

Ursula Hansen & Thorsten Hennig-Thurau

National Customer Satisfaction Indices: A Critical Investigation from an Application Perspective

Professor Dr. Ursula Hansen holds a Chair of Marketing at the University of Hannover, Niedersachsen, Germany, since 1975. She is author of several books on different fields of marketing theory (e.g. retailing, green marketing, marketing history), and her work has been published in numerous leading journals on marketing science including Marketing ZFP, Die Betriebswirtschaft and the Journal of Consumer Policy. She also is CEO of the IMUG institute and member of several institutions in the area of marketing. Contact address: Department of Marketing I: Marketing & Consumer Research, University of Hanover, Koenigsworther Platz 1, 30167 Hanover, Germany, Phone (++49) 511 762 5613, Fax (++49) 511 762 5630.

Dr. Thorsten Hennig-Thurau is an senior lecturer of marketing at the University of Hannover, Niedersachsen, Germany. His reseach interests include relationship marketing, services management, and the management of higher education. His work has been published by Psychology & Marketing, Die Betriebswirtschaft and Zeitschrift für Betriebswirtschaft, among others. Contact address: Department of Marketing I: Marketing & Consumer Research, University of Hanover, Koenigsworther Platz 1, 30167 Hanover, Germany, Phone (++49) 511 762 4343, Fax (++49) 511 762 5630, Email th@muk.ifb.uni-hannover.de.

Acknowledgement: The authors wish to thank Dr. Michael Korpiun, a former research assistant at their institute, for his constructive contributions to an earlier draft of this paper.

Paper to be published in Kunst, P., Lemmink, J., Stauss, B. (eds.): Quality Management in Services, Assen/Maastricht 1999, final draft

Abstract:

The development of **National Customer Satisfaction Indices** (NCSIs) has recently become a central area of research in service management and marketing. Despite an considerable amount on scientific and societal discussion, a massive lack of metatheoretical research in the area of NCSI development and usage exists. Based on the identification of several levels of NCSI based satisfaction comparisons the authors formulate a catalogue of general criteria which function as a frame of reference for the critical analysis and evaluation of NCSIs. With the **American Customer Satisfaction Index** and the **German Customer Satisfaction Barometer**, two well-known indices are than evaluated with regard to the criteria previously worked out. Several proposals are generated corresponding to the results of the analysis, which can be used as a guideline for future improvements of NCSI methodology and usage.

Keywords: National Customer Satisfaction Indices, Service Quality, American Customer Satisfaction Index, Deutsches Kundenbarometer, Relationship Quality

National Customer Satisfaction Indices: A Critical Investigation from an Application Perspective

1 Introduction

The development of **National Customer Satisfaction Indices** (NCSIs) has recently become a central area of research in service management and marketing. The term NCSIs is applied for satisfaction measurement and monitoring procedures that differ from company-related satisfaction research insofar as they incorporate customers' satisfaction judgements referring not only to one firm, but to a large number of companies in different sectors of business. Building on this, NCSIs allow insights into the state of a **nation's customers degree of satisfaction** with service offerings. Up to now, NCSIs are established in several countries including the United States ("American Customer Satisfaction Index"/ACSI, see Fornell et al. 1996) and Germany ("Deutsches Kundenbarometer"/DKB, see Meyer & Dornach 1996). Furthermore, the implementation of NCSIs is in a planning state in numerous other countries, e.g. in Austria, Switzerland and Israel.

Unfortunately, the present academic discussion of these nationwide tools for satisfaction measurement is limited. Until now, contributions are only made by institutionally involved researchers (e.g. the developers of NCSIs), but are missing from other (formally non-involved) academics. Corresponding to this, the scientific discussion on NCSIs is predominantly of a more or less descriptive kind, while a critical and metatheoretical perspective is merely taken. This absence of constructive critics may be explained by the early phase of the lifecycle of NCSI research, but, however, is contradictory to the outstanding theoretical and particularly societal status which is attached to NCSI results. As a consequence, the need for supplementing the existing "insight" contributions by critical and constructive "outside" analysis, which can lead to further refinements of the existing methodology of NCSIs must be postulated.

Thus, in our paper we address the following issues of NCSI research by taking such an "outside" perspective. First, we identify the different addressees of NCSI and specify their interest in NCSI methodology and/or results. Second, several levels of NCSI based satisfaction comparisons applied **in praxi** by the users of NCSIs are

differentiated (cross-sectional vs. longitudinal, intrabranche vs. interbranche etc.). The results of the analysis are transformed into a catalogue of general criteria which we use in the following as a frame of reference for the critical analysis and evaluation of NCSIs. As the American ACSI and the German DKB can be seen as highly influential for the development of further NCSIs, both are chosen as objects of our detailed investigation into the construction and application of nationwide satisfaction barometers. The two indices (ACSI and DKB) are then evaluated with regard to the criteria previously worked out. Finally, the deficiencies of both barometers identified here are used to develop proposals for the editors of NCSIs for future improvements of their methods.

2 Overview: On the Functionality of NCSIs

2.1 Identifying NCSI's Addressees

Information generated by NCSIs is used by a multitude of groups which maintain informational relations among each other. Figure 1 provides an overview of central addressees (and stakeholders, respectively) of NCSIs and systematizes their inter-relationships.

- take in Figure 1 approx. here -

Consumers and the government as decision makers in matters of economic policy are "end users" of NCSI data. For the **government**, the information on customer satisfaction complements existing knowledge on the welfare state of the nation. In this way, it tries to get a genuine and radically subjective impression of the extent of need fulfillment in the consumption society. For **individual consumers**, NCSIs offer a chance to compare their own company- or branch-related satisfaction judgements with societal assessments. Beyond this, NCSI results may also be considered as additional arguments in the customer's buying process. NCSI information is gathered by independent organizations, so that it may be used by customers as a guideline when deciding between alternative products and brands. In this case, the existence of NCSI data can contribute to a higher degree of market transparency.

Furthermore, the results of NCSIs are used by several companies and the media to increase their extent of goal fulfillment. In the case of **companies**, this means that the marketing department can get valuable information from NCSIs about the degree of achieved closeness to customers, which must be seen as crucial in many of today's market environments. Additionally, satisfaction data from NCSIs can play a key role in marketing controlling. Finally, media enterprises try to increase their circulation by reporting on the nation's degree of satisfaction, which corresponds to the informational needs of their customers.

2.2 A systematization of different comparisons implied by NCSIs

The calculation of satisfaction scores by NCSIs enables the researcher to analyze the data in a number of ways. Comparisons between different objects of the consumers' satisfaction level must be seen as the heart of the analysis. Fundamentally, these comparisons can be divided into intrabranched and interbranched comparisons. On the level of **intranched comparisons**, the following ways of analysis are possible: (a) The satisfaction data of company x generated by NCSI can be compared with its competitor within the same branched, that is company y. (b) The satisfaction data of company x can be related to the average satisfaction score of its branched. Both ways of analysis offer company x the possibility to define its position in the market. A common way of interpretation views comparatively high satisfaction scores as an expression of a high degree of customer and/or market orientation what is often taken as a synonym for competitiveness.

Interbranched comparisons of satisfaction scores generated by NCSIs open up further possibilities for data analysis. Especially three basic ways of analysis have to be mentioned here. First, the average satisfaction scores of different branches can be compared with each other, which may lead to the development of an interbranched satisfaction ranking. Second, the satisfaction score for company x of branched X can be compared with the average score of branched Y. Finally, company x may be interested in comparing its level of customer satisfaction with the satisfaction score of a specific company classified in a different branched.

