BENCHMARKING AND SUSTAINABLE TRANSPORT POLICY
- SOME LESSONS FROM THE BEST NETWORK -

Henrik Gudmundsson[a], Lucy Gordon[b], Andrew Wyatt[c]
[a](Corr. author) National Environmental Research Institute, and FLUX- Center for Transport Research. Present address: NERI, PO BOX 358 DK-4000 Roskilde, Denmark.
[b]14 Shelley Court, Tite Street, London SW3 4JB, UK[1]
[c]Oxford Policy Management, 6 St Aldates Courtyard, 38 St Aldates, Oxford, OX1 1BN, Tel: +44 1865 297 300 [2]
Email: hgu@dmu.dk; lucy_a_gordon@yahoo.co.uk; andrew.wyatt@opml.co.uk

Abstract

Benchmarking is a widely used method of comparing performances and practices in order to learn from the best. In 2000 the European Commission initiated research to explore benchmarking as a tool to promote policies for 'sustainable transport'. This paper reports findings and recommendations on how to address this challenge. The findings suggest that benchmarking is a valuable tool that may indeed help to move forward the transport policy agenda. However, there are major conditions and limitations. First of all it is not always so straightforward to delimit, measure and compare transport services in order to establish a clear benchmark. Secondly 'sustainable transport' evokes a broad range of concerns that are hard to address fully at the level of specific practices. Thirdly policies are not directly comparable across space and context. For these reasons attempting to benchmark 'sustainable transport policies' against one another would be a highly complex task, which is generally not advised. Several other ways in which benchmarking and policy can support one another are identified in the analysis. This leads to a range of recommended initiatives to exploit the benefits of benchmarking in transport while avoiding some of the lurking pitfalls and dead ends.

Keywords: Benchmarking; Sustainable transport; Policy; Performance; BEST
Topic: H1 Public Sector Performance

1. Introduction

Transport companies and organisations around the world have taken on board benchmarking as a method for performance comparison, improvement and learning. Benchmarking has been applied to both freight and passenger transport and with respect to a wide range of parameters from reliability, speed, and costs of transport services to safety, environmental impact and customer satisfaction. Benchmarking has proved its value as a management tool in many such cases (Fearnley et al 2002; ECMT 2000a).

More recently benchmarking has been proposed as an instrument to help move forward the agenda of the European Union's Common Transport Policy (Deiss 2000). The present paper will focus on this much more ambitious application of benchmarking.

---

[1] Formerly with OGM Consulting, Brussels
2 Formerly with CMPS - Centre for Management and Policy Studies, Ascot, England

The specific aim of this paper is to summarise findings concerning benchmarking in the policy context of 'European sustainable transport' as drawn from the EU sponsored thematic network BEST (Benchmarking European Sustainable Transport), which functioned from 2000-2003. While the notion of sustainable transport is contested and has not been rigorously defined (Greene 2001) it is often considered to envisage a need to balance increasing demands for mobility with the needs to ensure economic efficiency, environmental protection and social and regional cohesion. The application of benchmarking in this setting raises a number of questions such as: Can benchmarking be used for policy making to the same advantage as in private business? Can 'sustainable transport' be 'benchmarked' at all, and if so, how? Should the European Commission or other bodies instigate a program to promote 'sustainable transport benchmarking', and what should it include?

Giving practical recommendations to the European Commission on such questions was the difficult task of the present authors. The aim of this paper is to relay to the wider community of transport professionals some of the thinking behind these recommendations. The paper should thereby serve not only as advice to transport policy makers, but also help to move forward the analytical understanding of the strengths and limitations of benchmarking as a policy tool.

Methodologically, the paper builds on a series of six 2-day BEST conferences held in Brussels, Belgium in 2000 -2003. A broad range of transport policy makers, stakeholders, experts, and researchers attended the conferences. The paper draws from individual BEST conference presentations as well as comprehensive reports and conclusions from each conference, and the final set of recommendations from the BEST network to the EU Commission (see OGM (ed) 2003). Also included is experience from a parallel project called BOB (=Benchmarking of Benchmarking), in which three actual pilot transport benchmarking projects were undertaken to explore the general propositions from the BEST network.[3] All BEST and BOB material including conference presentations, reports and conclusions have been documented at the BEST website[4].

It should be noted that the analysis and interpretation of the BEST and BOB projects in this paper are strictly those of the present authors. The paper does not represent viewpoints of the European Commission or any other participants in the BEST conferences.

The remainder of the paper is divided into six sections. The following section 2 addresses the basic concept of benchmarking and the general conditions to be observed in order to use it successfully. The following three sections discuss the application of benchmarking in the BEST areas of concern, considering in turn the transport sector generally (section 3), the field of sustainable transport more specifically, and finally the context of policies and policy making (section 4). Based on this analysis five different linkages between policy and benchmarking are distinguished (section 5). For each of them key recommendations to the European Commission on how to proceed in order to support sustainable transport policy are substantiated. Section 6 exemplifies a proposed procedure to 'screen' the benchmarking potential of various sustainable transport policy issues. In Section 7 overall conclusions are drawn together and some research needs are summarised.

2. Benchmarking

Benchmarking is traditionally understood as a method of comparing the performances of similar organisations or processes, in order to learn from the best performers and thereby improve own performance. Since its origin as a management tool for industry in

[3] The three BOB pilot projects concerned: 1) Passenger rail (performance and institutions), 2) Airport ground accessibility, and 3) Professional Road Transport Safety

[4] At the website www.bestransport.org

the early 1980s benchmarking has come into widespread use across most business and policy sectors. A substantial literature on benchmarking has emerged, covering prescriptive as well as more analytical contributions (see e.g. Keegan 1998; Yasin 2002; Fearnley et al 2002). A few of the key concepts can be described as follows:

- A *benchmark* is a standard of excellence or performance against which other similar things are measured or judged. The standard is normally measured by indicators.
- *Best Practice* is the means (practice or activity) by which that benchmark level of performance is achieved, for example a particularly effective production method
- *Performance benchmarking* – involving comparisons of performance data for similar products or services. Considered as the 'basic' or 'early' benchmarking notion.
- *Process benchmarking* – involving qualitative comparisons of similar practices etc, across various organisations. More sophisticated; considered necessary to understand performance and learn from Best Practice
- *Functional benchmarking* – involving comparisons of practices or work processes for similar generic functions across entirely different sectors. More challenging still.

While technical concepts was not a key concern for the BEST it was nevertheless soon realised that a clear understanding of the method was needed in order to reap the full benefits from it. For example, comparisons of league tables or scorecards in a particular area, say just-in-time delivery, are not enough (= performance benchmarking). Successful exploitation of benchmarking requires also the understanding of *reasons behind variations in performances* (e.g. whether better just-in-time performance is explained by better practices or by differences in external conditions), and not least prepared*ness to implement relevant changes* of practice in one's own organisation on a continuous basis (= involving process benchmarking).

