



Institut für Volkswirtschaftslehre

Universität Augsburg

## Volkswirtschaftliche Diskussionsreihe

**Benchmarking in the Context of  
National Innovation Systems:  
Purpose and Pitfalls**

**Markus Balzat**

**Beitrag Nr.238, Mai 2003**

# **Benchmarking in the Context of National Innovation Systems: Purpose and Pitfalls**

**Markus Balzat**

University of Augsburg, Institute for Economics, Chair of Economics V  
Universitätsstraße 16, D - 86 135 Augsburg,  
E-Mail: markus.balzat@wiwi.uni-augsburg.de  
Phone: +49 821 598 4174

**May 2003**

## **Abstract**

Recent contributions to the literature of national innovation systems (NIS) reflect growing research interests in cross-country comparisons, in particular so in performance comparisons across national innovation systems. Although such comparisons differ significantly in terms of methodological aspects, most of them are labeled as benchmarking studies. But not only the use of this technical term can be problematical; the very task of comparing NIS is now under debate as well. It is the purpose of this paper to discuss the usefulness of benchmarking studies carried out within the framework of national innovation systems (NIS benchmarking).

*Keywords: innovation, national innovation systems, performance measurement  
(JEL Classification: O10, O30, P51)*

# Contents

<b>1. Introduction</b>	<b>4</b>
<b>2. On benchmarking and its way in the NIS approach</b>	<b>6</b>
2.1 What does benchmarking mean in different contexts?	6
2.2 Why did benchmarking enter the literature on NIS?	13
<b>3. Discussing the significance of benchmarking in the context of NIS</b>	<b>16</b>
3.1 The purpose of NIS benchmarking	16
3.2 Potential pitfalls of NIS benchmarking	19
<b>4. Conclusions</b>	<b>23</b>
<b>5. References</b>	<b>25</b>

## 1. Introduction<sup>1</sup>

The national innovation systems approach was introduced in the late 1980s (see Freeman (1987), Dosi et al. (1988)) and further elaborated in the years thereafter (see Lundvall (1992), Nelson (1993), Edquist (1997)). A national innovation system as a historically grown subsystem of the entire national economy consists of those organizations and institutions that play a major role in the innovative activity in a country. In the NIS approach, interaction on the organizational level as well as the interplay between organizations and institutions are given central interest.

The NIS approach has been applied to reveal the structure of and the main actors being involved in innovation processes in a couple of highly industrialized countries as well as in emerging countries.<sup>2</sup>

The systemic approach to innovation gives a realistic view of innovation processes and can hence serve as a suitable framework to derive innovation and technology policy measures. In fact, the approach is now widely used in the economics of innovation literature when focused on the determinants and the structure of technological change. In addition, the NIS concept is increasingly used by international organizations as an analytical framework and it enjoys growing acceptance among policymakers around the globe.

Together with the spread of the systemic concept of innovation, shifting research interests in its applications are observable. Early NIS studies put emphasis on the role of nation-specific factors spurring technical change and on nation-specific innovation patterns resulting from historical, political and cultural developments. On the contrary, recent trends in the literature show that international comparisons are now in the center of attention. These so-called "benchmarking studies" aim at identifying "best practice" policies and/or "best

---

<sup>1</sup> This paper has been presented at the 3<sup>rd</sup> European Meeting on Applied Evolutionary Economics (EMAE) held at the University of Augsburg, April 10-12, 2003. Valuable remarks by Jackie Krafft are deeply acknowledged. Furthermore, helpful comments on an earlier version of this paper by Arnold Wentzel are gratefully appreciated.

<sup>2</sup> See Nelson (1993) for a collection of these early NIS studies.

practice behavior" and then at deriving clear-cut and easy to implement policy measures. A main reason for this is that "[p]ublic policy is increasingly concerned about promoting innovation in order to stimulate economic growth"<sup>3</sup>, as for instance the Lisbon Summit of the European Union held in March 2000 underscored.

It follows from these points that we currently observe an intended convergence of two conflicting streams, namely a systemic perception of innovation processes with strongly country-specific features on the one hand and aims to obtain clear-cut policy advice by means of benchmarking exercises on the other hand.

Motivated by this issue, it is the purpose of this paper to discuss the usefulness of benchmarking studies when applied to the framework of national innovation systems. The contents of this paper are structured as follows: Based on a critical clarification of the term 'benchmarking', the following section explains why benchmarking methods became part of the analytical framework of national innovation systems. Section 3 then discusses the potential advantages and disadvantages of benchmarking studies in the context of NIS while the conclusions from this debate are presented in section 4.

---

<sup>3</sup> Kleinknecht (2000), p. 169.

## 2. On benchmarking and its way in the NIS approach

### 2.1 What does benchmarking mean in different contexts?

For at least three decades<sup>4</sup> now, the term 'benchmarking' is widely used in the private business sector by decision-makers and in the management literature. Also, there is a large body of literature in economics that deals with benchmarking exercises applied to public sector organizations. It is therefore not surprising that 'benchmarking' may have different connotations in different contexts. To come up with a uniformly valid definition of this technical term cannot be a rational objective. Instead, it appears helpful to elaborate some basic facets of 'benchmarking' together with differences between them.

