whose profits are repatriated overseas or elsewhere in the UK. In the case of UK-owned companies with headquarters outside Wales, taxation of profit is treated as a benefit, although one that passes to the UK exchequer. This reflects the political reality of decision-taking on funding of regional economic development. This rule is operationalised for the model simply by regarding gross profit as a benefit. This treatment also avoids the potential difficulty of distinguishing between taxes on profits of foreign- and UK-owned firms respectively.

14.2.5 Treatment of migration

In disregarding migration, the model proposed differs from the Irish appraisal system, which makes provision for additional tax revenue attributable to immigration resulting from projects. Arguably, the need may arise in the future to include a similar explicit counting of the tax benefits of in-migrants, but in the interests of simplicity this aspect is omitted from current treatment. One implication is that the model may therefore underestimate the scale of impacts generated by projects studied. For example, it does not treat as a benefit the multiplier effects of spending by in-migrants or of additional taxation revenue they produce. This goes some way to correct for disregarding the incidence of ‘optimism bias’ (11.12.1), as discussed below (14.3.2.4).
14.2.6 Indicators Generated

The model generates two indicators that can be used together to derive conclusions about the impact of a project or programme, subject to budget constraint faced by the agency:

- The ratio of discounted net benefit divided by discounted net cost; and
- Net Present Value.

Calculation of the ratio of discounted net benefit divided by discounted net cost permits the calculation of ratios that can be readily understood and compared. For example, ceteris paribus, a project showing a positive ratio of 3 (three times as much discounted net benefit as net cost over a given period of time) would generally be favoured over one with a ratio of 2. This process follows the practice recommended by Forfás (2003: 53) for agencies in Ireland and the approach of Swales (1997b: 863) who suggested subtracting costs from benefits, and then dividing the result by the wages of subsidised employees.

However, on its own, the ratio approach ignores the scale of benefits and costs. Projects tend to come in variety of sizes and time-scales, and to be discrete or 'lumpy'. Agencies face a capital allocation problem which may be conceptualised as maximisation of benefit on the lines of linear programming (Hill, 1989: 15-18; 39-40), although real-world behaviour is more complex (11.8 and 13.6). The model therefore also generates measures of Net Present Value (defined in 5.2.1).

The Benefit-Cost and Net Present Value indicators are used to produce two alternative rankings of projects or programmes to guide business planning.

14.3 Definition of key concepts

This section sets out the principal concepts underlying the model, grouped under (1) benefits, (2) costs and (3) items relevant to both benefits and costs, such as the rate of discount.
14.3.1 Benefits

The principal benefits counted are wage income and profit generated as a result of the project and its indirect impacts. Wage income is estimated by observing the number of jobs estimated to be created or safeguarded, and calculating likely total wage income based on median income for the sector as described below in 16.3. The model thus includes the following elements on its benefit side:

14.3.1.1 Jobs directly associated with the project \((J)\)

Jobs enter the model directly, rather than the wage bill as in the Irish model. Estimates of the jobs directly associated with the project may include a number of elements such as Jobs Created, Jobs Safeguarded or jobs associated with projects such as industrial property development. No distinction is made between these different categories of employment, although this refinement could be introduced if required by policy makers (for example, to take account of doubts sometimes expressed regarding the validity of measures of Jobs Safeguarded) (11.6.1).

14.3.1.2 Wages \((w)\)

In most cases, direct jobs will be associated with a single industrial sector, which enables estimation of the direct wage bill. Where a project involves a number of sectors, separate estimates may be made of the number of direct jobs in each sector and the wage bill for each calculated.

Application of a multiplier within the model to both direct jobs and direct profits avoids the entry of separate elements for direct and indirect wages, and direct and indirect profits. In this respect, it differs from the appraisal model recommended for use in the Republic of Ireland, which allows separate estimates of direct and indirect wages to enter the model (Forfás, 2003: 54).

As discussed above in describing the appraisal system in the Republic of Ireland (13.4), not all wages can be counted as a benefit of the project because of the
opportunity cost of labour. The necessary adjustment is made by subtracting the shadow wage part of the total wage bill.

The model uses wage data provided by the Annual Survey of Hours and Earnings (ASHE), a new survey developed to replace the New Earnings Survey (NES) and which is accessed via NOMIS (Office for National Statistics, 2005). The survey details Gross Annual Pay for all employee jobs by Standard Industrial Classification. This has a number of advantages. Information about wages is readily available and updated on a regular basis, whereas estimates of value added are usually calculated indirectly, using input-output tables.

McVittie (2005: 8) maintains that the ASHE estimates of average earnings are the only productivity-related variables for which a ‘reasonably reliable’ order of accuracy is achieved at the regional level: “Using average earnings as a target would allow more reliable tests for changes in growth trends than is the case for GVA-based measures, and would have the added benefit that average earnings are closely related to both productivity and welfare agendas.”

14.3.1.3 Profits directly associated with the project (II)

This element captures benefits from a project that cannot be directly associated with jobs created or safeguarded, an entity described as ‘profit benefit’ in the appraisal system used by the Welsh Development Agency (Williams, 2005: 6). Early analyses of UK regional policy were criticised for failing to include profit alongside labour income as a component of total additional income associated with job creation (Schofield, 1976: 182).

The use of the multiplier within the model itself avoids the need for entry of a separate item for indirect profits, which arise as a result of increased orders for supplier firms who benefit indirectly from the project.

No shadow price is attached to gross profits, in marked contrast to the appraisal system recommended in the Republic of Ireland (Forfás, 2003: 32). The rationale for this approach is set out in 14.3.1.8 below.
Arguably no allowance should be made for counting as benefits part of the profits made by UK- or overseas-owned firms with headquarters outside Wales. As noted above, however, in the case of UK-owned firms, the model treats profits repatriated outside Wales as a benefit.

14.3.1.4a Net grant deadweight ($\theta$)

The wages and profits benefits of a project are modified by an allowance made for net grant deadweight (11.11.1), which reflects the extent to which companies assisted would in any event have developed without grant aid or with lower grant aid. In the model, net grant deadweight is represented by the parameter $\theta$, whose impact is shown by applying $(1 - \theta)$ to the benefits.

14.3.1.4b Displacement ($d$)

In the same way, benefits are adjusted downwards by allowing for displacement, which is defined as the extent to which the benefits of a project are offset by reductions of output or employment elsewhere (HM Treasury, 2003: 54).

It is important to note that displacement elsewhere in the State (e.g. the UK) is not measured within the model. It may be argued that this is inconsistent with the treatment of profits described in 14.3.1.3 above; but this treatment reflects the value judgement inherent in regional economic policy that it is worth the burden of tax distortion in order to generate prosperity in disadvantaged areas. Displacement is represented by the parameter $d$, whose impact is shown by applying $(1 - d)$ to the benefits.

14.3.1.5 Multiplier Effect ($m$)

As discussed in 11.11, the full impact of a project or programme on the regional economy should take account of its various direct, indirect and induced effects.
The model makes use of published Type II multipliers, with values for 70 industry sectors drawn from the Welsh Input-Output tables produced by Cardiff Business School (WERU, 2003). The overall multiplier for all sectors is 1.45, with the highest estimate (for oil processing) at 2.58.

14.3.1.6 Leakage

In the same way as migration effects are disregarded, no specific adjustment is made for leakage effects, defined as “the number or proportion of outputs ... that benefit those outside of the project’s target area or group” (English Partnerships, 2004: 5), and which might be deducted from the gross direct effects. The regional multiplier effect makes allowance for leakage effects: ceteris paribus, the greater the leakage effects, the lower the regional multiplier.

14.3.1.7 Shadow Wage

As argued above (10.2.4), the shadow wage is used to ensure that the model takes into account the real resource cost to the economy of diverting labour to projects from other sectors. One of the departures proposed is extended use of the shadow wage to reflect policy objectives.

First, as in other models, the shadow wage is applied to adjust the benefit scored for regional variations in employment and prosperity. Thus different measures are used for three regions in the Republic of Ireland (Forfás, 2003: 37), where a reference or ‘benchmark’ shadow wage of 100 per cent is set for the most prosperous region, Dublin, with reductions for other regions. This baseline shadow wage applies to projects paying the average industrial wage for the region, and reduced if projects pay wages less than this (Forfás, 2003: 43). A similar conceptual approach is employed in the calculation of targets for gross value added in 22 unitary authority areas in Wales by the Welsh Development Agency (Williams, 2005: 8); in this case, however, rather than deducting the shadow wage, ex ante estimates of Gross Value Added generated by projects are compared with the targets based on local wages to provide measures of variance.
A further adjustment may be made to favour support for projects that, as discussed above (6.6.2), encourage substantial research and development and training. In general empirical literature finds that social rates of return to R&D are substantially above private rates of return (Griffith, 2000: 1). These findings are summarised by Griliches (1992):

“In spite (of many) difficulties, there has been a significant number reasonably well-done studies, all pointing in the same direction: R&D spillovers are present, their magnitude may be quite large, and social rates of return remain significantly above private rates.”

Special support for such projects has been justified because of their high degree of external spillover effects (Jones and Williams, 1998). Such projects would generally be expected to offer higher than average wages, and would therefore in any event tend to score highly in the appraisal process. However, the process favoured in the Republic of Ireland goes further than this in reducing the level of the shadow wage to reflect the magnitude of externalities involved (Barry et al., 2002: 45-46): “Within limits, the shadow wage is increased if projects pay above the average wage and reduced if projects pay below the average wage” (Forfás, 2003: 43).

Further adjustments in the shadow wage could be made to reflect clear gains for defined policy objectives such as sustainability, social inclusion, entrepreneurship or promotion of the Welsh language. Alternatively, appraisal and evaluation techniques may restrict themselves to more strictly economic objectives, while providing for ‘non-economic’ goals elsewhere in the appraisal process. The model could thus be developed the model to take account of these more complex policy objectives, although the thesis restricts its own empirical investigation to more conventional economic analysis.

In constructing shadow wage estimates, a number of potential options and sources of data could be taken up to construct local standard shadow wage rates:

1). One option is be to set the shadow wage at the level of unemployment benefit, i.e. at the same level across Wales (and the rest of the UK). This approach rests on the argument that if you create 100 local jobs you increase income in the area by 100 x
(wage minus unemployment benefit). From the point of view of the local area, this is valid; but from a wider territorial viewpoint (e.g. that of the exchequer), unemployment benefit has to be paid for from taxation.