To reflect these insights the BEST network adopted a comprehensive definition of benchmarking from the European Commissions Directorate General for Enterprise:

"…the continuous, systematic process for comparing performances of organisations, functions or processes against the "best in the world", aiming not only to match those performance levels, but to exceed them" (CEC 1996).

From this (prescriptive) point of view benchmarking can be understood as a process that ideally should involve the following main steps:

- Planning (selecting the subject to benchmark, identifying partners in the process, defining the objective of the exercise);
- Analysis (data collection, understanding own performance, making comparisons, identifying best performers);
- Integration (review of results, and integration or translation of transferable results into own organisational context);
- Action (implementation of actions and monitoring to ensure continuous improvement);

These steps have been laid out in a comprehensive *benchmarking guide,* one of the main outputs of the BOB and BEST projects.[5]

If performed well along these lines benchmarking may provide several benefits: First of all it enables a systematic approach to understanding, comparing and improving practices and performances of organisations. Secondly benchmarking is not based on abstract aims but on learning from what has successfully been done in practice elsewhere. For those involved, benchmarking is therefore likely to inject realism into the perception of one's own performance as well as into visions of potential improvements. Generally speaking benchmarking can help to increase reflexivity (self-awareness and even self-criticism)

[5] To be put on the BEST website during 2004. (http://www.bestransport.org/cadrebest.html)

through learning from other organisations and to build a performance-based culture in the organisations and sectors that brings it to use.

However experience obtained through the BEST and BOB projects also revealed that benchmarking does not always deliver the benefits assumed in the prescriptive view. Some of the key observed *conditions* for successful benchmarking are:

- As an intensive, multi-staged process it demands significant commitment of resources (human and financial) by participant organisations and acceptance of a time-scale, which is likely to run into years rather than months.
- Participants involved in a benchmarking exercise must share a vision of what constitutes success. On this basis the best performing organisations be identified, their practices analysed and lessons learnt by other participants.
- There must be access to the appropriate skills to identify, collect, analyse and understand extensive sets of data.
- A high level of trust between participants is required (or has to develop), both to share information with others and to rely on the data provided by others
- Sufficient autonomy to take necessary action is needed since the benchmarking process is not complete until action has been taken, and its effectiveness evaluated.

If such conditions are not observed it may severely restrict the potential benefits of the method and even reduce the chances of producing a meaningful result. This further implies that benchmarking may not always be the most feasible approach to improve performance. Rather benchmarking should be considered as one distinct method among a series of options that would include also e.g. the use of quality assessment, performance indicators, rankings, evaluations, peer reviews, etc. These various performance management methods may serve as alternatives to or forerunners for benchmarking but should not be confused with it.

3. Benchmarking transport

Considering the above opportunities and conditions, what can be said about benchmarking in the transport sector? As the BEST process has clearly demonstrated, transport represents an almost infinite range of possible applications. About 40 examples of benchmarking projects across the whole transport field were presented at the six BEST conferences. Some projects were regional or national in scope while several others involved comparisons across countries within Europe or on a global scale.

Table 1 illustrates a range of the themes and projects that were covered at the BEST conferences while Figure 1 exemplifies three specific applications (Please consult the Appendix for acronyms and the BEST website for further context).

While the wide range of projects does bear evidence to the general applicability of the benchmarking tool in transport it does not in itself prove that benchmarking can be used with universal success across the whole transport sector. From the discussions in BEST we have extracted three general observations in this regard:

1) Success of transport benchmarking projects was most often described in terms of *qualitative* benefits, e.g. increased awareness of own performance, stronger documentation of certain performance elements, more focus on customer needs, and better knowledge of alternative practices. *Quantitative improvements* in performance directly ascribed to implementing the results of benchmarking efforts were not extensively documented (but see Wallis 2001, Oakland 2003 for examples). More often benchmarking projects helped to quantify performance improvements that could not necessarily be ascribed to the benchmarking exercise per se. Direct quantitative *comparisons* of performance were often found to be difficult, or only possible in relative terms.

Table 1 Examples of benchmarking themes and criteria dimensions

Themes	Criteria dimensions (ex.)	Project examples
Passenger transport		
Urban public transport systems	Cost-effectiveness, service provision, accessibility, prices, customer satisfaction, etc	Citizens' Network; B.E.S.T; Emilia-Romagna Region
Metro systems of major cities	Efficiency, reliability, safety, financial management, etc	COMET; NOVA
Passenger railways	Standards of punctuality, passenger transport growth, etc	BOB-Passenger rail; Dutch Railways; Railbench
International airports	Customer satisfaction (service parameters ranging from signposting to shopping)	IATA Airport Monitor; Vancouver Airport
Ground accessibility to airports	Rail access to airports, travel time, parking provisions, etc.	BOB-Air; IARO
Freight transport		
Container terminals/ports	Efficiency to support intermodal transport chains	OECD
Freight Transport Service	Delivery time, consignment care, corporate efficiency, etc	ESC/FTA
Transport policies		
Transport infrastructures	Capacity utilisation, time delays, price/performance relationships, etc	Dutch Ministry of Transport
Metropolitan Transport Planning	Quality of planning processes	Federal Department of Transportation, USA
National Cycling policies in Europe	The use of targets, policy measures, planning processes for cycling policy	Natcyp
Integrated transport policies in European countries	Congestion, road safety, environment, exclusion, health	CfiT – UK
Transport safety policies	Driver training; Road safety culture of transport companies	BOB-Road
CO2 reduction strategies in road transport	Effectiveness of various policy instruments	ECN/COWI

2) Only a few of the presented transport benchmarking projects have been running continuously for several years, indicating that success in this respect is possible (e.g. COMET, B.E.S.T, IATA Airport Monitor) but not universal. This appears to occur more frequently if there is a foundation on which to build in advance (in terms of pre-existing collaboration, performance data, etc), if time and financial support are allowed to overcome initial difficulties, and if strong marketing of results to clients is undertaken. Mostly these projects have been initiated by strong (transport) organisations with continuous professional support. The typical focus is on conventional performance criteria for ´standard´ modal operations, such as urban transport companies and airports.

