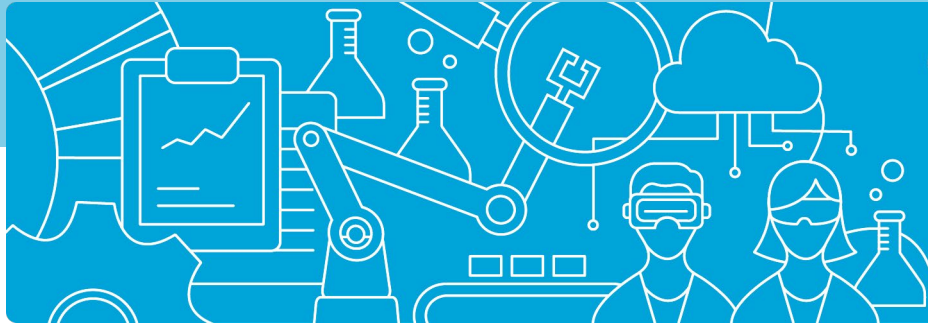


Studie zum deutschen Innovationssystem | Nr. 11-2024



Stefan Buechele, Guido Buenstorf, Uwe Cantner, Lukas Dreier, Petra Meurer, Liam Paul Neumann

Commissioned Project Evaluations of Research and Innovation Policy in Germany: A Review

U N I K A S S E L
V E R S I T Ä T



**FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA**

Durchführende Institute

Universität Kassel, Fachbereich Wirtschaftswissenschaften,
Fachgebiet Wirtschaftspolitik, Innovation und Entrepreneurship,
und International Center for Higher Education Research (INCHER)
Mönchebergstraße 17, 34109 Kassel
www.uni-kassel.de

Friedrich-Schiller-Universität Jena
Lehrstuhl für Volkswirtschaftslehre/Mikroökonomik
Carl-Zeiss-Straße 3, 07743 Jena
www.uni-jena.de

Studien zum deutschen Innovationssystem

Nr. 11-2024
ISSN 1613-4338

Stand

Februar 2024

Herausgeberin

Expertenkommission Forschung und Innovation (EFI)

Geschäftsstelle

Pariser Platz 6 | 10117 Berlin
www.e-fi.de

Alle Rechte vorbehalten, insbesondere das Recht auf Vervielfältigung und Verbreitung sowie die Übersetzung. Kein Teil des Werkes darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der EFI oder der Institute reproduziert oder unter Verwendung elektronischer Systeme gespeichert, verarbeitet, vervielfältigt oder verbreitet werden.

1. Introduction

The evaluation of policy interventions is an essential tool for reviewing and optimizing policies, especially in the field of research and innovation (R&I) policy (BMW¹ 2013). Policy evaluations involve systematic analyses to assess the design, implementation, and impact of policy interventions. Their aim is to provide decision makers with evidence-based insights to enhance the achievement of desired objectives.

Monitoring and evaluation serve multiple purposes (Gertler et al. 2016): Monitoring of policy interventions through the continuous collection and analysis of information on key indicators aims to assess the quality of the implementation and to track the development of interventions over time. In contrast, policy evaluations are conducted at specific points in time, often with contributions from external experts to ensure an impartial perspective (Gertler et al. 2016). These policy evaluations not only review whether the intervention was successfully implemented but also determine, in the case of impact assessments, whether associated effects can be attributed to the program, i.e. whether the program was effective.

Irrespective of the specific policy domain, policy evaluations are important to inform evidence-based policy-making and to assess the impacts of interventions. In the realm of R&I policy, the use of evaluations to facilitate policy learning is particularly important, as R&I policy uses a wide variety of funding approaches, is highly context-specific, and new expectations are raised in the course of tackling major societal challenges. For example, R&I policy operates in the areas of basic research, applied research and product development exhibiting varying degrees of uncertainty, and it addresses different technologies, sectors and regions. In addition, new challenges in R&I policy, such as mission-oriented policies focusing on the directionality of R&I activities, also require new policy designs.

The German Federal Budget Code (Bundeshaushaltsordnung, BHO¹) mandates that all measures with an impact on the federal budget be subject to an appropriate economic feasibility study. According to an administrative regulation linked to the BHO², this economic feasibility study has to include a performance review. The performance review aims to determine whether and to what extent the intended objectives of a policy intervention have been achieved, what the economic efficiency is and whether the policy intervention was actually the cause of the achievement of the objectives. Furthermore, performance reviews should also help to identify needs and opportunities for adjusting the policy intervention.³

There is broad consensus among researchers that to establish the effectiveness of policy interventions, their causal effects on the intended outcomes have to be assessed. This in turn requires the use of suitable experimental or quasi-experimental research designs (e.g. Shadish et al. 2002; BMW¹ 2013; Verein für Socialpolitik 2015; Athey and Imbens 2017). Prior evidence suggests, however, that this consensus is not (yet) reflected in the current practice of policy evaluation (Collin et al. 2022). To explore the state of evaluation practice in the realm of R&I policy in Germany, we systematically searched for published policy evaluation studies commissioned by the two main federal ministries active in this policy realm: the Federal Ministry of Education and Research (BMBF) and the Federal Ministry for Economic Affairs

¹ Cf. <https://www.gesetze-im-internet.de/bho/>.

² Cf. https://www.verwaltungsvorschriften-im-internet.de/bsvwvbund_14032001_DokNr20110981762.htm.

³ The original German version states: “Die Erfolgskontrolle ist ein systematisches Prüfungsverfahren. Sie dient dazu, während der Durchführung (begleitende Erfolgskontrolle) und nach Abschluss (abschließende Erfolgskontrolle) einer finanzwirksamen Maßnahme ausgehend von der Planung festzustellen, ob und in welchem Ausmaß die angestrebten Ziele erreicht wurden, ob die finanzwirksame Maßnahme ursächlich für die Zielerreichung war und ob die finanzwirksame Maßnahme wirtschaftlich war. Erfolgskontrollen sollen auch dazu führen, dass Bedarfe und Möglichkeiten des Um- bzw. Nachsteuerens rechtzeitig erkannt werden.” Cf. https://www.verwaltungsvorschriften-im-internet.de/bsvwvbund_14032001_DokNr20110981762.htm.

and Climate Action (BMWK). We were able to retrieve 81 policy evaluation studies published from 2009 to 2023. For this sample, we explored to what extent research designs allowing for the identification of causal effects (in short: causal analysis) were applied, how evaluation results were interpreted, and what obstacles made it difficult for evaluators to perform proper causal analysis.

Our findings reveal a mixed picture. A majority (73%) of the analyzed studies interpreted their findings on intervention outcomes as causal effects of the policy intervention. However, a substantial share of the latter (76%) did not employ suitable methods of causal analysis based on experimental or quasi-experimental research designs, which raises questions about the validity of these interpretations. Furthermore, all policy evaluation studies that assess the effectiveness of the analyzed policy interventions reported positive or at least partly positive effects. Over time, an increasing emphasis on impact assessments and a more frequent application of methods of causal analysis can be discerned. Policy evaluation studies in our sample were conducted by a diverse set of organizations ranging from non-university public research organizations (PROs) and universities to private-sector companies such as consulting firms.

