



# Why are there so few hard facts about the impact of cluster policies in Germany? A critical review of evaluation studies

Michael Rothgang<sup>1</sup> · Bernhard Lageman<sup>1</sup> · Anne-Marie Scholz<sup>2</sup>

Received: 31 March 2020 / Accepted: 1 January 2021 / Published online: 11 March 2021

© The Author(s) 2021

## Abstract

Based on empirical findings on the effects of cluster policies in Germany, this paper scrutinises the available knowledge on cluster policies impact. There is a growing body of insights on direct effects of policy measures on cluster actors, cluster organisations and innovation networks of the promoted clusters. For some industries such as biotechnology, there are indications that cluster policies had a substantial influence on the formation of new firms and emerging sectoral structures. While the available information seems to support the hypothesis that cluster policies can provide positive impulses for the development of clusters, the actual knowledge on far-reaching impacts of cluster policies on economic structures and processes is still rather limited. The paper asks for the reasons of this knowledge gap between expectations placed in cluster policies and the available evidence on their impact. We identify five reasons: (i) problems in addressing the systemic nature of cluster policy interventions and their effects, (ii) deficiencies regarding the methodologies used, (iii) a lacking informational basis, (iv) practical contexts (e.g., a lack of interest of policy makers) leading to deficiencies in incentive mechanisms and (v) the limited transferability of evaluation results to other cluster policy contexts. For future evaluations, we propose among others the use of system-related approaches to impact analyses based on mixed-method designs as well as comparative case studies based on new methods like process tracing. In order to improve the incentives for evaluators, an increasing awareness of policy makers about the relevance of evaluation studies would be important.

**Keywords** Cluster policy · Policy impact · Evaluation · Methodology · External validity

**JEL classification** O 21 · O38 · R 58

---

Why are there so few hard facts about the impact of cluster policies in Germany? A critical review of evaluation studies

---

✉ Michael Rothgang  
michael.rothgang@rwi-essen.de

Extended author information available on the last page of the article

## 1 Introduction

In recent years, the calls from many areas of public policy for a rigorous evidence-basing have been getting louder (e.g. Haskins and Baron 2011; Boockmann et al. 2014; Burda et al. 2014). Cluster policies are not exempt from this general trend. As cluster policies have been in place for three decades in most developed countries, a considerable body of knowledge about the impact of these policies should have accumulated by now. While some of the more visible cluster programmes that receive a lot of public funding have become the object of evaluation studies, we still do not know much about the actual impact of cluster policies.

This critical review uses Germany's cluster policy to investigate what is known about the impact of cluster policy, and to determine why knowledge gaps exist.<sup>1</sup> While most relevant publications address the issue of knowledge gaps rather casually, some studies try to assess the current state of research on cluster policy impact. The authors of these studies, who deal more thoroughly with the issue of impact, are sceptical about the available knowledge on the effects of this type of policy intervention in market processes (Uyarra and Ramlogan 2016; Fornahl et al. 2015; Andersson et al. 2004; Kiese 2008, 2017; Fromhold-Eisebith and Eisebith 2008a, b; Lindquist et al. 2013). This leads to the suspicion that there is a gap between the high expectations placed on cluster policies and the actual knowledge about their impact.

Germany is well-suited for researching cluster policies because scholars can refer to in-depth information about these interventions. There is also a considerable range of different types of cluster policies in Germany that have been applied over time. In fact, Germany has a roughly 25-year history of diverse cluster programmes developed under the aegis of federal or state governments. In this context, it should be possible to arrive at a clear understanding of the relationship between the expected and the realised impacts of these policies. In addition, because Germany's portfolio of cluster policy measures is, in many respects, similar in structure to the cluster policies in most EU member countries, the results of studies dealing with the impact of cluster policies in Germany should allow comparisons with the experiences collected elsewhere.

The term "cluster" has attracted the attention of policy makers and researchers from a variety of disciplines (cf. on the remarkable rise of the cluster narrative, Vicente 2016: 4). As a result, the term is now open to different interpretations and definitions. In this paper, we rely on the short version of Porter's rather general, but among researchers widely accepted definition:

"A cluster is a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities." (Porter 1998: 199)

While a cluster is marked by certain structural characteristics, the term "cluster policy" refers to action taken by government agencies to influence the development of clusters. Over the past decades, an increasingly wide range of different policy approaches has

<sup>1</sup> First ideas for this paper were presented in Vienna in a poster at the Austrian EU Council Presidency Conference on "Impact of R&I Policy at the Crossroads of Policy Design, Implementation and Evaluation" on the 5th and 6th of November 2018.

been denoted as cluster policies (Lindquist et al. 2013; Fromhold-Eisebith and Eisebith 2008a). For this reason, we will use the plural term (similar to Kiese 2012, 2017; Lindquist et al. 2013). Consequently, for the purpose of our paper, a cluster policy is any programme that is either primarily directed towards the promotion of industrial clusters, or at least contains an explicit component of cluster promotion in the framework of another policy area such as technological, industrial or regional development.

The literature dealing with clusters presents a variety of perspectives. For example, Fornahl et al. (2015) take a look at the rationale and practical concepts of cluster policy, whereas the most recent article by Kiese (2017) focuses on the impact and evaluation of German cluster policy. The studies of Kramer (2008), Wessels (2008) and Schmiedeberg (2010) deal with the practical or methodological challenges of cluster policy evaluation. Each of these papers mentions specific issues found in cluster policy evaluations. Uyarra/Ramlogan (2016) take an international bird's eye view of the impact of cluster policy based on a review of existing studies. This contribution covers many of the issues already discussed by other authors to assess the strengths and weaknesses of cluster policy impact evaluations and explores the causes of identified deficits.

The structure of our paper is as follows. Section 2 addresses our underlying understanding of “impact” and discusses why impact evaluations are important. Section 3 introduces some basic facts about cluster policies in Germany. We also sketch the portfolio of relevant evaluation studies and the studies we selected for this analysis. Finally, this section also presents what our analysis reveals about the impact of cluster policy identified in these studies. Section 4 discusses reasons for the present unsatisfactory state of knowledge on cluster policy impact. Section 5 summarises and discusses starting points for the advancement of cluster policy evaluations.

## 2 The challenge of evaluating the impact of cluster policies

### 2.1 The concept of “impact”

In this paper, we use the term “impact” to describe the results of cluster policy interventions, regardless of the level of effectiveness. Impact can refer to any significant influence on a cluster-related entity, whether it be a project family, a cluster firm or research organisation, a whole cluster, a region, a technology or an economy. In this way, impact refers to all subsequent stages following the implementation activities, preferably “final” stages, of longer chains of effects triggered by the original policy intervention and can include significant unintended results. In impact evaluations, “final stages” can only be defined pragmatically because the effects of a support impulse can reach far into the future, as in the case of the establishment of technological or spatial path dependencies.

Our usage of the term is compatible with its use for the last four stages in the schematic sequence of programme effects applied in evaluations of EU programmes: resource inputs > activities > outputs > results > intermediate impacts > final impacts (European Commission 2004: 11, 2006:7). This logical framework is often applied in the analysis of complex programmes (see also W.K. Kellogg Foundation 2004).

Following a common practice in the research literature, we could use the terms “impact(s)” and “outcome(s)” interchangeably (for the widespread confusion in the use of these terms in the evaluation literature see Belcher and Palenberg 2018). Faced with the choice between “impact(s)” and “outcome(s)”, the usage of both terms in the Merriam-Webster Online Dictionary<sup>2</sup> would suggest “impact” for our subject associating its use with direct, striking and forceful influences of a matter A on a matter B.

In the recent econometric literature, the term “impact” is sometimes exclusively used to denote direct causal effects of programmes that were analysed using RCTs (randomised controlled trials) or quasi-experimental research designs (e.g., Gertler et al. 2011; Haskins and Baron 2011; Boockmann et al. 2014). However, this would automatically exclude all causal statements that are based on other methods. Using “impact” in this sense also restricts its application to a few selected dimensions of cluster development for which statistical data are accessible.

One of the reasons why there are competing definitions of the term “impact” is because there are competing interpretations of causality (e.g., Belcher and Palenberg 2018 cite “systemic causality”). Another reason is because the term is used to describe different stages of longer effect chains. Our use of the term is identical to that used by Uyarra and Ramlogan (2016) in their meta-analysis of studies on cluster policy and the “Handbook of Innovation Policy Impact” (Edler et al. 2016), in which their contribution was published.<sup>3</sup>

In our discussion, we also use the term “effects”. The Merriam-Webster Online Dictionary subsumes under this term “something that inevitably follows an antecedent (such as a cause or agent)”.<sup>4</sup> We define cluster promotion intervention “effects” as “the totality of results that are triggered by the initial impulse of the intervention”. This definition implies, similar to the impact definition, a multi-level perspective on programme results. Effects may occur on the project level, the level of cluster initiatives and/or clusters, or at the level of regional or national economies and in technological evolution. Thus, in our usage of terminology, the concept of “effects” always includes impact(s) on all levels of a cluster programme and all stages of the effect chains. Additionally, this delineation also includes the practical steps in a programme’s implementation process that are not classified under “impact”.

Regardless of how impact(s), outcome(s) or effect(s) are interpreted, or whether “hard” or “soft” research methods are applied, the fundamental idea behind evaluation studies is the *counterfactual situation*. A study attempting to identify and evaluate the impact of an intervention must tackle the question: What would the present situation be if this intervention in the market process had not taken place? From this perspective, the object of interest is the “additionality” observed in the present situation triggered by the impulse of the intervention (Bauer et al. 2009). Because the counterfactual situation is always hypothetical and cannot be observed, the researcher has to apply a convincing method to identify the underlying causal links. The standard econometric method used to solve this problem is

<sup>2</sup> <https://www.merriam-webster.com/dictionary/outcome> (resp. impact), accessed on 16 January 2020.

<sup>3</sup> In the recent economic literature, a very generous use of the term “impact” can generally be observed. For example, Shiller (2019) as well as Case and Deaton (2020) use this term very broadly.

<sup>4</sup> <https://www.merriam-webster.com/dictionary/effect>, accessed on 20 August 2020. At the same time, this expression is used to connote numerous other meanings that do not apply to our context.

RCTs, mentioned above. But the question has to be addressed, whether and in what instances the assumptions that underly the application of RCTs are valid when evaluating complex programmes. While econometric methods are, in principle, superior for identifying programme effects, problems arise in their application. These problems are caused not only by data quality and availability, but also especially by the complexity of the effects created by policy impulses, and the problems of attributing the results to influencing factors. Thus, the question arises, what the application of different methods (quantitative and qualitative) either separately or in combination can contribute to our understanding of causal mechanisms in the evaluation of cluster policies. Against that background, both the use of the relevant terms and the methods used inevitably evoke paradigm debates in social and economic research.

## 2.2 Why do we need impact evaluations?

In recent years, industrial countries are placing more emphasis on evidence-based policies. Structural policies, including cluster policies, no longer have the luxury of escaping scientific scrutiny. While the call for impact analyses now includes structural policies, there are certain complications that are intrinsic to cluster policy evaluation projects that we will address below.

Different kinds of cluster policy evaluations include *ex ante*, accompanying, *ex post* evaluations, process analyses of implementation processes and participatory evaluations. All of these evaluation approaches play specific roles in the evaluation of cluster programmes (Diez 2001; Pawson 2013, Andersson et al. 2004: 119). Impact evaluations are only one of several different complementary evaluation approaches. There are many cluster policy evaluations that do not place the final programme impact at the centre of the study.

Yet, whatever methodological or institutional difficulties arise in developing promising evaluation approaches, the call for evidence-based cluster policies is widely shared by practitioners and researchers who are active in this field. The presence of difficulties does not justify the renunciation of the analysis of cluster policy impact. The more information we can gather on the impact of policy measures will increase the policy's legitimacy and improve the ability of policymakers to develop appropriate policy designs. In sum, impact evaluations should fulfil the following tasks:

- Determine what works and what does not work in both large and small cluster policy interventions, and enable policy makers to make choices from the portfolio of conceivable cluster policy approaches in a trial-and-error process.
- Increase our knowledge about the relative suitability of cluster approaches to achieve long-term political objectives.
- Evaluate how efficiently resources provided by cluster policies are used, and draw comparisons with the use of resources in other areas such as regional development policy, technology policy and industrial policy.
- Provide results that are either positive or negative with respect to the effectiveness of different cluster policy approaches, in a way that will increase the available knowledge on these policies.