In essence, benchmarking involves a performance comparison between various units of analysis that have similar goals or missions. The performance evaluation entails the gathering, utilization, and interpretation of relevant data. It is thus sought to identify the very unit which shows the best performance and which defines a 'benchmark' or 'best practice'. A benchmark sets a reference point "against which performance can be measured or assessed"<sup>5</sup>. Benchmarks can relate to a single indicator but also to a combination of various performance indicators. But it is a common feature of benchmarking studies that they are grounded on purely empirical comparisons, which leads to *relative* performance measures. Put differently, the identified reference point reflects the best *observed* performance, not to a (theoretically) optimal performance measure.<sup>6</sup>

As the corresponding literature demonstrates, benchmarking techniques have so far been applied to a large variety of criteria. In the business sector and also in the public sector, benchmarking studies attempt to capture and compare the

---

<sup>4</sup> One of the first and undeniably one of the most prominent benchmarking tests in the private business sector has been set off at Xerox Corp. in the USA in the year 1979 (see Camp (1989)).

<sup>5</sup> European Commission (2002), p. ix.

<sup>6</sup> Benchmarking thus differs from the kind of comparisons that are made in mainstream economic theory. There, observed market outcomes are compared with theoretically optimal situations, outcomes or processes, so that an *absolute* measure of performance is referred to as a yardstick.

organization and above all the efficiency of value-creating processes.<sup>7</sup> Generally speaking, efficiency is defined as the ratio of one or various output variable(s) divided by one or several input(s).

But typically, corporate benchmarking processes do not stop once the search for best practice is completed. On the basis of the gained results, subsequent phases involve the search for and critical reflection of inefficiencies as well as the modification of existing routines with the aim to increase efficiency (see chart 1 below).

So overall, by initiating learning processes<sup>8</sup> through careful comparisons, benchmarking is meant to pave the way for adjustments that lead to improved performance. However, these latter phases are not made explicit in the following, frequently cited definition by Camp (1989): "Benchmarking is the continuous process of measuring products, services and practices against [...] competitors". While the goal to reach best observed performance is already an ambitious one, some authors even go one step further and argue that it should even be the aim of any benchmarked unit to outperform existing best practice.<sup>9</sup>

But even though the just mentioned elements constitute the core of any benchmarking exercise, corporate benchmarking studies may differ in various respects.

Apart from differences in length and complexity, a crucial distinction exists between internal and external benchmarking, whereas the latter category can be classified further into competitive and functional benchmarking.<sup>10</sup> Functional benchmarking means that the composition of a benchmarking study's sample is

---

<sup>7</sup> Further examples are product design, product quality (referring to consumer taste and demand), and the rate of return of financial products like investment funds or stocks. As this list of examples illustrates, benchmarking always implies that efficiency is measured, either directly or implicitly.

<sup>8</sup> It has been emphasized by various authors in the literature that the triggering off of learning processes is a key aspect of benchmarking studies. Zairi and Leonard ((1994), p. 26), for instance, claim: "[B]enchmarking is very much an opportunity for an organization to learn from the experience of others". In the case of policy evaluation (or "policy benchmarking"), the OECD (1998) defines "best policy practice as a learning tool rather than a normative concept" (OECD (1998), p. 29).

<sup>9</sup> See Zairi and Leonard (1994), p. 26. In the literature on NIS benchmarking, though, less ambitious targets are set. For instance, the EU views its benchmarking project of national research policies as one of many means to achieve economic convergence among its member states, a target which is anchored in its founding treaties of the year 1957 already.

<sup>10</sup> See Baurle (1996), p. 9.

drawn from a certain process or a certain link in the value chain. This implies that the units of analysis do not necessarily have to operate in the same sector of the economy. It is this openness of benchmarking towards processes observable in various industries that makes it go beyond the management tool of competitive analysis.<sup>11</sup> An example of a functional benchmarking study is the famous Xerox case. Consequently, the definition "benchmarking is the search for industry's best practices"<sup>12</sup> can merely be of limited validity.

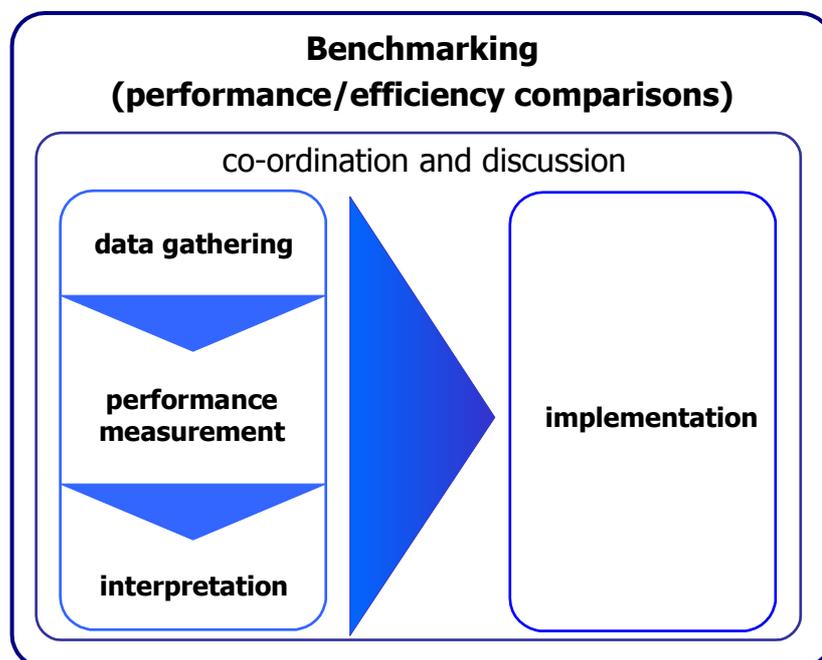


Chart 1: The central elements of benchmarking processes.<sup>13</sup>

As far as so-called benchmarking studies in the context of innovation systems are concerned, these differ in various respects from corporate benchmarking exercises.

<sup>11</sup> See Zairi and Leonard (1994), p. 56.