These findings indicate that rigorous methods of impact assessment have not been in widespread use in the evaluation of R&I policies in Germany. This limits the ability of policymakers to learn from past policy interventions, and, therefore, improve the effectiveness of future public R&I support programs.

2. Theoretical Background

2.1 Policy Evaluations

Monitoring and evaluation are of crucial importance in R&I policy, as they allow policy makers and program managers to optimize policy support programs and to review the effectiveness and efficiency of their policy support programs. Policy support programs are effective when they achieve their intended objectives and efficient when there are no alternatives for realizing these objectives at a lower cost.

Monitoring describes the process of continuous collection and analysis of information about a policy intervention, with key indicators (Khandker et al. 2010) being employed to compare actual results with planned results to assess the quality of intervention implementation. This involves the use of data such as characteristics of individual participants, information about applications and funding approval, as well as the situation of beneficiaries at the end of the policy intervention and the associated costs (ILO 2013). Monitoring addresses the question of how a policy intervention evolves over time. A reliable monitoring system that collects data is crucial for future policy evaluations (ILO 2013).

Policy evaluation is a structured and impartial examination of various aspects of a policy intervention including its design, execution, and outcomes, to determine its overall value or performance. The goal is to provide reliable insights to policy makers and program managers, enabling them to enhance the attainment of desired objectives. Unlike continuous monitoring, policy evaluations are conducted at specific points in time. They frequently involve input from external technical experts to provide an impartial perspective (Gertler et al. 2016).

Following the International Labour Organization (ILO 2013), there are two main evaluation types: performance evaluations and impact assessments. Performance evaluations assess the quality of service delivery and intervention outcomes, focusing on short-term and medium-term results. They use information collected through monitoring activities. However, unlike monitoring, performance evaluation aims to determine whether the desired goals of the intervention are fulfilled, providing a broader analysis (ILO 2013).

Impact assessments aim to identify changes in outcomes that are directly attributable to the intervention (Khandker et al. 2010; Gertler et al. 2016). They assess the potential outcomes that would have occurred if beneficiaries had not participated in the program. The main feature of impact assessments is the

intention to establish causality between the program and specific outcomes (ILO 2013). Impact assessments therefore address the question of what the causal effect of a policy intervention is on an outcome of interest.

Impact assessments are often used to support evidence-based policy making (Gertler et al. 2016). A classic example of impact assessment is analyzing the effects of a labor market program on the employment rates of participants. By employing methods of causal analysis based on suitable research designs such as randomized control trials or regression discontinuity approaches, impact assessments aim to construct counterfactual outcomes that would have been realized in the absence of the intervention (see Section 2.3).

2.2 Summative and Formative Evaluations

Evaluations can be differentiated by two roles: summative versus formative evaluations (Scriven 1991). Summative evaluation refers to assessing the effectiveness of an intervention (Edler et al. 2012), usually conducted ex-post (Stockmann 2006). An example of summative evaluation is the assessment of learning outcomes at the end of educational public support program. Summative evaluations often utilize quantitative methods and statistical analyses to assess the interventions (Edler et al. 2012).

In contrast, formative evaluation aims to improve a policy intervention during its implementation phase (Scriven 1991). It is process-oriented and focuses on how a policy intervention operates and how it can be developed further. Formative evaluations are typically conducted to collect feedback, identify issues in implementation, and suggest adjustments (Stockmann 2006; Edler et al. 2012). An example of formative evaluation is the continuous assessment of a curriculum in a school to adjust teaching methods and optimize student learning. Formative evaluations often use qualitative methods such as interviews, observations, and focus groups to gain detailed insights into the experiences of participants and the dynamics of the intervention (Edler et al. 2012). The main difference between formative and summative evaluations lies in their objectives, i.e. assessment (summative) or improvement (formative) (Edler et al. 2012).

2.3 Basic Concepts of Causal Analysis

Methods of causal analysis are used in the social sciences to determine cause-and-effect relationships between variables (Woodward 2007). This is particularly vital in policy evaluation (Falck et al. 2015; Collin et al. 2022), such as in assessing whether a public R&I support program (the independent variable) has a causal effect on the desired outcome (the dependent variable). Understanding the impact of interventions on outcomes can inform decision-making and improve future strategies, thereby inducing policy learning.

Central to causal analysis is the concept of the counterfactual scenario, which represents the hypothetical outcome of the treatment group had the intervention not been applied. This concept is crucial for understanding the causal effect of an intervention. However, as counterfactual outcomes for the treatment group are by necessity unobservable, in causal analysis they are approximated by outcomes of suitable reference units – the control group.

At the forefront of causal analysis are experimental designs, which are characterized by their use of randomization (e.g. via lotteries) to assign participants to treatment and control groups. In an experimental research design, the treatment group receives the intervention under investigation while the control group does not. Randomization can ensure that both known and unknown factors that could influence the outcome are distributed evenly across treatment and control groups. This equal distribution helps in identifying the causal effect of the intervention, by isolating this effect from other influences.

As experimental research designs allow for a high degree of control over external variables, they provide a robust framework for identifying causal effects (DiNardo and Lee 2011).

However, the real world often presents scenarios where such controlled randomization is not feasible, mostly because the researcher (e.g. evaluator) cannot directly control the assignment to treatment and control groups. In such contexts, researchers turn to quasi-experimental research designs (see e.g. Imbens and Wooldridge 2009). Quasi-experimental designs exploit natural or quasi-random variations in treatment assignment for the inference of causal effects. Key quasi-experimental designs⁴ include Difference-in-Difference (DiD), Regression Discontinuity Design (RDD), and Synthetic Control Methods (SCM). DiD leverages data from multiple (at least two) time periods, at least once before and once after the intervention, comparing the changes in outcomes over time between the treatment group and a comparable control group. RDD can be employed when the assignment to the treatment is based on a cutoff value of a continuous assignment variable (such as age, size or distance). This approach compares units just above and just below the threshold, assuming that those near the cutoff are similar in all respects except for the treatment received. SCMs construct a synthetic control group by combining pre-treatment data from multiple untreated units to replicate the characteristics of a treated unit. Comparison of the outcome of the treated unit with the counterfactual outcome of the synthetic control group provides an estimate of the treatment effect.

3. Method

For analyzing the use of causal analysis in evaluations of R&I policy, we collected data on 81 evaluation studies commissioned by the BMBF and the BMWK. A list of the studies in our sample can be found in the Appendix.

The examined policy evaluation studies exhibit a high degree of heterogeneity, both in terms of comprehensiveness, as short and long versions of studies are considered, and in terms of the underlying units of evaluation, which encompass public support programs and institutions dedicated to specific tasks. The analyzed studies span from 2009 to 2023.