## 3 Cluster policies in Germany: policy area, evaluations, impact

### 3.1 Policy area overview

According to the constitutional rules of Germany, the federal states are supposed to be responsible for the promotion of clusters that are situated in their territory, while federal institutions are expected to concentrate on matters of national importance. In practice, abstract constitutional rules do not clearly reveal the relevant delimitations of federal and state engagements in cluster policy. Prominent industrial clusters are spatially bound to either the territory of a single federal state, or in some cases may spill over into the territory of neighbouring states. Given this reality, the definition of national relevance might be controversial.

Cluster policy is implemented by institutions at all levels of public administration: the federal republic, the federal states, regional administrative authorities and local authorities. The European Union also contributes to relevant state programmes. Thus, cluster policy in Germany is a multi-level enterprise that is not hierarchically organised. Although this suggests a high degree of coordination between the relevant actors operating at the different administration levels, there is no close coordination between the responsible authorities, nor a “master” cluster policy plan.

Hence, German cluster policies operate in a diverse environment with a wide range of approaches that have more or less ambitious targets, and use different instruments. In fact, these policies are closely linked to other policy areas such as technological, regional or industrial policy. Yet, we can distinguish between pure cluster programmes and programmes that only feature a cluster component. In this context, the affinity of technological cluster and network programmes is particularly striking. Both the cluster and the network components merge seamlessly into one another.

### 3.2 Cluster policy evaluations

Our discussion is based on two categories of research on the economic impact of cluster policy measures: (i) Research that is commissioned by a responsible authority and is carried out by external independent research institutions, and (ii) Academic research that primarily pursues an academic purpose and is intended for publication in scientific journals or books. In the case of German cluster policy evaluations, many of the published purely academic studies are based on commissioned studies and rely on the data collected in the course of these commissioned studies. Because the generation of relevant datasets is time-consuming and expensive, scholars interested in researching cluster policy have no alternative but to refer to the relevant research data generated by commissioned studies.

While federal cluster programmes have received a great deal of attention from the German cluster research community, state programmes have received far less attention. An exception being the *Cluster Initiative Bavaria*, which received substantial financial resources, and has been analysed not only by a commissioned evaluation project, but also by a free academic research project (see Section 3.5). Each of the 16 federal states manages its own cluster programme, or an explicit cluster component that may be part of a larger structural policy program. Both cases, independent cluster programmes and cluster components in broader programmes, are included in the cluster policies in our

delimitation. If financial resources from EU structural funds are used, which is normally the case, the responsible state ministries (mostly ministries of economic affairs) are obliged to commission external analysts to conduct evaluations.

We have identified four major groups of empirical studies:

- (1) Comprehensive commissioned evaluation studies that analyse the impact of policy interventions of cluster programmes in the framework of an ex-post or ongoing evaluation (e.g., Rothgang et al. 2014a, b).
- (2) Studies that use a set of relevant regional data to conduct an econometric analysis based on a few core indicators to determine the overall impact of a policy intervention (Lehmann and Menter 2017; Falck et al. 2008).
- (3) Studies dealing with selected aspects of cluster policy impact using methods of econometric causal analysis (e.g., Engel et al. 2017; Nestle 2011).
- (4) Econometric analyses of specific aspects of the impacts of promoted clusters, such as the development of ensuing innovation networks (e.g., Graf and Broekel 2020; Cantner et al. 2013; Wolf et al. 2017), or studies that use qualitative methods to analyse special dimensions of cluster development, for instance, the importance of cluster initiatives for regional economic development (Kiese and Hundt 2014).

In light of the substantial heterogeneity of cluster policies, it is clear that a uniform methodological approach to the analysis of the impact of cluster policy measures would be entirely inappropriate. Financially limited interventions, which influence promoted clusters to a hardly identifiable degree, coexist with comprehensive programmes that are expected to have substantial impacts in several dimensions, including: the evolution of clusters, cluster firms, cluster regions, promoted technologies (in the case of technology clusters), and in the spatial and sectoral structures of an entire economy, such as employment, economic growth and welfare.

The heterogeneity of cluster policy programmes implies the use of specific evaluation tools that must be tailored to specific situations. Programmes that attract a lot of funding are also more likely to receive funds designated for evaluation, and also to attract the interest of academic researchers. So, it is not astonishing that in the research literature, there is a bias in favour of larger programmes. Our analysis also focuses on the evaluations of larger programmes, because they have a better chance of leaving a permanent footprint on economic structures. Nonetheless, problems that evaluators have to deal with are similar for large or small programmes, and the issue of programme heterogeneity does not affect our findings in a significant manner, because the programme evaluations we included in our analysis have significant commonalities. Two identifying characteristics are the promotion of cooperative research projects and the intermediary role of cluster initiatives, often set up specifically to enable cluster actors to participate in the programme. Programmes that do not share these characteristics are not considered in our paper for impact evaluations.

### 3.3 Selection of evaluation studies

Our contribution focuses on the analysis of the impact of cluster policy in Germany, i.e., it deals with a narrow segment from a much larger body of literature. A recent



Google Scholar search revealed 4260 titles when the keywords “cluster policy” and “Germany” were chosen together. This figure is deceptive, because although all data sets from the search contain the keyword “Germany”, the vast majority of these studies deal with cluster policy in general or with the cluster policies of other countries, citing Germany as a reference country in the text. There are also many double counts because working paper versions of published articles are also recorded. Using Scopus as the search engine with the same keywords reveals substantially fewer titles (28), making the number of academic publications that directly address cluster policy issues more manageable.<sup>5</sup>

Although we do not claim to have included all the relevant work in our analysis, we have tried to present a representative cross-section. Our search included the Scopus database, the Internet, and the reference lists of the available studies that address related questions. We considered both academic works and contract research studies, and have ensured that the key methodological approaches used in impact evaluations are represented.

Other criteria such as year of publication, type of cluster policy evaluated or organisational affiliations of authors did not play a role in the selection process. In view of the research topic, a strong disciplinary bias in favour of economics could not be avoided. Table 3 in the Appendix gives an overview of the selection. Several of the listed papers deal with the same programme but were included because they clearly addressed different aspects of impact. We would also have liked to have included cluster policy measures that are carried out below the level of federal states. Unfortunately, a comprehensive mapping, let alone systematic evaluation of these activities has not been undertaken (Kiese and Hundt 2014: 121).

Many of the reviewed academic publications on cluster policy impact are based on data and other research materials that the authors collected in the course of a larger commissioned evaluation project conducted on behalf of the responsible federal or, in the case of Bavaria, state ministry. Hence, contract researchers and academic researchers are often the same individuals. Exceptions to this rule are the cited studies of Falck et al. (2008, 2010), Lehmann and Menter (2017) and Nestle (2011). The link between commissioned evaluations and free academic research that uses the collected research material should not present a conflict of interest, because the commissioned evaluations were carried out by independent renowned research organisations (ROs) and universities, whereby the terms and conditions of the research contract mostly permit the free scientific use of the data collected by the involved researchers.

### 3.4 What we know about cluster policies impacts

We can classify our knowledge about the impact of cluster policies in a variety of ways. For example, the impact might be time sensitive, with a short, medium or long-term time-horizon. We could also refer to the above-mentioned classification scheme used in evaluations conducted under the aegis of the European Commission ( 2004: 11, 2006: 7). For our purpose, a classification based on the level of impact—micro, meso,

<sup>5</sup> Google Scholar was accessed on November 30, 2020, Scopus on November, 26, 2020. The Scopus search referred to the fields “title”, “abstract” and “keywords”, search terms were (“cluster policy” OR “cluster policies”) AND “Germany”.



macro—is the most useful because it reveals the complex and systemic nature of state intervention in cluster development.<sup>6</sup> Our discussion does, however, also include aspects of time sensitivity and the logic of action.

It might seem obvious that the impact of cluster funding should be based on the success criteria and individual programme targets set by the funding governmental agency. In principle, evaluators should be able to measure and evaluate programme targets by using a variety of indicators. However, two factors suggest that the impact of a programme should not be determined only by the achievement of the success criteria defined by the programme designers. First, many cluster programmes, especially the more ambitious ones, operate in a field that is characterised by a high degree of uncertainty. This was the case, for instance, with the BioRegio programme that supported emerging biotech clusters (Graf and Broekel 2020; Dohse 2000).<sup>7</sup> Second, the success criteria stipulated in programme documents is often the result of a complex negotiation and decision-making process. In many cases, designing a programme based on operable indicators is not a high priority for the decision makers. Instead, programme documents sometimes contain vague, overly precise, contradictory and/or unrealistic formulations of programme objectives. Evaluating subsequent developments based on such targets is difficult, at best. Whatever the case, in order to gain an unbiased perception of the programme's impact, all programme effects on specific actors, as well as the effects on the economic and technological environment should be evaluated by taking the viewpoint of an actor external to the programmes.

Based on our analysis of evaluation studies, Table 1 presents not only our current state of knowledge, but also what we do not know. The middle column lists impacts that are addressed in the mentioned studies. References to the sources are indicated by the numbers in brackets that refer to Table 3 in the Appendix. The right-hand column lists impacts that are, for the most part, not addressed in the studies, but that from a systems analysis perspective could emerge at micro, meso or macro levels. Because cluster programmes simultaneously influence individual cluster actors, clusters as well as their environment, Table 1 also classifies programme impacts based on these three levels defined by our classification scheme.

### 3.4.1 Micro-level impacts

Under micro-level impacts, we summarise all effects triggered by the original intervention on the individual cluster initiatives, on the firms and research organisations that take part in the cluster initiative as individual organisations, as well as relevant firms and ROs that did not participate. The micro-level comprises all organisational units, including individuals who work in the organisation, as well as all research projects that are funded by the programme.

<sup>6</sup> Other related distinctions known from cluster research could be conceivable, such as the distinction between cluster-internal and cluster-external effects, or effects on horizontal and vertical relationships between actors. Due to the kinds of analysis found in the studies on cluster policy impact, however, the distinction micro/meso/macro in our view make it easier to systemise the results.

<sup>7</sup> Another example of an open-ended technology programme that bears structural similarities to technology-related cluster programs is the promotion of the DARPA agency described by Mazzucato (2014), from whose activities the Internet has emerged. When DARPA was founded, nobody would have seriously considered defining the development of an “Internet” as a success criterion for the agency.