To make it even more complex, each of the before mentioned comparisons based on NCSI data can principally be done as cross-sectional analysis or, if data collection is carried out at different points in time, as longitudinal analysis. Longitudinal data from

NCSIs can be helpful with regard to both controlling companies' and other institutions' activities and the proactive identification of changes in consumer behavior. Figure 2 illustrates the variety of optional comparisons on the basis of NCSI data.

- take in Figure 2 approx. here -

Difficult issues are in particular (however not exclusively) associated with interbranch comparisons. Which conditions do allow comparisons of satisfaction scores (on the company level or on the aggregated level of means) among different branches? As Anderson (1993, p. 101) stresses, these comparisons "require not only measurement, but also an understanding of how industry differences might affect satisfaction". The complexity shown above is of great importance especially, since it severely contrasts with the common impression of NCSI data as being easy-to-understand and self-explanatory. Furthermore, the neutrality of NCSI organizers and the incorporation of customer judgements seem to guarantee a high degree of objectivity. The contrast of putatively easy understanding on the one hand and complexity on the other hand bears the danger of misinterpretation for several stakeholders. For example, companies may continue to offer their products instead of modifying them, and consumers may be misled by NCSI information about a single company or a whole branch. As a consequence, specific requirements have to be formulated with regard to the provision and use of information generated by NCSIs. The following section of the paper represents an attempt in this direction.

3 Formulating Criteria for the Evaluation of NCSIs

Theoretical foundation. The validity of results from empirical research is closely related to the theoretical basis of the included constructs. Building on this, three sub-criteria for the construction of NCSIs can be derived. (a) The measurement of the construct customer satisfaction is to be based on a consistent theoretical framework that must be in accordance with the current state of knowledge in marketing science. (b) Closely related is the demand for an integration of meaningful parts of the theoretical framework when constructing the NCSI. This criterion is of particular importance when the NCSI is used for the development of expertise on a certain

company's success and competitiveness respectively. (c) Finally, the quality of NCSI data is combined inseparably with a valid operationalization of the incorporated constructs and especially the construct of customer satisfaction.

Aggregation of data. In the case of NCSIs, the existence of multiple stakeholders implies several steps of data aggregation. Thereby, data aggregation on the one hand has to be concrete enough for analysis on the single company level and on the other hand must enable the analysis on the branch or even national level. As a consequence, the complex aggregational procedure has to be carried out according to the following aspects. (a) The method of data aggregation must be in accordance with the requirements of the different stakeholders which also implies validity on the highest (national) level of aggregation. (b) To ensure the expressiveness of any comparison on the intrabranch level, the companies need to be classified in the "correct" branch. (c) If comparisons on the interbranch level are intended, the organizers of NCSIs must also prove the adequacy of their procedure in this area of research.

Validity of research design. As in other areas of social research, the expressiveness of NCSI data is closely related to the validity of the research design used for empirical data collection. At a more concrete level, three facets have to be mentioned here that determine the validity of the research method. (a) For an adequate interpretation of satisfaction scores of different companies or branches, information about the significance level is needed. We know that the significance criterion itself has its weaknesses for which it is criticized in parts of the social science literature, but in any case it provides important information ("better-than-nothing"). (b) Another aspect that comes along with the significance criterion is the sample size. An increase of the sample size represents an approximation of the population and simultaneously a decrease of sampling error (see Churchill 1979, p. 65, for the relationship between sampling error validity in empirical research). With regard to the different levels of analysis in NCSIs, the sample must be large enough to allow inferences to the consumers' satisfaction with the nation's supply with products, with single branches, and even with specific companies. (c) Eventually, an additional indicator for the quality of the data collection process is the length of the interviewing procedure. As numerous studies indicate, there exists a nonlinear relationship between the length of the interview and the validity of data (Hammann &

Erichson 1994, p. 95). After exceeding a threshold value at 30 to 45 minutes length, the validity of the data drops (Böhler 1992, p. 91; Berekoven, Eckert & Ellenrieder 1991, p. 100). Consequently, we demand that the design of the interviewing procedure for NCSI has to take notice of this relationship.

Inter-temporal comparability of results. In the case of longitudinal research, the usage of NCSI is committed to an inter-temporal comparability of results. To ensure the existence of this inter-temporal comparability, two premises have to be fulfilled. (a) The methods used for data collection, data aggregation, and the calculation of satisfaction scores must be left unchanged or must at least be comparable over time. (b) In addition to the stability of the research design, the comparability of data gathered at different points in time also depends on the development of several exogenous variables. These variables are much more difficult (or even impossible) to control for the organizers of NCSIs, but may strongly influence the development of satisfaction scores over time. Thus, it must be seen as a necessary requirement to consider these potential influential factors which would lead to a more experimental research design.

Transparency of methods and results. As illustrated above, NCSIs are complex measurement instruments with a number of snares. Thus, in addition to the previous aspects related to the generation of data, another criterion concerns the way of the editors' **communication** on NCSIs. An appropriate extent of transparency is needed here because of the danger of faulty and misleading interpretations. On a more concrete level, a key aspect of transparency refers to the necessity of combining the communication on NCSI data with information on the underlying limitations of methods used to generate these data.

Compatibility with economic statistics. Finally, from the existence of the government as a stakeholder of NCSI information, another requirement is derived. To ensure the usability of NCSI data on the macro economic (or governmental) level, the conceptualization of satisfaction scores has to be compatible with economic statistics. For this, (a) the aggregational procedure and (b) the structuring of branches are of special relevance.

4 A Criteria-Based Investigation into the Construction and Application of Selected NCSIs

4.1 The American Customer Satisfaction Index (ACSI) and the German Customer Satisfaction Barometer (DKB) as Reference Objects

The American ACSI and the German DKB are special types of NCSIs. Both are chosen here particularly for two reasons. First, the ACSI and the DKB are at the center of scholarly discussion in the area of NCSI. Their methodology is taken as basis for the development of other NCSIs in several countries, such as Austria, Switzerland, and Taiwan. Second, apart from their common objective of nationwide satisfaction measurement, a number of differences between both instruments exists. These differences illustrate the multitude of degrees of freedom in the design of NCSIs.

The development of the ACSI can be traced back to the conception of the Swedish Barometer SCSB, which was implemented in 1989, and is inseparably connected with the names of Claes Fornell and Eugene Anderson (NERA 1991). The ACSI was implemented in the United States in 1994 as result from a co-operation project of the American Society of Quality and the National Quality Research Center of the University of Michigan Business School. Today, the findings for selected branches are published quarterly. The ACSI measures the customers satisfaction with private and public service offerings, but also includes all kinds of durable and nondurable goods (see figure 3).

- take in Figure 3 approx. here -

The German NCSI, the DKB, was founded in 1992 on the initiative of the German Marketing Association and was sponsored by the German Post AG. Since then, it has been carried out annually. The methodology of the instrument was developed and revised in the meantime by the German marketing academics Anton Meyer and Frank Dornach. The DKB concentrates on private and public services from 36 branches of industry. Nonetheless, some manifest products (e.g., computer hardware) are also part of the sample.

4.2 A Criteria-Based Comparison of ACSI and DKB

4.2.1 Theoretical Foundation

Consistent theoretical framework

With regard to the existence and design of a theoretical framework, the ACSI and DKB differ fundamentally. The ACSI is based on a complex structural model which is shown in figure 4. The model contains five constructs in addition to customer satisfaction of which three (perceived quality, customer expectations, perceived value) are interpreted as antecedents of satisfaction (see for perceived value Holbrook 1994). These three variables are also treated as interrelated, with the perceived value construct being influenced by the customer's quality perception and his/her expectations. The two other variables (customer complaints and customer loyalty) are conceptualized as consequences of satisfaction.