<sup>12</sup> Camp (1989), p. xi.

<sup>13</sup> Source: Own illustration. Alternative versions of a stylized benchmarking process in the context of innovation systems can be found in Lundvall and Tomlinson ((2002), p. 220) as well as in Barré ((2001), p. 265).

Primarily, NIS benchmarking studies do not always concentrate on efficiency as defined above. That is because the relation between inputs and outputs is a highly complex one in the case of innovation systems. Some authors even argue that the consideration of input-output relationships necessarily reflects a linear view of the innovation process that is contradictory to a systemic perception of innovation. Therefore, systemic benchmarking may well focus on the determinants of innovative activity alone. These include predominantly financial efforts devoted to research and development, but also more complicated issues as the evaluation of policy measures. Equally, some NIS comparisons have only concentrated on the outcomes of innovative action which can be measured in terms of patents or the share of new products in total products supplied.

Yet, as recent research work has shown, the concept of efficiency measurement has also been introduced to the NIS approach, even though dissimilar definitions of efficiency exist. The most common definition related proxies of innovative output to proxy variables of innovative input. Deviating from this, it has been claimed that "it is economic rather than technological performance which matters when measuring the efficiency of an innovation system"<sup>14</sup>. Efficiency defined in such a way means that innovative efforts are directly related to measures of real economic performance like GDP (per capita) growth. However, it needs to be pointed out here that Paasi's definition of efficiency is problematic for numerous reasons: It has been clarified by leading scholars in the field that the NIS concept is not about the impact of innovations on economic growth but rather on the determinants of innovative action and of technical advance in a country. Since even the linkage between R&D spending and innovative success is far from being fully understood or easily measurable, it appears thus to be far-fetched to establish a direct tie between innovative input and real economic performance. If the efficiency of an NIS is to be a meaningful expression at all, it should rather reflect the relationship between innovative efforts and innovative success. Alternatively, a measure of efficiency of innovation systems could reflect (1) the fitting between technological and industrial structures

---

<sup>14</sup> Paasi (1998), p. 221.

of an economy and the institutional framework<sup>15</sup>, or (2) the linkages between the building blocks of an innovation system like those between scientific research and the private business sector.

NIS benchmarking differs further from firm-level benchmarking because studies belonging to the former category generally do not go beyond measurement and interpretation of gained empirical results. Analysts can only make policy recommendations, and unfortunately NIS benchmarking studies rarely provide any information concerning the implementation of performance-enhancing adjustments and their success. This is somehow natural when the experts in charge of the carrying out of the comparative study are not assigned with the realization of performance-improving modifications of the systems analyzed, i.e. when the 'benchmarking' exercise comes to a halt once the empirical analysis is completed. Put differently, there can easily be a discrepancy between economic scholars and policymakers, or between analysis and corresponding implementation.

So, in benchmarking studies of economic (sub-)systems, the group of responsible actors - including economists, other experts, and policymakers - is far more heterogeneous than in corporate benchmarking processes.<sup>16</sup> This aspect is likely to have a negative impact on the sustainability and on the pace of benchmarking processes on the national level.

A further distinction between corporate benchmarking and benchmarking of innovation systems emerges when each of these two types of benchmarking studies are confronted with the expression 'performance measurement'. Benchmarking and performance measurement are not the same thing on the level of business firms. While the latter is argued to be an internally oriented tool focusing primarily on effectiveness and less on competitiveness, benchmarking is for the most part externally oriented ensuring that performance is

---

<sup>15</sup> See Johnson (1997) or Edquist and Johnson (1997) on this aspect.

<sup>16</sup> This has also been pointed out by the European Commission who claims that national benchmarking processes "involve the active and continuing commitment of many actors, notably policy makers, experts, national statistical services, and the European Commission services" (European Commission (2001), p. 3).

translated into competitiveness with special emphasis on the wants and viewpoints of (potential) customers.<sup>17</sup>

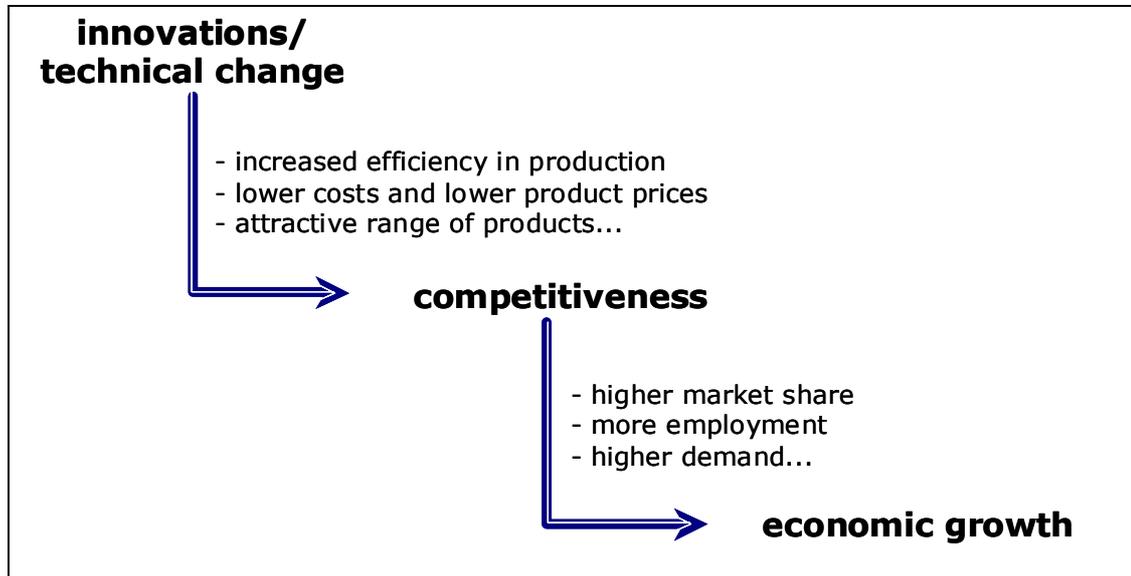


Chart 2: The significance of innovations and technical change for economic growth.<sup>18</sup>