2). A further possibility would be to use gross value added data available for the 22 unitary authority areas in Wales. This is the option used by the Welsh Development Agency (Williams, 2005) and reflects the greater welfare resulting from establishing jobs with a given level of wages in an area of lower overall income.

3). The thesis however uses average gross weekly earnings for Welsh unitary authorities to construct shadow wage baselines. The rationale for this is the same as that proposed for using wages for calculation of benefit (in 14.3.1.2), namely that the data is readily available and updated regularly, whereas estimates of value added are usually calculated indirectly.

14.3.1.8 Shadow pricing and profit

The model makes no adjustment to reduce Profit Benefit (Π) for example by means of shadow pricing profits, in the same way that the shadow wage is used to reduce the benefit attributed to jobs. This treatment may be viewed as controversial in the light of the discussion of shadow pricing within cost-benefit analysis. Thus, in the Republic of Ireland, such treatment has been regarded as orthodoxy since the 1990s. As well as including increased tax revenues in the calculation, this required appraisers to take account of the fact that grants cost the economy more than pound-for-pound; thus, a premium factor (known as the shadow price of public funds) was applied to all government revenues and expenditures. This was estimated in 1998 to be 1.5, down from over 2 in the mid-1980s (Honohan 1998: 3); a level that may be regarded as surprisingly high.

This general approach is continued as an integral part of the appraisal system advocated by Forfás (2003), which however proposed a further reduction in the level set for the shadow price from 100 per cent (implying that Irish profits were not counted as a direct benefit) to the same level as the adjusted shadow wage. Little justification is offered either for imposing a shadow price on profits in the first place or for the change in the value set:
"In the absence of any alternative plausible approach, it is proposed to set the shadow price of Irish profits equal to the shadow wage to the extent to which highly profitable projects pay relatively high wages, their profits will be given a greater weight in the CBA.” (Forfás, 2002: 32).

Theoretically, it is hard to see what case exists on the benefits side of the equation for applying a shadow price to profits. The argument for setting a shadow wage rests fundamentally on the concept of opportunity cost; that there is a real resource cost borne by the economy involved in switching workers to employment in the aided project. Hence, there is an argument for fixing the shadow wage at 100 per cent at full employment unless there are beneficial externalities, as argued by Barry et al. (2002: 44). No such logic applies to profit benefit generated by an aided project: a firm accruing higher profit does not necessarily mean a loss of profit by other firms. In fact, the estimate of displacement is intended to pick up any such loss that does occur on the benefits side.

Whatever the theoretical argument, however, it must be recognised that this treatment has a major consequences for the value generated by the model. For example, if Project A created 200 jobs in research and development, the wage benefit generated (200 x £27,632.80 = £5,526,560) will be reduced by the amount of the shadow wage; whereas if Project B produces an identical amount of profit benefit no such deduction is made. Furthermore Pickernell (1999: 5-6) suggests that Wales should avoid the situation noted by Roper (1993) in Northern Ireland, where grant aid directly increased company profits, rather than improved growth, competitiveness or employment creation. Another concern is that profit benefit might be lost by leakage, for example in the case of firms headquartered outside Wales or with a substantial number of non-resident shareholders.

14.3.2 Items not counted as benefits

Certain elements arguably meriting inclusion as benefits within the model are excluded.

277
14.3.2.1 Imports and exports and finance

Thus no allowance is made for possible effects on imports, exports and hence balance of payments. This follows the conclusions of Moore and Rhodes (1973) and Schofield (1976: 183) that such effects are likely to be insignificant. For the same reason, following Schofield (1976: 183) and as argued in 10.3.2 in examining the rationale for shadow pricing, no assessment is made of possible losses occasioned by the method of financing policy through taxation or borrowing.

14.3.2.2 Tax revenue

As noted above, tax revenue is not specifically scored as a benefit because wage and profit benefit are counted inclusive of taxation.

14.3.2.3 Loans and equity injections

Consideration might also be given to applying a formula for the market value of any loan or equity injection. The value of the loan or equity injection will be less than the amount lent or invested, reflecting the associated level of risk. In the Republic of Ireland it is suggested that the market value of a £1 injection is 0.50p, implying that the grant equivalent of a £1 equity injection would be 0.50p (Forfás, 2003: 47). Application of the model to equity investments would require a comparable adjustment.

14.3.2.4 Optimism bias

No provision is made within the model to adjust for optimism bias, defined as “a demonstrated, systematic, tendency for project appraisers to be overly optimistic” (HM Treasury, 2003: 29). The recommendation by HM Treasury (2003: 5) is that practitioners in the UK should “adjust for risk and optimism to provide the Base Case”, providing the best estimate of the costs and benefits of each option considered, and consider the impacts of changes in key variables and of different future scenarios on the Base Case.
However, it may be argued that such a process is best carried out by adjusting raw (ex ante) outputs. ‘Burying’ a corrective device within an appraisal model carries the risk of encouraging exaggeration of likely outputs, with the perverse consequence of increasing optimism bias still further. No provision is made for optimism bias within the models prescribed by Forfás (2003) for Irish agencies or by English Partnerships (2004: 25) in calculating additionality of regeneration projects by RDAs in England (English Partnerships, 2004).

14.3.3 Costs

The ‘grant’ item in the cost side of the model is set for the purpose of this thesis at the cost of the project or programme. The rationale for this treatment is that the prime purpose of the model is to provide agencies with a straightforward means of comparing the relative potential and actual benefits and costs of projects undertaken, rather than make the more complex calculations entailed by factoring in e.g. the overhead costs of running the agency, the possible distortionary effect of taxation, or reducing cost by the estimate for net grant deadweight.

This treatment involves a number of drastic steps taken in the interests of simplicity.

14.3.3.1 Administration and compliance costs

First, additional costs, such as those of that administration or compliance, are not included.

14.3.3.2 Shadow price of public funds

Secondly, (arguably the most controversial feature of the model’s treatment of costs) no additional ‘shadow price’ of public funds is imposed on cost, as is the practice recommended in the Republic of Ireland. Thus the value for φ in the formal presentation of the model is set at zero. However, sensitivity analysis is carried out to investigate the potential impact of variation (16.2.4).
14.3.3.3 Transfer payments

Thirdly, no deduction is made to reflect that proportion of cost that is accounted for by net grant deadweight and regarded as a transfer payment rather than real resource cost (Hill and Williams, 2005).

14.3.4 Discount rate

Benefits and costs entering the model are spread over time and converted to a net present value by applying an appropriate rate of discount. Such discounting is a normal feature of evaluative models and is applied to this model notwithstanding various caveats already made (10.3.1).

While this is common practice in the UK, it should not pass without critical reflection, as suggested above (12.3.5). It could be argued that the low value of the discount rate (3.5 per cent) recommended in the UK by HM Treasury (2003) for use by public sector organisations is inappropriate in making ex ante appraisals of economic development projects where the magnitude of the discount adjustment is likely to be well within the error inherent in making estimates of the outcomes. On the other hand, it may equally be suggested that setting a high discount rate would be prejudicial to the needs of future generations (10.3.1).

This thesis follows Treasury guidelines in using 3.5 per cent for discount rates to enable across-the-board comparisons with other UK evaluations.

In the context of UK regional policy, the most appropriate single point in time to choose for discounting the costs and benefits of an individual project is generally considered to be the date of the formal offer of grant or assistance (King, 1990: 56).
14.4 Specification of the Model

This section details the model used in appraisal and evaluation.

The Model

Benefit/Cost Ratio = \( \frac{PVB}{PVC} \)

Net Present Value = \( PVB - PVC \)

\[ PVB = \text{present value of benefits} = \sum_{t}^{t} B_t / (1 + r)^t \]

\[ PVC = \text{present value of costs} = \sum_{t}^{i} C_t / (1 + r)^t \]

Benefits

Benefits \( B \) includes labour benefits in the form of jobs (created, safeguarded or accommodated) and other value added benefit in the form of profits \( \Pi \).

\[ B = (J) \left[ (1 - \theta)(1 - d)(1 + m)(w - w') \right] + (J) \left[ \left( \frac{\Pi}{J} \right) (1 - \theta)(1 - d)(1 + m) \right] \]

\[ \therefore B = ((J)(w - w') + (J)\left( \frac{\Pi}{J} \right) ) (1 - \theta)(1 - d)(1 + m) \]

This is the operational form of the benefits side of the model used in Excel calculations of data.

For simplification, this can be written as

\[ B = J \left( Z(w - w') + \left( \frac{\Pi}{J} \right) Z \right) \]
where $Z = (1-\theta)(1-d)(1+m)$

or $B = JZ[(w-w') + \pi]$

where $w =$ wage per job
$\pi =$ profit per job

$\therefore B = JZ[(w + \pi - w')]$

Costs

$C = (1 + \varphi) \text{Grant} + \text{Admin, Compliance etc}$

N.B. Zero values are set both for Admin, Compliance) and for $\varphi$, the latter setting making no provision for shadow pricing of public funds.

$r$ is the discount rate
$\theta$ is grant deadweight
$\varphi$ is tax deadweight
$w$ is the average market wage
$w'$ is the shadow wage ($w'', w'''$ for further adjusted shadow wage rates)

$J$ is direct jobs associated with the project
$\Pi$ is profit benefit
$d$ is displacement
$m$ is the multiplier effect
14.5 Sources of data

The data used as variables and the sources of parameter values are set out in Table 14.5, with reference to discussion in the text in the left-hand column.

Table 14.5 Variables and sources of parameter values

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jobs directly associated with the project ( J ) (14.3.1.1)</td>
<td>WDA PerMIS system</td>
</tr>
<tr>
<td>Wages ( w ) (14.3.1.2)</td>
<td>Annual Survey of Hours and Earnings (ASHE) Table 5.1a Weekly pay - Gross (£) - For full-time employee jobs: United Kingdom, 2004. NOMIS (Office for National Statistics, 2005).</td>
</tr>
<tr>
<td>Profit benefit directly associated with the project ( D ) (14.3.1.3)</td>
<td>WDA PerMIS system</td>
</tr>
<tr>
<td>Net grant deadweight ( \theta ) (14.3.1.4a)</td>
<td>Mid-Term Evaluation of the EU Objective 1 Programme in Wales (Old Bell 3, 2006: 84)</td>
</tr>
<tr>
<td>Displacement ( d ) (14.3.1.4b)</td>
<td>(Old Bell 3, 2006: 84)</td>
</tr>
<tr>
<td>Multiplier Effect ( m ) (14.3.1.5)</td>
<td>(Old Bell 3, 2006: 84)</td>
</tr>
<tr>
<td>Shadow Wage ( \tau ) (14.3.1.7)</td>
<td>Average gross weekly earnings for Welsh Unitary Authorities</td>
</tr>
<tr>
<td>Rate of discount ( r ) (14.3.4)</td>
<td>HM Treasury (2003)</td>
</tr>
</tbody>
</table>

The estimates for jobs, wages and profit benefit vary with the project, while figures for average gross weekly earnings for Welsh Unitary Authorities used in calculating the Shadow Wage appear in the Excel tables attached.

