

Svein Bråthen

Essays on economic appraisal of transport infrastructure
Examples from aviation and fixed fjord links

Dissertation submitted to the Norwegian University of Science and Technology (NTNU),
Faculty of Civil and Environmental Engineering, Department of Transport Engineering
for the degree of Dr. ing.

Essays on economic appraisal of transport infrastructure
Examples from aviation and fixed fjord links

Svein Bråthen

Dissertation submitted to the Norwegian University of Science and Technology (NTNU),
Faculty of Civil and Environmental Engineering, Department of Transport Engineering
for the degree of Dr. ing.

Preface and acknowledgements

Over the years, the economics of transportation has accumulated a vast amount of research. It is thus with great humbleness that I submit my dissertation within this field. However, I do hope that improvements of the existing economic appraisal methods and insight regarding the way transport infrastructure affects economic growth at the local level, are matters of interest to the research community as well as to practitioners.

This dissertation is submitted to the Norwegian University of Science and Technology, Faculty of Civil and Environmental Engineering, Department of Transport Engineering (NTNU/DTE) for the degree of Dr. ing. It consists of seven essays, of which three papers are published in international journals. The papers form the core of the dissertation. An introductory essay discusses aspects of cost-benefit analysis and macro productivity studies, while the three remaining essays are prologues to each of the published papers. The papers appear as in the journals, which means that the reference list formats may vary.

Many have contributed during this process. First of all, I would like to thank my supervisors. My main supervisor, professor Tore Ö. Sager at NTNU/DTE, has indeed provided encouraging and constructive instructions. My assistant supervisor, professor Arild Hervik at Molde University College has been a never-resting source of inspiration for years. Without his encouraging enthusiasm, confidence and gentle inducement in the first place, I doubt that this dissertation had come into existence.

This work is primarily financed through a scholarship granted by Molde University College (MC), and by financial contributions from Møre Research, Molde (MR). The former principal at MC, professor Anders Dedekam jr. initiated the scheme for financing of Ph.D work that has benefited me and others. His personal support on various occasions has also been important. His successor, Martin Risnes is apparently maintaining these attitudes. The executive team at MR has also contributed financially to a liveable environment.

Paper One, *Economic Appraisal in Norwegian Aviation*, was written with Harald M. Hjelle at MC and Knut S. Eriksen and Marit Killi at Institute of Transport Economics, Oslo, in a fruitful co-operation. The Norwegian Air Traffic and Airport Management (NATAM), represented by senior advisor Knut Fuglum, was the employer of the main research project.

Paper Two, *Strait Crossings and Economic Development. Developing economic impact assessment by means of ex post analyses* was written with my assistant supervisor, Arild Hervik. The Norwegian Research Council (NFR/Protrans) funded this research. Erik Nettet at Ålesund College has been a member of the research team. The paper has benefited considerably from his modelling skills.

Paper Three, *Do Fixed Links Affect Local Industry? A Case Study*, was written on the basis of a study that was funded by the Norwegian Research Council (NFR/Logitrans). As for Paper Two, NFR represented by program executive Torstein L. Garnaas provided financial support.

Senior research adviser James Odeck at the Norwegian Public Roads Directorate has been a creative discussion partner on many of the topics in the dissertation. Senior research officer Sverre Strand at the Institute of Transport Economics has participated actively in discussions on parts of the prologue to Paper One. Associate professor Bo T. Kalsaas at NTNU introduced me with great enthusiasm to the world of qualitative methods, of which one is tentatively applied in Paper Three.

The excellent library at MC has indeed been very helpful. A number of colleagues at MC have contributed in professional discussions and in maintaining a reasonable momentum during this process. In particular, assistant professors Harald M. Hjelle and Öyvind Sunde, and associate professor Arne Lökktangen have provided valuable support.

Supportive family and friends have indeed earned my gratitude. In particular, I want to express my profound gratitude to Vegard, Anne Margrethe and Eldbjörg, for enduring my cognitive dissonance at times when my ambitions have not been met by my abilities.

Molde University College, February 2001

Svein Bråthen

A. Summary and conclusions

A.1 General

This dissertation focuses on economic appraisal of transport infrastructure investments. The main objective is to expand the knowledge about how improvements in transport infrastructure can influence economic efficiency and economic development. Generally, the focus throughout the dissertation is on assessing economic *benefits* of transportation. The cost side of the story is only briefly discussed.

There are several methods in use for economic impact assessment, and cost-benefit analysis (CBA) is the most widespread one. CBA is recommended as the most convenient method for evaluating transport infrastructure investments at the project level, but there is a potential for improvements. *Inter alia*, this potential is acknowledged within the following three areas. First, CBA in transport has traditionally been carried out primarily in the road transport sector. The method has a potential for being adapted and applied in other parts of the transport sector (as well as in other sectors). Second, there are special issues regarding larger changes in the transport system, and it should be clarified how the CBA is able to reveal the true benefits of such changes. Third, there have been doubts about whether the CBA has been able to capture the real effects of infrastructure investments and make clear how local economic development is affected. These three topics are addressed in the published papers that are parts of this dissertation:

1. *Economic Appraisal in Norwegian Aviation*. Published in Journal of Air Transport Management Vol. 6 No. 3 2000.
2. *Strait Crossings and Economic Development. Developing economic impact assessment by means of ex post analyses*. Published in Transport Policy Vol. 4 No. 4 1997.
3. *Do Fixed Links Affect Local Industry? A Case Study*. Published in Journal of Transport Geography Vol. 9 No. 1 2001.

The dissertation also contains discussions on some of the theoretical and empirical issues that are treated in the papers. These discussions take place in prologues preceding each paper. As a generic problem regarding the communication between experts and decision-

makers, better ways of presenting the CBA results have also been requested. This topic is briefly discussed in Paper One.

In the introduction to the dissertation, a brief description of how the economic impact assessments (EIA) are regarded and used by the decision-makers is given. The main part of the introduction is an assessment of the differences between cost-benefit analyses (CBA) based on micro-economic theory, and macro production functions. This is done as to indicate why developments of CBA are regarded as being relevant for improving the EIA procedures at the project level.

Compared with the CBA, the macro studies set out to give a more comprehensive assessment of the economic effects of transport infrastructure. The claimed reason was that productivity effects from spillover and externalities should be captured in a better way. The results from these studies have been questioned, but despite the lack of consensus regarding the macro studies' abilities in this respect, a need for a more thorough assessment of the economic effects of transport infrastructure has been acknowledged.

The rest of this summary consists of three main parts, one for each paper. Each of these parts contains separate summaries of the prologue and the paper, as well as proposals for further research.

A.2 Paper One

The prologue to Paper One presents the underlying theoretical framework, and extends the discussion of several topics like aggregation of individual preferences and the classification of economic impacts. The basis for economic appraisal in Norwegian aviation is cost-benefit analysis (CBA). First, the prologue provides a formal presentation of the general CBA concept including the underlying assumptions from neo-classic microeconomics. Second, various consumer surplus measures are examined, and Hicks' compensated demand and Marshallian demand are discussed. Third, approaches to the aggregation of the individuals' willingness to pay (WTP) and the CBA's role in the decision-making process are briefly commented upon. Fourth, the classification of impacts from aviation investments is discussed. The decision-makers are often confused about what effects that are actually taken into consideration in the CBA. A classification that may improve the understanding of this matter is presented. The prologue ends with a

description of the impact assessment procedure within aviation, and a brief discussion on the use of cost-effectiveness analysis (CEA).

Paper One discusses how the CBA can be applied as an economic impact assessment (EIA) approach for projects within The Norwegian Civil Aviation Authority's (NCAA)¹ jurisdiction. As a rule, measures like e.g. airfield investments entail intangible effects that are not included in a financial analysis, because the pricing of aviation services does rarely reflect changes in economic welfare. Even if there still are problems to be solved, regarding e.g. the value of time in multi-modal passenger transport, the method is applicable to a wide scope of projects in aviation. Two examples illustrate the decision dilemma where unprofitable financial results are achieved on the one hand, and socially profitable economic outcomes on the other.

The NCAA is responsible for its own financial viability through revenues from user charges and other commercial activities. At the same time, the NCAA's role in the

Norwegian transport sector calls for a broader economic appraisal of projects within its jurisdiction. Making priorities subject to a full economic impact analysis instead of a more narrow financial cash flow approach, may significantly alter the portfolio of projects that reaches implementation. The dilemma for the NCAA is obvious: One may very well face situations in which the financial analysis gives recommendations contrary to the results of the economic impact assessment. The significant economic effects from aviation may call for some kind of public-private partnership, to be able to take the welfare effects properly into account.

The EIA method that was developed here is now in use within the NCAA.

Proposals for further research: In Norway as well as in most other countries, there is so far no tradition for optimal peak pricing concerning e.g. slot allocation. If the prices are below marginal costs, then the "excess demand" signal a need for extended capacity that may not be founded in the users' WTP. Thus, a risk of over-investing is present if the traffic forecasts are based on a sub-optimal pricing regime, which is often the case. Little empirical research has been done on the importance of this issue in air transport as well as other modes, and further work would certainly be welcomed.

¹ Since Paper One was published, the NCAA has changed its name to Norwegian Air Traffic and Airport Management (NATAM). For consistency with the published paper, NCAA will be used here.

Empirical evidence on the relevant risk premiums for typical aviation projects is lacking, and should be the subject of further analysis. A search for “market copies” of the different project categories to be able to extract the risk premium from the stock exchange, may be a fruitful approach.

The factor prices on noise that are used in Paper One are assumed to be valid for noise increases as well as decreases. The validity of this assumption is highly questionable and should be subjected to further research. Other factor prices, like the economic costs of aircraft operations, are also encumbered with substantial uncertainty.

A.3 Paper Two

The prologue to Paper Two presents a method for comparing the economic profitability of a ferry operation to that of a fixed link, as a backdrop for illustrating how the results from the *ex post* analyses in Paper Two can be applied. Fixed links versus continued ferry operations should in this analytical framework be considered simultaneously, in order to identify the most profitable fjord crossing concept. The theoretical process is as follows: In the analytical process one has to look for the marginally profitable improvement of ferry operations. One may identify a bundle of profitable measures to improve the design of these operations. If a fixed link (bridge or tunnel) is still profitable when the marginally profitable ferry service improvement is implemented, then the fixed link should be chosen as the most profitable design. If the fixed link turns out to be unprofitable, then a continued but improved ferry operation should be preferred from a social point of view. The inclusion of network effects like congestion into the analyses is also commented upon.

Factor prices like e.g. the value of time and the ferry operating costs are discussed. In most cases, these elements may affect the outcome of the analyses. It is shown that today's use of factor prices is encumbered with substantial uncertainty.

The prologue also assesses various aspects of infrastructure financing. First, the Norwegian scheme for mixed private and public financing is presented both with respect to fjord crossings and toll rings around the major cities. Second, the question of whether private sector financing possibilities has affected the road planning and decision-making process is assessed. These aspects lead to the third point, namely how experiences with financial difficulties through a period of economic recession in the beginning of the 90's influence

the way *risk* is handled today. The local and central governments role as a guarantee for private loans is also considered. Fourth, private and public road financing is discussed theoretically, and research on the economic aspects of financing regime is presented.

Paper Two focuses on ex post analyses of five case studies of larger transport infrastructure projects. The purpose is to assess the measurable economic profitability actually occurring from new infrastructure. It turned out to be substantial differences in benefits ex post compared with the ex ante analyses. The ‘inconvenience cost’ saved by the road users comes in addition to the traditional value of time benefits. The paper also examines the profitability of private versus public funding and carries some considerations on regional impacts of transport infrastructure. A short presentation of a pilot study of two of the fixed links is given to elicit the influence on local industry. Informal results support the theory of forward and backward linkages developed in new economic geography, and that aspect is taken further in Paper Three.

These ex post analyses are at variance with the ex ante appraisals, and have proven to be valuable instruments for informing appraisal practice. The Norwegian Public Roads Directorate has adapted these results for use in economic impact assessments of fixed links.

Proposals for further research: Thorough knowledge of the transport market structure in the region is very important for enhancing the quality of economic impact assessments. Such knowledge may be obtained by further ex ante/ex post analyses of projects already implemented. There are reasons to emphasise the need for further research on this matter, to gain experience from a larger number of infrastructure projects.

A.4 Paper Three

The prologue to Paper Three discusses issues related to transport infrastructure and economic development. In order to establish causal relationships between infrastructure expansion and economic growth, transport infrastructure must act as an intermediate input in private production processes. Actual transport costs for people and goods can thus be reduced, and beneficial network activities can be facilitated. Well-functioning transport infrastructure is therefore supposed to have a positive impact on economic growth by stimulating the production of final goods and services that use the new transport infrastructure capital as a significant input factor. This is achieved either directly through

reduced transport costs or through various network effects like larger labour markets, exploitation of scale economies, and knowledge spillovers.

There are three fundamental premises for economic impact assessment of transport infrastructure. First, the investment should be effective in economic terms. Second, the causal linkage between infrastructure investment and economic growth should be manifest in changes in transport-economic behaviour. This means that economic development depends on the reactions of economic agents like households and firms to the changes in the transport network. Third, transport improvements that affect the above mentioned variables must eventually be transformed into measurable economic benefits. These benefits comprise improved factor productivity, larger output, increased demand for inputs, increased property values, and greater demand for consumer goods. This means that if the transport system can be viewed as a constraint on the economic activity of households and firms, then the possibility of economic development is present if this constraint is relaxed by increased supply of transport infrastructure services. These economic benefits may only be measurable in the long run.

The prologue also discusses the transport infrastructure's role in regional interaction. A network connects economic agents. The selected firms' network activities are one of the main issues in Paper Three. The strength of the economic network affects to what extent the companies are able to do technical and organisational development and innovations. Well-functioning transport and communications are necessary to make such networks viable. The prologue ends with a brief discussion of whether the economic effects connected to the presence of imperfect competition, technical externalities and pecuniary externalities are included in a traditional cost-benefit analysis.

Paper Three focuses on the way transport infrastructure can initiate cumulative economic growth processes at the regional level. The theoretical foundation for the study is external effects as the vehicle for cumulative growth. The 'realist' approach² as a research method is applied for gaining a comprehensive understanding of transport infrastructure investments' effects on local industry. This approach offers a systematic way of building causality models.

² The 'realist approach' seeks to analyse how and why a phenomenon take place, by means of the underlying structures and mechanisms that can establish a contingent relation between events (reductions in transport costs) and causes (e.g. changes in the network of customers and suppliers). The causal link between events and causes can be explained by a set of underlying *structures and mechanisms*.

Several cases were studied on the basis of the realist model. The study focuses on four firms' relations with R&D milieus, contacts with suppliers ('upstream firms') and customers ('downstream firms') together with the size and function of the local labour market. The firms were situated in the vicinity of two of the fixed links (the Ålesund connection and the Askøy bridge) studied in Paper Two. One experience from the interviews was that even if the quantified effects of improved infrastructure appear to be small when considered from recent growth theory, all the respondents claimed that infrastructure is a very important factor in a long run evolution perspective, both for the firms and the local community. One explanation, based on economic rationality, is that the infrastructure serves as a public good that can be demanded at virtually no costs. Because the tolled fixed links replace ferry connections with fares of approximately the same magnitude as the tolls, it is straightforward that continued ferry operations imply higher perceived costs because of ferry fare collection during an 'infinite' time period. After a payback period of 15 to 20 years, the toll collection on the fixed links ceases.

The respondents expressed the importance of the firms' deep roots in a stable local community, where the connection to the mainland gives options for local expansion, using and developing the local labour force through a 'learning by doing' evolutionary process. For most of the firms in this study, the asserted ability to maintain production without moving to the mainland supports the view that reduced transport costs may be of importance to achieve decentralised production structures.

Essentially, the competitive environment and the number and structure of firms, suppliers and customers turned out to be mainly unchanged, virtually leaving no significant effects concerning changes in pecuniary or pure technological externalities affecting these firms. The small tendencies that were observed concerning labour pooling and changes in the competitive milieu represent events that took place only occasionally. The conclusion is that in these cases, a possible local economic development does not seem to be driven by factors included in the growth theories that are used in the paper. One reasonable explanation is that the rather small industrial networks may be below the 'critical mass' needed to benefit from such scale effects. The firms seemed already to be quite well connected ex ante through the relatively frequent ferry services. Essentially, the economic benefits of connecting smaller regional units seem to be captured by a thorough analysis of the changes in transportation costs. Also, the differences between the two fixed links concerning the inconvenience cost (see Paper Two) could not be traced in these firms' adjustments towards the improved infrastructure.

Proposals for further research: The interviews in Paper Three were carried out by means of a structured guide (appendix A in the paper), with the design based on the theoretical framework. The respondents seemed to be well familiar with the issues that were raised in the interviews. They were able to provide clear-cut answers, albeit the network effects were claimed to be modest. There is a case for trying out a similar case study concept on larger and denser industrial networks that are exposed to significant transport infrastructure changes. A thorough understanding of the firms' long-term adjustments to changes in the transportation network may provide useful input to the development of appraisal practices.

A.5 A concluding overview of the dissertation

The dissertation consists of an introduction and three main parts. The introduction indicates why a thorough understanding of economic effects at the project level is necessary, and that there is a potential for improving current appraisal practice. The main idea behind the three main parts is to add knowledge to an established EIA framework within transport, by:

1. Extending the present framework to encompass new fields.
2. Making the analyses more complete by widening the scope of user benefits.
3. Providing a better understanding of how local firms adjust to improved infrastructure.

The first main part is an application of a well-known method for economic appraisal in the transport sector, but applied on a field that has weak traditions for economic appraisals that comprise more than financial cash flow analyses. The prologue presents and discusses the welfare economics that is fundamental in EIA. Paper One applies this theoretical framework to the field of aviation services.

The second main part attempts to bring the traditional CBA framework a bit further by presenting an *ex post* analysis of fixed fjord links that have been operating for a number of years. The prologue to Paper Two examines EIA for fjord crossings as a backdrop for Paper Two. In the paper, the theoretical framework from welfare economics is used to widen the scope of the CBA in this kind of infrastructure investments. This is done by identifying a new component that should be included in the assessment of road users'

benefits of larger infrastructure projects like fixed fjord links. The component is termed the 'inconvenience cost' (IC) of scheduled ferry operations. The conclusions in Paper Two are not universal in the sense that the IC cannot be applied to each and every fixed link. However, the IC is clearly identified in four of the five cases examined here. On a general basis, the results indicate that planners should be aware of the IC during the assessment process. The second main part also discusses how the selection between public and private financing may affect project profitability.

The third main part sets out to provide a deeper understanding of how local firms adjust to improved infrastructure. To what extent the cost-benefit analyses in the transport sector are able to capture the 'true' benefits, from the industry in particular, has been subject to both public and professional debates. The identification of the IC in Paper Two raised curiosity with respect to this matter, even though the IC was identified in connection with the opening of the fixed links, and not as long term effects for the adjacent industry. The prologue to Paper Three assesses the theoretical links between transport infrastructure and economic development. The paper is trying to assess the kinds of underlying structures and mechanisms that affect the local firms' adjustments. Firms adjacent to two of the fixed links that were included in Paper Two have been subjected to in-depth interviews. These firms were selected from a segment of firms that in a pilot survey did report significant network benefits from the new infrastructure with 24h accessibility. Results from the pilot survey are reported in Paper Two.

The conclusion from the in-depth interviews in Paper Three is that in these cases, a possible local economic development does not seem to be driven by factors included in the growth theories that are presented in the prologue and in the paper. One reasonable explanation is that the rather small industrial networks may be below the 'critical mass' needed to benefit from e.g. economies of scale. In assessing this kind of infrastructure, the effects on local industrial networks appear to be small. Hence, this study does not lend support to the view that a broader framework than the CBA is needed to capture the effects of industrial linkages and scale economies. However, the conclusion is drawn from an assessment of firms in small industrial networks, and cannot be generalised to larger networks with extensive interactions between economic agents.

While Paper One applies existing EIA methodology to a "new" area for transport infrastructure assessments, Paper Two and Three attempt to broaden the knowledge regarding how transport infrastructure affects local economic agents, to inform future assessment practice. Paper Two provides a case for an extended view on benefits from fixed fjord

links. Paper Three indicates that when the experiences from Paper Two are carefully incorporated into the cost-benefit framework, then there is no obvious case for investigating further economic benefits in cases where the industrial networks appear to be small.

Contents

Preface and acknowledgements

A. Summary and conclusions	i
A.1 General	i
A.2 Paper One	ii
A.3 Paper Two	iv
A.4 Paper Three	v
A.5 A concluding overview of the dissertation	viii

1. Economic appraisal of transport infrastructure: an introduction	1
1.1 Problem statement and main contents of the dissertation	2
1.2 Transport infrastructure assessment – a potential for improvements	3
1.3 Cost-benefit analysis and economic assessment – conceptual issues	5
1.3.1 Introduction	5
1.3.2 Measuring economic benefits by means of production functions ..	6
1.3.3 Measuring economic benefits by using cost-benefit analysis	9
1.3.4 Conclusions	15
References	17

2. Prologue to Paper One: Economic appraisal in Norwegian aviation	19
2.1 Introduction	20
2.2 The cost-benefit analysis and the measurement of welfare change	21
2.2.1 Formal derivation of the cost-benefit analysis concept	21
2.2.2 Consumer surplus measures in cost-benefit analysis	24
2.3 Cost-benefit analysis and the distribution of utility	28
2.4 Real and reflective impacts	34
2.5 A cost-benefit framework for aviation	40
2.6 Cost effectiveness analysis	46
References	48

3. Paper One: ECONOMIC APPRAISAL IN NORWEGIAN AVIATION	51
1. Introduction	53
2. Cost-benefit analysis (CBA) in the aviation sector – an overview	54
3. Factor prices	58
4. CBA for aviation projects – methodological topics	63
5. CBA in practice: Two examples with simulated numbers	73
6. Conclusions	79
References	80
4. Prologue to Paper Two: Strait crossings and economic development.		
Developing economic impact assessments by means of <i>ex post</i> analyses	83
4.1 Introduction	84
4.2 Ferry operations or fixed links – some comments	85
4.2.1 A model for selecting the optimal mode for fjord crossings	85
4.2.2 A few comments on factor prices	90
4.2.3 Selecting ferries or fixed links – a few conclusions	94
4.3 Toll financing of fjord crossings	95
4.3.1 The Norwegian scheme for financing toll roads	95
4.3.2 The scheme for mixed public and private financing	96
4.3.3 Private financing and the decision-making process	97
4.3.4 Risk assessment in toll financed projects	98
4.3.5 Risk sharing – a practical approach	99
4.3.6 Costs of tolling and costs of public funding	101
References	104
5. Paper Two: STRAIT CROSSINGS AND ECONOMIC DEVELOPMENT.		
Developing economic impact assessments by means of <i>ex post</i> analyses	107
1. Introduction	109
2. The theoretical model	110
3. Some results from <i>ex post</i> analyses	112
4. Private or public funding?	115
5. Regional impacts and <i>ex post</i> analyses	117
6. Conclusions	122
References	123
Appendix 1	125
Appendix 2	133

6. Prologue to Paper Three: Do fixed links affect local industry?	
A Norwegian case study	137
6.1 Introduction	138
6.2 Transport infrastructure and economic development	140
6.3 The role of transport infrastructure in production processes	145
6.4 The transport infrastructure's role in regional interaction	147
6.5 Cost-benefit analyses and economic development	152
6.6 Circular and cumulative causation	155
6.7 The case study approach in Paper Three, and implications of the findings	160
6.8 Concluding remarks	162
References	163
7. Paper Three: DO FIXED LINKS AFFECT LOCAL INDUSTRY?	
A Norwegian case study	167
1. Introduction	169
2. Transport and regional growth – theoretical approaches	171
2.1 The theory of circular and cumulative causation	171
2.2 Endogenous growth theory	173
2.3 The new economic geography: Non-convexities and pecuniary external effects	174
3. Realism as a research methodology for transport systems analysis	176
4. New infrastructure's influence on four firms	183
4.1 Background	183
4.2 The case study	184
5. Summary and conclusions	193
References	195
Appendix 1	198

**Economic appraisal of transport infrastructure:
an introduction**

1. Economic appraisal of transport infrastructure: an introduction

In this chapter, an introduction to the dissertation is provided, with the problem statement and the main contents. A brief description of how the economic impact assessments (EIA) are regarded and used by the decision-makers is given. The main part of the chapter is an assessment of the differences between cost-benefit analyses (CBA) based on micro-economic theory and macro production functions. This is done as to indicate why developments of CBA are regarded as being relevant for improving the EIA procedures at the project level.

Compared with the CBA, the macro studies set out to give a more comprehensive assessment of the economic effects of transport infrastructure. The claimed reason was that productivity effects from spillover and externalities should be captured in a better way. The results from these studies have been questioned, but despite the lack of consensus regarding the macro studies' abilities in this respect, a need for a more thorough assessment of the economic effects of transport infrastructure has been acknowledged. While Paper One applies CBA to a "new" area for economic analyses of transport infrastructure investments, Paper Two and Three attempts to broaden the knowledge regarding how transport infrastructure affect local economic agents, to inform future assessment practice.

1.1 *Problem statement and main contents of the dissertation*

This dissertation focuses on economic appraisal of transport infrastructure investments at the project level. The main objective is to add knowledge regarding how improvements in transport infrastructure can influence economic efficiency and economic development. Generally, the focus throughout the dissertation is on assessing economic *benefits* of transportation. The cost side is only briefly discussed.

There are several methods in use for economic appraisal, among which the cost-benefit analysis (CBA) is the most widely used. For a review of the methods, see EC DG VII (1994). The EC report advocates CBA as the most appropriate method for assessing transport infrastructure investments at the project level, but it also states that there is a potential for improvement. To increase the quality of applied CBA as an instrument for project assessment is the main issue here.

Inter alia, the potential for improvement is acknowledged within the following three areas: First, CBA has traditionally been carried out in the road transport sector. The method has a potential of being adapted and applied in other parts of the transport sector (as well as in other sectors). Second, there are special issues regarding larger changes in the transport system, and it should be clarified how today's CBA is able to reveal the true benefits of such changes. Larger improvements in accessibility like 24h access instead of depending on scheduled transport services can be mentioned as one example. Third, there have been doubts about whether the CBA has been able to capture the real effects with respect to local economic development. These three topics are addressed in the following published papers that are parts of this dissertation:

1. *Economic Appraisal in Norwegian Aviation*. Published in the Journal of Air Transport Management Vol. 6 No. 3 2000.
2. *Strait Crossings and Economic Development. Developing economic impact assessment by means of ex post analyses*. Published in Transport Policy Vol. 4 No. 4 1997.
3. *Do Strait Crossings Affect Local Industry? A Case Study*. Published in the Journal of Transport Geography Vol. 9 No. 1 2001.

Theoretical and empirical issues that are treated in the papers are discussed in prologues preceding each paper. As a generic problem regarding the communication between experts and decision-makers, better ways of presenting the CBA results have also been requested. The 'meta' problem of improving the presentation of CBA results is addressed in Paper One and the adjunct prologue.

1.2 Transport infrastructure assessment - a potential for improvements

For society, economic effects of infrastructure improvements are mainly connected to a perceived reduction in transport costs. A net reduction can, at least theoretically, initiate regional economic growth. The transport costs consist of time costs, vehicle operating costs, tolls/fares and the inconvenience cost. The latter is the cost of being dependent on scheduled transport.¹ The sum of these cost elements is usually termed *generalised costs*.

A common supposition, at least in the decision-making sphere, is that transport infrastructure serves a special role as catalyst for economic activity. This strong belief has materialised in a number of visionary infrastructure investment programs on national and international levels, e.g. the EC's TEN networks. However, by using assessment methods

¹ The inconveniences of scheduled ferry transport are addressed in Paper Two.

such as cost-benefit analyses (CBA), it is not unusual that at least parts of such programs turn out to be economically unprofitable. Historically, economic impact assessment has played a modest role in explaining the composition of national investment programs in Norway (Odeck 1991). Economically unprofitable projects and programs have been approved with reference to the benefits for economic development and hence regional balance.

Among economists, the traditional view is that if the transport costs reductions are properly measured, then the real effects are captured. For a definition of 'real effects', see Chapter Two. This view relies on the assumption that competitive market structures are present. SACTRA (1999) refers to the need for investigating possible bias in the CBA results when this assumption is violated. Previous research, e.g. Mohring and Williamson (1969) has shown, however, that under reasonable assumptions regarding the industrial structure with respect to scale economics, the bias in a well-informed CBA tends to be small. Their conclusion was that there is much more to gain from improving the empirical assessment of road user benefits than to try correcting for the absence of competitive markets. Jansson (1992) regards the viewpoint that the road user benefits underrate the economic effects as belief in 'the grand transportation mystique' that cannot be justified on economic grounds. However, using production functions, studies have indicated that transport infrastructure may have such a catalytic effect on economic development (Aschauer (1989, 1991), Munnell (1990), SOU 1991:22 (1991)). These analyses give a significantly higher economic profitability than the aggregated CBA results indicate. However, these studies have been criticised both on methodological and empirical grounds for overestimating the economic benefits. The discussion of the difference between aggregated production function studies and the CBA will be resumed in Section 1.3.

Within the political sphere in Norway, a rather common viewpoint has been that the CBA's ability to capture economic development effects is limited (Nyborg 1996). Political doubts have evolved especially during the last decade. One interpretation is that the number of economically profitable infrastructure projects in rural areas tends to be decreasing, and the need for legitimation of rural infrastructure projects is thus growing. At the same time, the belief in infrastructure as a way of removing functional obstacles to economic cohesion and growth is as present as ever. The reason for politicians to focus on inadequacies in the CBA, may be twofold. First, the CBA method may be insufficiently reported to the decision-makers. There may e.g. be some confusion about the assessment of real impacts vs. distributive effects. As an example, job creation in the affected regions is very often a matter of concern. In many cases however, relocation of employment may be the main effect, with few new jobs as a net result. Second, there may be shortcomings

in CBA which are intuitively recognised as important by the decision-makers, but not properly dealt with. Such shortcomings may be both theoretical and empirical; some of these are addressed in this dissertation.

1.3 Cost-benefit analysis and economic assessment - conceptual issues.²

This section examines whether macro studies or micro-oriented CBAs are the best to reveal the 'true' benefits of infrastructure investments. After the introduction in 1.3.1, the section is organised as follows: Section 1.3.2 and 1.3.3 deal respectively with the production function and cost-benefit analysis approach to assess the real effects of transport infrastructure investments. Section 1.3.4 concludes.

1.3.1 Introduction

The providers of transport infrastructure face a complex environment for selecting projects, regarding both the assessment procedures and the policy environment. On one hand, there are advocates that regard transport investments as catalysts for industry, even if the economic impact assessment methods should prove the projects to be unprofitable. On the other hand, there may be cases where transport investments prove to be highly profitable, but where scarce funding is a problem.

In many countries, transport infrastructure investments are under pressure from cutbacks in public spending. Different forms of public-private partnerships (PPP) have emerged to maintain and in many cases increase the investment rates. One reason why infrastructure investments are in focus is that they are perceived as important generic instruments for enhancing economic development and regional balance. There are two schools of thought regarding economic assessment of transport infrastructure. Most of the economists that have engaged in transport planning, have regarded the incremental 'social surplus' measured on the actual project as a theoretically convenient welfare measure. The social surplus is usually measured and valued by means of a cost-benefit analysis (CBA). A competing school of thought has however claimed that the benefits from transport investments comprise several impacts that are not captured in a CBA. The viewpoint is that transport infrastructure investments have large 'development effects' as real economic effects, e.g. regional effects emerging from localisation of firms.

² This section is a revised version of the paper '*Productivity studies versus cost-benefit analyses - which is the best indicator for returns to transport investments?*' which was presented at the European Transport Conference 2000, 11-13 september 2000, Homerton College, Cambridge UK.

As mentioned in section 1.2, analyses emerged around 1990 that used a production function approach to produce some striking results regarding high returns to transport investments. In a seminal article, Aschauer (1989) used macro product functions to estimate an apparently strong and positive relation between infrastructure investments and economic cycles, proposing that the results could well support an expansive infrastructure policy. The macro productivity studies approach was chosen to assess the 'total' benefits including the spillover effects that were not assumed to be captured by the CBA. The backbone of Aschauer's analytic approach was its alleged ability to measure the spillover benefits that were not assumed to be captured by project specific cost-benefit analyses. This work, and other studies along the same path, has later been criticised for data weaknesses and collinearity problems (see e.g. Hulten and Schwab (1993)).

At the same time, other authors (e.g. Forkenbrock and Foster 1990) have advocated project specific cost-benefit analyses (CBA) as the policy relevant analytical approach. The scarcity of public funds and the increasing role for PPP call for robust assessment methods. The reason is not as much an economic as a financial one: If the road users' willingness to pay for improved infrastructure is overstated in the assessments, then the investors in a PPP may face financial deficits.

1.3.2 Measuring economic benefits by means of production functions

The political concern and criticism regarding the CBAs ability to capture all the economic costs and benefits from improved transport have especially been directed towards the long-term effects for the prosperity of national industries. One has attempted to develop ways of dealing with the assumed spillover from transport investments to the rest of the economy, partly to be able to provide scientific answers to this criticism. It has been presupposed that the infrastructure's contribution to productivity can be captured through the estimation of macro production functions. Most studies take a production function like this one as the point of departure:

$$Q = f(K,L,G)$$

As a Cobb-Douglas function:

$$Q = L^a K^b G^c$$

where Q = GNP or GRP (output)
 L = Labour input
 K = Private sector capital inputs
 G = Public sector capital inputs (can be specified in segments, e.g.
 G_{TI} = transport infrastructure capital stock)
 a, b, c = Coefficients (to be estimated as elasticities)

Aschauer's macro time series regressions found c to be positive in the range of 0.38 to 0.56, measured on public capital and not transport infrastructure capital in particular. Gramlich (1994) used the result to show that, if correct, the marginal productivity of public infrastructure capital turned out to be 'pretty stratospheric' with an annual return of 100 per cent per annum or more, which seems implausible.

The most important elements in the critique of these macro productivity models are:

- There may be problems with the breakdown of public capital into different kinds of transport infrastructure. The definition of the term 'infrastructure capital' is in itself a matter for discussion. The 'reconstruction cost approach' uses the construction value, while the 'market value approach' will use the discounted returns in terms of user benefits. These approaches are principally different. As an example, a scarcely used high quality infrastructure will probably have a lower value in the 'market' than in the 'reconstruction' approach. However, the 'reconstruction costs' approach is relevant when assessing the productivity effects of investments actually made.
- A lot of the studies that are carried out comprise 'transport infrastructure' as such, with limited validity regarding mode specific productivity effects.
- It is a complex matter to include all elements that influence private sector productivity. Not only has investment activity in other sectors like education and health care an influence, but also the composition of the consumption bundle may have significant impact on productivity.
- Generally, there is a strong correlation between infrastructure investments and other factors that influence the economic development. A regression analysis of a national or regional production function uses a measure of output (or output divided by inputs) as the dependent variable, with a measure of the public capital stock as an independent variable (Boarnet 1997a). However, the public capital variable must be exogenous to the economy for the analysis to become valid. The problem is that infrastructure investments may follow the output or be influenced by the nation's or

the region's prosperity, i.e. when the productivity increases, so does national income, which in turn makes it possible to augment the infrastructure investments. The direction of causality can thus be ambiguous. The case might just as well be that productivity growth has given room for infrastructure investments, and not vice versa. Hulten and Schwab (1993) among others support the view that the Aschauer-like studies cannot verify a high productive effect from new infrastructure investments.

- Macro productivity studies are in nature *ex post* analyses, based on aggregated historical data. A diminishing marginal utility of transport infrastructure may make it difficult to extrapolate from historical evidence into the future.
- The results from these macro analyses have the *actual* investment policy as the point of departure. Odeck (1996) has shown that the investment policy is not necessarily based on the recommendations from CBA, but to a greater extent on distributive considerations like regional balance. Given that significant and reliable productivity effects from transport infrastructure can be identified, these effects are based on the *actual* infrastructure policy, and not the *optimal* one with respect to economic profitability.
- Unless the analyses are based on productivity from large infrastructure projects at the regional level, the results from such analyses are not applicable for individual infrastructure projects.

Boarnet (1997a) states two reasons to be cautious when discussing the design of infrastructure policy from macro productivity studies. First, the production function approach based on aggregate data cannot give any insight into the profitability of particular projects. Broad programs to implement infrastructure packages based on the recommendations from macro analyses would most probably include unprofitable projects. Second, the validity of the macro analyses depends on present and historical evidence regarding the *use* of the infrastructure capital stock, i.e. the services that the capital stock actually has been producing. Because almost all research available on public capital has measured only the value of the *stock* rather than the *services* produced, the policy recommendations are likely to overlook the possibilities for achieving a more efficient use of the existing capital stock. The results from estimation of production functions for a sample of California counties (Boarnet (1997b) indicate that productivity gain from more efficient use of the transport infrastructure (congestion reduction) will be more efficient than increasing the infrastructure capital stock.

Jansson (1993) discusses the importance of the industrial structure in the area, especially when doing cross-sectional analyses over several regions of the link between transport infrastructure and economic growth. Based on Walters (1968), Jansson emphasises one assumption that should be met before any causal relationship between transport infrastructure and regional productivity can be established statistically. The aggregated production in the different regions must be from the same kind of production, or at least be a collection of a comparable set of products that can be expected to use infrastructure as a production factor in the same way. The production technology should also be comparable. If different products or technologies appear, a cross-sectional study will provide information about how different products or technologies affect the regional productivity, but be inconclusive about the productivity effects of transport infrastructure.

Infrastructure services are a production factor for the industries in the influence area, and thus it should be possible to observe the value of the infrastructure through the industry's willingness to pay (WTP) for these services. Accordingly, infrastructure services are input in leisure and commuting trips. In micro based analyses (CBA), one attempts to measure the WTP for both business trips and private trips. At first sight, one should expect that $\partial K_P / \partial G_{TI}$ estimated from adequate data by using the equation above would correspond to the discounted net benefits of the 'average' marginal infrastructure project, calculated by means of CBA (Jansson 1992). The next section will discuss if this is necessarily the case.

Thus far, a conclusion may be that the macro approach tends to capture the productivity effects of transport infrastructure that affect economic development at a national or regional level, but that the statistical problems are severe. Theoretically, the difference in benefits from these studies and the aggregated net benefits from CBA of the projects that are carried out during the same time period, is supposed to represent the spillover benefits that are not captured on the micro (project) level. One would expect productivity spillovers if new projects improve accessibility throughout a large network. From a literature review, Boarnet (1997a) found that in the United States, the network externalities are likely to be small, at least after 1970. The next section will discuss the cost-benefit framework compared to the macro productivity studies.

1.3.3 Measuring economic benefits by using cost-benefit analyses

Cost-benefit analysis (CBA) has been the most common method to be used for assessing the benefits of transport investments at the project level. For a general introduction, see Nas (1996) and for applications to transport, see e.g. Button (1993) and Paper One in this

dissertation. Under assumptions of competitive markets, the point of departure is that the benefits of transport infrastructure can be measured 'on the road' by the changes in time costs, fares and vehicle operating costs. These cost elements express how changes in the transport system affect the users' use of real resources. Using correct demand elasticities, the changes in traffic volumes should reflect the changes in user costs. The valuation of goods done by the market's willingness to pay expressed by prices in competitive markets, is an analogy. The external time costs (during congestion) and environmental costs are also taken into consideration when the benefits are compared to the investments and operating costs of the infrastructure. In the absence of competitive markets things get more complicated. This will be briefly discussed below.

To illustrate why there are reasons to expect different results from CBA and macro productivity studies, imagine that a transport investment is carried out that alters the economic equilibrium for the firms or persons affected by this investment. Figure 1.1 illustrates a situation with the existing equilibrium (point A) and a set of possible new equilibria (points C, D and E) resulting from a decrease in transport costs, in this case rotating the individual's budget line from XY towards XZ (Mohring 1976).

A reduction in transport costs creates an increase in the individuals' real income. The initial real income is illustrated by the budget line XY. Every point on the budget line constitutes a possible combination of leisure and manufactured goods. Maximum leisure time is given by the intercept with the vertical Y axis (e.g. 24 hours minus 7 hours sleep). The *ex ante* equilibrium is given by point A where the budget line is tangent to the

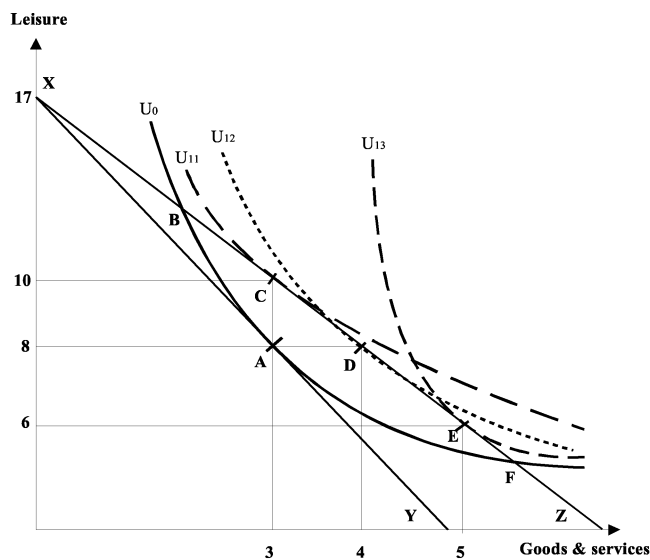


Figure 1.1 A set of responses to changes in transport costs

indifference curve U_0 , where 8 hours of leisure and 3 units of goods/services are chosen. The real effect from reduced transport costs is given by the movement of the budget line towards XZ. The budget line rotates around X, which is the maximum leisure time constraint. The maximum amount of available goods and services is however increasing with augmented real disposable income.

The possible area of new equilibria between the budget line XZ and a new indifference curve is limited by BF. Outside this area, the indifference curves will cross, which is incompatible with utility maximising. However, the indifference curves may have different shapes. If the individual's utility function implies that reduced transport costs entail an adjustment towards point C, then leisure increases to 10 hours, while the demand for goods and services remains unchanged. If the new equilibrium turns out to be in point D, then the increase in real income will be used to increase the demand for goods and services to four units, while the amount of leisure will remain unchanged. If the adjustment is in point E, then leisure will actually decrease by two hours and the demand for goods and services will increase to 5.

The observations here are that the shape of the utility functions determines to what extent reduced transport costs can result in augmented GNP. The equilibrium in point C will hardly entail any changes in the demand for goods and services, while an adjustment towards point D or E will result in increased GNP. These relations can be illustrated by a traditional price/quantity relationship as in Figure 1.2.

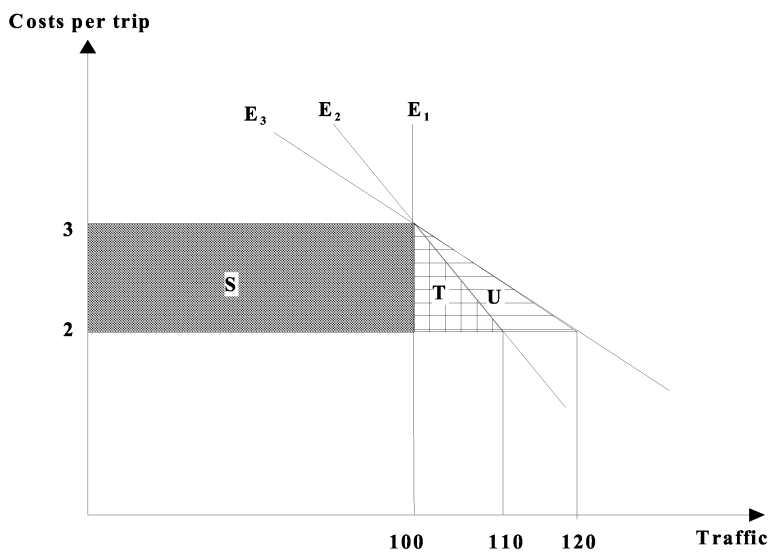


Figure 1.2 Changes in consumer surplus from reduced transport costs

Now if the new equilibrium turns out to be point C in Figure 1.1, the increase in leisure time does not result in any increase in travel activity (no extra goods or services demanded, not even transport services). The change in consumer's surplus (ΔCS) can be identified as the area S in Figure 1.2. Thus, an increase in real benefits (leisure time) appears through the ΔCS , but no growth in GNP can be identified. An equilibrium point D in Figure 1.2 is supposed to give an increase in traffic volume from 100 to 110, and correspondingly a ΔCS equal to S+T. The equilibrium point E corresponds to a traffic volume of 120, and a ΔCS corresponding to S+T+U.

It can easily be seen that there are principal differences in measuring the benefits by changes in consumer's surplus and by changes in GNP by means of production functions. If reduced transport costs entail increased leisure time only, consumer's surplus will be enhanced, but the GNP will not be augmented (prospective changes in labour productivity from increased leisure time is ignored here!). Generally, the more inelastic demand for goods and services, the higher is the possibility that the ΔCS will surpass the ΔGNP .

Mohring (1976) and Mohring and Williamson (1969) discussed whether the entire economic benefits from transport investments can be identified through the road user benefits, or if there are reasons to expect spillover effects that have to be captured by macro production functions. Their conclusion was that the direct effects from both reduced transport costs for existing transport flows and induced effects from reorganisation can be captured by the area under the Hicksian traffic demand curves. The income effects in the transport markets determine to what extent the compensated demand deviates from actual (Marshallian) demand³.

Mohring and Williamson (1969) claimed that for most practical purposes adequate traffic forecasting and proper estimates of the changes in transport costs⁴ would capture the essential parts of the direct economic benefits. This meant that the essential effects were captured without taking the induced traffic from reorganisation into consideration. The scale effects are relevant because it can be assumed that increasing economies of scale entails an increased ratio of transport costs to total costs. Hence the reduced transport costs' impact on total costs increases with increases in manufacturing scale economies. By simulating various levels of such scale effects, they found that a 25% reduction in transport costs captured between 88% and 93% of the transport improvement benefits, measured as benefits for the existing transport activities. According to these authors

³ For a discussion on the properties of Hicksian demand and Marshallian demand, see Prologue One.

⁴ Changes in transport costs mean changes in the road users' generalised travel cost, i.e. time costs, vehicle operating costs, tolls and other cost elements like the inconvenience cost (IC). For an assessment of IC, see Paper Two.

opinions, and if these simulations sustain generalisation to the real world, then a modest share of the benefits would be lost if the benefits for the induced traffic were not quantified. However, the induced traffic in selected fixed links in Norway have recently exceeded 50% (Sandvik 2000), which indicate that real world situations may call for careful assessment of the benefits from induced traffic, particularly in cases where the transport costs are significantly affected. Such assessments are a part of the current appraisal practice.

SACTRA (1999) focuses on whether the presence of external effects and imperfect competition allow CBA to capture all real effects of an infrastructure investment or not. These matters will be briefly discussed in turn with respect to how they may affect the quality of the CBA. External effects represent use of real resources, and may entail spillover effects from the use of the infrastructure. *Negative externalities* are counterproductive in the sense that costs from an action adversely affect third parties' utility functions or production functions. Because these externalities occur due to the fact that the full costs are not covered through prices, a sub optimal use of the good will result. However, the effects of such external effects are to a certain extent included in the CBA by the use of shadow prices. Examples are external time costs in congested hours, noise and various local and global emissions. There may be imperfections regarding the determination of both prices and quantities, but these matters are subject to continuous refinements.

Positive externalities are real resources that are initially used and paid for by one agent. However subsequently, the use of these resources entails positive effects for other agents as well. Such goods are often characterised as *pure public goods*, meaning that their nature of non-rivalry and non-excludability makes it difficult to charge other users. Hence, the true benefits will exceed the WTP for the initial use, resulting in underproduction of the good compared to the socially optimal level. Such situations may be cases for public provision of goods, like e.g. R&D programs. If the transport systems entail such effects, e.g. spreading of knowledge and hence augmenting human capital by means of a physical network, then a case for governmental intervention may be present in terms of increased public spending. Theoretically, such effects can be identified through a real increase in the value of time for the road users. The effects can be included in a CBA if the real increase in long term user benefits is identified, and if forecasts that are able to capture the resulting traffic growth effects are applied.

Positive external effects are normally not included in the CBA because of significant identification problems. Paper Three in this dissertation assess the interactions within

local industrial networks that had benefited from improved transport infrastructure, and no evidence was found indicating that the positive externalities were significant.