In contrast, performance measurement and benchmarking applied to innovative activity can go hand in hand. It follows from the benchmarking of innovative efforts and outcomes of nations that competitiveness is also dealt with - even though this is not necessarily done in an explicit manner. That is because on the one hand, innovations determine the competitiveness of nations which in turn fosters economic growth (see chart 2 above for a simplified illustration of the linkage between innovations and economic growth). Hence, it can be claimed that a nation's innovation system is the engine of growth of the entire economy. On the other hand, as has been argued in the literature on national innovation systems, the NIS approach has not been developed as a means to study the relationship between innovations and growth. Rather, the NIS approach is a means to focus on the patterns and the determinants of innovation processes from the perspective of nation-states.

<sup>17</sup> See Zairi and Leonard (1994), chapter 10.

<sup>18</sup> Source: Own illustration, inspired by McKinsey ((2002), p. 12).

All these points show that the use of the term 'benchmarking' has different connotations in economics and in the corporate world. It seems problematic that 'benchmarking' has recently been transferred in an uncritical way from the management literature to economics, and especially so to the approach of national innovation systems. For reasons of clarity, it might sometimes be better to omit the term 'benchmarking' in the context of NIS, at least in those cases where the studies do not comprise a continuous co-operation between analysts and decision-makers. Alternatively, expressions like 'the empirical analysis of...' or 'the empirical treatment of...' could be used there instead. A further argument for the omission of the technical term 'benchmarking' in the context of innovation systems is that in corporate benchmarking studies the focus is often just on details of the best practice routines while further differences amongst the units of analysis are abstracted from.<sup>19</sup> However, it is not desirable to proceed similarly in system-level comparisons since this would turn the NIS concept upside down.<sup>20</sup>

To sum up, the principal means of benchmarking tests include the collection of relevant data, the selection and application of techniques required to analyze the data, and the implementation of measures that (are expected to) contribute to efficiency enhancements. This latter stage can either entail the modification of existing or the introduction of new processes and structures. These means of benchmarking studies need to be distinguished from their purpose<sup>21</sup>, which is the improvement of efficiency or other measures of performance on the one hand and the initiation of learning processes on the other hand.

---

<sup>19</sup> See Heindl (1999), p. 7.

<sup>20</sup> See Balzat (2002).

<sup>21</sup> This has also been accentuated by Heindl (1999), pp. 8-9.

## 2.2 Why did benchmarking enter the literature on NIS?

In the previous subsection where definitions of the term 'benchmarking' have been discussed and the fundamentals of benchmarking studies have been presented, a broader perspective has deliberately been chosen: It encompassed both business studies that are normally initiated by managers and studies of innovation systems carried out by economic scholars (and being partially initiated by policymakers). In that way, it became evident that on the one hand these types benchmarking studies share basic similarities: the measurement, comparison and enhancement of performance. On the other hand, they differ significantly in their set-up. Furthermore, the subsection above already foreshadowed difficulties that may arise when benchmarking studies are carried out beyond a sample of business firms that can be compared with each other in a fairly uncomplicated manner.

So, before discussing strengths and weaknesses of benchmarking exercises of innovation systems (see the subsequent section), it needs to be explained why benchmarking became part of the systemic approach to innovation, and thus of the economics of innovation and technical change literature.

Political agreements triggered the carrying out of national benchmarking studies. Most importantly, the European Union urged its Commission to work together with the EU-15 countries in order to "develop indicators and a methodology for the benchmarking of national research policies"<sup>22</sup>. This plan of work, being a result of the Lisbon Summit of the European Union<sup>23</sup> held in March 2000, was agreed upon by the Council of Research Ministers of the EU on June 15, 2000.

---

<sup>22</sup> European Commission (2000), p. 3.

<sup>23</sup> At this summit, the EU defined various ambitious goals to be realized within the first decade of the 21st century. For instance, these include the aim "to become the most competitive and dynamic knowledge-based economy in the world. To achieve this goal, [...] the European Council decided to introduce a new open method of co-ordination, of which benchmarking will be one of the key tools" (European Commission (2001), p. 3).

One facet of this larger project has been a cross-country comparison of the relations between the private business sector and scientific research bodies. This work has been done by the Austrian Federal Ministry of Economy and Labor in collaboration with the European Commission and its Directorate General for Entrepreneurship.<sup>24</sup>

In addition to this, the OECD has been working for ten years or so on a broad project with the focus on the employment situation in various of its member countries.<sup>25</sup> One fragment of the so-called "OECD Jobs Study" was the finding out of best practice policies related to technology and innovation.

Besides these international benchmarking projects, there are also national initiatives resulting in benchmarking studies of economic activity. In Germany, for instance, trade unions, employers' associations and the federal government formed an "Alliance for Jobs, Education and Competitiveness" in the year 1998.<sup>26,27</sup> To gain insights into the strengths and weaknesses of the German labor market and into the determinants of its performance, the members of this pact decided to set up a task force (named "Benchmarking Group") with the mission to "gather and analyze all relevant data with regard to the performance of the German labor market"<sup>28</sup>.