- take in Figure 4 approx. here -

However, the validity of the postulated framework must be questioned in some points. With reference to recent studies, especially the interpretation of the customer's quality perceptions as an antecedent of customer satisfaction must be viewed as critical. As Hennig-Thurau and Klee show on the basis of a comprehensive literature review, satisfaction is appropriately regarded as "a shortterm emotional state that results from an intrapersonal comparison of the customers' expectations with the evaluation of a single... service encounter" (Hennig-Thurau & Klee 1997, p. 744), while the customer's quality perception is seen as a higher-order and more stable variable, which is built mainly on previous experiences of (dis)satisfaction related to discrete transactional episodes. Consequently, the causal relationship between satisfaction and perceived quality is opposite to what is postulated in the ACSI model: The emotional state of satisfaction leads to an overall, global, and cognitively dominated attitude about service quality (see also Dabholkar 1993; Bitner & Hubbert 1994; Bolton & Drew 1991, 1994). The most problematic aspect of the Fornell et al. (1996) model is the implicit understanding of quality perception as a quasi-objective construct. That is clearly illustrated when Strandvik and Liljander

(1995, p. 119) explain the state of being dissatisfied despite high quality by arguing that the product or service "does not fit the customer's preferences".

In opposition to the ACSI, the DKB is not based on a theoretical framework. In addition, the organizers of the DKB do not differentiate systematically between customer satisfaction and other related constructs. In the case of satisfaction and quality perception, they justified their position by quoting an empirical investigation carried out on their own, which shows no significant differences between both constructs (Meyer & Ertl 1996, p. 214). However, this finding must be questioned for it ignores the close relation between the empirical operationalization of constructs and their theoretical conceptualization (e.g. Homburg & Giering 1996).

Integration of Meaningful Parts of the Framework

In addition to customer satisfaction, both indices analyzed here collect information on the degree of customer loyalty and the consumers' complaining behavior. Both indices assume a direct and linear relationship between satisfaction and loyalty. This assumption, however, is inconsistent with recent theoretical and empirical research which shows that the influence of satisfaction on customer loyalty is rather restricted (Stauss & Neuhaus 1996; Hennig-Thurau & Klee 1997, p. 738 f.). Thus, researchers have argued for the integration of additional variables. In the context of relationship marketing literature, particularly the importance of quality perception, trust, and commitment are seen as crucial for customer loyalty or retention. A model which integrates all of these constructs was presented by Hennig-Thurau and Klee (1997, see also Hennig-Thurau 1998). In their model, these variables are taken as components of relationship quality, which itself is determined by the customer's level of satisfaction (see figure 5). The consideration of such a "meta-construct" relationship quality within the construction of NCSIs could contribute to increasing the level of explained variance.

- take in Figure 5 approx. here -

With regard to the influence of satisfaction and loyalty on profitability and turnover as indicators of economic success, both indices assume a significant and positive relation between these variables. While in the case of the ACSI this assumption is subject of empirical investigation (Anderson et al. 1997), Meyer and Dornach as

organizers of the DKB do not undertake detailed investigations in this area and refer to secondary statistical data instead (Meyer & Dornach 1995b, p. 40ff.). To get a more differentiated understanding of the relationship between psychographic and economic target variables, an integration of the latter into NCSIs' variable system can be seen as desirable.

Valid Operationalization of the Incorporated Constructs

Another difference between both indices lies in the operationalization and measurement of the satisfaction construct. By treating the customers' expectations as a separate variable, the ACSI takes into account the heterogeneity of expectations concerning different kinds of services and goods. This procedure enables the organizers of the ACSI to find out about the level of performance of the branches incorporated in the investigation. For example, Fornell et al. came to the conclusion that regarding the economy of the United States "customer expectations play a greater role in sectors in which variance in production and consumption is relatively low" (Fornell et al. 1996, S. 7). Satisfaction is measured here on a multi-item scale with three items (overall satisfaction, expectancy disconfirmation, performance vs. ideal product) because the construct is seen (in accordance with the relevant literature) as "not directly observable" (Fornell 1992, p. 12). This can be regarded principally as an adequate measurement procedure. However, the authors of the ACSI do not publish information on the reliability of the scale (e.g., coefficient alpha) which prevents us from a more substantial evaluation. With the exception of complaining behavior, the other variables in the conceptual ACSI model shown in figure 4 are also measured on multi-item scales (Fornell et al. 1996, p. 10).

In the case of the DKB, expectations and performance assessments are separated only up to a certain degree. Since 1995, the DKB includes autonomous items for the measurement of customer expectations, however, this information is not integrated into the calculation procedure but has a complementary status only. The satisfaction construct itself is measured on a single-item scale ("How satisfied are you with the services or goods of XY at all?") which maybe is the most serious deficiency of the DKB methodology. The problems associated with such a procedure have been intensively discussed in literature and should not be recapitulated here. A trenchant quote comes from Jacoby when he states: "How comfortable would we feel having

our intelligence assessed on the basis of our response to a single question?" (Jacoby 1978, p. 66; see also Churchill 1979, p. 66).

Recent findings of customer satisfaction research, especially on the predictive dimension of the satisfaction construct, are not considered neither by the ACSI nor the DKB. Especially the following two aspects can be expected to be fruitful for the further development of satisfaction measurement by NCSI (for an overview, see Hennig-Thurau & Klee 1997). First, some authors have mentioned that the influence of customer satisfaction (and the customer's quality perception respectively) on customer loyalty varies with the aspiration level of the customer (Stauss & Neuhaus 1996). As a consequence, several different states of satisfaction (e.g., "progressive satisfaction", "resigned satisfaction") can be distinguished. Each of them has a specific relation with customer loyalty. Second, another approach tries to differentiate satisfaction scores depending on the level of customer involvement (Bloemer & Poiesz 1989). This approach leads to the fundamental distinction between manifest satisfaction (corresponding to high involvement settings) and latent satisfaction (corresponding to low involvement settings). While in the case of manifest satisfaction a stable and significant influence of customer satisfaction on repurchasing behavior exists, such a relation does not exist when the customer has low or even no involvement in the product s/he has to assess (latent satisfaction). As with the different states of satisfaction, a differentiation of NCSI data with regard to the involvement level may rise the validity of measurement enormously.

4.2.2 Aggregation of Data

Method of Aggregation

Also the method of data aggregation differs significantly. The development of the ACSI was accompanied by an intensive scientific discussion on the permissibility of data aggregation in the case of individual satisfaction scores (Fornell et al. 1992; Anderson 1992; NERA 1991; Johnson & Fornell 1991; Fornell 1990). The search for an adequate way of calculating satisfaction scores for branches and the nation as a whole was the central issue of this discussion (Anderson 1992, p. 93 ff.) and led to an integration of the business administration and the economic perspective. The developed procedure was validated on empirical data of the Swedish barometer from

1989 to 1991. Furthermore, the method was modified to reduce distortions caused by the over- or underrepresentation of certain branches (Fornell et al. 1992).

On the company level, the index calculation is based on the multi-item measurement of satisfaction described above. In the next step, the company related satisfaction scores are aggregated to a satisfaction score for the respective branch. Here, aggregation implies the multiplication of each company related score with a specific weighting factor. This weighting factor represents the percentual contribution of a single company to the total turnover of the branch. The weighting procedure of the ACSI is criticized by Meyer and Ertl (1996, p. 224), who argue that "the satisfaction judgement of a Mercedes driver is of higher importance for the branch satisfaction score as the one of a person who drives a VW Golf". However, such a procedure is advantageous with regard to the criterion of compatibility with economic statistics (see below). Furthermore, branches are summed up to more abstract **sectors of industry** (e.g., services, manufacturing durables) using the same weighting procedure again. This time, the weighting factors are derived from the macro economic system of national accounts.

