Preamble

The purpose of empirical ex post evaluations (impact studies) is to quantify the impact of various types of interventions, mostly government interventions, by measuring how effectively such an intervention accomplishes the desired objectives and comparing its effectiveness with its costs (efficiency). This is why impact studies are a key element in the evaluation of policy interventions and for political decisions.

The Verein für Socialpolitik (VfS), the association of German-speaking economists, is addressing these guidelines and recommendations to impact analysis providers as well as to those political and administrative decision makers responsible for planning government interventions and commissioning evaluation studies. The VfS is seeking to enhance an informed dialogue on the effectiveness of government policies and thus to contribute to improving the use of existing knowledge and scientific methods.

These guidelines are especially intended to serve as a transparent foundation for interpreting, assessing and discussing the quality of evaluation studies. Their purpose is to provide guidance to academics, policymakers, lawmakers and the general public for evaluating empirical statements on policy measures. These guidelines are centred on the principles of methodological quality, transparency, objectivity and independence, which are universal irrespective of the specific area of application and the concrete methodology being used. They focus on empirical methods that use micro data to study the effect of measures already conducted and thus to perform ex post evaluations.

For a variety of reasons, not all evaluation studies will satisfy the criteria listed in the guidelines. These reasons can be anything from an inadequate data set to an insufficient legal basis for an evaluation; also, there could be a particular (historical) evaluation situation. In principle, however, the availability of micro data on individuals, households, companies or other entities has improved, both in Germany and internationally over the past few years. At the same time, improvements are still necessary in some areas.

The guidelines are followed by recommendations for political and administrative decision makers responsible for planning and conducting policy interventions and for commissioning impact studies. The recommendations particularly emphasise the need for pre-defining verifiable objectives of policy interventions, the use of a valid methodology, the publication of the results and the facilitation of independent replication studies. Moreover, the VfS is encouraging policymakers, lawmakers and administrative agencies to improve the political and legal framework in a way that will create an even better data environment for impact analyses while ensuring extensive data protection.
The Verein für Socialpolitik's guidelines for *ex post* evaluations based on micro data

**General information**

1. Evaluations (impact studies) should be transparent and comprehensive. Their underlying assumptions should be outlined clearly. The methodological limitations of evaluation studies should be discussed extensively.
2. The author(s) of the study should fully disclose any potential conflicts of interest.
3. Impact studies should be unbiased and open-ended.
4. The results of evaluation studies should be presented in a neutral and non-judgmental manner. Attention should be paid to the distinction between describing the facts and interpreting and assessing the results. When publishing an evaluation study, it is necessary to indicate whether publication required the prior consent of a commissioning party or a third party.
5. Should it prove impossible to satisfy the criteria of the guidelines, good reasons should be given in order to ensure the transparency of the limitations and the data and methodology used.

**Methodology**

The primary challenge in drafting evaluation studies rests in finding methodologies suited to identifying the causal effects of an intervention and to evaluate its effectiveness and efficiency. Depending on the topic of the study and the availability of data, different approaches and methodologies can serve this goal. The assumptions allowing a causal interpretation of the empirical results always need to be outlined clearly, as do the premises under which the derived findings on the effectiveness of the intervention are valid. The plausibility of these assumptions should be substantiated and discussed.

When choosing and presenting the methodology, the following aspects should be borne in mind.

6. The choice of methodology should meet the highest scientific and ethical standards. Wherever this is possible and reasonable, impact analyses should be conducted using (quasi-)experimental methods. The analyses should be designed such that groups that are (more heavily) affected by a measure (the treatment group) can be compared with groups that are not (or less severely) affected (the control group). Descriptive analyses of the data may serve as a useful complement.
7. The comparability of the treatment and control groups should be discussed extensively. In particular, potential (self-)selection bias of treated and control groups should be discussed in detail, and a potential change in the composition of the treatment and control groups over the course of the evaluation period should be examined. During the conceptual phase of an intervention, it may make sense to plan multiple designs of control groups.
8. Evaluation studies should use appropriate outcome measures which enable the quantification of the impact of a measure in a reasonable manner. Evaluations should distinguish between short-term, medium-term and long-term impacts and also between local and global impacts of a measure, as well as, if necessary, address the sustainability of the impact.
9. The results of impact analyses should be discussed carefully.
   a. Methodological differences between the current study and earlier studies should be discussed, especially results differ.
   b. The sensitivity of the results to key decisions which affect the methodology of the analysis and the data needs to be documented and discussed.
   c. The internal and external validity of the results, i.e., the correctness of the findings and their applicability in other situations, should be discussed.
10. Where legally possible, the data sets, computer software and other documentation used should be made available so that the results can be replicated easily. If this is not possible, the reasons for this should be given. The data access channels should be discussed and documented.
**Recommendations to political and administrative decision makers**

Political and administrative decision makers can contribute greatly to making better use of available knowledge and scientific methodology in evaluation studies. The earlier an evaluation is planned, the more readily and cost-effectively a high-quality analysis can be conducted. The quality of impact analyses can be decisively affected by taking the evaluation design into account as early as in the planning stages of a measure and the awarding of a contract.

1. When planning a measure, its intended objective should be named and fixed. Targets and measurable criteria for the success of the intervention should already be defined ex ante.

2. In order to reduce the costs of gathering data on key variables, the evaluation of an intervention should already be planned before its introduction, and the necessary data should be gathered during the implementation phase. In particular, it should be assured that sufficient information on suitable control groups can be made available.

3. Evaluation studies should be tendered publicly.

4. When awarding contracts, attention should be paid to the use of a valid methodology (in accordance with the guidelines).

5. Evaluations should be published in a timely manner in order to promote a broad academic and public discourse.

6. The data sets, program codes and any further documentation used should be made available so that other researchers can independently replicate and verify the results, subject to the appropriate data protection requirements.

7. Should current data protection legislation inappropriately obstruct the use of data for evaluations and the replication of such studies, a revision of the political and legal framework needs to be discussed.