Consequently, the fulfillment of these political targets required a combination of benchmarking techniques with innovative activities including innovation policy measures. At the same time, economic research in the field of innovation highlighted the significance of a systemic approach to innovation and thus rejected the linear model of innovation processes. To put all that into a nutshell, it is unsurprising that those economic scholars who were assigned with the realization

---

<sup>24</sup> See Polt et al. (2001).

<sup>25</sup> As it is made explicit, this OECD project is related to the results of the G7 summit held in Lille in the year 1996 (see OECD (1998), p.3).

<sup>26</sup> For the sake of abbreviation, Germans call this agreement usually just "Bündnis für Arbeit" (or, translated into English language: "Alliance for Jobs").

<sup>27</sup> The German "Alliance for Jobs" has been criticized sharply ever since, though. That is because its partners seemed to be unwilling to come to terms with each other. Therefore, it has become impossible to make much progress, i.e. to agree upon and bring about urgently needed structural reforms of the German labor market. As a consequence of this standstill, the alliance is currently prone to break apart.

<sup>28</sup> Eichhorst et al. (2001), p. 1.

of the just mentioned benchmarking tasks referred to the concept of national innovation systems as the framework for the analysis.<sup>29</sup>

Apart from the portrayed political background, research interests of scholars in the economics of innovation literature can serve as an explanation for the developed linkage between benchmarking and the NIS approach. Even though it is widely accepted that the systemic concept of innovation processes is able to give a realistic picture of innovation patterns,<sup>30</sup> it still has some shortcomings. Above all, the concept does not provide an apparatus to carry out system comparisons in a formalized way. Accordingly, early studies of national innovation systems focused on the description of historically grown structures of the very system under consideration.<sup>31</sup> Part of the research work on national innovation systems can hence be seen as an attempt to broaden the NIS concept by applying it to system (performance) comparisons.<sup>32</sup> For instance, Niosi (2002) suggests that the "[e]fficiency and effectiveness [of innovation systems] have to be seized [...] through careful empirical analysis and comparison of [organizations] with similar missions"<sup>33</sup>. Doing this, it is also sought to contribute to a better understanding of national disparities in terms of innovativeness, economic growth and employment.

---

<sup>29</sup> See e.g. European Commission (2002).

<sup>30</sup> This is true not only for the concept of *national* innovation systems, but also for related - but otherwise defined - versions of innovation systems like regional, sectoral, or technological systems.

<sup>31</sup> See the broad collection of NIS studies in Nelson (1993).

<sup>32</sup> See Balzat (2002), pp. 22-28.

<sup>33</sup> Niosi (2002), p. 296.

### **3. Discussing the significance of benchmarking in the context of NIS**

#### **3.1 The purpose of NIS benchmarking**

First of all, it is important to note that NIS benchmarking must *not* be understood as a normative concept. This follows from the very fundamentals of the NIS approach<sup>34</sup>, in particular from the consideration of nation-specific institutional set-ups, and from the crucial role ascribed to them in limiting and enhancing innovative activity. Consequently, NIS benchmarking rather needs to be understood as an information tool which can be useful for scholars in the field and for policymakers alike. This is certainly a key function of benchmarking in the context of innovation systems. So despite the problems that are brought about by the concept of 'best practice' (which will be presented in more detail in the following subsection), the results gained in benchmarking studies "can be used to guide policy learning, and as an input in discussions among [policymakers]"<sup>35</sup>. This implies that the outcomes of empirical studies can assist policymakers in preparing and designing measures to improve the performance of innovation systems. Correspondingly, Tsipouri (2001) claims that based on the revealed efficiency differences, policymakers should "learn and develop their own intuition and consensus"<sup>36</sup>.

This usefulness of systemic benchmarking stems mainly from the revealing of strengths and weaknesses of the systems analyzed as a result of international comparisons. To meet the objective of identifying national strengths and weaknesses, it is a prerequisite to have access to relevant and reliable data. As there can be various obstacles to this preliminary step, researchers need to filter out comparable data, harmonize existing data (e.g. adjust it for structural disparities) or even collect new empirical data. All three tasks are positive side-effects

---

<sup>34</sup> For a detailed presentation of the fundamental principles of the NIS approach, see Edquist (1997), pp. 15-29.

<sup>35</sup> Polt et al. (2001), p. 257. A similar opinion has been put forward by Barré (2001), p. 265.

<sup>36</sup> Tsipouri (2001), p. 300.

of NIS benchmarking.<sup>37</sup> In this way, benchmarking goes beyond the analysis of already existing data and contributes to a better understanding of innovation processes.

In order to obtain meaningful results, it is vital to use various indicators rather than a single indicator. It has repeatedly been emphasized that this aspect exerts a strong influence on the quality of benchmarking exercises since "any relevant concept of productivity cannot be characterized by the ratio of one output variable amongst many"<sup>38</sup>. Moreover, one has to be careful when it comes to the selection of this set of indicators since this has a large impact on the explanatory value of the study.<sup>39</sup>

Due to the identification of strengths and weaknesses of the systems under consideration, NIS benchmarking is one way to bring about processes of learning by comparing which have been defined as follows in the context of innovation systems: "Learning by comparing includes the international comparison of indicators, the use of simple statistical techniques to map causalities and the qualitative comparison of systems. Used in such a context, benchmarking has the potential to significantly improve the effectiveness of institutions within national innovation systems by creating a mutual learning environment."<sup>40</sup>

Recent empirical studies focused on the stock of existing knowledge as well as on the building up of new knowledge through learning processes because both are decisive inputs to an NIS. However, there are "various kinds of channels for exchanging [...] knowledge"<sup>41</sup> in innovation systems, either within one particular component or between two or more components. An example of the case where knowledge is exchanged within one element of the system are inter-firm collaborations. Industry-science relations may illustrate the case where knowl-

---

<sup>37</sup> This is precisely what happened in the aftermath of the Lisbon Summit of the EU where it was agreed upon to initiate a EU-wide benchmarking project of the success of national research policies (see European Commission (2001), pp. 6-8).