3) Several other transport benchmarking projects were completed as one-off exercises, including many of the policy-driven, experimental or intermodal projects (e.g. BOB-Pilots, Natcyp, CfiT, CO2-policy, etc). One-time benchmarking projects do not strictly fulfil the methodological requirement calling for a process of continued improvement efforts, but most of these projects appear to have produced useful information anyway. Several of these projects were, however, modified along the way, typically narrowing down significantly the original ambitions. Modifications were often related to breaches in the

general conditions for success mentioned in section 2 such as lacking availability of comparable data (or lacking capacity to generate them), differences in vision, concepts or purpose, or lack of trust. These experiences confirm that a wholesale promotion of benchmarking is not generally advised: It does require much effort, and can lead to frustrations.

The Citizens' Network Benchmarking Initiative, sponsored by the European Commission, involved European cities and regions in the evaluation and comparison of the performance of local transport networks. The participants were representatives of local and regional authorities. The first phase involved 15 cities; in the second phase approximately 40 cities and regions joined. Around 40 common performance indicators were defined, ranging from number of trips per inhabitant, to cost of car use, to length of bicycle networks, to average speed for public transport. In addition working groups were set up across cities to identify potentially transferable best practices in specific areas such as accessibility, intermodality, and quality and contracts in public transport. A key message was that some valuable learning was enabled despite the fact that fully comparable quantitative indicators could in many cases not be established.
A project of the European Shippers Council (ESC) and the Freight Transport Association (FTA), defined a set of Service Performance Indicators (SPI) to benchmark the performance of freight transport supply chains. Participants were mainly shippers and manufacturers. Originally focussing on the performance of the air freight industry only, the approach has evolved to encompass approximately 15 generic indicators across all modes of freight transport (road, deep-sea, short-sea, rail, air, inland waterway etc.). The indicators were defined for four main areas: goods arrival on time; arrival without damages; compliance with legal requirements and efficiency in terms meeting the objectives of the different parties in the transactions. A key message from the project was the need to focus on a limited range of clear indicators to remain operational.
The BOB- railway pilot project aimed to assess to what extent benchmarking can be used to increase the effectiveness of railway services, both from the point of view of operators and transport authorities. The participants were rail authorities and operators from 14 countries. Main themes included delays/punctuality, passenger safety, customer satisfaction and growth in rail patronage. Difficulties were encountered in defining indicators and providing comparable data for all areas. The best results were achieved concerning the two areas punctuality and growth in passenger patronage, where a learning potential was identified. A key result of the project was a detailed analysis on causes of delays in passenger railways. This analysis defined common measures of delay and clarified the roles of the different actors (authority, operator, infrastructure provider, regulator) etc.

Figure 1. Three examples of transport benchmarking

Some of these observations 1) – 3) may be due to particular conditions for benchmarking in the transport sector. The BEST process identified the following range of characteristics of the transport sector that appear to represent special challenges:

a) Transport is a service industry where many of the services are intangible and specific to time and place. This can make performance comparisons of quality, service etc. more difficult than in products sectors. This is often further complicated by different delimitations of transport modes, urban catchment areas etc. To the extent that common definitions and measures of e.g. service quality have already been set up, benchmarking may be greatly facilitated.[6]

[6] Examples include quality criteria defined for public passenger service (European Standard EN 13816) and Service Performance Indicators of freight transport as defined by the European Shippers Council (ESC) – see OGM (ed.) 2003 for references.

b) Comparisons of performance quality can also be difficult due to culturally dependent perceptions of mobility and related factors. Objective parameters such as 'minutes of delay' may therefore be difficult to interpret in an (international) comparative context, while subjective parameters such as customer satisfaction may be difficult to link with specific operational variables.

c) The transport sector functions widely on a modal basis. This is clearly an obstacle to promoting policy objectives such as customer and user focus, intermodality and change of modal split towards rail and short-sea shipping. Conversely, in an intermodal perspective it can be difficult to define and delimit the transport chains over which performance should be measured and compared.

d) The transport sector is the cause of substantial amounts of external effects in terms of congestion, traffic accidents and pollution. As transport organisations often have little incentive to internalise such effects by themselves, there is also limited motivation to collect data and analyse performance in those areas, *even if information on them may exist.*

e) Extensive restructuring is under way in the transport industry, in terms of devolution, privatisation etc. This means that conditions of trust may be difficult to maintain where former collaborators are changed into competitors or client/server relations.

f) Transport policies represent a particularly challenging area for benchmarking. Major problems arise in trying to define a clear and legitimate object to benchmark, and in the attempts to transfer results and practices across contexts. This problem is addressed in section 5 of this paper.

Further research would be needed to try to actually explain success or failure of transport benchmarking efforts by way of the specific sector features identified in the above. Such a task would greatly benefit from a cross-examination of the material on transport benchmarking identified and collected in the BEST and BOB projects, and only briefly reviewed here. Hypotheses for this research concerning the success or failure of transport benchmarking projects could be derived in two dimensions, namely,

1) The extent to which projects have adhered to key steps in the general benchmarking prescriptions as laid out in the BOB benchmarking guide (see section 2 above), and
2) The extent to which project outcomes have been affected by the particular conditions of transport as laid down mentioned in points a) – f) above.

4. Benchmarking sustainable transport

The BEST project was intended to address Benchmarking *Sustainable* European Transport, echoing the overall aim of European transport policies. However, participants in the BEST process encountered difficulties in integrating the concept of sustainable transport in a benchmarking context. Initially two basic approaches were identified:

The *first approach* is holistic in subscribing to the general notion of sustainable development as an effort to integrate economic, social and environmental goals in an overall transport strategy (CEC 2001b; Greene 2001; ECMT 2000b). In this approach long-term goals and objectives would be defined and measured by indicators for the three dimensions, and these objectives and indicators would again serve as key criteria in establishing what is 'best practice' in terms of sustainability. An important element of this holistic approach would be that the multiple goals and dimensions of sustainability should be pursued together, while optimising in one dimension alone (e.g. economic performance or environmental protection only) is discouraged. However, this top-down approach would ideally require that goals of sustainable transport were defined in advance, a condition that is rarely fulfilled today. Moreover, the approach would not help very much to identify which particular practices should be considered for benchmarking. In other words the holistic approach in itself is insufficient for sustainable transport *benchmarking*.

The *second approach* is more pragmatic as it refers to the implementation of policies that are believed to contribute towards sustainability even though the goals may not have been explicitly spelled out. A key example would be the EU Transport White Paper (CEC 2001a), which identifies a range of relevant policy objectives, for instance measures to:

- Shift the balance of modes towards rail, public transport, and cycling;
- Develop intermodal capacity;
- Remove bottlenecks to effective market integration;
- Reduce the number of traffic accidents;
- Internalise infrastructure and external costs in pricing and charging structures.