As explained in 16.2, for experimental purposes, estimates for net grant deadweight, displacement and Type II Multiplier effects are taken from the Mid-Term Evaluation of the EU Objective 1 Programme in Wales (Old Bell 3, 2006: 84). However, it is envisaged that future study could enable the introduction of more accurate, sector-specific measures. Estimates of Type II Multiplier effects are already available by industrial sector.

283
14.6 Limitations of methodology

The main conceptual and operational assumptions on which the model is based are set out in Table 14.6, together with reference to the part of the text that discusses their rationale and whether they are subject to sensitivity analysis.

Table 14.6: Key Assumptions

<table>
<thead>
<tr>
<th>Section</th>
<th>Assumption</th>
<th>Sensitivity Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.2.3</td>
<td>No allowance for macroeconomic effects (migration, distortionary effects of taxation or government borrowing, impact on balance of payments through exports or imports)</td>
<td>None</td>
</tr>
<tr>
<td>14.2.5</td>
<td>Designation of specific territory (Wales)</td>
<td>None</td>
</tr>
<tr>
<td>14.2.4</td>
<td>Treatment of gross profit as benefit (for UK and foreign-owned companies)</td>
<td>None</td>
</tr>
<tr>
<td>14.3.1.1</td>
<td>No distinction between Jobs Created and Jobs Safeguarded</td>
<td>None</td>
</tr>
<tr>
<td>14.3.1.7</td>
<td>Shadow wage</td>
<td>Sensitivity Analysis</td>
</tr>
<tr>
<td>14.3.1.3</td>
<td>No shadow price attached to Gross Profit (on benefit side)</td>
<td>None</td>
</tr>
<tr>
<td>14.3.1.6</td>
<td>Leakage effects incorporated within regional multiplier</td>
<td>Sensitivity Analysis</td>
</tr>
<tr>
<td>14.3.2.3</td>
<td>No specific allowance made for risk</td>
<td>None</td>
</tr>
<tr>
<td>14.3.13</td>
<td>No allowance made for 'shadow price' of public funds (on cost side)</td>
<td>Sensitivity Analysis</td>
</tr>
</tbody>
</table>

Four broad groups of limitations to the methodology are distinguished.

The first group concerns the theoretical basis on which the model is constructed. First, no allowance is made for macroeconomic effects, even within the target area of the project. Secondly, the effectiveness of the model rests on the soundness of the underlying programme theory and, as noted, there are a number of basic criticisms that can be made of the cost-benefit approach. For example, the presumption against the interests of future generations (10.3.1), is less acceptable in an era concerned with climate change and the need for development to be sustainable.
The second group concerns the limitations imposed by the designation of territory within which benefits are scored. As noted (14.2.4), no allowance is made for impacts outside the target area. While this is reasonable in appraising or evaluating a policy geared to assisting specific regions, a wider framework needs to pick up benefits or disbenefits to other areas which may be highly relevant for policy makers engaged in discussions e.g. with the Treasury or in international forums such as the European Commission or World Trade Organisation.

The third group of limitations concerns the accuracy of external parameters used in calculations, such as wage data that helps estimate gross benefit.

Finally, the predictions made by the model depends on the data fed into it, in this case the ex ante estimates of jobs and profit. These estimates are widely considered to be affected by error and more systematically by the phenomenon of optimism bias.

14.7 Ethical issues

No specific ethical issues are identified in the evaluative process required to operate the model.

Output data is collected as a matter of routine by regional development agencies, and provision of such information is generally required of clients as a condition for support. Government departments and agencies frequently publish the methods by which they evaluate projects, although it is not usual to find detailed cost-benefit and other calculations in the public domain.

The data used in this thesis were made available by the Welsh Development Agency (now part of the Welsh Assembly Government’s Department for the Economy and Transport).
14.8 Conclusions

This Chapter set out the form of the model, listing the benefits and costs treated as variables (14.3). The model is specified in 14.4; sources of data are provided in 14.5 and key assumptions listed in 14.6, which also summarised limitations to the methodology.

The model produces two evaluative guides, a measure of Net Present Value and a Benefit-Cost ratio. As argued in 14.2.6, both of these are required to guide agencies' decisions in a situation of budget constraint.
Chapter 15 Analysis of Data

15.1 Introduction

Chapter 15 applies this analysis to a number of projects, with two goals:

(1) To investigate the scope for using such a model to analyse controversial regional policy issues e.g. the use of such subsidies as Regional Selective Assistance, foreign direct investment versus support for indigenous businesses; innovation versus more general business support; and the ‘Gazelles versus Trundlers’ argument – whether public support should be targeted at potentially fast-growth firms rather than all businesses (Hill and Williams, 2003).

(2) To assess briefly whether the economic appraisal process can factor in other duties laid upon agencies, such as the requirement to “develop sustainability appraisal tools and use them in the appraisal of programmes and projects” (WDA, 2005: 11).

This analysis shows how the model responds to changes in estimated variables. As further information becomes available about the various components, such as grant deadweight, displacement and the multiplier, so the model will produce more accurate predictions of benefit.
Section 15.2 specifies high and low scenarios for a number of key variables. Section 15.3 describes the projects appraised using the model and presents the outturns for the two chosen indicators, the Benefit-Cost Ratio and Net Present Value.

As described in detail in Appendix 15, the model draws in WDA output data which can be scored for value added (13.2.5). This data takes the form of jobs (whether created, safeguarded or potential jobs accommodated) and Profit Benefit.

Jobs are converted to value added by grossing up with wage data, and then netting down by subtracting that part of jobs benefit accounted for by the shadow wage. Profit Benefit is counted directly as contributing to benefit.

Both types of benefit are summed and discounted; this produces ‘Total Discounted Benefit’. This figure is then grossed up by the multiplier effect, and netted down for net grant deadweight and displacement. Estimates of project cost are also discounted, and estimates produced of the Benefit-Cost Ratio and Net Present Value, with calculations of High, Low and Median Case estimates.

15.2 High and low case scenarios

Different assumptions regarding “high” and “low” scenarios are built in for a number of variables whose magnitudes are uncertain. In setting values for a high and low case, this thesis selects a range of deadweight, displacement and multiplier factors.

First, to establish base case estimates, it takes as a template the levels chosen by the Mid-Term Evaluation of the EU Objective 1 Programme in Wales (Old Bell 3, 2006: 84). These are in turn based on benchmark estimates from the evaluation of the Yorkshire Humberside Objective 2 programme 1994-1996 (Munday and Williams, 2006) and the most recent UK RSA evaluation (Arup, 2000).

The initial assumptions set are:

- High case (net grant deadweight or gross non-additionality 0.35, displacement 0.23, multiplier 1.25)
- Low case (net grant deadweight 0.48, displacement 0.35, multiplier 1.16)
These estimates are used to generate the outcomes shown in 15.3.

At the second stage, refined estimates of the multiplier effect are introduced to indicate the impact on the outcomes for the three projects studied.

15.3 Projects appraised using the model

This section describes three projects appraised using the model and following the assumptions set out above (15.2). Each project is described, indicating its broad objectives, or 'strategic fit'. 'High' and 'low' case scenarios are adopted, as discussed above (15.2), to reflect the fact that a range of values for net grant deadweight, displacement and multiplier effect should be considered. Application of the model then produces 'high' and 'low' case values for key parameters, and a median value is also generated.

It is conceded that the selection of these projects is not random. As explained in 11.2 above, one limiting factor is the need for ex ante data on outputs that may be measured for their contribution to value added. This reduces the number of projects available more than might be supposed. Not only are obviously 'capacity building' projects excluded, but a large number of projects that should generate jobs and profit are also ruled out (for example, most of those receiving assistance with technology and innovation and financial support from Finance Wales). One possible explanation is that a number of project teams took the understandable option of estimating intermediate outputs (e.g. number of businesses assisted) rather than the riskier course of predicting job and profit creation.
15.3.1  @Wales Digital Media Initiative (1236.2)

The first project studied is that used in the pilot, the @Wales Digital Media Initiative, which has three goals:

- Promotion of the Welsh Digital Media industry;
- Support of digital media clusters; and
- Early-stage incubation for digital media companies.

The @Wales centre is located in Cardiff Bay, and specifically target and encourages entrepreneurship in the sector across Wales (WDA, 2006: Project 1236.2). The centre has secured a BREEAM (or (Building Research Establishment Ltd. Environmental Assessment Method) rating of 'excellent', the highest category of sustainable development, while use of ICT reduces the need for physical travel and encourages local solutions. It was ranked eighth in a global survey of fifty entrants in the Best Science Incubator Award 2006 (Science Alliance, 2007).

In its Key Activity Summary Sheet (KASS) records, the projects lists a number of ‘capacity building’ outputs, some clearly related to capacity (such as assisting community enterprises within the sector), others that might be supposed to enhance value added, including attracting new businesses to Wales. However, it also scores for two specifically value added measures – Jobs Created and Jobs Safeguarded – although not for Profit Benefit.

The cost-benefit outturns produced by the model are shown in Table 15.3.1.1. For this project, discounted benefits for each year for all but one of Wales' 22 Unitary Authorities are available, whereas other projects report for the Objective 1 and non-Objective 1 regions only. Further details of the method of calculation are given in Appendix 15, which describes the process followed in the Excel worksheets.
Table 15.3.1.1: @Wales Digital Media Initiative (Model Outturns)

<table>
<thead>
<tr>
<th>Case</th>
<th>Model Outturns</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benefit-Cost Ratio</td>
<td>Net Present Value (£)</td>
<td></td>
</tr>
<tr>
<td>High Case</td>
<td>4.40</td>
<td>7,987,478</td>
<td></td>
</tr>
<tr>
<td>Low Case</td>
<td>2.76</td>
<td>4,128,810</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>3.58</td>
<td>6,058,144</td>
<td></td>
</tr>
<tr>
<td>High Case (with 1.449 Multiplier)</td>
<td>5.10</td>
<td>9,633,078</td>
<td></td>
</tr>
<tr>
<td>Low Case (with 1.449 Multiplier)</td>
<td>3.44</td>
<td>5,742,728</td>
<td></td>
</tr>
<tr>
<td>Median (with 1.449 Multiplier)</td>
<td>4.27</td>
<td>7,687,903</td>
<td></td>
</tr>
</tbody>
</table>

Subject to the initial experimental assumptions detailed in this chapter (15.2), the @Wales Digital Media Initiative was forecast to generate a positive median Benefit-Cost Ratio of 3.6 and a Net Present Value of £6.1 million. This may be compared with the estimate by Swales (1997a) that, on an ex ante basis, the maximum cost per job that should be paid in grant or subsidy in the UK was 2.45 times the annual wage for a job that would operate unchanged for 10 years.