3.1 Data Collection

Website searches were conducted specifically for policy evaluations studies of the BMBF and the BMWK that are listed in the federal budget plans. The policy evaluation studies were downloaded manually from the websites of the two ministries, the responsible project management agencies, the evaluating institutions as well as, in individual cases, other organizations, resulting in a total of 132 documents. Upon closer examination, 51 of these documents posed certain issues including duplicates, documents that did not actually refer to evaluation studies, brief summaries if a long version was available, studies lacking sufficient detail on their methodology, and early-stage policy evaluation studies that did not yet include quantitative assessments. All in this sense problematic documents were excluded from the sample and are therefore not considered in the subsequent analysis.

3.2 Coding

Guided by our research objectives we developed an initial coding scheme. To assess the coding scheme's effectiveness, we conducted an initial test coding of six randomly selected policy evaluation studies, by

⁴ For classic application of these designs see e.g. Card and Krueger (1994) or Abadie et al. (2010). For overviews, see e.g. Imbens and Wooldridge (2009), DiNardo and Lee (2011) or Abadie (2021).

two separate coders. Based on the initial test coding, the coding scheme was refined to address the identified concerns.

For descriptive coding and categorization, information such as project name, start and end dates, project funder and type of evaluator were taken from the title pages of the policy evaluation studies. In cases where this did not suffice, information was retrieved from the remainder of the policy evaluation study or the project's official website. For coding items that required a more nuanced understanding of the policy evaluation study, such as its type, the empirical methods utilized, and whether the policy evaluation study made causal claims, coders systematically reviewed the executive summary, conclusion, and any methodology chapters. Additionally, the table of contents were used to identify other chapters that might contain relevant information, which were then analyzed.

All policy evaluation studies underwent independent coding at two separate sites. Once both sites completed coding, any disparities in the codified items were resolved through a collaborative meeting. Instances where discrepancies proved challenging to reconcile were addressed by having coders review relevant sections of the respective policy evaluation study that led to the coding deviations.

3.3 Data Description

We compiled a total of 81 policy evaluation studies, focusing on the evaluation of public R&I support programs. Of these studies, 27 relate to the area of responsibility of the German Federal Ministry of Education and Research (BMBF) and 54 to the area of responsibility of the Federal Ministry for Economic Affairs and Climate Action (BMWK).

The timeline of these studies spans from 2009 to 2023. The median publication year of 2018 indicates a more recent focus on policy evaluation. Most of these studies, specifically 74 out of 81, concentrated on evaluating public support programs (e.g. funding measures), reflecting the federal government's significant emphasis on funding as a key policy tool. The remaining seven studies assessed various institutions, such as competence centers and working groups, offering insights into the structural and administrative aspects of policy implementation.

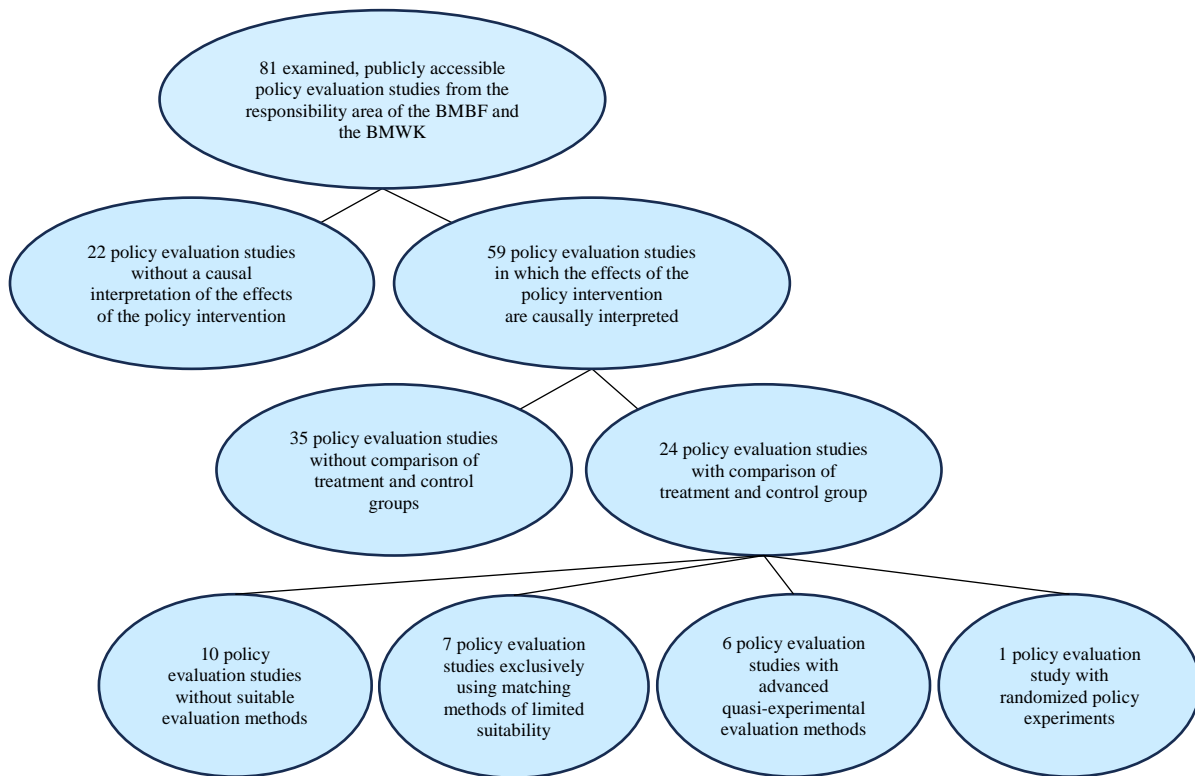
Studies in our sample also varied in terms of format. In total, 16 studies were presented as concise summaries, allowing quick understanding of key findings, while 65 studies were detailed, providing an in-depth view of methodologies and results.

4. Results

4.1 Use of Methods of Causal Analysis

In a first step, we analyzed the policy evaluation studies regarding their causal interpretation of the results and the methods they used (see Figure 1 and Table 1 for an overview). In the 81 policy evaluation studies, 59 studies (73%) interpreted the analyzed outcomes as causal effects. The remaining 22 studies (27%) did not quantitatively measure effects or interpret their findings as causal. Among the 59 studies which interpreted the effects as causal, only 24 compared treatment and control groups. Within this group, ten studies did not employ experimental or quasi-experimental research designs. Only a single study was based on randomized policy experiments. Another six studies used quasi-experimental designs such as Difference-in-Differences (five studies) and Regression Discontinuity (one study). Seven studies exclusively used Matching procedures. However, the causal interpretability of studies solely using Matching procedures is debatable, given the inherent limitations of this method.

Figure 1: Examination results of evaluation studies from the areas of responsibility of the BMBF and the BMWK



Taken together, we find that of the 81 policy evaluation studies, only 14 studies used (some) methods of causal analysis alongside causal interpretations. In contrast, 45 studies interpreted their results as causal despite using designs or methods that do not support such conclusions. This corresponds to 56% of all 81 analyzed studies and 76% of those 59 studies that interpreted the effects of the evaluated interventions as causal.