**Table 1** What is known and what should be known about cluster policy impact in Germany

Impact on	Available empirical evidence on intended and non-intended impact <sup>1</sup>	Open questions <sup>2</sup>
<b>Micro-level Impacts</b>		
Cluster initiatives as catalysts of cluster actors' cooperation	Strong trigger effects on development of cluster initiative during the funding period, many cluster initiatives survive after the phasing out of the promotion program, activities of cluster organisations during funding period well documented in many cases (23, 29, 33)	Medium-term and long-term survival of cluster initiatives and their impact on the evolution of the cluster and its innovation networks
Organisational design and evolution of the cluster initiative	Information on the development of the legal forms, organisational structures and resource bases of the cluster initiatives (33)	Information missing on long-term survival record of the cluster organisations, power relations within cluster organisations and their relationships with single cluster actors (firms, ROs)
R&D expenditure of firms participating in joint projects	Participating firms raised their R&D expenses slightly, additionality proven in the case of R&D subsidies for collaborative projects (7, 14, 15, 16)	No substantial information about long-term effects on R&D expenditure
Supported research projects	Implementation as a rule successfully realised, information on project output such as patent applications, follow-up-projects, innovations (33, 38)	Missing information on impact of funded projects on firm output and performance as well on mid- or long-term evolution of research portfolio of firms
Incumbent firms (cluster firms which participate in cluster initiative)	In most cases, no relevant effects on firms that are formal members of cluster initiative, but do not benefit from sponsored R&D projects (33); behavioural changes of a part of firms participating in cluster projects, partially information on product and process innovation based on the sponsored projects (5, 6, 24, 25, 33)	No evidence with regard to medium-term and long-term behavioural effects on the performance of firms measured by indicators such as sales, productivity, technological capabilities
Research organisations (ROs)	Behavioural changes of ROs participating in cluster initiative, information on the evolution of the research portfolio of the ROs (33)	Information on the medium-term and long-term development of the participating ROs and possible influences of cluster promotion is largely missing, no systematic analyses
Increase of cluster actors' technological competence	Vast majority of projects carried out as planned, much information on progress, results and consequences (such as follow-up publications, patent applications or innovations) of promoted R&D projects (33)	Lack of information on the impact of sponsored projects on the technology base of the participating firms and evolution of the respective technology field
<b>Meso-level impacts</b>		
Short- and medium-term cluster performance (2–5 years)	Whereas the development of cluster initiatives is well documented in many cases, far less is known about the short and medium development of cluster populations (actors	Information about the long-term development of cluster population largely missing, available information exclusively concentrated on actors who participate in state-funded projects,

**Table 1** (continued)

Impact on	Available empirical evidence on intended and non-intended impact <sup>1</sup>	Open questions <sup>2</sup>
	participating or non-participating in cluster initiative) (33)	all cluster actors, i.e. also non-participating should be included
Long-term cluster performance (> 5 years)	Promotion of biotech clusters by several federal programmes spurred on the emergence and evolution of a prospering German biotech branch originally based on start-ups (8, 9, 10, 11, 38, 24, 38)	Apart from descriptive ex-post studies on biotech clusters no evidence of the long-term development of promoted clusters
National and international visibility of cluster	Strong evidence for increased national and international visibility in the case of federal cluster programmes (LECC), based on stakeholder and peer information (33)	Very limited knowledge about sustainability of image gains by cluster promotion
Start-ups in cluster/cluster area/-sector	Sporadic information on cluster-induced start-ups in the respective observation periods, no cross-cluster analyses (2, 33)	Information missing on medium-term and long-term development and survival records of cluster start-ups, relationships between start-ups and incumbent cluster firms
Relations between member firms of cluster initiative	Apart from joint projects which are restricted to a narrow circle of cluster firms, improved knowledge about other cluster firms, economic benefits very doubtful (6, 25, 33, 39)	No evidence about changes of mid- and long-term relations between members of cluster initiatives available
Innovation networks	Cluster promotion leads to the formation of new networks and the expansion of existing ones, changed structural characteristics of existing networks, specific information from selected clusters (5, 6, 24, 25, 33); positive effects of participation in cluster initiative on building of trust (32)	No information about the medium-term and long-term development of the affected networks, impact of extended and transformed networks on innovation performance unclear
Knowledge transfer between firms	Participation in cluster initiative favours open innovation among members in microsystems technology; non-members tend to prefer commercial solutions (28)	Reproducibility for other fields of technology; no knowledge on long-term impact of open innovation resp. commercialised knowledge transfer on technology and sectoral development
Technological evolution in regional/national economy	Impulses on technology development probable in some cases, anecdotal evidence (11, 25)	Missing knowledge about the long-term development of the relevant technology domain and influences from cluster promotion in critical phases of the development of this technology
Regional economy	Positive impact of cluster promotion on respective regions assumed by many authors, evidence on causal effects rather seldom (17, 18, 31.)	Knowledge about long-term impact of cluster policy on respective regions narrowly limited
Development of the relevant sectors	Apart from biotech industry (8, 9, 10, 11, 24) no systematic analyses available	Whereas long-term structural impact of cluster promotion on industrial sectors seem plausible from a theoretical point of view, this question has not

**Table 1** (continued)

Impact on	Available empirical evidence on intended and non-intended impact <sup>1</sup>	Open questions <sup>2</sup>
		been analysed in evaluative investigations
Macro-level impacts		
Effects on competing non-funded clusters	No systematic analysis available	Suspicion with regard to possible negative impact of cluster promotion expressed by some authors, evidence is missing
Interplay of different cluster policy approaches	Interplay of cluster and other policies as problem addressed (20, 26 27, 28)	In practice, different cluster programmes realised by different policy authorities support the same cluster sequentially or at the same time; the resulting total impact does not need to be the sum of the partial impact
Interplay of cluster policies with other fields of structural policy	Anecdotal evidence with regard to the relationships between cluster promotion and regional policy, first trials of quantification of effects on regional economy (17, 18, 19, 31)	Systematic analyses of the role and effects of cluster policy on industrial, regional and technology policy is missing
Efficiency of resource allocation	No studies	Comparisons with hypothetical alternative usages of applied resources in structural policies not available
National growth and welfare	Assumptions with regard to positive structural and welfare effects based on plausibility considerations (22, 25, 27, 28 31)	Exact knowledge on structural and welfare impact of cluster promotion and relative cost-benefit balance compared with other measures of structural policy missing

<sup>1</sup> In the left column, selected results from individual evaluation studies are shown, which are important from the point of view of our research question. The numbers in brackets refer to the research works listed in the list in Table 3 in the Annex. <sup>2</sup> The right-hand column lists individual aspects of the possible (intended and unintended) impact of cluster policy, to which the available evaluation studies have so far provided little information. Empirical findings were mirrored against expectations regarding desirable influences of cluster support on the development of clusters, their actors and their economic environment, which can be found in the theoretical cluster literature and have often been incorporated into programme documents on cluster support. Important sources in this context are Alecke et al. 2008; Andersson et al. 2004; Audretsch et al. 2016; Brenner et al. (2013); Brökel and Graf 2020; Crass et al. 2016; Eickelpasch et al. 2004; Falck et al. 2011; Ketels 2014; Porter 1998, 1999; Santner 2018; Segers 2016; Sölvell and Williams 2013; Uyarra and Ramlogan 2016; Vicente 2016

Evidence for the influence of a cluster policy can be found in the establishment of cluster initiatives that represent a group of related firms and research institutions. The actors in these cluster initiatives work together to define joint strategies, promote and sometimes implement joint projects. The initiative is supposed to represent the cluster vis-à-vis third parties, in particular public authorities. While information is available on the short-term development of cluster initiatives, information on their medium-term and long-term developments is scarce. There is a fair amount of information about the relative success of cluster projects, especially R&D projects that have been co-financed by cluster programmes. This information gives the impression that the relative success or failure of a sponsored project, especially

of R&D projects, depends heavily on the level of risk connected with the respective projects. Studies dealing with the medium-term and long-term effects of cluster promotion on sponsored projects, on follow-up projects, firm sales and innovations are largely missing.

The information about the impact of cluster projects on participating firms is limited. While there is some evidence that cluster projects have a positive influence on patent applications, sales and product and process innovations of participating firms, there is little evidence for medium-term and long-term influences on firm evolution. Some econometric findings suggest positive influences of cluster programmes on firm performance, but the assumed causal relations deserve further research. A general finding of evaluative works is that cluster programmes have a stronger impact on SMEs than on large enterprises.

An area that deserves more attention in the empirical research is the short-, mid- and long-term impacts of programmes on participating research organisations, namely universities and independent research institutes. There is some evidence for input additionality. While there is some evidence that the establishment and activities of cluster organisations that include ROs would not have taken place without the support, the findings about the funding of promoted R&D projects are less clear. It appears that programme initiatives, on average, do not crowd out private R&D investments. There is evidence that public project funding resulted in moderate increases of average firm R&D expenditure (Rothgang et al. 2014a, b; Engel et al. 2017). The results in terms of output additionality on the micro-level are less clear. There are empirical findings with respect to the programme effect on activities of cluster initiatives and partial findings about the output of research projects that otherwise probably would not have been realised, and to behavioural changes of cluster firms and ROs.

The question of output additionality is more complex than input additionality in that an additional effect is not necessarily linked to a net gain in terms of programme objectives. It is difficult to determine if a successful R&D project would have been carried out without public funding at some point in time by another company without public funding. Cluster firms that do not participate in cluster promotion might develop just as successfully as those clusters firms that receive public funding. Then, the question is not just the observed output additionality, but to what extent this additionality is associated with a benefit.

### 3.4.2 Meso-level impacts

Under meso-level impacts of programme-induced influences on cluster evolution, we include: induced demographic changes in the size and structure of the cluster firm population, effects on the evolution of the technology that is the focus of the sponsored cluster, influences on the innovation networks within and around the cluster, effects on the regional and sectoral development as well as the international visibility of the cluster. The inclusion of “international visibility” among meso-level impacts may seem surprising at first glance, but an international comparison of the cluster allows evaluators to draw on the assessments of a global expert community.

The few systematic studies on the development of promoted clusters concentrate on selected sectors (especially the biotech sector), or deliver as descriptive case studies empirical finding on the analysed cluster. Some evaluations contain partial information on start-ups that were established after the onset of the programme. There is also

information on cluster populations (firms and ROs) that specifies the organisations participating in the cluster initiative as more or less active members. Relevant organisations that do not participate in the cluster initiative, although they are located in the cluster area, are rarely taken into consideration. Again, evidence dealing with the evolution of the sponsored clusters over several years is largely missing. An exception are two commissioned studies on the development of the biotech sector (Dohse 2000, 2007; Dohse and Staehler 2008). These studies try to establish causal links between the promotional impulses and observed development on the basis of methods of structural analysis.

Most available information on meso-level impacts is on the development of innovation networks triggered by the programme impulses. Studies based on social network analysis show that the co-financing of R&D projects leads to a significant expansion and structural changes of the innovation networks within the promoted clusters. There is evidence that cluster policies provide the opportunity for SMEs to participate in research networks, and lead to the central network actors playing a more pronounced role (Cantner et al. 2013; Töpfer et al. 2017).

The pioneering work of a recent publication (Graf and Broekel 2020) analyses the long-term effects of cluster funding on innovation networks. The authors analyse the short- and long-term effects of the BioRegio contest on innovation networks in the promoted clusters. They use biotech clusters that were not promoted under BioRegio, but received funds from sectoral biotech promotion programmes, as a comparison group. The initial impact on the networks of the BioRegio clusters was strong, but not sustainable. The long-term effects were similar in both groups with slight advantages in favour of the clusters supported by sectoral subsidies (Graf and Broekel 2020: 10-11). The study covers a central aspect of BioRegio support, although it does not claim to provide a conclusive overall assessment of the BioRegio contest, nor is it possible to generalise the results to innovation networks of clusters in other industries because of the different sectoral innovation systems.

Thus, in the case of output additionality on the meso-level, there is strong evidence concerning the extension and transformation of existing cluster innovation networks (Graf and Broekel 2020; Cantner et al. 2013; Rothgang et al. 2014a, b). However, the question remains whether the newly created networks always provide added value for the actors involved. This question is particularly justified if the participating companies were not free in the selection of their cooperation partners by programme regulations.

The available evidence also shows significant differences in the meso-level impacts that are particularly noticeable between different industries or fields of technology (Dohse; Fromhold-Eisebith 2014: 78). It remains unclear, however, whether these differences are primarily due to special features of the respective industry, to characteristics of the respective support programmes, or to other factors.

Information on the consequences of cluster promotion on the cluster's region is largely missing, apart from isolated, mostly positive, general statements in some evaluation studies. Although some econometric analyses based on accessible regional statistics and firm databases have given positive testimony to the respective aspects of cluster policies, the causal interpretation of the results remains unresolved.

### 3.4.3 Macro-level impacts

Macro-level impacts of cluster policies include: effects on sectoral development and non-funded clusters, the interplay of competing federal and state-level cluster policies, efficiency of resource allocation in all cluster programmes and the impact on national growth and welfare. Our knowledge about this type of cluster policy impact is still very limited. Therefore, it is hardly possible to make serious statements about the cost-benefit ratios of cluster policies compared with other structural policies (e.g. conventional industrial and technological policies).