If *imperfect competition* is present, then transport infrastructure improvements can contribute to a more competitive environment because interaction among firms is facilitated. This may change the local price vector. Intuitively, we will now describe how this may influence the cost-benefit analysis. First, the transport improvement will have a 'first order' impact on the transport costs, which normally is an important parameter in the traffic forecasts. Second, both transport system costs and transport user costs may be reduced because of increased competition. Third, if the industrial environment becomes more competitive, then reduced prices on goods and services in the area will contribute to a reduction of 'dead weight loss' and hence increase the social surplus. Thus, if enhanced competition is a result of the infrastructure investment, then there are reasons to believe that CBA will underreport the true benefits. However, to assess the effects of imperfect competition require a vast amount of information regarding price/quantity vectors *ex ante* and *ex post*. After having examined recent studies on this topic, SACTRA (1999) concludes (section 4.75 and 4.76):

'It seems unlikely that we should dismiss imperfect competition as *always* unimportant in cost-benefit analysis. The *prima facie* evidence suggests that price-cost margins are, on average, likely to be too large to allow this, and in addition, dropping the assumptions of perfect competition and constant returns to scale opens up a wide range of further reasons why the ratio of total economic benefits to transport benefits may, on occasions, be quite far from one. Nevertheless, *typically*, although the assumption of perfect competition in non-transport-using sector introduces errors, they may be quite small relative to those involved in estimating conventional transport benefits.in the present state of knowledge, we are nowhere near to being able to offer precise estimates.' (of the deviation between the benefits from CBA and the benefits when imperfect competition are taken into consideration, *this author's remark*)

SACTRA concludes that there is no empirical evidence that can legitimate an inclusion of additional benefits of this kind, and that CBA is the best there is for impact assessments of transport infrastructure at the project level. However, SACTRA recommends the development of general equilibrium models (GEMs) to assess the problem with price changes and benefit calculations in situations with imperfect competition. A brief theoretical assessment of CBA in situations with imperfect competition is given in the Prologue to Paper Three, section 6.4.

1.3.4 Conclusions

There are two schools of thought regarding economic assessment of transport infrastructure. The social surplus is usually measured and valued by means of a cost-benefit analysis (CBA). A competing school of thought has however claimed that the benefit from transport investments comprises several impacts that are not captured in a CBA. The viewpoint is that transport infrastructure investments have large 'development effects' as real economic effects, e.g. regional effects emerging from localisation of firms.

There are principal differences in measuring the benefits of transport infrastructure by changes in social surplus by means of cost-benefit analysis (CBA), and by changes in gross national (or regional) product (GNP) by means of production functions. For example, if reduced transport costs entails increased leisure time only, consumer's surplus (CS) will be enhanced, but the GNP will not be augmented. Generally, the more inelastic demand for goods and services, the higher is the possibility that the ΔCS will surpass the ΔGNP .

Negative external effects are to a certain extent included in the CBA by using shadow prices. *Positive external effects* from 'facilitated networking activities' like common labour markets are normally not included in the CBA because of significant identification problems.

If *imperfect competition* is present, then transport infrastructure improvements can contribute to a more competitive environment because interaction among firms is facilitated. SACTRA recommends the development of general equilibrium models (GEM) to assess the problem with benefit calculations in situations with imperfect competition. Nevertheless, *typically*, although the assumption of perfect competition in non-transport-using sector introduces errors, they may be quite small relative to those involved in estimating conventional transport benefits (SACTRA 1999).

Despite these weaknesses, CBA seems to be most convenient for impact assessments of transport infrastructure at the project level. There are two reasons to be cautious when discussing the design of infrastructure policy from macro productivity studies. First, the production function approach based on aggregate data cannot give any insight into the profitability of particular projects. Second, the macro analyses is based on historical evidence regarding the use of the infrastructure capital stock, i.e. the services that the capital stock actually have been producing. Because almost all research available on public capital has measured only the value of the stock rather than the services produced,

the policy recommendations are likely to overlook the possibilities of achieving a more efficient use of the existing capital stock.

As opposed to the aggregated and historical macro productivity studies, CBA is traditionally carried out *ex ante*, based on expected future development and WTP in specific transport markets. These analyses may of course also be biased. A useful approach to improvements in the CBA method is to carry out *ex post* CBA on implemented transport projects. The *ex post* studies can then be compared to the *ex ante* analyses, making it possible to identify and correct biases. This may be a fruitful way of improving future CBA practice. *Ex post* analyses were also one of the recommendations from the EC's 4th framework programme (EC DGVII 1994). An example of such *ex post* analyses can be found in Paper Two in this dissertation.

References

- Aschauer, D. A. (1989) Is public expenditure productive? *Journal of Monetary Economics* **23** (2) 177-200.
- Aschauer, D. A. (1991) *Public investment and private sector growth: the economic benefit of reducing America's "Third deficit"*. Economic Policy Institute, Washington DC.
- Boarnet, M. G. (1997a) Highways and economic productivity: interpreting recent evidence. *Journal of Planning Literature* **11** (4) 476-486.
- Boarnet, M. G. (1997b) Infrastructure services and the productivity of public capital: the case of streets and highways. *National Tax Journal* **50** (1) 39-57.
- Button, K. J. (1993) *Transport economics, 2nd ed.* Edward Elgar, Aldershot, UK.
- EC DGVII (1994) *Cost benefit and multi-criteria analysis for new road construction*. Report no. EURET/385/94, European Commission DGVII, Brussels.
- Forkenbrock, D. J. and Foster N. J. S. (1990) Economic benefits of a corridor highway investment. *Transportation Research* **24A** (4) 303-312.
- Gramlich, E. M. (1994): Infrastructure investment: a review essay. *Journal of Economic Literature* **32** (3) 1176-1196.
- Hulten, C.R. and Schwab, R. M. (1993) Infrastructure spending: where do we go from here? *National Tax Journal* **46** (3) 261-273.
- Jansson, J. O. (1992) 'Kan ett makroperspektiv tillföra allokeringspolitiken visavi transportsektorn någonting?' in Hesselborn P. O. (ed) *Infrastruktur och samhällsekonomi* TFB report 1992:21, Transportforskningsberedningen, Stockholm 1-20.
- Mohring, H. (1976) *Transportation economics*. Ballinger, Cambridge, Massachusetts.

- Mohring, H. and Williamson, H. F. (1969) Scale and 'industrial reorganisation' economies of transport improvements. *Journal of Transport Economics and Policy* **3** (3) 251-271.
- Munnell, A. H. (1990) Why has productivity growth declined? Productivity and public investment. *New England Economic Review* (Jan/Feb) 3-22.
- Nas, T. F. (1996) *Cost-benefit analysis*. Sage publications, London.
- Nyborg, K. (1996) *Some Norwegian politicians' use of cost-benefit analysis*. Discussion paper no. 169, Statistics Norway, Research department, Oslo.
- Odeck J. (1996) Ranking of regional road investment in Norway - does socioeconomic analysis matter? *Transportation* **23** (2) 123-140.
- SACTRA (1999) *Transport and the economy*. The Standing Advisory Committee on Trunk Road Assessment. Department of the Environment, Transport and the Regions, London.
- Sandvik, E. (2000) *Trafikkendringer ved avslutning av bompengeneinnkreving*. Student report, Molde University College, Molde, Norway.
- Walters, A. A. (1968) *An introduction to econometrics*. Macmillan, London.

**Prologue to Paper One:
Economic appraisal in Norwegian aviation**

2. Prologue to Paper One: Economic appraisal in Norwegian aviation

2.1 Introduction

Paper One presents a method for economic appraisal within Norwegian aviation. The basis for the method of appraisal is cost-benefit analysis (CBA). While economic impact assessments are frequently used in the road sector, the use of CBA in the aviation sector has been far more limited. One of the reasons may be that the investments and operations within aviation have been subject to financial responsibility. If the revenues from a project cover the costs, then the project has conventionally been considered as profitable. The paper shows why this decision criterion may fail to reflect economic efficiency, and two examples are provided to illustrate why more comprehensive economic impact analyses are necessary. The paper gives an overview of the proposed framework for CBA in Norwegian aviation and discusses some important factor prices used in CBA. Selected methodological topics are assessed, like interdependency problems among projects, inter-modal transport issues and project risk assessment. Conventionally, the aviation authorities have financial responsibility for their activities. Because economic profitable projects may fail to prove financial viability, it will be important to balance the financial profitability criterion for project selection against a selection criterion based on economic efficiency.

The introductory sections 2.1-2.4 in this prologue present the underlying theoretical framework for the paper and extend the discussion of several topics, like aggregation of individual preferences and the classification of impacts. First, this prologue will provide a formal presentation of the general CBA concept including the underlying assumptions. Second, various consumer surplus measures are examined. Third, approaches regarding aggregation of the individuals' willingness to pay (WTP) and the CBA's role in the decision-making process will be briefly commented. Fourth, the classification of impacts from aviation investments will be discussed. The decision-makers are often confused about what effects that are considered in a CBA. A classification that may improve the understanding of this matter is presented. The introduction will end with a description of the impact assessment procedure within aviation, and a brief discussion on the use of cost effectiveness analysis (CEA) in this sector.

2.2 *The cost-benefit analysis and the measurement of welfare change*

2.2.1 Formal derivation of the cost-benefit analysis concept

The cost-benefit criteria can be formally derived by means of neo-classical microeconomics. The micro-economic framework uses the following assumptions of perfect competition:

1. Freely available information about market transactions (prices, volumes), consumers/firms maximise their own benefits.
2. Each firm is a price taker - it offers a homogenous product to many buyers at the equilibrium market price. Monopolistic behaviour is ruled out by assumption.
3. No barriers to entry. (Free entry and exit). Resources are completely mobile, and firms enter on expected profits and exit on expected loss.
4. All factors of production are privately owned.

A formal derivation of the welfare maximisation and cost-benefit analysis concept can be carried out like this:

Society comprises N individuals ($i = 1 \dots n$), each with a utility function of the form:

$$(1) \quad U_i = U_i(X_1, X_2, \dots, X_m)$$

where X_1 is the amount of good 1 available to individual i . Since individual i supplies some goods and receives others, some of the X_1, \dots, X_m will be negative (inputs) while others are positive (outputs).

The income of individual i is Y_i :

$$Y_i = \sum_{j=1}^m P_j \cdot X_j \quad p_j \text{ is the price of good } j.$$

$$(2) \quad Y_i - \sum_{j=1}^m P_j \cdot X_j = 0 \quad \text{The entire budget is used.}$$

The individual's utility is maximised when the following Lagrangean expression (equation 3) is maximised:

$$\text{Max}W = U_i(X_1, \dots, X_m) \quad \text{s.t.} \quad \sum_{j=1}^m P_j \cdot X_j = Y_i$$

$$(3) \quad Lw = U_i(X_1, \dots, X_m) + \lambda(Y_i - \sum_{j=1}^m P_j \cdot X_j)$$

where λ is the Lagrangean multiplier. The relevant maximising condition is:

$$\frac{\partial W}{\partial X_j} = \frac{\partial U_i}{\partial X_j} - \lambda_i \cdot P_j = 0$$

↓

$$(4) \quad \lambda_i = \frac{\frac{\partial U_i}{\partial X_j}}{P_j}, \quad \text{and}$$

$$(5) \quad P_j \lambda_i = \frac{\partial U_i}{\partial X_j}$$

$\frac{\partial W}{\partial Y_i} = \lambda_i$ where λ_i is the marginal utility, or the marginal welfare change, of income for individual i.

Any change of income to individual i will comprise changes in the goods and inputs X_j . It follows that an increase ΔX_1 in the amount of X_1 available to i will raise i's utility by:

$$\frac{\partial U_i}{\partial X_1} \cdot \Delta X_1$$

and similarly for all goods comprising the change in income. The change in income will change the utility of individual i by

$$(6) \quad \Delta U_i = \frac{\partial U_i}{\partial X_1} \cdot \Delta X_1 + \dots + \frac{\partial U_i}{\partial X_m} \cdot \Delta X_m$$

Substituting (5) into (6) we get:

$$(7) \quad \Delta U_i = \lambda_i \cdot P_1 \cdot \Delta X_1 + \dots + \lambda_i \cdot P_m \cdot \Delta X_m$$

that is;

$$(8) \Delta U_i = \lambda_i \cdot \sum_{j=1}^m P_j \cdot \Delta X_j$$

The equations 7 and 8 say that the utility from the increment of income is equal to the change in income multiplied by the marginal utility of income. To find the change in social welfare from a rise in national income, it can be assumed that:

$$(9) \Delta SW = \Delta U_1 + \Delta U_2 + \dots + \Delta U_n = \sum_{i=1}^n \Delta U_i$$

i.e. the individuals' utility can be aggregated.

This gives:

$$(10) \Delta SW = \sum_{i=1}^n \sum_{j=1}^m \lambda_i \cdot P_j \cdot \Delta X_{ij} \quad \text{see (7)}$$

The change in social welfare can now be illustrated in terms of costs and benefits. Costs are inputs, outputs are benefits. (10) above can be expanded to

$$(11) \Delta SW = \sum_{i=1}^n \cdot \left[\sum_{j=1}^k \lambda_i \cdot P_j \cdot \Delta b_{ij} - \sum_{j=1}^m \lambda_i \cdot P_j \cdot \Delta c_{ij} \right]$$

If now the marginal utility of income is assumed equal for all persons, so that

$$\lambda_1 = \lambda_2 = \dots = \lambda_n = \lambda$$

and if the change in national product is distributed across the community, then (11) can be simplified to

$$(12) \Delta SW = \lambda \cdot \left[\sum_{j=1}^k P_j \cdot \Delta b_{ij} - \sum_{j=1}^m P_j \cdot \Delta c_{ij} \right]$$

Since the absolute magnitude of SW is not relevant, equation (12) can be written as

$$\Delta SW = \sum_{j=1}^k P_j \cdot \Delta b_{ij} - \sum_{j=1}^m P_j \cdot \Delta c_{ij}$$

or

$$(13) \Delta SW = \sum_{j=1}^k B_j - \sum_{j=1}^m C_j \text{ if we set } B_j = p_j \cdot \Delta b_j \quad \text{and} \quad C_j = p_j \cdot \Delta c_j$$

This is a general result. Aspects regarding the assumptions of competitive markets and the marginal utility of income will be discussed later on. Once the cost and benefit elements are identified, the *valuation* of these becomes the issue at stake in terms of how social welfare is affected by public projects.

2.2.2 Consumer surplus measures in cost-benefit analysis

The determination of the B_j 's in equation 13 in the previous section will now be considered. The consumer and producer surplus are the relevant welfare measures. *Consumer surplus (CS)* can be defined as the difference between the maximum amount that an individual would be willing to pay for a good, and the amount actually paid. Figure 2.1 illustrates the Marshallian CS measure, where money income is assumed to be constant, independent of the price level.

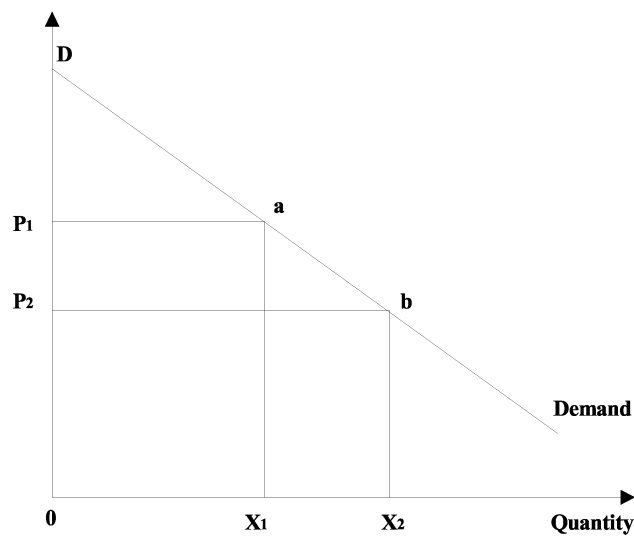


Figure 2.1 Consumer surplus

Here, the area $0DaX_1$ equals maximum willingness to pay (WTP) for X_1 units, while the area $0P_1aX_1$ corresponds to the amount actually paid for X_1 units. The difference P_1Da equals the consumer surplus (CS). A price decrease to P_2 gives $CS = P_2Db$, and the additional $CS = \Delta CS = P_1ab P_2$. The Marshallian surplus measures both the substitution and

income effect, meaning that the welfare effect from the price change alone may be biased. To be able to separate the welfare effects from the price change, Hicks' compensated demand curves (figure 2.2 and 2.3) are used.

These curves assume that the consumer's *real* income is constant, and illustrates the effect of a price change alone. The Hicks-compensated surpluses are termed Equivalent

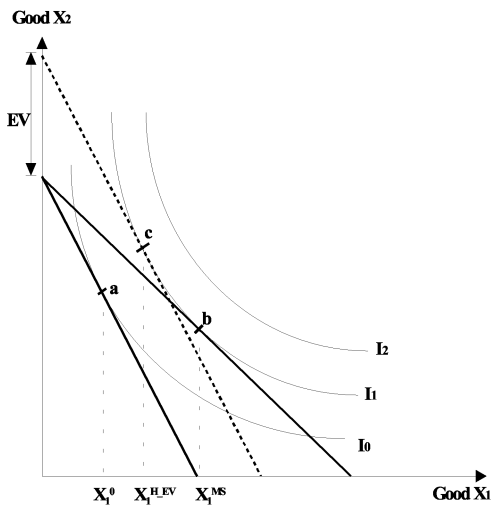


Figure 2.2 Equivalent variation

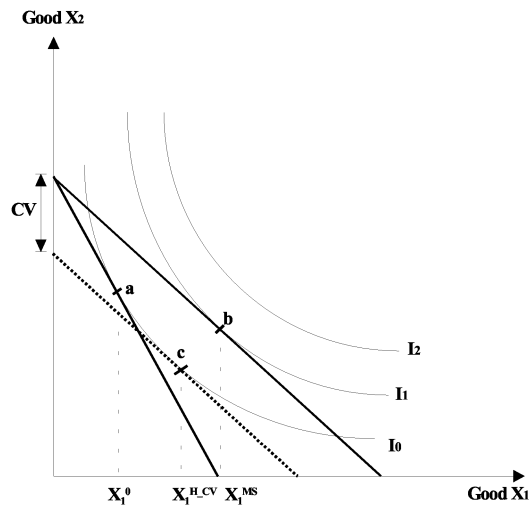


Figure 2.3 Compensating variation

Variation (EV) and Compensating variation (CV). These measures can be considered as follows (adapted from Shone 1981):

Equivalent variation, EV: Consider a price fall as in figure 2.2. The budget line rotates outwards, hits I_1 (higher utility level) at b. The individual is equally well off at point c (still at indifference level I_1). EV is now the exact amount of money that is needed to compensate the individual who does not buy at the lower price. The compensation is necessary to be able to get to the higher utility level I_1 , which is the post-price utility level in case of a price fall. In the case of a *price rise*, EV is then the maximum amount of money that the individual is willing to pay to be exempted from the higher price, but remain at the lower utility level I_0 , which now is the post-price level.

Compensating variation, CV: Consider a price fall as in figure 2.3. The individual is equally well off at point c as before (which was in point a), i.e. at utility level I_0 which is the pre-price utility level. CV represents the amount of income that could be taken away from the individual (the individual's willingness to pay), while maintaining the same level

of welfare as the individual had before the price fall. In case of a *price rise*, CV is the amount of money that the individual would be willing to accept to be able to maintain the pre-price utility level, which is now I_1 .

EV and CV are measures of an individual's willingness to pay (WTP) for, or willingness to accept (WTA) to forgo, a welfare change, here expressed in terms of a price change. *The monetary value* of a given welfare change is the issue at stake. Figure 2.4 illustrates the 3 different welfare measures and the corresponding demand functions.

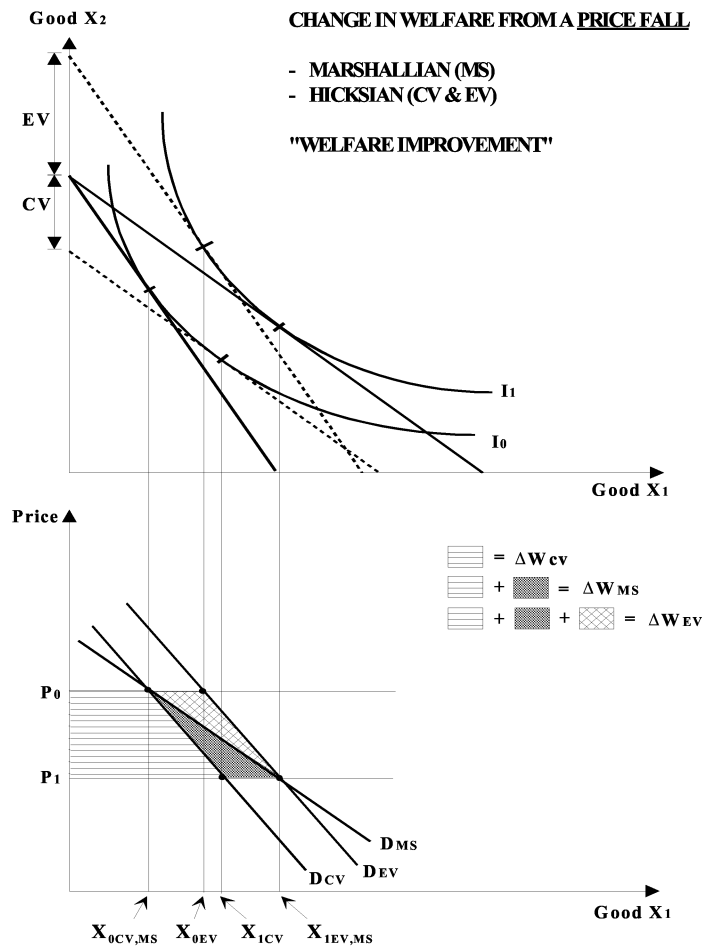


Figure 2.4 Welfare measures from a price change

The area ΔW_{CV} is the compensated variation welfare measure, and denotes the willingness to pay for the good. ΔW_{MS} is the uncompensated Marshallian demand, while ΔW_{EV} is the equivalent variation welfare measure. It is apparent from the figure that when significant

income effects are present, the CBA of a given project may give different results depending on the applied welfare measure. The Hicksian welfare measures takes only the substitution effect into account, and are the precise measures of a welfare change. Such a WTP measure (here: CV) will reduce the *ex post* demand compared to the Marshallian surplus measure and hence constitute a lower economic welfare measure. If good X_1 is a transport infrastructure improvement, then the benefits for the induced traffic will be lower because of the lower traffic level, the traffic increases to X_{1CV} instead of $X_{1EV,MS}$. A WTA measure will entail a higher level of welfare change because the *ex ante* marginal utility level is supposed to be on a higher level. The individual is *compensated* to be able to be on the same utility level without the welfare improvement as would have been the case with the welfare improvement.

An example from aviation (see Paper One, example two) can illustrate the difference between CV and EV, in a case with reduced air transport costs: The difference between WTP and WTA. Assume that a new Instrument Landing System (ILS) is planned on an airport. The welfare improvements consist of reduced travelling time and enhanced regularity. A survey among people using the airport is carried out by means of the contingent valuation method (for details on this method, see Mitchell and Carson 1989), to elicit the individuals' valuation of the infrastructure upgrading in terms of welfare change. The main question can schematically (and simplified) be posed in two ways: 1) What is the maximum amount of money that you are willing to pay (in terms of increased fares) to get improved regularity and reduced travel time? 2) What is the minimum compensation that you would accept to continue with the existing ILS system instead of getting a new one? The answer to the first question gives the individual's WTP measure by using the CV welfare measure. CV gives the amount of income that can be collected from the users if the new ILS is provided, to keep the respondent on the same utility level as before the welfare change, i.e. as if the old system should still be operating. The answer to the second question gives the individual's WTA by using the EV measure. EV gives the incremental income needed as compensation for the individual to be equally well off with the present ILS *as if* the new one was implemented. The WTP measure uses today's utility level as the point of departure (I_0 in fig. 2.4), while the WTA measure refers to the utility level that will be reached with the new ILS (I_1 in fig. 2.4). Both WTP and WTA are theoretically correct measures of the welfare change. It will be a matter of policy judgements to decide which one to use. If the use of 'best possible' infrastructure technology is regarded as a *right* that people should have free of charge, then the WTA welfare measure appears to be the most adequate one here. If improved infrastructure is regarded as something people should pay for having access to, then WTP will be the appropriate measure.

In practice however, the Marshallian surplus measure is the one that is used because Marshallian uncompensated demand is, in principle, observable (Hines 1999). Thus, the Marshallian measure is applied in Paper One. If the income effect of the price change is zero, then the three consumer surplus measures are identical. The problem with the Hicks-compensated demand measures is empirical because these demand curves are not easily observable. There is a sizeable literature that discusses methods for separating income and price effects based on information about consumer behaviour, see e.g. Mohring (1971). Traditionally, within transportation cost-benefit analyses are based on uncompensated (Marshallian) demand, where the traffic forecasts determine the volumes (based on changes in private transport costs). To what extent Marshallian demand deviates from Hicksian demand, depends on the income effect in the markets under consideration. Hines (op cit.) claims that although the Hicksian welfare measures seem theoretically more appealing, and even if there may be situations where Marshallian and Hicks compensated demand measures differ significantly, the Marshallian surplus measure constitutes a reasonable approximation that covers the needs in economic impact assessments.

In the literature, there have been concerns about spillover effects from transport infrastructure investments. As discussed in Chapter One, Mohring and Williamson (1969) showed that the incremental increased area under a Hicksian demand curve from reduced transport costs constituted the benefits for a consumer whereas the transport cost savings corresponded to the exact measure of the benefits for a producing firm.

2.3 Cost-benefit analysis and the distribution of utility

The comparison and aggregation of individual utility to represent social welfare is a controversial issue that deserves a couple of comments. The hallmark of transport infrastructure investments is that some individuals will benefit and some will lose. When individuals have conflictive interests, some kind of aggregation of their preferences has to take place before a collective choice is made (Sager 1998). Arrow (1963) showed that no aggregation device can jointly satisfy the following four conditions (applied from NOU 1997:27 1997):

1. The device should be able to process any logically consistent sets of individual preferences.
2. If every individual prefers x to y , then the society must prefer x to y .

3. The choice among alternatives must be based on how the individuals rank these alternatives, and not on how they rank irrelevant alternatives.
4. No individual has the power of dictatorship such that if the individual ranks x to y , then the society ranks x to y .

In general, the theorem says that it is impossible to aggregate individual preferences to represent a suitable social welfare function. Such representation would require that collective choice be both fair from a democratic point of view and logical (Sager 1998). As a contrast, the cost-benefit analyses depend on aggregation of individual utility, which are based on utilitarian theory, rooted back to Jeremy Bentham (1748-1832). It is therefore important to assess how the use of CBA may violate the conditions listed by Arrow. Arrow's impossibility theorem will, however, not be discussed thoroughly here, but a couple of points will be mentioned. According to Sager (1998), a planning agency in the public sector like the NCAA may reinforce the tendency to relax condition (1) above. This is so because the communication from planners to decision-makers is based on the use of a technical framework for impact assessments (e.g a CBA), instead of the outcome of a communication with the public where the preferences can be expressed. On the other hand, relaxing condition 1 seems necessary both because of the vast amount of information needed to fulfil it and because the ranking and handling of individual preferences with different strengths remains unsolved. However, there is a risk that analytic problem solving may replace political bargaining as the bureaucratic ideal (Sager 1998). The extreme scenario in this respect emerges if the decision-makers hand over their authority to the planners. The effect is that condition (4) is violated because the technocrats function as a dictator, either by interpreting the analytic results in accordance with their perception of the public interest, or by letting the outcome of the CBA (the Net Present Value) determine the decision.

Transparency in the communication between planners and decision-makers is important with respect to the way the results from the CBA are presented. With all its shortcomings, careful assessment of relevant project alternatives and their impacts together with a comprehensive listing of the gainers and the losers may strengthen the CBA's role as a robust *part* of the basis for decision-making. The method and technique presented in Paper One has these ambitions.

The positive net present value as an investment criterion states that as long as the benefits outweigh the costs, then the project is profitable. This *compensation criterion* (or the Kaldor-Hicks criterion) states that the project is profitable if the winners can compensate the losers, and still be better off. The compensation does not have to be carried out in

practice, a redistribution of utility between groups is supposed to be a matter for distributive policies, and not a part of the economic impact assessment. This criterion will be discussed a bit further below. However, the aggregation of individual utility may call for some kind of moderation to be satisfactory with respect to equity considerations. In his 'Theory of Justice', Rawls (1973) assessed the question of the individuals' rights, where every individual should be entitled to a set of basic rights, even if this should imply that wealthier groups had to reduce their welfare. Rawls defined for the sake of argument a special environment for agreement on the 'contract' that should define these rights. It should be designed by letting all members of society act as if they were behind a 'veil of ignorance', meaning that they did not know what position they were going to possess in the future society. The just distribution of goods and services is the one that would be chosen, on the condition that the members of society are left in ignorance about their future position in the society.

The CBA concept does not assess the matter of equity. There are no mechanisms that exclude projects where e.g. small benefits for the majority outweigh severe loss for the minority. From this fact, Nash et al. (1972) claim that any criterion for choosing between policies must depend on some set of value judgements, which determine what effects are to count as benefits and costs, and how they are to be traded off against each other. In the case of cost-benefit analysis, the common value judgement is that the individuals' preferences as producers or consumers are adopted as measures of value. Thus, costs and benefits are defined and measured in terms of consumer and producer surplus. In principle, this reflects the democratic principle that the individuals' votes should inform collective decision-making. As mentioned, the sum of these votes may not be appropriate on equity reasons. Nash et al. (op cit.) propose four classes of difficulties to be assessed:

a) *Whose* preferences should count? Should the preferences of future generations count? Can limits be defended, regarding peoples age (e.g children) and status (mentally ill, criminals)? What is defined as the influence area of the project - is it those who are directly affected or all those who have some kind of interests? E.g. maintaining a regional air transport network affects the users directly, but may also have an existence value for the entire society. Is the national utility the relevant one, or should the benefits or costs for foreign citizens also be taken into consideration?

b) *Which* preferences should count? Traditionally, the line is drawn between subjective and social preferences. Subjective preferences are reflected in the market behaviour, while the social preferences reflects desired policy as a result of political processes. The latter may influence the individual's preference set. Market failures like e.g. imperfect

information and free rider problems regarding the provision of public goods may however make it necessary to correct market behaviour.

c) *When* should individual preferences count? When considering the choice between "pure" market goods with no externalities, aggregated individual preferences can be considered as an adequate measure. Essential public goods provision should be evaluated carefully with respect to the individual's ability to make complex judgements. In such cases, standards or regulations set by the authorities may be one way to moderate the deficiency of consumer's sovereignty. This can be the case when dealing with biodiversity or certain kinds of nonrenewable resources where the WTP/WTA measures seem inappropriate. In such cases, a cost efficiency analysis (section 2.7) may be carried out to assess how such standards can be achieved in a cost efficient way.

d) *How* should individual preferences be aggregated? The Kaldor-Hicks criterion is commonly used in cost-benefit analyses. In short, the criterion states that the gainers of a project should be able to compensate the losers, and still be better off. If compensating variation (CV) is used as a consumer surplus measure, the criterion can be formulated as:

$$\max \sum CV_{Gi} - \sum CV_{Lj}$$

where CV_{Gi} is the welfare gain for individual i

CV_{Lj} is the welfare loss for individual j

This criterion can be used if the initial income distribution is socially acceptable, and when the project itself does not influence this income distribution - or when the government can redistribute income in a non-distortional way after the project is implemented.

Another aggregation criterion (see e.g. Dasgupta and Pearce 1978) is the aggregation of the individuals' marginal utilities from the change in provision of transport infrastructure services. However, the individuals' utility function is still the basic premise for both valuation and aggregation. This criterion can be expressed as the 'utilitarian surplus':

$$\max \sum CV_{Gi} y_i - \sum CV_{Lj} y_j$$

where y_i is the marginal utility of income of the i 'th gainer

y_j is the marginal utility of income of the j 'th loser

Equity weights may also be considered from a utilitarian perspective, where we may prefer to assign an extra amount of utility to the less favoured groups that are affected by the project:

$$\max \sum CV_{Gi} y_i e_i - \sum CV_{Lj} y_j e_j$$

where e_i is the equity weight given to the i 'th gainer

e_j is the equity weight given to the j 'th loser.

It is worth noting that the use of equity weights is a way of circumventing or smoothing the principle of letting the individuals' marginal utility be the 'vote'. The equity weights are assigned by the society based on equity considerations. By doing this, distributive aspects enter the cost-benefit analysis directly. In this way, the society's welfare function is incorporated directly into the aggregation of individual preferences. The use of equity weights is however highly disputable. Jara-Diaz and Farah (1988) discuss technical aspects of the valuation of user benefits in transport systems, but do not assess the aggregation of preferences and the interpersonal comparison. They suggest as a practical approach to reporting the users' benefits in a disaggregated fashion, to make the decision-makers aware of the distribution of gains and losses. In The Norwegian Standing Advisory Committee on Cost-benefit Analysis' (NSAC) 1997 report (NOU 1997:27), it is recommended that the analyst refrains from assigning equity weights on the ground that they represent subjectivity that are not founded in the principle of consumer's sovereignty. Besides, the decision-making is often a result of a *negotiation process* where there may be several decision-making levels involved. These levels may have different perceptions regarding equity. Thus, initial equity weigh assignment by the analyst does not seem appropriate. NSAC's recommendation is that the distributive effects be described thoroughly to enable the decision-makers to take them properly into consideration. The advice is also to assess project profitability under different equity constraints to see how sensitive the project's profitability is with respect to different views on equity, and to present the means that can be used for compensation.

When assessing almost any project of some magnitude within the aviation sector, equity rapidly become an important issue. In the decision-making process, considerations regarding equity and specific local effects like e.g. employment and localisation are often regarded at least as important as the economic profitability expressed by net present value and/or the cost/benefit ratio. In the road sector, studies have shown that the link between economic profitability and the choice of projects is not particularly strong (Odeck 1996). This is partly due to distributive policies, but also to the fact that the CBA's ability to

capture the economic effects is not acknowledged by the decision-makers (Nyborg 1996). The Kaldor-Hicks criterion is applied in the road sector analyses, without any equity considerations. Formally, the introduction of social 'welfare functions' may, and most probably will, cause deviations between project rankings based on profitability alone and the actual rankings as an outcome of a political process. Figure 2.5 explains why.

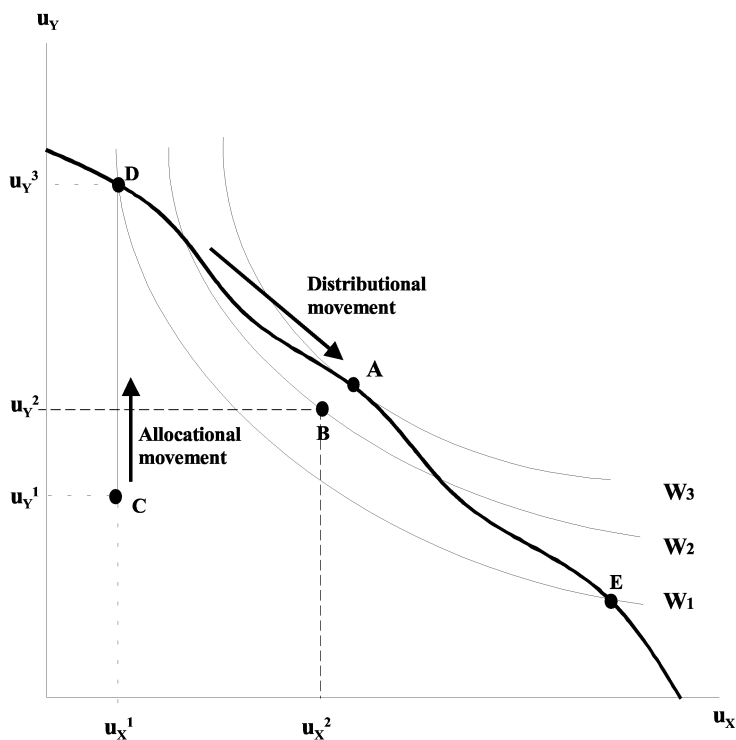


Figure 2.5 First and second best projects

The figure shows a simplified situation where two individuals, X and Y are affected by a project. The frontier is the 'grand utility frontier', and gives the Pareto optimal distribution of utility (u_X and u_Y) between X and Y. The frontier satisfies the conditions where the marginal rate of transformation (MRT) equals marginal rate of substitution (MRS). There is efficiency in production and consumption, thus nothing can be gained by reallocating resources. C is the initial situation (with corresponding utility distribution u_X^1 and u_Y^1), which one seeks to improve. Improving the use of available resources implies a move towards the utility frontier. This can be done by i.e. reducing air transport costs by means of some infrastructure improvement. Implementing a profitable project may result in situation D, where the situation for X remains u_X^1 , while the utility for individual Y increases significantly to u_Y^3 . This is a Pareto improvement because no one has become worse off. If the project had resulted in situation E, then a Kaldor-Hicks improvement had

taken place, because the utility for Y had been reduced. The indifference curves $W_{1...3}$ are the social indifference curves, where the society is indifferent regarding the utility distribution between X and Y. A movement along the frontier from D (or E) to A means that a project is selected that meets both the efficiency and equity norms. The adjustment takes place on the higher welfare level W_3 instead of W_1 . But one could also imagine projects that give situation B which is off the efficiency frontier (less profitable than projects A, D, E) but giving more welfare than D (or E). A project that takes the society to B is more preferred than an alternative project aiming at D or E because the former, despite its lower net present value, is superior in terms of social welfare rankings. It is worth noting that 'social welfare functions' are conceptual tools. In practice, they are results of political discourses rather than mathematical specifications.

The next two sections will discuss distinctions between real and distributive impacts, and how changes in transport costs can be reflected in terms of location and employment effects. Decision-makers within Norwegian aviation have signalled the need for a broader assessment regarding the effects from infrastructure investment, like being able to answer questions like: 'What does e.g. travel time savings and investments actually mean in terms of local changes in employment and localisation of industry?' The underlying condition here is the need for information to be able to evaluate projects in a more comprehensive way against political ends. A conceptual framework is presented as a possible starting point for a further discussion.

2.4 Real and reflective impacts

In the discussion on whether the cost-benefit analyses are able to capture all the real benefits of a project, it may be constructive to define what kinds of impacts that generate changes in supply and use of real resources. A discussion of the terms 'real impacts' and 'distributive impacts' follows, and a couple of examples are provided. Subsequently, the concept of *reflective impacts* is outlined. Reflective impacts are observable impacts on employment, industry and residents from changes in transport costs. This hierarchy of effects may contribute to an improved understanding of how impacts of transport investments can be categorised.

Real impacts from a project represent the changes in supply and use of real resources within the economy. Real resources are goods that have an economic value, whether they are monetised or not. As an example, better accessibility can contribute to economic development because existing production factors can be used more efficiently (because of reduced time costs and vehicle operating costs) or attract new production factors through

reallocation to locations with highest returns. This effect can alter the competitive environment for local industry. One consequence may be increased efficiency through exploitation of increasing returns to scale (IRS), and/or a sharpening of the local competitive environment where a previously protected industry is exposed to competition. This is further discussed in Paper Three. Another example may be that changes in accessibility do not necessarily result in increased use of transport services. Such effects may be termed Option Values. Within rural aviation, there are e.g. indications that the possibility of fast transport to get acute medical treatment has a real value beyond what is reflected in observed travel behaviour (Bråthen and Hervik 1992).

Distributive impacts represent a pure distribution of resources within an economy. The loss for those who have to renounce real resources (the 'losers') are offset by the gain for those who gets access to the same resources (the 'winners'). The net change in the supply and use of resources are thus zero. However, these distributive effects may be of significant concern in a political context, because transfers of welfare between groups and/or regions are held against a more or less explicitly formulated set of social objectives.

The distinction between distributive and real impacts becomes more complicated if the 'winners' and the 'losers' have asymmetric marginal utility from the same amount of identical goods. This is the case if e.g. the gain of extra income in one region is higher than the loss from renouncing the same amount of income in another region, e.g. as a result of moving a company from a 'rich' location to a 'poorer' one. If such imbalance exists - here as a result of diminishing marginal utility of money - then an apparently distributive impact may also have effects in real terms.

Reflective impacts are considered as the observable consequences of the real impacts from a given action. Such real impacts may e.g. be reduced transport costs, investment costs and operating costs. The reflective impacts comprise the impacts for firms, residents and employment in the private and public sector in the influence area. Investments are quantified in monetary terms. An investment entails the use of labour during construction, which are then considered as a reflective impact of the investment. The costs of using the labour force are a part of the cost-benefit analysis and enter the analysis via the investment costs. However, if the labour force have to be relocated, and if the relocation costs are not internalised in the CBA through the wages, then relocation costs are an additional real effect. If applicable, these costs should be included in the CBA. As elaborated below, the local demand effects of the use of labour are also a part of the reflective impacts.

The presentation of the reflective impacts raises the question of double-counting the benefits and the costs in the decision-making process. A very important matter is thus to present the reflective impacts, without real and reflected effects being added together in the back of the decision-makers' minds. Of course, when 'extending' the impact assessment beyond cost-benefit analysis in this way, the risk of such double counting is present. However, the political concern stays apparent and calls for a thorough description of what will be the social consequences of the infrastructure investments. The relevance of the focus on reflective impacts as a part of an economic impact assessment may be illustrated as follows:

Assume that two infrastructure investments are assessed by cost-benefit analyses. The same agency is responsible for decision-making in both cases, and the projects compete for the same funding, which leaves possibilities for only one of the projects. Both projects are profitable in terms of equal net present value, and the C/B ratios also turn out to be equal. Thus, profitability and ranking are equal. No clear recommendations can thus be made on basis of the CBA. The difference between the projects may however be the reflective impacts. If e.g. one of the projects reduces local unemployment while the other benefits tourism in an area with a high employment rate *ex ante*, there may be a case for selecting the first project.

Reflective impacts may be classified in various ways. For airport expansions, ACI (1993) uses three categories, namely direct, indirect and multiplier impacts. The direct impacts are related to activities in the airfield. The indirect impacts are linked to airport supply while the multiplier effects are caused by the direct and indirect impacts.

It may be convenient to use two main groups of reflective impacts to distinguish the real and distributive impacts, namely *direct and indirect reflective impacts*. The indirect impacts are divided into three subgroups. The reflective impacts can be divided into four functional categories (Bråthen and Strand 2000):

- *Direct reflective impacts* are functionally limited to operations of the infrastructure, the transport modes and the traffic management services, like e.g. Air Traffic Management systems.
- *Second order reflective impacts* are connected to enterprises that serve the users of the transport systems. These firms comprise travel agencies, hotels and restaurants, trade, parking and tourism in general. Effects in other parts of the transport network are also defined as a second order reflective impact.

- *Third order reflective impacts* comprise the direct effects (employment, housing) that results from relocating as an effect of changes in transport costs. These enterprises have no deliveries linked to the transport system or the transport users, they relocate because of changes in accessibility to other markets. Third order effects may also be termed *catalytic effects*, because they emerge as relocation effects without any direct relations ex ante except being users of the transport system.
- *Fourth order reflective impacts* are the multiplier effects of the direct, second and third order reflective impacts. They comprise the activity level in terms of e.g. employment, trade and health care that are generated as a result of these impacts. Environmental effects are also defined as a fourth order reflective impact.

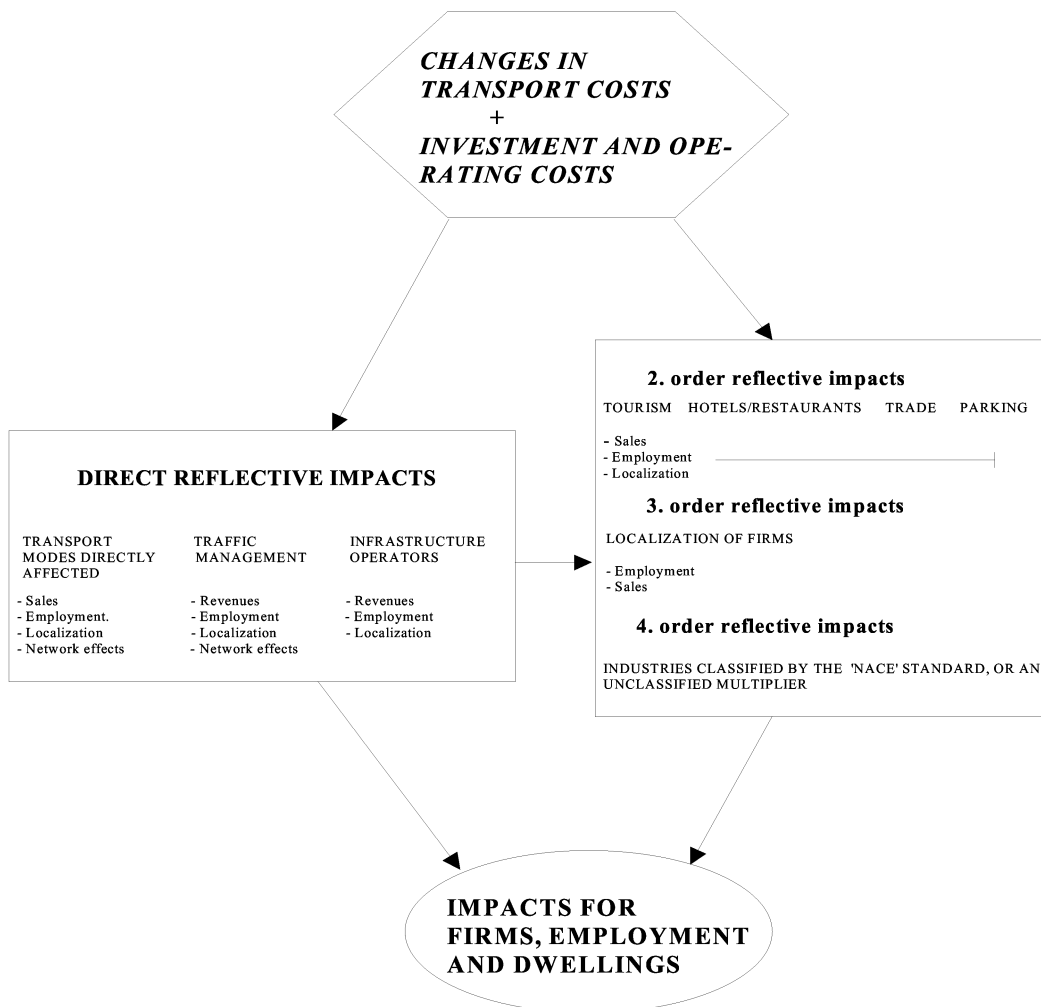


Figure 2.6 Real and reflective impacts

Figure 2.6 illustrates the relations between real and reflective impacts, where the impact categories that will be of greatest interest regarding aviation infrastructure investments are presented. As far as possible, one should identify in numerical terms the relevant reflective impacts (changes in employment, localisation etc.) and at the same time include the real effects as values in a cost-benefit assessment. As already underlined: one important issue to be aware of, is the risk of double counting the real economic and the reflective impacts. The reflective impacts are meant as supplementary information to the decision-makers of how the real effects are reflected in terms of e.g. employment and location.

Reflective impacts and traffic forecasts

A careful assessment of the reflective impacts may also provide useful information about how a given project may affect the market for aviation services. Often, the traffic forecasts do not treat local adaptation to changes in the transport system adequately (Strand 1999). Some of the reflective impacts may probably have feedback effects on the demand for air services. There are reasons to stress that a thorough traffic forecast as a basis for the cost-benefit analysis should capture the long run effects of the investment, including the way changes in transport costs influence the spatial equilibrium and hence local demand. Figure 2.7 illustrates some of the complex relationships between variables affecting air transport demand.

Forecasting models consist both of those factors that are directly or indirectly affected by the changes in transport costs and other factors that are independent of the investment but with a potentially significant influence. Figure 2.7 contains some of the short and long run relationships that are considered as the most important ones¹. In the short run, changes in travel costs have direct impacts on the travel activity. If these changes are of such magnitude that they significantly affect the framework under which individuals and firms are operating, then long run adjustments (e.g. localisation) may take place and thus affect the long run growth in the aviation market. Normally, such effects are captured through short run and long run elasticities. These elasticities are normally dependent upon local conditions. In the figure, such 'secondary' effects are shown by the links between 'changes in air travel costs' via 'environment for firms/individuals' to 'air traffic demand'. Changes

¹ In the figure, the GNP growth is illustrated as affecting local employment, preferences, travel behaviour and prices/taxes. In the real world, all the elements in the figure may influence GNP through productivity impacts.

in e.g. localisation cause employment effects that may be beneficial to parts of the aviation network, perhaps as transfer effects from other locations. The real effects are then the differences in transport costs when travelling from these two locations, with corrections for relocation costs. In the road sector, such effects have been given attention. In Boarnet (1998), the main conclusion is that one should be aware of negative output spillovers, where counties with rich street and highway infrastructure benefit at the expense of the surrounding poorer counties. Conventionally, such effects are regarded as distributional but there may be real components present in terms of changes in factor prices in the affected areas.