<sup>38</sup> Barré (2001), p. 259. Similar arguments have been made by Eichhorst et al. ((2001), p. 4) or Lundvall and Tomlinson ((2002), p. 217).

<sup>39</sup> The selection of the "right" set of indicators is of course based on subjective decisions. These decisions are hence always questionable, especially when they are not clearly justified.

<sup>40</sup> Lundvall and Tomlinson (2002), p. 225.

<sup>41</sup> Polt et al. (2001), p. 249.

edge is exchanged between two building blocks of an innovation system. But these different mechanisms for transferring or building up knowledge vary in strength and efficiency. International comparisons that attempt to measure these knowledge channels are a means to reveal performance differences regarding the transfer or the creation of knowledge in different systems.

But in any event, the sample of countries has to be composed carefully. That is because it is first a difficult exercise to study innovation processes that are organized in dissimilar ways in different countries, and second because the derivation of technology policy measures is even more difficult. Therefore, the value of NIS benchmarking studies is likely to rise if the sample of countries is to some extent homogeneous, for instance with respect to social values, political aims, and to the level of economic development.<sup>42</sup> Likewise, empirical studies have shown that highly advanced countries face similar challenges that result from comparable demographic developments, as well as from technological and structural change. Since adapting to these challenges is a not an easy task, there is "tremendous scope for mutual learning"<sup>43</sup>.

Accordingly, benchmarking constitutes a learning tool in two further respects that are both part of the learning-by-comparing process. First, it opens up the opportunity to learn from own experience, and second from the experience of others. This is especially true when the analysis of foreign innovation systems is based on sound interpretations while nation-specific contexts are taken into account. Then, systemic comparisons can lead to context-adapted modifications of innovation systems. Put differently, when the gained empirical insights of such comparisons do not result in a simple process of copying best practices, benchmarking may indeed become a reasonable analytical tool.<sup>44</sup>

---

<sup>42</sup> This has also been underlined by Eichhorst et al. (2001), p. 2.

<sup>43</sup> OECD (1998), p. 3. Evidently, these challenges have been posing serious problems to the some of the larger member countries of the European Union, as high levels of unemployment and low economic growth rates illustrate.

<sup>44</sup> It is a typical feature of benchmarking that it does not lead to simple replications of best practice. Instead, benchmarking is meant to lead to performance-enhancing changes that consider carefully existing structures and routines. Because of this, benchmarking has even been labeled as a "creativity technique" (see Bäurle (1996), p. 13).

All the preceding remarks on NIS benchmarking show that it is not the purpose of international comparisons of innovative activity to derive water-proof (i.e. failure-resistant) policy recommendations.<sup>45</sup> Rather, it is intended to present empirical evidence of innovative activity in the countries analyzed, especially of the strengths and weaknesses in the organization of innovation processes from the perspective of nation-states. In doing so, learning processes of various kinds can be initiated. These learning processes in turn are a sensible means to derive and implement adequate policy measures.

### 3.2 Potential pitfalls of NIS benchmarking

Even though the subsection above presented various arguments in favor of the usefulness of NIS benchmarking, the validity of most of these arguments was closely tied with the fulfillment of certain quality criteria concerning the set-up of the benchmarking study and the interpretation of the empirical results gained.

But still, the idea of carrying out cross-country comparisons of NIS can be questioned. Some of the central arguments against such comparisons will therefore be explained in order to provide the reader with a more extensive discussion of this issue.

First of all, the study of innovative activity per se can be criticized as being fraught with fundamental problems. These stem mostly from the limited explanatory value of many indicators of innovative action that are commonly used in the literature.<sup>46</sup> The limited explanatory value of these indicators can be due to at least three aspects: First, one has to be aware of the fact that all these indicators can only be proxies of innovative action. This implies that parts of the relevant phenomena always remain below the surface of empirical analysis

---

<sup>45</sup> See Eichhorst et al. (2001), p.3.

<sup>46</sup> A helpful discussion of various indicators of innovative activity can be found in Kleinknecht (2000).

while other parts may be distorted by and strongly correlate with activities beyond innovative action. Second, it has been underscored that "there can be big variations in the quality and usability of data series across countries"<sup>47</sup>. This means that innovation indicators collected in various countries cannot be assumed to be harmonized in terms of the underlying measurement techniques and in terms of the cut-off of the samples included. Unfortunately, these problems can even arise if one makes use of data collected by large international organizations like the OECD or the EU. Third, even if the recommendations that have been given earlier<sup>48</sup> are followed, the selection of indicators can only be improved but it can never be perfect. That is because it cannot be taken for granted that "each country will give the same priority to different performance indicators"<sup>49</sup>.

In addition to these possible drawbacks resulting from the availability and quality of data, the measurement technique itself is an issue of debate. It has been mentioned earlier that some benchmarking exercises rest upon inputs and/or outputs, i.e. the determinants and/or the outcomes of innovative activity alone. The findings of such empirical studies can certainly provide valuable information. However, neither innovative inputs nor innovative outputs can be captured exclusively because there are interdependencies at all stages of innovative activity. Therefore, mere input-output studies cannot disentangle the strength and the importance of linkages between the various actors involved in innovation processes, unless the output variables are as such are sought to reflect such linkages. But, as has been argued, it is precisely these linkages between the components of a national innovation system that have a major impact on its performance.<sup>50</sup>

Furthermore, it is a complex task to take into account a country's institutions, let alone consider or even measure international differences in the institutional environments in NIS benchmarking studies. Implicitly though, revealed per-

---

<sup>47</sup> Smith (2001), p. 275.