The next stage is to introduce a refined estimate of the Type II multiplier effect by replacing the experimental estimates (1.25 for the High Case and 1.16 for the Low Case) by the appropriate estimate of Type II Multipliers for 70 sectors from the Welsh Input-Output Project Update (WERU, 2003), which in this case is 1.449 (for the New Media sector). The impact on the outcomes is shown in the lower half of Table 15.3.1.1. The result of this change, an increase of 16 per cent in the multiplier, produces a linear increase of 16 per cent in the Benefit-Cost Ratio, an outcome replicated in the sensitivity analysis carried out in Chapter 16.

These outturns are now compared with the predictions made using the WDA’s own appraisal device (WDA, 2006: Project 1236.2). The WDA’s evaluative system
generates a figure for total gross value added which is derived from (1) estimated employment by industrial sector and (2) profit attributable to the project (in the case of Wales, Profit Benefit is zero). Once value added has been calculated, the project is rated by value added per £ of Agency spend. This allows different projects to be compared, as consideration of absolute value added alone would not take due account of the budget constraint.

A value added target is also set for locations across Wales at Unitary Authority level on the basis of value added per job (Williams, 2005: 11). This makes it possible to see if the value added per £ spend is higher or lower than other projects, and also whether the value added per job is above or below the average value added target for the specific location. The WDA system records:

- Percentage Variance between (1) value added per head and (2) value added per head target; and
- Value added per £ Agency spend.

It then ranks each of the two series, and sums the two rank numbers for each project, so producing a combined ranked order which can be used in business planning (Williams, 2005: 14). The following table details a number of outputs predicted by the WDA’s system for the Wales Digital Media Initiative.

Table 15.3.1.2: Wales Digital Media Initiative (WDA Value Added Predictions)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Outturn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value Added Total (£m)</td>
<td>207.14</td>
</tr>
<tr>
<td>Value Added per £ Agency Spend</td>
<td>46.52</td>
</tr>
<tr>
<td>Value Added per Head (£)</td>
<td>47270</td>
</tr>
<tr>
<td>Value Added per Head Target (£)</td>
<td>34864</td>
</tr>
<tr>
<td>% Variance</td>
<td>35.6</td>
</tr>
</tbody>
</table>

While the theoretical roots of the Agency’s approach are similar (13.2.5), the actual outturns are likely to vary sharply, as no deduction was made by the WDA via the shadow wage to reflect the opportunity cost to the economy of diverting labour from existing employment. This can be seen by comparing the two preceding tables; with
the thesis model predicting a median Net Present Value of £7.7 million compared with the WDA’s estimate of £207.14 million.

While the WDA model provides a valuable means of comparing outputs across projects, the thesis model indicates a closer measure of the real benefit by taking into account the true opportunity cost for the economy as a whole. This has the advantage of enabling decision makers facing a budget constraint to assess the cut-off point at which projects should no longer be promoted.

Clearly therefore introduction of the shadow wage (and with it the key feature of cost-benefit methodology) is an important feature of the approach proposed by this model. An increase in the shadow wage, ceteris paribus, will reduce the value added by the project (Taylor, 2002: 194); and, as the next chapter shows, project outputs are highly sensitive to variations in the shadow wage (15.2.4).

15.3.2 Accelerate Wales Clusters initiative (1080.2)

The second project studied is the Accelerate Wales Clusters initiative, whose aim is to improve the performance of Welsh automotive businesses and other performance engineering sectors (WDA, 2006: Project 1080.2). It is an EU Objective 1 funded programme approved in June 2004 and set to finish in 2007. One feature of the initiative was a dedicated support infrastructure with specialist managers focused on emerging technologies and high growth niche markets and close involvement with universities and colleges. This differed markedly from a previous project also entitled Accelerate Wales, which was designed to promote supply chain linkages.

As with the @Wales Digital Media Initiative, the Accelerate Wales Clusters KASS return claims a number of ‘capacity building’ outputs through group seminars and other network events that encourage knowledge transfer, experience sharing and promoting the release of employees for workplace training. However, it also scores value added outputs for jobs created and safeguarded.

The cost-benefit outturns are shown in Table 15.3.2.1. One notable difference is that in this project, a substantial proportion of the benefit is made up by Profit Benefit
rather than by new or safeguarded jobs. Therefore, the outturns are boosted by the fact that no shadow pricing is imposed on this part of the benefit, as explained in 14.3.1.8.

Table 15.3.2.1: Accelerate Wales Clusters (Model Outturns)

<table>
<thead>
<tr>
<th>Case</th>
<th>Benefit-Cost Ratio</th>
<th>Net Present Value (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Case</td>
<td>3.72</td>
<td>5,487,053</td>
</tr>
<tr>
<td>Low Case</td>
<td>2.33</td>
<td>2,685,146</td>
</tr>
<tr>
<td>Median Case</td>
<td>3.02</td>
<td>4,086,100</td>
</tr>
<tr>
<td>High Case (with 1.484 Multiplier)</td>
<td>4.41</td>
<td>6,893,870</td>
</tr>
<tr>
<td>Low Case (with 1.484 Multiplier)</td>
<td>2.98</td>
<td>4,000,162</td>
</tr>
<tr>
<td>Median (with 1.484 Multiplier)</td>
<td>3.70</td>
<td>5,447,016</td>
</tr>
</tbody>
</table>

The Type II Multiplier (1.484) for the vehicle manufacturing sector (WERU, 2003) is higher than the standard ‘high case’. A calculation using a multiplier of 1.484 produces a new high case values for the Benefit-Cost Ratio (4.41) and Net Present Value (£6.9 million). The median values reflect the midway point between this sectoral ‘high case’ and the ‘low case’. The median forecast is that Accelerate Wales will generate a positive Benefit-Cost Ratio of 3.7 and a Net Present Value of £5.5 million.

Table 15.3.2.2 details outputs predicted by the WDA’s system for Accelerate Wales Clusters. Again, the two methodologies give very different output measures, with the thesis model predicting a median Net Present Value of £5.5 million compared with the WDA’s value added estimate of £60.9 million. Once more, as later sensitivity analysis shows, application of the shadow wage proves to be a major influence on the thesis model outturns.
Table 15.3.2.2: Accelerate Wales Clusters (WDA Value Added Predictions)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Outturn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value Added Total (£m)</td>
<td>60.92</td>
</tr>
<tr>
<td>Value Added per £ Agency Spend</td>
<td>29.39</td>
</tr>
<tr>
<td>Value Added per Head (£)</td>
<td>37623</td>
</tr>
<tr>
<td>Value Added per Head Target (£)</td>
<td>30311</td>
</tr>
<tr>
<td>% Variance</td>
<td>24.1</td>
</tr>
</tbody>
</table>

15.3.3 Business Process Excellence (1026.2)

The Business Process Excellence project is listed under the WDA Business Plan heading of Supply Chains and Networks and is targeted at medium businesses (of between 50-249 employees). Its main focus is on business cost reduction, which is described as having a low priority for companies (‘a Cinderella compared with increased sales’) but is often a highly effective method of improving profit (WDA, 2006: Project 1026.1). The main output was anticipated to be Profit Benefit, with no job outputs forecast; and the ex ante data studied is that recorded for the 2006/07 Business Plan.

The multiplier value of 1.586 is that for ‘other manufacturing’.

Table 15.3.3.1: Business Process Excellence (Model Outturns)

<table>
<thead>
<tr>
<th>Case</th>
<th>Model Outturns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benefit-Cost Ratio</td>
</tr>
<tr>
<td>High Case</td>
<td>4.02</td>
</tr>
<tr>
<td>Low Case</td>
<td>2.52</td>
</tr>
<tr>
<td>Median Case</td>
<td>3.27</td>
</tr>
<tr>
<td>High Case (with 1.586 Multiplier)</td>
<td>4.78</td>
</tr>
<tr>
<td>Low Case (with 1.586 Multiplier)</td>
<td>3.31</td>
</tr>
<tr>
<td>Median Case (using High Case with 1.586 Multiplier)</td>
<td>4.04</td>
</tr>
</tbody>
</table>
Comparison of the model’s predictions with those of the WDA in Table 15.3.3.2 once more shows that the model’s median Net Present Value prediction (£3.6 million) is significantly less than the value added estimate by the WDA’s appraisal. The sensitivity analysis carried out in the following chapter indicates that the thesis model’s use of the shadow wage exercises a strong effect on output measures, attributable to increased employment. However in the case of Business Process Excellence, the main benefit is that generated by Profit Benefit, for which the model makes no downward shadow price adjustment. The next section explores what other factors may account for large variations in outturn.

Table 15.3.3.2: Business Process Excellence (WDA Value Added Predictions)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Outturn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value Added Total (£m)</td>
<td>11.74</td>
</tr>
<tr>
<td>Value Added per £ Agency Spend</td>
<td>18.35</td>
</tr>
<tr>
<td>Value Added per Head (£)</td>
<td>39277</td>
</tr>
<tr>
<td>Value Added per Head Target (£)</td>
<td>34417</td>
</tr>
<tr>
<td>% Variance</td>
<td>14.1</td>
</tr>
</tbody>
</table>

15.4 Project Rank Comparison

As noted above, it is likely that the outturns generated by the model are likely to differ from those of the WDA. Nevertheless, differences in the rankings of project output according to the thesis model as opposed to the WDA’s measure of variance of value added per head would require examination. A priori, as shown in the sensitivity analysis, such a difference would be most likely explained by some combination of:

a) variation in the wage data used in this model and the Agency’s value added data;

b) a difference between the thesis model’s shadow wage and the equivalent WDA measure (percentage variance between the outturn and the target for Value Added per head); or

c) absence from the thesis model of shadow pricing on Profit Benefit.
This Table compares the alternative rankings produced by the model (using the median case benefit-cost ratio and Net Present Value) and the WDA’s appraisal system, using both (a) the measure of variance between (i) value added per head and (ii) value added per job targets set for locations across Wales at Unitary Authority level and (b) the Agency’s Value Added / £ Spend calculation (Williams, 2005: 10).