At least, first attempts have been made, using econometric tools, to examine the effects of cluster policy at the macroeconomic level. Falck et al. (2010) conducted a research project on the effects and determinants of innovation in Bavaria that was funded by the German Science Foundation. Using a difference-in-difference approach, they analysed the “Cluster Initiative” established in 1999 by the state government of Bavaria. They found that this cluster policy programme increased the likelihood that the firms became innovators in the target industries by 4.6 to 5.7 percentage points. At the same time, the R&D spending of the firms decreased by 19.4% (Falck et al. 2010: 580). The authors attribute this decrease to companies’ improved access to external know-how (cooperation with public research organisations and access to suitable R&D personnel) as a result of the funding organisation.

In another paper, Lehmann and Menter (2017) investigate effects of the federal “Leading-Edge Clusters Competition” on regional GDP growth. Using panel data for the 150 German labour market regions, the authors applied a difference-in-difference estimation to evaluate the effects of this programme on the regions. They found that regional clusters might significantly increase the absolute regional wealth, and that cluster policy had a positive effect on regional economic performance (Lehmann and Menter 2017: 26). On this basis, they conclude that the support of clusters is “an efficient and adequate political instrument to foster regional wealth and corresponding regional endowment” (ibid.). These results seem to contradict the findings of Rothgang et al. (2014a, b) who identify a substantially smaller programme effect at the micro at meso level.

Despite the reported positive results, it remains unclear whether cluster policies fostered regional and national growth and prosperity, as well as technological progress, and how they compare with other structural policy measures. A larger number of studies would be necessary to confirm the general impact patterns found, preferably based on other approaches. The question of whether and how support for selected clusters affects the development of competing, unsupported clusters remains open. Furthermore, it is not surprising that there are no analyses of the economic efficiency of resource allocation in cluster support, as knowledge about its actual impact is limited.

The research literature provides diverse body of evidence on the overall impact of cluster policies. Apart from the authors quoted above, some researchers express reservations about the cluster concept in general, and the resulting policy approaches (e.g. Martin and Sunley 2003). The majority of cluster researchers give cluster policy a predominantly positive assessment, although criticism of individual aspects remains (Kiese 2008, 2012, 2017; Fromhold-Eisebith 2014; Fromhold-Eisebith and Eisebith 2008a, for an international perspective Uyarra and Ramlogan 2016; Lindquist et al.



2013; Ketels 2014). It should be noted that many of these statements are not only based on evidence, but also on intuitive assessments.

## 4 Why do we know so few facts about the impact of cluster policies?

The discussion in this and the following section relies on our analysis of the evaluation studies on cluster policy in Germany. Some aspects we highlight are rather practical in nature and, therefore, receive little attention in evaluation studies and are rarely considered in overall accounts of cluster policies. Nevertheless, these aspects are important for understanding the state of our knowledge on the impact of cluster policies.

### 4.1 Systemic nature of cluster policy interventions

The firms, research and educational organisations, public authorities and social institutions located in a cluster are related with each other in various ways by their participation in local business, information and knowledge exchange networks (Andersson et al. 2004; Bathelt and Glückler 2018; Porter 1998, 1999). The nucleus of a cluster is formed by an industry or group of related industries, including related service providers and public institutions. In large urban agglomerations, firm clusters of different industries coexist alongside each other, overlap and are mutually related to each other. What all clusters have in common, regardless of their industry/technology orientation, is that they form localised, more or less complex systems. Since clusters play an important role in innovation processes, it is appropriate to view them analytically as clearly identifiable innovation systems.

For this reason, cluster policies are ultimately intended to influence the development of a complex entity. Cluster policies might attempt to positively influence the dynamics of the cluster system by compensating for identified systemic weaknesses, often called “system failures”. These policies might also attempt to foster the development of new clusters, or use the system properties of clusters as a lever to promote a region or an industry.

In all cases, cluster policies do not primarily address single economic actors, but smaller or larger groups of interacting firms and related organisations. Single cluster organisations, firms or research organisations that receive government funding to carry out cluster-related projects, are involved in their capacity as members of a superordinate entity—a cluster or network—consisting of a group of interacting market and non-market actors. The expectation of cluster policy makers is that the impulse directed at the individual and responsible actors will capture the greater collective behind them to produce a welfare effect. Thus, a systemic component is naturally present in the “genes” of cluster policies.

The systemic character of cluster policy impact has certain consequences for its measurement (Rothgang et al. 2017). Not only are multiple indicators needed to identify different kinds of impact, changes in impact indicators created by the trajectory from the initial policy impulse to final results are also influenced by emergent processes. There is also substantial *ex ante* uncertainty about the actual impact of a cluster programme. Furthermore, we can identify substantial time spans from the initial policy impulse to a possible observed programme impact.

The spectrum of possible measures bearing the cluster policy label ranges from easily manageable measures that promote a cluster organisation and its management, to ambitious cluster programmes aiming at several simultaneous targets and involving considerable financial expenditures. This latter type of cluster programme can be quite complex, as it consists of many individual measures with a specific impulse that are combined under the umbrella of the programme. Some individual cluster policy measures can be rather simple and inexpensive, even if the underlying larger cluster narrative aims at a complex cluster entity. Examples are using public funds to set up cluster-related discussion groups, or finance additional training programs for cluster managers.

The question of which systemic impacts arise from an individual cluster funding project, or an individual cluster programme, can be answered by more or less ambitious investigation procedures. Sometimes, even simple intuitive assessments of the systemic relevance of a programme based on available evidence and past experiences with similar cluster programmes may be helpful. Demanding resource-intensive systemic investigations are only appropriate for ambitious and well-funded programmes.

Although evaluations generally recognise the systemic character of a cluster, they usually concentrate on only a few selected aspects of cluster development and performance (Uyarra and Ramlogan 2016: 225) that are accessible on the basis of the available information and data. In contrast, the systemic nature of clusters calls for a holistic approach, or at least a holistic perspective.

#### 4.2 Methodological toolkit

Given the systemic nature of cluster promotion, cluster evaluators must determine the most appropriate methodological toolkits to use in their evaluation of cluster policy impacts. Cluster policy evaluations use the whole range of quantitative and qualitative methods typically used to evaluate public programmes (see for a general presentation, Schmiedeberg 2010). Recently, clients of contract studies at the national and European level have been attaching more importance to quantitative methods, and have an affinity for RCTs. However, in Germany, conventional methods (both quantitative and qualitative) dominate.

Based on our observations about the systemic character of cluster policy impact, the methods used in cluster policy evaluation should address at least one of the following aspects, or a combination of any number of the listed aspects:

- The methods should address the causality problem in a suitable way.
- The methods should be able to deal with the suboptimal data availability, and situations where only a few observations are available.
- The applied methods should, in their combination, be able to deal with the multidimensionality of programme impulse and outcomes.
- The methods should be suited to increase our knowledge about the transmission processes from the initial policy impulse to programme impacts.
- The methods should consider the time-sensitive nature of programme impacts, and use indicators to measure an impact when it might actually be expected.

Any assessment of the relative strengths and weaknesses of different methodological instruments should consider the diversity of the cluster programmes discussed above.

The literature recognises the need for program-specific research designs (Kiese 2008). Hence, it is not astonishing that many authors favour a mixed method design that combines quantitative and qualitative methods tailored to the specificities of the evaluated programme (Andersson et al. 2004; Fromhold-Eisebith 2014; Kiese 2008, 2017; Fornahl et al. 2015: 95; Sölvell and Williams 2013). Larger contract research evaluations are always confronted with the task of tailoring the mix of methods to the characteristics of the programme. The evaluation of the Leading-Edge Cluster Competition (Rothgang et al. 2014a, b) and, although not explicitly designated as mixed methods design, the evaluation of the BioRegio and BioProfile programmes (Staehler et al. 2006) are examples of this practice.

There is a strong argument for combining quantitative and qualitative methods. Both methods have specific strengths and weaknesses that, to some extent, complement and counterbalance each other when they are combined (Plano Clark and Creswell 2008). As Kuhlmann (DeGEval 2016a) argued, it would be wrong to establish a hierarchical order among research methods e.g. that quantitative methods are always superior to qualitative ones. When different methods are applied to examine the object of investigation, a variety of perspectives permit the merging of findings through triangulation (Flick 2008). These synergistic effects might bring us closer to achieving our goal of increasing our knowledge about impacts. However, the mix of methods cannot compensate for deficits in the available portfolio of methods.

Table 2 outlines the strengths and weaknesses of the prevalent methods of data collection and analysis used in cluster evaluation. Most of the methods outlined in Table 2 are also the focus of recent methodological discussions conducted by German (DeGEval) and Austrian (fteval) evaluation organisations. The concrete mix of methods used varies greatly from evaluation to evaluation. The choice of methods in contract research depends on the contractually negotiated conditions, the size of the resource base available and the methodological competencies and preferences of the evaluators.

The methods outlined in Table 2 do not include required groundwork activities, such as project administration, organisational preparation of surveys (e.g., initiating contacts with the cluster initiatives) or collection of baseline data in the absence of an *ex ante* investigation. In some cases, the line between quantitative and qualitative analysis blurs. For example, the use of survey data in an econometric analysis requires the previous successful solution of the qualitative tasks. Here, however, the qualitative component is only a preparatory tool of a quantitative analysis approach, not an equal counterpart in a triangulation of different perspectives on the object of investigation.

As the assessment shows, each of the methods used has its individual weaknesses and strengths. Thus, evaluative works on cluster programmes should not be based solely on the available repertoire of rigorous econometric methods suited to identify causal relationships. At the same time, qualitative methods like case study research can lead to overly optimistic assessments. Even if the notorious deficiency of information on the long-term development of promoted clusters could be overcome, which seems improbable, the complexity of the effects triggered by the more ambitious cluster programmes would be too substantial to be grasped by the use of the present quantitative or qualitative state-of-the-art methods. Section 5 will discuss some additional methods that could supplement the methodological toolkit that is presently in use.

**Table 2** Strengths and weaknesses of the current portfolio of methods applied in cluster evaluations

Methods	Strengths	Weaknesses
1 Methods of data collection - apart from the use of relevant official statistics and research data bases -		
Monitoring of the activities of the cluster organisation	The establishment of reliable monitoring systems is a prerequisite for solid evaluation work	In most cases, evaluators' resources are too limited to fulfil this task, self-monitoring of cluster initiatives is only of limited scientific value
Standardised surveys of cluster actors or cluster organisations	High participation ratios probable as long as addressees (firms, ROs) receive subventions, easy to use instrument, which is suitable to collect a lot of information, possibly on a repeated basis	Cluster actors who do not participate in the cluster initiative hardly accessible, information affecting core interests of the organisation reluctantly transmitted
Stakeholder interviews with cluster actors	Access to deep information that would not be accessible in other ways	Stakeholders benefiting from the programme interested in a positive evaluation of the results
Interviews with programme designers and responsible authorities	Access to insider knowledge, which can be helpful in reducing information asymmetry between programme managers and evaluators	Denial of access to critical aspects of programming and implementation
2 Qualitative methods of data analysis		
Document analyses on cluster-related written information	Indispensable source of information on cluster initiatives, esp. with regard to the baseline assessment in cluster promotion	Idealised self-image of the "cluster" seen with the eyes of cluster initiative organisers, selection bias of delivered fragmentary information
Comparative analyses of interview results	No comparable specific information can be obtained through document analysis, statistics or own surveys	It is difficult to avoid subjective elements in the analysis, competent interpretations require skill and know-how on the part of the evaluators
Cluster case studies	Especially useful for the analysis of long-term cluster evolution, provided that baseline information and data about cluster development are available	Quality strongly dependent on the degree of investigation efforts, thorough case study very resource-intensive; results usually not transferable to other models of cluster support and other cluster sectors/markets. Quality of the message on causal effects cannot be verified, dependent on the accuracy of the picture shown
3 Quantitative methods of data analysis		
Benchmarking on level of clusters or cluster subsystems	Multiple application possibilities at levels of clusters, cluster initiatives and cluster actors, comparisons of different cluster programmes	It is difficult to find suitable benchmarks among other promoted or non-promoted clusters
Social Network Analysis (SNA)	Indispensable method for the analysis of innovation and other cluster networks	It is difficult to convince network actors to participate, no assessment of performance aspects of networking possible