This approach indeed opens a wide range of possible practical applications of benchmarking. However it also has its limitations. First of all, it is not clear how or how much the specific policy measure will actually contribute to sustainability in each of its dimensions. Such an assessment would require the introduction of several criteria taking into account both the full range of impacts as well as the role of each measure in the transport system as a whole. Otherwise the term 'sustainable' would become quite meaningless, since there is no guarantee that e.g. any efforts to remove bottlenecks would benefit the environment. This approach is therefore insufficient for *sustainable* transport benchmarking.

Recurring confrontations of the two types of approach in the BEST process made it clear that some level of tension is unavoidable between normative, holistic and strongly integrative notions such as sustainable development on the one hand, and the need for focus, manageability and incremental progress inherent in the successful application of tools like benchmarking on the other. Two potential 'pitfalls' define the road ahead:

1) Benchmarking efforts are at constant risk for becoming submerged under a mass of data. This hazard is likely to be even more pronounced when dealing with complex issues like sustainability. Losing track of practical experience to learn from is a highly probable result.

2) The need for focus and manageability in conventional benchmarking may jeopardise the comprehensive view implied in sustainable development. This may in the worst case install a 'tunnel-vision' for instance if 'best practices' are identified only by the narrowest range of economic performance criteria, leaving the wider environmental concerns of out the equation.

That the latter should be a genuine concern was clearly disclosed in the BEST process, where one review for instance revealed that the environmental dimension has not been widely adopted in mainstream transport benchmarking so far: While a few benchmarking projects specifically addressing the environment have been undertaken, environmental criteria and indicators are generally not yet well integrated in the measurement of transport performance in general (see Table 2, and further in Gudmundsson 2003). This is often justified by limited data availability, but is may also have to do with limited awareness of the data, which actually exist as well as lacking experience to exploit them in the sector.

Table 2. Performance Indicators in selected European transport projects presented at BEST

Project	Number of Performance indicators	Of which environmental
ISOTOPE	9	(1)
EQUIP	111	(2)
Quattro	11	0
CoMET /NOVA	32	0
B.E.S.T. (project)	36	1
Citizens Network	39	(2)
Three BOB pilots	[not calculated]	0

Note: Values in brackets refer to projects where environmental indicators have been considered but not included in actual measurements. For Acronyms and project names see the Appendix
Table adapted from: Gudmundsson (2003)
Sources: www.bestransport.org - www.eltis.org - www.europrojects.ie/equip

One recommendation concerning benchmarking for sustainable transport was therefore to generally support and promote the integration of environmental issues, data and indicators into the conceptualisation and operationalisation of 'mainstream' transport performance, and to raise awareness of environmental data and indicators already available. A concrete step at the European level would be to introduce benchmarking to support the so-called Cardiff-Strategy for the integration of environmental protection in the transport policy, initiated at the EU's Cardiff Summit in 1998 (Council (Transport) 1999). The operational indicators defined within this strategy framework (embodied in the so-called Transport Environment Reporting Mechanism or 'TERM' (EEA 2000)) could be explored as measures for transport benchmarking in this respect. While this would not necessarily overcome the tensions between the holistic approach and the manageability approach, it may at least help to explore the role of benchmarking in the service of sustainability further.

5. Benchmarking in a policy context

A key task of the present analysis has been to identify the most effective role for benchmarking in relation to European sustainable transport *policy*, that is, how *policy making* can help the transport sector exploit benchmarking and how benchmarking can help develop and implement key sustainable *transport policies*.

A general discussion in the BEST process has been concerned with the extent to which benchmarking can be used to support and improve policy making, in the same way that it has helped to improve management and performance in the private business context where the method originates. The following reasons for a cautious approach to 'policy benchmarking' came out (see further Wyatt 2002):

- Policy settings differ widely (e.g. in terms of history, culture, geography, institutions or political constellations). This will often make direct policy comparisons and transfer highly problematic, to say the least
- Policy often involves several different goals and objectives. A wide range of stakeholders may have legitimate interests in the way it is conceived and implemented. This means, that what is 'best practice' may well be disputed by different stakeholders and members of the society concerned.
- Policy involves different stages and activities (objective setting, analysis of alternatives, decision, implementation, evaluation, etc). The role and use of benchmarking may shift over the course of policy development, as the composition of 'players' and interests change (e.g. between politicians, management, stakeholders, experts)
- Policymaking comes in different styles. Benchmarking may be easier to exploit in a setting of 'rationalist' policy planning, than in settings dominated by power struggles or led by visions of charismatic politicians (May 2002)
- Public policy is expected to take into account the external effects of policies to a higher degree than private organisations. Otherwise there is a risk of creating new externalities or even policy failures, generating new problems even as existing ones are addressed. This can further increase the complexity and interpretation of information

- Policy is exposed to media interest. While this may help to propagate important benchmarking results to a wider audience it may also sometimes distort the attention and sacrifice a potential for genuine learning
- Policy implementation is typically more complex than management. In a democratic society policy makers very often can only aim to establish conditions, which will influence the autonomous decisions taken by private individuals and entities.

While these reservations taken together may appear restrictive, they do certainly not prohibit the exploitation of benchmarking for policy purposes. Rather the BEST process has made it clear that there is not one, but many possible roles for benchmarking in the policy context. More specifically five different types of relation between policy and benchmarking have been identified and considered:

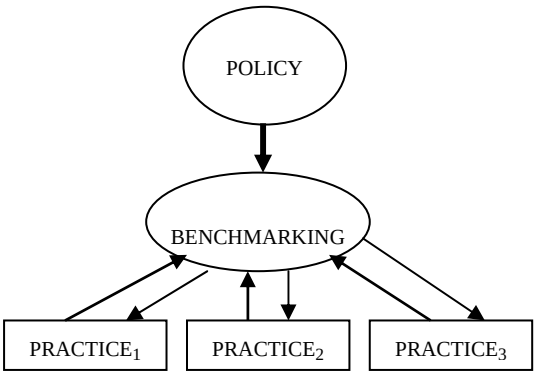
1) Policy *on* benchmarking, involving strategies to help and promote the proper use of benchmarking by stakeholders in the sector generally
2) Benchmarking *into* policy, concerning efforts to exploit results from already existing benchmarking studies in the development of policies
3) Benchmarking *for* policy, instigating original benchmarking projects with the specific purpose to support the development of particular polices
4) Benchmarking *of* policy, comparing performance of policies to one another in order to identify and/or transfer results
5) Benchmarking of policy making, comparing policy making processes in order to identify and/or transfer best practices in process terms

Figure 2 aims to illustrate the logic in each type of relation. The following sections will for each relation provide a recommended approach to EU transport policy makers in pursuit of sustainable transport objectives.