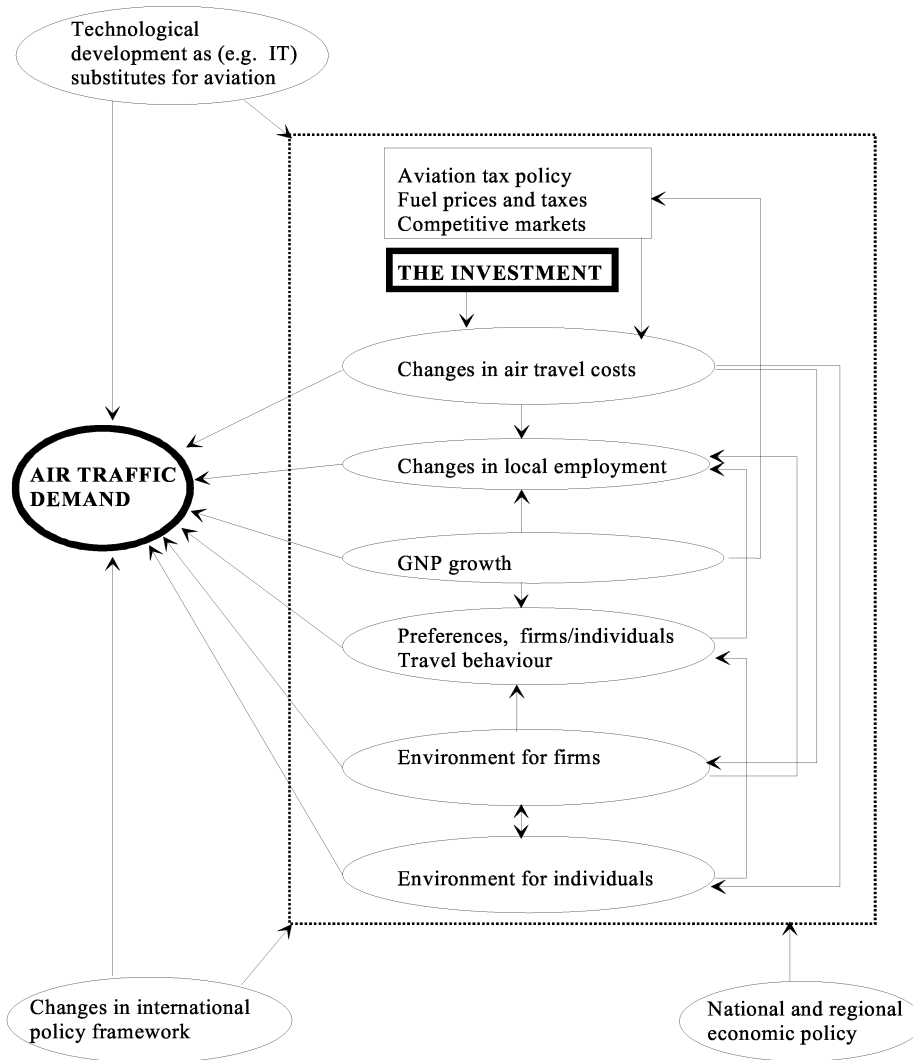


Figure 2.7 Driving forces in traffic forecasts

Figure 2.7 shows that in addition to the transport cost effects, air transport users may also be affected by conditions that are determined exogenously, i.e. outside the air transport system. Exogenous factors at the national level are related to economic policy regarding e.g. taxes and regional balance. In a world of intertwined economies, the development in international economy and policy will most certainly affect both domestic and international aviation, independent of national policies. Development of potentially competing technologies that may reduce the need for travel or change the travel pattern (e.g. information technologies, IT), may also be an exogenous element that complicates air traffic forecasting.

Complex interactions are the rule rather than the exception in traffic forecasts. A general challenge in forecasting is to establish models that are closed or semi-closed, i.e. models that contain all relevant explanatory variables, at least in the short run. Open models are not able to take 'control' of the dependent variable because explanatory variables are left out for various reasons. In a closed model, the deviation between the real development and the forecasts will be caused by the fact that one or more of the explanatory variables have had a different course than what is used as the point of departure in the traffic forecasts. A semi-closed model that is valid in the short run will probably be the best achievable practice. In the long run, the model becomes open in the sense that it will hardly be possible to include all the relevant explanatory variables. As an example already mentioned, the development in information technology and its long run influence on travel activity is largely open to conjectures. The lack of long run closed models leaves a fundamental problem to be solved because forecasts are made for predicting long run effects of transport system improvements². A thorough analysis of the characteristics of the local environments to be affected by the aviation investment is one way to improve the predictive power of the forecasting models to be used as an important prerequisite in the cost-benefit analyses.

2.5 A cost-benefit framework for aviation

In Paper One, a method for CBA in aviation is described, where a few important aspects regarding e.g. system limits are considered. In this section, a procedure for carrying out CBA is described, and selected elements are discussed in more detail. To ensure that the analysis is carried out in accordance with best practice, it is important to stick to a procedure that establishes a consistent framework for analysis across projects. Figure 2.8

² Paper Three gives a further elaboration on closed and open models.

illustrates the procedure that is suggested for used within the Norwegian Civil Aviation Authority (NCAA)³.

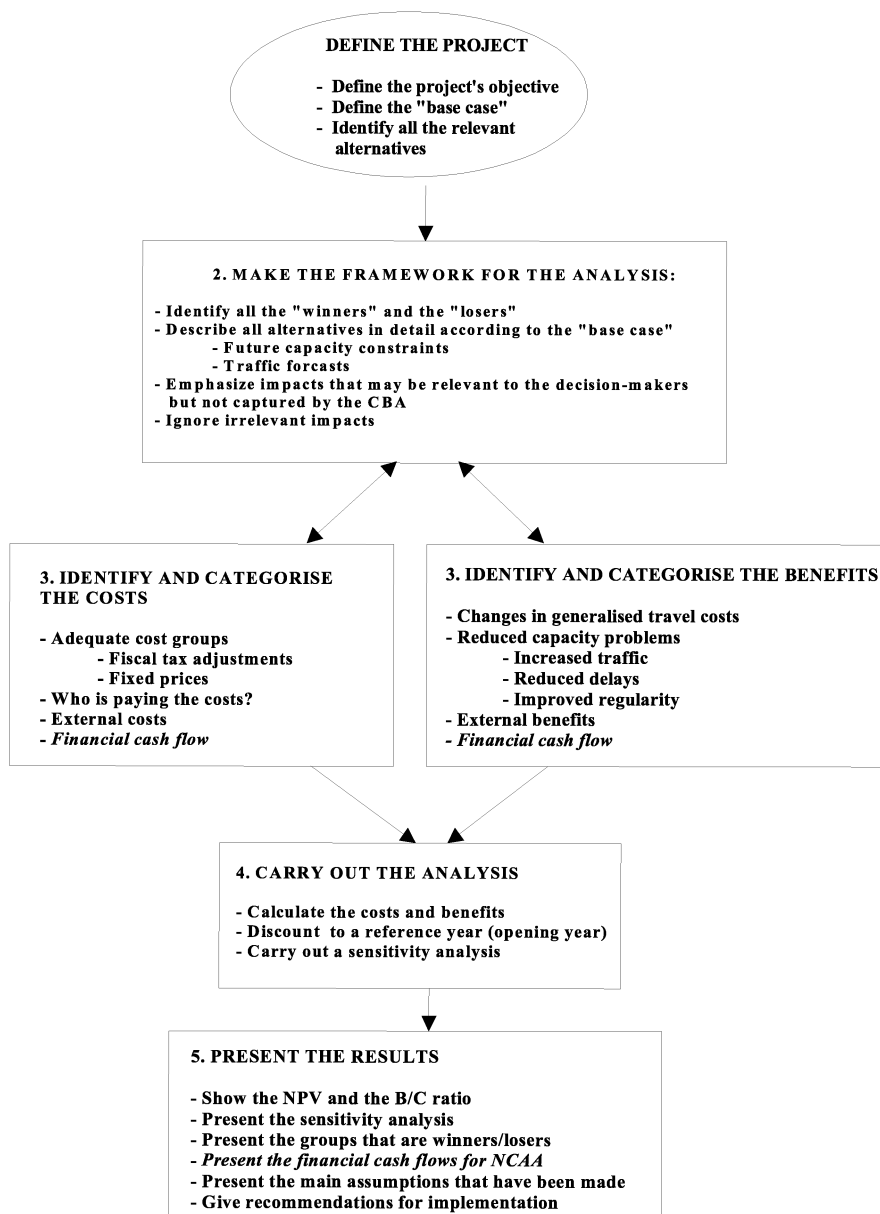


Figure 2.8 The CBA procedure for aviation

³ Since Paper One was published, the NCAA has changed its name to Norwegian Air Traffic and Airport Management (NATAM). For consistency with the published paper, NCAA will be used here.

First, the project is to be defined with main objectives and alternatives as outlined in Paper One. The next step is to make the framework for the analysis where the stakeholders are identified, and where the impacts from the various alternatives are described in detail, based on traffic forecasts that should be made with respect to capacity constraints affecting the growth rate. It is important to rule out irrelevant impacts. As an example, changes in the price of real estate are irrelevant (double counting) if the changes in transport costs are taken into consideration. The change in estate prices is a reflective impact of changes in the transport costs.

The third and perhaps most complex step is to identify the impacts in terms of real costs and benefits for each alternative. This step has to be seen in interaction with step 2, because the traffic forecasting models are dependent on the specification of transport costs. The real costs are adjusted for fiscal taxes, and external costs like emission and noise nuisance costs are included. For a discussion on the concept of external costs, see e.g. Button (1993). One main component in the CBA is the change in transport costs. These costs include value of time (VOT), ticket costs and other monetary costs. Usually, these costs are added together to express the *generalised travel costs*, where the VOT is expressed in monetary terms. For a discussion on the theoretical basis for including monetary VOT in the transport costs, see e.g. Bruzelius (1979) and Bates (1992).

Questions have been raised regarding the empirical estimations of VOT. In addition to the debate on methodological approaches (e.g. revealed preferences versus stated preferences research design), both the value of short versus long travel time segments and the variations in VOT between transport modes have been subject to discussion. Gårder (1989) summarises important parts of this discussion, and so does Killi (1999).

One of the main assumptions in CBA is that prices are determined in a competitive environment. Within aviation, taxes are levied for various reasons, which violates this assumption. As most other aviation authorities in Europe (Fewings et al. 1998), NCAA has a financial responsibility for its own activities. The cost structure within important parts of the aviation system (e.g. uncongested airport infrastructure and air traffic management systems) is characterised by increasing returns to scale (IRS). Figure 2.9 shows the possible pricing strategies with IRS.

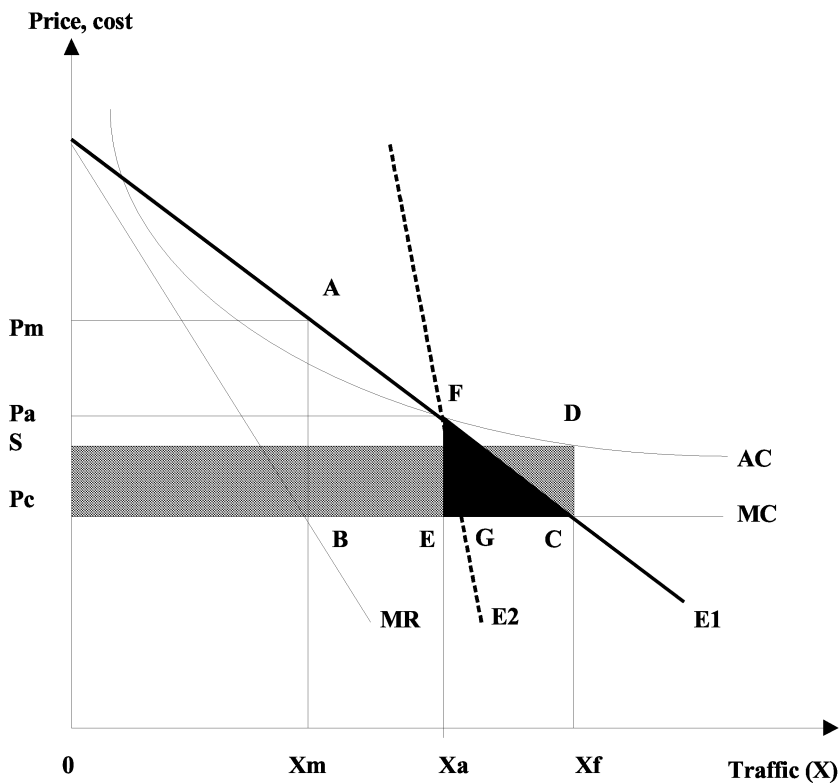


Figure 2.9 IRS, pricing strategies and dead weight loss

For a monopolist, profit maximising behaviour gives the price P_m and the corresponding traffic volume X_m . According to the theory of monopolistic behaviour, the profit maximising price is determined where the marginal revenue (MR) equals marginal costs (MC). Normally, to ensure cost coverage⁴ the monopolist is subject to regulations which aims at a pricing regime where the price equals average costs (P_a). Price P_c equals marginal cost, which maximises economic efficiency, but leaves a financial deficit equal to the shaded area P_cSDC . For the aviation authorities, in most cases the NCAA's air charges are submitted to the Ministry for approval, and the principle of $P_a = AC$ is applied. In a case with elastic demand, the dead weight loss from deterred traffic is the area ECF in figure 2.9, which is reduced to the area EGF with a less elastic demand. In practice, the air charges are differentiated among aircraft categories, but not for each airport or leg. There are thus reasons to believe that the air ticket costs used in the CBA

⁴ There are certain problems with the regulation of monopolies that will not be discussed in detail. These are connected to the level of production costs where asymmetric information between the monopolist and the regulator gives limited possibilities to control the amount of resources that is actually needed for a given level of production. For a discussion, see Tirole (1990).

are biased compared to using the marginal cost principle. This bias may go both ways, as long as congestion charges are not levied either. As a practical approximation the AC principle is used in CBA within aviation, and estimation of the biases should be subject to further research within aviation as well as in transport in general. This element is briefly discussed in Chapter One.

The fourth step in figure 2.8 is to carry out the analysis through a traditional discounting procedure:

$$NPV = -I_0 + R \cdot (1+i)^{-n} + \sum_{t=1}^n b_t \cdot (1+i)^{-t}$$

where NPV = net present value

I_0 = investment costs

R = residual value of the investment(s) at the end of the economic life span

b = annual net benefits

i = discount rate

n = economic life span (25 years)

t = years ($t=1..25$)

The level of the discount rate is briefly discussed in Paper One, where the distinction between systematic and non-systematic risk is assessed. *Non-systematic* risk is connected to the outcome of the specific project in question, and *systematic risk* is due to macroeconomic factors like general business cycles, fluctuations in the oil prices etc. and can thus not be diversified. The non-systematic risk is considered as irrelevant when analysing public projects because it can be *diversified*, and because each project's share of the total budget is small. NCAA is a publicly owned private enterprise with a substantial number of smaller investment projects running simultaneously, and quite few larger ones, which can give reasons to question the diversity properties. However, the risk regarding the investment costs can to some extent be taken into consideration by using *expected cost values*. The main risk premium is thus the systematic (unavoidable) risk component. Based on these principles, the standing advisory committee on cost-benefit analysis in Norway (NOU 1997:27) has assigned typical public investments to risk categories with different discount rates. The risk-free discount rate is 3.5%, and the determination of the risk component is based on the degree of correlation with business cycles.

Examples:

Public bus transport:	$3.5\% + 1.5\% = 5\%$
Roads:	$3.5\% + 3.5\% = 7\%$
Airports:	$3.5\% + 4.5\% = 8\%$

The differences in the choice of discount rate can be explained as follows: Public bus transport is a typical *reversible* investment, while a road investment is *irreversible*. Airport investments are also mainly irreversible, and the profitability of these projects are even more exposed to business cycles than roads⁵. This means that future cash flows for aviation projects will have relatively less weight in the discounting procedure than e.g. public bus transport. Other things equal, the demand for the future net benefits are higher for aviation projects to become profitable, because of the higher systematic risk exposure.

The fifth step in figure 2.8 is to present the results to the decision-makers⁶. In Paper One, the presentation design is based on the Planning Balance Sheet (PBS) method for impact assessment, but applied for presenting CBA results only. PBS can be seen as a structured way of presenting the results, showing the impacts on various affected groups. The technique involves setting down, in tabular form, all of the pros and cons associated with alternative investment options (Lichfield et al. 1975). However, the 'original' PBS method allows for the use of physical values and, if quantification is not possible, ordinal indices or scales can be used to measure qualitative impacts. The accounts are subdivided to show the effect of different projects on the groups affected. Thus, the impacts of alternative projects are to be presented systematically with respect to (groups of) individuals, private firms and the public administration. This offers guidance to equity considerations. Because the accounts can be compared with pre-determined planning objectives that are selected as reflective of community preferences. These objectives can be taken to express the social welfare function discussed above.

In the accounting, two main groups are considered:

(a) “The producers”, those who are engaged in construction activities and in operating the project (for instance road maintenance).

(b) “The consumers”, those who have benefits or losses from the services that the project provide (for instance travel time savers, those who have to renounce land for road use).

⁵ The exposure to business cycles can be expressed in terms of the income elasticity $E_i = \frac{\Delta x \cdot R}{x \cdot \Delta R}$ where x is quantity and R is income. If $E_i > 1$ (elastic as for air transport), then the exposure to business cycles is significant.

⁶ According to figure 2.8, elements from the financial analysis are also presented. This is due to NCAA's role as an agency with financial responsibilities, which is discussed in Paper One.

For the monetary impacts, the following table can illustrate the applied PBS principle:

Economic benefits and costs	Project 1				Project 2			
	Cost		Benefit		Cost		Benefit	
	C ^{*)}	A	C	A	C	A	C	A
The “producers”	n	n	n	n				
Group P1 NCAA	u	u	u	u				
Group P2 Local government	m	m	m	m				
Group P3 National government	b	b	b	b				
.....	e	e	e	e				
The “consumers”	r	r	r	r				
Group C1.1 Business travellers	s	s	s	s				
Group C1.2 Leisure travellers								
Group C2.1 Emissions								
Group C2.2 Noise nuisance								
Group C3.1 Airport shuttle services								
.....								

*) If desirable, the cost components can be divided into C=NPV of investments, A=NVP of annual operating costs.

Table 1 PBS for monetary impacts

Table 1 shows the economic costs and benefits, distributed among different interest groups affected by the project. Through adequate breakdown levels, the table may also illustrate that costs for one group may be benefits for others. The net present value and the cost-benefit ratio should be presented as main elements in the analysis. The two examples in Paper One illustrate one way of using this framework to support decision-making within aviation. The PBS method will not be further discussed here.

2.6 Cost effectiveness analysis

The cost effectiveness (CEA) method can be used when projects are unambiguously motivated from standards or regulations, and where the benefits are not monetised. If the standards have to be met, this should be done in a cost efficient way (NOU 1997:27 1997). Because such situations may emerge within aviation, a few comments on the CEA method will be made here.

The costs (investments and operating costs as well as other economic costs⁷) of each project should be identified. Here, the impacts are expressed in physical terms, e.g. units of reduced emissions to meet exogenously given standards or regulations. The impacts measured in resulting physical units from a given amount of monetary efforts give information concerning the *efficiency* of the project under consideration.

Figure 2.10 illustrates a CEA chart, with the impact (in terms of physical measures) on the Y axis and costs on the X axis.

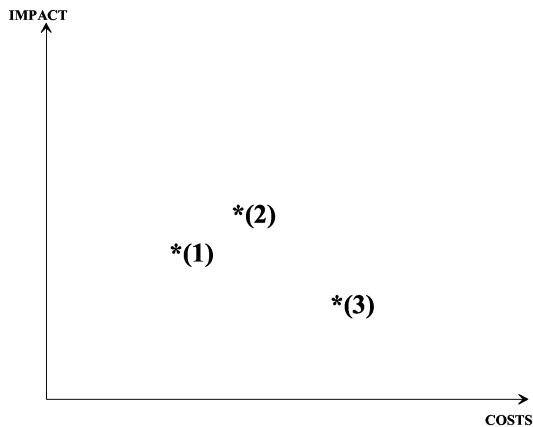


Figure 2.10 CEA-chart

Project 3 is less favourable than project 1 and 2, it gives less impacts at the highest costs. The choice between project 1 and 2 depends on the “impact per monetary unit” (the cost/efficiency ratio).

The main difference between CBA and CEA is that CEA does not have any compound decision-making criterion that can handle a variety of impacts measured on a unidimensional scale, like the net present value. CEA may be useful when assessing projects with unidimensional physical impacts. If the impacts are more complex, then CEA will not be able to support decision-making in a robust manner.

⁷ It is worth noting that the cost side may include changes in transport cost and other economic costs that are normally encompassed in a CBA. Different ways of operating aircraft may affect the emission levels, but also travel costs and noise levels.

References

- ACI Europe (1993) *The economic impact study kit*. European Region of Airports Council International, Brussels.
- Boarnet, M. G. (1998) Spillovers and the locational effects of public infrastructure. *Journal of Regional Science* **38** (3) 381-400.
- Bråthen, S. and Hervik, A. (1992) *Nytten av kortbanenettet*. Working paper M9204 Møreforskning Molde, Norway.
- Bråthen, S. and Strand, S. (2000) *Luftfartens ringvirkninger. Et forsøk på systematisering. Forprosjekt*. Working paper, Møreforskning Molde and The Institute of Transport Economics, Norway.
- Dasgupta, A. K. and Pearce, D. W. (1978) *Cost-benefit analysis. Theory and practice*. Macmillan Press, London and Basingstoke.
- Fewings, R., Morrell, P., Pagliari, R., Whelan, C. and Williams, G. (1998) *A comparative study of European airport provision*. Air Transport Group, College of Aeronautics, Cranfield University, Cranfield.
- Gårder, P. (1989) *The value of short time periods: literature review*. Report 72, Kungliga Tekniska Högskolan, Stockholm.
- Hines, J. R. (1999) Three sides of Harberger triangles. *Journal of Economic Perspectives* **13** (2) 167-188.
- Jara-Diaz, S. and Farah, M. (1988) Valuation of user's benefits in transport systems. *Transport Reviews* **8** (3) 197-218.
- Killi, M. (1999) *Anbefalte tidsverdier i persontransport*. Report 459/1999, Institute of Transport Economics, Norway.

- Lichfield, N., Kettle, P. and Whitbread, M. (1975) *Evaluation in the planning process*. The Urban and Regional Planning Series, Pergamon Press, Oxford.
- Mitchell, R. C. and Carson, R. T. (1989) *Using surveys to value public goods. The contingent valuation method*. Resources for the Future, Washington DC.
- Mohring, H. (1971) Alternative welfare gain and loss measures. *Western Economic Journal* **9** (4) 808-827.
- Mohring, H. and Williamson, H. F. (1969) Scale and "industrial reorganisation" economies of transport improvements. *Journal of Transport Economics and Policy* **3** (3) 251-271.
- Nash, C., Pearce, D. and Stanley, J. (1975) An evaluation of cost-benefit analysis criteria. *Scottish Journal of Political Economy* **22** (2) 121-134.
- NOU 1997:27 (1997) *Nytte-kostnadsanalyser*. Finans- og tolldepartementet, Oslo.
- Nyborg, K. (1996) *Some Norwegian politicians' use of cost-benefit analysis*. Discussion paper no. 169, Statistics Norway, Research department, Oslo.
- Odeck J. (1996) Ranking of regional road investment in Norway - does socioeconomic analysis matter? *Transportation* **23** (2) 123-140.
- Sager, T. (1998) Meaningful collective choice? : public planning and Arrow's theorem. *Progress in Planning* **50** (2) 75-140.
- Shone, R. (1981) *Applications in intermediate microeconomics*. Martin Robertson, Oxford.
- Strand, S. (1999) Airport-specific traffic forecasts: the resultant of local and non-local forces. *Journal of Transport Geography* **7** (1) 17-29.
- Tirole, J. (1990) *The theory of industrial organisation*. MIT Press, Cambridge, Massachusetts.

3

Paper One:

ECONOMIC APPRAISAL IN NORWEGIAN AVIATION

Svein Bråthen, Molde University College, Norway

Knut S Eriksen, Norwegian Institute of Transport Economics, Norway

Harald M Hjelle, Molde University College, Norway.

Marit Killi, Norwegian Institute of Transport Economics, Norway.

Reprinted from Journal of Air Transport Management Vol. 6, No. 3 pp. 153-166.

© 2000, with permission from Elsevier Science.

Paper One:

ECONOMIC APPRAISAL IN NORWEGIAN AVIATION

Svein Bråthen, Molde University College, Norway

Knut S Eriksen, Norwegian Institute of Transport Economics, Norway

Harald M Hjelle, Molde University College, Norway.

Marit Killi, Norwegian Institute of Transport Economics, Norway.

ABSTRACT: The Norwegian Civil Aviation Administration (NCAA) is responsible for investments and operations regarding the Norwegian airport system, and the air traffic management systems. This paper presents methods for economic appraisal in the aviation sector, proposed for use within the NCAA's jurisdiction. Methodological problems will be addressed, such as the handling of project risk and uncertainty, projects with mutual dependency, and projects influencing accident risks. The cost-benefit analysis methodology proposed here may give conflicting recommendations compared with financial analyses. This raises decision-making problems for public agencies like the NCAA.

1. Introduction

This paper describes a model for economic appraisal of aviation planning and operation. The method is based on cost-benefit analysis (CBA). It is developed for Norwegian aviation, but may be applied on aviation planning in general. The Norwegian Civil Aviation Administration (NCAA) is responsible for investments and operations regarding the Norwegian airport system, and the air traffic management (ATM) systems. The airport network for scheduled transport consists of 26 short take-off and landing airports, and 15 full length runway airports. The number of passengers arrived/departed on domestic routes amounts to approximately 19 millions per year for a population of 4.5 millions.

In its efforts to optimise the transport system as a whole, the Ministry of Transport has called for actions to improve cross modal planning, meaning that the methodology with respect to analytical work should be harmonised. Aviation plans, particularly those affecting the infrastructure, shall be reported to the Ministry, to be co-ordinated or linked with other transport sectors. Even if the aviation sector is mainly disconnected from public funding, it is still under political supervision.

While the major part of road and rail infrastructure is publicly funded, the NCAA's funds are defined by aviation taxes and commercial activities related to the airports. The only public sector engagement in aviation is through funding of public service obligation (PSO) routes. NCAA's revenues amounted to approximately NOK 3 billions in 1998. Traditionally, investments in the aviation sector have mainly been made on the basis of financial cash flow analysis combined with some sort of political judgement. Numerous investments are also determined by guidelines and requirements established by institutions like ICAO and ECAC. The fact that a substantial share of NCAA's investments affects other parties, should indicate that there is a need for a broader perspective. Typical non-financial benefits are:

- Reductions in travel time
- Reductions in accident risk
- Reductions in environmental hazards

Many investments may actually represent a financial loss for the NCAA, but still be economically profitable when benefits for the operators and the passengers are taken into account. Making priorities subject to a full economic impact analysis instead of a more narrow financial cash flow approach, may significantly alter the portfolio of projects that reaches implementation. The decision-making dilemma is evident: Despite the positive externalities, the NCAA has to finance its activities exclusively by user charges and other commercial activities. The civil aviation authorities in other countries are facing similar situations: Their role is a mixture of running a business enterprise and to provide public services.

While economic impact assessments are frequently used in the road sector, the use of CBA in the aviation sector has been far more limited. The next three sections present issues related to the development of a CBA tool for the NCAA. Section 2 gives an overview of the proposed framework for CBA in Norwegian aviation. Section 3 discusses some important factor prices used in CBA. Section 4 deals with selected methodological topics, while section 5 gives two examples to illustrate why more comprehensive impact analyses are necessary. Section 6 concludes the paper.

2. Cost-Benefit analysis (CBA) in the aviation sector – an overview

The main purpose of CBA is to provide a basis for the choice and the ranking of projects. Thus, it is important that the analyses are comparable across different projects. CBA gives the net changes in the use of real resources, measured in monetary terms. The most important economic effects to be analysed within the aviation sector are:

- Changes in generalised travel costs, often in terms of changes in time costs. If present, changes in air ticket costs and transport costs to/from airports should also be taken into consideration.
- Environmental effects, like noise and emissions.
- Effects on safety.
- Investments and operating costs.

Network effects are relevant when flight corridors and routes are affected. These include effects on the generalised travel costs for passengers in other parts of the network, as well as for different operators.

THE INITIAL SITUATION (without the project):

Passengers		Airlines		NCAA		The State	
Benefits	Costs	Revenues	Costs	Revenues	Costs	Revenues	Costs
WTP	Tickets Time costs Other travel expenses	Tickets. (subsidies)	Operating+ capital costs Air charges CO ₂ -tax Seat tax (PSO-tax)	Air charges Other commercial activities	Operating + capital costs CNS/ATM	Seat tax CO ₂ tax (PSO-tax)	(PSO routes) Other purposes

NEW PROJECT:

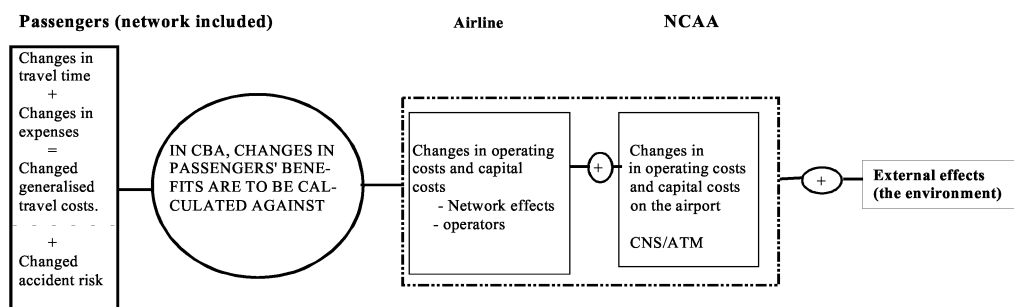


Figure 1 The flow of economic resources in the aviation sector

Figure 1 presents the flow of economic resources between agents in the aviation sector. The upper part of the figure describes the present flows in the Norwegian aviation sector. The system is driven by the passengers' willingness to pay (WTP). One part of the total WTP is the actual payments made through the air tickets. These payments constitute the revenue side for the airlines. The cost side for the airlines consists of various charges for NCAA's services, taxes to the government, and costs concerning operations and capital. The air charges constitute the revenue side for the NCAA, to cover the costs of different services. The revenues for the State consist of taxes for CO₂, and a seat tax. The PSO tax is in brackets because taxes to cover the expenses for operation of routes under Public Service Obligation are still not levied. The lower part of the figure shows the situation when a project alter the travel costs. The changes in generalised travel costs is the main factor on the benefit side, and should be traded off against the costs for the airlines and the NCAA, also accounting for the external effects. Possible network effects should also be a part of the picture.

A CBA should be considered because most often, changes in air ticket prices will not be able to internalise the changes in consumer surplus that is caused by the project. To illustrate this, Figure 2 shows a situation where the changes in air ticket revenues do not reflect the changes in economic benefits for the passengers. Consider two airports with different demand elasticity. A reduction in generalised travel costs from G_0 to G_1 gives an increase in consumer surplus (B) and increased air ticket revenues R. The difference in R does not coincide with the difference in B.

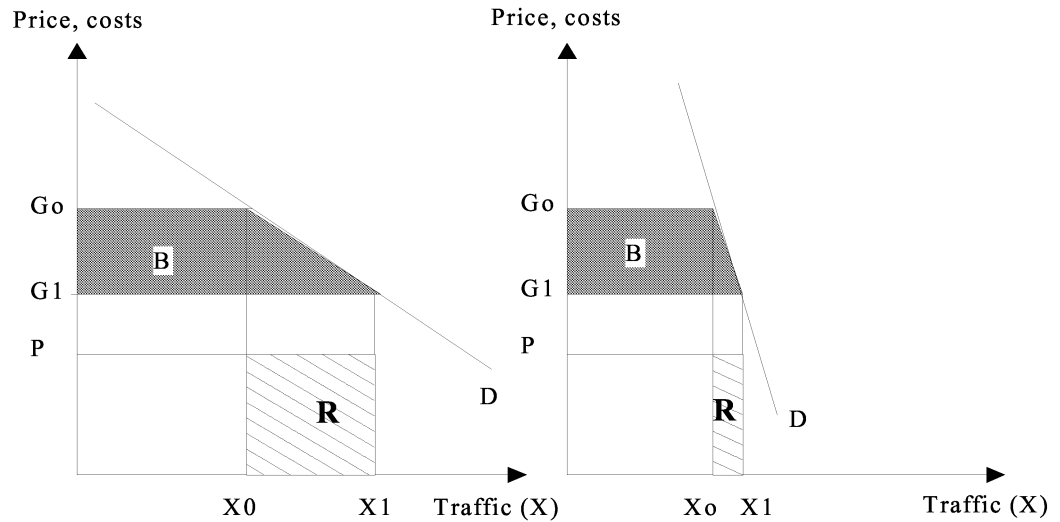


Figure 2 Revenues and consumer surplus

Here, the prices P are unchanged. This will be the practical situation in Norway, because the NCAA's charges are levied on the basis of *average cost pricing* for a network of airports, where these costs also comprise a mark-up to cover the costs of investments within a given time period. Thus, a marginal investment will not incur any change in the air charges. Even if the pricing regime varies among the European countries according to the way the civil aviation is organised (Fewings et al. 1998), there is a significant element of governmental price control based on some kind of average cost pricing. However, if the air charges can be currently adjusted to internalise the user benefits, then the changes in air ticket prices reflect the change in economic welfare, and a financial analysis will be sufficient. Likewise, if environmental effects like noise and emissions are correctly monetised, a corresponding change in air ticket revenues will internalise the value of these external effects. Say, if the costs of implementing damage-reducing measures like e.g. noise abatement are internalised in the air charges, then the WTP can be used directly to calculate whether the noise abatement expenses are being covered.

Such flexible pricing regimes are rarely present in real life. Within the NCAA's jurisdiction, the following CBA practices are recommended for the most common project categories:

1. *Projects motivated from standards or regulations.* If there are unambiguous links between the project and various standards/regulations, there is no point in doing a full CBA. A cost-efficiency analysis should in most cases be sufficient. However, there may be a case for CBA if there are alternative ways of meeting the standards, with different effects for the benefit side or the external cost side.
2. *Projects with environmental effects.* CBA should be done if there are effects for third parties that are not included in the NCAA's financial analyses. These effects may be either environmental externalities or travel costs. If the environmental effects are internalised through costs of recycling/disposal, or the implicit costs of choosing a more environmental-friendly design are captured in the financial analysis, then CBA may be omitted.
3. *CNS/ATM/ILS-services.* When considering improved equipment for navigation, communication and instrument landing, a CBA may be carried out if alternative routing of flights affect travel time, or if accident risk or the environment is affected. If standards or regulations demand such items, refer to point 1 above.
4. *Terminals.* CBA is relevant if the travel time is affected, if there are changes in airline costs (from e.g. delays or operating hours) that are not internalised, or if capacity constraints affect the adjacent transport network. Small projects are normally subject to financial analyses only.
5. *Extended runways.* These projects normally affect regularity, the size of aircraft to be served, and possibly the accident risk (however, one should be aware of "risk compensation" here). As a main rule, CBA should be carried out for such measures.

So far, the emerging CBA practice regarding aviation seem to be comparable to practice within e.g. the road sector. However, there are some methodological issues that should be focused. A couple of these issues are addressed in section 4. The next section discusses a bundle of factor prices concerning various intangibles.

3. Factor prices

Air transport projects do often affect various intangibles, like time, noise nuisance and various emissions. Monetised values are needed to assess intangibles in CBA, but generally there are either no existent markets for the effects to be analysed, or the market prices cannot be recognised as measures of the real values, due to monopolies or other market failures.

Value of time

The value of travel time saving (VOT) is frequently the largest element in CBA. As discussed in section 2, VOT is not commonly internalised in the air ticket costs, if the travel time is affected by a given measure. The purpose of the trip is an important criterion when calculating this effect. Business travels and private travels are the relevant segments chosen, and the VOT is commonly calculated by applying the average gross and net wage rates, respectively.

There have been several attempts to improve the VOT estimates. Among these is the Norwegian value of time study (Ramjerdi et al 1997). The methodology applied is two approaches to WTP analysis, Stated Preference (SP) and Transfer Price (TP). *Private travel* demand may be expressed as a function of generalised travel cost and analysed by means of a logit model. For *business travels*, attempts were made to estimate VOT by means of a revised version of “Hensher’s formula” (Hensher 1977). The underlying assumption is that both employer and employee will benefit from reduced travelling time.

	<i>Private travel</i>	<i>Business travel</i>
Travel time	173	213
Delays	260	320
Planned waiting time	19	71

Table 1 Value of time for air travels (NOK, 1998)

The estimate of VOT for business travels by air turned out to be higher than the marginal productivity of labour (the gross wage rate). From the Norwegian VOT study, it is

recommended that a conservative estimate, i.e. the average gross wage rate, should be applied till further analyses are carried out. The estimate for private travels is based on the results from the SP approach. In table 1, VOT used in CBA for air travels is shown.

The VOT for delays are applied from the equivalent Swedish value of time study (Algers & al 1995), and assumed to be 50 % higher than the VOT for travelling time.

The methodological approach in the Norwegian study has formerly been applied in similar studies in Britain (Bates & al 1987), The Netherlands (HCG 1990) and Sweden (Algers et al 1995). The results have been discussed with respect to methodological weaknesses. However, they are the most updated estimates available at present.

Accident costs

Accidents are rare in regulated air traffic. When they occur, however, the consequences are usually vast. Thus, *accident frequencies* as well as *accident costs* have to be looked into, to be able to include changes in accident risk into the economic appraisals.

Regarding *accident frequencies*, NCAA has analysed accident risks in Norwegian commercial aviation, comprising both route and charter traffic as well as other commercial traffic. From a data set consisting of all domestic flights for the period 1985–1994, accident risks are estimated (Fugleberg 1999), focusing on factors affecting the landing and take-off security. The effects of introducing certain security measures are estimated. These measures are *instrument landing systems (ILS)*, *radar surveillance* and *landing lights >720m*. In addition, security effects of terrain factors like hills or peaks are analysed. Additional estimates isolate the risk effects of scheduled¹ versus non-scheduled flights. On average, 16% of all accidents for scheduled flights are fatal.

The results show that the accident risk for scheduled flights, varies between 1.05 per million legs and 3.6 per million legs, depending upon combinations of the security measures. The effect of introducing e.g. radar surveillance (no ILS in use, which is the actual situation for many rural airports) is a reduction of accident risk from 3.0 per million to 1.5 per million. For non-scheduled flights, the accident risk varies between 2.1 per million and 62.9 per million. Introducing e.g. radar is estimated to reduce risk from 49.6 per million to 2.1 per million (still without ILS).

¹ Route and charter flights

The costs per accident consist of loss of *statistical lives*, *injury costs*, *material costs* and *administration costs*. However, injury costs are omitted because they are of minor importance in the aviation sector². From Elvik (1993), the cost of a statistical life is estimated to be NOK 17.0 millions (1998). In the CBA manual, material costs are set to 100% of aircraft value for fatal accidents, and 50% for non-fatal accidents. Average material accidents for scheduled flights are found to be NOK 44 millions, while average administrative costs are NOK 4 millions.

For an average accident on a given airport, a number of 60 passengers in the aircraft, a fatal accident risk of 0.16 and a mortality rate of 0.9 are used as an example (16% of the accidents are fatal, with 90% death rate). The material and administrative costs are as stated above. In this case, the average accident costs are:

Loss of statistical lives:	NOK	147millions
Material costs:	NOK	44millions
Administrative costs:	NOK	4 millions
 Total (1998):	 NOK	 195 millions

The average cost of one accident must be calculated separately for each airport. To calculate the expected reduction or increase in accident costs, the cost of one accident should be multiplied with the actual accident risk or by the change in risk due to the investment or the measure taken.

Environmental costs

Environment costs in the aviation sector consist of *noise* and *emissions* (local, regional and global effects). In addition we have pollution of subsoil water from de-icing. This is taken into account by adding the fully satisfying remedy costs to the cost side of the analysis.

Noise

Noise nuisance is an important source of environmental disturbance related to airfields, and is thus treated in a detailed way in the CBA manual. The basis is a WTP analysis from 1994 at Oslo airport Fornebu (now replaced by Gardermoen OSL). Here, noise

² For *fatal* accidents the risk of losing ones life is found to be about 0.9.

nuisance was significant over a long period of time (Thune-Larsen 1995). The applied method in the WTP study was conjoint analysis (CA) in combination with contingent valuation (CV), used to estimate the WTP for a 50 per cent reduction of subjective noise disturbance for citizens that suffered heavily from noise nuisance. The WTP for a 50 per cent subjective noise reduction amounted to NOK 3600 per person per annum for much annoyed persons within the noise zones. These zones comprise noise levels of 50 dBA equivalent flight noise (EFN) or more, in steps of 5 dBA. Assuming for practical purposes that the WTP is constant per percent change, a one per cent change in subjective noise disturbance has an annual value of NOK 72 per person per annum.

The result is assumed to be valid for noise increases as well as decreases, and the validity of this assumption should be subject to further research. First, there is reason to believe that the WTP is asymmetric (Sælensminde and Hammer 1994). Second, it is likely that the WTP is non-linear. One might expect a higher marginal WTP for changes in situations with high initial noise levels.

In general, 5% of the population that feel «much annoyed» by air-traffic noise inside zone 50–55 dBA, this share increasing gradually to 31% for zone 65 dBA and above. The change in subjective noise disturbance is assumed to be proportional to the logarithm of the noise pressure (dBA EFN). A map showing how the noise zones are affected by planned project, can be made. Then the change in number of people that are much affected can be identified for each noise zone. Subsequently, the value of the change in noise nuisance can be calculated. Alternatively, a more accurate registration of each household's noise disturbance, using e.g. GIS systems, can be applied.

Local and regional emissions

Local emissions up to 100 meters altitude are attributed to take-off and landing. Emissions from between 100 and 1000 meters altitude are usually classified as regional emissions. Some of these emissions are proportional to fuel consumption while others are not, and the most important categories are evaluated in the CBA. The analyses comprise the monetary values of NO_x, VOC (volatile organic components) and particles. Since there are no WTP-analyses or other recent relevant studies for Norway, values are applied from a meta-study done by ECMT (ECMT 1998).

The meta-study separates between densely and sparsely populated areas because the concentration of substances is important for the damage effect. The local and regional emissions may also have global effects, e.g. on the production of ozone or greenhouse

effects. Such effects are not considered in a local perspective. The numbers in table 2 give cost per kilogram emission.

	Densely populated areas	Sparsely populated areas
NOX	70	35
VOC	70	35
Particles	615	0

Table 2 Costs in NOK per kilogram from local and regional emissions to air.

The average emissions are calculated from the composition of the air fleet that uses the airport. The emission factors of each category of aircraft are known from other studies (e.g. Knudsen and Strömsøe 1990). Changes in aircraft movements and fleet composition are given by forecasts. The change in emission costs per year can then be calculated.

Global emissions

Global emissions cause depletion of the ozone layer and the greenhouse effect, of which the latter is at stake here. Among the gases that contribute to the greenhouse effect, carbon dioxide (CO₂) is the most important one. Here, CO₂ is considered as representative for all climate gases. All global climate costs are assigned to this gas, as it is assumed as an approximation that all relevant emissions are proportional to the fuel consumption. Some emission categories in the stratosphere are suspected to contribute more to climate changes than emissions near the ground level. Among these are NO_x and VOC. It is assumed that the prices for local and regional emissions partly compensate for that, as the ECMT-estimates partly cover the global effects as well.

It seems that neither damage costs nor WTP will give a satisfactory answer to what the best way of assessing the cost of CO₂ emissions will be. The «indirect political willingness to pay» approach is used here, i.e. agreements like Kyoto are considered as a “constraint”. The excise tax on CO₂ that will be necessary to fulfil Norway’s part of the Kyoto-agreement, are calculated, meaning that CO₂ emissions should be stabilised on the 1990-level from 2010 on. Estimates are made by means of the general equilibrium model

GODMOD (Jensen 1998), and the results show that it will be necessary to charge NOK 740 per tonne CO₂ in 2010 to stabilise on the 1990-level, for Norway's part. The main underlying assumption is that an international agreement forces all industrialised countries to charge CO₂ emissions to fulfil their part of the Kyoto agreement, within their own borders. If we assume that the charge is 0 in 1990, with a linear increment, the charge in year 2000 will be NOK 370. This may be seen as a compromise between the present and the future, not letting either of them take all the costs. However, the premises for these estimates deserve a couple of comments. First, international aviation is not a part of the Kyoto agreement. The emission costs from Norwegian domestic flights may not be too biased, but the costs for international flights are encumbered with substantial uncertainty. Second, the estimates are based on a situation without tradeable permits. Such a situation may, and probably will, reduce the emission costs with respect to these cost estimates. At the present stage however, trade of permits are not established.

The consumption of fuel per minute on a standard flight can be found as described above. The burning of a kilogram jet-fuel produces 3.15 kilograms CO₂. Knowing the average change in flight-time produced by the planned project makes it possible to calculate the CO₂-cost connected to specific projects, like e.g. changes in the ATC systems in Europe.

In the next section, some methodological topics concerning CBA in the aviation sector will be discussed.

4. CBA for aviation projects – methodological topics

The interdependency problem

Within the NCAA's jurisdiction, the value of a single project will quite frequently depend on whether other projects are implemented or not. To avoid ending up considering e.g. an entire airport system, there is a need for a practical way of dealing with this kind of interdependency. This problem will now be assessed in two ways. First, the nature of interdependency will be examined. Thereafter, some practical guidelines will be developed. The discussion is limited to projects with *sequential* interdependence.

One indicator of mutual dependency is prospective impacts in the traffic market. Figure 3 illustrates the problem of sequential interdependency, illustrated by the impacts from a stepwise upgrading of an airport.

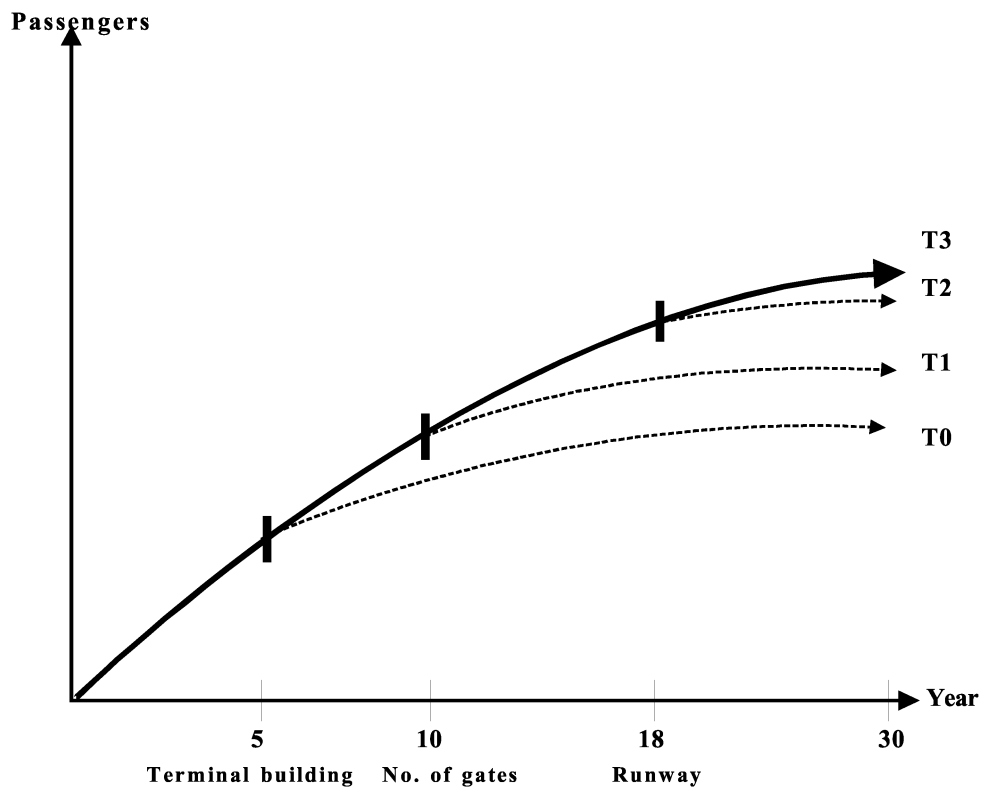


Figure 3 Traffic growth - sequential dependency

T₃ represents the traffic growth in a 30 year span, without any capacity constraints. In year 5, the lack of terminal throughput capacity limits the traffic growth to T₀. If the capacity is expanded, the growth rate will follow T₃ until the next capacity constraint occur, for instance the number of gates available. Without expansion, the traffic growth is limited to T₁. Correspondingly, if the need for extra runway length in year 18 is not met, the growth is limited to T₂. The figure shows that the benefits from the expanded terminal building in year 5 are limited if the number of gates remain unchanged after year 10. The traffic is assumed to be growing, but with an annually declining growth rate. This is so because even if there are no excess capacity during peak hours, there will be available capacity (e.g. slots) between peak hours to be used by people who are able to change travel behaviour.