**Table 2** (continued)

Methods	Strengths	Weaknesses
Randomised controlled trials (RCTs)/Quasi-experimental trials (QETs)	Method of choice, provided the conditions for its application are met	Conditions for application not given regarding central questions of impact evaluation; can only address some (rather few) effects; limited data availability; methods not applicable when there is a small number of cases
Other methods of statistical analysis	Multiple application possibilities, often in close connection with qualitative methods through triangulation Regional/sectoral/cluster-related statistical models	Many statistical methods do not fulfil requirements for rigorous causal analysis Suitable for comparative study of the development of clusters and cluster regions, where appropriate data are available
Results depend on a wide range of premises		
Cost-benefit analyses	Important instrument for comprehensive analysis of cluster programmes and comparisons with similar actions of structural policy	In order to obtain results, the use of strict premises is necessary

Sources: Ad 1: Methodologically and practically oriented literature on cluster policy and its evaluation (Schmiedeberg 2010; Lindquist et al. 2013); methodological sections of evaluation studies (for instance, Staehler et al. 2006); experience from the evaluation of various innovation policy programmes (e.g., Rothgang et al. 2014a, b)

Ad 2: Methodological literature (Kelle 2007; Plano Clark and Creswell 2008; Flick 2008; Lamnek 2005; fteval 2004)

Ad 3: Methodological literature (Bauer et al. 2009; Gertler et al. 2011; Haskins and Baron 2011; fteval 2004)

### 4.3 Informational basis

Under ideal conditions, evaluators of cluster policies would have full access to all relevant data on the activities directly triggered by a cluster policy intervention. These data would include time series data on the clusters and the cluster actors (firms and research organisations participating or not participating in the cluster initiative), data on realised innovations and patent applications as well as the economy of the cluster region. With this informational base, a wealth of baseline, process and “final” stage data on the promoted cluster and the market actors operating in its framework would be available to the evaluators. In reality, the situation for evaluators is quite different than this ideal situation (e.g. the descriptions of the informational situation in Kramer 2008; Dohse and Staehler 2008; Eickelpasch and Pfeiffer 2004; Fromhold-Eisebith and Eisebith 2008a, b). Evaluators are left with no choice other than to use the available information even if certain core dimensions of impact can only roughly be addressed. In such a situation, the evaluations give a rather incomplete picture of the different dimensions of programme impact.

Contracted evaluations rely heavily on the quantitative and qualitative information that is collected during the evaluation process in surveys and expert interviews. In some cases, evaluators develop monitoring systems to collect data delivered by the cluster organisation or the project agency (e.g. Rothgang et al. 2014a, b). Official statistics are often used when it is possible to do so with reasonable effort. In some more ambitious commissioned evaluation projects, researchers make use of scientific data bases like the worldwide patent database PATSTAT, the German federal funding data basis, the innovation panel data (CIS data), or the firm data base Amadeus to obtain information on the researched clusters and to construct control groups for econometric analyses (e.g. Engel et al. 2017). Some of the academic researchers using quantitative methods to investigate cluster policies create their own data bases or data collections during their involvement in commissioned evaluation projects (e.g. Cantner et al. 2013; Engel et al. 2017; Engel et al. 2015; Hinzmann et al. 2017; Engel et al. 2013; Fromhold-Eisebith and Eisebith 2008a, b; Dohse 2000, 2005, 2007; Dohse and Staehler 2008). Other researchers make use of information collected in academic studies (e.g. Kiese 2012; Nestle 2011; Falck et al. 2008, 2010).

Cluster actors are often reluctant to cooperate with evaluators, limiting access to internal firm data. This is the case even if the programme's promotion conditions explicitly require the provision of clearly defined data under specified conditions. While it is true that cluster actors are not an ideal source of information in the sense that they might make a too favourable judgment on the usefulness and the results of the cluster programme, their subjective appraisals are indispensable. In the absence of appropriate methods to determine the success of cluster promotion, the assessment of success or failure must necessarily be based on the judgments of stakeholders. Even if other methods are used, stakeholder opinions provide valuable supplementary and perhaps necessary corrective information (Lindquist et al. 2013: 7).

Summing up, the overall availability of information about cluster development triggered by the original promotion impulse is constrained by two key factors. The relevant data is scattered and difficult to compile efficiently. In addition, evaluators and researchers who choose to rely on insider information are often met with restraint by the core actors who dispose of important insider information. Both factors have an adverse effect on the general quality of evaluations.

#### 4.4 Practical contexts of evaluation studies

Essential for the understanding of scientific work is its context of origin. The first eye-catching fact about the research on cluster policy evaluations is that all comprehensive evaluation studies of public cluster programmes (e.g. Staehler et al. 2006; Dohse and Staehler 2008; Rothgang et al. 2014a, b), as well as a general assessment of cluster policy (Fornahl et al. 2015) have been produced in the context of official evaluation mandates. Many of the academic journal articles on the effects of German cluster policies were created as by-products of commissioned works, and used data collected in the course of commissioned evaluations. In practice, it is difficult to draw a clear line between commissioned studies and free academic work. In some cases, the connections are indirect and can hardly be discovered by outside observers when data from contract work forms the basis of subsequent scientific analyses.

Evaluation work in Germany, as in many other countries, is largely dependent on public financing and carried out by various types of independent research organisations or consulting firms on behalf of government agencies (for the general situation see, Astor et al. 2014). Among the contractors of commissioned evaluation projects, the main actors are publicly (co-)financed independent economic and innovation research institutes, a circle of private-sector consulting firms that specialised in innovation and technology policy topics and are active at either the national or European level, and national branches of global consulting companies.

The content of contract evaluations is always determined by agreed upon objectives and procedures, and is constrained by the available resources. The benefits of the evaluation should be in a reasonable proportion to the costs incurred (DeGEval 2016b: 40). Analyses are often limited because of a lack of resources. It should be noted that all contract research automatically brings a principal-agent problem into play (Schmiedeberg 2010: 391), which needs not diminish the scientific quality of such work, but must be taken into account and, although not absent in purely academic research contexts, is less acute there.

Contract evaluations deal with specific cluster programmes or include an assessment of the cluster policy component that is part of the program (as is the case in European structural policies). If the evaluation is financed by the European Union, evaluators are usually required to follow a uniform, standardised grid of indicators, research questions and methods. The European Commission justifies this approach by the legitimate wish to meet the challenges of enforcing minimum standards of evaluation and of producing comparable results for all member countries that take part in the respective European program.

The integration of evaluation research in rather narrow contractual frameworks has two consequences for the contents and course of evaluation processes: (i) The information wishes of the contractor determine the portfolio of research questions asked, and (ii) The procedure of the evaluation, not the least its timetable, is determined by the needs of the client. Thus, evaluators are required to deal primarily with questions of interest to the client who, in the case of programmes that are co-financed by the EU, are even obliged to work in the given framework of research questions (for the general approach see, European Commission 2004). The civil servants responsible for programme oversight have a strong interest in promptly receiving the analysis of the effects of the programme and information about its implementation. They are also looking for positive results. In programmes characterised by long-term effects, these requirements are in conflict with the researchers' objective possibilities to deliver solid information about the actual effects that are determined by the temporal logic of the effect chains of the programme impulse. The stipulated time schedule of the evaluation follows the requirements of the policy process and not the intrinsic methodological requirements of the analysis. This results in a tendency to take a short-term view on the programme. Results are expected before effects can actually be observed. Furthermore, the interest of decision makers in long-term retrospective evaluations of the impact of programmes that have been initiated under the responsibility of previous governments is limited.

Finally, a problem that complicates meta-analyses of commissioned evaluation studies of German cluster programmes should not remain unmentioned. In times of big data and the Internet, it should actually be easy to make all evaluation



studies commissioned by a ministry or another public organisation available in a freely accessible public database, regardless of the type and scope, or quality and timing of the products created. Unfortunately, this is not the case, because the federal and state governments have not yet been able to agree on the establishment of such a database.

#### 4.5 Limits of transferability of evaluation results

Given that there is no comprehensive ex-post analysis of a larger German cluster programme that would provide information on all essential aspects of the programme's impact, there are a number of individual studies that provide selective insights into the effects of individual programmes (e.g. Dohse 2000; Dohse and Staehler 2008; Eickelpasch and Pfeiffer 2004; Nestle 2011; Rothgang et al. 2014a, b). Yet, the extent to which the results obtained for one case, be it a programme, single cluster or policy approach, can be used to draw conclusions for other programmes, clusters or cluster policy approaches, is an open question. This question is especially pertinent given the conflicting general assessments of the usefulness of cluster policy measures in Germany that contain explicit and implicit judgements about cluster policy impact (positive, e.g. Dohse 2005; Falck, Heblich, Kipar 2008, 2010; Lehmann and Menter 2017; negative Fromhold-Eisebith 2014; Fromhold-Eisebith and Eisebith 2008a, b; centre position Kiese 2008, 2012, 2017).

On an abstract level, this question has been addressed in the methodological literature under the heading of the “external validity” of empirical findings. In the case of statistical analyses, the question is if it is admissible to generalise the results obtained from a sample of members of a population to the rest of the population. This requires, among other things, that the results identified in the evaluation can be replicated for groups beyond those evaluated (Gertler et al. 2011: 14). As Peters et al. (2015) show in their analysis of a comprehensive set of studies based on RCTs that were published between 2009 and 2014, the question of external validity was, in most cases, insufficiently answered or not considered at all. However, the programmes addressed in these RCTs are located in the areas of labour market policy, health care, social policy or micro-oriented development policy, and are aimed at individuals, families or households. Beneficiaries and comparison groups not subject to treatment can be selected from a relatively homogeneous population of individuals (families/household) and thus offer favourable conditions for RCTs.

In the case of complex cluster policy programmes, things are quite different. Cluster programmes tend to address complex units with pronounced individuality, for which there are at most a very small number of comparable counterparts that share noticeable structural similarities. In the context of qualitative studies, such as case studies in general, one should rather focus on the (partial) transferability of evaluation results to other specific cases. The stringency of statistical conclusions is excluded from the outset, hence, we are primarily interested in the transferability of evaluation results obtained for one cluster programme to other programmes (for external validity in qualitative research see, Onwuegbuzie and Johnson 2008). Even in cluster-related research contexts, where the conditions for statistical analysis are given, induction problems occur. For instance, Nestle (2011: 228, 233) warns against generalising his results.

The pervading problem of transferability/generalisability of evaluation findings in cluster policy results from individual characteristics of cluster programmes, cluster initiatives and clusters as well as external conditions:

- (1) Each cluster programme uses very different instruments, has different targets defined by programme designers, and has varying amounts of public funding. Cluster programmes also encounter a variety of industry, technological and market conditions. Experiences collected in one industry, such as biotech, might not be transferable to another industry, such as for example energy production.
- (2) When a cluster programme addresses clusters from different sectors, it inevitably encounters very different sectoral and regional innovation systems. What works well in one field of technology, may prove to be problematic in another (Fromhold-Eisebith and Eisebith 2008b).
- (3) Promoted clusters are in different phases of their development, as demonstrated and analysed in the research literature under the heading of “cluster life cycles” (Bergman 2008; Fornahl and Hassink 2017; Menzel and Fornahl 2007; Brenner and Schlump 2011). Thus, insights obtained for an emerging cluster might not be applicable to a mature cluster.
- (4) Cluster support presupposes the existence of cluster initiatives that act as intermediaries between state organisations and cluster actors. Success and failure of cluster support is also dependent on the capabilities and engagement of cluster organisers, facilitators, mediators and information brokers, who organise joint efforts and provide cohesion and trust-building among cluster actors.
- (5) There may be an imbalance of power between the participating actors in single cluster initiatives. For instance, the presence and relative weight of SMEs and large companies, of research organisations, or the possibility of a single actor playing a dominant role. These patterns, in respect to the distribution of power between different actors, can vary between different clusters, further complicating the transfer of experiences.
- (6) External circumstances, such as changing general political conditions or foreign trade factors, may exert decisive influences on the success or failure of a cluster promotion programme.