All 59 studies with causal interpretations reported either positive effects (40 studies) or at least partly positive effects (19 studies). We furthermore observed that 69 studies complemented their quantitative analyses with qualitative methods like document analysis and interviews, indicating efforts to provide a comprehensive evaluation.

Table 1: Policy evaluation studies interpretation of results

Number	Studies interpreting results ...	
	... as causal	... not as causal
	59	22
Studies using ...		
... causal analysis methods	14	
... no causal analysis methods	45	
Studies finding ...		
... positive effects	40	
... partly positive effects	19	

4.2 Evaluators

In the analysis of the 81 policy evaluation studies, evaluators were classified into various categories: PROs, private-sector companies (e.g., consulting firms), universities, government agencies (e.g., federal offices), and others. The distribution of involvement was as follows: PROs in 24 studies, private-sector companies in 62, universities in eleven, government agencies in seven, and other types of evaluators in four studies. We also found that 25 of the studies were collaboratively conducted by at least two different types of evaluators.

A notable observation from the analysis was the different inclination of the various evaluator types to use methods of causal analysis. PROs applied these methods in 42% of their studies, private-sector companies in about 18%, and universities in approximately 27%. Evaluations conducted by government agencies or other evaluators did not use causal analysis.

4.3 Temporal Trends

Analyzing the temporal trends in the policy evaluation studies published between 2009 and 2023, we observe some shifts over time (see Table 2). Of all 81 policy evaluation studies, 33 were published up to and including 2018, while 47 were published from 2019 onwards. For one study the publication date could not be determined. There was an increase in the commitment to perform impact assessments over time: approximately 67% of the studies up to 2018 had this commitment, compared to 87% of those published from 2019 onwards.

Regarding the application of methods of causal analysis, there was a notable increase as well: only 12% of studies up to 2018 used these methods and interpreted the results as causal, which went up to 21% in the studies from 2019 onwards. However, there was also an increase in studies making causal interpretations without appropriate causal analysis methods, from about 48% to 62%.

The involvement of the different organizational types of evaluators also evolved over time. The participation of PROs increased from seven (21%) policy evaluation studies up to 2018 to 17 (36%) studies since 2019. For private-sector companies, the increase was from 22 (67%) to 40 (85%) policy evaluation studies. Additionally, there was a notable rise in collaboration, from approximately 24% of all policy evaluation studies being collaborative up to 2018 to 36% since 2019.

Table 2: Temporal trends in policy evaluations studies

	Studies published ...	
	... until 2018	... from 2019
Number	33	47
Studies ...		
... with commitment to perform impact assessment	67%	87%
... using causal analysis methods	12%	21%
... interpreting effects as causal (without appropriate method)	48%	62%
... with collaboration	24%	36%

4.4 Barriers for Causal Analysis in Policy Evaluations

The challenges in employing causal analysis for policy evaluation are multifaceted, as evidenced by corresponding discussions in 32 out of 81 policy evaluation studies reviewed. These challenges substantially impair the ability to conduct robust and meaningful impact assessments.

One of the primary challenges is the lack of available data as well as the insufficient quality of data. This issue hampers the identification of appropriate control groups, which are essential for conducting meaningful impact assessments. Even when data might in principle be available at the onset of a policy measure, it may not be accessible or suitable for causal analysis when needed. Methodological constraints also pose relevant barriers. For instance, a meaningful comparison between treatment and control groups becomes untenable when the number of cases is too small. Additionally, in scenarios where all potential beneficiaries of a policy intervention are supported, establishing a control group becomes unfeasible. This limitation is further compounded when the expected effects of a policy intervention unfold over an extended period, making timely causal identification challenging. The evaluation of policy interventions often hinges on specific target variables, which must be measurable or at least estimable. In addition, unintended consequences, whether beneficial or detrimental, may not always be quantifiable, potentially leading to incomplete evaluations. Lastly, assessing the efficiency of policy interventions is even more demanding in terms of data availability, as it typically involves comparing the effects of various interventions. Although impact assessment can facilitate cost-benefit estimations in numerous instances, the overarching data requirements for efficiency assessment pose an additional challenge.

These obstacles underscore the critical need for robust methodological approaches and an encompassing data infrastructure to enhance the usefulness of policy evaluation studies.

5. Discussion

The above findings indicate that most policy evaluation studies we reviewed fall short of the methodological standards required for a substantial causal analysis. There has been a notable scarcity of randomized policy experiments. Furthermore, the employment of quasi-experimental research designs for causal analysis, beyond basic matching approaches, has been limited to a few instances, despite a slight uptick in their application more recently.

This gap in rigorous causal analysis undermines the ability of policymakers to draw lessons from the successes and failures of past initiatives, thereby hindering the improvement of future policy interventions. It is by no means a new or exclusively German issue. As early as 2013, the Scientific Advisory Board of the BMWi, the predecessor of today's BMWK, emphasized the importance of impact assessments for policy learning and outlined the requirements for evaluation (BMW 2013). A meta-evaluation from Sweden (Collin et al. 2022) analyzed 110 innovation policy evaluation studies published between 2005 and 2019. Similar to the results presented in our study, only nine out of the 110 studies applied methods of causal analysis. Results regarding evaluator types are likewise comparable. The majority of evaluators in Sweden were consulting firms, corresponding to our classification of private companies. Also similar to the results presented in our study, the analyzed evaluation studies in Collin et al. (2022) mainly attribute positive effects to the evaluation object. This contrasts with the broader scholarly literature, where policy interventions are often not assessed as being effective, despite a tendency in research studies to favor statistically significant results (see e.g., Vivalt (2020) or DellaVigna and Linos (2022)), due to the so-called publication bias (e.g. Franco et al. 2014).

Variations in the propensity of different evaluator types to employ methods of causal analysis are noteworthy. PROs, with a 42% application rate of these methods, as well as universities, with a 27% application rate, are mostly strongly focusing on methods of causal analysis. In contrast, private-sector companies, including consulting firms, applied these methods in only about 18% of their policy

evaluation studies. It is crucial to interpret these patterns with caution, as they likely reflect a complex interplay of factors including organizational priorities, differential resource availability and data access, varying levels of specialized methodological expertise for impact assessments, but also guidelines or requirements for the procurement of policy evaluation studies. Enhancing the capacity for applying methods of causal analysis across all evaluator types could involve targeted training and development initiatives, fostering collaborations across evaluator types, and encouraging a culture of continuous learning and methodological rigor in policy evaluation practices.

The temporal analysis of the studies, spanning from 2009 to 2023, reveals positive trends. There has been an overall increase in the commitment to perform impact assessments, which is particularly notable in the shift from 67% of studies aiming for impact assessments up to 2018, to 87% from 2019 onwards. This trend underscores a growing recognition of the importance of understanding the actual impacts of policy interventions. At the same time, the application of methods of causal analysis saw an increase from 12% of the policy evaluation studies published up to 2018 to 21% in the more recent studies. However, there was also a rise in studies making causal interpretations without employing appropriate causal analysis methods, from about 48% to 62%. This trend indicates the unabated need for a more stringent application of causal analysis methods in policy evaluations.