Turning first to the two sets of WDA estimates, it should be noted that while there is no theoretical reason for the consistent ranking of (1) Variance and (2) Value Added / £ Spend, the similarity between the two is not particularly surprising. Given an agency whose expenditure is constrained by a budget, it is likely that spending on the three projects studied would face similar scrutiny and control. Therefore, the larger the Variance (i.e. the greater the positive percentage difference between (1) value added per head and (2) the value added per head target), the greater the likely the amount of Value Added secured for each pound spent.

A more complex picture emerges with the estimates generated by the thesis model. As Table 15.4 shows, comparison of the thesis model with the WDA appraisal system produces the same rank order of projects for Net Present Value, but a different ranking for Benefit-Cost Ratio.

As previously argued (11.8), an agency with unlimited resources would be expected to seek to maximise Net Present Value, while introduction of a budget constraint increases the significance of the Benefit-Cost Ratio. In Table 15.4, while the @Wales Digital Media Initiative emerges as the most favoured project by NPV, application of the Benefit-Cost Ratio puts it in third place. By contrast, Business Process Excellence is third by NPV but first according to the Ratio.
Table 15.4: Project Rank Comparison

<table>
<thead>
<tr>
<th>Project</th>
<th>Thesis Model Benefit-Cost Ratio Ranking</th>
<th>Thesis Model Net Present Value Ranking</th>
<th>WDA Variance Ranking</th>
<th>WDA Value Added / £ Spend Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>@Wales Digital Media Initiative</td>
<td>3 (3.6)</td>
<td>1 (£6.1m)</td>
<td>1 (35.6)</td>
<td>1 (46.52)</td>
</tr>
<tr>
<td>Accelerate Wales Clusters</td>
<td>2 (3.7)</td>
<td>2 (£5.4m)</td>
<td>2 (24.1)</td>
<td>2 (23.39)</td>
</tr>
<tr>
<td>Business Process Excellence</td>
<td>1 (4.0)</td>
<td>3 (£3.6m)</td>
<td>3 (14.1)</td>
<td>3 (18.35)</td>
</tr>
</tbody>
</table>

As suggested in 11.6, the model is designed for use within a wider impact assessment framework in which agencies can consider a variety of issues alongside the economic outputs forecast by appraisal. Consideration of both NPV and the Benefit-Cost Ratio assists agencies by enabling meaningful comparison of the relative merits of projects. For example, ceteris paribus, a newly established RDA with a relatively small number of possible projects ready to begin would be attracted by the prospect of securing a large return in terms of Net Present Value. By contrast, an agency considering a large number of competing projects and facing a tight budget constraint would be impressed by projects promising a healthy Benefit-Cost Ratio; and the tighter the budget constraint, the more influential the Ratio would become.

By considering both measures, an agency is helped to choose projects that satisfy the objective of maximising social welfare within a particular budget constraint.

15.5 Conclusions

This analysis has observed the response of the model to different project datasets. Section 15.2 set high, low and median case scenarios, and 15.3 analysed outcomes of three WDA projects. The ranking of these outcomes by Benefit-Cost Ratio and Net Present Value was studied in 15.4, and compared with the equivalent measures produced by the WDA’s model.

The thesis model generates different ranking among projects analysed, depending on the selection of indicator (NPV or Benefit-Cost Ratio). Given that the model is
designed for use within a wider impact assessment framework, this feature is useful in that it helps decision takers to compare the relative merits of projects, with particular consideration for their budget constraint.

By comparison with the various methods reviewed in Chapter 13, this makes the proposed model better able to measure the impact of regional projects, in line with the objectives of this thesis (1.2). For agencies faced with taking decisions about a wide range of projects, with widely differing outcomes and time-scales, this gives a considerable advantage.
Chapter 16 Sensitivity Analysis

16.1 Introduction

The next stage in the development of the model is to investigate how it is likely to respond to changes in its underlying assumptions and its various parameters. This is explored by a process of sensitivity analysis, analysis of the effects on an appraisal of varying the projected values of important variables (ODPM, 2004b: 171).

16.2 Sensitivity Analysis and Treatment of Data

The objective of the exercise performed in this thesis is to gain insight into how variations in assumptions and parameters may impact upon the predictions of output made by the model. The analysis presented in this chapter (16.2.4) is carried out on the 'high case' data for the @Wales Digital Media Initiative project, described above (15.3.1). Calculations are made in Excel File 1, with comparable figures also presented for the 'low case' scenario in File 3. Details of how these files are constructed are provided in Appendix 15.

16.2.1 Advantages of Sensitivity Analysis

The technique of sensitivity analysis provides an important method for checking the quality of a given model, as well as a powerful device for checking the robustness and reliability of its analysis (Saltelli et al., 2000). It reveals how the model depends upon the information fed into it, upon its structure and upon the framing assumptions made to build it. Each alternative may be regarded as one of a number of mutually exclusive courses of action that may attain the objectives. If a small change in a parameter results in relatively large changes in the outcomes, the outcomes are said to be sensitive to that parameter. This may mean that the parameter has to be
determined very accurately or that the alternative has to be redesigned for low sensitivity. A similar process may be carried out in respect of the assumptions on which the model rests.

Sensitivity analysis was created to deal simply with uncertainties in the input variables and model parameters. Over the course of time the ideas have been extended to incorporate model conceptual uncertainty, i.e. uncertainty in model structures, assumptions and specifications. Sensitivity analysis may thus be used to increase confidence in the model and its predictions, by providing an understanding of how the model response variables respond to changes in the inputs, whether that involves data used to calibrate it, model structures or other factors, such as the model's independent variables.

Sensitivity analysis is thus closely linked to uncertainty analysis, another computational method, which aims to quantify the overall uncertainty associated with the response as a result of uncertainties in the model input (European Commission IPS, 2005). Uncertainty analysis has the objective of evaluating the uncertainty on the model response as a result of uncertainties both about the model input parameters (parametric uncertainty) and on the model form itself (structural uncertainty). It may be contrasted with contingency analysis, which explores the effect of change in the environment in which the alternatives are to function (a 'what-if' type of analysis, with the what-ifs being external to the alternative) (Web Dictionary of Cybernetics and Systems, 2005).

16.2.2 Methods of sensitivity analysis

There are a number of different methods for carrying out a sensitivity analysis, ranging from simple one-at-a-time methods to global, multivariate methods (Scott, 2003). The information revealed by this process can be valuable, as

- Different level of acceptance (by the decision-makers and stakeholders) may be attached to different types of uncertainty;
- Different uncertainties impact differently on the reliability, robustness and efficiency of the model (European Commission IPS, 2005).
16.2.3 Parameter estimates

The data used in this sensitivity analysis is the same as that set out in Table 14.5, with the same estimates used in the previous chapter (15.2). As noted in section 11.11, however, there is considerable uncertainty about levels of both displacement and grant deadweight. One major area for future research is to make estimates of displacement and deadweight for all sectors analogous those for Type II multipliers made by WERU (2003), which are used in this thesis, although the findings of this chapter suggest that impact of variation in both may be weaker than changes in other variables.

16.2.4 Findings

The findings are set out in Table 16.2.4 and reflect 10 per cent plus or minus variations from the base case in key assumptions and parameters, indicating the consequent percentage change in the two outturns of the model, the Benefit-Cost Ratio and Net Present Value. The table also records the impact of variation in a notional shadow price of public funds, although for the reasons set out in 14.3.3.2, no allowance is made in the model for general shadow pricing of public funds.

This analysis generates a series of estimates, which show wide variation in the point elasticity of the Benefit-Cost Ratio and Net Present Value in respect of different variables and parameters. For example, the Benefit-Cost Ratio shows a unit elasticity relationship to variations in the number of jobs attributed to the project, while there is a virtually unit elasticity in relation to Net Present Value. Similarly, varying the multiplier effect has an identical pair of unit and near unit elasticity effects on the two outcomes, matching the variation in jobs.
Table 16.2.4 Sensitivity analysis on parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cost-Benefit Outturns</th>
<th>Net Present Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benefit-Cost Ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Value</td>
<td>% Change</td>
</tr>
<tr>
<td>Base Case</td>
<td>5.11</td>
<td></td>
</tr>
<tr>
<td>Jobs +10%</td>
<td>5.61</td>
<td>10.0</td>
</tr>
<tr>
<td>Jobs -10%</td>
<td>4.59</td>
<td>-10.0</td>
</tr>
<tr>
<td>Wages (+10%)</td>
<td>7.28</td>
<td>42.80</td>
</tr>
<tr>
<td>Wages -10%</td>
<td>2.92</td>
<td>-42.80</td>
</tr>
<tr>
<td>Shadow wage (+10%)</td>
<td>3.43</td>
<td>-32.80</td>
</tr>
<tr>
<td>Shadow wage (-10%)</td>
<td>6.77</td>
<td>32.80</td>
</tr>
<tr>
<td>Rate of discount (+10% increase)</td>
<td>5.04</td>
<td>-1.28</td>
</tr>
<tr>
<td>Rate of discount (-10% decrease)</td>
<td>5.17</td>
<td>1.304</td>
</tr>
<tr>
<td>Multiplier (+10% increase)</td>
<td>5.61</td>
<td>10.0</td>
</tr>
<tr>
<td>Multiplier (-10% decrease)</td>
<td>4.59</td>
<td>-10.0</td>
</tr>
<tr>
<td>Deadweight (+10% increase)</td>
<td>4.82</td>
<td>-5.40</td>
</tr>
<tr>
<td>Deadweight (-10% decrease)</td>
<td>5.38</td>
<td>5.38</td>
</tr>
<tr>
<td>Displacement (+10%)</td>
<td>4.95</td>
<td>-2.99</td>
</tr>
<tr>
<td>Displacement (-10%)</td>
<td>5.25</td>
<td>2.99</td>
</tr>
<tr>
<td>Shadow Price of Public Funds (+10%)</td>
<td>4.64</td>
<td>-9.09</td>
</tr>
<tr>
<td>Shadow Price of Public Funds (-10%)</td>
<td>5.67</td>
<td>11.11</td>
</tr>
</tbody>
</table>
Diagram 16.2.4.1  Sensitivity analysis: Benefit Cost Ratio and Net Present Value (Impact of Plus 10% change in key variables)

Diagram 16.2.4.2  Sensitivity analysis: Benefit Cost Ratio and Net Present Value (Impact of Minus 10% change in key variables)
By contrast, the table shows that variations in net grant deadweight and displacement have a relatively inelastic impact on the two outcome measures. Thus a 10 per cent increase in net grant deadweight produces a reduction in the Benefit-Cost Ratio of only 5.3 per cent (a point elasticity of -0.53), while Net Present Value is reduced by 6.6 per cent (a point elasticity of -0.66).