None of these difficulties is a sufficient reason for abandoning comparative studies on the effects of cluster programmes. Similar problems arise in all sciences that rely heavily on descriptive work. However, it is important that two elemental facts be considered in comparative studies of the impact of cluster promotion. First, it is apparent that success or failure of cluster programmes is to a large extent context-dependent. Second, the existence of diverse idiosyncratic factors, not only within cluster programmes, but within the very clusters themselves, leads us to the conclusion that contingency plays a decisive role.

The limited transferability of evaluation results obtained for one cluster programme to other programmes, or from one supported cluster to another within the same programme, is one of the pitfalls in research practice. To a large extent, this limited transferability is responsible for the observation that it is rather difficult to assess the general impact patterns of cluster policies. In the following section, the question of what options are available here will be examined.

## 5 Summary and discussion

Our analysis shows that the research literature on cluster policies in Germany delivers important insights on selected cluster programmes, and on single aspects of the consequences of their implementation. However, convincing answers to some of the most pertinent questions regarding the general impact patterns and practical benefit of these programmes have not yet been delivered. Our results are in line with the findings of other authors who focus on the German context (Kiese 2008, 2012, 2017; Fromhold-Eisebith and Eisebith 2008a, b, Fromhold-Eisebith 2014), as well as of authors who analyse cluster policy in general (Andersson et al. 2004; Uyarra and Ramlogan 2016; Lindquist et al. 2013).

Our critical review of the results of existing studies reveals that the availability and certainty of statements describing the impact of programme interventions decrease along three axes: (i) Knowledge on impacts becomes more fragmentary with the transmission path's increasing distance from the original policy impulses; (ii) The evidence on effects decreases as the complexity of policy measures and transfer mechanisms increases; (iii) Because the effects unfold over the course of time, it follows from (i) and (ii) that the certainty of knowledge decreases as the temporal distance of the progression of action levels increases from the original intervention.

We identified the following main reasons for the unsatisfactory state of knowledge about the impact of cluster policies:

- (i) While cluster policies aim at a positive long-term development of the promoted clusters, which are complex entities marked by self-organisation and emergence, the systemic character of the subject and design of the policy intervention does not find the degree of attention in evaluations that it deserves.
- (ii) The methodological toolkit used, in most cases mixed-method designs with different weighting of quantitative and qualitative investigative elements, neglects systemically oriented investigation components and favours the research of those aspects that indicate short-term effects, while medium- to long-term effects are only partially addressed based on easily accessible data.
- (iii) The informational basis of most evaluations shows significant weaknesses, in the sense that important data is not or hardly accessed.
- (iv) Evaluation research on the effects of cluster policies is heavily dependent on research commissioned by public administrations. Hence, evaluation studies follow the logic of the policy process, rather than being primarily driven by the intrinsic impulse of researchers to increase the knowledge stock on programme impact.
- (v) The substantial heterogeneity of clusters, cluster environments and cluster programmes renders it very difficult to transfer conclusions from one programme to another. This contributes significantly to the present incoherent picture of cluster policies in Germany.

The lack of a desirable *systemic perspective* for policy interventions in complex environments is not only typical for evaluations of cluster policy, but also in other policy areas, as the recent example of the British Medical Research Council shows (MRC 2019). Innovation and technology policy are policy areas that, due to the

characteristics of the innovation process, are suited for the application of systemic approaches. Corresponding methodological proposals are closely linked to the approach of national innovation systems (NIS). Thus, researchers refer explicitly to the NIS approach when they advocate systemic evaluations (e.g. Edler and Fagerberg 2017).

However, the practical implementation of systemic evaluation approaches is difficult. This is indicated not only by the observation that this call for systemic evaluations is anything but new (Georghiou 1998; Perrin 2002; Arnold 2004), but also by the criticism of the general approach and vagueness of the innovation system concept. The partial reluctance of evaluators to adopt a systems approach is a result of their tendency to rely on equilibrium-oriented neoclassical economics. The critical content of innovation system theory has subsequently been lost in the widespread use of the approach by the OECD and World Bank (Chaminade et al. 2018: 17-18). Recently, the call for systemic and holistic approaches in innovation research and policy has become more pronounced (Borrás and Edquist 2019; Edler and Fagerberg 2017; Edler et al. 2016).

Whatever the strengths and weaknesses of the innovation systems approach, it is possible to use qualitatively oriented system analytical procedures in cluster evaluation, both with a higher and a lower degree of granularity. A system-oriented, holistic analysis might begin with the consideration of very simple relationships in the cluster by using a limited number of relevant indicators (Arnold 2004). In this case, the system observation would be on the coarsest level of granularity and the lowest level of detail. This would allow a relatively straightforward benchmarking with the development of other clusters in the same sector or technology field.

Furthermore, as in other scientific fields, it should be possible to develop mathematical models that depict typical complex cluster constellations and allow conclusions to be drawn on the basis of empirical material about the influence of different determinants of cluster development. Of course, the resources required for evaluations in general and model building in particular should be related to the level of ambition and financial weight of the programme (Gertler et al. 2011). At the same time, the informative value of economic models should not be overestimated, as Bradley and Untiedt (2007) show in relation to models of European cohesion policy.

In terms of *methodological tools*, our analysis shows that cluster policy evaluations use a wide range of generally accepted and proven quantitative and qualitative research methods. As mentioned above, this should be complemented by instruments that are related to cluster systems. The possibilities of triangulating quantitative and qualitative methods (Flick 2008), which should actually be anchored in mixed-method research designs, have not yet been fully explored in most cluster policy evaluations. The opportunities offered by benchmarking, which can be applied at all levels of the clusters and their development, are also underused (see for a positive example, Dohse 2000). However, benchmarking is often used in the context of comparing the activities of cluster initiatives, and is promoted as a tool for improving the management of cluster organisations (cf., for instance, the work performed by Lindquist et al. 2013).

Case studies are a well-established and indispensable instrument of cluster research, which is qualitative in nature, but also contains quantifiable elements. Given the amount of research involved, it is understandable that the existing case studies mostly refer to snapshots of a cluster's status and previous development (e.g. Bathelt and von Bernuth 2008). However, case studies do not consider the long-term development of a funded cluster beyond the funding period, an approach that could provide insight into the role of cluster funding in the development of the cluster. In this context, the application of process tracing on cluster development could prove to be a fruitful investigative approach. This approach has its origin in cognitive psychology, and is gaining attention in political science (Bennett and Checkel 2015). Indeed, its application has been proposed for impact evaluation (Bjurulf et al. 2012), but, to our knowledge, has not yet been applied on a wider basis in cluster evaluation research. In contrast to panel studies, process tracing would not be subject to the methodological flaw that panel participants (companies, individuals) can hardly be kept in line over longer periods of time, and dropped out participants cannot simply be replaced (Nestle 2011: 235).

Also, the potentials of the use of quantitative methods (RCTs) have not been exploited so far. Future applications of RCTs should take into account existing evidence about the time patterns of cluster policy impact in order to measure each indicator at the right point in time. Due to uncertainties about information value and quality of data, the results of RCTs should be complemented by other methods in order to be sure that the actual programme effects were estimated.

In times of big data, it seems strange that cluster policy evaluations suffer from a *lack of relevant data*. In principle, this problem has two sources. The development of an adequate database involves substantial resources, and the owners of the relevant data (companies, research organisations, etc.) are not willing to share their data with evaluators. The use of official statistics to fill the gaps is extremely limited. The classification systems found in official statistics lag behind the reality of economic structures. Also, the most relevant types of data are not gathered, and official statistics cannot provide the required individual data due to current data protection regulations. The data problem in cluster policy evaluations, which always includes relevant, but inaccessible tacit knowledge, can even be seen in the availability of data required for solid baseline surveys. Therefore, researchers who want to estimate the influence of cluster programmes on promoted clusters, affected industries, regions or technologies on the basis of the scant available information often fall back on intuitive assessments that use all available information, but also inevitably include speculative elements. Some of the most informative studies on German cluster policy fall into this category (e.g. Staehler et al. 2006; Dohse and Staehler 2008).

The dependence of cluster evaluators on official research commissions is not necessarily a problem for the development of cluster research. Admittedly, the establishment of relevant funding programmes for research on cluster (or general structural) policy evaluations could essentially contribute to the advancement of evaluation research, but the possible scope for cluster policy evaluations in the

framework of commissioned research is not yet fully exhausted. It would be important to raise the awareness of the responsible decision makers for the particular importance of adequate time horizons of evaluations that should correspond to the length and temporary structure of the relevant chains of effect. This would automatically lead to a stronger emphasis on ex-post evaluations, including long-term retrospective studies.

The objective limits of transferability of evaluation results could only be overcome by the production of more evaluation studies on cluster policies that meet the methodological requirements formulated above. In the synopsis of many evaluation studies on different programmes, more general insights about cluster policies might be carved out that would allow better assessments about what works in cluster policy and what does not work.

Cluster policy is a rapidly changing policy area that is increasingly losing its formerly clear contours. Recently, new political concepts have been put into practice, such as smart specialisation. These concepts take up essential aspects of the cluster narrative and simultaneously open new ways of promoting regional development. While the evaluation of programmes operating under the new label will have to adapt to the design of the respective programmes, probably little will change in terms of the fundamental challenges facing impact evaluation. Therefore, the ideas expressed here can be transferred to the evaluation of these programmes.

We did scrutinise whether the sample of articles used in this paper is representative for all evaluation studies. To the best of our knowledge, we identified a large share of all publicly available studies. However, there are some commissioned studies that are not available to the public. Due to our approach in identifying relevant research (see Section 1), we are quite sure that the picture we present here is representative of all publicly available evaluation studies on cluster policy. While the contribution of this paper lies in addressing various critical aspects in impact evaluation, the paper finds its limits in the specification of proposals for the further development of methodological approaches. This is precisely where there is a need for future evaluators to play a decisive role. A second path to more elaborate impact evaluations lies in the further development of the evaluation culture in Germany.

If we draw a balance of our critical review of German cluster policy evaluations, we find useful information on the effects of cluster policy, but also significant weaknesses. Against this backdrop, a methodological rethinking with respect to cluster policy impact analyses appears appropriate. Cluster policy, in its more promising variants, is essentially a systemic approach to policy. One core result of this literature survey is that the impact of a cluster programme can only occur after long periods of time. Hence, ex post evaluations should be an integral part of evaluation work. This long-term perspective can and should complement ongoing and mid-term evaluations. Long-term processes of the gradual emergence, growth, decline and transformation of clusters based on both high-quality case studies and quantitative approaches deserve far more attention than they have received so far.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Appendix

**Table 3** Analysed evaluation studies of German cluster policies, including meta-analyses (Full titles in list of references).