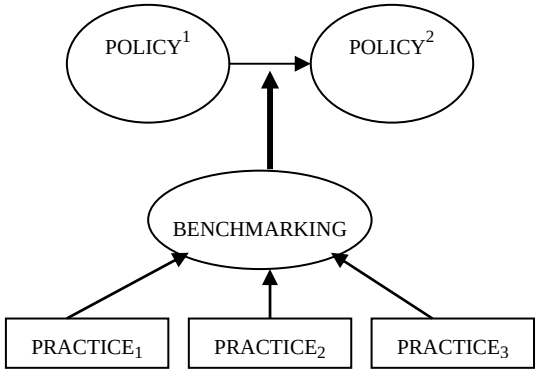
5.1. Policy on benchmarking

Benchmarking projects may be initiated by organisations entirely as a 'bottom-up' process with a focus on whatever topic is relevant for the particular organisations or businesses. Efforts to help and guide bottom-up benchmarking initiatives towards realisation and success can be described as a *policy on benchmarking*.

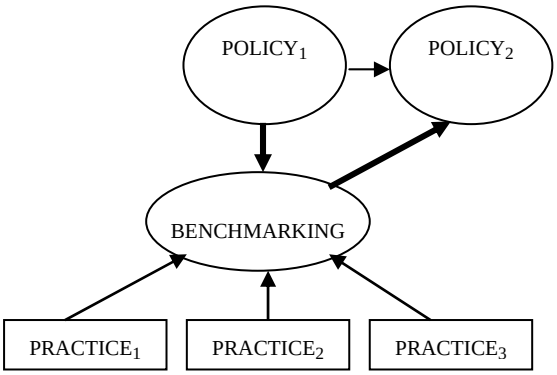
The general recommendation from the BEST network to the European Commission (and national policy makers) was to adopt a positive and supportive policy on benchmarking in the transport sector. More specifically this promotional effort could include support in the form of a Benchmarking Information Clearinghouse (e.g. a website), staging of regular conferences, provision of network support, and process facilitation.



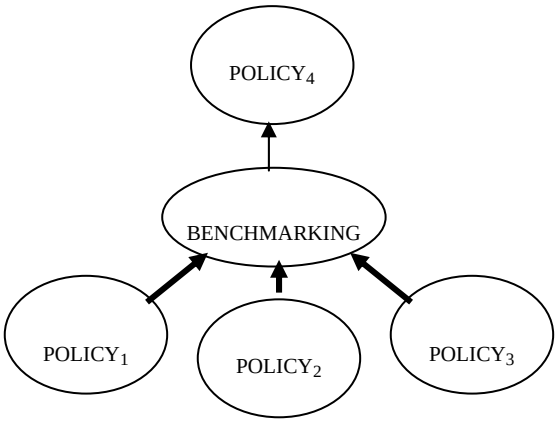
A) POLICY *ON* BENCHMARKING



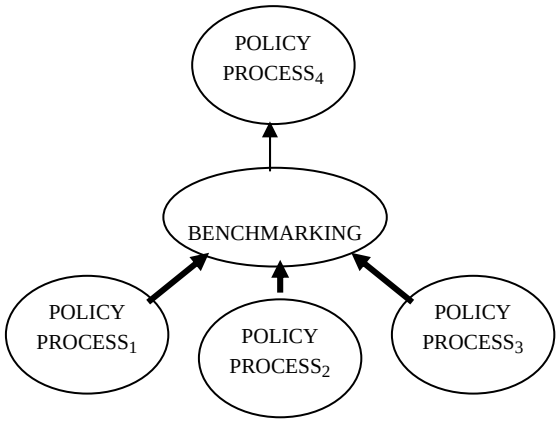
B) BENCHMARKING ***INTO*** POLICY



C) BENCHMARKING ***FOR*** POLICY



D) BENCHMARKING *OF* POLICY



E) BENCHMARKING OF POLICY ***MAKING***

Figure 2: Five different roles of benchmarking in relation to policy

A key element in the effort would be the elaboration of a guideline to ensure that the proper conditions for successful benchmarking are observed, as generally laid out in the BOB benchmarking guide. Moreover a guideline or checklist should be developed to ensure that relevant measures and indicators for all dimensions of sustainable transport (including the environment) are considered. A requirement to address the full set of dimensions could even be made for any transport benchmarking exercise that was to receive EU or national support. The latter would be relevant in particular to avoid performance or 'Best Practice' being defined only in very narrow terms, ignoring possible externalities.

5.2. Benchmarking into policy

Benchmarking into policy concerns how policy-relevant information and transferable lessons can be systematically distilled from studies that have already been defined or concluded, and then fed into policy making. This will entail at the very least validating the methodology used and the robustness of any statistical analysis. For meaningful lessons to be learnt by users other than the original participants, analysis should also go further to consider the conditions under which the transferability of policy lessons from one administrative or operational context to another can be assessed. A starting point for establishing an appropriate methodology for the comparative analysis and appraisal of transport benchmarking studies may be found in the systematic reviews undertaken by collaborations within medical and social policy research. These techniques are well

understood in the research communities in these fields, and have attracted a substantial literature in recent years (see e.g. Davies & Boruch 2001).

The recommendation to the European Commission was to define – in collaboration with interested Member States - areas of key policy interest where a number of independent benchmarking studies have already been conducted. Examples include urban public transport, intermodal supply chains, and performance of passenger railways. For each area a systematic review process could be instigated in order to distil possibly transferable policy relevant results. The results of such a review would provide input to the development of EU policies, targets, and standards to support the implementation of overarching policy objectives.

5.3. Benchmarking for policy

A second level of the above process would be to *initiate* benchmarking projects in particular areas in order to directly exploit and learn from them in the development of policies. This can be described as the use of *benchmarking for policy*. Arguably this represents the most interesting area of application.

Benchmarking studies may generally point towards aspects of the framework conditions within which operators function - for example regarding funding mechanisms, or the way in which regulatory or inspection regimes are implemented - that can be associated with high or low performance, and suggest areas for improvement. They may also indicate aspects of transport industry performance where national authorities can work with operators to raise performance to the levels equated with European or World best practice. One example from BEST is the BOB-rail project in which the analysis of causes of train delays could potentially lead to changes in the allocation of tasks between responsible actors and therefore to more effective contractual relations.

In order to be fully supportive of policy making, however, the indicators chosen and the 'best practices' identified for benchmarking should always reflect the wider context in which the policy is (to be) applied, including wider policy goals, stakeholders, and background factors. At the very least, it is important to avoid identifying or achieving 'best practice' through ignoring other relevant major policy concerns, e.g. ignoring environmental impacts in promoting superior economic performance or vice versa.