There is an even smaller impact by displacement (within the target territory only). A 10 per cent variation in displacement causes the Benefit-Cost Ratio to decline by 3.0 per cent (point elasticity, -0.3), while Net Present Value declines by 3.7 per cent (point elasticity, -0.37).

These findings should be interpreted with caution, given earlier discussion of the difficulties involved in making estimates of both deadweight and displacement (11.11). Clearly the absolute level (rather than changes) of deadweight and displacement is of relevance to the calculation of final outputs. It is interesting however that the model predicts that changes in the two are not likely to exert a decisive impact on project outputs.

Relatively low point elasticities also occur in respect of variation in any Shadow Price of Public Funds. Here the impact on the Benefit-Cost Ratio is near linear for a 10 per cent variation (point elasticity -0.91 for an increase and 1.11 for a decrease). As noted previously (13.5), the model in the form developed in Chapter 14 does not apply general shadow pricing for the use of taxpayers' money to fund regional development. It is reassuring that application of a shadow price would not be influential in determining outcomes.

Two variables stand out for their highly elastic impacts on both the Benefit-Cost Ratio and Net Present Value. One is the level of wages attributable to jobs associated with the project. A 10 per cent increase in wages leads to a 42.8 per cent increase in NPV, a point elasticity of 4.28, far higher than a 10 per cent increase in the number of jobs associated with the project. The reason for this striking difference is not difficult to identify. The benefit of each job associated with the project is reduced by the allowance made for opportunity cost to the economy through the shadow wage, which remains the same whatever the wages attributed to the project. Ceteris paribus,
imposition of the shadow wage on a high wage project creating 50 jobs will therefore result in deduction of the same amount of benefit as in the case of a low wage project creating 50 jobs.

The wage level of jobs created by the project is therefore of great significance to policy makers. The sensitivity analysis suggests that, ceteris paribus, a 10 per cent marginal increase in the wage level will be of greater benefit than a 10 per cent increase in the regional multiplier effect.

The second ‘strong’ variable is the shadow wage itself. As already observed (10.2.4), the higher the level set for the shadow wage, the greater the reduction in overall benefit. Table 16.2.4 shows that a 10 per cent increase in the shadow wage results in a 40.8 per cent decline in Net Present Value, a point elasticity of -4.1.

This powerful response points to two conclusions. The first is that opportunity cost of labour is an important aspect of regional policy, especially (as noted in 9.2), if an economy approaches full employment. Agencies that fail to allow for this real resource cost to the economy are likely to overestimate the benefit of their interventions.

The second conclusion is cautionary. Decision takers need to exercise care in calculating the shadow wage level applied locally. A modest change in the shadow wage leads to significant movements both for the Benefit-Cost Ratio and for Net Present Value. The thesis has already compared the different approaches pursued in the Republic of Ireland and Wales. In Ireland, the latest formulation of the economic model provides for only three regional variants across the territory of the Republic (13.4). In Wales, by contrast, the analogous Gross Value Added targets adopted by the Welsh Assembly Government are based on estimates for each of the 22 unitary authorities (13.2.5).

In deciding the territorial unit for calculation of the shadow wage, agencies need to consider more than availability of data. The large size of the three regions chosen in the Irish Republic may combine areas of prosperity with others of relative deprivation. On the other hand, the residence-based GVA measures chosen for Wales
(Williams, 2005: 11) for small local authorities may ignore the reality of people commuting across boundaries in search of well-paid employment. If the model is used for international comparisons of the success of regional economic projects, the selection of geographic area becomes crucial.

This finding of a high elasticity for both Benefit-Cost Ratio and Net Present Value to variations in key parameters and assumptions corresponds to the conclusions of Swales (1997a: 83), whose model was similarly found to be sensitive to the value of the shadow wage.

16.3 Conclusions

The results presented in Chapter 16 should be treated with caution. The sensitivity analysis shows how the outcomes produced by the model vary as different parameters and assumptions are changed. These results depend heavily on the estimates used, a number of which, as noted previously, are subject to considerable uncertainty.

Nevertheless, the exercise suggests two clear conclusions that are relevant to policymakers:

- The magnitude of the response to the change in project wage underlines the significance of securing high wage projects as opposed, for example, to low wage volume manufacturing and services;
- Evaluators need to be alert to the significant effect of varying the shadow wage, which as suggested may be used to favour projects that produce high externalities or other defined objectives (14.3.1.7).

Systematic application of a soundly based economic model helps direct the attention of decision makers to those features of agencies' interventions that are likely to make a real difference to regional prosperity.
CHAPTER 17 Conclusions

17.1 Introduction

This thesis provides a solution for agencies that seek to optimise their contribution to regional development. It suggests a technique of project appraisal that is both theoretically sound and of practical use, helping policy makers and practitioners to appraise, monitor and evaluate regional policy projects.

This chapter presents the conclusions of that search and suggests ways in which its ambitions may be carried forward. These are presented under the following topic headings:

- Relevance of research question
- Limitations of approach
- Implications for theory
- General implications for policy and practice
- Specific implications for stakeholders in Wales
- Implications for further research

17.8 Final Conclusions
17.2 Relevance of research question

The thesis demonstrated the scope of regional economic policy and the scale of resources devoted to it (1.1). Considerations of public spending and value for money alone therefore underline the importance of the research question, which investigated what model best helps agencies to optimise their contribution to regional development.

However the true significance of the question lies in whether regional economic policy delivers the benefits expected of it. A substantial literature reviewed above has investigated this issue, and considerable effort has also been made to appraise and evaluate policies and individual programmes and projects, for example, the comparison of ‘best practice’ programmes in various countries (Institute of Welsh Affairs, 2002).

By contrast, there is little evidence of any systematic effort to fix on an agreed optimal system of judging the effectiveness of projects. This is surprising given the emphasis laid on evaluation of structural fund programmes by the EU (European Commission, 2000). Indeed, the widespread presence of the structural funds as a driver of regional policy makes imperative an EU-wide evaluative system.

This thesis argues that an effective system of appraisal and evaluation should satisfy two key tests:
- It must be theoretically sound
- It must be clear and practicable

The study suggests that it is possible – at least for a range of projects whose outcomes are measurable in terms of value added – to use an economic model that assists decision takers to optimise their activities subject to a budget constraint. Whether this particular model or some other device is used, there is every reason to make ‘measurable’ projects subject to such a process, both in conducting ex ante appraisal and in carrying out ex post evaluation. The clear conclusion of the thesis is that unless such examination is conducted, some projects will be overvalued while the contribution of others will be
underestimated, resulting in an overall loss of social welfare. Therefore application of an appraisal system sensitive to real benefits and real resources cost will help ensure better decisions on investment of public money.

The thesis underlines the need for accurate and effective appraisal and evaluation, and the advantage of having a clear system to enable comparison across a range of programmes and projects.

17.3 Limitations of approach

The limitations imposed by underlying assumptions, designation of target territory and errors in parameters are reviewed in 14.6.

While advocating use of project level appraisal using value added measures where possible, it is clear that a substantial number of interventions cannot be realistically assessed by such means. As previously recorded, many actions designed to develop the regional economy are either long-term or indirect in nature (8.4.1). How then are policy makers to take account of their impacts and compare the value of investing resources in them rather than in measures that produce results sooner?

This thesis emphasises the need for decisions to be made in the wider context of an evaluative framework approach (11.9) which requires practitioners to take account of a wide range of policy issues. The model proposed plays a strong role within that process – but it does not relieve decision takers of the need to weigh up ‘value added’ outcomes against their social and economic background.

It is emphasised that the model is intended as a decision guide, not a rule; and that it should be used within an overall evaluative framework that ensures consideration of wider issues.
17.4 Implications for theory

It is beyond the remit of this thesis to make judgements on the controversial debate between neoclassical and evolutionary economists about the nature of equilibrium and the working of free markets reviewed in Part 2, and the implications for regional convergence or divergence.

However, there is a strong case for supposing that free markets may not always deliver a spatially optimum pattern of welfare, implying a potential case for advocating government intervention (7.7). Whether or not policy is effective depends in part on the soundness of the underlying programme theory.

One difficulty facing evaluators is the problem of dealing with time. One aspect of this is revealed by the application both of a shadow price of public funds and a discount rate as in the Republic of Ireland (Forfás, 2003). Hill and Williams (2005) argue that this may involve ‘double discounting’, because the magnitude of Net Present Value (NPV) depends in part on the theoretical rate of social opportunity cost (Schofield, 1989a: 88-89) (10.3.2): in other words, the way the discount rate is worked out already allows for the ‘tax deadweight’ cost of distortion. The nature of NPV is central to the concept of cost benefit analysis, and yet its theoretical roots appear to rest on decidedly shaky ground.

A thorough reworking of the theoretical roots of cost benefit analysis should be undertaken to identify the real economic foundations of discounting.

17.5 General implications for policy and practice

One advantage of the model is that it can be applied both to ex ante appraisal and ex post evaluation of projects that deliver measurable direct outputs of jobs and profit within a reasonable time-frame. One clear implication is that much more could be done to further understanding of what works best in certain contexts (14.2.1).
A standard international project appraisal system would be great assistance in achieving such understanding and optimising learning by evaluation.

Despite the current emphasis on Gross Value Added as a target indicator, it is difficult to ensure up-to-date GVA data. The use of wages as a key datum rather than GVA has the advantage of providing a more up-to-date and socially relevant indicator (14.3.1.2).

**Consideration should be given to the use of wages as a key target indicator.**

There is a practical argument for using company turnover as a useful indicator, which is more readily accessible than profit (11.6).

**Consideration should be given to requiring clients to release details of company turnover as a condition of grant support and business advice.**

The thesis concluded that adjustment of the shadow wage provides a practical and transparent way of favouring projects that deliver specific policy goals, e.g. sustainability, social inclusion and restoration of the Welsh language. This however requires adoption of a more rigorous and consistent cost benefit evaluative system.

**Policy makers should consider use of the shadow wage to encourage policy objectives.**

One surprising finding is the inconsistency with which UK practice treats the issue of imposing a ‘tax deadweight’ shadow price for using public funds (13.5). One UK Government Department continues to advocate such a charge despite its abandonment by the Treasury (13.2.1). This thesis argues that no shadow charge should be imposed on regional economic interventions unless applied across the board to all aspects of public spending (13.5).
A consistent approach is required on shadow pricing for using public funds.