In contrast to this complex and historically grown method, the data aggregation procedure of the DKB is rather straightforward. Although its organizers have taken notice of the Swedish experiences (Meyer & Ertl 1996), they do without any weighting procedure. Instead, unweighted means are aggregated on both the company level and the branch level. To improve the quality of data interpretation, in addition to the overall satisfaction score the customers' satisfaction with a number of attributes are measured. These attributes differ from branch to branch (e.g., employees' friendliness, quickness of task fulfillment, technical skills, see Meyer & Dornach 1995b, p. 89), but their identification is not explained by the organizers of the DKB and no evidence is given for their relevance for the customer's satisfaction judgement. The customers' satisfaction on the attribute level is not incorporated into the aggregation procedure and, therefore, has to be interpreted in isolation.

Classification of Companies and Intra-branch Comparison

The assignment of companies to certain branches forms the basis for **intra-branch** comparisons of customer satisfaction which can be seen as a central component of NCSI analysis. By the definition of branches a decision is made on the kind of companies which are interpreted as competitors within the same market. So, the

selection of comparison standards for a single enterprise is strongly influenced by this decision (see for the general relevance of **relative** satisfaction judgements Dick & Basu 1994).

The definition of branches and the assignment of companies to a specific branch is oriented in the case of the ACSI to the structure of the system of national accounts (ACSI). In the case of the DKB, no single formal criterion for structuring is given by the organizers but the result is close to popular systematization. Especially the aggregation of branches to the higher level construct sector implied by the ACSI methodology raises problems. For example, a "services" sector and a "finance and insurance" sector are distinguished, though both sectors are closely related with regard to several aspects, and therefore a distinction must be seen as highly arbitrary. Furthermore, the systematization of branches of both indices is partially superficial and vague. For instance, in the case of the ACSI household batteries and HiFi systems are both assigned to a common branch named "consumer electronics", and the DKB asks the consumers to evaluate a heterogeneous category (and branch, respectively) titled "vacational destinations".

The problems arising from these kinds of systematization are illustrated by a decision of the regional court Hamburg. According to this decision, the Techniker Krankenkasse is not allowed to communicate that they are the "top" of the branch of health insurance companies (LG Hamburg, AZ 315 O 53/97). The court argued that according to its view, the market offers of private health insurance companies on the one hand and public health insurance companies on the other hand are not comparable with each other.

Methodology of Interbranch Comparisons

The problem of heterogeneity described above within the context of intrabrand comparisons is intensified on the **interbranch level**. At the DKB, the aggregated satisfaction scores for the different branches are merged into a total rating of branches (see figure 6, which shows an excerpt from the 1997 rating of branches). Apart from the general dubiousness of such a ranking based on the inseparability of the evaluation of companies and the public image of the companies offerings (e.g., see the case of social organizations), additional problems coming along with such a way of analysis can be exemplified by means of the comparison between life insurance companies and companies offering computer hardware. For example,

while life insurances can be characterized through a large amount of credence qualities, computer hardware predominantly exists of search and experience qualities, which leads to different internal judgement processes. Furthermore, the comparison between monopolistic branches (e.g., the German Mail) and branches with high competition (e.g., parcel services) must be seen as inappropriate because of the different levels of customers expectations (see also below).

- take in Figure 6 approx. here -

The organizers of the ACSI do not calculate a ranking of the branches' satisfaction scores. Instead, the pros and cons of comparisons between different branches are intensively discussed and the specific features of certain branches are worked out (Fornell et al. 1996; Anderson 1993, S. 101). However, they keep themselves the option for interbranch comparisons, when they say: "We can determine how well a particular firm is doing relative to the best firms in its own industry, the best firms in other industries in that sector, or the best in nation as a whole. **Industries and sectors can be compared with one another in a similar fashion**" (Fornell et al. 1996, p. 12; italics added). Referring to this statement, the existence of something like "the best company in nation as a whole" shall be questioned here with vigor once again.

Anderson (1993, p. 93f.) indicates in this context that even the intensity of the relationship between customer satisfaction and loyalty may differ from branch to branch, which also effects the relative importance of satisfaction scores.¹ This founding is also supported by an empirical study of the authors, whose results are described in a prior version of this paper (Hansen, Korpiun & Hennig-Thurau 1998).

4.2.3 Validity of Research Design

Significance of Data

¹ Interestingly, this argument can also be found at Meyer & Dornach (1995a, p. 433).

For the addressees of NCSI data, the validity of satisfaction scores and company rankings based on these scores is of central relevance. As a general guideline, rankings or comparisons of companies are permissible only if they are **significant** from a statistical point of view. If the differences between the companies considered in the comparison are not significant, one has to assume that the measured result is an artifact and does not exist in reality. So, a follow-up study could lead to different or even reverse results. In the health insurance companies example cited above, the difference between the Techniker Krankenkasse as "best company" and other health insurance companies was rather small (0.02 points on a five-point scale in 1996), so that the competent regional court consequently has questioned the validity of any ranking based on these data.

Both indices inspected here publish no information on the significance of satisfaction comparisons neither on the branch nor on the company level what has to be criticized. This omission prevents the user from an assessment of the validity of any comparison or ranking as mentioned above. Closely related to this deficiency there is another weak point of the DKB: On the single-item scale used, the difference between the best and the worst satisfaction score on the level of branches is approximately only 0.2 points that implies the assumption that the used measurement scale has no good selectivity due to skewness. In the case of the ACSI, these problems do not occur to the same degree because it uses a different scale (ten-point instead of five-point scale; see Fornell 1992, p. 13).

Sample Size

Both barometers, the ACSI and the DKB, are based on large samples. In the case of the ACSI, data from more than 45.000 respondents is used to calculate the scores (Fornell et al. 1996, p. 13). For the DKB, a sample of approximately 37.000 consumers is contacted (DKB 1998b). As a consequence, analysis of aggregated levels (national level, sector level, branch level) will raise no validity problems. However, on the level of single enterprises, the DKB calculates satisfaction scores based on sample sizes of 100 (and even smaller, see Meyer & Dornach 1995b, p. 136), what limits the generalizability of the calculated results significantly. Here, an increase of the number of respondents per company is strongly required for getting valid information on this level of analysis. In the case of the ACSI, for the calculation of satisfaction scores on the company level approximately 250 interviews are

conducted per firm (Fornell et al. 1996, p. 9). From a methodological point of view, this can be seen as some kind of lower limit for the generation of valid data.

Length of Interviewing Procedure

For both instruments, there is no information published about the average length of the interviewing procedure. However, while in the case of the ACSI customers are only asked to evaluate companies from one or two branches, DKB respondents have to answer questions regarding their satisfactory state with up to nine different branches (Meyer & Dornach 1995c, p. 171). Unlike the ACSI, which includes only 25 standardized questions (17 structural, 8 demographic), the DKB covers a number of additional questions (Meyer & Dornach 1995a, p. 442 f.) which extends the interviewing procedure. Furthermore, in both cases data is predominantly collected by telephone interviews. This influences the relationship between interviewing length and validity, as Churchill states: "The general rule of thumb is that long questionnaires can be handled best by personal interview and **least well by telephone interview**" (Churchill 1991, p. 335; italics added). As a consequence, it must be supposed that the length of the interviewing procedure of the DKB exceeds the critical threshold thus having a negative impact on the validity of NCSI results.