The NCAA use guidelines for capacity expansion. If these guidelines take the form of mandatory regulations, then it can be said that a project A in year X will entail project B in year Y and project C in year Z. A strong sequential interdependency is established, implementing project B before project A is meaningless. This is a case for designing a *package* of projects, where the net present value for the entire package is calculated.

However, there are several points to be made when designing project packages:

1. Sequential project packages may be designed from standards that represent constraints regarding the planning and decision-making process. When the demand side is given, criteria are established to define the timing and sequence of the various measures to be taken. The projects in the package are resolved simultaneously.
2. If such constraints are not defined, one should assess different packages, and implement the most profitable one.
3. With respect to the *financial* side of the story, project packages makes it difficult to assess the financial profitability of single projects. This may be in conflict with a financial framework that deals with the profitability of single projects (as presently within the NCAA).

Figure 4 illustrates the investment costs and annual revenues from the project package in Figure 3. The situation is simplified in the sense that the annual revenues are constant when the capacity constraint is effective. The annual operating costs are omitted.

In Figure 4, area **D** illustrates the revenues in the base case. Area **C** is the additional revenues when the terminal is upgraded in year 5, area **B** is added revenues when the number of gates is increased in year 10, and **C** is added when the runway is upgraded in year 18. The **I**s are the different investments and **R** is the residual value of the investments after year 30 (the economic life span). This illustration applies for both the financial revenues and the passengers' benefits in a cost-benefit perspective. To design a framework for approaching this kind of interdependence problems in a practical way, the following guidelines have been recommended to the NCAA:

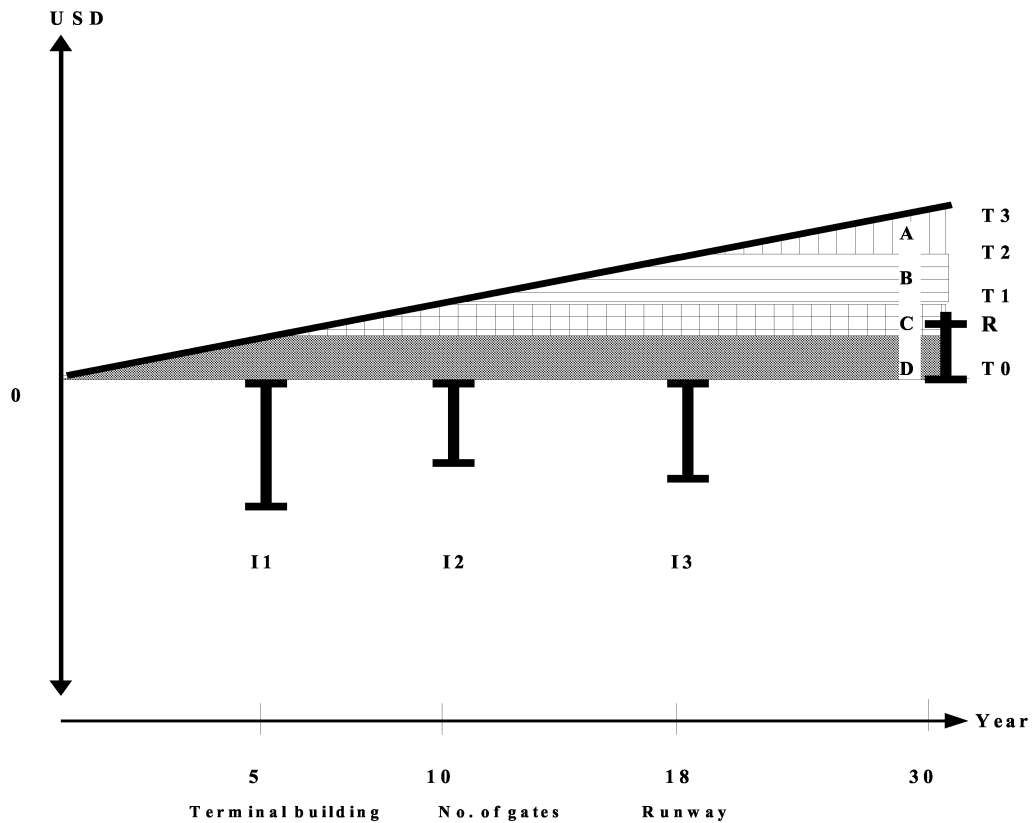


Figure 4 Costs and revenues - sequential dependency

- A. The case of *uncertain* dependency: If there are reasons to be in doubt whether the projects T_2 and T_3 will actually be carried out, then the recommendation is to consider the projects as separate and independent. The projects are then analysed with respect to declining traffic due to capacity constraints occurring at a later stage. Thus, the traffic volume that constitutes the market for e.g. project T_1 is defined by the area between the curves T_1 and T_2 in Figure 3.
- If T_1 is a profitable project on its own, then it may be implemented independently. Area C in Figure 4 is the relevant revenue area.
 - If profitability of T_1 over the life cycle is dependent of the implementation of T_2 and/or T_3 , then a strong connection through *binding decisions* on implementation of the entire profitable package should be established.

- However, linking of projects is not recommended if profitability of the package is dependent of projects to be implemented more than 10 years from now. Most often, it is not possible to make meaningful binding decisions for a very long time span. The decision tree in Figure 5 illustrates this situation.

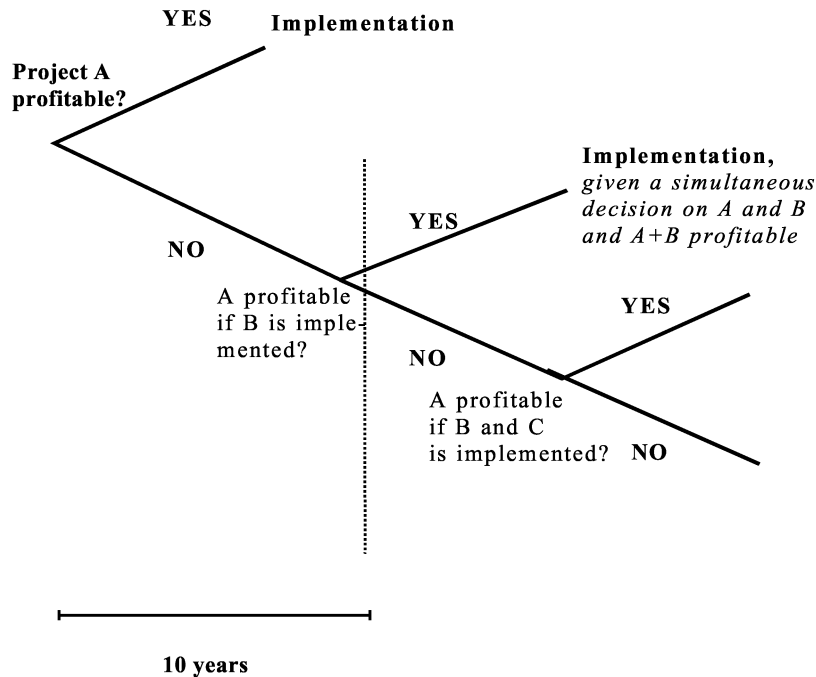


Figure 5 Decision tree - sequential dependency

- B. The case of *certain* dependency: If there is no doubt about the strong dependency, then the projects should be considered as a package:
1. Projects that are linked within a time span of 10 years are considered as one package.
 2. For projects to be implemented more than 10 years from now, the interdependency is characterised by such a degree of uncertainty that they should not be included in the analysis. In practice, capacity constraints that occur after year 10 are considered as effective.

The network problem and the value of time

Network considerations emerge frequently when doing economic appraisal of e.g. increased airport capacity. The network effects can be categorised as follows:

A. Effects within the airport network, affecting services for airlines and passengers.

B. Effects on other transport modes.

Point B will be commented upon here. The effects on other transport modes are mainly connected to traffic spill-over. When capacity constraints are reached during peak hours, several adjustments may take place. Some traffic will divert towards less attractive slots. Other parts of the traffic market will either be deterred because of higher generalised transport costs, or be diverted to other modes. These effects may seem trivial from a theoretical viewpoint, but there are some practical matters that should be carefully examined:

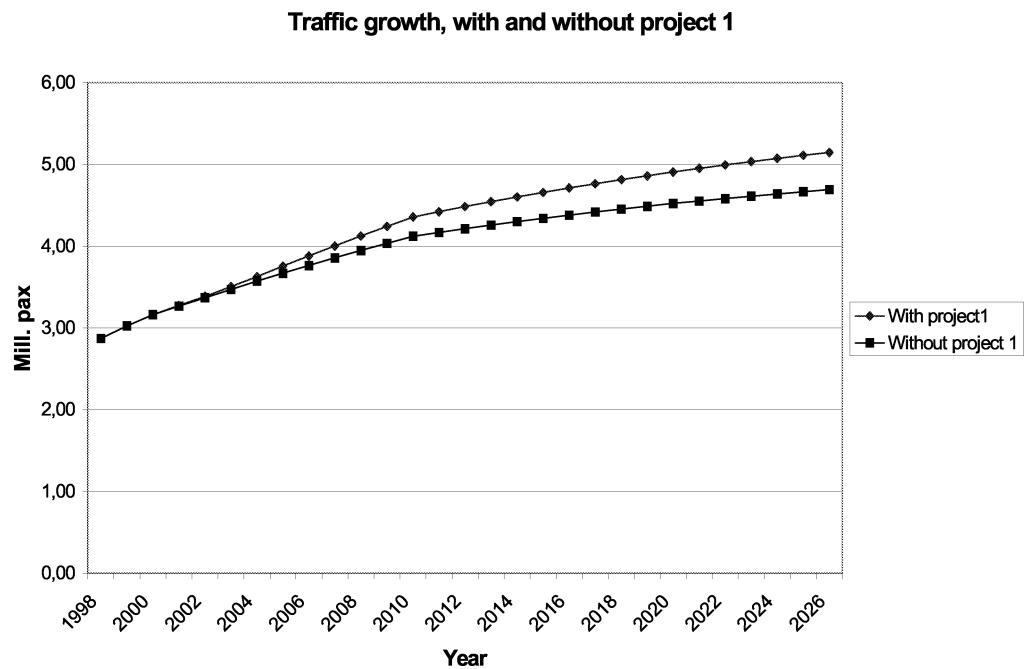


Figure 6 Traffic growth with and without constraints

1. *Traffic forecasts.* Figure 6 illustrates the number of passengers that constitute the market for improvements in e.g. airport infrastructure. The lower curve represents the anticipated traffic volumes that will use the airport under the capacity constraint, while the upper one indicates the forecasted number of passengers without the constraint. The distance between the two curves gives the forecasted amount of traffic that will benefit from the extended airport capacity, but which will either be diverted to other modes or deterred from travelling, without airport capacity expansion.
2. *Alternative transport modes, traffic deterrence and the travellers' benefits.* As an airport reaches the capacity constraint, transport costs increase abruptly. Figure 7 illustrates this situation. The marginal generalised travel costs for using the existing airport infrastructure is given by MC_A^1 . As capacity problems emerge (at volume X_0), MC increases towards infinity as the capacity constraint is reached (somewhere near X_1). The MC_R curve is the generalised travel costs of using alternative transport modes, e.g. road transport combined with the use of a nearby airport. This curve makes the generalised reservation costs for the air passenger. If the generalised travel costs by air exceed this level, some passenger will switch to the competing mode (here $X_2 - X_1$), others will give up travelling.

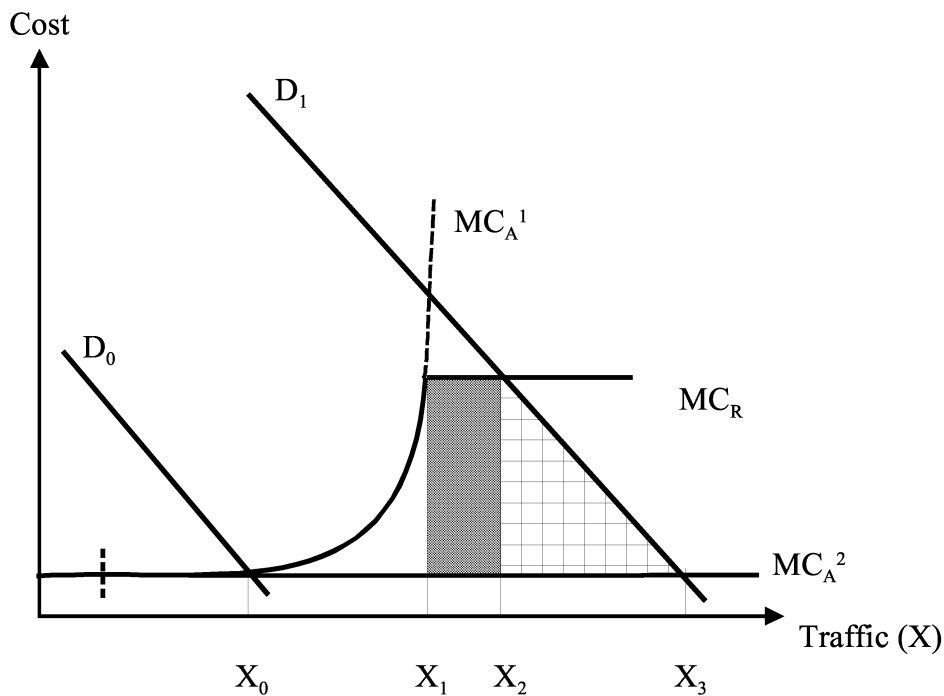


Figure 7 Benefits of increased airport capacity

The curve MC_A^2 is the generalised transport cost if the capacity is increased. For simplicity, this cost is represented as constant. The shaded and the hatched area illustrate the benefits from increased capacity. The shaded area represents the benefits for those who switch to competing modes, while the hatched area illustrates the benefits for the traffic that is deterred when the capacity constraint is in effect. In addition to these areas, there will be benefits from cost reduction for the traffic $X_1 - X_0$.

3. *The value of time in transport chains.* The value of time for a typical air passenger is significantly higher than for various surface transport modes, especially for leisure travels. Most VOT studies are directed towards partial analyses of travellers in different sectors, without studying the consumer's VOT through a chain of various modes. When using the recommended VOTs of different transport modes in an analysis that includes diversion of traffic from one mode to another, one might face a rather dramatic change in VOTs, which raises questions about the rationale for this big difference. After all, the same group of travellers is considered, although some differences may be explained from the characteristics of the transport mode. These questions are important when considering multi-modal passenger transport.

Defining the base case and the alternatives

The point of departure for a CBA is to define the project's main objective. It is important that the description of the objective does not include any specific means. As an example, if the take-off and landing routines on a given airport create unacceptable noise for the surrounding dwellings, the objective should be "reduce the noise by xx EFN" and not e.g. "build a mound to reduce aircraft noise by xx EFN). In this way, no relevant means to meet this objective are excluded. The definition of the base case is also a vital prerequisite for the CBA, because the projects or project alternatives are compared to it. The list of possible projects that meet the objectives should be complete. However, one can imagine that all the projects turn out to be unprofitable, leaving the base case as the most profitable one.

The base case should represent the expected situation during the economic life span without any changes, i.e. without any project being implemented. However, the base case should include projects that are already approved for implementation. Thus, the base case is not necessarily a 'do nothing' alternative. Correspondingly, optimising the use of the existing infrastructure should be included in the base case, as well as effects from

technological development. As an example, the number of aircraft movements on a given airport may represent a constraint that may be desirable to relax. An eventual trend towards larger aircraft will influence the forecasted aircraft movements and thus the need for increased capacity of this kind.

Ideally, the base case demand should reflect the market-clearing demand under a set of optimal prices. In Norway as well as in other countries, there is so far no tradition for optimal peak pricing concerning e.g. slot allocation. If the prices are below marginal costs, then the “excess demand” signalises a need for extended capacity that may not be founded in the users’ WTP. Thus, a risk of over-investing is present if the traffic forecasts are based on a sub-optimal pricing regime, which is often the case. To our knowledge, empirical examination of the importance of this issue in air transport as well as for other modes, e.g. within urban transport systems, is not available and may be subject to further research.

NCAA as a corporation

Up to now, NCAA has mainly operated on the basis of financial analyses, where the individual airport is the financial unit. There has been a weak tradition for consolidated account, where e.g. the entire network of airports is considered. If the situation of airport 1 influences the traffic volume on airport 2, this may give a net financial effect for the NCAA that is not represented by financial analyses at the airport level. As an example and as commented above, capacity constraints on airport 1 may transfer passengers to airport 2, which is situated in the vicinity. The revenue for the NCAA may be practically zero because the transferred passengers generate revenue to the NCAA via airport 2 with connected routes. The CBAs are in essence “global” analyses where such network effects should be taken into account. Financial analyses should also include the entire influence area to be able to capture the financial effects without double counting.

Project risk and uncertainty

Project risk and uncertainty is a major challenge for investors, regardless of whether the project is financed with private or public funds. Ideally, CBA should be dealing with *probabilities* and *expected values* rather than deterministic figures.

Traditionally, the treatment of project risk and uncertainty has been on a very crude level in most impact assessments. Recently, a national standing advisory committee on cost-benefit analyses (SAC) has given methodological recommendations for economic impact assessments in the Norwegian public sector, comprising the treatment of risk and uncertainty. Returns on private investments are usually a composite of two factors: A risk free rate of return plus a risk premium. There have been some disagreements about whether or not (or to what extent) one should include such a risk premium in the discount rate for public investment analysis. In order to discuss this question we need to divide the risk premium into two further components. *Non-systematic* risk is connected to the outcome of the specific project in question, and *systematic risk* is due to macroeconomic factors like general business cycles, fluctuations in the oil prices etc. The latter kind of risk is unavoidable and genuine even to a whole public sector. The former kind is considered as irrelevant when analysing public projects because it can be *diversified*, and because each project's share of the total budget is small. Additionally, the risk will be distributed over all taxpayers.

SAC argued, partly on this basis, that if all figures were calculated as *expected values*, the only relevant risk premium to include into the discount rate was the systematic risk component. The size of this component would in turn depend on to what extent the project returns were correlated with macroeconomic fluctuations. On the basis of such information, the recommendation was to assign typical public investments into risk categories with different discount rates. As for the NCAA, SAC's recommendations are not followed on this issue, for two reasons:

- The assumed diversity of project specific risk is not plausible for an agent characterised as a public enterprise. The risk is borne exclusively by the NCAA. Thus, the non-systematic risk is highly relevant to the NCAA as a decision-maker.
- Empirical evidence on the relevant risk premiums for typical NCAA projects is lacking, and should be subject to further analysis. A search for "market copies" of the different project categories to be able to extract the risk premium from the stock exchange, may be a fruitful approach.

The use of a single discount rate is thus recommended, with no risk-dependent discrimination between projects. However, two important steps towards a more satisfying treatment of project risk are taken. First, the use of *expected values* should replace the use of *most probable outcome*. The analyst has to assess the range of possible outcomes and the connected probabilities for these outcomes. Only rarely a formal probability

distribution will be available, but even the use of subjective probabilities will be an improvement. Second, the use of sensitivity analysis should be extended and improved, capturing realistic variations of significant, non-deterministic factors. Preferably, these sensitivities should be arranged into scenario analysis, where e.g. the effects of general business cycles can be illustrated with respect to factor prices and traffic forecasts combined.

5. CBA in practice: Two examples with simulated numbers

Example 1: Terminal at Stavanger airport (SVG)

The medium-sized Stavanger airport (SVG) face capacity constraints regarding terminal capacity. This is assumed to be the limiting growth constraint in the near future. The discounted investment costs of a new terminal are NOK 275 millions. The residual discounted value of the investment after 30 years is NOK 7 millions. The increased operating costs amount to NOK 170 millions, discounted over a 25 year span. Without expanded capacity, one segment of the growing number of passengers will be transferred to less attractive slots until these slots are full. The remaining segment will use the services of neighbouring airports, or refrain from travelling, causing increased transport costs and a welfare loss from traffic deterrence. The situation is described more formally in section 4. Figure 8 shows the SVG airport and the neighbouring airports. Without extended capacity, people use car to the airports in Kristiansand (KRS) and Haugesund (HAU), and air transport from there. Bergen (BGO) is situated too far to serve as an alternative airport. Passengers with Bergen as their final destination will now go there by car, or desist from travelling. In this case, the surface transport to nearby airports (KRS and HAU) takes between 2 and 3 hours, resulting in extra time costs, vehicle operating costs and accident costs as the most significant effects.

The main benefit of extended terminal capacity will be the travel time saving (TTS) of travelling from SVG instead of using HAU and KRS as alternative airports. In addition, TTS for travellers with KRS and BGO as final destinations are taken into consideration. To be able to calculate these benefits, traffic forecasts have to be made. In addition, the following factors have to be assessed:

- The travel purposes (business travels and others).
- The destinations from SVG.
- The level of diverted traffic, and the distribution between HAU and KRS.
- Time costs and other costs for surface transport to adjacent airports.
- The share of travels that are deterred because of higher transport costs.

In addition, delays for passengers and operators during peak hours can be expected. The benefits of avoiding delays may be significant, and should be given careful attention. The details in the forecasts of traffic volumes and delays are not examined in this paper.

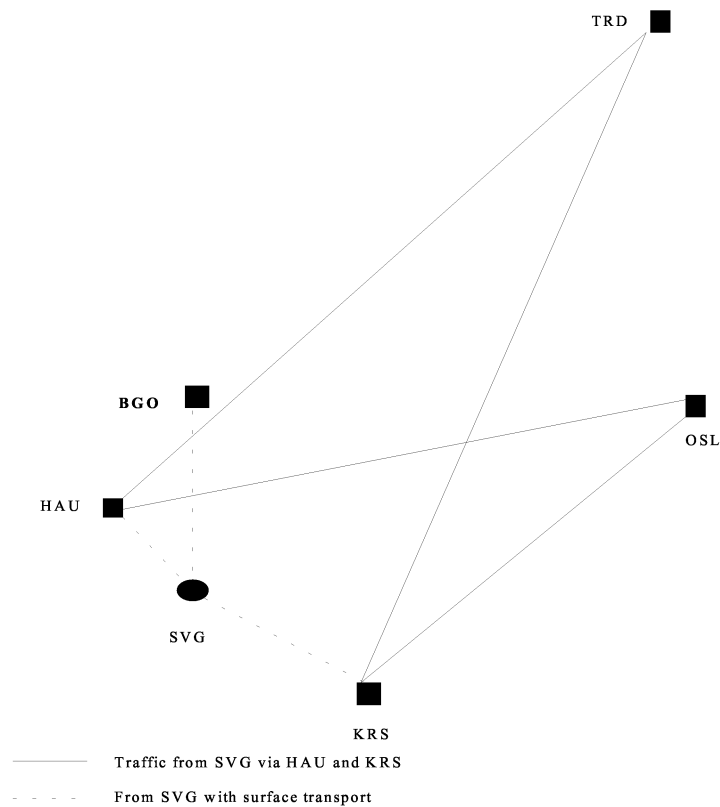


Figure 8 Alternative routes, travels from SVG

These benefits are held against the investment costs and operating costs for both the NCAA and the operators. Environmental effects and the value of changes in accident risks are also to be included. The accident risk assessment is done by comparing the risk change when passengers are transferred from air to surface transport. Especially when the actual surface transport is by car, a significant rise in accident risk in terms of expected loss of statistical lives has to be included in the calculations. All the economic effects within the project's life span are discounted, and the net present value is calculated, using an appropriate discount rate.

In addition to the net present value (NPV), the most important effects are presented for various agents like the NCAA, the State, the passengers, the operators and the environment. The presentation method is chosen to inform the decision-makers of both the economic and the financial effects. Table 3 illustrates the real economic effects, including external effects and excluding fiscal taxes. The financial cash flow is also illustrated. The resource flows and cash flows from Figure 1 are recognised.

The benefits adds up to NOK 582 millions, exceeding the costs of 541 millions. Thus, the net present value is NOK 41 millions and from an economic perspective, the project should be carried out. The NPV/Cost ratio should be > 0 , and is approximately 0.1 here. This ratio is a pure *ranking criterion*. The cost side in the NPV/C ratio is the payable costs for the NCAA, and gives NPV with respect to the use of NCAA's budget. However, the main decision criteria should be that $NPV > 0$.

The new SVG terminal appears to be economic profitable. However, the financial cash flow gives a negative result of NOK 329 millions, meaning that this project would not have passed a financial analysis. It is the travel time saving and the decrease in accident costs that give the most significant contribution to the difference between the economic and the financial assessment here. As pointed out in section 1, this kind of difference gives reasons to discuss the situation for self financing agents in markets where elements of intangible goods constitute a significant share of the economic impacts.

Costs and benefits SVG example	Economic impacts		Financial cashflow	
	Benefits	Costs	Revenue	Cost
C1 Passenger benefits				
C1.1 Business travels	370			
C1.2 Others	0			
C1.3 Delays	54			
C1.4 Ticket costs				1550
S1 Accident costs	137			
S1.1 Statistical lives	n.a.			
S1.2 Injuries	n.a.			
S1.3 Material damage and administration	n.a.			
E1 The environment				
E1.1 Local and regional emissions (NO _x , VOC, particles)		6		
E1.2 Global emissions (CO ₂)		36		
P1 The State/general taxes				
P1.1 Fiscal aviation taxes			92	
P1.2 VAT on NCAA investments and operating costs			62	
P2 Operators/airlines				
P2.1 Operating costs		122		122
P2.2 Aviation charges to the NCAA				95
P2.3 Fiscal aviation taxes				92
P2.4 Ticket revenues			1550	
P2.5 Delay costs	21		21	
<i>SUM financial cashflow operators/airlines</i>			<i>1571</i>	<i>309</i>
P3 Other commercial activities (shops, taxfree)			n.a.	22
P4 NCAA				
P4.1 Investments		223		275
P4.2 Residual value		-7		
P4.3 Maintenance and operations		161		171
P4.4 Revenues from aviation taxes			95	
P4.5 Other commercial activities			22	
<i>SUM financial cashflow NCAA</i>			<i>117</i>	<i>446</i>
<i>SUM economic impacts</i>	582	541		
Net present value (NPV)				41
NPV/Cost ratio				41/446 ≈ 0,1

Table 3 SVG example, economic impacts and financial cash flow (mill. NOK).

Example 2: Instrument landing system (ILS) at Molde airport (MOL)

A local airport (MOL) is served by two air services: One following a triangular route from the national hub (OSL) via MOL and the neighbouring town to the north (KSU) and back to OSL. This route is serviced by Boeing 737s with five flights per day. The other service is a coastal route between the regional hubs BGO and TRD, which is servicing MOL and KSU both on the northbound and the southbound route. This coastal route is serviced by Fokker F50s. The situation is illustrated in Figure 9. Currently, MOL has an ILS system for landings from the west. When western winds exceed 5 m/s the approach is done by a circulating manoeuvre from the west, followed by a landing from the east. Western winds totally dominate the weather-conditions in the area, meaning that the vast majority of landings are currently done from the east.

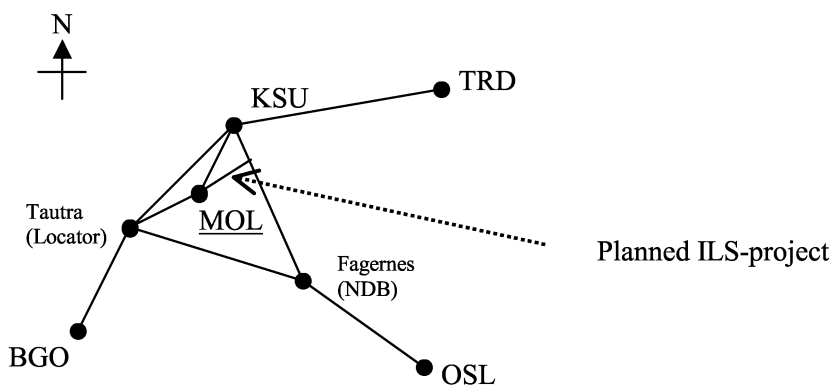


Figure 9 The flight connections serving MOL

The NCAA wants to carry out an impact analysis of building an ILS system for landings from the east. The necessary initial investment for such a system is estimated to be NOK 5 millions and the system is likely to have a economic life of 15 years. Within a planning horizon of 25 years this gives a discounted (7% discount rate) total investment of NOK 6.4 millions. The investment will after 25 year have a residual value of NOK 0.3 millions. Discounted maintenance and operating costs are estimated to amount to a total figure of NOK 2.4 millions.

ILS systems are usually motivated by their anticipated effect on flight safety. However, in this case the system will also have the most important effects on both the user benefit side, and on the operator side. Apart from the northbound BGO-MOL leg, approaching

the airport from the west represents a *detour* from all connected airports. This means that passengers currently face an estimated travel time loss of between 2.5 and 5 minutes compared to a direct approach from the east. Correspondingly, the operators would save aircraft operating costs from a shorter average leg, and consequently there would also be a positive environmental impact due to less fuel consumption.

The impacts are divided into effects for the NCAA, for the operators, for the passengers, for the operators and for the environment. In order to illustrate the financial side, figures are included that illustrate the cash flow for the operators and for the NCAA. Table 4 illustrates the real economic effects, including external effects and excluding fiscal taxes.

Costs & Benefits ILS example	Economic impacts		Financial cashflow	
	Benefits	Costs	Revenue	Costs
C1 Passenger benefits				
C1.1 Business journey	13.1			
C1.2 Other journeys	8.6			
S1 Accident costs				
S1.1 Statistical lives	0.3			
S1.2 Injuries				
S1.3 Material damage and administrative costs	0.1			
E1 The environment				
E1.1 Local and regional emissions (NOx,VOC,Part.)	2.4			
E1.2 Global emissions (CO2)	10.9			
P1 The State/general taxes				
P1.1 Fiscal aviation taxes			0.0	0.0
P1.2 VAT on NCAA investments and operating costs			1.3	0.1
P2 Operators/airlines				
P2.1 Operating costs	13.6		13.6	
P4 NCAA				
P4.1 Investments		4.9		6.1
P4.2 Residual value	0.2		0.3	
P4.3 Maintenance and operations		2.2		2.4
<i>SUM financial cashflow NCAA</i>			<i>0.3</i>	<i>8.4</i>
<i>SUM economic impacts</i>	<i>49.2</i>	<i>7.2</i>		
Net present value (NPV)		42.0		
NPV/Cost ratio		42.0/8.4		
		=4.98		

Table 4 MOL example, economic impacts and financial cash flow

The benefit side amounts to NOK 49.2 millions, and the costs totals NOK 7.2 millions. Thus, the project has a net present value of NOK 42 millions, and a net benefit/cost-ratio of almost 5. The ILS is evidently profitable in economic terms. It would probably be ranked among the top projects in comparison with other projects within a budget constraint. Generally, the ILS is mainly implemented as a safety measure. However in this example, the major benefits are travel time reductions and less environmental damage.

Here the financial analysis provides a contrary result. The project does not generate revenues at all, but causes extra investment and maintenance costs for the NCAA. However, if the cash flow analysis had included the airlines' cost savings, these savings would have exceeded the NCAA's cost increase. This example do give reasons to extend the scope of impact assessment from financial cash flow analyses towards a broader view provided by a full economic impact assessment.

6. Conclusions

In this paper, CBA as an economic impact assessment approach for projects within NCAA's jurisdiction, is reported. As a rule, measures like e.g. airfield investments entail intangible effects that are not included in a financial analysis, because the pricing of aviation services rarely do reflect changes in economic welfare. Even if there still are problems to be solved, regarding e.g. the value of time in multi-modal passenger transport, the method is applicable to a wide scope of projects within aviation. Two examples are presented, which are based on realistic, but not real numbers. However, they illustrate the decision dilemma where unprofitable financial results are achieved on one hand, and profitable socio-economic outcomes on the other.

The NCAA is responsible for its own financial viability through revenues from user charges and other commercial activities. At the same time, the NCAA's role in the Norwegian transport sector calls for a broader economic appraisal of projects within its jurisdiction. Making priorities subject to a full economic impact analysis instead of a more narrow financial cash flow approach, may significantly alter the portfolio of projects that reaches implementation. The dilemma for the NCAA is obvious: One may very well face situations where the financial analysis gives recommendations contrary to what can be learnt from the economic impact assessment. The significant elements of intangibles may call for some kind of public-private partnership, to be able to take the welfare effects properly into account.

References

- Algers, S., Lindqvist Dillèn, J., 1995. *Value of Time Study, Final Report, Part 1*, Transek. Sweden.
- Bates, J. et al., 1987. *The Value of Travel Time Savings*. ITS, Leeds University, England.
- Bråthen, S., Eriksen, K. S., Hjelle, H.M., Killi, M., 1999. *Economic appraisal in Norwegian aviation. Manual*. Møre Research/Institute of Transport Economics, Norway (in Norwegian, English summary).
- Bråthen, S., Eriksen, K. S., Hjelle, H.M., Killi, M., 1999. *Economic appraisal in Norwegian aviation. Examples*. Møre Research/Institute of Transport Economics, Norway (in Norwegian, English abstract).
- ECMT, 1998. *Efficient Transport for Europe – Policies for Internalisation of External Costs*. European Conference of Ministers of Transport, OECD.
- Elvik, R., 1993. *Økonomisk verdsetting av velferdstap ved trafikkulykker*. Oslo, TÖI-report 203/1993, Institute of Transport Economics, Oslo, Norway (in Norwegian).
- Federal Aviation Authorities (FAA), 1997: *FAA Airport Benefit-Cost Analysis Guidance*. Draft report, FAA Office of Aviation Policy and Plans.
- Fewings, R., Morrell, P., Pagliari, R., Whelan, C., Williams, G., 1998: *A Comparative Study of European Airport Provision*. Air Transport Group, College of Aeronautics, Cranfield University, UK.
- Fugleberg, O., 1999. *Estimering av risikoreduksjon av tiltak rettet mot flysikkerhet*. Working paper, NCAA, Norway (in Norwegian).
- Hensher, D. A., 1977. *Value of Business Travel Time*. Pergamon Press, London.
- Jensen, T., 1998. *Alternativ transportteknologi. Reduserte CO2-utslipp fra transportsektoren*. TÖI-report 413/1998. Institute of Transport Economics, Oslo, Norway (in Norwegian).

Knudsen, S., Strømsøe, S., 1990. *Kartlegging av utslipp til luft fra norsk sivil luftfart*. NILU OR 88/90 og 89/90 (in Norwegian).

Ramjerdi, F. et al., 1997. *The Norwegian Value of Time Study. Part 1*. TÖI-report 379/1997. Institute of Transport Economics, Oslo, Norway.

Sælensminde, K., Hammer, F., 1994. *Verdsetting av miljøgoder ved bruk av samvalgsanalyse*. TÖI-report 251/1994. Institute of Transport Economics, Oslo, Norway (in Norwegian).

Thune-Larsen, H., 1995. *Flystøyavgifter basert på betalingsvillighet*. TÖI-report 289/1995. Institute of Transport Economics, Oslo, Norway (in Norwegian).

Prologue to Paper Two:

Strait crossings and economic development

Developing economic impact assessments by means of *ex post* analyses

4. Prologue to Paper Two: Strait crossings and economic development. Developing economic impact assessments by means of *ex post* analyses

4.1 Introduction

Paper Two deals with the economic impacts of five strait crossings. It presents the results from an econometric analysis that identifies the road users' *ex post* willingness to pay (WTP) for a fixed link compared to the previous ferry operations. A model for assessing the feasibility of toll financing compared with public funding is also presented. In addition, aspects regarding regional impacts of strait crossings are examined.

In the *ex post* analyses, the cost-benefit analysis (CBA) is used as an economic impact assessment (EIA) method. The kind of *ex post* analysis that is applied here can be considered as a convenient way of informing future assessment practice, as proposed by EC DGVII (1994):

- The development of methods for quantitative assessment of the strategic economic impact of road and transport projects, distinguishing clearly those benefits that are genuinely additional to the transport user benefits, is an important priority.
- The development of ex-post evaluation methods to complement ex-ante appraisal is recommended both as a check on appraisal accuracy, and to inform future appraisal practice.

The use of 'true' WTP is an important element in economic impact assessments, and for optimising the overall provision of transport infrastructure. In these five cases, the assessment of the fixed link is done in comparison with ferry operations, because the point of departure is to replace a running ferry service with a fixed link. This raises a few questions regarding how such a comparison should be made. If there are elements in the WTP that are 'transport mode specific' (here: the choice between two kinds of 'infrastructure' for fjord crossings), then the factor prices used in the cost-benefit analyses are of course affected.

The rest of this introduction is organised so as to give a more profound insight into selected parts of the paper. Section 4.2 examines the EIA method for comparing the profitability of a ferry operation compared to a fixed link. Section 4.3 assesses various financing aspects. A broader discussion regarding regional impacts of fjord crossings is left to the prologue for Paper Three.

4.2 Ferry operations or fixed links - some comments

4.2.1 A model for selecting the optimal mode for fjord crossings

In Norway, there has been a lot of concern about the profitability of replacing ferry connections with fixed links. The usual procedure for comparing these systems has been to compare the fixed link with the existing ferry services, and to recommend the replacement if it turns out to be economically profitable. What follows is a theoretical model that considers the aspects of analysing a continued ferry operation as an alternative to a fixed link. The model can be useful in structuring the decision problem. The point is that these two technologies are analytically inter-related, any measures to improve the ferry connection will have an effect on the CBA results for the fixed link. The simultaneity problem can be illustrated by presenting two situations that are subsequently connected. The first one is the classical scheme where a fixed link is analysed, while the other deals with improved ferry operations. Savings in time costs, fares¹ and vehicle operating costs are compared with construction and operation costs. These costs elements are usually referred to as *generalised travel costs* (see e.g. Button 1993). The generalised costs can be defined as:

$$G = g(C_1, C_2, \dots, C_n) \quad (1)$$

where G is generalised costs and $C_1 \dots C_n$ denote the various cost elements. Figure 4.1 illustrates the concept of generalised travel cost saving for a fixed link.

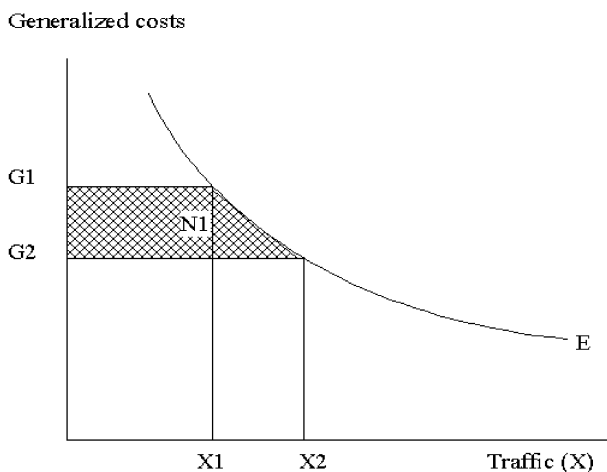


Figure 4.1 Generalised travel cost saving for a fixed link

Assume that the generalised travel costs are reduced from G_1 to G_2 if a ferry service is replaced by a fixed link. Traffic increases from X_1 til X_2 . The project is socially profitable if the net present value (NVP) > 0 :

$$NPV_1 = K_1 - \sum_{i=1}^n \left[B_{1i} \left(\frac{1+r_1}{1+r} \right)^i - D_{1i} \left(\frac{1+r_2}{1+r} \right)^i - E_{1i} \left(\frac{1+r_3}{1+r} \right)^i \right] + R_1 \quad (2)$$

where i = years
 n = economic life span (often 25 years)
 r = discount factor
 r_1 = annual benefits real growth rate
 r_2 = annual operation/maintenance costs real growth rate
 r_3 = annual environmental costs real growth rate
 B_{1i} = annual benefits of the fixed link compared with the ferry services
 D_{1i} = annual net operation costs (operation costs ferry minus operation costs for the fixed link)
 E_{1i} = annual net environmental costs (environmental costs ferry minus environmental costs for the fixed link)
 K_1 = investment costs (present value)
 R_1 = residual value of the investment after n years (present value)
 NPV_1 = net present value

Equation (2) allows for real growth rates on the various benefit and cost elements. Apart from the traffic growth rate (expressed through r_1), the real growth effects r_2 and r_3 are usually assumed to be zero. One exception is if the traffic growth raises demand for extended capacity or increased wear and tear. B_{1i} expresses the changes in generalised travel costs, as defined in equation (1). The net operating costs D_{1i} are most often negative when dealing with replacement of ferry services, because ferry operations are more costly². This means that D_{1i} often represents cost savings that give positive contributions to the NPV. The environmental costs are given a wide interpretation here. They include cost elements that are not internalised, like various emission and accident costs. The residual value reflects the market value of the investment beyond the economic life span. Normally, the economic life span of infrastructure is assumed to be 25 years, while the technical life span is supposed to

¹ For simplicity, it is assumed that the fares represent the marginal cost of ferry operations. In Norway, the ferry fares are determined partly to satisfy equity considerations. Some of the implications are commented upon below.

² The costs of ferry operations are substantial, and they can be a critical factor in CBA of fjord crossings.

be 40 years. The 15 years period of remaining operations has a value R_I that is included in the CBA³.

The traditional analytical framework seeks the most profitable solution for a fixed link through various technical designs, compared to the running ferry services as the base case. If the NPV exceeds 0, then the project is generally recommended. In Norway, there is no tradition for seeking the optimal design of ferry operations as an integrated part of the analysis regarding the optimal design of a strait crossing. Alternative ferry operations should however be treated as full alternatives, and one should evaluate if an improved operational design of the ferry connection is the best solution in terms of discounted social profitability.

Such an evaluation might reveal that improved ferry operations are more profitable than the existing services. If this turns out to be the case, then the initial level of generalised costs to which the fixed link is compared will decrease (reduced G_I in Figure 4.1). This simply means that the 'base case' for a fixed link should be the socially most profitable ferry service concept. More frequent ferry services will e.g. result in reduced time costs. If such a profitable operation design exists that reduces the generalised costs⁴, this will lead to lower benefits N_I in Figure 4.1. This may in turn imply that the fixed link can turn out to be unprofitable.

Thus, fixed links and continued ferry operations should in this analytical framework be considered simultaneously with the purpose as to identify the most profitable fjord crossing concept. The theoretical process is as follows: In the analytical process one has to look for profitable improvements until the last marginally profitable improvement is found. When this last measure is found then one may have a bundle of profitable measures to improve the design of the ferry operations. Now one has to consider the level of generalised costs G_I that should be the point of departure for determining N_I in Figure 4.1. If a fixed link (bridge or tunnel) is still profitable (the reduction in generalised costs for all cars still gives sufficient benefits to cover the costs) then the fixed link should analytically be chosen as the most profitable design. If the fixed link turns out to be unprofitable then a continued but improved ferry operation should be preferred from an economic point of view.

Figure 4.2 illustrates a scheme for analysing alternative designs for ferry operations. The area N_2 illustrates that the improvement will benefit both those travelling at the existing travel costs G_I (traffic X_0) as well as the generated traffic given by $X_I - X_0$. The ferry fare is given by price P . The area I gives the increase in revenues caused by the (bundle of)

³ In the Norwegian road sector, R is assumed to be 15/40 of the investment costs K for road pavements. R is discounted. It is a simplification that appears to be acceptable (NPRD 1995).

⁴ A similar effect may of course take place if the improvements affect D_i and E_i in equation 2, instead of or in addition to effects on G_I . This aspect will not be discussed any further here.

measures for improving the ferry operations. The figure can illustrate another point: The traditional analytical step has been to carry out a financial analysis to determine whether the area *I* (increased revenues) is sufficient to cover the augmented operation costs. The analysis is relevant for assessing the need for a raise in public expenditure and is of course of interest from a governmental point of view. A financial analysis is however not sufficient in an economic context, because the benefits from e.g. time cost reductions are not included in the financial analysis.⁵

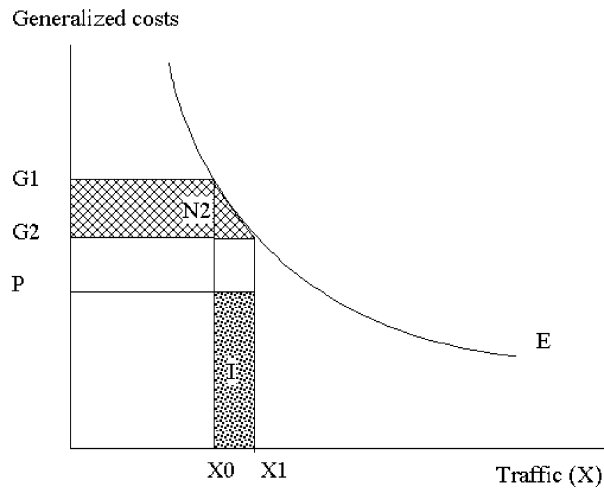


Figure 4.2 Reduction in generalized costs due to improved ferry operations

It should therefore be necessary to carry out a full CBA for the measures planned to improve the ferry operation design, i.e. a computation of the area *N₂* in Figure 4.2. The economic profitability of using improved ferry operations to reduce the generalised travellers' costs from *G₁* to *G₂* can be evaluated by using equation (3):

$$NPV_2 = K_2 - \sum_{i=1}^n \left[B_{2i} \left(\frac{1+r_4}{1+r} \right)^i - D_{2i} \left(\frac{1+r_5}{1+r} \right)^i - E_{2i} \left(\frac{1+r_6}{1+r} \right)^i \right] + R_2 \quad (3)$$

⁵ If all the economic effects of an infrastructure investment are internalised in the user charges, then there is a case for financial analyses only. The implication of the flat fare system for Norwegian ferries under the NPRD's jurisdiction, is further commented upon below. A general recognition is that the financial analyses do not include all relevant economic effects of transport infrastructure investments.

where i	= years
n	= economic life span (often 25 years)
r	= discount factor
r_4	= annual benefits real growth rate
r_5	= annual operation/maintenance costs real growth rate
r_6	= annual environmental costs real growth rate
B_{2i}	= annual benefits of the improved service compared with the existing one
D_{2i}	= annual net operation costs (the difference in operation costs between the planned and the existing ferry service)
E_{2i}	= annual net environmental costs
K_2	= investment/ferry acquisition costs (present value)
R_2	= residual value of the ferries after n years
NPV_2	= net present value

If the expression $NPV_2 > 0$ then the project is profitable. The calculations of the benefits of the fixed link must now be revised because the reduced generalised costs G_1 in the new 'base case' (improved ferry operations) will affect the profitability of the fixed link. G_2 in Figure 4.2 now corresponds to this reduced G_1 in Figure 4.1, which will be the new benchmark for assessing the change in user benefits from implementing a fixed link. If NPV_1 falls below 0 then improvement of ferry operation design is the most profitable measure, while the fixed link has become unprofitable.⁶

Among the 5 fixed links that are examined in Paper Two, two are feeding traffic to and from city centres (the Ålesund connection and the Askøy bridge). In the traffic forecasts, a certain amount of generated traffic⁷ is expected. The generated traffic may shift the demand for using the road network in the adjacent city. In a given situation with little excess capacity in the road network, a situation as described in Figure 4.3 may take place. The feeding of traffic into the city centre shifts the city traffic demand curves from D_0 to D_1 . The private generalised costs PGC is the generalised travel costs that the motorists face, while the social generalised costs SGC also include the external time costs that are caused by the congestion. Without any congestion charges the traffic increases from X_0 to X_1 . Because only the PGC are paid, the SGC are not fully covered by the road users. Thus, the

⁶ If now $NPV_2 > 0$ the condition of economic profitability for the fixed link can be expressed as:

$$\alpha = (B_1 - B_2)/(C_1 - C_2) > 1$$

where B is benefits (1=fixed link, 2=ferry) and C is costs. If $\alpha > 1$ then alternative 1 (the fixed link) is the most profitable. The equation expresses the ratio of change in user costs (benefits) to the change in system costs of the two alternatives.

⁷ Generated traffic can be defined as the traffic growth that is taking place within one year after opening, with corrections for the underlying traffic growth that would have taken place irrespective of the new infrastructure. Generated traffic takes place because of the shifted generalised travel costs. For fixed links, the generated traffic may normally imply a rise in the traffic volume of between 20 and 50%.

crosshatched area denotes the external economic costs. If such congestion in the adjacent road network takes place, then the value of this area should be subtracted from the NPV of the fixed link. If a congestion charge t (Figure 4.3) is levied (augmenting the generalised travel costs from G_1 to G_2 for the traffic X_2), then the traffic is reduced to X_2 , and the congestion costs are *internalised*. The generated traffic level for the fixed link is then reduced by $X_1 - X_2$, affecting the profitability of the fixed link. In the *ex post* analyses in Paper Two, the problem of increased congestion in the road network adjacent to the fixed links is not addressed.