6. Conclusion

Evaluation is of crucial importance in R&I policy in order to review the effectiveness and efficiency of policy interventions. The findings from evaluations enable continuous policy learning and ongoing improvement of policy interventions. They also provide important insights into the utilization of public funds. However, to enable best usage of evaluation results, their limits also have to be recognized. Therefore, all authors of evaluation studies should inform their readers about whether obtained results can actually be attributed to the evaluated policy intervention, or whether results need to be interpreted with greater caution.

A number of measures can be taken to improve the preconditions for informative policy evaluation studies. Among these, the availability of all required data is of particular importance for impact assessment. Indeed, the lack of data availability is often cited as an obstacle in policy evaluation studies. It is therefore important to facilitate access to this data for evaluation purposes and to enable comprehensive access to detailed data. Data access should also be provided for replication purposes. In this way, subsequent methodological advances can be utilized for policy evaluation, longer-term outcomes can be assessed, and meta-analyses are enabled.

Our analysis also showed that not all policy evaluation studies are published or easily accessible. In addition, often only short versions of evaluation studies are published, which do not sufficiently present all relevant information to the readers. To enable effective policy learning and to build on experience, all evaluation studies should be published in both short and long versions. Publication should be independent of whether or not the evaluated policy interventions were found to be effective. Policy can learn from successful as well as from less successful interventions.

References

- Abadie, A.; Diamond, A.; and Hainmueller, J. (2010): Synthetic Control Methods for Comparative Case Studies. Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*. 105(490). S. 493–505.
- Abadie, A. (2021): Using Synthetic Controls. Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*. 59(2). S. 391–425.
- Athey, S.; Imbens, G.W. (2017): The State of Applied Econometrics: Causality and Policy Evaluation. *Journal of Economic Perspectives*, 31(2), 3-32.
- BMWi – Bundesministerium für Wirtschaft und Energie (2013): Evaluierung wirtschaftspolitischer Fördermaßnahmen als Element einer evidenzbasierten Wirtschaftspolitik. Gutachten des Wissenschaftlichen Beirats beim Bundesministerium für Wirtschaft und Energie. Berlin.
- Card, D.; Krueger, A.B. (1994): Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review*. 84(4). S. 772–793.
- Collin, E.; Sandström, C.; Wennberg, K. (2022): Evaluating Evaluations of Innovation Policy: Exploring Reliability, Methods, and Conflicts of Interest. In: Wennberg, K.; Sandström, C. (Eds.) *Questioning the Entrepreneurial State*. *International Studies in Entrepreneurship*, vol 53. Springer, Cham.
- DellaVigna, S.; Linos, E. (2022): RCTs to Scale. *Comprehensive Evidence from Two Nudge Units*. *Econometrica*. 90(1), 81–116.
- DiNardo, J.; Lee, D. (2011): Program Evaluation and Research Designs. In: Ashenfelter, O.; Card, D. (Eds.): *Handbook of Labor Economics*, Vol 4. Elsevier, 463-536.
- Edler, J.; Berger, M.; Dinges, M.; Gök, A. (2012): The Practice of Evaluation in Innovation Policy in Europe. *Research Evaluation*, 21(3), 167-182.
- Falck, O.; Koenen, J.; Lohse, T. (2015): Die Bedeutung von Programmevaluation in der deutschen Innovationspolitik. *ifo Schnelldienst*, 68(22), 52-55.
- Franco, A.; Malhotra, N.; Simonovits, G. (2014): Publication Bias in the Social Sciences: Unlocking the File Drawer. *Science*, 345(6203), 1502-1505.
- Gertler, P.J.; Martinez, S.; Premand, P.; Rawlings, L.B.; Vermeersch, J.M.J. (2016): *Impact Evaluation in Practice*. Second Edition. Washington, DC: Inter-American Development Bank and World Bank.
- ILO – International Labour Organization (2013): *Monitoring and Evaluation of Youth Employment Programmes: A Learning Guide* / International Labour Office, Youth Employment Programme, Employment Policy Department.
- Imbens, G.W.; Wooldridge, J.M. (2009): Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*. 47(1). S. 5–86.
- Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2010): *Handbook on Impact Evaluation: Quantitative Methods and Practices*. World Bank Publications.
- Scriven, M. (1991): Beyond Formative and Summative Evaluation. In: McLaughlin, G.W.; Phillips, D.C. (Eds.) *Evaluation and Education: A Quarter Century*, 19–64. Chicago, IL: University of Chicago Press.

Shadish, W. R.; Cook, T. D.; Campbell, D. T. (2002): Experimental and Quasi-experimental Designs for Generalized Causal Inference. Boston, MA: Houghton Mifflin.

Stockmann, R. (2006): Evaluation in Deutschland. In: Stockmann, R. (Ed.): Evaluationsforschung: Grundlagen und ausgewählte Forschungsfelder. Waxmann Verlag.

Verein für Socialpolitik (2015): Guidelines and recommendations of the VfS for ex post evaluations. <https://www.socialpolitik.de/de/guidelines-and-recommendations-vfs-ex-post-evaluations>. (Last accessed: March 05, 2024.)

Vivalt, E. (2020): How Much Can We Generalize from Impact Evaluations? Journal of the European Economic Association. 18(6), 3045–3089.

Woodward, J. (2007): Causal Models in the Social Sciences. In: Turner, S.; Risjord, M. (Eds.): Handbook of the Philosophy of Science, Philosophy of Anthropology and Sociology, 157-210.

Appendix: Analyzed Evaluation Studies

Alt, M.; Berger, M.; Bersch, J.; De Monte, E.; Egel, J.; Gottschalk, S.; Lotz, G.; Nachtigall, H.; Treperman, J. (2022): Ex-ante Analyse zum Förderprogramm „INVEST – Zuschuss für Wagniskapital“. Mannheim, Berlin: ZEW; Technopolis Group.

Astor, M.; Becher, G.; Heinrich, S.; Riesenberger, D. (2014): KMU-innovativ IKT: Chancen für Unternehmen. Förderung im Schwerpunkt „Softwaresysteme und Wissenstechnologien“. Auftraggeber Bundesministerium für Bildung und Forschung. Berlin: Prognos AG.

BAFA – Bundesamt für Wirtschaft und Ausfuhrkontrolle (2014): Evaluation der Energiesparberatung vor Ort. Endbericht. Auftraggeber: Bundesstelle für Energieeffizienz im Auftrag des Bundesministeriums für Wirtschaft und Energie. Eschborn: BAFA.

Baldauf, M.; Biermann, U.; Böhmer, S.; Fittkau, J.; Gollan, M.; Lübbers, T.; Wittenberg, J. (2021): Evaluation der Fördermaßnahmen „EXIST-Gründerstipendium“ und „EXIST-Forschungstransfer“ des Bundesministeriums für Wirtschaft und Energie im Förderzeitraum 2014 bis 2018. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Hamburg: Ramboll Management Consulting.

Barckhausen, A.; Joest, S.; Koselleck, M.; Paulus, J.; Quezada, A. (2016): Evaluation des Förderprogramms „Investitionszuschüsse zum Einsatz hocheffizienter Querschnittstechnologien im Mittelstand“. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Berlin: dena.