4.2.4 Inter-temporal Comparability of Results

Continuity of Methods

Changes in the level of customer satisfaction over time may be used as an indicator for increases or decreases of market performance. Therefore, they principally can be interpreted as a key variable in a company's or the governmental controlling system. The condition of method continuity is met by both the ACSI and the DKB. The methodology of the ACSI has been slightly modified since its first use in 1994, but the central aspects were left unchanged. In the case of the DKB, the methodological foundation is also kept constant. Merely the number of considered firms and the quantity of additional information have been extended continuously by the organizers of the instrument (Meyer & Dornach 1995b).

Constancy of Exogenous Variables

Even in the case of methodological continuity, a number of obstacles are still related with longitudinal satisfaction research. Most of these obstacles have to do with the dynamic environment, which influences the customer's state of (dis)satisfaction over time. So, in recent years many efforts have been made by companies (e.g., total quality programs, relationship management tools) to reach a higher level of customer orientation and to get "closer to the customer". All these efforts seem to have been unsuccessful because till 1996 the degree of customer satisfaction has constantly fallen nationwide: between 1994 to 1996 the ASCI has lost 3,1% (NQRC 1997), and the results of the DKB show a quite similar development.²

While some researchers (e.g., Meyer & Dornach 1995b, p. 1 ff.) view this development as the true expression of a societal trend and consequently postulate measures to stop it, other scientists rather draw the attention to changes of exogenous variables in general and to the consumers rising aspiration level in particular. To explain this increase of the aspiration level, particularly the following arguments are stated frequently:

- Today consumers find more and more pleasure in the output of NCSIs. This kind of popularization is mainly based on the coverage of NCSI results in large parts of the media. As a consequence of the increased popularity of NCSIs, consumers are becoming more sensitive to customer satisfaction (Köhler 1996, p. 583; Nerdinger 1997). Although this may be intended by the organizers of NCSIs, the implications of this rise of customers expectations are integrated in the NCSI methodology at most partially. In the case of ACSI, the modeling of expectations as a separate construct can at least be seen as starting point.
- Changes of customer satisfaction over time can also occur when the products or services as reference objects are modified in a significant dimension (e.g. as a result of technical progress). With the modification of the product there usually goes along a justification of the internal expectation standard on the customers' side, which limits the validity of a comparison between former and today's levels of satisfaction.

² In 1997, for the first time since its renewal in 1992 the DKB has found an increase of customer satisfaction in some areas.

- Finally, customer expectations are implicitly tied to existing market structures (e.g., monopolistic or competitive markets). If these market structures change over time as it happens today in Germany with the telecommunications market, new competitors are appearing whose offerings also change the customers' expectation levels. Consequently, a comparison of satisfaction scores on markets with monopolistic conditions with scores on highly competitive markets must be questioned in general for its validity.

4.2.5 Transparency of Methods and Results

As illustrated above, several stakeholders and especially the media are taking notice of NCSI results. However, often only parts of the information necessary to develop a holistic and therefore true understanding of the relationships investigated are noted by the addressees. As a consequence, the actions taken by consumers or companies based on the received NCSI information may be frequently suboptimal or, in some cases, even of danger for the continuity of business.

Looking at ACSI and DKB, especially the latter has to be criticized with regard to this criterion. Here, several general statements can be found both in the official report and in scientific articles which do not cover the different restrictions of the methodology discussed in parts in this paper. For example, the validity problems arising from intra- and interbranch comparisons and from small sample sizes are not even mentioned. Closely related, companies are explicitly encouraged by the organizers to integrate the results of the DKB within their advertising strategy (see Meyer & Dornach 1995a, p. 450) but no additional information is given on how to handle the complexity of the data. Furthermore, the importance of the satisfaction construct for customer loyalty and economic success is not put into perspective by the authors of the DKB. Instead, satisfaction is rather treated as some kind of "magic formula" for a company's success. In the case of the ACSI, an extensive methodological report is available and the organizers have referred to at least some of the caveats related to NCSI in a number of scholarly articles.

Additionally, for both barometers (DKB and ACSI) relevant information on the methodology of the instrument has not been published yet (e.g., significance levels, reliability coefficients). So, summing up the things mentioned above, the organizers

and communicators of both barometers' data are asked to commit themselves to a higher degree of transparency and completeness with regard to the results and the methodology of NCSIs.

4.2.6 Compatibility with Economic Statistics

Because of the differing foci of both indices, it is not surprising that the ACSI has a higher level of compatibility with economic statistics in general. With regard to the aggregation of data, one can comment that as a consequence of the weighting procedure the satisfaction scores can easily be integrated into the national accounting system. In the case of the DKB, no weighting is performed so that a combination of results with economic statistics is hardly possible. With the structuring of branches it is almost the same: while the ACSI method is directly based on the multi-level standard industrial systematization (SIC), the less formally derived systematization of the DKB restricts the possibilities of cooperative analysis between the micro (marketing) and the macro (economics) level.

5 Summary and Implications for Further Research

NSCIs give information about the level of consumers' satisfaction and fulfill several needs. At the same time, the complexity of their methodology raises a number of questions regarding the validity of the generated results. Drawing upon the investigation undertaken here, some proposals shall be formulated that may contribute to the further development of the instruments. To go into details, the following aspects seem worth mentioning:

PROPOSITION 1: The measurement of customer satisfaction should be based on a theoretical framework (DKB) that integrates recent knowledge from different areas, but especially from relationship marketing research (ACSI). Central aspects should be an extended conceptualization of the satisfaction construct, and the integration of other determinants of customer loyalty, e.g. trust and commitment. In the case of the DKB, in addition the operationalization of the satisfaction construct should be changed into a multi-item approach replacing the single-item approach used up to now.

PROPOSITION 2: In the context of data aggregation, the existing systematizations of branches have to be object of modification. The research on market structure analysis (see Day, Shocker & Srivastava 1979; Myers & Tauber 1977; Bauer 1989) can be expected to turn out as a fruitful source for this. As we have shown in our analysis, interbranch comparison and especially rankings of branches are accompanied by a number of problems. Because of their questionable validity, their calculation and publication should generally be reconsidered.

PROPOSITION 3: The relevance of NCSIs should be increased by communicating additional information (e.g., significance level of satisfaction comparisons) with regard to the underlying methodology. Furthermore, the generalizability of all generated results should be tested with regard to the background of the sample size prior to publication.

PROPOSITION 4: Intertemporal comparisons of satisfaction scores should be complemented by information about central exogenous variables. Here, particularly the customers aspiration level is of great importance for the interpretation of longitudinal data.

PROPOSITION 5: The way of communication on NCSI results should consider the multifarious risks of misinterpretation. The organizers should prove responsibility for the generated data by informing their addressees on the existing methodological problems and limitations of their instrument.

PROPOSITION 6: As the actors of economic policy are NCSI stakeholders, it has to be checked whether the compatibility of the indices' methodologies with economic statistics can be heightened. By the way, this must be taken into account when modifying the systematizations of branches as it was suggested in proposition 2.