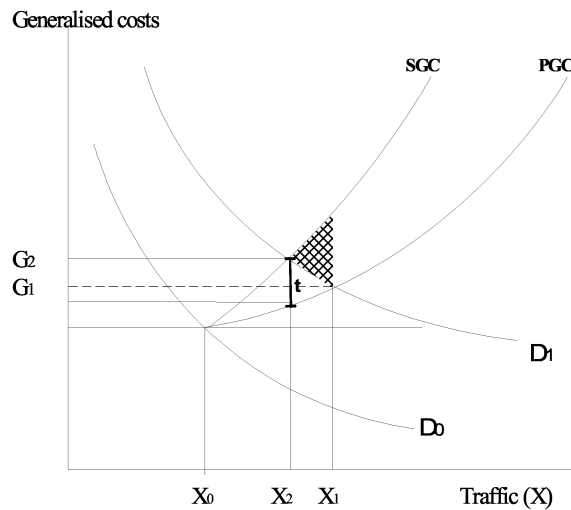


Figure 4.3 Generated traffic and congestion in the adjacent road network

4.2.2 A few comments on factor prices

The reliability of CBA depends on reliable factor prices. The issue of factor prices has to be carefully examined. The most significant items in the cost-benefit account are:

- The time costs for the travellers (in-vehicle time, waiting time)
- Inconvenience costs for scheduled transport
- Vehicle operating costs
- Environmental costs
- Ferry operating costs
- Construction and operating costs of the fixed link

The value of time (VOT) in CBA is based on extensive empirical research. The recent VOT that is used in Norwegian transportation can be found in Killi (1999), based on Ramjerdi et al. (1997). For a discussion of the theoretical background for these studies, see e.g. Bruzelius (1979). Bråthen (1995) showed that properties regarding specific fjord crossings might give a case for deviations from the VOT values that are conventionally used. If significant proportions of the transport market are holiday tourism, then the utility functions may not comprise WTP for faster fjord crossings but for the strictly opposite, namely (slow) ferries that allow preferred relaxed enjoyment of the fjord scenery. A survey among tourists in two fjord crossings (tourism constituted about 50% of the transport markets here) concluded that more than 50% preferred continued ferry operations instead of the fixed links that were being planned. There were reasons to interpret these expressed preferences as indicating that the value of travel time saving was assigned a value of zero. A significant reduction of user benefits was the resulting effect.⁸ In Figure 4.1 this means that G_1 is reduced, and that the area N_1 becomes smaller. These findings support the argument that a close examination of the local transport market is necessary when assessing economic impacts.

In Paper Two, the existence and level of the inconvenience cost (IC) is discussed. The IC may be a significant element when calculating the user benefits, and will hence be a part of B_{1i} in equation 2. Paper Two examines five fixed links, where the IC is significantly positive in four of them. The paper does neither give support to a *statistical generalisation* of the IC concept, nor does it explore the *nature* of the IC. In general, however, there are theoretical reasons to expect that scheduled transport has 'costs of inflexibility', where the efficient use of time is affected. Another aspect of scheduled transport is the perceived uncertainty regarding available capacity during peak hours for both the vessel itself and the adjacent road network. People have to allocate more travel time to deal with this kind of uncertainty. A third element in the IC is connected to increased 'mental distance' in connection with ferry operations. The value of knowing that e.g. police and health care services can be provided on an unscheduled basis, is also probably a part of the IC⁹. High-frequency scheduled ferry operations combined with a short-distance adjacent network with adequate capacity may make the IC approach zero. One of the five links (The Ålesund connection) meets these criteria. It is thus desirable to gain experience from *ex post* analyses

⁸ Neither of these fixed links (the Hardanger Bridge in Western Norway and the North Cape Connection) was economically profitable in the first place. NCC is however currently under construction. HB is postponed, partly as a consequence of these findings.

⁹ The concept of 'mental distance' is most probably also present for people that live in the area without (or very infrequently) using the ferry services. The value of knowing that such services can be provided swiftly is an 'optional value' that may not be observed fully through the road users' WTP. See Bråthen and Hervik (1992) for a discussion of optional values in the rural air transport market.

on a larger number of fixed links, and to carefully examine the properties of new projects to check whether such experience can be transferred.

Another important aspect when dealing with fjord crossings is the costs of ferry operations. They are usually computed from a standard model. Bråthen (1993) provides an example of such a model, which is used by the Norwegian Public Roads Directorate (NPRD 1995). It is based on a cost function consisting of annual sailing distance, fuel prices and crew costs. The annual sailing distance is based on the length of the fjord crossing, the departure frequency and the daily hours of operation. The departure frequency and duration of daily operation vary with traffic volume and type of connection (local road link, trunk road or urban area) and are given in the NPRD's guidelines for ferry operation services. Such guidelines raise a crucial point concerning the assessment of ferry operation costs in the CBA. The predicted traffic in the economic life span of the project and the service guidelines determine when to expand the capacity by means of increased number of departures and/or larger vessels. There is no evidence indicating that these guidelines for determining the quality of ferry operations are economically optimal.

It is most likely that the results of the CBA are sensitive to the kind of *investment criteria* that are used for expanding ferry capacity. The existing NPRD guidelines provide a set of rules that are imprecise regarding economic efficiency. In addition, a flat fare system is used that causes deviations from marginal costs. A fare system designed in accordance with the theory of pricing under capacity constraints (in Norway also under the theory of financial constraints as most ferry companies receive subsidies) would give the transport market proper signals of the costs of supplying additional capacity. The absence of such price signals may lead to excess capacity both initially and during the economic life span. In turn, this may lead to biased ferry operation costs in the CBA for both improved ferry operations and for the fixed links. Ferry costs are often a significant element in these assessments. This affects the B, C and E elements in equation (2) above and all the benefit and cost elements in equation (3). Figure 4.4 illustrates a typical demand situation for a ferry connection.

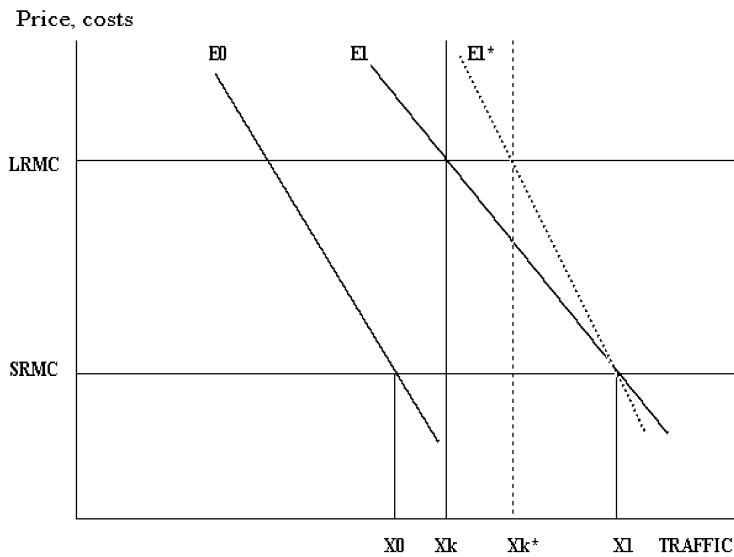


Figure 4.4 Peak and off-peak demand

E_0 is the demand in the off-peak period, while E_1 is the demand in the peak period. LRMC is the long run marginal cost of ferry operations (marginal costs including capacity costs), while SRMC is the short run marginal costs. If the ferry fares are considered equal to SRMC for simplicity, then the demand in the off-peak period will be X_0 . The demand in the peak period will be X_1 . Due to the capacity constraint given by X_k the observed excess demand (queue) in the peak period is $X_1 - X_k$. Since the efficient investment criterion is that a capacity increase is profitable when willingness to pay (WTP) exceeds LRMC then no increase in capacity is necessary from an economic point of view ($WTP = LRMC$ for demand E_1). Ideally, the supply of ferry capacity X_k should be where the market clears at price = LRMC.

If the queue $X_1 - X_k$ is observed and some benchmarks are established without knowing the WTP, by establishing guidelines like: "maximal queues to be allowed are $n\%$ of ferry capacity pr. month" then investments can be made, say, to expand capacity to X_{k^*} . In the illustrated demand situation this results in *excess capacity* $X_{k^*} - X_k$ from an economic efficiency point of view, because WTP is not present for the extra capacity costs (the demand curve runs below LRMC for $X > X_k$) even in the peak period E_1 .

If now the demand curve in the peak period is estimated to e.g. E_1^* , then an expansion in capacity to X_{k^*} would be efficient (X_{k^*} is the point where WTP marginally covers the LRMC). Note that the volume X_1 for price = SRMC is equal for both E_1 and E_1^* . The difference is the demand elasticity and hence the WTP. For illustration, E_1 can be viewed as peak load demand consisting of non-business traffic, while E_1^* can be seen as peak load

demand for business travellers with higher value of time. The short-run marginal cost curve (SRMC) indicates the optimum price level in the off-peak period.

The design of a fare system will often be a subject of political interest. A fare system based on efficient pricing has up to now not been feasible from equity considerations. Such a pricing regime would imply that prices should vary with time of day and between fjord crossings of the same sailing distance. It should be a task to calculate and bring forward the difference in ferry costs under today's flat fare system compared with the ferry costs under an economically efficient fare system.

4.2.3 Selecting ferries or fixed links - a few conclusions

The first part of Paper Two provides an *ex post* study of five fixed links along the Norwegian coastline. In four of them, a significant inconvenience cost (IC) was identified. When going from ferries to fixed links, the IC is an additional user benefit. Projects which at first appear as unprofitable may be recommended when the IC is included. In this section, a model is presented that shows why it is important to assess improved ferry services as alternatives to fixed links, to be able to select the most profitable mode for the fjord crossing. The IC may play an important role when doing these kinds of assessments. This introduction also shows that in addition to the IC, the value of time (VOT) may be biased if a large segment of the users (tourists) actually prefer to use time to experience the fjord scenery. Unfortunately, none of the typical scenic fjord connections were among the five cases in the *ex post* study. Hence, it is not possible to say anything about the level of IC in the cases where fjord cruise experiences may outweigh faster accessibility in the users' utility functions.

Another critical element in the CBA of fjord crossings is the costs of ferry operations. There might be a significant bias in the calculation of ferry operation costs if the criteria actually used for the supply of ferry operations diverge from the allocation of capacity based on economic criteria. Such a divergence might imply that 1) the supply of ferry services can initially be biased and 2) the forecasted needs for increased capacity during economic lifetime of the project are non-optimal, both regarding the scheduling and the size of the expansion. This implies that changes in capacity costs may enter the cash flow at the wrong time and at a biased level. There has been research on the potential for optimising the efficiency of ferry operations. Försund (1993) found that there might be an efficiency potential of about 30% in the overall Norwegian ferry operation measured against the best technology frontier. The authors pointed out that more research had to be made to get more reliable estimates for subgroups of vessels and ferry companies. It is worth noting that this report does not take the possibly biased capacity conditions from flat fares into

consideration. At that time, a pilot study reported by Bråthen (1994) indicated that capacity conditions were a minor problem in the Norwegian ferry sector. There are reasons to carry this study a bit further, because the results indicate that the situation may actually be as described in Figure 4.3 with excess capacity due to lack of capacity pricing. If this is the case, the potential for improved effectiveness may be higher than the indicated 30%. In fjord crossing projects where ceased ferry operations count for the largest share of the benefits, a significant sensitivity towards such bias may be present.

4.3 *Toll financing of fjord crossings*

4.3.1 The Norwegian scheme for financing of toll roads

Paper Two presents a model for selecting the optimal financing scheme for toll roads where the options are either public financing or private financing, or a mixture of these. In this section, a backdrop for the financing of fjord crossings is presented.

In Norway, tolls are used to supplement the scarce amount of government funds allocated for road infrastructure, allowing infrastructure projects to be implemented earlier than would be the case with government funds alone. All the projects are sanctioned by the Norwegian parliament, the Storting. The National Public Roads Directorate (NPRD) is responsible for planning, construction, maintenance and building of roads and toll collection facilities. Private companies are often responsible for toll collection and the provision of private funds.

Experience from private financing of public roads dates back to the 1930's. Today, there is a variety of projects with private financing or joint ventures of private and public funding in Norway. These projects comprise main roads, fjord crossings and toll rings around the major cities. A few of these have, in periods, faced economic problems. One project is literally bankrupt, but most of them are running quite well. However, important questions concerning equity aspects, the planning procedure, public and private risk taking and economically sound designs of the financial packages have been raised over the last ten years.

The remainder of section 4.3 is organised as follows: First, the Norwegian scheme for mixed private and public financing is presented, both with respect to fjord crossings and toll rings around the major cities. Second, whether private sector financing has influenced the road planning and decision-making process is assessed. These aspects lead to the third point, namely how experiences with financial difficulties through a period of economic

recession in the beginning of the 1990's influenced the way *risk* is handled today, also considering the local or central governments role as a guarantor for private loans. Fourth, a theoretical discussion on private and public road financing is given, and research on the economic aspects with respect to the choice of financing regime is briefly presented.

4.3.2. The scheme for mixed public and private financing

The use of tolls as a financial instrument for road construction in Norway has increased considerably in the last decade. From 1930 to 1980, less than 5 percent of the total Norwegian road budget came from tolls. Today, more than 25 percent of the total annual budget for road construction comes from fees collected from more than 30 toll projects spread throughout the country.

For each individual toll project, local authorities and private interests establish a dedicated toll company. The main objective of the company is to operate the road toll system and to administer the revenues. The Ministry of Transport and Communications defines the operating regulations. A toll company is thus a non-profit company where the investors receive no share dividends.

It may be hard to believe that private involvement is possible given the above principles. However, toll financing has to be seen in connection with the local initiative that is present in all toll projects in Norway. These local initiatives reflect the perception of the existing infrastructure as inadequate. Usually, the counties and municipalities affected by the planned infrastructure control the toll company as owners. The public ownership derives from the fact that most roads financed through toll revenues involve a share of public funding as well. Typically, the local authorities are guarantors for the private loans. The counties and municipalities are motivated by the desire to see a road project carried out, a project that might not have been realised without road tolls. Through the collection of supplementary toll revenue, local authorities can get the project on the priority list for national highways. The willingness to raise extra private funds often means extra government funds for that region. This desire for extra government funds leads to local political pressure on both local and central decision-makers.

4.3.3 Private financing and the decision-making process

One matter of concern has been to what extent the possibility of private financing may influence the decision-making process. Bråthen et al. (1995) discussed how the interplay between local and central government combined with the option of private financing could explain why the actual ranking of projects is not in accordance with economic efficiency criteria. The main reason for such deviation during the last decade may be found in the rigid structure of public road financing. A nearly fixed share of public funds has been allocated to each county during the last 20 years¹⁰. Within each county, project rankings are made. Especially in rural Norwegian counties there is scarcity of profitable projects.

Historically, there has been strong political belief in road infrastructure as a mean for achieving regional balance. Such equity considerations may at least partly explain the way public funds for roads have been allocated. Another explanation of sub-optimal project ranking may be that in the period 1994-97 the counties competed for funds granted in addition to the ordinary budget (NPRD 1993). These funds were allocated according to the projects' economic profitability and did not depend on the historical distribution of funds among counties. This allocation gave the counties incentives to transfer profitable projects to this competition for extra funds, and not enter them in the ordinary ranking procedure. A third explanation may be that private financing with little or no share of public money makes it hard for politicians to resist temptation to show their benevolence towards the local community.

There has been some concern about whether the options of private financing combined with strong local pressure for user payments, contribute to locked-up processes. Such processes may bind the Ministry and the Norwegian Parliament (NP) in their proposals and decisions. Nyborg (1996) have interviewed former representatives in the NP's Committee for Transport and Communications. One of the claims was that privately financed projects in Norway are often large and complex with a long time span for planning and a lot of work and money involved, and they benefit from articulated local enthusiasm. Thus, such essential features make them difficult to reject. The toll financing options are considered as a particularly favourable factor when the projects are given priority in the decision-making process.

¹⁰ The *aggregate level* of the public funds for roads has decreased by about 15 per cent during the period 1993-98.

4.3.4. Risk assessment in toll financed projects

Fjord crossings are mainly expensive projects with an array of risk elements both on the cost and benefit side that can influence both economic profitability and financial viability. The procedure for impact assessments outlined in section 4.1 comprises numerous factors that are relevant for risk assessments. Ideally, the risk analysis should focus on the main factors summarised below:

- The probability distribution of the critical factors determining the revenue forecasts. The most critical factor will be the traffic volumes and the WTP in the transport market. Robustness in modelling the transport market is crucial.
- The uncertainty in the investment costs, where the use of expected values is recommended (NOU 1997:27 1997).
- The probability distribution for factors like the market interest rate for private loans, the inflation rate and the development in personal income. This requires complex analyses comprising factors affecting these elements, like the excise and tax policy and international business cycles.

Equation (4) is a simple model for investment cost coverage. It describes the critical factors having direct impacts on the ability to cover the private investment costs (Hervik and Bråthen 1990):

$$Y_1 - \sum_{i=1}^N [X \cdot P \cdot \lambda \cdot q^i - (Y_i \cdot r_i + D_i)] = 0 \quad (4)$$

where

- Y_1 = investment costs
- X = traffic before opening
- P = unit tolls at opening
- λ = induced traffic factor
- q^i = “Payback multiplier”, a compound exponential growth factor (inflation and annual traffic growth, i = year 1...N)
- Y_i = remaining debt, year i
- r_i = nominal interest rate, year i
- D_i = operating costs year i borne by the private investment

The most critical element in equation 4 is the exponential growth factor, q^i , where the traffic growth rate together with the rate of inflation contribute heavily to debt payment

ability. E.g. if a period of 20 years is considered, an inflation rate of 2.5 per cent and a traffic growth rate of 2 per cent give a payback multiplier of 2.4 in year 20. These numbers reflect the situation since 1993 in the Norwegian economy. Going 15 years back, the Norwegian economy faced an inflation rate of nearly 10 percent and a traffic growth rate of 3 percent. These numbers give a multiplier of 12.2 in year 20. Projects that were planned with the financial assumptions from the mid 80's faced severe difficulties when facing the conditions in the 90's, with lower traffic growth and higher interest rate in real terms.

4.3.5 Risk sharing - a practical approach

The experiences gained from financial difficulties in a couple of fjord crossings during the past ten years give reason to recommend some kind of risk sharing between the public and private sector considering quasi-public goods like road infrastructure. If private investors are handling risk mainly caused by macro-economic factors, this violates the principle saying that the agents who will be exposed to the risk factors should to a certain extent be able to affect them. Risk connected to construction costs can be reduced by e.g. doing seismic analyses.

The role of the public authorities as a guarantor for loans taken by the toll companies has been questioned. An important intention of risk sharing arrangements has been that the agent's perception of risk should coincide with the 'true' financial risks. One problem with public guarantees is that the local agents' risk perception may be weakened. This is the main reason why state guarantees are not recommended. Guarantees at the county or municipality level are considered more appropriate because they enhance the risk perception at the local level. However, the magnitude of the infrastructure investment costs makes redemption of such guarantees prohibitive for the public economy at the local level. In addition, the conditions under which the projects may run into financial problems are correlated with a high probability of a recession in the economy at large, affecting the local public economy. The risk profile is thus inadequate because the factors that may cause financial problems at the project level may also cause severe problems for the local public guarantors. They may not be in the position to carry their financial responsibility without severe consequences for other activities like education and health services at the local community level.

To be able to do better risk assessment that reduced the probability of financial deficits, a *conditional reimbursement* model has been used as a vehicle for risk sharing between private agents and the authorities in Norway. The model has been approved by the

Ministry of Transport and Communications as a way of providing local guarantees within adequate financial limits (Samferdselsdepartementet 1991). The basis for agreement is that the project as well as the private financing concept is approved by the local public authorities. The main components in the model are: (a) The share of private financing in the projects is set from conservative estimates on WTP (reflecting the appropriate toll level) and the traffic volumes with respect to generated traffic and traffic growth. Thus, the critical factors in equation (4) are assessed. (b) The share of public financing from the Ministry may be advanced by the toll company with refunding from the authorities later on. (c) If the toll revenues exceed the careful estimates from (a), then the refunding in (b) is wholly or partly dropped. These public funds are instead used on other highly preferred projects *within* the local community that is acting as a guarantor for the private loans.

The absence of state guarantees combined with the possibility of using such 'surplus' funds at the local community level augments the incentives to make risk estimates. However, if the toll revenues turn out to be insufficient, no governmental grants will be provided by the Ministry. Thus, the system of conditional reimbursements carries no guarantees, but it has the following pros:

1. The local authorities (mainly the county) get incentives to assess the project risk.
2. Careful estimates are made regarding the revenue possibilities compared to costs.
3. The local authorities have to consider the toll financed project in connection with other projects receiving public grants.
4. If the risk assessment and the careful assessment of possible revenues turn out to be positive, the county may act as a guarantor.
5. A public guarantee by the county gives more favourable interest rates for private loans.

The main con seen from the local authorities' point of view is:

1. The careful estimates of the possible toll revenues often call for a higher share of public financing, thus causing a higher pressure on limited public funds. The result may be a postponed implementation of desired projects.

The basis for the projects' viability is partly the road users' WTP for the improved infrastructure. Thus the transport market assessments that are one of the main topics in Paper Two, is also highly relevant for the financial impact assessments and the design of the public-private partnership in this kind of road infrastructure financing.

4.3.6 Costs of tolling and costs of public funding

The choice of private versus public funding raises questions about the costs of funds. The demand function of the actual project constitutes the basis for calculating the efficiency loss if prices exceed marginal costs. Efficiency loss occurs when consumers with $WTP \geq MC$ are not served. The allocation loss together with the costs of running the toll collection system constitutes the real costs of private funding. Figure 4.5 shows the allocation loss from toll financing.

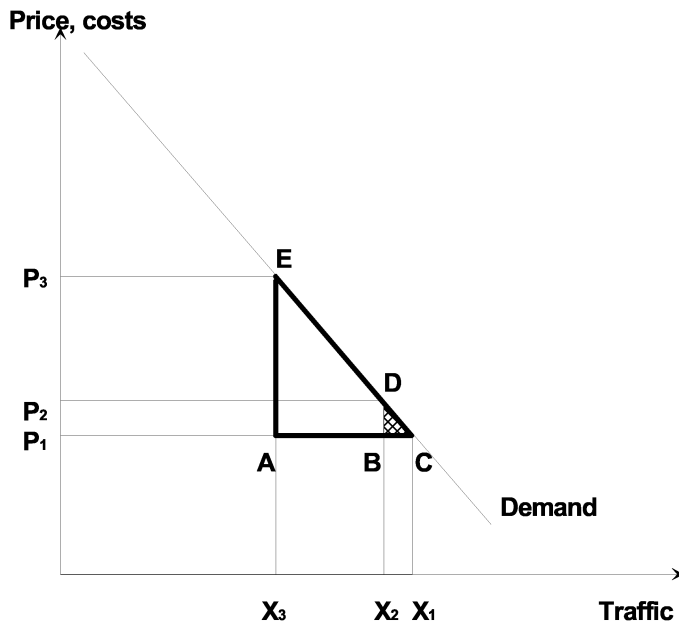


Figure 4.5 Allocation loss when price > MC

Price P_1 denotes the price level that equals marginal costs. Here, P_1 can be taken to represent the generalised travel costs, with full economic cost coverage¹¹. First, a situation can be considered where the tolls are set at a high level. This can be necessary for two reasons, namely to cover a large share of the investment costs and to compensate for a relatively modest traffic volume (as often is the case in fjord crossings). A higher price P_3 then entails traffic deterrence $X_1 - X_3$ and the allocation loss is equal to

$$A_t = \frac{(P_3 - P_1)(X_3 - X_1)}{2}$$

¹¹ All the economic costs of road usage are internalised.

A_t corresponds to the area AEC in Figure 4.5. If the tolls are set at the lower level P_2 then the allocation loss is reduced to the hatched area BDC. A lower level of tolls can be set when a larger traffic volume makes high tolls unnecessary or the toll financing constitute a minor share of the total investment costs. The elasticity of demand and the design of the discount system are elements that should be considered in order to minimise the allocation costs.

The toll collection costs and the allocation cost should be compared to the cost of public funding, which occurs because public money has to be collected through excises and taxes imposed on other sectors. This causes a situation where the consumers in other sectors face prices exceeding marginal costs. Figure 4.6 illustrates the principles behind costs of public funds.

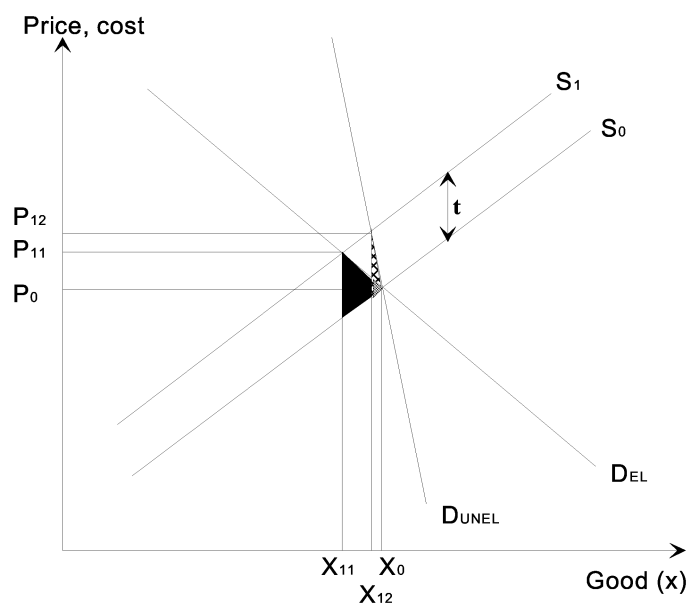


Figure 4.6. Allocation loss from taxation

A tax t is imposed, shifting the supply curve from S_0 to S_1 . In a case with elastic demand D_{EL} , the price rises from P_0 to P_{11} . The deterrence effect is given by $X_0 - X_{11}$, and the black+grey area constitutes the allocation loss. If the demand is less elastic (D_{UNEL}), then the price rises from P_0 to P_{12} , and the deterrence effect is given by $X_0 - X_{12}$. The allocation loss is significantly reduced (the grey+hatched area). It is thus apparent that the markets that are chosen for taxation should have inelastic properties.

The level of allocation costs has been subject to extensive research. Vennemo (1991) has estimated a cost of public funds to between 80 and 90% of the public expenditure if increased income tax is used, and between 35 and 40% if increased VAT is used to cover increased public expenses. According to these results, one monetary unit used in public funding costs between 1.35 and 1.8. There are other studies giving lower costs. The Norwegian Standing Advisory Committee on Cost-benefit Analysis recommends a cost of public funds of 1.2 (NOU 1997:27 1997).

In the selection of financing regime, these costs of public funding should be compared with the allocation loss and the collection costs that go with private toll financing. A model for doing this, based on estimated demand elasticities in the transport markets, is presented in Paper Two. This model also makes it possible to take the scheduling of the infrastructure investment into consideration with respect to the choice of financing regime. One common problem has been to assess the economic effects of postponements in public investment because of scarce funding, and to compare with toll financing which allows much faster implementation of e.g. fjord crossings or urban arterial road systems.

References

- Bruzelius, N. (1979) *The value of travel time: theory and measurement*. Croom Helm, London.
- Bråthen, S. and Hervik, A. (1992) *Nytten av kortbanenettet*. Working paper M9204, Möreforsking Molde, Molde, Norway.
- Bråthen, S. (1993) *Ferjekostnader i konsekvensanalyser*. Report, NPRD (The Norwegian Public Roads Directorate), Oslo.
- Bråthen, S. (1994) *Kapasitetsprising i ferjesektoren*. Working paper M9403, Möreforsking Molde, Molde, Norway.
- Bråthen, S. (1995) Ferjer – en attraksjon i turist-Norge? *Samferdsel* 34 (5) 8-11.
- Bråthen, S. and Hervik, A. (1992) *Kryssing av fjorder. Samfunnsøkonomiske lønnsomhetsberegninger ved ny infrastruktur*. Report 9205, Möreforsking Molde, Molde, Norway.
- Bråthen, S., Hervik, A., Nettet, E., Bræin, L., and Sunde, Ø. (1995) *Erfaringer med bompengeprojekter*. Report 9507, Möreforsking Molde, Molde, Norway.
- Button, K. J. (1993) *Transport economics, 2nd ed.* Edward Elgar, Aldershot, UK.
- EC DGVII (1994) *Cost-benefit and multi-criteria analysis for new road construction*. Report no. EURET/385/94, European Commission DGVII, Brussels.
- Försund, F. R. (1993) *Productivity growth in Norwegian ferries*. Reprint series no. 438, Department of Economics, University of Oslo, Oslo 352-373.
- Hervik, A. and Bråthen, S. (1990) *Økonomiske analyser av nye bru/tunnelprosjekter*. Report 9003, Möreforsking Molde, Molde, Norway.
- Killi, M. (1999) *Anbefalte tidsverdier i persontransport*. TÖI report 459/1999, Institute of Transport Economics, Oslo.
- NOU 1997:27 (1997): *Nytte-kostnadsanalyser*. Finans- og tolldepartementet, Oslo.

- NPRD (1995) *Håndbok 140 Konsekvensanalyser, del 1*. Norwegian Public Roads Directorate, Oslo.
- NPRD (1993) *NVVP 1994-97 Norsk veg- og vegtrafikkplan*. Norwegian Public Roads Directorate, Oslo.
- Nyborg, K. (1996) *Some Norwegian politicians' use of cost-benefit analysis*. Discussion paper no. 169, Statistics Norway, Research department, Oslo.
- Ramjerdi, F., Rand, L., Sætermo, I-A. F., Sælensminde, K. (1997) *The Norwegian value of time study*. TÖI report 379/1997 (2 vol.), Institute of Transport Economics, Oslo.
- Samferdselsdepartementet (1991) *St.meld. 46 1991/92 Bompengemeldingen*. Ministry of Transport and Communications, Oslo.
- Vennemo, H. (1991) *An applied general equilibrium assessment of the marginal cost of public funds in Norway*. Report no. 62, Statistics Norway, Research department, Oslo.

5

Paper Two:

STRAIT CROSSINGS AND ECONOMIC DEVELOPMENT

Developing economic impact assessment by means of *ex post* analyses

Svein Bråthen, Molde University College, Norway

Arild Hervik, Molde University College, Norway

*Reprinted from Transport Policy Vol. 4 (1997) No. 4 pp. 193-200.
© 1997, with permission from Elsevier Science*

Paper Two:

STRAIT CROSSINGS AND ECONOMIC DEVELOPMENT

Developing economic impact assessment by means of *ex post* analyses

Svein Bråthen, Molde University College, Norway

Arild Hervik, Molde University College, Norway

ABSTRACT: There is widespread scepticism among politicians and authorities concerning CBA as a main instrument for decision-making. This may cause possible bias towards non-quantitative factors as decisive. The challenge is to improve the methodological approach, carefully examining the real benefits for road users. One way of doing this is to develop *ex post* CBA to test how critical assumptions fits to reality. This paper focus on *ex post* analyses carried out with five case studies in larger infrastructure projects to assess the economic profitability actually occurring from new infrastructure. There are substantial differences in benefits *ex post* compared with the *ex ante* analyses, mainly explained by a shift parameter in willingness to pay (WTP) for the improvements. These 'Inconvenience Costs' saved by the road users add to the traditional value of time benefits. The paper also examines the profitability of private versus public funding and ends up with some considerations on regional impacts of transport infrastructure. A short presentation of a pilot study from two of the fixed links to elicit the influence on local industry is given. Some informal results supports the theory of forward and backward linkages from new economic geography.

1. Introduction

Various methods for evaluating the impacts of transport infrastructure are in practice, and there has been a lot of concern whether or not transport infrastructure causes significant benefits in the economy. Some (Aschauer 1989, 1991 among others) claim that there are significant impacts from infrastructure on output at the national level. These analyses have been heavily criticized (see Jansson 1991, Hulten & Schwab 1993 and Berechman 1994 among others), mainly due to the aggregation problems of infrastructure capital, the direction of causality and the problem of extrapolating history into the future. The last element may be relevant because of the diminishing marginal utility of infrastructure. There are also difficulties in transferring experiences from product functions based on aggregated time-series data to the micro (project) level.

There is a professional controversy between macro- and microbased methods for Economic Impact Assessments (EIA). A common objection against the Cost-Benefit Analysis (CBA) is the problem of including external effects (e.g. network effects) into the analytic framework. If significant, these should be captured in the benefit measured in the macro studies. On the other hand, the advantage of the CBA method is the ability to cover the distinctive characteristics of particular projects on the micro level.

The report from EU DGVI EURET/385/94 “Cost-Benefit and Multi-Criteria Analysis for New Road Construction” (ITS/Leeds), ends up with the following main conclusions (p.vi):

- The development of methods for quantitative assessment of the strategic economic impacts of road and transport projects, distinguishing clearly those benefits which are genuinely additional to the transport user benefits, is an important priority.
- The development of ex-post evaluation methods to complement ex-ante appraisal is recommended, both as a check on appraisal accuracy, and to inform future appraisal practice.

In Norway, recent experiences from new infrastructure projects show that the traffic volumes in some strait crossings have deviated from the predictions in the *ex ante* CBA. These deviations have gone both ways compared to predictions. This indicates a

willingness to pay (WTP) which is different from the *ex ante* calculations. Most of these infrastructure projects are funded by private capital, and the revenues depend heavily on the traffic volumes and the willingness to pay (WTP) for improved infrastructure services.

The paper starts with a theoretical model for explaining the deviations in traffic and how this influences the CBA. The demand curves for five case studies are estimated. The main part of the paper presents *ex post* CBA from these crossings. The profitability of private versus public spending are briefly presented. Some regional dimensions that might not be properly treated in the traditional CBA are discussed, and the paper ends up with a brief presentation of a pilot case study carried out to elicit how improved transport infrastructure may cause possible conditions for cumulative economic growth in the areas surrounding 2 of these fixed links.

2. The theoretical model

The model is a traditional market scheme with generalized travel costs (time costs, vehicle operating costs and tolls) along the Y axis and traffic volume on the X axis (Figure 1). The benefits of improved infrastructure in the traffic market are measured as changes in the generalized travel costs, i.e. the reduced travel costs result in a movement along the demand curve to a new equilibrium point. This equilibrium with corresponding traffic volume and generalized costs is calculated in the *ex ante* CBA. If the traffic volumes and traffic flows *ex post* corresponds with the *ex ante* estimates, then it is straightforward to see that the WTP for the new infrastructure corresponds with the results from the *ex ante* CBA, using standard value-of-time rates. Experiences from some larger infrastructure projects in Norway indicate that this is not the regular case. The most common situation is that the traffic volumes deviate significantly from the *ex ante* estimates.

One way to deal with the problem of WTP for improved infrastructure is to identify what we have termed the "inconvenience costs", or WTP for full travelling flexibility when comparing e.g. a fixed link with fixed-scheduled ferry operations. Ferry operations imply constraints on travelling behaviour, e.g. no 24-hour services, capacity constraints and so forth. There will be a WTP for not having to take these inconveniences into account which is not visible in the traditional CBA where standardized time costs are used. Full flexibility can *a priori* be considered of particular value to industries where punctuality

and quality in production are dependent on a high transport standard to maintain (or achieve) market positions. Such inconvenience costs are discussed in Jones and Hervik (1992).

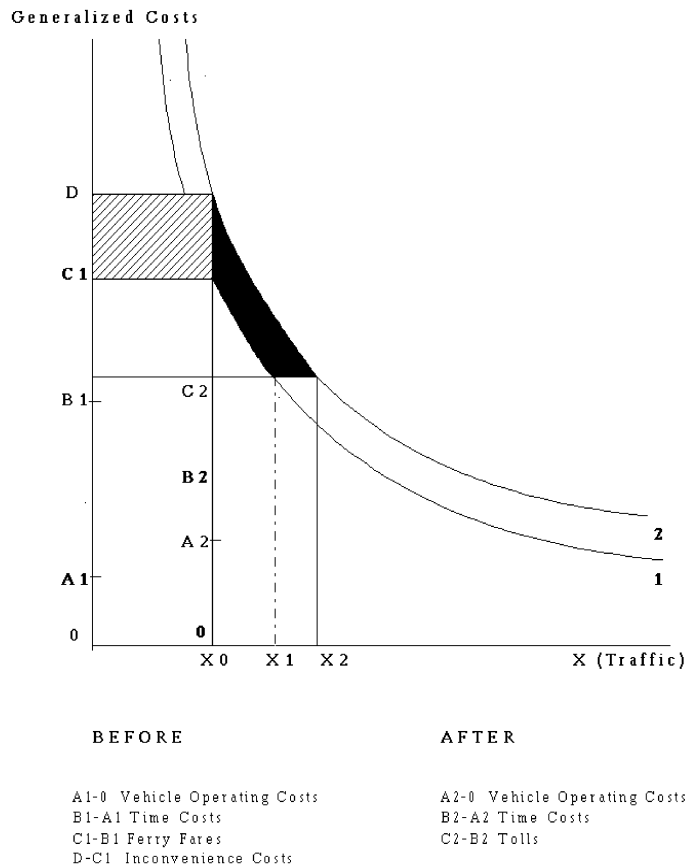


Figure 1 Inconvenience costs for existing and generated traffic

Figure 1 shows the concept of inconvenience costs, where the increase in traffic volume will be larger than what can be read from the initial demand curve estimated (curve 1), given the generalized travel costs reduction from C_1 to C_2 . The traffic observed is X_2 instead of the expected volume X_1 . The demand curve has shifted to the right (curve 2), leaving generated traffic $X_2 - X_0$ instead of $X_1 - X_0$. The shifted demand corresponds to the "true" generalized costs D instead of C_1 . The hatched area now illustrates the *inconvenience costs* for the initial traffic X_1 , and these costs are equal to the observed additional increase in WTP for improvements due to new infrastructure for the existing traffic. This value will also influence the calculated benefit for the generated traffic and cause a rise in user benefits depicted by the shaded area in Figure 1.

The generalized travel costs calculated in the original appraisal consists of the sum of travel time costs, both in-vehicle time and time on board the ferry; waiting time costs, vehicle operating costs and ferry fares. Thus, the value of waiting time is included as unweighted, i.e with the same value as in-vehicle time. Various studies have revealed a higher value for waiting time than for in-vehicle time, but these studies consider waiting time for other modes than ferries.

The estimated inconvenience costs thus encompasses the sum of any extra waiting time costs, planning/uncertainty costs of scheduled ferry services, value of 24 hour service, and other positive and negative factors associated with changes in comfort, e.g. possible discomfort in subsea tunnels as a negative value. These case studies do not allow separation of these factors.

3. Some results from *ex post* analyses

It is quite obvious from the figure that if the inconvenience costs of new or improved infrastructure is significant, this may have a substantial impact on the outcome of a CBA. In order to estimate the (implicit) inconvenience costs several case studies are carried out to study the course of these costs. The strait crossings which were the aim of this study are situated along the Norwegian coastline. These 5 infrastructure investment projects sum up to about 3 billion NOK, or 570 million USD. The cases studied are

- (1) The Ålesund connection, linking the city of Ålesund to suburban areas.
- (2) The Molde connection, linking the city of Molde to surrounding areas and improving access from the cities of Molde and Kristiansund to the central eastern Norway.
- (3) The Kristiansund connection, linking the cities of Kristiansund and Molde together.
- (4) The Askøy connection, connecting the city of Bergen to the suburban Askøy island.
- (5) The Helgeland bridge, linking Sandnessjøen to surrounding areas.

The results from this work give strong evidence that assumptions made concerning the inconvenience costs will play an important role concerning the output from CBA. In all 5 projects there are significant deviations in WTP *ex post* compared with the *ex ante* calculations. The econometric model (see Bråthen et al. 1995 for full presentation) showed a stable and invariant traffic demand equation for all connections. The model uses monthly data for the last 10 years (both during ferry operations and after the opening of the fixed link). Several price changes have taken place in this period. The model estimates the demand function in the before/after situation. The results obtained using dummy variables to model the structural change due to the opening of the fixed link show that there is a significant shift in the demand function, and not just a movement along a given initial demand curve due to lower generalized costs. Table 1 shows the price elasticities on generalized costs, and economic profitability with and without the estimated inconvenience costs (IC) depicted in Figure 1.

Project ¹⁾	Social surplus excluding IC (mill USD) <i>Ex ante</i>	Social surplus including IC (mill USD) <i>Ex post</i>	IC per vehicle (USD)	Rise in benefit with IC (%)	Price elasticity (t _{95%})
Ålesund (1987)	0	-56 ²⁾	-1.3	-25.5	-0.88 (-2.89)
Molde (1991)	22	28	1.0	+10.9	-0.90 (-4.09)
Kristiansund (1992)	0	54	5.2	+39.2	-0.79 (-4.86)
Bergen (1993)	67	168	3.4	+60.5	-0.75 (-2.86)
Helgeland (1991)	-9	2	3.0	+42.3	-0.81 (-3.52)

Table 1 Inconvenience costs and surplus in selected strait crossings

¹⁾ Opening year in parenthesis ²⁾ The inconvenience cost is negative

The inconvenience costs per vehicle varies between - 1.3 USD to + 5.2 USD and changes value of time estimates up to 60%. The shifted demand curve then represents an extra benefit pr. vehicle calculated *ex post* which is clearly positive, with the Ålesund

connection as an exception. This benefit increase, calculated in % of total benefits ranges from +10% to +60% for all projects except the Ålesund connection, which suffers a decrease of -25%. This fixed link represents a special case. The connection connects some "suburban islands" to a city, and the extra driving distance per car is about 5 kilometers, compared with the ferry services. Since this is a subsea tunnel with a significant vertical gradient, the driving conditions are time-consuming. There are negative inconvenience costs estimated to -1.3 USD implying that the consumers have a lower willingness to pay than traditional value-of-time estimates. The number of cars has increased, but the number of persons using the connection has decreased compared with the ferry services. One reason may be that there were frequent ferry services, connecting the suburban islands directly to the city centre, serving as a small-scale suburban public transport system less expensive to use than the new toll road.

The other connections show a significant increase in socio-economic surplus when taking the inconvenience costs into account. The Helgeland bridge turns from being unprofitable *ex ante* to be marginally profitable with the inconvenience costs included in the *ex post* analysis. Thus, adding changes in benefits of such magnitude to the ordinary benefit calculations based on the standard value-of-time rates may then have influence in the processes of planning and decision-making.

Most of the connections have a price elasticity (measured on generalized costs) in the interval of -0.8 to -0.9 for light traffic. These are mainstream results for this kind of infrastructure. The model has shown no significant results for goods transport. The exception is the Bergen connection, with an elasticity as high as -1.23 for HGVs. This result may be explained by a positive shift in the industrial activity in the region (a new plant for offshore industry established simultaneously), combined with a competitive goods transport market (sea). The elastic traffic market for HGVs means that if the revenues from the toll stations are below budgets, then there may be problems in using an increase in user charges for commercial vehicles as a measure for increasing revenues. This will also be the case for light traffic with price elasticities as high as -0.8 - -0.9, the income effect may cause the elasticity to exceed unity in case of a considerable price rise. The Ålesund connection is the only one where the company has run into economic difficulties, and where price increases were insufficient because of high elasticities. The toll revenues are now able to cover only half the investment costs.

4. Private or public funding?

Financing through user charges means costs connected to:

- Allocation loss through rejected traffic and prices far beyond marginal costs. User charges of some magnitude may cause significant traffic rejection since price elasticities are rather high.
- Collection costs. These costs are about 17% of the revenues as a national average in Norway, mostly varying between approx. 10% and 25%.

When deciding upon private or public funding, the costs of private funding should be compared with the efficiency loss from levying marginal taxes. In Norway, the cost of public funds are estimated between 1.2 and 1.8 (Vennemo 1991). Cost of funds of 1.25 is used as an approximative value here. As a base for the decision on private or public funding, one should regard the following critical aspects:

- The changes in generalized costs, and the level of the demand elasticities provide information to calculate deterred traffic under different user charges. Given a certain traffic system with a defined traffic market where different segments' WTP are known: How can a charging system be designed to minimize traffic rejection?
- The level of collection costs, and the cost implication of choosing certain collection systems. There is a trade-off between the costs of sophisticated collection systems and their possibilities of minimising allocation loss in the traffic market. For instance, electronic devices for collection may provide nearly perfect price discrimination causing minimisation of deterred traffic. However, the economies of scale are heavily present in such systems, and low traffic volumes may cause prohibitive collection costs.
- Can private funding make an early start, i.e. can the project be implemented now, instead of waiting to get a share of scarce public funds in the future, e.g. 10 years from now? May future growth in traffic create capacity costs (new and larger ferries) that can be avoided with an early start, or can a modest traffic growth and low WTP make it profitable to make a postponement? Theoretically, the optimal

time for implementation is when the traffic and benefits are exactly high enough to give a profit rate the first year equal to the official public discount rate.

Project	Private funding year 1 (present situation)	Public funding year 1	Public funding year 10
Ålesund	-56 (92%)	-107	-55
Molde	28 (100%)	31	21
Kristiansund	54 (60%)	56	34
Bergen	168 (92%)	176	122
Helgeland	2 (50%)	8	9

Table 2 Socio-economic surplus with different financing alternatives (actual private funding share in parenthesis), inconvenience costs included.

The decision problem in choosing between different kinds of funding is often connected to the question of early building with private capital vs. postponement and public funding. Since it is lack of public funding that ration profitable new infrastructure, the trade-off will be between «cost of funds» on one hand and the sum of allocation loss through traffic deterrence and collection costs for private funding on the other. Using results from the demand curve estimations and the *ex post* CBA, a simple iterative model is implemented to compare traffic rejection costs and collecting costs from private funding with public funding for a project opened in year 1 and a public funded project postponed for 10 years. The model calculates the socio-economic surplus with different funding, given the WTP including inconvenience costs, the change in generalized costs for different charge regimes, the traffic volumes for existing/generated traffic, the traffic demand elasticity, the collection costs and the cost of levying marginal taxes to provide public funding. Table 2 presents some of the results with the actual share of private funding chosen in each project. There are 3 alternatives: Private funding in year 1, public funding in year 1 and year 10 (postponed).

The *ex post* project analyses with estimations of the inconvenience costs show that public funding is marginally better than private funding for the four profitable projects, assuming cost of funds equal to 25%. The exception is the Ålesund connection, where the postponed public financing gives the highest socio-economic surplus. However, the Ålesund connection is unprofitable for all alternatives. All the others are profitable with both private and public financing. The figures also show that slightly higher "cost of funds" for public financing will make the private financed alternatives the most profitable solution. Given a profitable project, private financing with the economic loss as a sum of collection costs and traffic deterrence costs is not dramatically different from public financing, a reasonable "cost of funds" of public financing taken into consideration. The uncertainty here is mainly connected to traffic reactions when reducing user charges to zero (100% public funding), and the cost of public funding.

5. Regional impacts and *ex post* analyses

One important topic in the *ex post* analyses is the regional impacts of the new infrastructure. This issue is often of great concern to the politicians during the decision process and is ordinarily not in focus in the CBA. There is a need to present as a part of the CBA both the pure distribution effects and the induced regional economic growth due to the infrastructure project. In general, regional impacts are mid- and long-term effects of the infrastructure investments, even though some shifts in economic activity based on expectations may take place even before the new infrastructure is constructed. The eldest project among those analysed above has been in operation for 10 years, and for most of them the time has not yet come to perform thorough analyses. There are problems with separability over time - there are with few exceptions other events taking place in the same region which make it difficult to separate the regional effects stemming from the infrastructure improvement. A careful research design has to be implemented to capture the regional economic impacts. As one approach, Berechman (1994) advocates an intra-regional equilibrium model framework to explore how changes in transportation accessibility caused by infrastructure expansion affect decisions of firms under condition of externalities. These externalities are *transport congestion* (contraproductive) and *pure agglomeration economies*.

Reductions in generalized costs will probably increase the local firms' equity base through retained earnings. According to the imperfect equity theory (e.g. Greenwald, Salinger and

Stiglitz 1990) this may induce local business investments and thus increase the economic activity in the region. Most of this increase in economic activity in the region should result in increased traffic demand, taking place in a mid and long term perspective.

A reduction in the generalized costs for the industry through major infrastructure investments may have different results:

- A Travelling activity increases. It will be easier to maintain and establish industrial networks of customers and suppliers, it may e.g. be easier to penetrate markets (Mackie and Simon 1986). These effects will emerge rather momentarily (through generated traffic).
- B The travelling activity does not increase momentarily, but cheaper travels may allow resources to be reallocated to e.g. investments in human capital or machinery. There are reasons to believe that in the longer run this will imply increased economic activity with more travelling activity as a result. The result may also be increased returns on investments, with increased wages and thereby increased value of time and/or further increase in investments. The increase in investments may take place outside the region.
- C If the initial transport infrastructure (e.g. a ferry connection) was convenient, the reduction in generalized costs may be marginal. The impact on business travels may be small both in the short and long run.