Bauer, P.; Boockmann, B.; Brändle, T.; Gensicke, M.; Hartmann, J.; Kreider, I.; Pfeiffer, I.; Zühlke, A. (2019): Evaluation des Bundesprogramms Bildungsprämie (BIP). Beauftragt vom Bundesministerium für Bildung und Forschung. München, Nürnberg, Tübingen: Kantar; f-bb; IAW.

Berger, F.; Biela, J.; Merkl, F.; Kahl, J.; Klement, B.; Dornbusch, F. (2020): Begleitende Evaluation des Modellvorhabens „ZIM-Kooperationsnetzwerke International“ sowie Untersuchung der Förderung von internationaler Zusammenarbeit bei Forschung und Entwicklung (FuE) im Zentralen Innovationsprogramm Mittelstand (ZIM). Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Berlin, Leipzig: Technopolis Group; Fraunhofer IMW.

Berger, F.; Galati, N.; Merkl, F.; Geyer, A.; Frank, A.; Lehmann-Brauns, C. (2020b): Begleitforschung zur Fördermaßnahme „Starke Fachhochschulen – Impuls für die Region“ (FH-Impuls). Abschlussbericht. Frankfurt a. M., Wien, Berlin: Technopolis; inspire research; Stifterverband für die Deutsche Wissenschaft.

Berger, F.; Heimer, T.; Tafreschi, D.; Teichler, T. (2016): Evaluierung des Kompetenzzentrums Innovative Beschaffung. Studie i Auftrag des Bundesministeriums für Wirtschaft und Energie. Frankfurt a. M.: Technopolis Group.

Berger, F.; Strähle, T.; Pschorn, L.; Waiblinger, F.; Lotz, G. (2022): Begleitende Evaluation des Innovationsprogramms für Geschäftsmodelle und Pionierlösungen (IGP). Abschlussbericht. Berlin; FRankfurt a. M.: Technopolis Group.

Biela, J.; Warta, K.; Galati, N.; Zingerle, S.; Klein-Franke, S. (2022): Evaluation des Professorinnenprogramms des Bundes und der Länder. Dritte Programmphase und Gesamtevaluation. Evaluationsbericht. Frankfurt a. M.: Technopolis.

Bießlich, S.; Engelhardt, S.v.; Kaufmann, P.; Kerlen, C.; Kind, S.; Kofler, J.; Marcher, A.; Nindl, E.; Robeck, M.; Rodriguez Rivera, K.; Zinke, G. (2021): Evaluation des Nationalen Programms für Weltraum und Innovation. Im Auftrag des Bundesministeriums für Wirtschaft und Energie. Berlin, Edinburgh, Wien: iit; Kerlen Evaluation; KMU Forschung Austria GmbH.

Blind, K.; Kulicke, M.; Heimer, T.; Kreutzer, S.; Nachtigall, H.; Pschorn, L.; Rammer, C.; Müller, B. (2021): Querschnittelevaluation „Unterstützungslandschaft für innovative Gründungen“. Auftraggeber:

Bildungsministerium für Bildung und Forschung, Bundesministerium für Wirtschaft und Energie. Karlsruhe, Frankfurt, Berlin, Mannheim: Fraunhofer ISI; Technopolis Group; ZEW.

BMBF – Bundesministerium für Bildung und Forschung (Hrsg.) (2012): Evaluation des Förderprogramms InnoProfile. Zwischenergebnisse für 32 Initiativen der ersten und zweiten InnoProfile-Förderrunde. Berlin: BMBF.

Bocek, F.; Vollborth, T.; Wolff von der Sahl, J. (2023): Wirksamkeit der geförderten FuE-Projekte und Kooperationsnetzwerke des Zentralen Innovationsprogramms Mittelstand (ZIM). Fokus: 2020 beendete ZIM-Projekte sowie bis 2020 geförderte ZIM-Kooperationsnetzwerke. Ergebnisse der Befragung aus dem Jahr 2022. Expertise im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz. Eschborn: RKW Kompetenzzentrum.

Bothor, S.; Bubeck, S.; Eberl, T.; Fahl, U.; Gottschling, J.; Heyden, E.; Steurer, M.; Oertel, B.; Scharp, M.; Meyer, S.; Knoll, M.; (2014): Evaluation ausgewählter Maßnahmen zur Energiewende. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Stuttgart, Berlin: IER; igt.

Brachert, M.; Brautzsch, H.-U.; Dettmann, E.; Giebler, A.; Schneider, L.; Titze, M. (2020): „Evaluation der Gemeinschaftsaufgabe 'Verbesserung der regionalen Wirtschaftsstruktur' (GRW)“ durch einzelbetriebliche Erfolgskontrolle. Endbericht. Finanziert durch das Bundesministerium für Wirtschaft und Energie. Halle (Saale), Coburg: IWH; Hochschule Coburg.

Braeseke, G.; Rieckhoff, S.; Engelmann, F.; Lingott, N.; Dean, M.; Pörschmann-Schreiber, U. (2020): Evaluation der Förderung von Modellprojekten zur Gewinnung von jungen Menschen aus Vietnam zur Ausbildung in der Pflege in Deutschland. Kurzfassung. Im Auftrag des Bundesministeriums für Wirtschaft und Energie. Berlin: IGES.

Buhr, R.; Cleuvers, B.A. (2014): Übergreifende Evaluation der Begabtenförderung des BMBF im Programmfeld "Berufliche Talente". Berlin: iit; FiBS.

Bührer, S.; Walz, R.; Seus, S.; Astor, M.; Stehnen, T.; Malik, F. (2020): Evaluation der BMBF-Rahmenprogramme Forschung für die Nachhaltigkeit FONA 1 (2005-2009) und Forschung für Nachhaltige Entwicklungen FONA 2 (2010-2014). Abschlussbericht. Studie/Gutachten im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF). Karlsruhe, Berlin: Fraunhofer ISI; Prognos AG.

Caliendo, M.; Kritikos, A.; Künn, S.; Loersch, C.; Schröder, H.; Schütz, H. (2014): Evaluation der Programme „Gründercoaching Deutschland“ und „Gründercoaching Deutschland – Gründungen aus der Arbeitslosigkeit“. Im Auftrag des Bundesministeriums für Wirtschaft und Technologie und des Bundesministeriums für Arbeit und Soziales in Kooperation mit der Kreditanstalt für Wiederaufbau. Potsdam: Universität Potsdam; IZA; DIW; infas.

Caumanns, T.; Flambard, A.R.; Geschewski, A.; Greif, A.; Heidelberger, M.; Schumacher, G.; Weber, M. (2019): Ex-Ante Evaluation für ein Technologietransfer-Programm Leichtbau. Schlussbericht. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Jülich: PTJ.

Cemerin, M.; Nagel, E. (2009): Evaluation des Programms zur Förderung der Teilnahme junger innovativer Unternehmen an internationalen Leitmessen in Deutschland. Schlussbericht. Eschborn: BAFA.