<sup>48</sup> See p. 15 above.

<sup>49</sup> Lundvall and Tomlinson (2002), p. 214.

<sup>50</sup> See e.g. Lundvall and Tomlinson (2002), p. 216.

formance differences in the measured innovations indicators can be traced back to disparate institutional arrangements like differing incentive structures. Obviously, this way of dealing with institutions still makes it necessary to have a closer look at the institutional frameworks in the countries analyzed.

Beyond these issues, NIS benchmarking has been severely criticized because of its basic idea of searching for best practice. It has been argued repeatedly that the concept of best practice is inconsistent with the heterogeneity of national systems of innovation which has been observed in earlier studies. The line of reasoning here is that due to the dissimilarity of innovation systems with regard to their institutional as well as organizational structures, identified benchmarks have very limited relevance. For instance, it is claimed that "in diverse environments [...], there is no single best practice that leads to superior, quantifiable performance"<sup>51</sup> and that "[d]iversity implies that there is no single best way to do anything [...]" and so any overarching single 'guide' indicator for performance ought to be treated with suspicion<sup>52</sup>. It could be implied from such critical statements that NIS benchmarking is meaningless because different systems are not comparable. Obviously, corporate benchmarking is often confronted with the same objection when it is argued that different business firms cannot be compared with each other. The argument that the selected units of analysis are too dissimilar and that then the consequent identification of best practices would be of no value to the other units can 'knock out' any intended benchmarking study.<sup>53</sup>

However, viewed from a different perspective, it needs to be questioned if there really is room for learning processes (especially for learning by comparing and learning from others' experience) in a sample consisting of almost similar entities. It is a misinterpretation if the identification of best practice is understood as a first step which leads to the imitation of the benchmark. The idea of

---

<sup>51</sup> Tsipouri (2001), p. 300.

<sup>52</sup> Smith (2001), p. 276.

<sup>53</sup> See Kreuz (1995), p. 27. A note on corporate benchmarking: If researchers and decision-makers had stuck strictly to the argument from above, not a single functional benchmarking study would have been done on the firm-level so far because no firm is identical to another.

best practice needs to be understood in a less rigorous way. In the context of NIS benchmarking, it has been well explained that the "search for best practice is based on the identification of policies that 'work' in a specific country, and on an understanding of the general principles that can be derived from the observed experience. [...] [T]he search for best practice evolves [...] towards the prescription of 'context-related' good practices. The notion of best practice must be understood as a learning tool, rather than a normative concept"<sup>54</sup>.

A final point of criticism exposed here concerns the implementation of policy recommendations. It results from the theoretical roots of the systemic approaches to innovation, namely an evolutionary interpretation of economic change<sup>55</sup>. According to this line of theorizing, economies change permanently, they are never at rest. Hence, they do not reach a static point of equilibrium where all economic variables attain an optimal level.<sup>56</sup> With regard to practical policymaking, there are usually time lags between the identification of a suitable, innovation-enhancing policy framework (which could, for instance, result from NIS benchmarking studies) and the implementation of the same. Since in the meantime, however, economic structures will have changed, it is uncertain that the policy measures chosen will unfold their full effectiveness. Hence the combination of the dynamic nature of modern economic systems and delays in the process of policy formation and policy-making may reduce the very value of policy advice.

---

<sup>54</sup> OECD (1998), p. 21.

<sup>55</sup> See Nelson and Winter (1982) or Witt (1999). On the relation between evolutionary economic theories and the approach of innovation systems, see e.g. Saviotti (1997).

<sup>56</sup> See Metcalfe (2001), p. 561 and p. 576.

## 4. Conclusions

Applications of the concept of national innovation systems have so far contributed largely to the identification of the main determinants of the innovative success of nations and to the patterns of innovation processes on the national level. Here, special emphasis has been given to the role of nation-specific, historically grown (and thus path-dependent) institutional arrangements.

In view of the fundamental findings gained in single-country studies of innovation systems, and in view of the plausible assumptions in which the NIS approach is rooted, recent streams in the economics of innovation literature reflect growing interest in international comparisons. These are mainly initiated by policymakers or by policy-advising international organizations. On the basis of such cross-country comparisons which are often called "benchmarking studies", it is expected to obtain insights into the strengths and weaknesses of the systems analyzed. In doing so, these NIS benchmarking studies can be a useful source of information to derive innovation policy measures with the aim to improve the framework conditions of innovation systems. Further advantages of such comparisons have been outlined earlier. Most importantly, though, they stem from the improvement of mutual learning processes including learning from (others') experience.

However, as the objections raised in the discussion above show, there are various potential pitfalls when combining a systemic approach to innovation on the one hand with benchmarking techniques on the other. Already the integration of the corporate benchmarking terminology into the NIS framework is problematical and can result in misunderstandings. Correspondingly, the relevance of so-called best practice in the context of innovation systems has been a central issue of debate.

Finally, the discussion presented here it is not at all intended to discourage researchers from doing empirical analysis grounded on a systemic perception of innovation processes. Rather, the objective is to make researchers aware of

some of the principal drawbacks that may spring from various sources. Identifying and discussing these obstacles can turn out to be a decisive step towards eliminating them. The motivation to do so rests on the belief that it is still a meaningful effort to carry out cross-country comparisons in terms of the patterns and the success of innovative activity on the country-level.