References

- Anderson, E. W. (1993): Firm, Industry and National Indices of Customer Satisfaction. Implications for Services, in: *Advances in Services Marketing and Management*, Vol. 2, pp. 87-108.
- Anderson, E. W.; Fornell, C. & Rust, R. T. (1997): Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services, in: *Marketing Science*, Vol. 16 (in print).
- ASQ (1998): American Society for Quality, National Economy-Sectors and Industries, <http://www.asq.org/graphics/fig2a.gif>.
- Bauer, H.H. (1989): *Marktabgrenzung*, Berlin 1989.
- Berekoven, L.; Eckert, W. & Ellenrieder, P. (1991): *Marktforschung: Methodische Grundlagen und praktische Anwendung*, 5. Aufl., Wiesbaden 1991.
- Bitner, M. J. & Hubbert, A.R. (1994): Encounter Satisfaction Versus Overall Satisfaction Versus Quality: The Customer's Voice, in: Rust, R. T. & Oliver, R. L. (eds.): *Service Quality: New Directions in Theory and Practice*, Thousand Oaks, pp. 72-94.
- Böhler, H. (1992): *Marktforschung*, 2. Aufl., Stuttgart u. a. 1992.
- Bolton, R. N. & Drew, J. H. (1991): A Multistage Model of Customer's Assessments of Service Quality and Value, in: *Journal of Consumer Research*, Vol. 17 (October), pp. 375-384.
- Bolton, R. N. & Drew, J. H. (1994): Linking Customer Satisfaction to Service Operations and Outcomes, in: Rust, R. T.; Oliver, R. L. (eds.): *Service Quality: New Directions in Theory and Practice*, Thousand Oaks, pp. 173-200.
- Churchill, G. A. (1979): A Paradigm for Developing Better Measures of Marketing Constructs, in: *Journal of Marketing Research*, Vol. 16 (February), pp. 64-73.
- Dabholkar, P. A. (1993): Customer Satisfaction and Service Quality: Two Constructs or One?, in: Cravens, D. W.; Dickson, P. (eds.), 1993 AMA Educators' Proceedings, Chicago, Ill., pp. 10-18.
- Day, G.S.; Shocker, A.D. & Srivastava, R.K. (1979): Customer-Oriented Approaches to Identifying Product-Markets, in: *Journal of Marketing*, Vol. 43 (Fall), pp. 8-19.
- DKB (1998a): Das Ranking der Globalzufriedenheit 1997, http://www.postag.de/kundenbarometer/dkb_ranking97.html, 15.03.1998.
- DKB (1998b): Was ist das Deutsche Kundenbarometer?, http://www.postag.de/kundenbarometer/dkb_was97.html, 10.09.1998

- Fornell, C. (1992): A National Customer Satisfaction Barometer: The Swedish Experience, in: *Journal of Marketing*, Vol 56 (1992), January, pp. 6-21.
- Fornell, C.; Johnson, M. D.; Anderson, E. W.; Cha, J. & Bryant, E. (1996): The American Customer Satisfaction Index: Nature, Purpose, and Findings, in: *Journal of Marketing*, Vol. 60 (1996), October, pp. 7-18.
- Hammann, P.; Erichson, B. (1994): *Marktforschung*, 3. Aufl., Stuttgart, Jena 1994.
- Hansen, U.; Korpiun, M. & Hennig-Thurau, Th. (1998): Nationale Kundenzufriedenheitsindizes als Informationsgrundlage des Dienstleistungs-Management. Eine kritische Bestandsaufnahme, in: Bruhn, M. & Meffert, H. (eds.): *Dienstleistungs-Management*, Wiesbaden 1998 (in print)
- Hennig-Thurau, Th. (1998): Konsum-Kompetenz: Eine neue Zielgröße für das Management von Geschäftsbeziehungen, Frankfurt/M. u.a. 1998 (in print)
- Hennig-Thurau, Th. & Klee, A. (1997): The Impact of Customer Satisfaction and Relationship Quality on Customer Retention: A Critical Reassessment and Model Development, in: *Psychology & Marketing*, Vol. 14, Special Issue on Relationship Marketing, December (No. 8), pp. 737-764.
- Holbrook, M. B. (1994): The Nature of Customer Value, in: Rust, R. T.; Oliver, R. L. (eds.): *Service Quality: New Directions in Theory and Practice*, London 1994, pp. 21-71.
- Homburg, C. & Giering, A. (1996): Konzeptualisierung und Operationalisierung komplexer Konstrukte, in: *Marketing ZFP*, 18. Jg. (1996), H. 1, pp. 5-25
- Jacoby, J. (1978): Consumer Research: A State of the Art Review, in: *Journal of Marketing*, Vol. 42 (1978), April, pp. 87-96.
- Meyer, A. & Dornach, F. (1995a): Nationale Barometer zur Messung von Qualität und Kundenzufriedenheit bei Dienstleistungen, in: Bruhn, M. & Stauss, B. (eds.): *Dienstleistungsqualität: Methoden, Konzepte, Erfahrungen*, Wiesbaden 1995, pp. 429-453.
- Meyer, A. & Dornach, F. (1995b): Das Deutsche Kundenbarometer 1995 - Qualität und Zufriedenheit, hrsg. von Deutsche Marketing Vereinigung e. V. und Deutsche Bundespost Postdienst, Düsseldorf u. a. 1995.
- Meyer, A. & Dornach, F. (1995c): Das Deutsche Kundenbarometer - Qualität und Zufriedenheit, in: Simon, H./Homburg, Ch. (eds.): *Kundenzufriedenheit: Konzepte, Methoden, Erfahrungen*, Wiesbaden 1995, pp. 161-178.
- Meyer, A. & Dornach, F. (1996): Das Deutsche Kundenbarometer 1996 - Qualität und Zufriedenheit, hrsg. von Deutsche Marketing Vereinigung e. V. und Deutsche Bundespost Postdienst, Düsseldorf u. a. 1996.

- Meyer, A. & Ertl, R. (1996): Nationale Barometer zur Messung von Kundenzufriedenheit. Ein Vergleich ausgewählter grundlegender Aspekte des "Deutschen Kundenbarometers - Qualität und Zufriedenheit" und des "American Customer Satisfaction Index (ACSI)", in: Meyer, A. (ed.): Grundsatzfragen und Herausforderungen des Dienstleistungsmarketing, Wiesbaden 1996, pp. 201-231.
- Myers, J.H.& Tauber, E. (1977): Market Structure Analysis, Chicago 1977.
- NERA (National Economic Research Associates) (1991): Developing a National Quality Index: A Preliminary Study of Feasibility, New York 1991
- Stauss, B. & Neuhaus, P. (1996). The Dissatisfaction Potential of Satisfied Customers - The Qualitative Satisfaction Model. Paper presented at the French-German Workshop on Services Marketing Research, 28.2-1.3.1996 in Innsbruck, Austria.
- Strandvik, T. & Liljander, V. (1995). A Comparison of Episode Performance and Relationship Performance for a Discrete Service, in Kleinaltenkamp, M. (ed.): **Services Marketing: Concepts and Applications**, Wiesbaden 1995.