Experience from other countries shows that in most cases new infrastructure does not give new industrial growth in the short run as a main impact. The analyses carried out on strait crossings are of special interest because these strait crossings may imply a significant shift in the infrastructure supply which may give a substantial benefit to industry. The main assumptions to be fulfilled to expect extra generative effects are (Forslund and Karlsson 1991):

- a) If the initial situation is low standard infrastructure and a vital missing link will be eliminated, then this might give extra industrial growth (the falling marginal utility effect).

- b) If the region has a well developed industrial sector with growth potential then there is a better chance for new infrastructure to be the limiting factor to initiate the growth potential.
- c) Part of the industry should in some respect have transport infrastructure as critical to further development, and it should be possible to identify this critical issue.

The infrastructural effects on firm behaviour will, in turn, affect the travel behavior and possibly the value-of-time will change over time. There are four main groups of effects which may take place:

1. Extending supply of infrastructure may change the firms' investment behaviour in the area covered by the *ex ante* CBA calculations, thus affecting the local travel behaviour and/or the value-of time.
2. Local external effects may occur, because other firms in the area are affected through agglomeration effects. These effects can be observed through in-creased production/productivity, thus affecting traffic flows and/or value-of time. We will return to this below.
3. Global external effects may be present through "spin-offs", where increased profits and activity caused by extended infrastructure creates investments and thus possible agglomeration effects outside the local area. "Spillover effects" to other parts of the economy may also occur. These effects will not be reflected in the CBA, neither *ex ante* nor *ex post*.
4. Pure distributional effects are a common result of radical changes in transport infrastructure. These effects are connected to relocation of firms and households without any economic impact. These effects may be considered with substantial political interest.

One effect of improved links may be initialization of increased interaction between firms in the area, thus creating medium and long term growth processes. The "industrial growth" argument is very often used politically to promote fixed links. A pilot case study was carried out to find out whether firms in selected sectors were able to elicit such kind of possible effects from infrastructure that can influence the economic activity in the area.

90 firms in the area surrounding the Askøy and Ålesund connections were interviewed, using a structured questionnaire and phone interviews. The time span is not very long here, these fixed links have been operating for 5 and 10 years, respectively. The theoretical framework for the study is growth theory from the new economic geography, see Krugman (1991, 1995), Krugman and Venables (1993) among others. The theoretical framework is described very briefly below.

The new economic geography considers imperfect competition as the vehicle for economic growth. Krugman (1995) emphasized the following factors for concentration (“centripetal forces”) and regional economic growth:

- Natural facilities, such as harbours, inland waterways, attractive position.
- External effects through markets
 - *Forward linkages* towards the markets for manufactured goods.
 - *Backward linkages* towards the market for intermediate goods.
 - *A robust labour market* which allows “labour pooling”.

The forces for spreading of activities (“centrifugal forces”) are:

- Factors affecting prices, such as high land use costs, commuter costs and transport costs in urban areas, lower costs outside urban areas
- Negative external effects, such as external congestion costs, pollution in general.

Our pilot study focus on the centripetal forces, firms clustering together. Two kinds of clustering may exist. The one kind of accumulation may be of firms in extensive competition, producing e.g. close substitutes. Entry of firms thus intensify the competitive environment, worsening the situation for the existing firms. The other kind of clustering may be the one who leads to mutually positive processes between firms, thus creating a cumulative growth process. A fundamental assumption for this to happen is imperfect competition with increasing returns to scale (IRS) in production. The latter kind of processes is the focus of concern here.

The central issue is the forward and backward linkages in the production processes. If new firms enter the area (or existing firms expand), this may lead to a cost reduction for the existing firms as well. Then there are a positive backward linkage to the manufacturing costs for all firms in the cluster. If entry increases demand for products produced by the existing firms, there are a positive forward linkage to the market for manufactured goods.

These linkages occur as a consequence of externalities. There are both pure externalities and pecuniary externalities working through the market. Pure externalities in this context may come from mobility of the workforce, where returns on investments in human capital is received by other firms by circulation of employees. *Backward linkages* may come into existence if improved transport infrastructure expands the labour market within reasonable commuting distance. This may influence the growth rate through pure externalities as an increase in the return on human capital and labour pooling, the latter reducing the probability of workforce scarcity in the area. Pecuniar externalities working through markets depend on imperfect competition with IRS. If improved transport infrastructure increases the probability for new entries, this may augment the demand for intermediate supplies. Given IRS, then the increased demand for supplies will lower the costs of supplies for all manufacturing firms, thus increasing productivity and inducing growth. This is a positive backward linkage from the entry of firms to the production factor market. *Forward linkages* may occur if entry influences the market for finished goods for all firms in the cluster. When a firm enters a region, the demand for labour increases. Labour migration to the area may then lead to increased demand for finished goods for the existing manufacturing firms as well. With IRS, the growing demand may lead to a price decrease. This pecuniary external effect is working through the market for manufactured goods.

The transport costs are an important matter for generating possible agglomeration effects from imperfect competition. If transport costs are high, concentration in existing urban areas will most probably be the result. The tolerance for negative, centripetal effects such as congestion and pollution (diseconomies of urban scale) is higher if the transport costs in less densely populated areas are also high. Thus, a weaker transport network in rural areas may reinforce urban concentration. In Krugman's 3 region model (Krugman 1993) it is shown that if there are lower transport costs between region 1 and either 2 or 3 than between 2 and 3 then region 1 will serve as a "hub" who attracts activity to exploit increasing returns.

The results from the pilot study is informal, and the results presented here is indicative for future work. The main purpose was to study how improved transport infrastructure influenced contacts with markets for supplies, finished goods and labour. The value of travel time savings is ranked as the most important factor, reported by 37% of the respondents. This effect is normally included in the cost-benefit analyses for the transport infrastructure, and indicates perceived improved productivity through more effective transport. 30% reports that the ability to have better/more frequent contacts with the final goods market is an important matter for the firm's profitability. A more comprehensive labour market is rated as less important (an average score of 2 on a scale from 1 to 7). We also asked what kind of concrete actions which firms had made as a direct consequence of the fixed link. 10 per cent of the respondents reports increased investments, 19 per cent have increased their number of workers, 20 per cent had increased their activity, 30 per cent reports better profitability. As many as approx. one third of the respondents claimed that they had been able to build better networks towards customers and suppliers.

The informal results give *indications* on the kinds of pecuniary external effects known from theory, meaning that there may be an emerging cumulative growth process in these areas. One point worth noting is that the Ålesund link has been operating 5 years longer than the Askøy connection, but this study is not able to trace any noticeable differences between them concerning the potential for economic development.. The results shown in section 3 with negative economic profitability measured *ex ante* for the Ålesund link may be taken as a support to this, but there are differences in the industrial structure which makes a direct comparison methodologically inadequate. A thorough study looking into industrial structure, firm entries/exits and changes in demographic structures are forthcoming.

6. Conclusions

The outcome of the infrastructure investments is encumbered with considerable uncertainty. *Ex post* analyses have shown significant deviations from the *ex ante* appraisals, and have proven to be valuable instruments for informing appraisal practice.

Thorough knowledge of the traffic market structure in the region is very important in reducing this uncertainty. Such knowledge may be obtained by studies of projects already implemented through before/after analyses sketched above. There are reasons to

emphasize the need for further research on this matter. The *ex post* CBA approach carried out in several projects may be extended to cross-sectional analyses, as pointed out by Berecman (op.cit).

Thorough analyses of the regional industrial structure with respect to studies of cumulative growth caused by large infrastructure projects is recommended in places where such effects are traceable. These analyses may result in possibilities of generalization which might allow more reliable "ex ante" analyses of economic impacts from developing transport infrastructure.

References

Aschauer D. A. (1989) 'Is public expenditure productive?', *Journ. of Monetary Economics* 23, pp.177-200.

Aschauer D. A. (1993) 'Public capital and economic growth', *Public Policy Brief*, 4, pp.9-30.

Berecman Y. (1994) 'Urban and regional economic impacts of transportation investment: A critical assessment and proposed methodolog', *Transportation Research A*, vol. 28 A, 4, pp. 351-362.

Bråthen S., Hervik A., Nettet E., Bræin L. and Sunde Ø. (1995) *Experiences with toll financing* (in Norwegian). Report 9507 Møre Research/Molde College, Norway.

Forslund U. and Karlsson C. (1991) *Regional impacts of infrastructure - an empirical review* (in Swedish). ERU, Industridept Ds 1991:55, Sweden.

Greenwald B. C., Salinger M. A. and Stiglitz J. E. (1990) 'Imperfect capital markets and productivity growth' *Unpublished paper*, NBER.

Hervik A., Bråthen S. and Nettet E. (1994) *Industrial benefits from transportation investments* (in Norwegian). Working paper M9401 Møre Research/Molde College, Norway.

Hulten C. and Schwab R. (1993): 'Infrastructure spending: Where do we go from here?' *National Tax Journal Vol XLVI* 3.

Jansson J. O. (1991) 'Can a macro perspective add something to allocation policy versus the transport sector?' (in Swedish) In P-O Hesselborn (ed): *Infrastruktur och samhällsekonomi*. TBF-Report 1992:21, Sweden.

Jones P and Hervik A (1991) 'Restraining car traffic in European cities: An emerging role for road pricing' *Transportation Research no 26A* 1992.

Krugman P R (1991): *Geography and Trade*. Cambridge, Massachusetts: The MIT Press.

Krugman P R (1993): *The hub effect: or, threeness in interregional trade*. In Ethier, Helpman and Neary (eds.): *Theory, Policy and Dynamics in International Trade*.

Krugman P R and A Venables (1993): *Integration, specialization and adjustment*. NBER Working Paper No. 4559, 1993.

Mackie P J and Simon D (1986) 'Do road projects benefit industry?' *Journ. of Transport Economics and Policy* 9.

Vennemo H (1991) *An applied general equilibrium assessment of the marginal cost of public funds in Norway*. Statistics Norway, Oslo 1991.

Appendix 1: The econometric model

The estimation of inconvenience costs is made by means of econometric analyses on time series data from 1984-1995. This appendix shows the demand functions for the 5 fixed links that are presented in Paper Two, and the error correction model that is used to establish the new market equilibrium after the improved infrastructure is put into operation. The demand functions and the error correction model are described in Bråthen et al. (1995) and in Hervik and Nettet (1996). The results provide the elasticity measured on generalised travel costs (GC_t), the income elasticity measured from changes in wealth (W_t) and the shift in the traffic demand function (SHIFT) when the improved infrastructure is introduced. Here, wealth is used as a proxy for income, which may make the interpretation of the income elasticity a bit problematic. However, for these purposes the interest is directed towards the shift effect and the price elasticity measured on generalised costs. For each link, the inconvenience cost (IC) is also presented, measured in \$ (USD) as in table 1 in Paper Two. X_t is the demand in period t, i.e. the traffic volume measured as the number of cars per month during the time period.

Demand functions (long-term equilibria) for 5 fixed links

Ålesund

$$X_t = -0.88 GC_t - 0.27 SHIFT$$

(-2.89) (-6.20)

$$R^2 = 0.88 \quad AR(12,82) = 0.99$$
$$IC = -1.3 \text{ USD}$$

Molde

$$X_t = -0.90 GC_t + 0.18 SHIFT$$

(-4.09) (2.00)

$$R^2 = 0.87 \quad AR(12,102) = 1.91$$
$$IC = 1.0 \text{ USD}$$

Kristiansund

$$X_t = -0.79 GC_t + 0.11 W_t + 0.21 SHIFT$$

(-4.86) (2.41) (10.15)

$$R^2 = 0.90 \quad AR(12,90) = 1.10$$
$$IC = 5.2 \text{ USD}$$

Bergen

$$X_t = -0.75 GC_t + 0.49 W_t + 0.47 \text{ SHIFT}$$

(-2.86) (6.50) (8.33)

$$R^2 = 0.93 \quad \text{AR}(12,103) = 0.87$$
$$\text{IC} = 3.4 \text{ USD}$$

Helgeland

$$X_t = -0.81 GC_t + 0.16 \text{ SHIFT}$$

(-3.52) (4.25)

$$R^2 = 0.92 \quad \text{AR}(12,90) = 0.73$$
$$\text{IC} = 3.0 \text{ USD}$$

The model for adjustment to the long-term equilibrium

The equations above denote the long-term solutions in these transport markets, which are the established equilibrium adaptations to the new infrastructure. When modelling transport demand on monthly data, it is important to control the dynamic structure that leads to these equilibria. Therefore, the econometric models are specified as error correction models, where the short- and long-term effects are separated. The long-term effects are the relevant ones for interpreting the adaptations according to economic theory. The short-term effects in error correction models may be difficult to identify and to explain from economic theory. Error correction models with stationary variables are used to identify the partial effects of each variable. Regressions with non-stationary variables may give spurious relationships (Granger and Newbold 1974). To eliminate the non-stationarity the variables may be differentiated before they enter the model. The explanatory variables have to be stationary, either by themselves or in a combination with other non-stationary variables. The latter is called *cointegration*.

For Ålesund, Molde, Bergen and Helgeland, the variables X_t , W_t and GC_t are differentiated over 12 months (annual change) to capture stochastic seasonality, and to make the data series stationary. For Kristiansund, the same variables are differentiated monthly. In addition to stochastic seasonality, the models show sign of deterministic seasonality. Thus, seasonal dummy variables are included in the empirical models. For Ålesund, the monthly air traffic (*airp*) is also included, because the fixed link serves a commercial airport.

The general error correction model is formulated as:

$$\Delta_m X_t = C + T + \gamma_i \sum_{i=1}^n \Delta_m X_{t-i} + \mu_j \sum_{j=0}^n \Delta_m GC_{t-j} + \lambda_k \sum_{k=0}^n \Delta_m W_{t-k} + \alpha X_{t-m} \\ + \beta_1 GC_{t-m} + \beta_2 W_{t-m} + \beta_3 SHIFT + \beta_i Si + \beta_h Dh + u_t$$

where	Δ_m	= change over m months ($\Delta_{12}X_t = X_t - X_{t-12}$)
	X	= traffic volume
	C	= constant
	T	= deterministic trend
	GC	= generalised travel costs
	W	= wealth
	$SHIFT$	= shift in the traffic demand function
	Si	= deterministic seasonal dummy month i (e.g. $i=8$ means a seasonal fluctuation in august each year)
	Dh	= special event h (e.g. D_{925} is representing a special event in May 1992)
	u_t	= stochastic error term
	$\gamma, \mu, \lambda, \alpha, \beta$	= parameters to be estimated by OLS

The long-term solutions represent the solving of the dynamic model when the adjustments have taken place. This is also the most interesting solution to discuss against economic theory. The numbers in brackets under the long-term coefficients are the t-values. R^2 is the multiple correlation coefficient while AR (n,k) is an F-distributed Lagrange multiplier test for autocorrelated residuals of an order up to n , based on an estimation of the dynamic model.

For each of the fixed links, the inconvenience cost (IC) was calculated in Norwegian kroner (NOK), transformed to USD as per ultimo 1996. In brief, the calculation of IC is carried out as follows: First, the change in the generalised cost logarithm is identified from the long-term solution (using the SHIFT coefficient). The GC logarithm immediately before the shift is known from the data. Second, the GC logarithm after the shift is calculated by adding up these two logarithms. Third, the GC logarithm before and after the shift is transformed to monetary terms. The difference in monetary terms before and after the shift (i.e. opening of the new infrastructure) denotes the IC in NOK (or USD₁₉₉₆ as presented here).

The final error correction models are presented in Tables 1-5 below. The numbers in brackets below the coefficients are the standard deviations. The diagnostics in the lower part of the tables comprise R^2 , σ (estimated standard deviation), AUTO F(k,1) (the Lagrange multiplier test (F test) for autocorrelation of order 1..k), NORM $\chi^2(2)$ (chi-square distributed test for normal residuals) and ARCH F(ij) (F-test for i-th order heteroscedasticity).

	Lag								
	0	1	2	3	4	5	6	8	12
c	4.15 (0.93)								
$\Delta_{12}x$		0.04 (0.07)	-0.02 (0.07)	0.11 (0.07)	0.08 (0.06)	0.15 (0.06)	0.09 (0.06)		
$\Delta_{12}GC$	-0.55 (0.17)	0.39 (0.19)	-0.69 (0.20)	0.16 (0.21)	0.11 (0.21)	-0.39 (0.21)	0.31 (0.15)		
$\Delta_{12}airp$	0.38 (0.07)	-0.11 (0.06)	0.03 (0.05)	-0.03 (0.05)	-0.11 (0.05)	-0.08 (0.05)	-0.03 (0.05)		
x									-0.42 (0.05)
GC									-0.37 (0.13)
airp									0.20 (0.05)
S6	-0.008 (0.01)								
S7	-0.03 (0.01)								
S8	-0.02 (0.01)								
SHIFT	-0.11 (0.02)								
D8711	-0.27 (0.04)								
D892	-0.12 (0.05)								

DIAGNOSTICS:

$R^2 = 0.87$ $\sigma = 0.04$ $AUTO F(12,82) = 0.99$ $NORM \chi^2(2) = 0.03$
 $ARCH F(12,70) = 0.74$

Table 1 Demand relations for Ålesund. Dependent variable: Annual growth in car traffic ($\Delta_{12}x_t$). Time period: 1984.7 – 1994.10.

	Lag												
	0	1	2	3	4	5	6					12	
c	3.83 (0.79)												
$\Delta_{12}x$		0.32 (0.08)	0.08 (0.05)										
$\Delta_{12}GC$	-1.35 (0.18)	0.57 (0.22)											
x													-0.30 (0.06)
GC													-0.27 (0.08)
S6	0.06 (0.02)												
S7	0.24 (0.06)												
S8	0.18 (0.05)												
SHIFT	0.05 (0.03)												
D915	-0.23 (0.06)												
D925	0.27 (0.07)												

DIAGNOSTICS:

$R^2 = 0.87$ $\sigma = 0.06$ $AUTO F(12,102) = 1.91$ $NORM \chi^2(2) = 3.88$
 $ARCH F(12,90) = 1.09$

Table 2 Demand relations for Molde. Dependent variable: Annual growth in car traffic ($\Delta_{12}x_t$). Time period: 1984.6 – 1994.9.

	Lag									
	0	1	2	3	4	5	6	8	12	
c	12.16 (0.99)									
Δx		0.38 (0.07)	0.25 (0.07)	0.17 (0.07)	0.18 (0.06)	0.15 (0.07)				
ΔGC		0.59 (0.18)	0.84 (0.12)	0.74 (0.14)						
x										-0.95 (0.07)
GC										-0.75 (0.17)
w										0.10 (0.04)
S6	0.07 (0.02)									
S7	0.19 (0.03)									
S8	0.16 (0.04)									
SHIFT	0.20 (0.02)									
D927	0.20 (0.05)									
D929	-0.35 (0.05)									
T	0.003 (0.0002)									

DIAGNOSTICS:

$R^2 = 0.90$ $\sigma = 0.04$ $AUTO F(12,90) = 1.09$ $NORM \chi^2(2) = 0.30$
 $ARCH F(12,78) = 1.12$

Table 3 Demand relations for Kristiansund. Dependent variable: Annual growth in car traffic (Δx_t). Time period: 1983.7 – 1994.9.

	Lag											
	0	1	2	3	4	5	6	8	12			
c	4.19 (0.92)											
$\Delta_{12}x$		0.19 (0.07)	0.11 (0.06)									
$\Delta_{12}GC$	-0.62 (0.18)											
x												-0.53 (0.05)
GC												-0.40 (0.14)
w												0.26 (0.04)
S6	0.04 (0.01)											
S7	-0.05 (0.01)											
S8	0.02 (0.01)											
SHIFT	0.25 (0.03)											
T	0.001 (0.0004)											

DIAGNOSTICS:

$R^2 = 0.93$ $\sigma = 0.04$ $AUTO F(12,103) = 0.87$ $NORM \chi^2(2) = 10.93$
 $ARCH F(12,91) = 0.36$

Table 4 Demand relations for Bergen. Dependent variable: Annual growth in car traffic ($\Delta_{12}x_t$). Time period: 1984.3 – 1994.9.

	Lag									
	0	1	2	3	4	5	6	8	12	
c	14.42 (1.32)									
$\Delta_{12}x$		0.06 (0.03)								
$\Delta_{12}GC$	-1.73 (0.40)		1.02 (0.58)	-0.99 (0.51)						
$\Delta_{12}w$				-0.77 (0.30)	0.78 (0.30)					
x										-1.09 (0.05)
GC										-0.88 (0.25)
SHIFT	0.17 (0.04)									
D872	-1.34 (0.08)									
T	0.003 (0.0006)									

DIAGNOSTICS:

$R^2 = 0.92$ $\sigma = 0.07$ $AUTO F(12,90) = 0.73$ $NORM \chi^2(2) = 11.96$
 $ARCH F(12,78) = 0.23$

Table 5 Demand relations for Helgeland. Dependent variable: Annual growth in car traffic ($\Delta_{12}x_t$). Time period: 1984.7 – 1994.9.

References

Bråthen, S., Hervik, A., Nettet, E., Bræin, L., and Sunde, Ø. (1995) *Erfaringer med bompengeprojekter*. Report 9507 (appendix), Möreforsking Molde, Molde, Norway

Hervik, A. and Nettet, E. (1996) *Willingness to pay for improved transport infrastructure. An econometric study of five Norwegian projects*. Working paper M9612, Möreforsking Molde, Molde, Norway.

Granger, C.W.J. and Newbold, P. (1974) Spurious regressions in economics. *Journal of Econometrics* 2 (2) 111-120.

Appendix 2: Inconvenience costs

In Paper Two, the inconvenience cost (IC) is the willingness to pay (WTP) for full travelling flexibility. The IC is identified as a shift in the traveller's demand function, done by using econometric time series analysis. In the analysis, the value of travel time (VOT) is a part of the generalised travel costs. The value of travel time savings after opening of the fixed link, is calculated as usual (VOT • actual time savings), using the same VOT as in economic impact assessments in the Norwegian road sector. The same VOT is used through the entire time series, both before and after the opening of the fixed link. The econometric analysis reveals that the traffic demand function for the fixed link appears to contain a component that manifests itself as a shift in the traveller's demand curve, compared with the previous ferry operation. The interpretation is that the value of travel time savings appears to be a too myopic representation of the 'true' benefit of a fixed link. The intuitive explanation is that the IC represents the value of 'unscheduled transport flexibility'. This flexibility seems to be the main difference in transport quality compared to the previous ferry operations. What follows is a brief examination of the IC concept.

For a representative individual, the utility function *without* IC can be expressed as:

$$U_i = U_i(X_1, X_2, \dots, X_n) \quad (1)$$

where U_i = utility for individual i
 X_1 = fjord crossings from A to B
 $X_2 \dots X_n$ = all other goods and services

Hence, the demand function for fjord crossings can be expressed as:

$$X_1 = X_1(P_1, P_2, \dots, P_m, R) \quad (2)$$

where X_1 = transport demand
 P_1 = the price of using a typical fjord crossing, including time costs, tolls and vehicle operating costs
 $P_2 \dots P_m$ = the prices of all other goods and services that influence transport demand
 R = income

If the goods X_2, \dots, X_n are regarded as a 'composite good' X_c (without price change), then equation (1) can be reduced to:

$$U_i = U(X_1, X_c) \quad (3)$$

Figure 1 shows the indifference curves in this two-good case. The budget lines with the given budget $R = P_1 X_1 + P_c X_c$ are inserted, where the price P_1 for fjord crossings is allowed to vary.

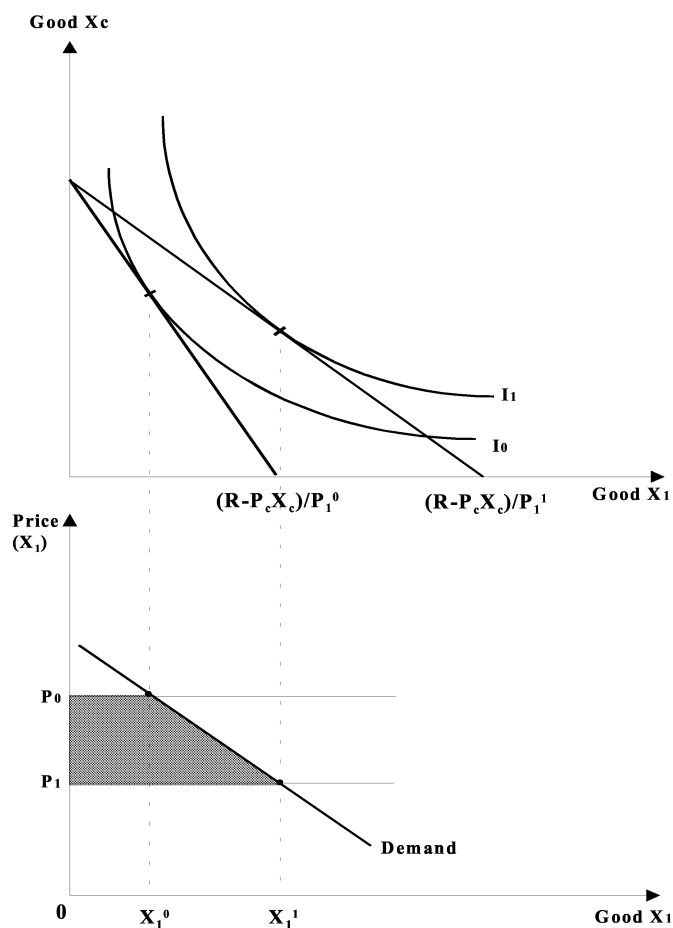


Figure 1 The benefits of reduced transport costs (good X_1)

For simplicity, the linear Marshallian demand is the one that is illustrated here. For a general discussion on Marshallian and Hicksian demand, see the Prologue to Paper One.

The price of X_1 (i.e. the transport cost) is reduced from P^0 to P^1 , and hence the budget line rotates outwards. The corresponding quantities of good X_1 are $X_1^A=(R-P_cX_c)/P^0$ and $X_1^B=(R-P_cX_c)/P^1$ if the entire budget is used on X_1 . However, the individual prefers to consume an amount of X_c as well. The indifference curves derived from Eq. (3) illustrate the combinations of goods X_1 and X_c , that leave the individual indifferent. Because some of the income is used on the composite good X_c , the resulting volumes of X_1 turn out to be X_1^0 and X_1^1 , with the price fall from P_0 to P_1 as shown in the lower part of Figure 1. In the case with a fixed link replacing ferry operations, the predicted *ex post* traffic volume corresponds to X_1^1 (which may or may not correspond to the real traffic volume *ex post*). The IC is not included in the analysis at this stage. The shaded area is the increased consumer surplus from the transport cost reduction.

If now the IC is introduced as an extra element in the individual's utility function and as a part of the transport costs, then Equation (3) can be expanded to:

$$U_i = U((X_1 + IC), X_c) \quad (4)$$

The IC component in the utility function affects the marginal rate of substitution between the use of this transport system, and the use of the composite good X_c . Thus, the shape of the indifference curves is affected, as shown in Figure 2.

The indifference curves are now I_0^2 and I_1^2 . If the IC affects the indifference chart in this way, the *ex post* demand for X_1 will be X_1^{IC} instead of X_1^1 , and the increased consumer surplus will be equal to the sum of the hatched and shaded areas. In the fixed links case, if now the IC is included, the traffic *ex post* is X_1^{IC} , which equals the observed traffic volume after opening of the fixed link¹. The traffic volume *ex ante* (with the ferry operations) is still X_1^0 .

¹ This research was initiated by the fact that in some of the fixed links, the *ex post* traffic volumes deviated significantly from the *ex ante* forecasts.

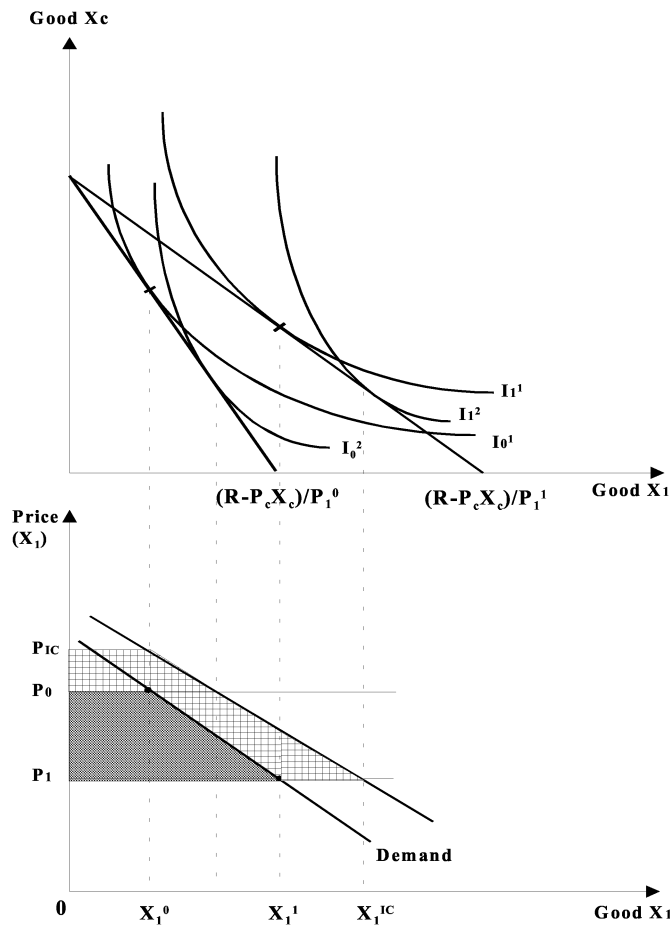


Figure 2 The benefits of reduced transport costs (good X₁), with IC

The econometric *ex post* analysis revealed a shifted demand as a partial effect when the fixed link was opened, which supports the theoretical framework presented here. In Figure 2, the value of the IC is $(P_{1C} - P_0)$. The IC is interpreted as an extra cost component by using ferry transport. The fixed link reduced the transport costs to P_1 . Hence, the total benefits for the individual is the consumer surplus equal to the shaded + hatched area in Figure 2.

PROLOGUE to Paper Three:

**Do fixed links affect local industry?
A Norwegian case study**

6. PROLOGUE to Paper Three: Do fixed links affect local industry? A Norwegian case study

6.1 Introduction

As a backdrop for Paper Three, this prologue briefly discusses the complexity in the interactions between transport infrastructure and economic development. The need for explanations of the economic linkages and impacts of transport policy on private sector economic performance has been acknowledged within the literature. Cost-benefit studies have traditionally not been able to provide *complete* estimates of the payoffs of transport projects because they may fail to account for adjustments during the projects' economic life span. These permanent or transitory adjustments may take place in the logistics system in the private sector either as an evolutionary process in general or as a consequence of the investment. The cost-benefit analysis (CBA) does not typically address issues associated with spillover benefits, or network synergy that accrue (Gillen 1996).

The individual firms' adjustment to improved infrastructure is discussed in Paper Three. The study objects are 4 firms located adjacent to new fixed links that have replaced ferry services. These links are connecting island communities to the mainland. *A priori*, one could expect that the most significant improvement would be that the firms on the connected islands get 24h access to the rest of the transport network, and that this would be beneficial in terms of being able to develop industrial networks and improved logistics. Such improvements could be cost saving connected to e.g. holding of stocks and leaner production processes because of 24h accessibility.

Transport infrastructure is a production factor in the production of goods and services. To recapitulate from Chapter One, the macro production studies indicated that infrastructure capital could be an important factor for enhancing economic growth. However, there did not seem to be any consensus to what extent these studies were able to trace the 'true' returns from infrastructure capital at an aggregated level. Theoretically, the macro studies should be able to capture the productivity effects on output (measured e.g. in GNP), but be unable to capture the consumer surplus effects properly. The macro studies do not provide any understanding of what specific mechanisms that are present between economic agents in a local industrial environment. Insight into how improved transport infrastructure affects local firms may be useful for understanding local growth processes.

Regional and national economic growth depends on complex processes. The decisions of economic agents are inextricably intertwined. These processes and the way transport infrastructure affects them is discussed in Paper Three. It seems apparent that these interdependencies should be taken into account in order to explain the process of development. The traditional, neoclassic view of regional development is, however, that production factors like labour and capital are perfectly mobile between regions. The flow of capital and labour are supposed to flow towards the regions offering the highest returns.

One of the purposes of Paper Three is to see if improved infrastructure may affect the local productivity. According to Oum et al. (1992), a gross of aggregate measures of productivity can generally be decomposed into a number of sources including:

1. Exploitation of economies of scale or network size;
2. Exploitation of economies of traffic density;
3. Better exploitation of indivisible inputs;
4. Elimination of allocatively inefficient input combinations;
5. Elimination of technical ('X') inefficiencies;
6. Increased quality of inputs and outputs;
7. Reductions in negative externalities, such as air pollution, noise, etc.

The focus in Paper Three is mainly on the sources 1 and 6, where the four firms adjacent to new transport infrastructure have been asked about their adjustment processes. The main topics are whether the infrastructure allows a possible better access to suppliers and customers, if their R&D activities with respect to improving the quality of products and processes are affected, and if their use of labour in terms of recruitment and exchange is influenced.

The rest of this prologue is organised as follows: Section 6.2 discusses issues related to transport infrastructure and economic development. Section 6.3 looks briefly into the transport infrastructure's role in the firms' production processes and section 6.4 assess transport infrastructure and regional interaction. Section 6.5 discusses whether cost-benefit analysis may fail to cover the 'true' benefits of transport infrastructure investments that has large impacts on the local economy. Section 6.6 concludes the prologue.

6.2 *Transport infrastructure and economic development*

Economic development depends primarily on locational advantage, whether it is between regions or nations (Eberts 1990). Firms seek areas that offer greater opportunities for economic profits. Transport infrastructure can enhance these opportunities either by increasing productivity or by reducing factor costs. Augmenting the efficiency of production factors and providing attractive environment for firms and households (that may even keep households willing to reside despite lower wages compared to other residential areas) are examples of productivity and factor costs effects. According to Eberts (1990), regions characterised by an environment conducive to further activity are likely to benefit from infrastructure investments. Such an environment can be characterised by abundance of skilled labour and access to cheap power and other supplies. In addition, according to Forslund and Karlsson (1990), if the infrastructure investment removes functional bottlenecks, then the infrastructure investment may enhance growth.

The degree to which an improvement in transport infrastructure affects economic development is dependent on the economic and demographic characteristics of the region where the improvement takes place. From a theoretical standpoint, the analysis of the impact of infrastructure investments should consider the nature of the local economy. There are three fundamental premises for economic impact assessment of transport infrastructure (Banister and Berechman 2000). First, the investment should be effective. This means that the investment should have tangible impacts on the performance of the transport network. Investments that improve the organisation and provision of transport services in an area, is regarded as effective if they generate measurable benefits that are sufficient to outweigh the incremental investment and operating/maintenance costs. Second, the causal linkage between infrastructure investment and economic growth must be manifest in changes of transport-economic behaviour. This means that economic development depends on economic agents like households and firms reacting to the changes in the transport network. In the short run, this means changes in travel variables like trip generation and mode choice, while the localisation decisions of households and firms and the real estate prices are affected in the longer run. Such observable changes from improved accessibility are necessary to foster economic development. Third, transport improvements that affect the above mentioned variables must eventually be transformed into measurable economic benefits. According to Mohring (1976), whether these effects are fully included in a CBA or not depends on whether the prospective real changes in transport costs are taken into consideration.

These benefits comprise improved factor productivity, larger output, increased demand for inputs, increased property values and greater demand for consumer goods. This means that if the transport system can be viewed as a constraint on the economic activity for households and firms, then the possibility of economic development is present if this constraint is relaxed by transport infrastructure improvements¹.

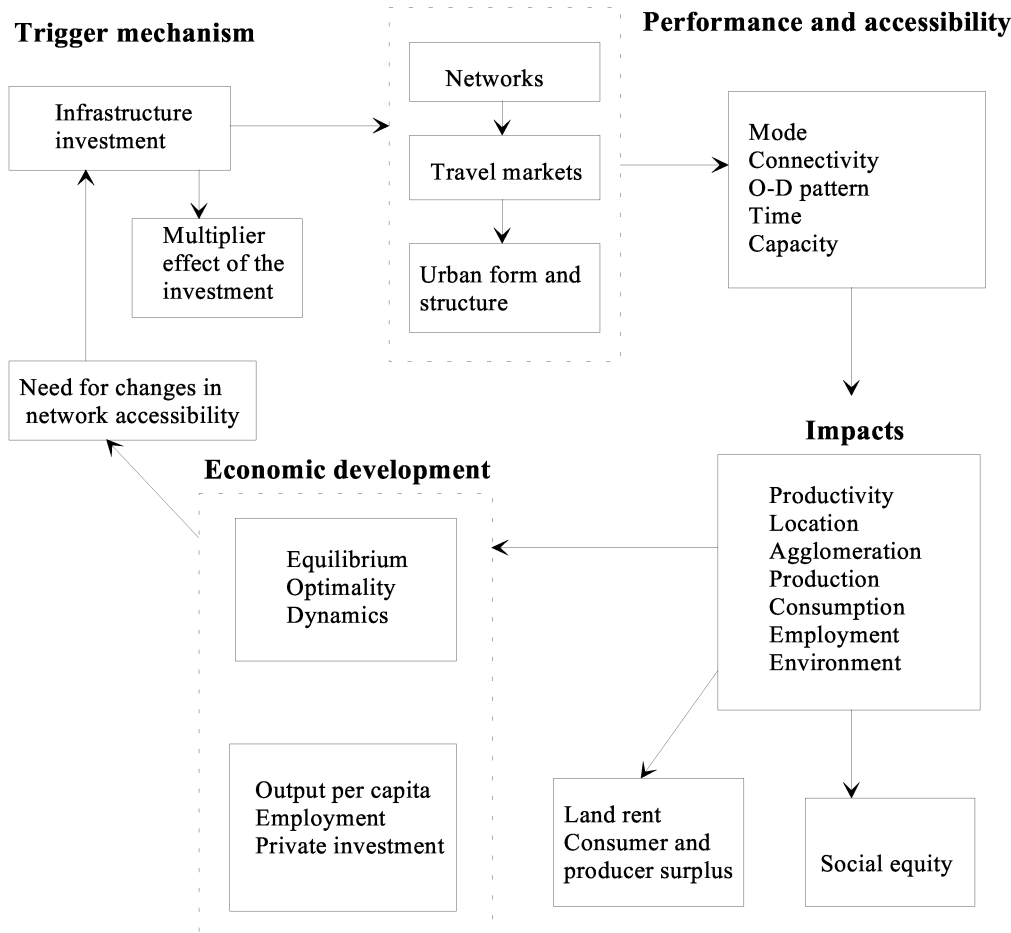


Figure 6.1 Relationships between transport infrastructure investment and economic development (adapted from Banister and Berechman 2000)

The model outlined in Figure 3 of Paper Three provides a framework for enhancing the understanding of the interplay between transport infrastructure improvements and the local economic structure at a rather detailed level. Figure 6.1 presents the causal

¹ Boarnet (1998) focuses on more efficient use of the existing infrastructure as a more successful way of improving efficiency than investments in new infrastructure.

relationships between transport infrastructure and economic development. This model illustrates important underlying variables of the circular and cumulative development illustrated in Figure 1 in the paper. The ‘trigger mechanism’ is the improvement in transport infrastructure. This affects the transport networks’ ‘performance and accessibility’ in various ways, expected to be followed by changes in travel behaviour and localisation of firms and households. The ‘location and real effects’ of these changes are expressed through changes in productivity, location and so forth. The economic effects of the changes in the transport network are expressed mainly through the consumer surplus (see the Prologue to Paper One) and producer surplus. Land rents express expected gains from reduced transport cost, and can be seen as a part of the consumer and producer surplus. Counting both land rent changes and these surpluses may entail double-counting. ‘Economic development’ results from productivity effects, hereby increasing regional output. In turn, this may cause a need for further infrastructure investments. However, these basic connections may or may not be significant, dependent on the local industrial milieu and whether the infrastructure serves as a real bottleneck or not.

Storper (1997) describes elements in a theoretical framework comprising *transaction costs* and *evolutionary economics* for explaining economic development. He regards these as two fundamentally different theoretical concepts. The transaction cost approach assumes allocation of activities (also in spatial terms) based on cost minimisation. Here, a stochastic element is present because allocation is a result of the minimisation of expected transaction costs. This stochastic element may have importance if the transport systems are perceived as having influence on the connection between accessibility and the possibility of managing production plans and commodity flows based on just-in-time (JIT) logistics. Evolutionary economics on the other hand does not have cost minimisation as the point of departure, but is more concerned with factors that may influence and hinder the firms’ ability to innovate and to achieve a cost efficient production.

According to Storper (1997), recent challenges within regional economics are connected to the ability of defining and explaining the underlying forces contributing to territorial specialisation and economic development in a world where physical transportation and electronic communication connected to goods and services are facilitated. These commodity flows are to a growing extent managed according to generic scientific principles. The implication may be that the use of the same principles in applied logistics may make it more difficult to use logistics as a competitive factor, except the aspects that are connected to the nature of differences in distance to markets. The reason why regions and regional differences are relevant as units of analysis may be sought along two axes:

1. Re-specialisation and de-standardisation of production factors and finished goods entail - *ceteris paribus* - that transaction costs increase because the sets of customers, suppliers and production processes are growing.
2. Increased agglomeration may cause organisational and technological learning effects:
 - *The specific one*: Input/output relations between firms may create buyer/supplier relationships that are important for increasing knowledge.
 - *The generic one*: The non-market related mutual relationships that are a part of economic and organisational learning processes and cooperation. An example is 'labour pooling', where the economies of density in the labour force is exploited through the possibility of 'rotating' between different firms in related sectors.

If such network effects are present, then the region can be considered as the unit within which it is desirable to develop the necessary infrastructure to exploit these benefits. Storper (1997) claims that the theories and models that predict strong globalisation and the end of the 'economies of proximity', are most likely able to capture only fractions of the broad picture. Such models are most often based on trade equilibrium assessments, where the traded goods are assumed to have non-spatial properties, i.e. they may be produced where e.g. the labour costs are minimised. In fact, a dual development may take place: Knowledge-based production of goods and services may benefit from regional networks of firms in mutual buyer/supplier relationships, while more standardised, "low-tech" (perhaps labour intensive) production may face increased globalisation. Such tendencies are already observed in the Norwegian shipbuilding industry. Low-tech labour intensive hull production is exposed to outsourcing towards areas with lower labour costs (Poland, Romania), while consultant services and high-tech equipping is delivered by the national industrial marine network.

The formation of regional development theories has shown that this field can be approached in several ways. Over the years, a compound paradigm has emerged, where elements from regional economics and economic geography are prominent. In this paradigm, the question of economic development in regional, national and global perspectives is broken down along several important theoretical and empirical axes, as points of departure for model building. An important element in this multidisciplinary paradigm, is the 'holy trinity' of technology, organisation and territory, and the interplay between these, as illustrated in Figure 6.2.

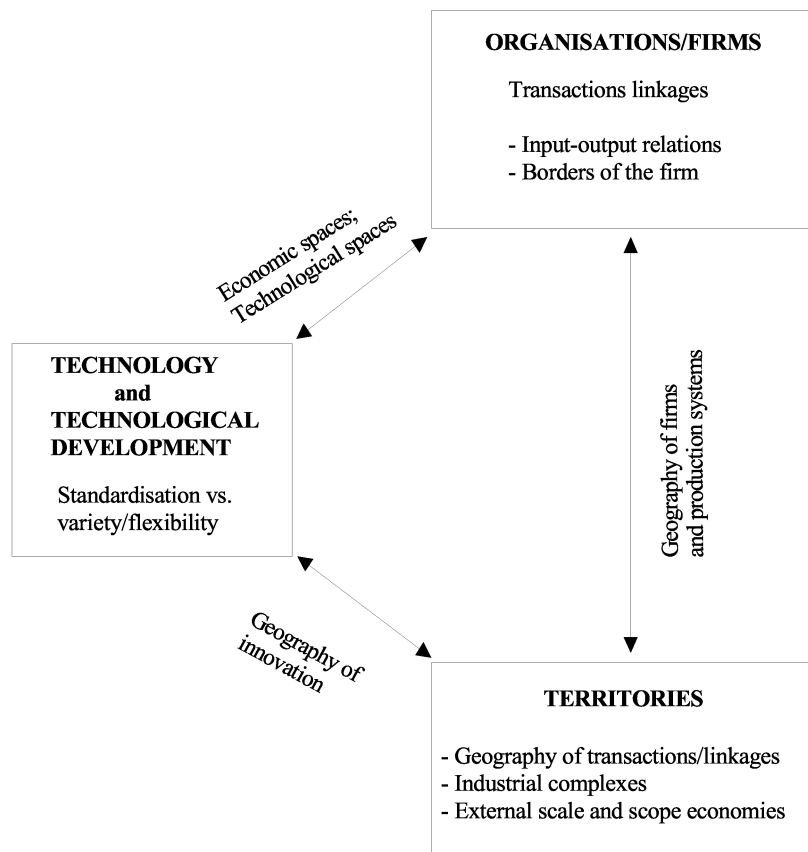


Figure 6.2 The 'holy trinity' of relations for regional development (Storper 1997)

Technology and technological development are essential premises for regional economic development. R&D and innovations occur from local processes, and whether such developments take place or not depends on the region's 'innovation capital', or the technological and organisational culture that might be present. Such processes may alter the conditions for regional development. These processes comprise activities that are initiated within the region, and the ones that take place in competing regions.

Organisations/firms are dependent on adequate accessibility in terms of short transport distances for reliable supplies and deliveries in general. They may also be interconnected in networks that are dependent of certain proximity in the geographic sense. Such networks exist because there are arenas established which serve the need for knowledge exchange and because there may be economies of scale and/or scope in the local buyer/seller relationships.

Territories (be it urban or rural areas) can be characterised by varying intensity in interactions and spillover effects between firms, technologies, and production factors.

The understanding of the potential that may be present for each part of this ‘trinity’ and the interplay between the three parts of the regional economy can be fruitful when assessing how transport infrastructure can influence regional economic growth. The industrial technology in a region can be dominated by either standardised, scale-dependent production technology, or technology that allows flexible specialisation and thus a potential for greater product diversity. The first category *may* make it difficult for new entrants, and the growth potential may be present only (at least in the short run) if the improved transport infrastructure allows the scale-dependent industry to strengthen or maintain its competitive position. The second category represents a more diverse industrial culture. There is also a difference between technology that is specific for certain kinds of production and thus more exposed to market demand, and more generic technology by which a wider array of products can be made. The latter has a lower degree of path dependency in the sense that it allows for market driven adjustments towards leaner production, without being too dependent on historical decisions regarding the choice of production technology. If the infrastructure facilitates networking activities as discussed in Paper Three, then a case for regional growth may be present. The latter category was *a priori* expected to be present in the case studies reported in this paper.

6.3 *The role of transport infrastructure in production processes*

In general, economic growth is defined as an increase in output per capita. Alternatively, economic growth can be defined as a rise in input factor productivity. Another interpretation of economic growth is in terms of increased variety of the goods and services produced and also in terms of an expanded number of firms producing them (Berechman 1994).

In order to establish causal relationships between infrastructure improvement and economic growth, transport infrastructure should serve as an intermediate input to private production processes. Actual transport costs for people and goods can thus be reduced, and beneficial network activities can be facilitated. Well-functioning transport infrastructure is therefore supposed to have a positive impact on economic growth by stimulating the production of final goods and services that use the new transport infrastructure capital as a significant input factor. This is achieved either directly through

reduced transport costs or through various network effects like larger labour markets, exploitation of scale economies and knowledge spillovers.

The transport capital stock should be complementary to the private capital stock in order to foster economic development through further private capital investments or better utilisation of the existing capital stock, i.e. enhancing the marginal productivity of capital². If transport capital is a *substitute* to private production processes, economic productivity may decline if additional transport infrastructure investments are made. Berechman (1994) presents a simple example where private investments in telecommunication facilities may enable employees to have flexible working locations (e.g. at home) and hours, thereby increasing their productivity. If transport infrastructure and telecommunications are substitutes, added investments in transport infrastructure may negate the positive effects of telecommunication on labour productivity.

Berechman (1994) also discusses the three categories of impacts that are generated by infrastructure investments, namely investment effects, accessibility, and external effects. The *investment effects* refer to the economic multiplier effects from the transport investment itself. The temporary effects regarding employment, increased regional income and consumption are the measured outcomes. The *accessibility effects* refer to the increase (or even decrease) in spatial mobility from the infrastructure expansion. Effects on travel time, vehicle operating costs and traffic flows are the measured outcomes. Paper Two assesses the inconvenience cost concept as an additional accessibility benefit when scheduled ferry services are replaced by fixed links. The *external effects* are conventionally related to environmental nuisances from the additional infrastructure and its use. These effects comprise mainly noise nuisance, emissions, and landscape intrusion. In addition, external time costs related to congestion affects accessibility. In Paper Three, external effects working through markets (pecuniary effects) are examined. If improved transportation makes it possible for a set of firms with IRS³ cost structure to improve their use of the production facilities, price decreases may result if the competitive environment is satisfactory. These aspects are also discussed in Section 6.5.