Clemens, M.; Schulze Düding, J.; Goerge, M.; Kemfert, C.; Michelsen, C.; Neuhoff, J. (2021): Quantitative und qualitative Wirkungsanalyse der Maßnahmen des Deutschen Aufbau- und Resilienzplans (DARP). Endbericht. Aktualisierte Version. Kurzexpertise im Auftrag des Bundesfinanzministeriums. Berlin: DIW.

Daimer, S.; Berghäuser, H.; Braun, A.; Korte, S.; Kripp, K.; Kulicke, M.; Meyer, N.; Rijkers-Defrasne, S. (2014): Begleitende Evaluierung der Fördermaßnahme „Validierung des Innovationspotenzials wissenschaftlicher Forschung – VIP“. Abschlussbericht. Berlin; Düsseldorf: Fraunhofer ISI; VDI TZ.

- Durdel, A.; Morgenstern, J.; Boll, P.; Tölle, J.; Heiberg, J. (2023): Evaluation zum Bürgerrat Forschung und Citizen Panel. Abschlussbericht – Zusammenfassung. Beauftragt vom Bundesministerium für Bildung und Forschung. Hamburg: Ramboll Management Consulting.
- Ekert, S.; Schüren, V.; Bode, A. (2016): Evaluation des Programms go-cluster des Bundesministeriums für Wirtschaft und Energie (BMWi). Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Berlin, Bickenbach: InterVal; CONABO.
- Engelhardt, S.v.; Cremers, N.; Kerlen, C.; Kilic, Ü.; Künzel, M.; Robeck, M.; Rodriguez Rivera, K.; Sobek, S.; Wangler, L.; Zinke, G. (2023): Evaluation des Rahmenprogramms der Bundesregierung Forschung für die zivile Sicherheit 2018-2023. Im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF). Berlin, Edinburgh: IIT; Kerlen Evaluation Ltd.
- Galati, N.; Nieth, L.; Stalla, M.; Berger, F. (2021): Evaluation der Beteiligung an den Erasmus+ Kooperationsprojekten im deutschen Hochschulbereich und der projektbezogenen Informations- und Beratungsarbeit der NA DAAD. Erasmus+ Programmgeneration 2014-2020. Frankfurt a. M.: Technopolis.
- Geyer, A.; Berger, F.; Dudenbostel, T. (2016): Empirische Begleitung des Programms „Forschung an Fachhochschulen“. Abschlussbericht. Wien: Technopolis Group.
- Geyer, A.; Heimer, T.; Treperman, J. (2016): Evaluation des High-Tech-Gründerfonds. Schlussbericht. Wien: Technopolis Group.
- Görlitz, K.; Tamm, M.; Becker, C.; Lübbers, T.; Schiel, S.; Schröder, H. (2012): Datenmonitoring und Evaluation des Programms „Bildungsprämie“. Endbericht. Forschungsprojekt im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF). Essen: RWI; GIB; infas.
- Grebe, T.; Schüren, V.; Ekert, S. (2016): Evaluierung des Markterschließungsprogramms für kleine und mittlere Unternehmen. Kurzfassung des Endberichts. Auftraggeber: Bundesamt für Wirtschaft und Ausfuhrkontrolle. Berlin: InterVal.
- Günther, J.; Nulsch, N.; Lang, C.; Schwartz, M.; Hornyeh, C.; Rammer, C.; Hud, M.; Schliesser, P. (2012): Evaluierung des BMWi-Programms „FuE-Förderung gemeinnütziger externer Industrieforschungseinrichtungen Ostdeutschlands – Innovationskompetenz Ost (INNO-KOM-Ost)“ einschließlich des Modellvorhabens Investitionszuschuss technische Infrastruktur. Studie im Auftrag des Bundesministeriums für Wirtschaft und Technologie. Halle (Saale): ZEW; IWH.
- Guidehouse Energy Germany GmbH (2020): Evaluation der Beihilfen für indirekte Kosten des EU-Emissionshandels im Rahmen der Strompreiskompensation für den Zeitraum 2013-2017. Bericht an: Bundesministerium für Wirtschaft und Energie. Köln: Guidehouse.
- Hany, E. (o.J.): Begleitende Evaluierung der ideellen Fördermaßnahmen der Begabtenförderungswerke im Rahmen der Begabtenförderung im Hochschulbereich. Abschlussbericht. Erfurt: Universität Erfurt.
- Happ, A. (2018): Evaluation des BQ-Portals. Endbericht. Auftraggeber: Bundesministeriums für Wirtschaft und Energie. Eschborn (BAFA).
- Happ, A.; Lüders, B.L. (2014): Evaluation der Arbeitsgemeinschaft für wirtschaftliche Verwaltung e.V. (AWV). Auftraggeber: Bundesministerium für Wirtschaft und Energie. Eschborn: BAFA.
- Heinrich, S.; Hoch, M.; Thamling, N.; Thormeyer, C.; Tschumi, D.; Rau, D.; Weinert, K. (2020): Evaluation der KfW-Förderprogramme EBS NWG für den Förderjahrgang 2015. Evaluation der KfW-Förderprogramme zum energieeffizienten Bauen und Sanieren für Nichtwohngebäude (EBS NWG) im Förderzeitraum 2015 bis 2018. Basel u.a.: Prognos AG.
- Heinrich, S.; Langreder, N.; Thormeyer, C.; Grodeke, A.-M.; Holm, A.; Kokolsky, K. (2022): Evaluation der Förderprogramme EBS WG im Förderzeitraum 2018. Evaluation der Förderprogramme „Energieeffizient Bauen und Sanieren“ für Wohngebäude (EBS WG) als Teil CO₂-

Gebäudesanierungsprogramms des BMWi im Förderzeitraum 2018 bis 2020. Im Auftrag des Bundesministeriums für Wirtschaft und Energie. Basel u.a.: Prognos AG.

Heinrich, S.; Langreder, N.; Thormeyer, C.; Seefeldt, F.; Brischke, L.-A.; Pehnt, M. (2019): Ex post-Analyse des Pilotprogramms STEP up! Beitrag zur Erfolgskontrolle finanzwirksamer Maßnahmen für das Pilotprogramm „Stromeinsparungen im Rahmen wettbewerblicher Ausschreibungen: Stromeffizienzpotentiale nutzen“ (STEP up!). Im Auftrag des Bundesstelle für Energieeffizienz im BAFA. Basel u.a.: ifeu; Prognos AG.

Heinrich, S.; Saad, N.; Grodeke, A.-M.; Hoch, M.; Langreder, N.; Spillmann, T.; Thamling, N.; Theodorou, K.; Thormeyer, C.; Weinert, K. (2022): Evaluation des Förderprogramms KfW 433. Evaluation des Förderprogramms Energieeffizient Bauen und Sanieren - Zuschuss Brennstoffzelle (KfW 433) im Förderzeitraum 2016 bis 2020. Im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz. Basel u.a.: Prognos AG.