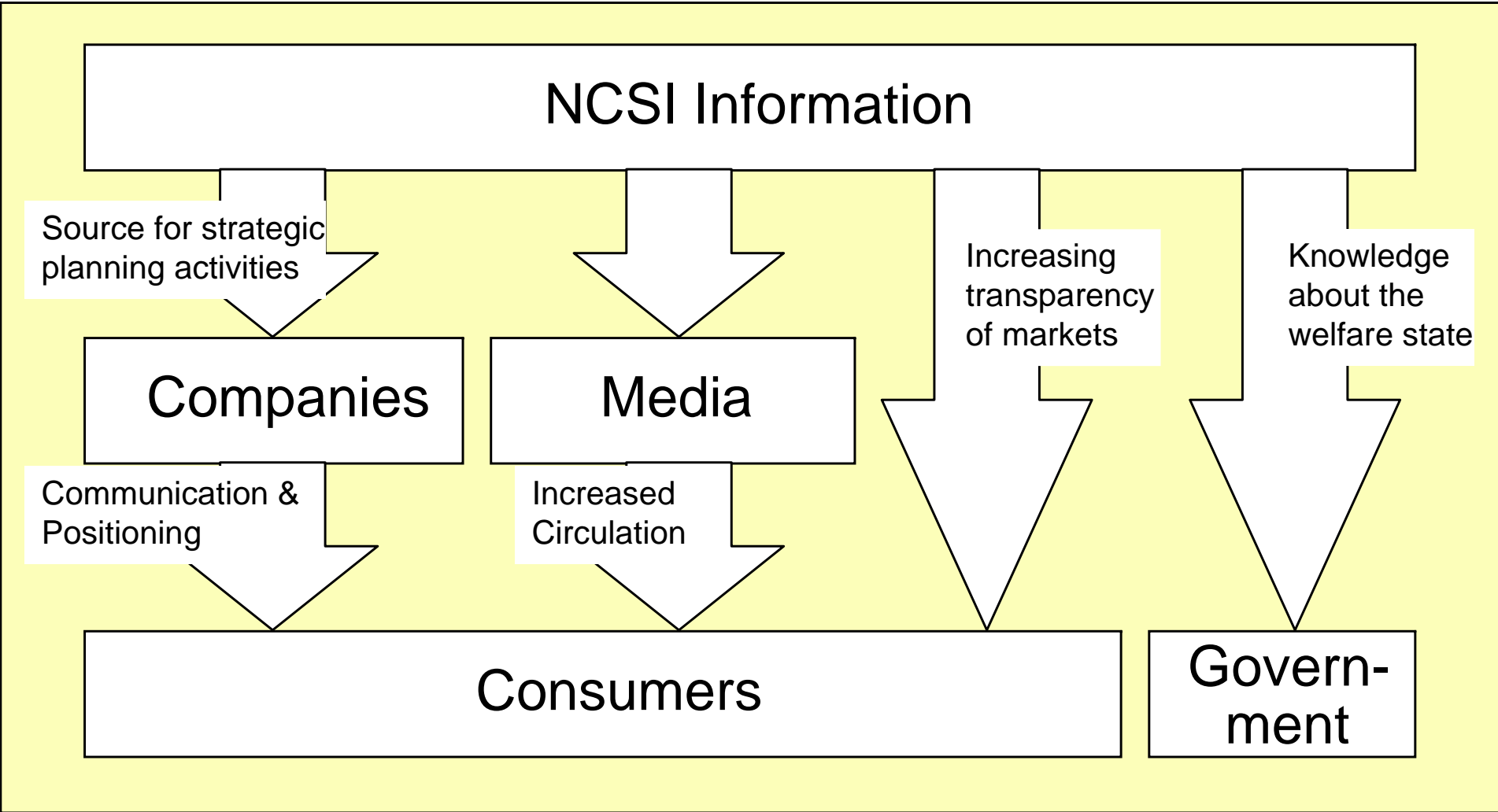


Figure 1: A systematization of central stakeholders of NSCIs

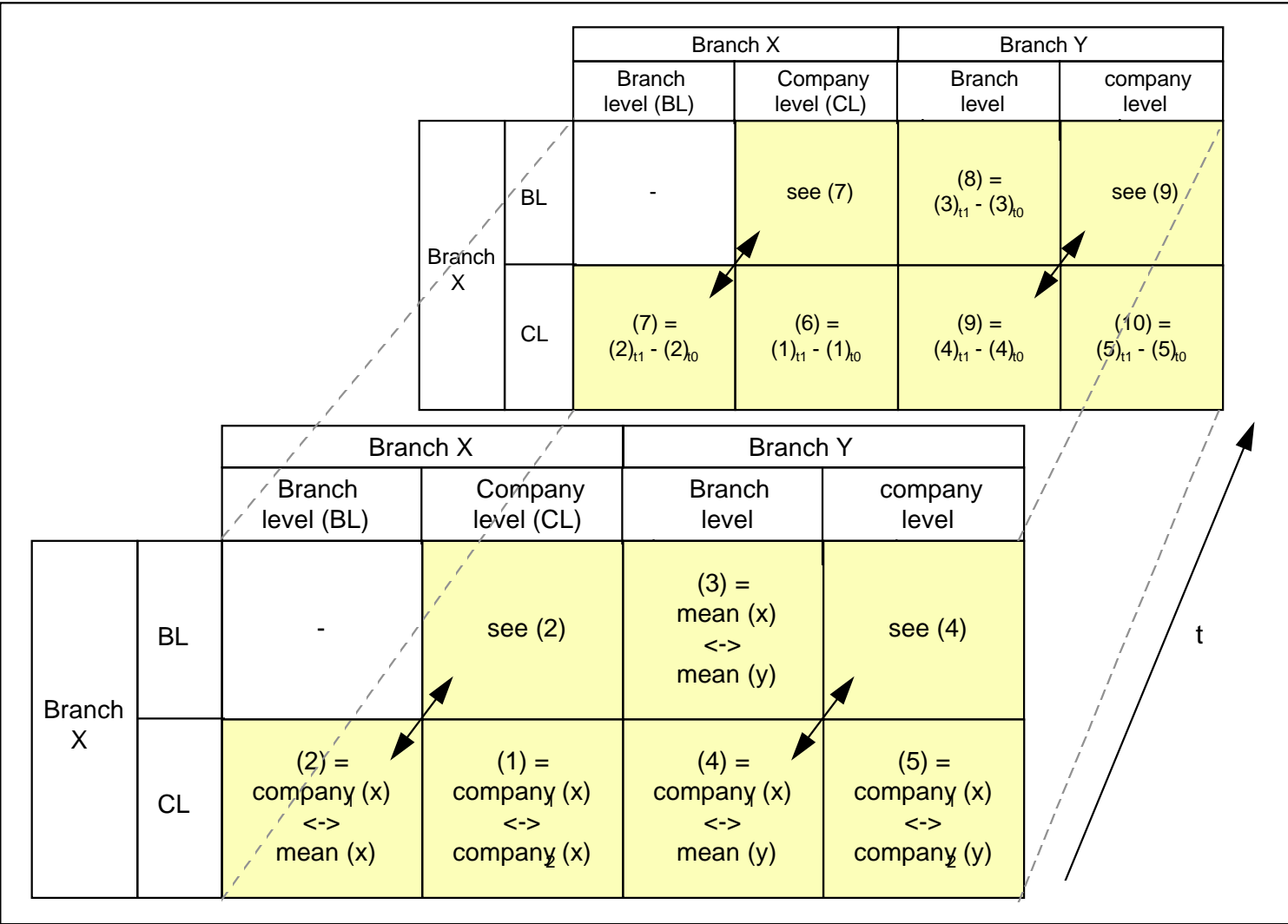


Figure 2: Optional comparisons based on NCSI data

ACSI: National Economy—Sectors and Industries

ACSI

Manufacturing Nondurables 8.3% (SIC 2)	Manufacturing Durables 9.6% (SIC 3)	Transportation, Communications, Utilities 8.8% (SIC 4)	Retail 9.3% (SIC 5)	Finance, Insurance 6.6% (SIC 6)	Services 9.0% (SIC 7,8)	Public Administration, Government 12.6% (SIC 9)
Apparel Athletic shoes Beer Cigarettes Food (6 types) Gasoline Newspapers Personal care Soft Drinks \$654 B	Automobiles Consumer electronics Household appliances PCs/printers \$439 B	Airlines Broadcasting/TV Electric service Express delivery Phone/ long-distance U.S. Postal Service \$450 B	Department stores Discount stores Restaurants/ fast food Supermarkets \$375 B	Banks/ commercial insurance/life insurance/ property \$659 B	Hospitals Hotels Motion pictures \$48 B	Garbage/trash service (central city, suburb) IRS \$52 B

Size of each of 7 sectors shown as % of GDP. Sales figures reflect total sales for the selected U.S. firms. Total sales of selected firms are 58% of GDP for these 7 sectors and 43% of national GDP.