Table 3 Examples of key sustainable transport policy related goals and indicators

Decoupling of transport growth from economic growth
- Volume of freight transport relative to GDP
- Volume of passenger transport relative to GDP

Change modal split of freight and passenger transport (return to 1998 levels by 2010)
- Modal split of freight transport – percentage share of road
- Modal split of passengers transport – percentage share of cars

Traffic fatalities (Halving number between 2000 and 2010)
- Road fatalities per million persons

Improving Urban Air Quality
- Population exposure to air pollution by ozone (Urban air quality)
- Population exposure to air pollution by particulate matter (PM10)

Contribution to Kyoto target
- Transport emissions of greenhouse gasses by mode

Reduction of Vehicle Pollution
- Transport emissions of air pollutants

Reducing noise from road, railways and aviation
(no clear indicator defined yet)

The approach recommended to the European Commission in terms of benchmarking for sustainable transport policies was to start with the identification of (5-10) key policy objectives for which benchmarking could be explored. The potential policy areas could be drawn from key documents such as the Transport White Paper (CEC 2001a), the Sustainable Development (CEC 2001b) strategy, and the Strategy for integration of the environment. (Council (Transport) 1999). A list of key objectives and related indicators are distilled from these documents in table 3. The further selection of policy areas to benchmark should be carried out in consultation with EU Member States and Accession Countries, as well as relevant organisations. The selection of issues should be supported by a screening device to identify areas for which benchmarking may be a feasible approach. Such a device is exemplified below in section 6. Benchmarking groups could be set up for each policy area identified. Participation would be on a voluntary basis, although participants would have to meet set criteria for participation (availability of adequate resources, commitment to attend meetings and share data etc.). Where appropriate, representatives of other sectors – e.g. environment – and countries – e.g. USA, Canada, and Japan - would be invited to participate.

5.4. Benchmarking of policy

The possibility of the *benchmarking of policy* has also been examined in the course of the BEST project. Benchmarking of policy implies the application of benchmarking methodologies to compare the performance of different policies adopted by different administrations.

There are, however, considerable difficulties with this notion. First of all the essentially managerial methodologies of benchmarking do not necessarily transpose easily into the higher levels of complexity and contestability in the policy realm, as indicated by the points of reservation at the beginning of section 5. For a start, there are problems of definition. A policy cannot be regarded separately from its objective: it can only be understood in terms of the course of action adopted by a government to achieve some particular end, and that objective itself reflects the values, preferences and priorities of the society in which it has arisen. It is therefore likely that different administrations will be pursuing objectives that differ in varying degrees, even if they are all related to a common overarching aim (for example as regards environmental sustainability). Whilst it might in theory be possible to benchmark the effectiveness of different policies in delivering the ultimate aim, it is unlikely that agreement could be gained on the terms of the comparison. Alternatively, if precise equivalence between policy objectives could be established, it might be possible to compare the relative effectiveness of the different policy instruments employed in different administrations to achieve the shared objective. Even then, careful allowance would have to be made for differences in the wide range of social, economic, cultural, institutional, technological and geographical factors that may have had a bearing on success or failure.

In view of these complications, benchmarking of policies is not generally recommended. It appears that the notion of the benchmarking of policy may simply stretch the concepts and language of benchmarking too far. There is a well-established discipline

of policy evaluation within which the expertise to deal with these issues has developed, and it would probably be more productive to re-frame questions about policy benchmarking in terms of the comparative evaluation of particular policies or instruments (Wyatt 2002).

5.5. Benchmarking of policy making and management

Instead of the benchmarking of policies in terms of their contents, performance or outcomes, it would potentially be valuable to benchmark the processes of policy making in different administrations. This can be regarded as the *benchmarking of policy making*. Such an exercise would be conducted as a form of process rather than performance benchmarking; it would not focus directly on the effectiveness of the policies but on the quality of the processes in place in different administrations to support decision making and implement decisions taken (see e.g. Lyons 2002 for an example from USA and the Natcyp project for the European case of cycling policies).

General issues that might be considered for benchmarking of policy making include consultation, engagement of civil society and other stakeholders, use of evidence and expert advice, means of ensuring policy coherence across the whole of government, and the use of risk assessment and risk management techniques. Concepts such as the policy cycle may be employed to identify overall steps in which particular management practices or institutional designs could become the subject of benchmarking. Still, it is not clear how easy it would be to secure wide agreement on the criteria for good policy making, as convergence between the administrative cultures of Europe is at a relatively early stage.

One recommendation to the European Commission, which was also mentioned in section 4 on sustainable transport, was to use benchmarking to support in the process of integrating environmental factors into national transport policy making. Another general recommendation was to include counterparts from other regions of the world. The US Department of Transportation has extensive experience in monitoring the performance of transport policies and processes at federal and state level. The Canadian Ministry of Transport has given special attention to monitoring the implementation and management of sustainable transport strategies.

6. A scoping procedure

Clearly not all transport policy topics would lend themselves to benchmarking in the sense that it has been defined and explored here. Starting from key policy objectives and indicators such as those mentioned in table 3, a screening process has therefore been proposed to the European Commissions in order to help select the more from the less promising options. The process would basically consult the following criteria:

- High policy relevance
- Availability of indicators to measure performance
- A learning potential due to variations in performance
- A clear role for identifiable practices in achieving good performance.

If a positive reply can be given to all steps, then the opportunities for added policy value of benchmarking might be good. If not, other opportunities to improve performance or policy outcomes may be pursued instead. The diagram in Figure 3 illustrates this screening process. To exemplify the procedure we will briefly look into a hypothetical situation where *modal split* is considered as a potential policy priority for benchmarking. This is how the procedure would work:

(1) High policy relevance?

Yes, to influence the modal split of both freight and passenger transport is a key European policy concern. The overall European aim is that the share of road transport in 2010 is not greater than in 1998 (CEC 2001a).

(2) Indicators available?

Relatively comparable European data on modal shares of transport exist, and the modal split has even been included among the structural indicators monitored in the so-called Lisbon Process (CEC 2002a). The data may not be perfect, but there is a basis to build on.

(3) Variations in performance?

The structural indicators and other sources reveal that there is a considerable variation in the modal composition of the freight transport flows in the Member States, whereas the modal split is quite even for passenger transport. This suggests a focus on freight.

(4) Variations due to practices?

To address this step it would be necessary to identify more detailed *policy related practices* that could potentially contribute to a change in modal spilt. In the particular case of modal split policy we could for instance look for practices among the so-called intermodal freight transport actions envisaged in the EU's Marco Polo programme [7]

(5) Exploring benchmarking options

In this final step an assessment would be made of each identified relevant practice in terms of the *operational feasibility* of undertaking a benchmarking study for the particular practice, consulting conditions such as those listed in Section 2 of this paper (e.g. commitment of resources by participant organisations, shared visions of success, capacity to analyse and understand extensive sets of data, high level of trust, autonomy, etc)

[7] E. g: Procedures in sea and inland ports to integrate short-sea shipping and inland waterway into the transport chain; Pricing, procedures and methods in the terminal haul; Training centres for rail, inland waterway and freight forwarding professionals. (CEC 2002b).