Figure 6.3 summarises the categories of effects from transport infrastructure investments.

² The marginal productivity of human capital is also relevant here.

³ IRS = increasing returns to scale.

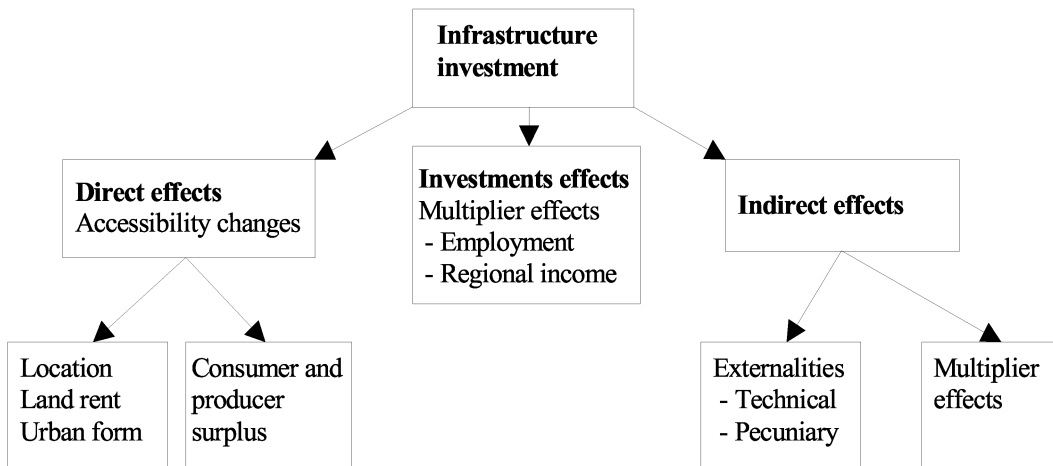


Figure 6.3 Effects of transport infrastructure investments

The prologue to Paper One discusses the way these effects can materialise within the transport investment's influence area. Paper Three assesses in particular the way infrastructure integrates in the firms' buyer/seller relationships and the way R&D activities are affected.

6.4 The transport infrastructure's role in regional interaction

A transport network connects economic agents. The selected firms' network activities are one of the main issues in Paper Three. An economic link in a network is characterised by having an economic value that is not directly priced. Furthermore, it has a certain durability, and it emerges as a result of investments in interpersonal networks, buyer/seller relations and/or infrastructure. Economic relations can be formalised by contracts, or they can be kept informal. A network is not based on temporary, sporadic contacts. An economic network supports the following interactions between firms (Johansson and Karlsson 1994):

1. Company A supplies company B with equipment, technology and other know-how.
2. Companies A and B co-operate on technology and purchasing of goods and services.
3. Companies A and B co-operate on the export markets.
4. Companies A and B have an R&D co-operation.

The strength of the economic network affects to what extent the companies are able to do technical and organisational development and innovations. Regional milieus with a larger number of agents and a broad set of activities may have a higher potential for growth. Well-functioning transport and communications networks are necessary to make such networks viable. The level and distribution of infrastructure services are influencing the localisation decisions regarding establishment of new activities and expanding the existing ones (firms and households). These decisions affect the regional economic development. Such adjustments evolve gradually.

Adjustments to new infrastructure are different. Induced traffic on a new road link may occur immediately as a consequence of improved accessibility, while localisation and relationships between companies are long term network effects, which may have an adjustment period of 10 years or more (Banister and Berechman 2000). Johansson and Karlsson (1994) emphasise the following elements that may influence the way transport infrastructure affects the long-term adjustments:

1. The infrastructure affects the availability of different resources. It affects the factor mobility and thereby the possibility of achieving the best composition of different production factors.
2. The infrastructure affects the costs of different production factors, and the access to and integration with markets for finished goods.
3. The infrastructure affects the possibility of establishing innovative R&D networks and the costs of maintaining the existing ones.
4. In general, and as a synthesis of the first three points above, the infrastructure affects the costs of establishing and operating economic networks.

Paper Three examines in detail how the substitution of ferry services with fixed links has contributed to possible changes in the way firms adapted to the new situation. The firms studied had *a priori* indicated significant network benefits. In theory, improved infrastructure may be a two-edged sword for firms and regions. The following simple hypothetical example illustrates this, from a neoclassical point of view. Let A and B be two regions which produces a homogenous commodity with costs C_A in region A and C_B in region B. The transport costs per kilometre is t . The two region centres is situated with a distance M kilometres. The market for region A has its boundary m_A kilometres towards B, and correspondingly the boundary for region B is m_B . The m 's are determined by the production costs and the transport costs. Thus,

$$(1) \quad m_A + m_B = M$$

The boundary is determined where

$$(2) \quad C_A + tm_A = C_B + tm_B$$

By means of (1) and (2), m_A can be expressed as

$$(3) \quad m_A = 0.5 \left[M + \frac{C_B - C_A}{t} \right]$$

If the production costs in region A is relatively cheaper, then m_A will increase if infrastructure reduces the costs of transport. If A is a depressed area but with lower production costs, then transport infrastructure could assist in expanding its market potential. On the other hand, if A is a more prosperous region than B, then the improved transport infrastructure may worsen the situation for the lagging region.

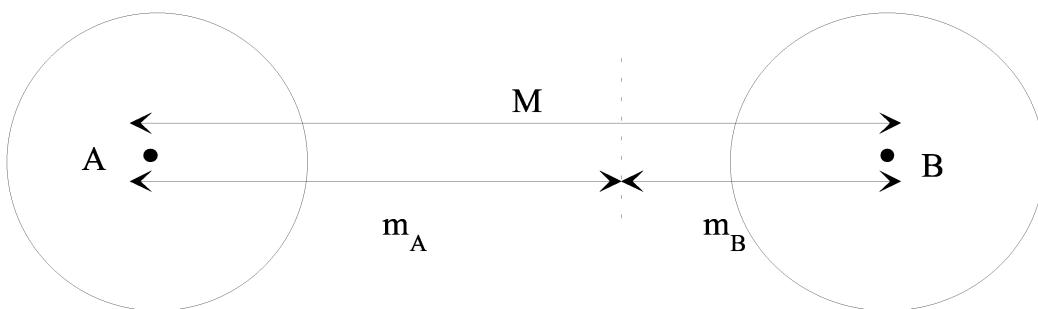


Figure 6.4 Market areas served by regions A and B.

Figure 6.4 shows a situation where $C_A < C_B$, giving $m_A > m_B$. This model is strongly simplified (two regions, one homogenous good, equal transport costs), but it shows how reduced transport costs may affect the access to markets, and that the productivity in the regions' industries may decide whether improved transport infrastructure may enhance prosperity or poverty. However, regions do not normally specialise in single commodities, but produce a range of goods. Thus, the final effect of the improved transport infrastructure depends upon relative production costs and the importance of transport in the costs functions for the various commodities.

The main questions regarding the link between transport infrastructure and regional development are:

1. How do changes in the transport network influence the economic development in the affected area?
2. How does the economic development affect the passenger and freight volumes by mode?

Two main categories of numeric models can be helpful in answering these questions, namely partial models for transport use (for assessing specific projects) and more comprehensive models, e.g. general equilibrium models with spatial adjustments. The transport models comprise different groups of transport users with specific preferences, different transport modes with mode-specific user costs and other properties. The travel behaviour is mainly determined by the generalised travel costs, like travel time costs, waiting time costs and payable costs. The trip generation mainly takes place in the different nodes in the network. The travel activity is surveyed, together with the transport users' reactions to changes in transport costs and income. If changes in the transport costs take place, then the models may provide information about the impacts for the different modes with respect to changes in transport volumes. This information is one of the core inputs to the CBA. Whether or not these simpler models may fail to reflect real world situations is discussed in Section 6.5 below.

Simpler models are used for smaller networks and gravitation models are of this kind. Spatial models are used to study the spatial effects of firms and residents, and the properties of a set of alternative nodes for localisation are described. Multi-layer models can handle the interplay between transport infrastructure investments, the spatial distribution of activities, and the corresponding travel pattern. Multi-sector models can be used to estimate the production and demand in each node, and for calculating the commodity flow between each pair of nodes. These models are often expressed as general equilibrium models (GEM), where the infrastructure's effect on the distribution of activities within the transport network is assessed (see e.g. Bröcker 1998). As a contrast to the CBA factor prices which is conventionally determined as if a competitive economy represents reality, GEMs may capture price changes from imperfect competition, like increasing returns to scale (IRS) in the local firms' cost functions. The common hallmark of these models is that they tend to provide information about the various users' demand for the use of nodes and modes within the transport network. However, one critical comment that can be made against these models is that they may fail to capture the qualitative properties of people's behaviour, such as 'cultural' path dependencies and satisficing as opposed to maximising behaviour. GEMs are used to provide a comprehensive modelling of how an impact in an economic system (e.g. transport infrastructure in a regional economy) may shift the regional economic equilibrium in

terms of changes in prices and quantities. Such modelling can be applied especially in cases where the new infrastructure causes a new organisation of the region's economy. GEMs may provide a more comprehensive analysis of the real effects⁴ than usually provided by a CBA. The latter is based on traffic forecasts and fixed factor prices directly connected to the use of the transport infrastructure (including some of the external effects).⁵

A careful description of the regional transport network and the supply and demand structures is needed to assess the effects of transport investments by use of GEM models. They are also used in studies of international trade aiming at calculating the effects of removing or introducing trade barriers. Experiences from international trade theory are highly relevant when assessing transport barriers, see e.g. Krugman (1991).

A basic assumption when designing GEMs is that the variables (factor prices and corresponding quantities) converge. The awareness of the time dependency in these convergence processes is important. Westin (1995) presents a few examples of fast, medium and slow adjustments, which are shown in Figure 6.5.

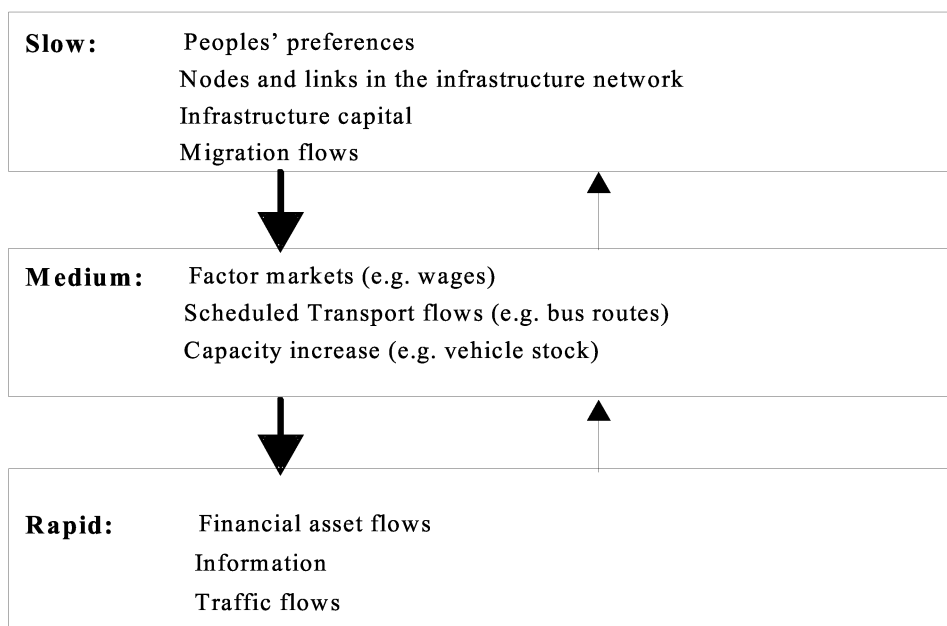


Figure 6.5 Time perspectives for processes affecting transport flows and regional development (adapted from Westin 1995).

⁴ The term 'real effects' denotes the net use of economic resources.

⁵ See the Prologue to Paper One for a more comprehensive discussion of CBA properties.

Changes in the users' preferences, the stock of infrastructure capital and the nodes and links that the infrastructure defines, make up the slow adjustment processes. The structure of the factor markets, the route network and the vehicle stock can be considered as having a medium adjustment process. The daily traffic flow's adjustments to e.g. congestion, the flow of capital within and between regions and the flows of information belong to the rapid adjustment processes. These time perspectives are not meant to have general validity. For example, in deregulated markets adjustments in the route network may take place on short notice. The adaptation to new information networks can be slow in the beginning (due to e.g. lack of available knowledge and hardware), but information technology may pass thresholds where things start to develop rapidly. Information technology and networks may thus be either in the slow or in the rapid category, depending on the region.

A main point in the figure is that the premises for changes are defined 'top-down', shown by the bold arrows. This means that the slow adjustments give premises for the faster processes to take place. Infrastructure capital may e.g. be needed for route adjustments. The thin arrows illustrate possible feedback from the faster processes. E.g. if optimisation of a route network takes place because of infrastructure investments, this may in turn increase the demand and hence the need for more infrastructure.

6.5 *Cost-benefit analyses and economic development*

There are mainly two questions that can be raised with respect to whether or not traditional CBA captures the whole set of benefits from transport investments if they enhance regional cohesion. The first question is whether the industries are sufficiently distorted in terms of imperfect competition. If so, an induced quantity change may have significant economic value different from what can be observed through the factor prices used in a traditional CBA. The second question is whether externalities⁶ and increasing returns to scale are present that may cause cumulative growth. These questions will now be discussed in turn.

If improved transport infrastructure contributes to price reductions in case of imperfect competition, then this is an efficiency gain that is not readily captured through the fixed factor prices usually applied in a CBA. The presence of imperfect competition means that firms set the price above their marginal production costs. This creates an additional

⁶ Negative external effects of a counterproductive nature, like environmental externalities and external time costs from congestion may also be present. These are not discussed here.

source of potential welfare gain from a transport project. The argument can be illustrated as in Figure 6.6. The figure is adapted from Venables and Gasiorek (1999). The interpretations of the figure given here are, however, a bit different from those of the authors.

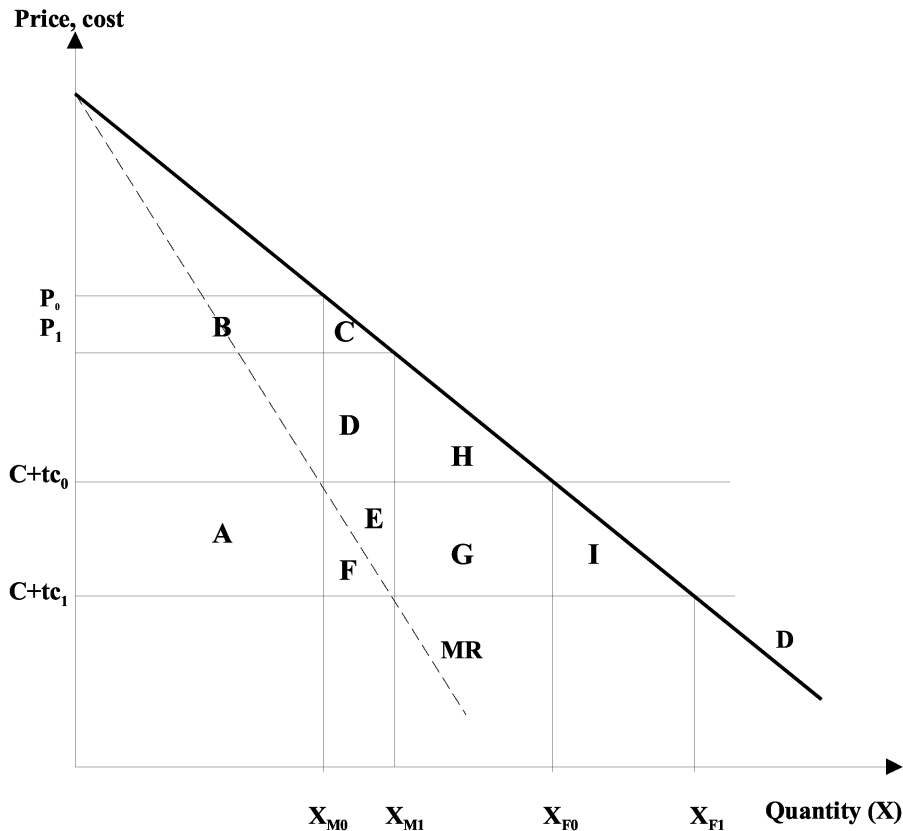


Figure 6.6 Impact of a road improvement under imperfect competition.

The welfare effect of the transport cost reduction can be viewed in different perspectives. The first perspective is as follows: If the price reduction $P_0 - P_1$ on the goods is taken to represent the economic effect of the transport cost reduction, then area $B+C$ is the observed 'welfare effect' from the quantity increase. The second perspective is to consider the observed reductions in (generalised) transport costs only. The reduction from tc_0 to tc_1 gives the area $A+F$ as the welfare effect. This is in line with conventional CBA practice described in the Prologue to paper One, section 2.5. However, since the economic surplus (welfare) is measured as the area between the demand curve D and the new cost curve $C+tc_1$, a third perspective is the most relevant. The welfare effect of the increased output (the 'trade effect') is represented by the area $C+D+E+F$. In addition, area A is the welfare effect from the direct cost saving for the original quantity traded. If the market situation is

characterised by imperfect competition, then the two first perspectives (based on price reductions and transport cost reductions) will fail to represent the economic welfare gain. The areas ACDEF represent the correct measure of change in economic welfare. It is apparent that the information needed for assessing these effects consist of three elements: The change in transport costs and the corresponding change in output, the commodity price level and the price change from transport cost reductions. Figure 6.6 also illustrates that if the transport cost reduction exposes the monopolistic market to competition where the equilibrium commodity price is $C + tc_1$, then the areas GIH should be added to ACDEF as a welfare gain.

The second question is related to the first. Given the presence of technical or pecuniary externalities, it is possible that 'cumulative causation' mechanisms may operate. This means that it may be the case that the effects of an economic change in one activity are not necessary dampened as they spread through the economy, but could possibly be amplified (Venables and Gasiorek 1999). One possible mechanism is that there are technological externalities between firms - for example knowledge spillovers - which create interdependence between the locations of firms. If the new transport infrastructure attracts one firm to a location, then the technological externalities can lead to other firms being attracted because they benefit from these spillovers. A second mechanism is working through pecuniary externalities⁷ of the firm. Such externalities arise when there are increasing returns in the production processes and where increased demand facilitated by the new transport infrastructure makes production less costly. As an example, if demand for a given commodity rises and causes a new entry in this industry, then increased demand for intermediate goods to produce that commodity may increase competition among the supplying firms and hence cause a wider variety of less costly intermediate goods. This in turn causes a pecuniary externality through price reductions that are beneficial for the "downstream" industry and may lead to further expansion of this industry. Thus, the reduced prices of intermediate goods may cause increased demand for finished goods that again raises demand for intermediate goods, and so forth. Such a cumulative process may enhance economic growth. These mechanisms are discussed further in Paper Three.

The two questions - whether or not market imperfections and/or externalities are present - challenge the conventional CBA based on fixed factor prices. There may be a case for a broader assessment of transport infrastructure projects if they are suspected to influence the local economy in a way that raises these questions. Paper Three focus on the possible existence of externalities in firms adjacent to new fixed fjord links.

⁷ The term 'Pecuniary external effects' refers to situations where actions cause effects on third parties, and where these effects emerge as changes in prices of supplies and finished goods.

6.6 *Circular and cumulative causation*

The theory of circular and cumulative causation (CC) is the theoretical superstructure of Paper Three, with Myrdal (1957) as a prominent contributor. Even if Myrdal is rightly connected with CC and contributed significantly to its development, the initial modern CC framework was exposed in Young (1928). The ideas can however be traced back to Marx, who identified the urban poor and their situation in many nineteenth-century English cities (Vickerman 1980). Marx was not concerned with the neo-classical situation where surplus labour was absorbed from surrounding rural areas until the balance of marginal products was achieved. Marx discussed capitalist concentration, the advantages accruing to large units of production and the external economies of agglomeration, attracting increasingly specialised labour. The extent of that specialisation and of growth depended crucially on the quality of communications. However, according to Vickerman, one difference between CC and Marx's analysis is that the labour migration was not necessarily just a response to higher wages, there was also a strong push element associated with the mechanisation of agriculture and unemployment in rural areas.

Another important contributor to the CC is Kaldor (1970). Explicitly, Kaldor stated that CC is

'nothing else but the existing of increasing returns to scale - using that term in the broadest sense - in processing activities. These are not just the economies of large-scale production, commonly considered, but the cumulative advantage accruing from the growth of the industry itself - the development of skill and know-how; the opportunities for easy communications of ideas and experience; the opportunity of ever-increasing differentiation of processes and of specialisation in human activities'.

Myrdal identified a diverse range of dynamic mechanisms that produce circular and cumulative processes. But he did not give priority to any particular mechanism in his methodological work. Different mechanisms should be selected that best explain a particular circumstance. However, manufacturing was a vital force in the growth and development processes, with its ability to generate the increasing returns, external economies, productivity gains and investments that are the sources of per capita growth. Myrdal was also concerned with the importance of non-economic factors (in the neo-classical sense), and suggested that the omission of these factors represented principal shortcomings in economic theory (Myrdal 1957, p. 30). In Myrdal's terms, non-economic factors comprise elements like the organisation of society and constraints imposed by religion and culture. With reference to the model shown in Paper Three, Figure 3, factors like innovation abilities and preferences that diverge from neo-classical cost minimising

behaviour may influence transport demand and hence the benefits of improved transport infrastructure, and may be categorised as non-economic in the Myrdal sense. The call for analyses of political and institutional effects in the study of development in regions and nations is prominently present in Myrdal's work.

At the micro level, the qualitative aspects of local firm's adaptation to improved transport infrastructure may be significant factors in revealing the "true" benefits of such improvements, acting as Myrdal-type non-economic factors. E.g. if there is a culture for cross-regional cooperation, then increasing returns to scale (IRS) connected to the stock of knowledge or 'human capital' can be exploited by means of reduced travel resistance. This may also cause exploitation of IRS connected to the capital equipment. Paper Three approaches the question whether IRS can be traced in a selection of local firms adjacent to two fixed links in Western Norway.

Endogenous growth theory and circular and cumulative causation.

External economies are central to endogenous growth theory (Toner 1999). Like neoclassical growth theory, endogenous growth theory identifies technological innovation as the principal source of increase in per capita output. In the neoclassical world, the technological innovation rate is exogenously given. As a contrast, endogenous growth theory locates the sources of technological change mainly within the economic system. The driving forces behind the rate of output growth in the neoclassical economic growth models are the exogenous labour force increase and the exogenous technical change. Within endogenous growth theory, the primary externality invoked is that of knowledge. The particular form of knowledge that is focused is the kind that gives rise to technological change or improved productivity. Knowledge is generated within the system as product or by-product of economic activity. According to Toner (1999), the sources of knowledge include investments in R&D, 'human capital', physical capital as well as 'learning by doing'. The central claim of endogenous growth theory is that knowledge is at least partly a public good since it is to varying degrees non-excludable and non-rivalrous in consumption. In the former case, use of the good by one firm does not preclude its use by another. In the latter case, patents may to a certain extent exclude other uses, but mostly for a limited period of time. Nevertheless, the mobility of the labour force may allow diffusion of 'derived' knowledge, despite the patents. The non-excludability of knowledge can be claimed from the limited abilities to control the intellectual property rights.

The economic significance of these factors is that there are no (or minimal) costs in the use of the current stock of knowledge to expand output, meaning that there are no opportunity costs of using the current stock of knowledge because there is non-rivalry in consumption (Romer 1996). Conventional neoclassical production functions are linearly homogeneous, so that a doubling of output requires a doubling of input (for an examination of neoclassical production functions and growth models, see e.g. Armstrong and Taylor 1993). Non-rivalrous inputs means that there are scale economies in production since an increase in output can occur without an increase (or with less than proportionate increase) in the use of these inputs. Another point is that the non-excludability of knowledge provides an important externality through the costless transformation of information to other economic agents. 'Technological spillover' allows the social rate of return from technological innovation to exceed the private return.

In the endogenous growth models, knowledge and the cumulative expansion of the knowledge base in the economy sustain both capital accumulation and growth (Grossman and Helpman 1994). This is so because increased knowledge may increase the need for capital investments to convert knowledge to physical products. In addition, the existing capital base may be used more intensively because people find new ways of doing things, e.g. through more rational production processes and organisational innovations. Knowledge may diffuse between economic agents mainly in two ways, 'reversed engineering' and mobility of the workforce, as described in Paper Three.

The endogenous growth theory has focused on positive technological externalities caused by diffusion of knowledge. If such positive externalities are significant for economic growth, this may be a case for public intervention by e.g. R&D funding or infrastructure to facilitate the interaction between economic agents. Figure 6.7 illustrates the traditional concept of positive externalities, where the R&D sector is used as an example. The MPB curve is private demand for R&D services. The social demand (marginal social benefits (MSB) curve) is to the right in the figure, indicating that the society's willingness to pay (WTP) for R&D services (P^*) exceeds the private WTP (P^1). The socially desirable volume Q^* exceeds the volume Q^1 which is demanded in the private market. This is so because the external effects are not included in the investors' production functions. Without some kind of public sector engagement, the volume of R&D services may become suboptimal.

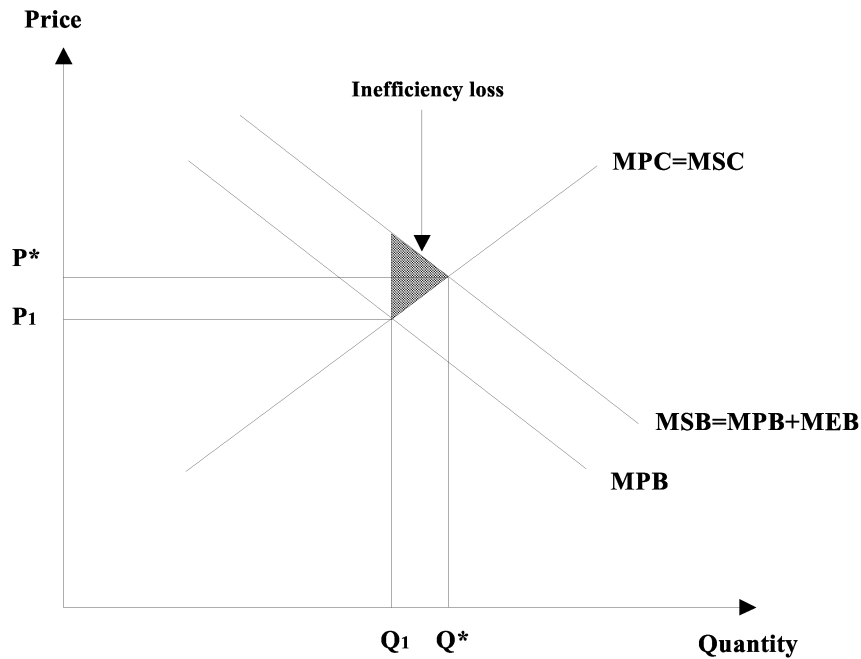


Figure 6.7 Positive technological externalities

In endogenous growth models, imperfect competition is assumed in the supply of new technologies, while the rest of the economy (markets for labour, final goods, capital) is assumed perfectly competitive. The circular and cumulative causation theory has however not given rise to a formal (mathematical) modelling approach. The reason is that the conditions give rise to *pecuniary economies* and the non-convexities of these economies are incompatible with an equilibrium approach. These externalities will be discussed in the next section.

According to Romer (1994), the new growth theory carries over from neoclassical economics the highly aggregated view of the economy. The generation of knowledge and technological external economies is treated at a high level of abstraction where it is either a by-product of production or the outcome of R&D services. In the same paper, Romer stated that endogenous growth theory is complementary to, but different from, the study of R&D or productivity at the level of the industry or the firm. Paper Three addresses the concepts of R&D and technological externalities, but at the level of the firm. The aim is to transform the ideas from endogenous growth theory into an analysis of specific firms to investigate whether improved transport infrastructure facilitates and enhances such diffusion processes between local firms. The theory of circular and cumulative causation

is more concerned with a detailed and specific understanding of the underlying structures and mechanisms that may cause economic development. Elements from endogenous growth theory may augment the understanding of the interplay between transport infrastructure improvements and local industry at the individual firm level.

'New economic geography' and circular and cumulative causation

Another theoretical element that is used in Paper Three, is the 'new economic geography', which is considered as complementary to the endogenous growth theory when establishing a framework for how local firms may adapt to improved transport infrastructure. The common hallmark of these theories is the focus on the role of external effects in economic development.

In the 'new economic geography', externalities may be pecuniary, rather than technological. Such externalities may be transmitted through market transactions (Venables and Gasiorek 1999). They arise, for example, when industry (or firm) A invests in a new plant which cheapens its supplies to industry (or firm) B, which in turn increases its demand for A's output. The circumstances in which pecuniary economies may arise are when the output of industry (or firm) A expands, and when the interactions between industry A and surrounding industries take the following forms (Toner 1999):

1. The output of industry B is complementary to A.
2. The output of industry C is a substitute for inputs into the production of A.
3. The output of industry B is used in the production of industry A.
4. The demand for industry D is increased due to a rise in income following increased output of industry A.

Given the presence of externalities - technological or pecuniary - it is possible that cumulative causation mechanisms may operate. This means that it may be the case that the effects of an economic change in one activity are not necessarily dampened as they spread through the economy, but could possibly be amplified. The process in which pecuniary externalities can cause cumulative causation to take place, is described briefly in section 6.5.

If firms operate under increasing returns to scale, then reductions in transport costs may facilitate reorganisation of production into larger units supplying wider geographical areas. Multi-plant firms may find it convenient to centralise production to a smaller number of plants. The mark-ups over marginal costs may be reduced, and hence increase

the overall sales because of price reductions. Exploiting the increasing returns to scale by increasing outputs may provide a source of welfare gain, but the spatial distribution of activities may be affected.

In Paper Three, forward and backward linkages are discussed, where reducing the transaction costs between firms may lead to expansion of industries. When transport costs are very high, firms tend to locate wherever there are customers. As transport costs fall, firms become more footloose and better able to exploit the advantage of proximity to suppliers and customers. It is therefore possible that improved transport links may facilitate a phenomenon of industrial agglomeration. Firms in an industry (or complementary industries) cluster together, and benefit from the forward and backward linkages coming from neighbouring upstream and downstream activities (Venables and Gasiorek 1999). Forward and backward linkages are closely coupled with distortions in the economy such as those created by imperfect competition.

One of the purposes of this section has been to advocate the importance of thorough knowledge of firm-specific behaviour to understand how improved transport infrastructure influences firm behaviour. The development effects that may be inherent in scale economies have hitherto been focused. However, scale economies may not only provide a local development potential, but also cause competitive disadvantages of regions. Where scale economies are strong, firms can enjoy a degree of protection from new entrants. Where transport costs are high, firms will be protected from external competition from regions that may have greater scale or agglomeration economies (SACTRA 1999). Reduced transport costs may hence cause external firms to undercut the costs of local firms. The overall impact may be indeterminate, depending on the relative degree of scale economies, the size of the local market, the local labour market, and the extent of transport infrastructure improvements. In Paper Three, the scope is limited to a qualitative identification of the external effects from endogenous growth theory and 'new economic geography'.

6.7 The case study approach in Paper Three, and implications of the findings

Four firms adjacent to two of the five fixed links that were included in Paper Two have been subjected to in-depth interviews in a multiple case study. One aim of the study has been to investigate whether there could be traced any effects of technical or pecuniary externalities from improved transport infrastructure in firms that had reported significant

network benefits from improved transport infrastructure. Another purpose was to get a better understanding of how firms changed their real world behaviour with respect to various networking activities in markets for supplies, finished goods, R&D services and labour.

The case study inquiry copes with the technically distinctive situation where there are reasons to expect many more variables of interest than data points (Yin 1994). For the study of the firms' adjustment to improved transport infrastructure, the anticipated high level of complexity in the adjustment process raised the need for assessment of a high number of variables to approach an understanding of this process. With reference to the realist research strategy, the selected cases are not claimed to be representative for the whole population of firms adjacent to the fixed links, or to firms near fixed links in general. However, the firms that are included in the study were selected from a segment of firms that in a pilot survey did report significant network benefits from the new fixed links⁸. Reported network effects with respect to contacts with suppliers, customers and R&D services should, *a priori*, give reasons to expect traces of pecuniary and technical external effects. From the pilot survey results, the four firms were selected as critical cases where the phenomena in question could be expected.

Implications of the findings

Paper Three attempts to provide a deeper understanding of how local firms adjust to improved infrastructure. The extent to which the cost-benefit analyses in the transport sector are able to capture the 'true' benefits, from the industry in particular, has been subject to both public and professional debates. The identification of the inconvenience cost (IC) in Paper Two raised curiosity with respect to this matter, even though the IC was identified in connection with the opening of the fixed links, and not as a long term effect for the adjacent industry. Paper Three sets out to assess the kinds of underlying structures and mechanisms that affect the local firms' adjustments.

The conclusion from the in-depth interviews is that in these cases, a possible local economic development does not seem to be driven by factors included in the growth theories that are presented in the prologue and in the paper. One reasonable explanation is that the rather small industrial networks may be below the 'critical mass' needed to benefit from e.g. economies of scale. In assessing this kind of infrastructure, the effects on local industrial networks appear to be small. Hence, this study does not lend support to the

⁸ Results from the pilot survey are reported in Paper Two.

view that a broader framework than the CBA is needed to capture the effects of industrial linkages and scale economies. However, the conclusion is drawn from an assessment of firms in small industrial networks, and cannot be generalised to larger networks with extensive interactions between economic agents.

Paper Three indicates that when the experiences from Paper Two are carefully incorporated into the cost-benefit framework, then there is no obvious case for investigating further economic benefits in cases where the industrial networks appear to be small.

6.8 *Concluding remarks*

In this prologue, the infrastructure's complex role in economic development is assessed as a backdrop for Paper Three. This complexity can be assessed in many ways. The formal, neoclassical models of adjustment to a changing economic environment may overlook path dependencies like local identity and industrial culture, as well as lack of information and perceived uncertainty regarding e.g. the economic returns from relocating of economic activities. These are factors that may slow down the mobility of labour and capital. A thorough understanding of what mechanisms that determine how firms adjust to improved infrastructure can be important in a couple of ways. First, it can provide information about whether there are spillover effects through network activities that may not be included in an ordinary CBA, like pecuniary external effects where better transport may make it possible to exploit economies of scale. Second, it may provide information of the economic agents' utility functions, thereby making it possible to provide better modelling of locating behaviour.

References

- Armstrong, H. and Taylor, J. (1993) *Regional economics and policy*. 2nd ed. Harvester Wheatsheaf, New York.
- Banister, D. and Berechman, J. (2000) *Transport investment and economic development*. UCL Press, London.
- Berechman, J. (1994) Urban and regional economic impacts of transportation investment: a critical assessment and proposed methodology. *Transportation Research* **28A** (4) 351-362.
- Boarnet, M. G. (1998) Spillovers and the locational effects of public infrastructure. *Journal of Regional Science* **38** (3) 381-400.
- Bröcker, J. (1998) Operational spatial computable general equilibrium modeling. *Annals of Regional Science* **32** (3) 367-387.
- Button, K. J. (1993) *Transport economics*, 2nd ed. Edward Elgar, Aldershot, UK.
- Eberts, R. W. (1990) Public infrastructure and regional economic development. *Economic Review* **26** (1) 15-27.
- Forslund, U. and Karlsson, C. (1991) *Infrastrukturens regionale effekter. En genomgång av empiriska studier*. ERU report 1991:55, Allmänna Förlaget, Stockholm.
- Gillen, D. W. (1996) Transportation infrastructure and economic development. *Logistics and Transportation Review* **32** (1) 39-62.
- Grossman, G. M. and Helpman E. (1994) Endogenous innovation in the theory of growth. *Journal of Economic Perspectives* **8** (1) 23-44.
- Johansson, B. and Karlsson, C. (1994) Transportation infrastructure for the Mälars region. *Regional Studies* **23** (2) 169-185.
- Kaldor, N. (1970) The case for regional policies. *Scottish Journal of Political Economy* **17** (4) 337-348.

- Kreps, D. (1990) *Microeconomic theory*. Harvester Wheatsheaf, New York.
- Krugman, P.R. (1991) *Geography and trade*. MIT Press, Cambridge, Massachusetts.
- Mohring, H. (1976) *Transportation economics*. Ballinger, Cambridge, Massachusetts.
- Myrdal, G. (1957) *Economic theory and under-developed regions*. Gerald Duckworth, London.
- Oum, T. H., Tretheway, M. W. and Waters II, W. G. (1992): Concepts, methods and purposes of productivity measurement in transportation. *Transportation Research* **26A** (6) 493-505.
- Romer, P. M. (1994) The origins of economic growth. *Journal of Economic Perspectives* **8** (1) 3-22.
- Romer, P. M. (1996) Why, indeed, in America? Theory, history and the origins of modern economic growth. *American Economic Review* **86** (2) 202-206.
- SACTRA (1999) *Transport and the economy*. Standing Advisory Committee on Trunk Road Assessment. Department of the Environment, Transport and the Regions, London.
- Storper, M. (1997) *The regional world. Territorial development in a global economy*. Guilford Press, New York.
- Toner, P. (1999) *Main currents in cumulative causation*. Macmillan, London.
- Venables, A. and Gasiorek, M. (1999) *The welfare implications of transport improvements in the presence of market failure. The incidence of imperfect competition in UK sectors and regions*. Report to the Standing Advisory Committee on Trunk Road Assessment. Department of the Environment, Transport and the Regions. London.
- Vickerman, R. W. (1980) *Spatial economic behaviour: the microeconomic foundations of urban and transport economics*. Macmillan, London.

Westin, L. (1995) *En modell för integrerad analys av regional struktur och transportflöden*. CERUM report no 4. Centre for Regional Science, Umeå University, Umeå, Sweden.

Yin, R. K. (1994) *Case study research. Design and methods*. 2nd ed. Applied Social Research Methods Series, Volume 5. SAGE, Thousand Oaks, California.

Young A. (1928) Increasing returns and economic progress. *Economic Journal* **38** (December) 527-542.

7

Paper Three:

DO FIXED LINKS AFFECT LOCAL INDUSTRY?

A Norwegian case study.

Svein Bråthen, Molde University College, Norway

*Reprinted from Journal of Transport Geography Vol. 9 (2001) No. 1.
© 2001, with permission from Elsevier Science*

Paper Three:

DO FIXED LINKS AFFECT LOCAL INDUSTRY?

A Norwegian case study.

Svein Bråthen, Molde University College, Norway

ABSTRACT: The way transport infrastructure may affect factors that can initiate cumulative economic growth processes at the regional level, is focused in this paper. The theoretical foundation for the study is external effects as the vehicle for cumulative growth. A multiple case study of four firms near two fjord crossings is reported. The findings indicate that the networks of firms benefiting from these fixed links are too small to expect improved infrastructure to serve as a “big push” for cumulative growth. However, the interviewed executives perceive the fixed links as important improvements compared to the previous ferry services.

1. Introduction

In Norway as in other countries, the role of transportation infrastructure in economic development has been focused. From the public debate it is apparent that the questions of regional balance and equity have been important political issues that seems to be generic factors across developed countries. Trans European Networks (TEN) are proposed on the grounds that they will contribute to economic cohesion within the EU. The European Commission (1994) emphasised the need for parallel development of secondary networks and access points to major networks.

Aschauer (1989, 1991) contributed significantly to revive the debate on the link between transportation infrastructure and economic growth, when he estimated substantial returns on infrastructure investments by using econometric time series analyses on macro production functions. Even if the results have been criticised (see for instance Hulten and Schwab (1993), these analyses fit nicely with the “common political sense” that there is “something more to infrastructure” on the benefit side than the impacts that can be quantified as economic returns in a CBA. However, studies that have used modern econometrics (e.g. Harmatuck 1996) indicate that the results from earlier Aschauer type studies on the infrastructure’s effect on GNP are overstated.

One main concern in transport policy design is of course to carry out transportation infrastructure investments that are beneficial to the society. The controversial point is how these benefits can, and should, be measured. From an economist's point of view, the primary desired effect is the reductions in transport costs, especially for the industry in the affected area. This is the point of departure for the impact assessments (IA). As an average, the travel time saving counts for between 50% and 80% of the benefits of a typical road investment (Gårder 1989).

This paper deals with the effects of replacing ferry operations with fixed links, in this case between islands and mainland cities. Building a fixed link means that unscheduled, 24h access is introduced. Bråthen and Hervik (1997) have shown that for five fixed links, the willingness to pay (WTP) for improved infrastructure also carried an element of what can be termed *inconvenience costs (IC)*. IC can be interpreted as the WTP for various accessibility benefits like 24h access, regularity in external logistics, transport quality and so forth. IC are added to the conventional transport costs, and form a part of the generalised travel costs. This cost element was of a considerable magnitude but with

significant variations between the links. The IC added between 20% and 60% to the value of time benefits in all cases except one, where the IC were slightly negative. This link was special, because the ferry operations served as a public transport system from the island community directly towards the city centre. A further discussion on these differences can be found in Bråthen and Hervik (1997). It seemed to be a link between how the fixed link affected travel cost savings, and the level of the estimated IC. A larger number of cases is required to estimate this link statistically and should be subject to further research.

The significant IC in four of these five links, indicated that the ferry connections were functional bottlenecks. This result raised the question of how firms adapt to improved infrastructure in the longer run. The main focus here is to shed light on whether this adjustment carries mechanisms that can entail cumulative economic growth. To improve this understanding, a multiple case study using four firms as units was carried out, focusing on firms situated in the proximity of two of the fixed links. One of the links, the Ålesund connection, was the one with slightly negative IC, indicating a smaller improvement in transport quality than expected *ex ante*. The other fixed link, the Askøy-Bergen connection had a rather high IC. These links were selected to be able to contrast the firms' situation with and without significant IC. The fixed links connect islands of approximately 6000 (Ålesund) and 20 000 (Askøy-Bergen) inhabitants to cities and other municipalities on the mainland.

From theory, it is well known that if there is a local industrial milieu present with a development potential *and* if the present transportation infrastructure constitutes a real bottleneck to development, then regional economic growth can be expected to take place in case of infrastructure improvements (Forslund and Karlsson 1991). The connections between productivity, localisation and transport infrastructure are usually modelled from neoclassical theory, where cost minimising and/or profit maximising behaviour based on perfect information about all relevant outcomes are common assumptions. In real world contexts, agglomeration economies, historic path dependence and bounded rationality complicate the neoclassical approach. Mohring (1993) addresses this problem when he states that very little can be said "... unless a great deal is known about ...(households and)...the production processes of business firms". According to SANCTRA (1999), there are ambiguous effects of transport infrastructure on industrial development. These matters are indeed quite complex, and in this paper I limit the scope to examine if there are indications of external economies for local industry that are caused by improved infrastructure.

The rest of the paper is organised as follows: Section 2 presents briefly the theoretical framework for the study. Section 3 describes *realism* as a research method applied on transport systems. Section 4 describes the setting in which the study is carried out, and presents the case interviews. Section 5 concludes the paper.

2. Transport and regional growth - theoretical approaches

Several theoretical approaches can be used for examining the interplay between transport and regional economic development, with focus on local industry. I will focus on three of them. These are Myrdal's theory of circular and cumulative causation (Myrdal 1957), endogenous growth theory, and 'new economic geography'. In my opinion, important elements in Myrdal's theory are deepened by means of endogenous growth theory and 'new economic geography'.

2.1 The theory of circular and cumulative causation

Myrdal (1957) focused on various factors affecting economic growth, including a discussion on how changes in infrastructure may influence cumulative and circular growth processes.

Keeble (1967) stated, on the basis of Myrdal's work, that once particular regions have by virtue of initial advantages moved ahead of others, new increments of activity and growth will tend to be concentrated in the already-expanding regions rather than in the remaining areas of the country. Fig. 1 shows possible causes and effects when the infrastructure improvement influences the conditions for local industry (e.g. conditions as described in section 3). Improved infrastructure in an area may attract new firms, which in turn expands the demand for labour and capital. In Fig. 1, there are four main loops that describe the circular and cumulative effects. First, attraction of new firms increase the population and hence the local supply of skilled labour. Second, the expansion of local employment and population attracts capital and firms to serve the increased local demand for goods and services. Third, secondary industries may evolve to support new and existent primary industries. Both the second and the third loop may support the development of external economies for existent and new enterprises. The external

economies evolve from the increased number of interacting economic agents. The dotted line from infrastructure to external economies indicates the infrastructure's role in reducing 'the costs of interaction' between agents. I will briefly return to the nature of these interactions in the following sections when the endogenous growth theory and the new economic geography are discussed. The fourth loop illustrates how the increased economic activity augments the public sector revenues, allowing for investments in public infrastructure. These investments again drive forward localisation effects and reduced interaction and transaction costs between agents. One of the main empirical issues that Myrdal do not deal with, is related to the 'critical mass problem', i.e. what level of economic activity that has to be present to make these processes actually take place.

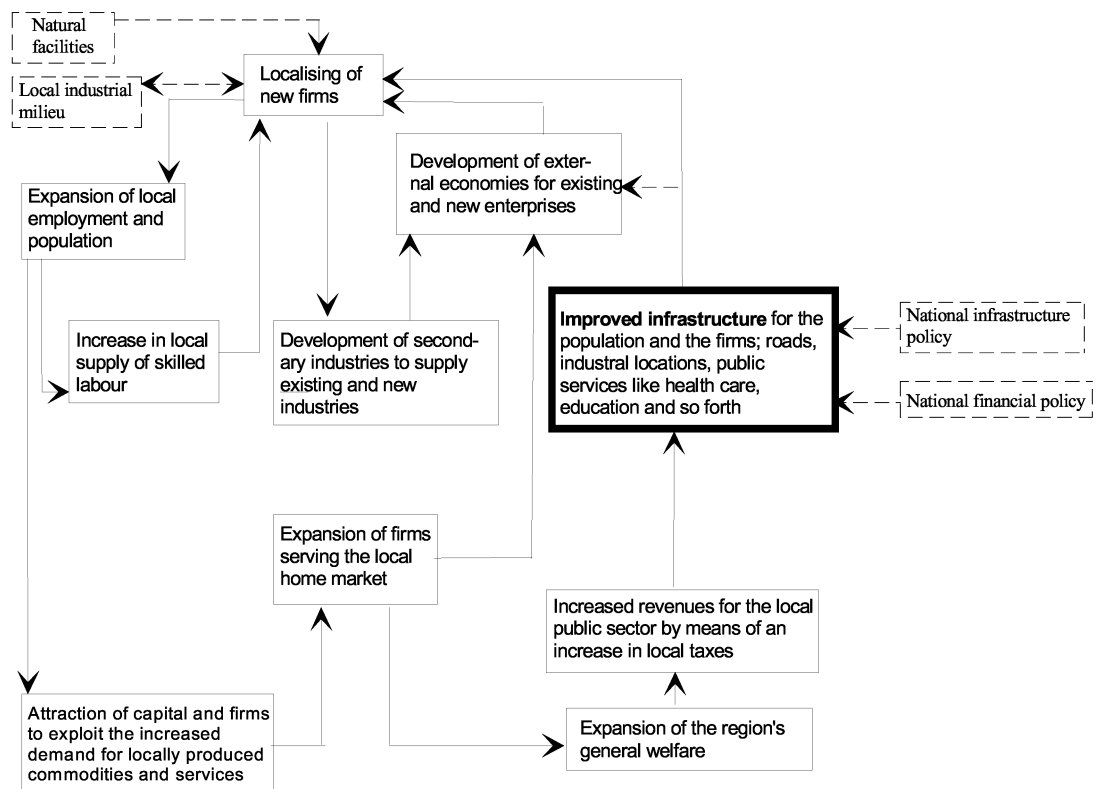


Fig. 1 Infrastructure and cumulative growth.

The variables that can be influenced by government measures are marked with dotted lines in the right hand side of Fig. 1. National financial policy influences both the overall level of public funding of e.g. transport infrastructure, while the national infrastructure policy determines the use of public funding as a mean for achieving regional balance. In

addition, there are other means for supporting regional industry, like R&D funding and differentiated taxes.

The infrastructure impact on transport costs and thereby the economic agents' *localisation decision* is the driving force in the model. The model is positive in the sense that it develops a scheme that illustrates a *plausible* cause/effect chain between initial investments and regional or national industrial development. However, Myrdal does not address *how* the increase in infrastructure quality may affect firms in terms of productivity, innovation and localisation. This problem has later been addressed by the 'endogenous growth theory' (Romer 1986 and others) and 'the new economic geography' (Krugman 1991 and others) through focus on external economies. These theoretical developments can be considered as contributions to the closure of Myrdal's model regarding the transport infrastructure's role in a regional growth process, and I will examine them in the following sections.