Figure 3: Branches included in the American Customer Satisfaction Index (ASQ 1998)

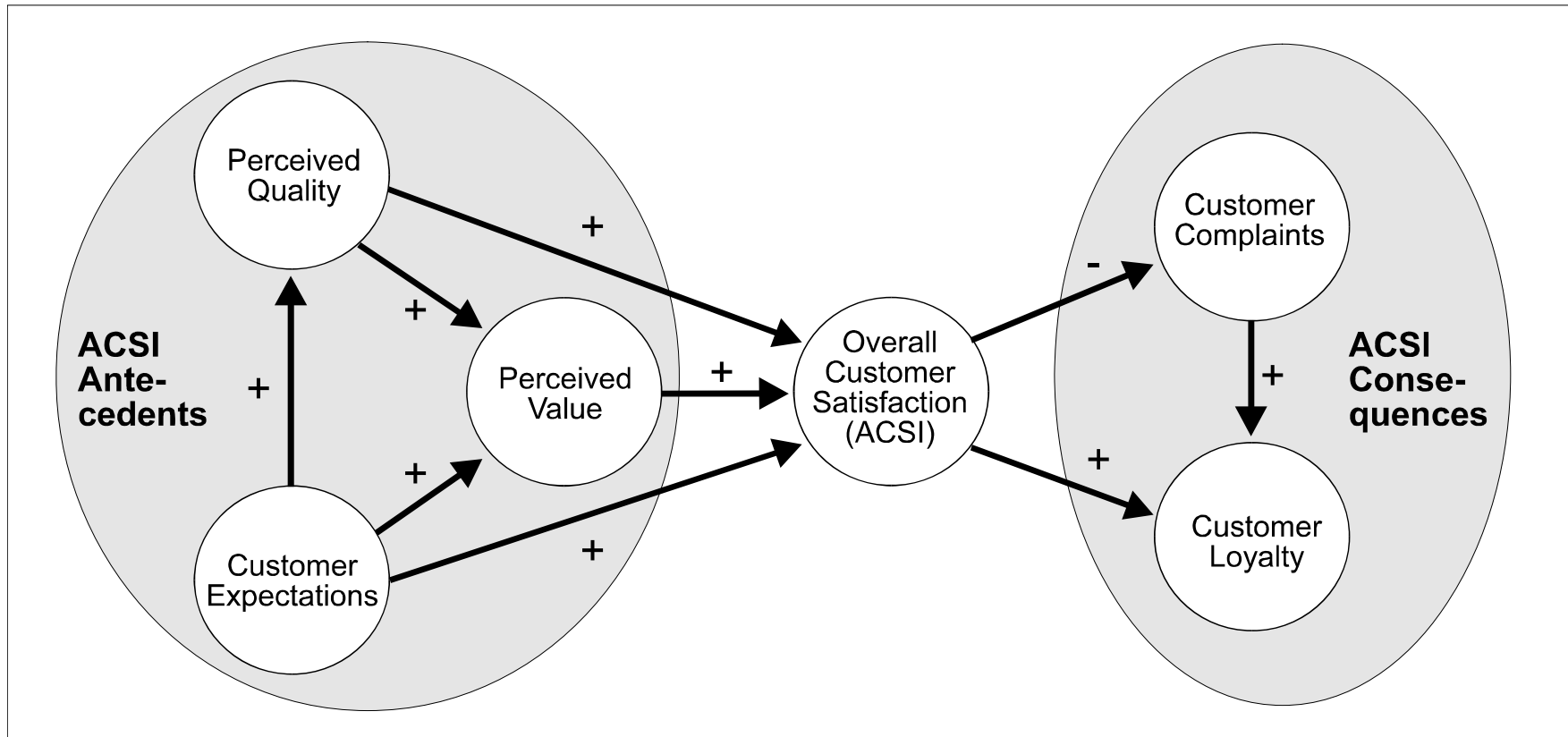


Figure 4: Structural model of the American Customer Satisfaction Index (Fornell et al. 1996, p. 8)

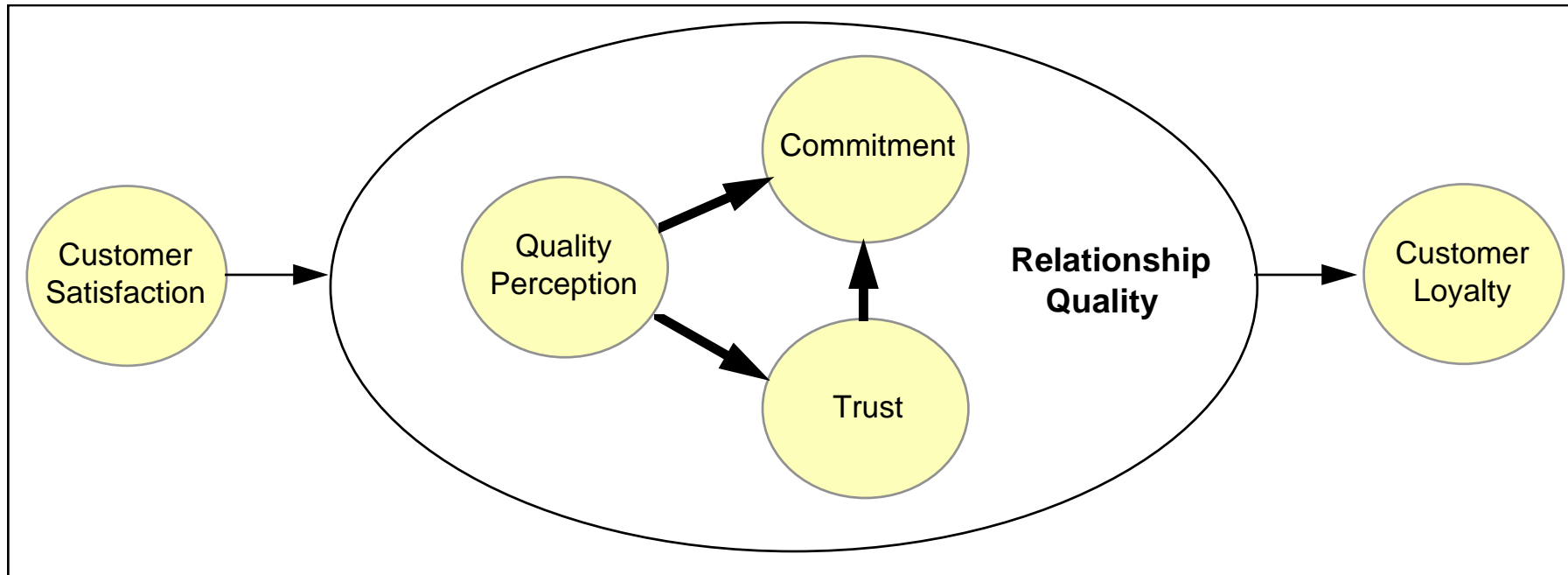


Figure 5: Relationship quality as moderator between customer satisfaction and customer loyalty (Hennig-Thurau & Klee 1997, p. 742)

vacational destinations	1,99
pharmacies	2,21
travel companies	2,23
broadcast stations	2,25
credit card companies	2,25
motoring organizations	2,27
car insurance companies	2,28
airlines (private customers)	2,30
garages	2,31
cellular phone companies (hardware)	2,36
mail order firms	2,37
electrical appliances companies (customer service)	2,38
health care insurance companies	2,39
closing traders	2,40
banking companies	2,41
social and ecological organizations	2,41
supermarkets/food shops	2,41
drug stores	2,42
computer companies (hardware)	2,48
computer companies (operating systems)	2,49
building societies	2,53
life insurance companies	2,53
energy suppliers	2,54
do-it-yourself markets	2,58

Figure 6: DKB 1997 rating of branches (excerpt; DKB 1998a)