Despite the theoretical arguments that underpin the decision not to shadow price profits (14.3.1.8), it would be perverse if the decision guide encouraged projects that resulted in leakage of benefit to outside headquarters or shareholders.

Further consideration needs to be given to the treatment of Profit Benefit, and research should investigate the case for shadow pricing profit in the same way as wages (13.3).

Sensitivity analysis of the model (Chapter 16) pointed to two conclusions:

- The significance of securing high wage projects;
- The strong effect of varying the shadow wage.

The importance of high wage projects is not surprising, and is echoed in the policy objectives of governments and agencies, as in the case of the Republic of Ireland (13.4). Given this context, it is encouraging that the model responds actively to variation in wage benefit.

The second conclusion however is something of an amber warning light to evaluators, especially if, as recommended, the shadow wage is used to favour interventions that generate substantial externalities or fulfil objectives such as sustainability. This topic is considered further in its Welsh context (17.6).

Two alternative rankings for projects are produced by the model, one by Benefit-Cost Ratio and the other by Net Present Value.

Ceteris paribus, the tighter the budget constraint, the more likely it is that the Benefit-Cost Ratio will be the dominant consideration, because agencies will be
concerned to achieve the greatest marginal return on each pound invested in projects.

The thesis discussed the difficulty of using cost-benefit techniques to projects whose outcomes are long-term or indirect in nature, such as capacity building measures (8.4.1). The role of the model developed would be as a decision guide that forms part of a wider evaluative framework (11.9). Agencies' overall evaluative frameworks will attempt to assess the effect of such initiatives; but further work is needed to produce reliable and comparable measures of their effectiveness.

Further work is required to develop the model as a means of assessing the contribution made to social welfare by different regional policies and projects.

17.6 Specific implications for stakeholders in Wales

The development of the model has conclusions of particular significance for policy makers and practitioners in Wales. The thesis demonstrates that the economic model can be applied to a considerable range of projects; namely those that produce measurable outputs in terms of jobs or profit. As previously indicated (13.2.5), it could be used to estimate the effectiveness of a greater number of projects than were scored for value added by the former Welsh Development Agency.

Since April 2006, the Agency has been wound up and its operations incorporated as part of the Welsh Assembly Government. The Agency's decision taking functions, previously vested in its Board and Chief Executive, have now passed to the Assembly Government's Minister for the Economy and Transport. This integration has produced major changes in governance, with major decisions now taken at ministerial level, rather than as previously by the WDA Board, the Agency’s Chief Executive and senior officials at regional and departmental level.
Critics have suggested that future intervention would become more risk averse (Morgan, 2004) although former Assembly Government economic development minister Andrew Davies insisted that decision-making and budget-holding would be delegated to the enlarged department’s officials (Davies, 2006). No detailed analysis of these arrangements has yet been published, and it is difficult to establish whether the process of taking decisions has become more centralised or more averse to bearing risk.

The thesis noted the tendency in the WDA to refrain from appraising the ‘Value Added’ outputs of projects likely to generate jobs and profits within a short or medium term time-frame (15.3). It also noted that Scottish Enterprise no longer use economic models to compare value added by individual projects (Richmond, 2007). This retreat from appraisal at project level is also evident in Wales, where appraisal for the 2008/09 business planning process will be pitched at programme level, reflecting a grouping of projects (on average, 15) (Department for the Economy and Transport, 2007).

Clearly practitioners need to strike a balance between the effort and resources required to appraise at project level and the advantage delivered by accurate forecasting of future benefit. It is not clear, however, whether this retreat is a consequence of the cost and time involved in project level appraisal. The aggregation involved in higher level appraisal runs the risk of losing the essential detailed lessons learnt at the level of individual projects.

While judgement on the overall impact of this reform is beyond the remit of this thesis, it is clear that the case for clarity in appraisal and evaluation is as strong as ever. Indeed, if the process of business planning and decision taking has become more concentrated at the centre, a clear evaluative framework with robust project appraisal appears more rather than less necessary.

Both the analysis of data and the subsequent test of sensitivity indicate that the regional multiplier effect has near unit elasticity impact on the Benefit-Cost Ratio and Net Present Value, while net grant deadweight and displacement have a relatively inelastic impact on
both these outcome measures (16.2.4). By contrast, both the wage level of jobs and the shadow wage have a strong impact.

As suggested above, the model emphasises the case for favouring projects that are likely to generate high wages. To some extent, these conclusions may be viewed as a vindication of the current policy of the Welsh Assembly Government, whose strategy identifies ten “existing growth sectors and clusters within the Welsh economy” for sector-specific business support (Welsh Assembly Government 2006: 1).

However, it may also be argued that the advantage of targeting high wage sectors should also form an integral part of the appraisal system, as well as policy. It remains to be seen whether the current direction of evaluative policy – away from project appraisal and towards higher level appraisal – will assist such integration. In theory, there is no reason why appraisal at a higher, programme level should not be designed to favour high wage projects. It is likely, however, that many programmes will include projects with a considerable span of value-added outputs. In practice, therefore, higher level appraisal of wider programmes may make it more difficult to distinguish and encourage those particular projects that generate high wages.

Regional policy designed to deliver the European Union’s Lisbon Agenda goals of jobs and growth is expanding apace (1.1). Yet there is no evidence of debate among policy makers or practitioners about the merits or otherwise of giving up the scoring of individual projects for their contribution to value added.

**Policy makers should review the wider consequences of appraising at programme rather than project level.**

The strong effect of varying the shadow wage has a number of implications for policy makers. As argued above (12.4), the need to calculate real opportunity cost of labour for the economy as a whole increases as full employment is approached. At the same time, governments pursue various objectives, some of them in essence not ‘economic’.
Thus the shadow wage may be used as an instrument for ensuring that appraisals favour those policies that help deliver such 'non-economic goals' – for example, the 'cross-cutting' objectives of promoting sustainability, social inclusion and the Welsh language adopted by the Welsh Assembly Government (13.4). Such objectives may well be delivered in part through regional economic policy but fall outside its conventional territorial goals.

The implication for stakeholders in Wales is that the shadow wage could become an instrument of considerable significance. Among policy makers and practitioners, fixing shadow wage levels for different regions (and perhaps, for sectors generating major externalities (6.6.2), different industries) could well become a topic for lively debate, carrying political implications. Such debate characterises decisions on interest rates taken by the Monetary Policy Committee of the Bank of England. While variations in shadow wage levels are unlikely to attract the attention that attends the Bank's decisions, agencies would wish the levels set to remain in force for longer periods of time than is the case with interest rates.

**Further debate is needed among policy makers, practitioners and the academic community on the relevance of using shadow wage levels to guide decisions on delivery of government programmes.**

### 17.7 Implications for further research

The thesis makes frequent reference to the difficulty of estimating the true additionality of regional economic policy. Some at least of this difficulty could be relieved by reliable estimates for problematic parameters, especially net grant deadweight and displacement. While values for Type II Multipliers are readily available, grant deadweight and displacement pose a greater challenge, in part because they are likely to vary with industrial sector and regional context. Various calculations have been made (11.11.14), but there is no single source of information available.
The sensitivity analysis carried out in Chapter 16, however, suggests relatively low levels of elasticity of the Benefit-Cost Ratio and Net Present Value to variations in net grant deadweight and displacement; so that absence of a full range of estimates should not be a crucial obstruction to use of the model.

By contrast, the high elasticity found in response to variations in wage levels and in the shadow wage make them prime candidates for the close interest of policy makers and practitioners. Fortunately, neither variation is subject to the problems encountered in estimating net grant deadweight and displacement. As argued above (14.3.1.2), robust data exists for wage levels, while application of the shadow wage to the model is straightforward. This represents a positive finding for policy makers, indicating that, in this area at least, no major conceptual problem stands in the way of adopting a sound system of appraisal and evaluation.

**Research is required into the incidence and magnitudes of net grant deadweight and displacement, and to identify gaps in knowledge, but this should not prevent development of an evaluative economic model for regional projects.**
17.8 Final conclusions

Given the substantial sums spent on regional economic policy, it may be surprising that no generally agreed system has been introduced to compare projects supported by public agencies. Although regional development agencies exchange information and best practice, this co-operation has not to date extended to development of an approved model to guide their endeavours (13.3).

This thesis suggests that it is well worth developing an instrument for use in deciding allocation of resources between competing projects.
Appendix 15 Excel File Guide

Part I: Summary of Worksheets

Seven Excel files are appended to the thesis, six of which detail 'High Case' and 'Low Case' scenarios (16.2) for each of the three projects studied (16.3). The contents of the other file, File 2, are described below.

In the case of one of the projects (@Wales Digital Media Initiative, Worksheet 1 Benefit-Cost), sensitivity analysis has been carried out to analyse how variations in assumptions and parameters impact on the predictions of output made by the model (16.2.4). Worksheet 1 for the other two projects is limited to the main outcomes (of the Benefit-Cost Ratio and Net Present Value) with high, low and median outcomes reproduced for convenience.

File 1 provides the High Case estimates for the @Wales Digital Media Initiative project, which was described above (15.3.1). In Worksheet 1, note that the two 'calculator' cells (A8 and H8) provide read-outs of the Benefit-Cost Ratio (yellow) and Net Present Value (purple).

Worksheet 1 also sets out the calculations carried out for the sensitivity analysis, with results for 10 per cent variations of key assumptions and parameters appearing in columns B-G and K-N respectively. Please note that the sensitivity analysis is performed only for this set of estimates.

File 2 presents the outcomes of File 1 in the form of a Table (reproduced in the thesis as Table 16.2.4), with additional estimates of point elasticity. The outcomes of the sensitivity analysis for the Benefit-Cost Ratio and Net Present Value are also presented in the form of charts.

File 3 shows comparable returns for the 'Low Case' scenario for the @Wales Digital Media Initiative. Note that the File is set to show outcomes for the 'refined' 1.449 Type II Multiplier for the New Media sector from Worksheet 5 Cell C20 (Industrial
Sectors) rather than the initial experimental value of 1.16 available in Worksheet 4 Weekly Pay Cell O105.

File 4 shows the High Case estimates for the Accelerate Wales Clusters initiative (15.3.2). Note that this file is set to show outcomes for the ‘refined’ Type II Multiplier (estimated at 1.484) for the vehicle manufacturing sector.