No.	Author(s)	Year of publication	Shortened title/topic(s)	Major methods used
01	Alecke et al.	2008	Some unpleasant macroeconomic evidence for regional cluster policy	Econometric, regional model, meso-level
02	Audretsch, Lehmann, Menter	2016	Public cluster policy and new venture creation	Econometric, macro-level
03	Brenner, Emmerich, Schlump	2013	Regional effects of a cluster-oriented measure	Econometric, meso-level
04	Brenner, Schlump	2011	Policy measures and their effects in the different phases of the cluster life cycle	Meta-analysis
05	Brökel, Graf	2020	Innovationspolitik und Netzwerke	Social network analysis meso-level
06	Cantner, Graf, Hinzmann	2013	Policy induced innovation networks: the case of LECC	Social network analysis meso-level
07	Crass, Rammer, Aschhoff	2016	Geographical clustering and the effectiveness of public innovation programs	Econometric, meso- level
08	Dohse	2000	Technology policy and the regions – the case of BioRegio Contest	Descriptive statistics, qualitative methods, meso-level
09	Dohse	2005	Clusterorientierte Technologiepolitik in Deutschland	Meta-analysis
10	Dohse	2007	Cluster-based technology policy	Descriptive statistics, qualitative methods, meso-level
11	Dohse, Staehler	2008	BioRegio, BioProfile and the rise of the German Biotech industry	Descriptive statistics, qualitative methods, micro- and meso-level
12	Eickelpasch, Pfeiffer	2004	InnoRegio: Unternehmen beurteilen die Wirkung	Expert interviews micro- and meso-level
13	Eickelpasch, Pfeiffer, Pfirmann	2004	Zur Evaluierung der Förderung regionaler Innovationsnetzwerke	Meta-analysis, expert interviews, micro- and meso-level
14	Engel, Eckl, Rothgang	2017	R&D funding and private R&D	Difference-in-difference approach, micro-level
15	Engel et al.	2013	Does cluster policy trigger R&D activity?	Difference-in-difference analyses, micro-level
16	Engel, Rothgang, Eckl	2013	Systemic aspects of R&D policy	Difference-in-difference approach, micro-level
17	Falck, Heblich, Kipar	2008	The extension of clusters: difference-in-difference evidence from the Bavarian state-wide cluster policy	Difference-in-difference approach, micro- and macro-level
18	Falck, Heblich, Kipar	2010	Industrial innovation: direct evidence from a cluster-oriented policy	



**Table 3** (continued)

No.	Author(s)	Year of publication	Shortened title/topic(s)	Major methods used
				Difference-in-difference approach, micro- and macro-level
19	Falck, Kipar, Wößmann	2011	Innovationstätigkeit von Unternehmen: Die Rolle von Qualifikationen, Kooperationen u. Clusterpolitik	Econometric, micro- and macro-level
20	Fornahl et al.	2015	Cluster als Paradigma der Innovationspolitik	Meta-analysis
21	Fromhold-Eisebith	2014	Erfolgsgeschichte oder Modeerscheinung? Clusterpolitik im Spannungsfeld von Theorie und Praxis	Meta-analysis
22	Fromhold-Eisebith, Eisebith	2008a	Clusterförderung auf dem Prüfstand. Eine kritische Analyse	Meta-analysis
23	Fromhold-Eisebith, Eisebith	2008b	Looking behind facades: Evaluating effects of (automotive) cluster promotion	Descriptive statistics, qualitative methods, micro- and meso-level
24	Graf, Brökel	2020	A shot in the dark? Policy influence on cluster networks	Social network analysis, micro- and meso-level
25	Hinzmann, Cantner, Graf	2017	The role of geographical proximity for project performance	Social network analysis, micro- and meso-level
26	Kiese	2008	Die Clusterpolitik deutscher Länder rund um Regionen als Herausforderung für die Evaluation	Meta-analysis
27	Kiese	2012	Regionale Clusterpolitik in Deutschland	Meta-analysis
28	Kiese	2017	Regional cluster policies in Germany: challenges, impacts and evaluation practices	Meta-analysis
29	Kiese, Hundt	2014	Cluster policies, organising capacity and evaluation practices	Meta-analysis, qualitative methods
30	Kramer	2008	Die Evaluierung von Clusterpolitik	Meta-analysis
31	Lehmann, Menter	2017	Cluster policy and performance	Econometric model, meso- and macro-level
32	Nestle	2011	Open innovation in cluster: eine Wirkungsanalyse zu Clusterinitiativen	Difference-in-difference approach, micro- and meso-level
33	Rothgang et al.	2014a	Begleitende Evaluierung des Förderinstruments "Spitzencluster-Wettbewerb"	Mixed-methods design, micro- and meso-level
34	Rothgang, Dehio, Lageman	2017	Analysing the effects of cluster policy	Meta-analysis
35	Santner	2018	Cluster-internal and external drivers of cluster renewal: evidence from two German	Descriptive statistics, qualitative methods
36	Schmiedeberg	2010	Evaluation of cluster policy	Meta-analysis
37	Segers	2016	Lessons from the biotechnology clusters in Belgium and Germany	Descriptive statistics, qualitative methods
38	Stachler, Dohse, Cook	2006	Evaluation der Fördermaßnahmen BioRegio und BioProfile	Mixed-methods design, micro- and meso-level

**Table 3** (continued)

No.	Author(s)	Year of publication	Shortened title/topic(s)	Major methods used
39	Töpfer, Cantner, Graf	2017	Structural dynamics of innovation networks	Social network analysis, micro- and meso-level
40	Wolf et al.	2017	Cluster ambidexterity towards exploration and exploitation	Econometric meso-level

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alecke B, Alsleben C, Scharr F, Untiedt G (2008) Geographic concentration of sectors in the German economy: Some unpleasant macroeconomic evidence for regional cluster policy. In: Blien U, Maier G (eds) The economics of regional clusters. Networks, technology and policy, New Horizons in Regional Science, pp 209–233
- Andersson T, Schwaag Serger S, Sörvik J, Wise Hansson E (2004) The cluster policies whitebook. IKED International Organization for Knowledge Economy and Enterprise Development, Holmberg
- Arnold E (2004) Evaluating research and innovation policy: a systems world needs systems evaluations. Res Evaluation 13(1):3–17 Download from <https://academic.oup.com/rev/article-abstract/13/1/3/1593255>, on 24 February 2020
- Astor M, Fischl I, Hoffmann J, Koglin G, Kulicke M, Sheikh S, Wessels J, Whitelegg K (2014) Evaluation von Forschungs- und Technologiepolitik in Deutschland und Österreich – ein Überblick. In: Böttcher W, Kerlen C, Maats P, Schwab O, Sheikh S, DeGEval-Vorstand (eds) Evaluation in Deutschland und Österreich. Stand und Entwicklungsperspektiven in den Arbeitsfeldern der DeGEval – Gesellschaft für Evaluation. Waxmann, Münster, pp 139–149
- Audretsch DB, Lehmann EE, Menter M (2016) Public cluster policy and new venture creation. *Economia e Politica Industriale* 43(4):357–381. <https://doi.org/10.1007/s40812-016-0050-9>
- Bathelt H, von Bernuth C (2008) Leipzig's media cluster. In: Karlsson C (ed) Handbook of research on innovation and clusters. Cases and policies. Edward Elgar, Cheltenham, pp 270–282
- Bathelt H, Glückler J (2018) [2002], Wirtschaftsgeographie. Ökonomische Beziehungen in räumlicher Perspektive. 4., vollst. überarb. u. erw. Auflage. Eugen Ulmer, Stuttgart
- Bauer TK, Fertig M, Schmidt CM (2009) Empirische Wirtschaftsforschung. Eine Einführung. Springer, Berlin
- Belcher B, Palenberg M (2018) Effects and impacts of development interventions: toward conceptual clarity. *Am J Eval* 39(4):478–495
- Bennett A, Checkel JT (eds) (2015) Process tracing. From metaphor to analytical tool. Cambridge University Press, Cambridge
- Bergman EM (2008) Cluster life-cycles: an emerging synthesis. In: Karlsson C (ed) Handbook of research on cluster theory. Edward Elgar, Cheltenham, pp 114–132
- Bjurulf S, Vedung E, Larsson CG (2012) A triangulation approach to impact evaluation. *Evaluation* 19(1):56–73
- Bockmann B, Buch CM, Schnitzer M (2014) Evidenzbasierte Wirtschaftspolitik in Deutschland: Defizite und Potenziale, IZA Standpunkte Nr. 68. IZA, Bonn

- Borrás S, Edquist C (2019) Holistic innovation policy. Theoretical foundations, policy problems, and instrumental choices. Oxford University Press, Oxford
- Bradley J, Untiedt G (2007) Do economic models tell us anything useful about cohesion policy impacts? GEFRA Working Paper: July 2007 – Nr. 3. Münster
- Brenner T, Emmerich CE, Schlump C (2013) Regional effects of a cluster-oriented policy measure – the case of the InnoRegio program in Germany. Working Papers on Innovation and Space, marburg geography # 05.13
- Brenner T, Schlump C (2011) Policy measures and their effects in the different phases of the cluster life cycle. *Reg Stud* 45(10):1363–1386
- Brókel T, Graf H (2020) Innovationspolitik und Netzwerke. In: Stemer J (ed) *Handbuch Innovative Wirtschaftsförderung*. Springer, Wiesbaden. [https://doi.org/10.1007/978-3-658-21597-2\\_28-1](https://doi.org/10.1007/978-3-658-21597-2_28-1)
- Burda M, Schnitzer M, Fitzenberger B, Schmidt C (2014) Evidenzbasierte Wirtschaftspolitik. Thesenpapier zur Kerntagung des Vereins für Socialpolitik in Hamburg 2014. Download from [https://www.socialpolitik.de/docs/2014/Thesen\\_Evidenzbasierte\\_Wirtschaftspolitik.pdf](https://www.socialpolitik.de/docs/2014/Thesen_Evidenzbasierte_Wirtschaftspolitik.pdf), on 10 November 2019
- Cantner U, Graf H, Hinzmann S (2013) Policy induced innovation networks: the case of the German 'Leading-Edge Cluster competition'. Chapter 19. In: Scherngell T (ed) *The geography of networks and R&D collaborations*. Springer, Heidelberg. [https://doi.org/10.1007/978-3-319-02699-2\\_18](https://doi.org/10.1007/978-3-319-02699-2_18)
- Case A, Deaton A (2020) *Deaths of despair and the future of capitalism*. Princeton University Press, Princeton & Oxford
- Chaminade C, Lundvall B-A, Haneef S (2018) *Advanced introduction to national innovation systems*. Edward Elgar, Northampton
- Crass D, Rammer C, Aschhoff B (2016) Geographical clustering and the effectiveness of public innovation programs, ZEW Discussion Paper 16-083. ZEW, Mannheim
- DeGEval (2016a) Evaluation und Wissenschaft. Positionspapier 08 des Vorstandes der DeGEval – Gesellschaft für Innovation. Mainz, <https://www.org/publikationen/positionspapiere/>
- DeGEval (2016b) Standards für Evaluationen. In: Erste Revision. DeGEval – Gesellschaft für Evaluation, Mainz [www.degeval.de](http://www.degeval.de). Accessed 16 Jan 2020
- Diez MA (2001) The evaluation of regional innovation and cluster policies: towards a participatory Approach. *Eur Plan Stud* 9(7):907–923
- Dohse D (2000) Technology policy and the regions – the case of the BioRegio contest. *Res Policy* 29(9)
- Dohse D (2005) Clusterorientierte Technologiepolitik in Deutschland. *Konzepte und Erfahrungen. Technologiefolgeabschätzung Theorie und Praxis* 14(1)
- Dohse D (2007) Cluster-based technology policy – the German experience. *Ind Innov* 14(1):69–94
- Dohse D, Staehler T (2008) BioRegio, BioProfile and the rise of the German biotech industry, Kiel Working Paper 1456. IfW, Kiel
- Edler J, Fagerberg J (2017) Innovation policy: what, why, and how. *Oxf Rev Econ Policy* 33(1):2–23 Download from: <https://academic.oup.com/oxrep/article-abstract/33/1/2/2972712>, on 19 November 2019.
- Edler J, Cunningham P, Gök A, Shapira P (eds) (2016) *Handbook of innovation policy impact*. Edward Elgar, London
- Eickelpasch A, Pfeiffer I, Pfirrmann O (2004) Zur Evaluierung der Förderung regionaler Innovationsnetzwerke – Ansatz und Erfahrungen am Beispiel des InnoRegio-Programms. *Zeitsch Evaluat* 3:181–198
- Eickelpasch A, Pfeiffer I (2004) InnoRegio: Unternehmen beurteilen die Wirkung des Förderprogramms insgesamt positiv. *DIW Wochenbericht* 23:331–337
- Engel D, Rothgang M, Eckl V (2015) Systemic aspects of R&D policy – subsidies for R&D collaborations and their effects on Private R&D. *Ruhr economic papers* 587. RWI, Essen
- Engel D, Mitze T, Patuelli R, Reinkowski J (2013) Does cluster policy trigger R&D activity? Evidence from German biotech contests. *Eur Plan Stud* 21(11):1735–1759
- Engel D, Eckl V, Rothgang M (2017) R&D funding and private R&D: empirical evidence on the impact of the leading-edge cluster competition. *J Technol Transf.* <https://doi.org/10.1007/s10961-017-9609-5>
- European Commission, DG Budget – Evaluation Unit (2004) *Evaluating EU activities*. In: *A practical guide for the Commission Services*. European Communities, Luxembourg
- European Commission, Directorate-General Regional Policy (2006), *Indicative guidelines on evaluation methods: monitoring and evaluation indicators*. The New Programming Period 2007-2013. Working Document No. 2, [https://ec.europa.eu/regional\\_policy/sources/docoffic/cocof/2006/cocof\\_06\\_0010\\_00\\_en.pdf](https://ec.europa.eu/regional_policy/sources/docoffic/cocof/2006/cocof_06_0010_00_en.pdf),
- Falck O, Heblich S, Kipar S (2008) The extension of clusters: difference-in-difference evidence from the Bavarian state-wide cluster policy, *Jena Economic Research Papers* # 2008-073. Friedrich-Schiller University, Max Planck Institute of Economics, Jena