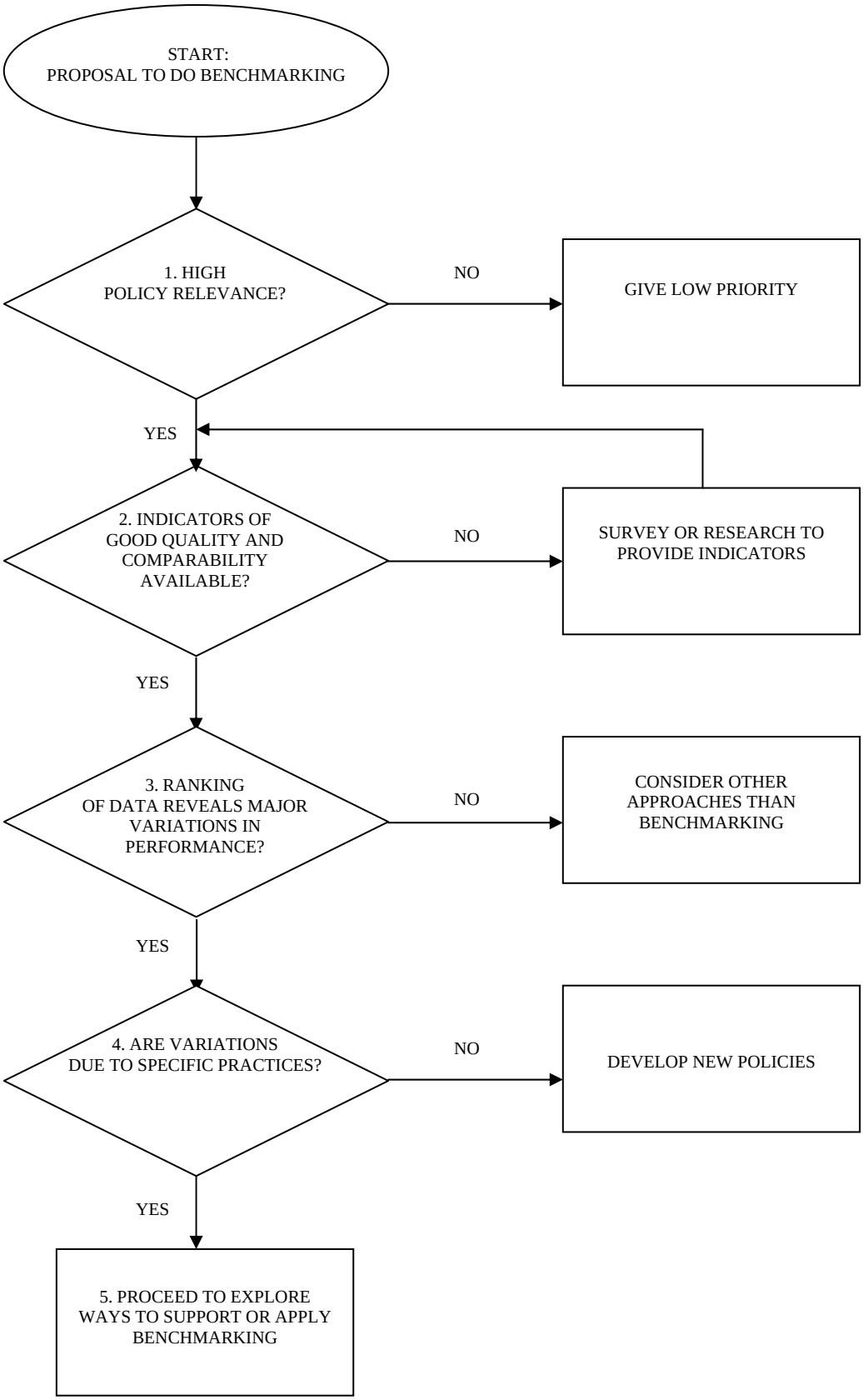


Figure 3 Screening chart to select topics for benchmarking

7. Summary of findings and recommendations

We conclude that benchmarking is a potentially useful instrument that should be considered as part of a wider effort to promote performance improvement instruments in the transport sector. Benchmarking needs to be laid out carefully, and should not necessarily be considered the first option. Benchmarking efforts should, where possible, build on existing work to define and measure performance outputs (in terms of service quality, reliability, environmental performance etc).

Benchmarking can be applied in the transport sector, as in any other sector. Performance benchmarking of transport services and operations can however be difficult due to problems of defining, delivering and comparing transport services. Learning opportunities can be restricted by a lack of understanding of the links between performances and practices. Narrowing the focus of benchmarking to only one or a few parameters can be necessary (and productive) but will on the other hand often limit the relevance of the policy support, especially in terms of wider objectives and externalities relevant for sustainability policy. There is still a limited understanding of the particular conditions for successful benchmarking of transport arising from the special characteristics of the sector (service industry) and its current situation (e.g. modal traditions, undergoing institutional changes). A cross analysis of the projects presented in BEST/BOB would be a way to move this understanding forward. The overall policy recommendation is not to pursue transport benchmarking in isolation but as part of a broader package to promote measurement, management, experimentation and research in transport performance.

Applicability of benchmarking for sustainable transport depends strongly on how this notion is defined. Ideally the pursuit of sustainable transport should consider performance in the economic, social and environmental dimensions in an integrated way. On the other hand benchmarking of transport could rarely proceed if the operational availability of multi-dimensional information was a sine qua non. One recommendation to move forward is to at least insure that environmental objectives, measures and indicators are integrated into mainstream transport benchmarking at all levels, drawing on already existing indicators systems (such as TERM) and improved guidelines and methodologies.

Policy making can benefit from benchmarking both through the process of building commonly accepted standards and indicators of comparison, and by helping to identify good transferable practices, instead of ‘reinventing the wheel’. However there are limitations to the applicability of benchmarking in a policy setting due to generally high complexity, contestability and contingency. Substantial difficulties can be encountered e.g. in the identification of common objectives and standards, in the distinction between structurally-determined and policy-determined causes for variations in performance, in the commitment to open and unprejudiced learning, and the capacity to continuously implement new experience. Attempting to benchmark the performance of policies against one another ‘policy benchmarking’ is therefore generally not advised.

Four other distinct ways to obtain mutual support between policy and benchmarking have been identified. For each type specific recommended approaches to pursue sustainable transport policy objectives were put forward:

A *policy on benchmarking* should support voluntary bottom-up benchmarking initiatives, and provide guidance to enhance their quality and general relevance.