File 5 shows the Low Case estimates for the Accelerate Wales Clusters initiative, again set to show outcomes for the ‘refined’ Type II Multiplier (estimated at 1.484).

File 6 shows the High Case estimates for the Business Process Excellence project (15.3.3), with the file showing the outcomes for the ‘refined’ Type II Multiplier (estimated at 1.586) for ‘other manufacturing’.

File 7 shows the Low Case estimates for the Business Process Excellence project, again showing outcomes for the ‘other manufacturing’ Type II Multiplier (estimated at 1.586).

Part II Contents of Worksheets

This section describes the contents of the various Worksheets in greater detail. It is generally easier to go straight to Part III to follow through the steps taken in Excel calculations.

Worksheet 1 Benefit-Cost:

This Worksheet produces calculations of the Benefit-Cost Ratio and Net Present Value. The estimates appear in the ‘Calculator’ cells (A8 and H8 respectively), and the figure in this cell will vary with the magnitudes set for the variables listed in Table 14.5. Calculations may be made of the High, Low and Median Case estimates (recorded in Row 34 of the ‘High Case’ file) (15.2 and 15.3). In the attached Excel files, the value for the Type II multiplier effect is set at the ‘refined estimates’ made for the three projects studied, using the estimates for each sector listed in Worksheet 5. This can be readily switched to the experimental estimate described in section
15.2, which is used to produce initial estimates of the Benefit-Cost Ratio and NPV (and again in the sensitivity analysis conducted in Chapter 16).

**Worksheet 2 Discounted Benefit:**

Worksheet 2 presents details of discounted benefit data.

In the case of the @Wales Digital Media initiative, the total benefit for each year for every Unitary Authority is discounted, and the sum total for the whole period appears in Column K. The totals are then in turn grossed up to reflect the Multiplier effect, and netted down for net grant deadweight and displacement. To a considerable extent these two processes tend to cancel each other out.

Other projects, such as the Accelerate Wales Clusters initiative (16.3.2), may report only for the Objective 1 and non-Objective 1 areas. Calculations are made to produce comparable results for these areas in the appropriate worksheets.

**Worksheet 3 Location of Outputs:**

Output data (in the case of the @Wales Media Digital initiative) appears by relevant unitary authorities for the whole period only (not for each year, for which an estimate is made). Outputs counted for value added (jobs and profit benefit) by the WDA are in bold type. To estimate total wages, each jobs figure is then multiplied by the relevant median wages figure in Worksheet 4 (Weekly Pay) (with the results for e.g. Blaenau Gwent in Cell K6).

Welsh mean and median wage figures are available for most industrial sectors from ASHE figures produced by National Statistics (2005) (14.3.1.2), and so the Welsh median figure is used where available rather than constructing estimates from the UK figures. Where there is no median figure for Wales, an estimate is made by multiplying the UK median by the overall Welsh ratio of median / mean for ‘All manufacturing’ or ‘All services’. This is necessary in the case of the computer and related activities sector which is used in appraising the @Wales Digital Media Initiative (calculation in Worksheet 4 K105).
The job benefits for 'Future Years' have been divided equally between 4 years (columns W-Z) and discounted.

Summing all job benefits gives a measure for 'Total Jobs Benefit' (Column K), which is then netted (in Column L) by subtracting the jobs grossed up by the shadow wage figure (which is derived from an estimate of gross weekly earnings for each unitary authority (in Worksheet 6) (multiplied by the number of jobs and by 52 weeks to give an annual shadow wage target).

This job benefit figure (net of the shadow wage) is then added to the figure for Profit Benefit (in this case zero) with results in Column N for each unitary authority. Note that Profit Benefit is not reduced by applying any shadow price. The percentage share of the Welsh total for each unitary authority is then worked out in Column O. Note that the job outputs for 'Prior Years' are discounted at zero per cent and counted as a benefit.

Next an estimate for the annual Unitary Authority total benefit is worked out. The absolute jobs output share of each year is set out (in Row 8 from Column R) (this will have to be changed later to include Profit Benefit). This amount is then divided up by the share of the Welsh total for each unitary authority in Column O. This produces estimates of total benefit for each year for every Unitary Authority. These totals are then ready for discounting in Worksheet 2.

For theoretical reasons (no opportunity cost) no shadow price deduction is made from Profit Benefit (14.3.1.8). This has the apparently perverse result of favouring projects that generate Profit Benefit rather than jobs.

**Worksheet 4 Weekly Pay:**

Annual Gross Pay for all employee jobs (UK and Wales) (mean and median) is taken from the 2005 Annual Survey of Hours and Earnings, Office for National Statistics. Experimental estimates of Net Grant Deadweight, Displacement and Type II multipliers are also inserted (in Row 105, M-N). This estimates may be developed to
provide detailed estimates for each industrial sector; and in the refined estimates made for the three projects studied, the Type II multipliers listed for each sector in Worksheet 5 are used.

**Worksheet 5 Industrial Sectors:**

This provides GDP per FTE and Type II Multipliers for 70 sectors (WERU, 2003). Note the use of the following sectoral multipliers for the refined estimates made for the three projects described in detail in Chapter 15.

- @Wales Digital Media Initiative: New media (including software development)
- Accelerate Wales Clusters: Automotive and other performance engineering sectors
- Business Process Excellence: Cost reduction for medium-sized businesses (50-249 employees)

**Worksheet 6 Gross Weekly Earnings:**

Table of average gross weekly earnings for all Welsh Unitary Authorities.

**Worksheet 7:**

This provides details of annual cost from the WDA’s PerMIS system (11.8). This is discounted in Row 16. Although the model does not impose a general Shadow Price of Public Funds (14.3.3.2), sensitivity analysis is carried out to investigate the potential impact by making 10 per cent plus and minus variations in the project cost figures provided in Row 16 which are discounted to provide the overall cost estimate in Cell K16.
Worksheet 9 Annual Outputs:

This provides additional all-Wales value added output data for e.g. the @Wales Digital Media Initiative to that in Worksheet 3. This sheet provides the percentage of (all-Wales) Jobs Created, Jobs Safeguarded and Profit Benefit in each year. These percentages are then applied to the Total Jobs and Profit Benefit figure in Worksheet 3 (Column N) to make annual estimates of outputs for each unitary authority.

Part III: Changing the Data

When entering new data, these changes are made:

Table 15.A.1 Changing the Data

<table>
<thead>
<tr>
<th>Operation</th>
<th>Worksheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enter new per annum Jobs and Profit Benefit data from PerMIS Outputs</td>
<td>3. Location of Outputs</td>
</tr>
<tr>
<td>Wages – Change (and multiply by 52) either</td>
<td></td>
</tr>
<tr>
<td>(i) Selected sector Welsh median if available (Column H) OR</td>
<td></td>
</tr>
<tr>
<td>(ii) Estimated median (Column J)</td>
<td></td>
</tr>
<tr>
<td>(There should be no need to change the shadow wage which</td>
<td></td>
</tr>
<tr>
<td>remains the same for each Unitary Authority and EU Objective area)</td>
<td></td>
</tr>
<tr>
<td>Insert new cost data from PerMIS Financial</td>
<td>7. Annual Cost</td>
</tr>
<tr>
<td>If available insert new Deadweight, Displacement and Type II Multiplier</td>
<td></td>
</tr>
<tr>
<td>The discount count rate (set at 0.035) may be amended</td>
<td>2. Discounted Benefit (Cell A3)</td>
</tr>
</tbody>
</table>

Table 15A2: Projects Appraised by Cost-Benefit Model

<table>
<thead>
<tr>
<th>KASS Ref.</th>
<th>System Status</th>
<th>Activity</th>
<th>Plan Year</th>
<th>WDA Region /Business Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1236.1</td>
<td>Approved in Business Plan</td>
<td>@Wales Digital Media Initiative</td>
<td>2005/2006</td>
<td>South</td>
</tr>
<tr>
<td>1080.1</td>
<td>Submitted to Business Plan</td>
<td>Accelerate Wales Clusters</td>
<td>2005/2006</td>
<td>South</td>
</tr>
<tr>
<td>1026.2</td>
<td>Approved in Business Plan</td>
<td>Business Process Excellence</td>
<td>2005/2006</td>
<td>South</td>
</tr>
</tbody>
</table>
References


329


Cardiff University Centre for Advanced Studies (1994)
http://www.cf.ac.uk/cass/projects/ctp.html


331


Department for the Economy and Transport (2007) email from to the author from the Strategy and Review Department, 1 November.


335


Econterms (2006b) http://economics.about.com/cs/economicsglossary/g/externality.htm


http://www.sussex.ac.uk/Units/spru/events/ocs/viewabstract.php?id=18


Feldstein M. and Horioka C. (1980) *Domestic Saving and International Capital*


csfinfo.com/documents/publications/conference/LouvainworkshopIIADHegarty.ppt


http://helios.bto.ed.ac.uk/resman/nrm/CBA/revision/tsld007.htm


Studies Association, Reinventing Regions in the Global Economy, University of Pisa. April.


In-the-money.com (2005) [http://www.in-the-money.com/glossarynet/capital.htm](http://www.in-the-money.com/glossarynet/capital.htm)


Institute of Welsh Affairs (2002) *Competing with the World – A study of the economic development strategies of some of the world’s richest regions*.


http://www.regional-studies-assoc.ac.uk/events/lisbon07/papers/lawless.pdf


Netherlands Economic Institute (1972) Kampen en Zwolle na tien jaar stimulering. Rotterdam.


356


http://www.scotland.gov.uk/Topics/Statistics/14713/460


http://www.psy.vanderbilt.edu/faculty/sohee/sledge.htm


http://www.le.ac.uk/economics/research/RePEc/lec/leecon/econ02-11.pdf


http://plato.stanford.edu/contents.html


Welsh Development Agency (WDA) (2001a) Address by Brian Howes, Chairman, Finance Wales, Carmarthen Regional Roadshow. 6 November.

Welsh Development Agency (2001b) Speech by the Chairman of the WDA, Sir David Rowe-Beddoe, at the UK-China Western Provinces Development Conference, Chongqing, China, 12 November.


Excel Files

The CD that accompanies the thesis contains seven Excel files, six of which (File 1 and Files 3-7) detail ‘High Case’ and Low Case’ scenarios (16.2) for each of the three projects studied (16.3).

File 2 presents the outcomes of File 1 in the form of a Table (reproduced in the thesis as Table 16.2.4), with additional estimates of point elasticity.

A guide to the contents of the files and their operation is described in Appendix 15.