2.2 *Endogenous growth theory*

Endogenous growth theory focuses on the increase in *human capital* and technological growth through R&D and 'rich' labour markets. Local industrial milieus with a high number of firms may benefit from the spreading of knowledge to other firms through physical products, R&D products and personnel mobility. The R&D products (knowledge, design) are non-rival in consumption, meaning that they can be used simultaneously by several companies. However, the R&D products may be partly excludable by legal protection. Even if R&D products are protected, the mobility of personnel transfers knowledge between firms. The costs structure of R&D products is increasing returns to scale (IRS), meaning that the initial costs are high, but the marginal costs are low, even close to zero. Design development may serve as an example. Xeroxing a design is mainly free, developing the design is in most cases very expensive. Transfer of knowledge is a typical *technological externality* because the firms can gain access to ideas, processes and products for only a fraction of the development costs.

Monopolistic competition is a competitive environment with incentives for extensive R&D activities. Intuitively, rivalling monopolists have both the economic resource base (through accumulated profits or expected future profits) and they possess a strong interest in R&D to maintain or achieve market positions. It is obvious that if non-rivalries exist,

the neo-classical assumptions regarding perfectly competitive environments have to be relaxed; it must be possible to withdraw from competition by making R&D work wholly or partly excludable. Total exclusion is hardly probable, and two cases may indicate why:

1. '*Reversed engineering*' or 'innovation by imitation' means that other firms gain legal access to innovations by simply buying and examine closely the actual commodities. Knowledge moves between firms independent of physical proximity. The car industry is a well-known example.

2. *Mobility of knowledge* means that the increase in 'human capital' from R&D work creates 'spillover' effects to the rest of the economy because people change jobs. Even if firms protect their innovations, people move and supply other firms with knowledge as a production factor far below the initial development costs. Clusters within the electronics and information technology may have benefited from such agglomerations of knowledge.

Through these activities, there are reasons to expect an increase in the number of non-standardised products. The key workforce is supposed to move between enterprises, allowing human capital accumulation through externalities, i.e. human capital as a production factor far below development costs. According to Myrdal's model reflected in Fig. 1, the development of external economies as a consequence of the increased number of firms and expansion of local employment and population can be analytically assessed by the technological externalities from endogenous growth theory. However, the term 'external economies' from Myrdal may also comprise external effects working directly through the market prices. Such effects will be discussed in the next section.

2.3 *The 'new economic geography': Non-convexities and pecuniary¹ external effects*

Krugman (1995) emphasised external effects through markets as important for regional development. Such *pecuniary externalities* occur if new firms enter the area (or existent firms expand) as a result of reduced transport costs, and these entries subsequently lead to a cost reduction for the existent firms as well. By definition, this kind of *pecuniary externalities* are effects from an agent's behaviour that impose costs or benefits on other

¹ The term 'Pecuniary external effects' refers to situations where actions cause effects on third parties, and where these effects emerge as changes in prices of supplies and finished goods. Latin *pecuniarius* from *pecunia*: money.

parties that can be measured through a change in the *market prices* for goods and services. Pecuniary effects can have both real and distributional characteristics.² Cost reductions from a more effective scale of production, is a real effect and is the issue at stake here.

A fundamental assumption for cost reductions to happen is imperfect competition with IRS in production, resulting in reduced production costs, caused by augmented demand. In addition, improved transport networks may create a 'richer' labour market, which allows 'labour pooling' in the sense that counter-cyclical industries may benefit from exchange of workforce, and hence be able to exploit a potential flexibility in the local labour market. The exchange of workforce links up with the endogenous growth theory, where the spreading of knowledge is supposed to take place through workforce mobility.

Forward and backward linkages in the production processes are two kinds of pecuniary effects that are considered as important when cumulative growth processes are to be explained. If new entries augment the demand for finished goods or services produced by the existent firms, then there are positive *forward linkages* to the market for manufactured goods. If new entries improve the exploitation of the supplying industry's production capacity, the result may be positive *backward linkage* by reducing the cost of supplies for all firms in the cluster. Prices go down because of IRS, increased production and monopolistic competition.

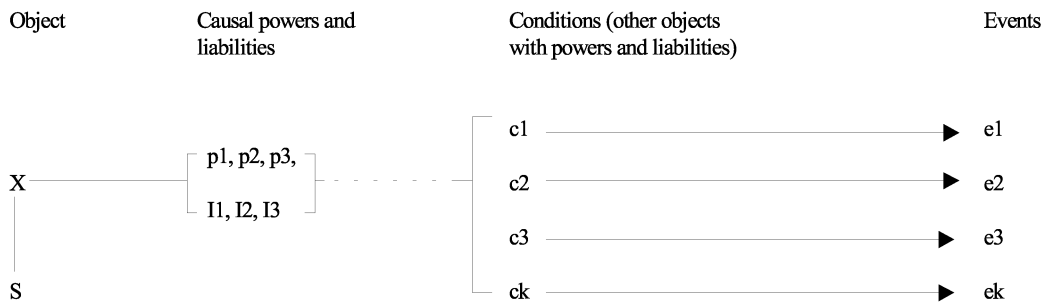
In this section, I have briefly presented three theoretical approaches that may shed light on the connection between transport infrastructure and local industrial development, and the framework presented here will constitute the basis for the case study. The main perspective is to understand how firms adapt to a presumptively significant improvement in transport infrastructure. The feedback from infrastructure to the industrial milieu is the main issue, including the transport infrastructure's effect on the relations with the local labour market. The theory of circular and cumulative causation is the principal framework, with emphasis on external economies. The endogenous growth theory and the new economic geography are considered as complementary to understand the role and the nature of the external economies in Myrdal's model. The next section will describe *realism* as a research method to model the interplay between transport infrastructure and industrial development, with the theoretical backdrop as described in this section.

² Pecuniary externalities are also referred to as distributional transfers between economic agents. Prices rise in one place, with an offsetting reduction elsewhere. See e.g. Nas (1997). This definition is not applied here.

3. Realism as research methodology for transport systems analysis

The main issue here is to consider what benefits improved infrastructure can give to trade and industry in a specific area. The analyses in Bråthen et al. (1997) support that the road user benefits are contingent upon local conditions that can be expected to vary between areas and kinds of infrastructure projects. The focus here is on the relationships between changes in transport costs and how firms in the adjacent area adapt to these changes, in terms of effects on the relationship between suppliers and customers. Labour market effects are also considered because of the increased commuting area for the firms. In the following, a methodology that can be used for approaching these issues is presented.

The ‘realist approach’ seeks to analyse how and why a phenomenon take place, by means of the underlying structures and mechanisms that can establish a contingent relation between events (reductions in transport costs) and causes (e.g. changes in the network of customers and suppliers). The causal link between events and causes can be explained by a set of underlying *structures and mechanisms*. A *structure* can be defined as the composition of the object to be studied. For simplicity, the object can be termed *X* (see Fig. 2). In this case, *X* represents the infrastructure in a given fixed link, e.g a scheduled ferry service or a fixed link with 24h service. The structure of the crossing can be described by the characteristics of the different infrastructure designs that are relevant for understanding how a change in *X* may affect local industry. Such characteristics are connected to e.g. departure frequencies for the *ex ante* ferry services, toll fees, vertical gradients in subsea tunnels (affecting the travel time for especially heavy goods vehicles) et cetera. The *mechanisms* describes how a change in *X* can determine the causal powers and liabilities that is relevant to the analysis. A *causal power* may in this sense be interpreted as one factor that can enhance the ability for firms in the area to extend their markets, or enhance the area’s attractiveness for new entrants. A *liability* may e.g. be increased competition from firms in other regions because of the area’s augmented exposure to competition as a result of reduced transport costs. A change in the infrastructure will mechanically change the travel time directly. The travel time change, expressed by the value of travel time saving, represents one of the causal powers that may trigger off economic development in the area.



Object X, having structure S necessarily possessing causal powers (p) and liabilities (I) under specific conditions (c) will:

- (c1) not be activated, hence producing no change - e1
- (c2) produce change of type e2
- (c3) produce change of type e3, etc.

_____ = necessary condition
 ----- = contingent relation

Fig. 2 The structures of causal explanation.
 (Source: Sayer 1992)

Changes in X can be considered as changes in transport infrastructure with matching causal powers and liabilities, for instance saved travel time, increased competition etc. which may or may not trigger off events (for instance local industrial growth or decline). These events depend on to what extent conditions like for instance an industrial growth potential in certain industries is present. Hence, depending on conditions, the operation of the same mechanism can produce quite different results, or alternatively, different mechanisms may produce the same result. One mechanism that may work as a substitution (or even a complement) to improved transport infrastructure is discretionary grants given to support local industry.

Structures and mechanisms as underlying conditions are perceived as necessary for the phenomenon to take place. But they may not be sufficient i.e. the same structures and mechanisms may or may not lead to the actual events. This means that certain *conditions* have to be fulfilled. When studying the relationship between transport infrastructure improvements and local economic growth, there are reasons to expect that certain charac-

teristics of the local industrial milieu constitute important prerequisites for growth. The following statements illustrates conditions that are assumed to be important for transport infrastructure to entail local economic growth (Forslund and Karlsson 1991):

- (a) If the initial situation is low standard infrastructure and a vital bottleneck or missing link will be eliminated, then this might give extra industrial growth.
- (b) If the region has a well-developed industrial sector with growth potential, then there is a better chance for the lack of new infrastructure to be the limiting factor to initiate the growth potential.
- (c) Part of the industry should in some respect have transport infrastructure as critical to further development, and it should be possible to identify this critical issue.

In the realist approach, an observable phenomenon is considered as a result of a *combination* of underlying structures and mechanisms and what can be termed contextual specific matters. To ask for the cause of something is to ask ‘what makes it happen’, what ‘produces’, ‘generates’, ‘creates’ or ‘determines’ it or, more weakly, what ‘enables’ or ‘leads to’ it (Sayer 1992). In this context, the concern is to identify the conditions that may trigger off local industrial growth as a consequence of improved infrastructure. Structures, mechanisms and endogenous conditions should be termed as a basis for investigations into causal relationships.

In general, whether a causal power or liability is activated or not at any occasion depends on conditions whose presence and configuration are contingent. Sayer (1992) claims that the relationship between causal powers or mechanisms and their effects is not fixed, but contingent. In a transport policy perspective, a mechanism like travel time saving will entail local industrial development when the underlying conditions are certainly those that will produce the desired effects. Whether improved transport infrastructure has significant influence on industrial development may e.g. depend on the factors (a)-(c) above.

This case study is an attempt to look into factors that can cause improved infrastructure to generate cumulative benefits for trade and industry. A qualitative study may often shed light on causal relationships that may bring about events that are actually wanted. Conversely, qualitative studies may also be used for cancelling out causal relationships

that are irrelevant. The following conditions must be met to determine a causal relationship:

1. *Intrinsic conditions*: There must be no change or qualitative variation in the object possessing the causal powers if mechanisms are to operate consistently. As an example, when studying cause and effects of infrastructure, a change in transportation technology shifting the balance between modes may blur the causal relationship between infrastructure provision and demand for transport services. One example is if a shift towards information technology takes place to replace physical transportation regarding e.g. to keep up or establish personal networks within related businesses.
2. *Extrinsic conditions*: The relationship between the causal mechanism and its external conditions which influences its operation and effects, must be constant if its outcome is to be predicted by given relations. If the use of infrastructure is dependent of e.g. the situation for the firms in the area (which are affected by e.g. world market prices and/or changes in production technology), then a study limited to the relations between transport costs and traffic volumes cannot be expected to manifest regularity.

If both these conditions are met, a *closed* system exists in which causal relationships exist and regularities are produced, meaning that we can 'control' the relationships and be precise about the causes and effects. It is quite obvious that these assumptions are most frequently violated, leaving us with open systems of short-lived and unstable cause-effect relationships. Within open systems however, closed or quasi-closed systems may exist on a local or regional level or within a limited time span, where one mechanism dominates or overrides the effects of others.

Transportation is influenced by a lot of factors, and some of them are shown in Fig. 3 below. Hence, most often we deal with open systems when analysing causes and effects. At the most, we can hope for quasi-closed systems where causal relations are restricted in time and space (Sayer 1992) i.e. that the variables in the model represents structures, mechanisms and conditions that can be controlled, at least in the short run. Even if the two conditions for closure of the system is violated, there are some procedures in social science to take partial account of these problems. One is to control for the effects of other causal relationships, as is done for instance in a typical econometric analysis of transport demand.

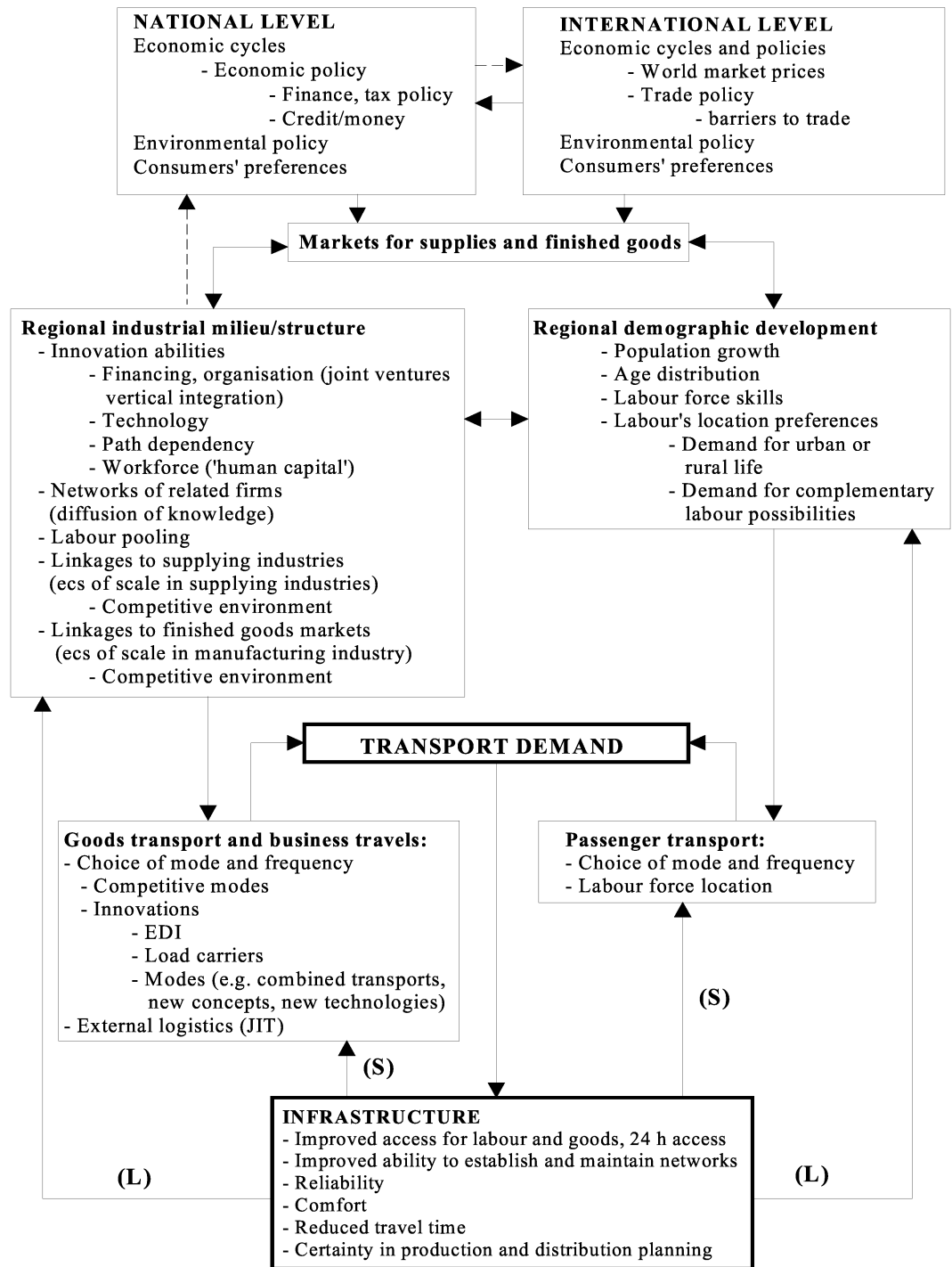


Fig. 3 Demand for transport infrastructure services

Another way is to *assume* that the systems are closed, as we do when economic models are limited to comprise variables (1,...,k), even if a number of n variables constitute the possibility set of influencing variables.

When modelling the demand for transportation infrastructure services, one has to make this kind of assumptions and limitations to construct a quasi-closed system as a basis for demand modelling. In the discussion of models that can give an approximation towards closed or quasi-closed systems, I will refer to “quasi-closed systems/models” as the best real world approximation. Models reflecting such systems are at best stable in the short run. Fig. 3 illustrates a model of k variables determining the demand for transport infrastructure services that attempts to be quasi-closed.

The fundamental driving forces behind the industry’s demand for transport is the international and national markets for various commodities. Besides being influenced by the consumers’ preferences, these markets are affected by economic cycles, trade barriers, economic and environmental policies and so forth. Next, the local industry’s ability to compete constitutes an important prerequisite for the development of local transport demand. *Inter alia*, these competitive abilities are dependent on the local industrial milieu with respect to innovation abilities, the labour market and the network of supplying and manufacturing firms. This chain of conditions spills out in the transport market through demand for transport of commodities and persons, which again may determine how infrastructure improvements affect the transport market. The short run effect of reduced transport costs (marked (S) in Fig. 3) gives a direct effect on transport demand. A model that predicts these direct effects may be quasi-closed in the very short run by including variables like transport costs and the road user’s income. Fig. 3 illustrates that in the longer run, the infrastructure change may have effects on the composition of the industrial structure in the region concerning, for example, the competitive environment and localisation of new firms, and also on the demographic development in the area. These feedback effects are marked (L) in the figure. Especially when there is a large shift in the infrastructure quality, these repercussions may become significant. The dotted arrows in the figure represent *weak* repercussions. In most cases, a change in the local industrial milieu may have significant local effects, but weaker influence on the national economy. Correspondingly, moderate changes in the national economy have minor impacts on the global economy. Thus, we can assume that the demand effects from developments in national and global economy are exogenously given.

The relevance of open and quasi-closed systems regarding transport and regional development is that infrastructure's influence on regional development as an *event* is conditional upon other factors, like growth and composition of national and international market demand for the regional industry. Therefore, the way the regional industries' transport demand fluctuates with the demand for supplies and manufactured goods should be subject to attention. As seen from Fig. 3, the national and international demand depends on a number of relationships considering economic policy, trade barriers, environmental policies and so forth. Intuitively, as a basis for a quasi-closed model, it can be assumed that transport infrastructure on a regional level do not have significant impact on national and international demand, at least not in the short run.

By means of the theoretical framework outlined in Section 2, the case study will explore how transport infrastructure influences the local industrial milieu. The theoretical framework may provide guidelines for structuring models that are quasi-closed. As shown in Fig. 3, this will include innovation abilities, networking and possible expanding of markets for supplies and finished goods. These are complex mechanisms where the infrastructure is just one of several conditions that can eventually enforce local economic growth. To be able to distinguish between these conditions, case studies on the company level can be one way to map how the local industrial milieu can be influenced by improved transport infrastructure.

Another subject to consider is whether transport infrastructure may affect the local labour market. Such effects may be an extension of the labour market, with easier access for a larger workforce than before. This may facilitate the possibility to make 'labour pooling', i.e. to exchange workers when needed. A change from scheduled ferry operation during daytime (18 h) to a fixed link with 24h access may facilitate networking also in the labour market.

The next section will describe the design of the qualitative case study and the selection of cases. The point of departure is the model for a quasi-closed system concerning the interplay between transport infrastructure and the local industry, as illustrated in Fig. 3.

4. New infrastructure's influence on four firms

4.1 Background

As mentioned in Section 1, Bråthen and Hervik (1997) presented *ex post* benefit-cost analyses (CBA) of five fixed links, based on econometric estimates of the traffic demand functions. This CBA method is in accordance with the recommendations from the Norwegian road authorities, and conforms with practice in, for example, the UK. These analyses were compared with the *ex ante* CBA that were done during the project planning process. These fixed links have replaced ferry connections during a time span of ten years. An informal survey among 100 firms was also presented, to study the firms' adaptation to the improved infrastructure regarding supplier/customer networks. The firms were selected among industries that *a priori* were expected to have the greatest benefits from transport infrastructure improvements. The phone survey was carried out among firms at the islands that were connected to the mainland by the Ålesund and Askøy fixed links. The industries comprised competitive industries (fish processing, engineering, furniture and construction) located on the islands.

The survey showed that 10 per cent of the firms reported a perceived significant improvement in especially market related network activities (contacts with suppliers and customers) because of the improvements in infrastructure. These firms had relatively high rankings regarding whether the improved infrastructure had affected the relations with suppliers and customers. From theory, the perceived improvements could indicate technological externalities through spreading of knowledge working through the firms' production functions. It could also be a sign of pecuniary externalities working through markets. This result gave reasons to do a more thorough examination of how firms in the area adapted to the new infrastructure.

The two fixed links have common hallmarks. They connect islands to mainland cities. Both islands are characterised by extensive commuting towards the city centre, thus serving as "suburban islands". The investments are financed by tolls levied on the road users. A single ticket for car costs about NOK 50 (at Askøy a return ticket is paid). Fig. 4 illustrates the position of the fixed links, which are indicated by the bold lines in the enlarged sections.



Fig. 4 The Bergen/Askøy and Ålesund connections
 (Source: Norwegian Mapping Authorities)

4.2 The case study

A multiple case study approach is used in this paper. To have a closer look into if there is possible to trace any indications on cumulative growth in these areas, in-depth interviews with respondents from four firms were carried out. These firms were selected from the

100 firms in the survey, and they were chosen among the 10 per cent segment that reported high network benefits from the new fixed links. The firms had local ownership, with the executives as partners. The executives were interviewed by means of a structured guide. Two of these firms were located at the islands served by the Ålesund connection, and the two other were situated at Askøy, the island which is linked to the city of Bergen by the Askøy bridge.

According to Yin (1994), there are two general strategies for analysing case study evidence. One is use a proposed theoretical framework as an analytical base, while the other is to develop a framework for systematic *description* of the case. The first one is used here, namely to base the design of the case study on the theoretical framework in Section 2. In Section 3, a causal model for the link between transport infrastructure and regional development is presented. The case study follows this model (see Fig. 3) by focusing on whether the networking activities between firms change in a way that may support the theory of circular and cumulative causation, driven by the external economies outlined in Section 2. The conditions, structures and mechanisms are complex, and this case study approach is an attempt to map these elements. An analytic strategy for identifying such causal links is to look for “matching patterns”, i.e. in a systematic way investigate if it is possible to collect evidence to support the theoretical framework. There are reasons to underline that studies of this kind may only provide an *analytic generalisation* to theory, by saying that these cases support the theoretical propositions or not. It is however not possible to make *statistical generalisations* by saying that the observed effects are representative for fixed links in general.

During the theoretical review given in Section 2 and the suggested interplay between transport infrastructure and regional development from Section 3, we stated the reasons for focusing on the following three conditions:

1. The change in the structure of supplying firms. How is the structure of supplying goods after the fixed links with respect to:
 - The number and sites of supplying firms.
 - The price of the supplies.
 - Change in collaboration/’networking’ activities.

- The holding of trading stock. Are trading stock reduced because of increased certainty in deliverance?
2. The change in the structure of customers with respect to
- The number and sites of customers.
 - How has the competitive environment been influenced?
 - Change in collaboration/'networking' activities.
3. Infrastructure's influence on the use of R&D services.
4. The functioning of the labour market after the fixed link is built. How does fixed links influence:
- Labour market size, where do the firm recruit additional or replaced workforce from? What kind of workforce is recruited?
 - Do this workforce provide external knowledge from relevant businesses that adds to the firms' knowledge base?

The Ålesund connection

The link connects the island municipality of Giske with the mainland city of Ålesund. The link was opened in 1987. The infrastructure investment consists of three subsea tunnels and one bridge. The construction costs amounted to £50 million. According to the CBA in Bråthen and Hervik (1997), the link turned out to be unprofitable, with a deficit of about £35 million. Ålesund and Giske has 38 000 and 6300 inhabitants, respectively. Two conditions on demography are worth noting:

- Giske consists of three communities formerly linked together with relatively low-service ferry operations (Giske, Godøy and Valderøy/Vigra). These communities are now linked together with one bridge and a subsea tunnel with free 24h access (no tolls). The tolls are only levied to/from the mainland.

- The airport serving the Ålesund area (approximately 150 000 inhabitants) is located off the mainland, at Vigra. Even if the airport is closed during night-time, there are reasons to believe that unscheduled services to the mainland and the adjacent islands combined with the options on 24h airport services (for instance urgent freight services) may be valuable to the industry. However, there was a frequent ferry service operating previously.

Firm 1: Engineering company.

The firm employs approximately 15 man-years and is organised as a family-owned company. It is located at the island of Valderøy. The products are mainly mobile rock crushing plants. The firm has no departments elsewhere.

Relations with suppliers. Reduction in the stock of supplying goods is to some extent present with respect to gadgets, which now may be collected momentarily by improved access to the city. For larger supplies, the firm has a diversified structure of supplying firms, a number of them is located in Switzerland, Germany and the UK. There is rarely contact with the supplying firms directly, the contact takes place via agents that visit the respondent once or twice a year to ‘discuss actual solutions’. Standard supplying items are mainly used in production, however there has been a lot of work laid down in selecting the most suitable items. The transportation infrastructure has not affected the competitive environment or changed the structure of suppliers.

Customers. The structure or number of customers is reported unchanged after the fixed link was built. The mobile crushing plants are transported in sections that are mounted together on the customer’s site, which most often is well outside the region. The ferries did not represent any problems considering axle loads or size of the loads. The constraints on weight and size are determined by lane width, bridges and subcrossings elsewhere in the road network. Hence, there is no reason to claim that contacts with the market for finished goods have been influenced by the infrastructure in this case.

R&D. The supply of R&D services are based on *ad hoc* relationships where engineering services are hired from firms in the city of Ålesund. It was claimed that it is easier to have access to such services without the scheduled ferry services. With the reported frequency of contacts, (‘occasionally’) it is not easy to see that the ferry services could limit the use of these services.

The firm has no co-operation with other R&D services. The knowledge transfer to other firms may however take place through the engineering consultancy that is hired, and also by 'reversed engineering'. There are no reasons to say that the infrastructure has influenced the intensity of such transfers.

Labour. The labour force has very little turnover. The executives have been in the firm for a number of years. Thus, the exchange of labour among firms in terms of job change is very modest. However, what is reported to take place, is temporary exchange of workers. Particularly in periods with a high activity level but with different seasonal peaks, this kind of exchange may take place. The exchange with e.g. shipyards on the mainland is reported to enhance knowledge with respect to welding processes etc., and is easier to accomplish at present than with the ferry services because of the flexibility and 24 h operations. There are indications of technological externalities, by means of enhanced mobility of the human capital within the local cluster of mechanical engineering.

Conclusion. It seems that workforce mobility facilitated by the improved infrastructure is what can be linked to the theory of cumulative growth by the presence of externalities. However, this effect is hard to quantify.

Firm 2: Producer of prefab kitchens.

The firm employs 43 man-years and is organised as a family-owned company. It is located at the island of Giske, 10-15 minutes drive from the city of Ålesund. The firm has no departments elsewhere. The products are mainly prefab kitchens.

Supplying firms. The firm has a diversified structure of supplying firms, on a national and international level. However, the transportation infrastructure cannot be said to have affected the competitive environment or change in the structure of suppliers. A couple of long haul trucks deliver every week. Some of the deliveries arrive at night, but they are unloaded within ordinary working hours. Thus, the 24h flexibility is of minor importance, it was claimed that 'the trucks could as well have had their waiting time on the mainland, and then taken the first ferry in the morning'. There have been minor changes in the supplier structure after the infrastructure improvements. One exception is that kitchen units are delivered more frequently than before, but this is not claimed to be a result of the fixed link.

Customers. There are not reported any changes in the structure or number of customers after the fixed link was built, and there is no perceived change in the competitive environment. The ferries did not represent any problems considering axle loads or size of the loads, but they represented a more severe barrier than for firms at the islands of Vigra and Valderøy, because of the lower frequency in the ferry services. The respondents emphasised the flexibility that is perceived as a prerequisite for the necessary frequent transports of finished goods.

R&D and product development. This firm has no regular or ‘frequent ad hoc’ co-operation with R&D services. R&D on material technology take place in the supplying industry. Product development takes place within the company and is done by the executive team. However, one important factor in the development process is the input from the distributors. One of the executives has the responsibility for visiting distributors twice a year. This seems to be a typical ‘interactive innovation process’ where innovations take place in interactive systems where market information, technical competence, informal practical knowledge and R&D are important inputs in the innovation model. This is as opposed to the linear innovation model, where R&D and large firms play the major role in the innovation process (Asheim and Isaksen 1996). The transfer of knowledge to other firms may take place through ‘reversed engineering’, but there are no reasons to say that the infrastructure has influenced the intensity of such transfers.

Labour. The labour force is stable, with very little turnover. Nearly 100% of the workforce are living within the municipality. The executives have been in the firm for a number of years. Thus, the exchange of labour among firms in terms of job change is very modest. There is no exchange of workforce with firms in the same or related businesses.

Conclusion. The executives claim that the significant growth that the firm has experienced for the last 10 years would not have been present with the ferry connections still operating. According to the respondents, either the firm would have been smaller, or it would have been relocated to the mainland. For this firm, one economic effect of improved infrastructure may be that the transaction costs of relocation is avoided. It is worth noting that the island of Giske had a significantly lower frequency on the ferry services to the mainland than the Valderøy isle, where firm 1 is located. There may be local forward linkages to the market for finished goods in this case, if the expansion experienced over the last 10 years has lead to a larger local economic activity which may have increased the demand for finished goods in other sectors. However, we do not know whether this may lead to increased product diversity and reduced prices because of

economies of scale. Thus, the result cannot be generalised to the theory of ‘new economic geography’

From the theory of cumulative growth, high transport costs may contribute to an increased centralisation of economic activities. In this case, reduced transport costs in terms of increased flexibility seem to have contributed to a decentralised production.

The Askøy connection.

The bridge connects the municipality of Askøy with the city of Bergen at the mainland, and was opened in 1992. Askøy has 19 200 inhabitants, while Bergen has 225 000. The construction costs amounted to £55 millions. According to the CBA in Bråthen and Hervik (1997), the link turned out to be very profitable with a surplus of around £110 millions. The tolls have also contributed to construction of the main arterial road network on the Bergen side, from the bridge towards the city centre. In addition to the fixed link, a fast passenger ferry service operates between Askøy and Bergen to serve the commuters.

Firm 3: Fish processing

The firm employs approximately 35 man-years and is organised as a family-owned company. It is located at the northern part of Askøy, 45-50 minutes drive from the city of Bergen. The firm has no departments elsewhere. The products are mainly smoked salmon.

Supplying firms. The transportation infrastructure has not affected the competitive environment or change in the structure of suppliers.

Customers. No changes in the structure or number of customers caused by the fixed link are reported. The entire production is for export. The firm claims that their new load carriers (40 foot containers and long haul trucks) could not be used on the ferries. However, the modern ferries with load restrictions similar to the trunk road system could not have been a real constraint here. Contacts with customers abroad are made easier because of improved access to the Bergen airport.

R&D and product development. This firm has no regular or frequent ad hoc co-operation with R&D services. The product development was taking place 'in house' based on signals from the markets.

Labour. The labour force has little turnover. The majority of the workforce is living within the municipality. The executive is the founder of the firm. The exchange of labour among firms in terms of job change is very modest. There is no exchange of workforce with firms in the same or related businesses.

Conclusion. The respondent was very concerned with the long run perspectives, concerning access to markets, the size of the labour market and the general perception of belonging to the urban community on the mainland. The concern about the bridge as a necessity for 'Askøy not becoming a future hinterland' was very tangible. The reduced planning uncertainty concerning ferry capacity and scheduled services was mentioned as a significant 'mental' factor affecting both production costs and accessibility to customers. These elements are not easy, or even possible, to quantify, but indicates that the 'inconvenience costs' related to unscheduled services may be significant. As mentioned, the inconvenience costs are estimated to be significant for the Askøy bridge. It seems that reduced uncertainty and perceived reduced 'mental distance' from the improved infrastructure, having possibly affected the decisions made concerning an upgraded plant, is what we can link to the theory of cumulative growth. However, these kinds of effects have not been possible to quantify in this study. Like for other firms in this study, reduced transport costs are claimed to keep up production outside more central areas.

Firm 4: Concrete industry

The firm employs approximately 18 man-years and is organised as a family-owned company. It is located at the northern part of Askøy, 45-50 minutes drive from the city of Bergen. The firm has no departments elsewhere. The products are mainly various products from concrete, like pipes and pavements.

Supplying firms. All supplies are transported by trucks, except sand, which is shipped. It was reported that 'stockout' of cement occurred from time to time because of delays or under-capacity in the ferry system, causing occasionally increases in the production costs. A silo for receiving shipped cement has not been considered as profitable. Such costs could be considered as the 'upper limit' for the value of eliminated stockout. There are no

indications that the transportation infrastructure has affected the competitive environment, or changes in the structure of suppliers.

Customers. Changes in the structure or number of customers caused by the fixed link have not been reported. The public sector is the main customer. The executive claimed that the tolls for HGVs is of such magnitude that tendered contracts in the Bergen area carrying marginal profits have been lost. However, the total transport costs towards Bergen have been reduced. There are 4 firms competing in this market. If the competitive environment have been affected by reduced transport costs, resulting in a price decrease of the manufactured goods, then pecuniary effects from economies of scale may be present.

R&D and product development. The firm is member of AS Alfanor, an R&D and marketing organisation owned by 13 of the leading concrete product firms in Norway. The companies are scattered throughout the country, and only occasionally they are in direct competition with each other. There has not been possible to identify whether the R&D and marketing co-operation have resulted in local monopolies or price-fixing behaviour. The R&D activity in this network have not been influenced by the fixed link.

Labour. The labour force is stable, and the majority of the workforce is living within the municipality. The exchange of labour among firms in terms of job change is very modest. There is no temporary exchange of workforce with firms in the same or related businesses.

Conclusion. Like for firm 4, the respondent was concerned with the long run perspectives, concerning access to markets, the size of the labour market and the general perception of belonging to the urban community on the mainland. The reduced planning uncertainty concerning ferry capacity and scheduled services was mentioned as a significant factor affecting the accessibility to customers. The deliverance to the mainland is often characterised by a 'matrix' of other contractors being dependent on punctual services to avoid costly waiting time, for example, on a building site. Such costs are usually not internalised. These elements are not easy, or even possible, to quantify or monetise, but indicate that the notion of 'inconvenience costs' related to unscheduled services is relevant. There may be local forward linkages to the market for finished goods in this case. If the infrastructure contributes to increased competition in the markets for finished goods, leading to exploitation of IRS, there may be local forward linkages to the market

for finished goods. However, this study is not able to provide answers to whether this may lead to increased product diversity and reduced prices because of IRS.

5. Summary and conclusions

In this paper, I have by means of an *ex post* case study focused on how fixed fjord links have affected local firms, by carrying out in-depth interviews with four firms adjacent to two crossings. The firms were selected from a segment that had reported significant benefits of a kind that indicated external economies to be present.

Section 2 discusses the theoretical framework for the study, starting with a review of Myrdal's theory of cumulative causation. Subsequently, endogenous growth theory and the 'new economic geography' were considered as complementary to the cumulative causation theory. The understanding of 'external economies' as a growth factor in cumulative causation models can be enhanced by introducing the diffusion of knowledge, 'innovation through imitation' and pecuniary effects from exploiting prospective increasing returns to scale (IRS) in upstream and downstream firms.

Section 3 adapts *realism* as a research method for gaining a comprehensive understanding of transport infrastructure investments' effects on local industry. The 'realist' approach offers a systematic approach to the building of causality models. Such a model is presented as a framework for the case study.

Section 4 presents a multiple case study that was carried out on the basis of the realist model. The study focus on the four firms' relations with R&D milieus, contacts with suppliers ('upstream firms') and customers ('downstream firms') together with the size and function of the local labour market. One experience from these interviews was that even if the quantified effects of improved infrastructure appear to be small when considered from recent growth theory, all the respondents claimed that infrastructure is a very important factor in a long run evolution perspective, for the firms and the local community. One explanation, based on economic rationality, is that the infrastructure serves as a public good that can be demanded at virtually no costs. Because the tolled fixed links replace ferry connections with fares of approximately the same magnitude as the tolls, it is straightforward that continued ferry operations imply higher perceived costs because of ferry fare collection during an 'infinite' time period. After a payback period of 15 to 20 years, the toll collection on the fixed links ceases.

The interviews were carried out by means of the guide given in Appendix 1. The design was based on the theoretical framework given in Section 2. The respondents seemed to be familiar with the issues that were raised in the interviews. They were able to provide clear-cut answers, albeit the network effects were claimed to be modest. There may be a case for trying out a similar case study concept on larger and denser industrial networks that are exposed to significant transport infrastructure changes.

Martin (1999) have discussed the growth theories focusing on IRS as the vehicle for explaining economic change (the 'new economic geography' in particular), and their ability to capture how firms adapt to a changing economic environment. He arrived at the conclusion that the theoretical focus was too narrow, he says (p.75): ... there are [also] several limits to such a narrow approach. For one thing, it means that 'messy' social, cultural and institutional factors involved in spatial development are neglected.[But] it is precisely the social, institutional, cultural and political *embeddedness* of local and regional economics that can play a key role in determining the possibilities for or constraints on development, and thus why spatial agglomeration of economic activity occurs in particular places and not in others".

In this study, the respondents expressed the importance of the firms' deep roots in a stable local community, where the connection to the mainland gave options for local expansion, using and developing the local labour force through a 'learning by doing' evolutionary process. The localisation effects of the fixed links were not measured for the communities at large. For most of the firms in this study, the asserted ability to maintain production without moving to the mainland supports the view that reduced transport costs may be of importance to achieve decentralised production structures.

Essentially, the competitive environment and the number and structure of firms, suppliers and customers turned out to be mainly unchanged, virtually leaving no significant effects concerning changes in pecuniary or pure technological externalities affecting these firms. The small tendencies that were observed concerning labour pooling and changes in the competitive milieu, represented events which took place only occasionally. My conclusion is that in these cases, a possible local economic development does not seem to be driven by factors included in the growth theories described in section 2. One reasonable explanation is that the rather small industrial networks may be below the 'critical mass' to be able to benefit from such scale effects. The firms seemed already to be quite well connected *ex ante* through the relatively frequent ferry services. Essentially, the economic benefits of connecting smaller regional units seem to be captured by a thorough

analysis of the changes in transportation costs. Also, the differences between these links concerning the inconvenience costs could not be traced in these firms' adjustments towards the improved infrastructure.

Experience from EU's 3rd framework programme (EURET/385/94) advocate improved project-specific appraisal methods in the selection of transport infrastructure projects:

- *The development of methods for quantitative assessment of the strategic economic impacts of road and transport projects, distinguishing clearly those benefits which are genuinely additional to the transport user benefits, is an important priority.*
- *The development of ex-post evaluation methods to complement ex-ante appraisal is recommended, both as a check on appraisal accuracy, and to inform future appraisal practice.*

A thorough understanding of the firms' long-term adjustments to changes in the transportation network may provide useful input in development of appraisal practices. Further research is required regarding *ex post* evaluations of transport links.

References

Aschauer, D.A., 1989: Is Public Expenditure Productive? *Journal of Monetary Economics* 23, 177-200.

Aschauer, D.A., 1991: Public Investments and Private Sector Growth: The Economic Benefits of Reducing America's "Third Deficit". Economic Policy Institute, Washington DC.

Asheim, B.T., Isaksen, A., 1996: Location, agglomeration and innovation: towards regional innovation systems in Norway. STEP report no 13/96, The STEP group, Oslo.

Bråthen, S., Hervik, A., 1997: Strait Crossings and Economic Development. *Transport Policy*, 4 (4) 193-200.

EURET/385/94, 1994: Cost-benefit and Multi-Criteria Analysis for New Road Construction. EU, DG VII.

European Commission, 1994: Europe 2000+. Co-operation for European territorial development.

Forslund, U., Karlsson, C., 1991: Regional Impacts of Infrastructure. Ministry of Industry, Sweden (In Swedish).

Gårder, P., 1989: The value of short time periods. Literature review. KTH, Meddelande 72, Stockholm, Sweden.

Harmatuck, D. J., 1996: The Influence of Transportation Infrastructure on Economic Development. *The Logistics and Transportation Review*, 32 (1), 63-76.

Hulten, C.R., Schwab, R.M., 1993: Infrastructure Spending: Where Do We Go from Here? *National Tax Journal*, XLVI (3).

Keeble, D. E., 1967: Models of Economic Development. In: Chorley, R. J., Haggett, P. (Eds.): *Models in Geography*. Methuen & Co, London, UK.

Krugman, P. R., 1991: Increasing returns and economic geography. *Journal of Political Economy*, 99 (3), 483-499.

Krugman, P. R., 1995: Urban Concentration: The Role of Increasing Returns and Transport Costs. *Proceedings of the World Bank Annual Conference on Development Economics 1994*. The World Bank.

Martin, R., 1999: *The new 'geographic turn' in economics: some critical reflections*. *Cambridge Journal of Economics* 1999 (23), 65-91.

Mohring, H., 1993: Maximizing, measuring, and not double counting transportation improvement benefits - a primer on closed-economy and open-economy cost-benefit-analysis. *Transportation research B* 27 (6), 413-424.

Myrdal, G., 1957: *Economic Theory and Under-Developed Regions*. Gerald Duckworth & Co. Ltd. London, UK.

Nas, T. F., 1997: Cost-benefit Analysis. Theory and Application. SAGE Publications, Inc. USA.

Romer, P., 1986: Capital accumulation and increasing returns. Ph.D thesis, Univ. of Chicago.

SANCTRA, 1999: Transport and the Economy. The Standing Advisory Committee on Trunk Road Assessment. Department of the Environment, Transport and the Regions, UK.

Sayer, A., 1992: *Method in Social Science*. Routledge, London/New York.

Yin, R. K., 1994: Case Study Research. Design and Methods. 2nd ed. SAGE Publications.

APPENDIX 1: Case study questions and a sketch of possible answers.

In this section, relevant topics that were covered in interviews with firm executives are presented as a guide for in-depth interviews with open-ended questions. The indicated *possible answers* were meant as inspiration for follow-up questions.

A.1. Recruitment of key personnel, R&D activities

1. Has the improved infrastructure affected the recruitment to technical appointments?

- YES (go to 1a)
- NO

1a. Are people recruited from other firms in the same or related business to this kind of jobs?

- YES
 - Executives?
 - Professionals?
- Has this been a deliberate policy?
 - If so, why? (possible answers:)
 - We have not considered business-specific knowledge as necessary, but we have tried to recruit people that may infuse new ways of ‘doing things’.
 - Transfer of experiences from other sectors is also valuable to us.
 - Which sectors, and why?
 - Have these new employees in technical positions imposed any changes in the production processes? How?

2. Has the improved infrastructure affected the recruitment to *administrative* appointments?

- YES (go to 2a)
- NO

2a. Are people recruited from other firms in the same or related business to this kind of jobs?

- YES
 - Executives?
 - Professionals?

 - Has this been a deliberate policy?
 - If so, why? (*possible answers:*)
 - We have not considered business-specific knowledge as necessary, but we have tried to recruit people that may infuse new ways of 'doing things'.
 - Transfer of experiences from other sectors is also valuable to us.
 - Which sectors, and why?
 - Have these new employees in administrative positions carried any changes in the administrative processes? How?

3. *Has the improved infrastructure affected the use of hired consultants?*

- YES. Why? (*possible answers:*)
 - Easier access to these services, and too expensive to increase our own staff.
 - Consultants gain experience from several firms, and these experiences are beneficial for us. If so, how?
- NO

4. *Has the improved infrastructure affected other ways of exchanging technological ideas?*

- YES. In what fields? (*possible answers:*)
 - Technology transfer by selling technology or licensing. If so, how?
 - Co-operation on use of R&D on technology. If so, how?
 - Employee exchange programs with firms in related businesses. How is this done, and why?
- NO

5. *Is the firm doing R&D work?*

- YES. In what fields? (*possible answers:*)
 - Product technology
 - The production process/technology
 - Product design
 - The production process
 - Internal organising of the firm

- How is this R&D work carried out? (*possible answers:*)
 - Co-operation with other firms or between in-firm divisions
 - Contact or co-operation with R&D milieus (research firms, consultants, universities)
 - Within the firm only

- You have now expressed how your firm organises its R&D activity. On a scale from 1 to 7 (1=not important, 4=medium importance and 7=very important), how would you rank the transport infrastructure's influence on how your firm organise and carry out the R&D activity?

- NO, this firm has no R&D activity.

A.2. *Development of relations with suppliers (the firm/respondent as a buyer)*

6. Has improved infrastructure made your firm taking specific actions concerning the firm's suppliers? (*possible answers:*)

- YES, increased co-operation on purchasing/production planning and development/other development issues (e.g. upgrading of personnel)
 - Exactly how has improved transportation infrastructure influenced this?
(*possible answers:*)
 - Suppliers are "closer" than before (mental distance), it is easier to plan through the working day (flexibility), transportation is cheaper (time savings), a more well functioning labour market has made us able to exploit these possibilities.

- Has this co-operation entailed cost reduction on supplies?
Exactly what is the main reason for this?
- Our firm has merged with/entered into an agreement with suppliers to prevent that our competitive advantages/knowledge are transferred to competitors.
- On a scale from 1 to 7 (1=not important , 4=medium importance and 7=very important), how would you rank the transport infrastructure's influence on how your firm organise and carry out this activity?
- NO, the transport infrastructure has not brought about any changes regarding the relationship with our supplying firms. Why? (*possible answers*)
 - We do not want to enter into relationships that allow anyone to look into our competitive advantages on this matter. Our suppliers may learn, and by generic product development transfer this knowledge on to our competitors.

A.3. Development of relations with customers (the firm/respondent as a seller)

7. Has improved infrastructure made your firm taking specific actions concerning the firm's customers? (*possible answers:*)

- YES, increased co-operation on purchasing/production planning and development/other development issues (e.g. upgrading of personnel)
 - Exactly how has improved transportation infrastructure influenced this?
(*possible answers:*)
 - Customers are "closer" than before (mental distance), it is easier to plan through the working day (flexibility), transportation is cheaper (time savings), a more well functioning labour market has made us able to exploit these possibilities.
 - Has this co-operation entailed cost reduction on your products?
Exactly what is the main reason for this?
 - The competitive milieu has become more aggressive, therefore we want to build 'strategic alliances'.
- How has the transportation infrastructure improvements influenced the relationship with your customers? (*possible answers:*)

- Improved deliverance regularity and frequency, flexibility, service as competitive factors.
 - Our firm has merged with/entered into an agreement with customers to prevent that our competitive advantages/knowledge is transferred to competitors.
- On a scale from 1 to 7 (1=not important , 4=medium importance and 7=very important), how would you rank the transport infrastructure's influence on how your firm organise and carry out its relations with customers?
- NO, the transportation infrastructure has not brought about any changes in our relationship with customers. Why? (*possible answers:*)
- We do not want to enter into relationships that allow anyone to look into our competitive advantages on this matter. Our customers may learn, and via generic product development transfer this knowledge on to our competitors.

A.4. The firm's buyer-supplier structure.

8. What kind of structure is present regarding supplying goods?
- Kind of deliverance? (goods types)
 - One or several suppliers? (no of suppliers)
- What has happened with this structure after the infrastructure improvements? (*possible answers*)
- Other suppliers than before?
 - Increased no of suppliers?
 - Within the local district?
 - Outside the district?
 - Why this change? (improved deliverance regularity etc.)
- Has the change in no of or kind of suppliers influenced the prices of supplying goods? If so, how?
- Does the firm use more goods than before? If so, why?
- Has the firm's stock of supplies been affected? If so, how?
- You have now expressed how your firm's relationships with suppliers have changed with respect to no of and kind of supplying firm(s). On a scale from 1 to 7 (1=not

important , 4=medium importance and 7=very important), how would you rank the transport infrastructure's influence on this matter?

9. What kind of structural relationship is present versus your customers?

- One or several customers?
- What has happened with this structure after the infrastructure improvements? (*possible follow-ups*)
 - Other customers than before?
 - Increased no of customers?
 - Within the local district?
 - Outside the district?
 - Why these changes?
 - Has the change in no of customers influenced the prices of your products?
How?
 - Is the firm selling more than before? If so, why?
 - Has the firm's stock of finished goods been affected? If so, how?
- You have now expressed how your firm's relationships with customers have changed with respect to no of customers. On a scale from 1 to 7 (1=not important , 4=medium importance and 7=very important), how would you rank the transport infrastructure's influence on this matter?

A.5. *The firm's relationship with the labour market - the broad picture*

- How do your firm experience that improved transportation infrastructure has influenced recruitment in general?
- How do your firm experience that improved transportation infrastructure has influenced recruitment for accompanying persons (e.g. spouses)?