Benchmarking into policy could be pursued through systematic reviews of existing studies in areas such as urban public transport.

Benchmarking for policy should be pursued focussing on key issues of European policy concern, such as intermodality, externalities, bottlenecks and traffic safety. Each area should be scoped for benchmarking feasibility in terms of available indicators and

identifiable practices. For each relevant topic defined a multi-national team would be set up to explore benchmarking, while ensuring commitment, resources and time.

Finally *benchmarking of policy making* could be considered as an opportunity to improve governance processes in various segments of the policy cycle.

As the BEST process has revealed benchmarking is not a panacea for performance deficits in the transport sector, let alone for the achievement of sustainable transport. Nevertheless it certainly deserves to be explored further, both as a practical tool to enhance collaboration between transport organisations and as a means for researching the intricate links between knowledge generation, practice and policy making in the transport sector.

Acknowledgements

The authors wish to thank the European Commission's Directorate General for Energy and Transport, which funded the BEST project. All participants in the BEST and BOB projects are thanked for their invaluable contributions, which made writing this paper possible. The authors remain solely responsible for all interpretations and remaining errors in the paper.

References

CEC, 2002a. Structural indicators. Communication from the Commission. COM (2002) 551 final. Commission of the European Communities. Brussels, 16.10.2002

CEC, 2002b. Proposal for a regulation of the European parliament and of the Council on the granting of Community financial assistance to improve the environmental performance of the freight transport system. COM (2002) 54 final2002/0038 (COD). Brussels

CEC, 2001a. European transport policy for 2010: time to decide. White Paper. COM(2001) 370. Commission of the European Communities. Brussels

CEC, 2001b. A Sustainable Europe for a Better World: A European Union Strategy for Sustainable Development (Commission's proposal to the Gothenburg European Council). Communication from the Commission. COM (2001) 264 final Brussels

CEC, 1996. Benchmarking the Competitiveness of European Industry (COM (96) 463 DG Enterprise), Bulletin EU 10-1996, point 1.3.69.

Council (Transport), 1999. Council Strategy on the integration of environment and sustainable development into the transport policy submitted by the "Transport" council to the European Council of Helsinki. Trans 197/Env 335: Brussels

Davies, P., Boruch, R., 2001. The Campbell Collaboration: Does for Public Policy What Cochrane Does for Health. British Medical Journal, 323 (7308) 294-295

Deiss, R., 2000. Benchmarking European Transport. pp. 35-81 in: ECMT. Transport Benchmarking: Methodologies, Applications and Data Needs. Proceedings of the Paris Conference, November 1999. OECD Studies, 1 (9) OECD, Paris.

ECMT, 2000a. Transport Benchmarking: Methodologies, Applications and Data Needs. Proceedings of the Paris Conference, November 1999. OECD Studies, 1 (9) OECD, Organisation for Economic Co-operation and Development, Paris.

ECMT, 2000b. Sustainable Transport Policies. European Conference of Ministers of Transport. Paris

ECN/COWI, 2003. International CO2 policy benchmark for the road transport sector. Results of a pilot study.

EEA, 2000. Are we moving in the right direction? Indicators on transport and environment integration in the EU. TERM 2000. Environmental issues series No 12. European Environment Agency, Copenhagen.

Fearnley, N; Gordon, L, de Vlieger, J.J., 2002. Benchmarking transport policy: the use of benchmarking in effectively developing and implementing transport policy. Presented at European Transport Conference, 9-11 September 2002, Homerton College, Cambridge

Greene, D.L., 2001. Sustainable Transportation. pp. 15335-39 In: Baltes, Paul B. & Smelser, Neil J. (eds.) The International Encyclopedia of the Social & Behavioral Sciences. Elsevier Science Ltd, Oxford

Gudmundsson, H., 2003 Benchmarking & European Sustainable Transport Policies. World Transport Policy & Practice, 9 (1) 24-31.

Keegan, R. (ed.), 1998. Benchmarking Facts. A European Perspective. The European Company Benchmarking Forum, Dublin.

Lyons, W.M., 2002. The U.S. Framework for Transportation The U.S. Framework for Transportation Policy: The Role of Benchmarking. Presentation to BEST Conference 5: Benchmarking Transport Policy, Brussels June 11, 2002

May, A., 2002. Developing Sustainable Urban Land Use – Transport Strategies: A Possible Role for Benchmarking. BEST conference 'The application of benchmarking in transport policy', Brussels, June 10-11, 2002.

Oakland, J., 2003. Benchmarking to Achieve Organisational Excellence' Keynote speech at the Final BEST conference, Brussels, March 10-11, 2003.

OGM (ed.), 2003. The Role of Benchmarking in Facilitating the Successful Implementation of European Sustainable Transport Strategies: Final Recommendations to the European Commission, resulting from the BEST Thematic Network. Deliverable 12, BEST, Contract N°2000-TN.10011. OGM, Bruxelles.

Yasin, M.M., 2002. The theory and practice of benchmarking: then and now. Benchmarking: An International Journal, 9 (3) 217-243.

Wallis, I., 2001. Bus cost benchmarking 'down under': Driving efficiency improvements, Paper presented at the 4th BEST conference, Brussels, October 4-5 2001

Wyatt, A., 2002. International Comparisons in Policy Making. Best Conference 5: Benchmarking Transport Policy. Brussels, 10-11 June 2002

Appendix. Acronyms and abbreviations

BEST:	Benchmarking European Sustainable Transport.
B.E.S.T:	(‘Scandinavian project’) = Benchmarking European Service of public Transport. Nordic/European project comparing public attitudes to public transport service in 9 cities
BOB:	Benchmarking of Benchmarking. Three BOB Pilots conducted in parallel to the BEST network: 1) Passenger rail (performance and institutions), 2) Airport accessibility, and 3) Professional Road Transport Safety
CfiT:	Commission for Integrated Transport, UK.
Citizens’ Network:	EU sponsored Benchmarking Initiative at local level (40 cities/regions)
COMET/ NOVA:	Private ’Benchmarking clubs’ for Metro companies in 9+7 cities worldwide
EQUIP	Benchmarking quality in urban passenger transport. EU funded Research project
ESC	European Shippers Council
FTA	Freight Transport Association
IARO	International Air Rail Organisation
IATA	International Air Transport Association
ISOTOPE	Improved Structure and Organisation for Transport Operations of Passengers in Europe. EU funded Research Project.
NATCYP	National Cycling Policy Benchmarking Program
QUATTRO	Quality Approach in Tendering urban public TRansport Operations. EU funded Research project