- Falck O, Heblich S, Kipar S (2010) Industrial innovation: direct evidence from a cluster-oriented policy. *Reg Sci Urban Econ* 40:574–582
- Falck O, Kipar S, Wößmann L (2011) Innovationstätigkeit von Unternehmen: Die Rolle von Qualifikationen, Kooperationen und Clusterpolitik, Ifo Beiträge zur Wirtschaftsforschung. ifo-Institut, München
- Flick U (2008) [2004], *Triangulation. Eine Einführung. Qualitative Sozialforschung*, Bd. 12. 2. Auflage. VS Verlag für Sozialwissenschaften, Wiesbaden
- Fornahl D, Hassink R (eds) (2017) *The life cycle of clusters. A policy perspective*. Edward Elgar, Cheltenham & Northampton
- Fornahl D, Heimer T, Campen A, Talmon-Gros L, Treperman J, Gerken u M v P, Knop S, Reinecke I, Schrapers M (2015) Cluster als Paradigma der Innovationspolitik – Eine erfolgreiche Anwendung von Theorie und Praxis? Studien zum deutschen Innovationssystem 13-2015. CRIE – Centre for Regional and Innovation Economics, Universität Bremen, Technopolis Deutschland GmbH, Bremen & Frankfurt
- Fromhold-Eisebith M (2014) Erfolgsgeschichte oder Modeerscheinung? Clusterpolitik im Spannungsfeld von Theorie und Praxis. In: RC Beck, RG Heinze and J Schmid (eds) *Zukunft der Wirtschaftsförderung*. Baden-Baden: Nomos, pp 67–88
- Fromhold-Eisebith M, Eisebith G (2008a) Clusterförderung auf dem Prüfstand. Eine kritische Analyse. *Z Wirtsch* 52(2-3):79–94
- Fromhold-Eisebith M, Eisebith G (2008b) Looking behind facades. Evaluating effects of (Automotive) cluster promotion. *Reg Stud* 42(10):1343–1356
- fteval (2004), Eckpunkte der Evaluierung von Forschungs- und Technologiepolitik. Endbericht („Modul“) im Rahmen des Projektes Zusammenarbeit des Rates für Forschung und Technologieentwicklung mit der Plattform Forschungs- und Technologieevaluierung. Wien: fteval – Plattform Forschungs- und Technologieevaluierung GesbR, [www.fteval.at](http://www.fteval.at)
- Georgiou L (1998) Issues in the evaluation of innovation and technology policy. *Evaluation* 1(4):37–51
- Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ (2011), *Impact Evaluation in Practice*. Washington DC: The World Bank. Download from <http://www.worldbank.org/ieinpractice>, on 20 January 2020
- Graf H, Broekel T (2020) A shot in the dark? Policy influence on cluster networks. *Res Policy* 49(3):103920
- Haskins R, Baron J (2011) Building the connection between policy and evidence. The Obama evidence-based initiatives. NESTA, London
- Hinzmann S, Cantner U, Graf H (2017) The role of geographical proximity for project performance: evidence from the German Leading-Edge Cluster Competition. *J Technol Transf*. <https://doi.org/10.1007/s10961-017-9600-1>
- Kelle U (2007) *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte*. VS Verlag für Sozialwissenschaften, Wiesbaden
- Ketels C (2014) Ein Jahrzehnt Clusterpolitik und –forschung: Implikationen für eine moderne, clusterorientierte Wirtschaftsförderung. In: RC Beck, RG Heinze and J Schmid (eds) *Zukunft der Wirtschaftsförderung*. Baden-Baden: Nomos, pp 45–64
- Kiese M (2008) In: Wessels J (ed) *Die Clusterpolitik deutscher Länder und Regionen als Herausforderung für die Evaluationen*, pp 27–38
- Kiese M (2012) Regionale Clusterpolitik in Deutschland. Bestandsaufnahme und interregionaler Vergleich im Spannungsfeld von Theorie und Praxis. *Metropolis*, Marburg
- Kiese M (2017) Regional cluster policies in Germany: challenges, impacts and evaluation practices. *J Technol Transf*. <https://doi.org/10.1007/s10961-017-9589-5>
- Kiese M, Hundt C (2014) Cluster policies, organising capacity and regional resilience: evidence from German case studies. *Raumforsch Raumordn* 72:117–131
- Kramer J-P (2008), *Die Evaluierung von Clusterpolitik. Zwischen theoretischem Anspruch und praktischer Realität*. Diplomarbeit am Institut für Wirtschafts- und Kulturgeographie der Leibniz Universität Hannover. München, Ravensburg: GRIN Verlag
- Lamnek S (2005) [1988], *Qualitative Sozialforschung. Lehrbuch*. 4., vollst. überarb. Auflage. Weinheim. Basel, Beltz Verlag
- Lehmann EE, Menter M (2017) Public cluster policy and performance. *J Technol Transf*. <https://doi.org/10.1007/s10961-017-9626-4>
- Lindquist G, Ketels C, Sölvell Ö (2013) *The cluster initiative Greenbook 2.0*. Ivory Tower Publishers, Stockholm
- Martin R, Sunley P (2003) Deconstructing clusters: chaotic concept or policy panacea? *J Econ Geogr* 3:5–35
- Mazzucato M (2014) [2013], *The Entrepreneurial state. Debunking public vs. private sectors myths*. Rev. ed. Anthem Press, Delhi

- Menzel M-P, Fornahl D (2007) Cluster life cycles – dimensions and rationales of cluster development. *Ind Corp Chang* 19:205–238
- MRC – Medical Research Council (2019), Developing and evaluating complex interventions: Following considerable development in the field since 2006, MRC and NIHR have jointly commissioned an update of this guidance to be published in 2019. Download from [www.mrc.ac.uk/complexinterventionsguidance](http://www.mrc.ac.uk/complexinterventionsguidance), on 19 February 2020
- Nestle V (2011) Open Innovation im Cluster: Eine Wirkungsanalyse zu Clusterinitiativen in forschungsintensiven Industrien. In: *Innovationsmanagement und Entrepreneurship*, Dissertation EBGV für Wirtschaft und Recht i.Gr. 2010. Gabler, Wiesbaden
- Onwuegbuzie, A.J. and R.B. Johnson (2008), Types of legitimation (Validity) in mixed methods research. Reprinted from: Onwuegbuzie, A.J. and R.B. Johnson (2006), The validity issue in mixed research. *Research in the schools* 13(1): 48-63. In: V.L. Plano-Clark and J.W. Creswell (eds.), 271-298
- Pawson R (2013) *The science of evaluation. A realist manifesto*. Sage, London
- Perrin B (2002) How to – and how not to – evaluate innovation. *Evaluation* 8(1):13–28
- Peters J, Langbein J, Roberts G (2015) Policy evaluation, randomized controlled trials, and external validity – a systematic review. *Ruhr Economic Papers* # 589. RWI, Essen
- Plano Clark VL, Creswell JW (2008) *The Mixed Methods Reader*. SAGE Publications, Los Angeles
- Porter ME (1998) Clusters and the new economics of competition. *Harvard Business Review*, Nov-Dec.:77-90
- Porter ME (1999) [1991], Nationale Wettbewerbsvorteile. Erfolgreich konkurrieren auf dem Weltmarkt. Management-Bibliothek. Wirtschaftsverlag Ueberreuter, Wien
- Rothgang M, Cantner U, Dehio J, Engel D, Fertig M, Graf H, Hinzmann S, Linshalm E, Ploder M, Scholz A, Töpfer S (2014a), Begleitende Evaluierung des Förderinstruments „Spitzencluster-Wettbewerb“ des BMBF. Forschungsvorhaben des Bundesministeriums für Bildung und Forschung. Abschlussbericht. RWI, ISG, Joanneum Research, LS Mikroökonomik Jena. RWI, Essen:
- Rothgang M, Cantner U, Dehio J, Engel D, Fertig M, Graf H, Hinzmann S, Linshalm E, Ploder M, Scholz A, Töpfer S (2014b) Begleitende Evaluierung des Förderinstruments „Spitzencluster-Wettbewerb“ des BMBF. Abschlussbericht. Kurzfassung. RWI-Materialien 83. Essen: RWI.
- Rothgang M, Dehio J, Lageman B (2017) Analysing the effects of cluster policy: What can we learn from the German leading-edge cluster competition? *J Technol Transf*. <https://doi.org/10.1007/s10961-017-9616-6>
- Santner D (2018) Cluster-internal and external drivers of cluster renewal: evidence from two German agricultural engineering case studies. *Eur Plan Stud* 26(1):174–191 <https://www.tandfonline.com/loi/ceps20>
- Schmiedeberg C (2010) Evaluation of cluster policy: a methodological overview. *Evaluation* 16(4):389–412
- Segers J-P (2016) Regional systems of innovation: lessons from the biotechnology clusters in Belgium and Germany. *J Small Bus Entrep* 28(2):133–149
- Shiller RJ (2019) *Narrative Economics. How stories go viral & drive major economic events*. Princeton University Press, Princeton & Oxford
- Stähle T, Dohse D, Cooke P (2006) Evaluation der Fördermaßnahmen Bioregio und BioProfile. Bericht im Auftrag des Bundesministeriums für Bildung und Forschung. CIR - Consulting für Innovations- und Regionalanalysen, ifW – Institut für Weltwirtschaft, CASS – Centre for Advanced Studies in the Social Sciences. Kiel and Cardiff
- Sölvell Ö, Williams M (2013) *Building the cluster commons*. In: An evaluation of 12 cluster organizations in Sweden 2005-2012. Ivory Tower Publishers, Stockholm
- Töpfer S, Cantner U, Graf H (2017) Structural dynamics of innovation networks in German leading-edge cluster. *J Technol Transf*. <https://doi.org/10.1007/s10961-017-9642-4>
- Uyarra E, Ramlogan R (2016) In: Edler J, Cunningham P, Gök A, Shapira P (eds) *The impact of cluster policy on innovation*, pp 197–231
- Vicente J (2016) *Économie des clusters*. Éd. La Découverte, Paris
- W.K. Kellogg Foundation (2004) *Logic model development guide. Using logic models to bring together planning, evaluation, and action*. Battle Creek, Michigan
- Wessels J (2008) Cluster- und Netzwerkevaluation. Aktuelle Beispiele aus der Praxis. AK Forschungs-, Technologie- und Innovationspolitik der DeGEVal in Kooperation mit dem Institut für Innovation und Technik, Berlin
- Wolf T, Cantner U, Graf H, Rothgang M (2017) Cluster ambidexterity towards exploration and exploitation: strategies and cluster management. *J Technol Transf*. <https://doi.org/10.1007/s10961-017-9617-5>

## Affiliations

Michael Rothgang<sup>1</sup> · Bernhard Lageman<sup>1</sup> · Anne-Marie Scholz<sup>2</sup>

Bernhard Lageman  
bernhard.lageman@t-online.de

Anne-Marie Scholz  
scholz@isg-institut.de

<sup>1</sup> RWI-Leibniz Institute for Economic Research, Essen, Germany

<sup>2</sup> ISG Institute for Social Research, Cologne, Germany