## 5. References

- Balzat, M., 2002. *The Theoretical Basis and the Empirical Treatment of National Innovation Systems*. University of Augsburg (Germany), Economics Department, Discussion Paper Series, Number 232, December 2002.
- Barré, R., 2001. Sense and nonsense of S&T productivity indicators. In: *Science and Public Policy* 28(4), pp. 259-266.
- Bäurle, I., 1996. *Resourcenbasierter Ansatz und Benchmarking: zwei widersprüchliche Strömungen?*. Diskussionsbeiträge der Wirtschaftswissenschaftlichen Fakultät der Universität Ingolstadt (Germany), Number 79.
- Camp, R.C., 1989. *Benchmarking: The Search for Industry Best Practice that Lead to Superior Performance*. ASQC Quality Press, Wisconsin.
- Dosi, G., Freeman, Ch., Nelson, R.R., Silverberg, G., Soete, L. (eds.), 1988. *Technical Change and Economic Theory*. Pinter, London.
- Edquist, Ch. (ed.), 1997. *Systems of Innovation: Technologies, Institutions and Organizations*. Pinter, London.
- Edquist, Ch., Johnson, B., 1997. Institutions and Organizations in Systems of Innovation. In: Edquist, Ch. (ed.), 1997.
- Eichhorst, W., Profit, S., Thode E. (eds.), 2001. *Benchmarking Deutschland: Arbeitsmarkt und Beschäftigung, Bericht der Arbeitsgruppe Benchmarking und der Bertelsmann Stiftung*. Springer-Verlag, Berlin.
- European Commission, 2000. *Towards a European Research Area: Science, Technology and Innovation, Key Figures 2000*. European Commission, Brussels.
- European Commission, 2001. *Towards a European Research Area: Key Figures 2001, Special Edition: Indicators for benchmarking of national research policies*. European Commission, Brussels.
- European Commission, 2002. *Benchmarking national research policies: The Impact of RTD on Competitiveness and Employment (IRCE)*. Report can be downloaded form the world wide web in pdf-format. The address is: [http://www.edis.sk/ekes/kneldok/dokument/bench\\_irce\\_0802.pdf](http://www.edis.sk/ekes/kneldok/dokument/bench_irce_0802.pdf)
- Freeman, Ch., 1987. *Technology and Economic Performance: Lessons from Japan*. Pinter, London.
- Heindl, H., 1999. *Benchmarking Best Practices*. Working Paper Series of the University of Wuppertal, No. 191. Wuppertal, Germany.
- Johnson, B., 1997. Introduction to Part I. In: Edquist, Ch. (ed.), 1997.

- Kleinknecht, A., 2000. Indicators of Manufacturing and Service Innovation: Their Strengths and Weaknesses. In: Metcalfe, J.S. and Miles, I. (eds.), 2000, *Innovation Systems in the Service Economy: Measurement and Case Study Analysis*. Kluwer, Boston.
- Kreuz, W. (ed.), 1995. *Mit Benchmarking zur Weltspitze aufsteigen*. Verlag Moderne Industrie, Landsberg am Lech, Germany.
- Lundvall, B.-Å. (ed.), 1992. *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. Pinter, London.
- Lundvall, B.-Å., Tomlinson, M., 2002. International benchmarking as a policy learning tool. In: Rodrigues, M.J. (ed.), 2002, *The New Knowledge Economy in Europe: A Strategy for International Competitiveness and Social Cohesion*. Edward Elgar, Cheltenham.
- McKinsey Global Institute, 2002. *Reaching higher productivity growth in France and Germany*. Report can be downloaded in pdf-format from the world wide web. The address (as of March 31, 2003) is:  
[http://www.mckinsey.com/knowledge/mgi/reports/pdfs/europe/overview\\_english.pdf](http://www.mckinsey.com/knowledge/mgi/reports/pdfs/europe/overview_english.pdf)
- Metcalfe, J.S., 2001. Institutions and Progress. In: *Industrial and Corporate Change*, Vol. 10, pp. 561-586.
- Nelson, R.R. (ed.), 1993. *National Innovation Systems: A Comparative Analysis*. Oxford University Press, Oxford.
- Nelson, R.R., Winter, S.G., 1982. *An Evolutionary Theory of Economic Change*. Bellknap Press, London.
- Niosi, J., 2002. National systems of innovations are "x-efficient" (and x-effective) - Why some are slow learners. In: *Research Policy* 31(2), pp. 291-302.
- OECD, 1998. *Technology, Productivity and Job Creation: Best Policy Practices*. OECD, Paris.
- Paasi, M., 1998. Efficiency of innovation systems in the transition countries. In: *Economic Systems* 22(3), pp. 217-234.
- Polt, W. et al., 2001. Benchmarking industry-science relations: the role of framework conditions. In: *Science and Public Policy* 28(4), pp. 247-258.
- Saviotti, P.P., 1997. Innovation Systems and Evolutionary Theories. In: Edquist, Ch. (ed.), *Systems of Innovation: Technologies, Institutions and Organizations*. Pinter, London.
- Smith, K., 2001. Cross-country comparisons: Comparing economic performance in the presence of diversity. In: *Science and Public Policy* 28(4), pp. 267-276.
- Tsipouri, L., 2001. Can we benchmark the contribution of research and development investment to growth and competitiveness?. In: *Science and Public Policy* 28(4), pp. 295-302.

Witt, U. (ed.), 1993. *Evolutionary Economics*. Edward Elgar, Aldershot.

Zairi, M., Leonard, P., 1994. *Practical Benchmarking: The Complete Guide*.  
Chapman and Hall, London.