Hirzel, S.; Antoni, O.; Heinrich, S.; Jessing, D.; Paar, A.; Pehnt, M.; Schломann, B.; Voswinkel, B. (2019): Evaluierung und Weiterentwicklung des Energieeffizienzfonds. Abschlussbericht – Langfassung. Karlsruhe, Heidelberg, Basel, Würzburg: Fraunhofer ISI; ifeu; Prognos AG; Stiftung Umweltenergierecht.

Illge, L.; Oertel, B.; Kenkmann, T.; Schumacher, K. (2020): Evaluation der Maßnahme „Nationales Effizienzlabel für Heizungsanlagen“. Zusammenfassung. Berlin: igt; Öko-Institut e.V.

Jähne, J.C.; del Estal, M.; Lemper, B.; Kramer, H.; Stein, M. (2017): Evaluierung des Forschungsprogramms „Maritime Technologien der nächsten Generation“. Kurzfassung des Abschlussberichts. Berlin: BearingPoint.

Jarren, O.; Baumbach, S.; Graf, F.W.; Mau, S.; Mittler, B.; Neitzel, S. (2017): Evaluierung des BMBF-Rahmenprogramms „Geistes-, Kultur- und Sozialwissenschaften“. Abschlussbericht. Berlin, Graz, Wien: Evaconsult; Joanneum Research.

Jung, D.; Zimmermann, F. (2018): Managerfortbildungsprogramm des Bundesministeriums für Wirtschaft und Energie. Evaluierung der Durchführungsperiode 2010-2012. Für die Deutsche Gesellschaft für internationale Zusammenarbeit (GIZ). Berlin: denkmodell GmbH.

Kaufmann, P.; Bittschi, B.; Depner, H.; Fischl, I.; Kaufmann, J.; Nindl, E.; Ruhland, S.; Sellner, R.; Struß, V.; Vollborth, T.; Wolff von der Sahl (2019): Evaluation des Zentralen Innovationsprogramms Mittelstand (ZIM). Richtlinie 2015. Endbericht. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Wien: KMU Forschung Austria.

Keil, J.; Steffen, H.; Theunissen, K.; Hagedorn, K. (2019): Evaluierung des Förderprogrammes „INVEST - Zuschuss für Wagniskapital“. Abschlussbericht. Im Auftrag des Bundesministeriums für Wirtschaft und Energie. Dortmund: ExperConsult.

Kreft, U.; Liesenfeld, J.; Uske, H. (2019): „Gemeinsam in die Ausbildung“ (GidA). Schlussbericht der wissenschaftlichen Begleitung/Evaluation zum Modellprojekt der Otto Benecke Stiftung e.V. Gefördert vom Bundesministerium für Wirtschaft und Energie. Duisburg: RISP.

Kuckshinrichs, W.; Aniello, G. (2018): Wirkungen der KfW-Programme „Energieeffizient Bauen“, „Energieeffizient Sanieren“, "IKK/IKU – Energieeffizient Bauen und Sanieren" und „KfW-Energieeffizienzprogramm – Energieeffizient Bauen und Sanieren" auf öffentliche Haushalte im Förderjahr 2016. Kurzgutachten der KfW. Jülich: Forschungszentrum Jülich; IEK-STE.

Kulicke, M.; Dornbusch, F.; Berghäuser, H.; Blind, K.; Kaiser, T.; Krukenberg, E. (2014): Erfolgskontrolle des Programms SIGNO – „Schutz von Ideen für die Gewerbliche Nutzung“ des Bundesministeriums für Wirtschaft und Energie. Endbericht. Öffentliche Fassung. Karlsruhe, Berlin: Fraunhofer ISI; Fraunhofer FOKUS.

- Kulicke, M.; Meyer, N.; Stahlecker, T.; Jackwerth, T. (2019): Evaluation des Programms WIPANO – „Wissens- und Technologietransfer durch Patente und Normen“. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Karlsruhe: Fraunhofer ISI.
- Lehmann, K.; Knittel, T.; Mohr, S.; Nell, P. (2018): Kultur macht stark. Bündnisse für Bildung (2013 bis 2017). Beauftragt vom Bundesministerium für Bildung und Forschung. Freiburg, Düsseldorf: Prognos AG.
- Lübbers, T.; Böhmer, S.; Morgenstern, J.; Baldauf, M.; Sassmann, Y.; Adam, H. (2021): Evaluation des Förderprogramms „go-digital“. Endbericht. Hamburg: Ramboll Management Consulting.
- Michels, J.; Gaedckens, C.; Wallau, F.; Ebert, L.; Gordon, J.; Vasilakis, K.; Bunge, L.; Kuhlenkamp, F. (2021): Evaluation der Industriellen Gemeinschaftsforschung (IGF) des Bundesministeriums für Wirtschaft und Energie im Förderzeitraum 01.09.2017 bis 31.12.2020. Abschlussbericht zur Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Köln: Kienbaum.
- Michels, J.; Wallau, F.; Ebert, L.; Gordon, J.; Hoffman, A.; Risch, L. (2021): Evaluation des Förderprogrammes zur FuE-Förderung gemeinnütziger externer Industrieforschungseinrichtungen – Innovationskompetenz (INNO-KOM). Abschlussbericht zur Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Köln: Kienbaum.
- Neßler, C.; Schmidt, U.; Wagner, L. (2023): Zwischenevaluation „Leistung macht Schule“. Zusammenfassung der Ergebnisse. Mainz: ZQ; JGU.
- Piégsa, A.; Kirchner, A.; Koziel, S.; Kreidelmeyer, S. (2018): Evaluierung der Zielwerte der Vereinbarung zwischen der Regierung der Bundesrepublik Deutschland und der deutschen Wirtschaft zur Steigerung der Energieeffizienz vom 1.8.2012 für die Bezugsjahre 2017-2020. Basel: Prognos AG.
- Schlomann, B.; Neusel, L.; Heinrich, S.; Grodeke, A.-M.; Radgen, P.; Grundler, A.T.A.; Schumacher, K.; Ludig, S. (2023): Evaluation der „Bundesförderung für Energieeffizienz in der Wirtschaft“ (Zuschuss und Kredit/Förderwettbewerb). Projekt für das Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA). Karlsruhe, Basel, Stuttgart, Berlin: Fraunhofer ISI; Prognos AG; IER Universität Stuttgart; Öko-Institut e.V.
- PwC – PricewaterhouseCoopers (2017): Evaluation der Energieeinsparberatung und der Energie-Checks der Verbraucherzentralen für das Bundesamt für Wirtschaft und Ausfuhrkontrolle. Endbericht. Frankfurt: PwC.
- PwC – PricewaterhouseCoopers (2018): Evaluierung der Förderprogramme „Energieberatung im Mittelstand“ und „Energieberatung für Nichtwohngebäude von Kommunen und gemeinnützigen Organisationen“ für das Bundesamt für Wirtschaft und Ausfuhrkontrolle. Endbericht. Frankfurt: PwC.
- PwC – PricewaterhouseCoopers (2019): Evaluation der Energieberatung für Wohngebäude. Für das Bundesamt für Wirtschaft und Ausfuhrkontrolle. Frankfurt: PwC.
- Ramboll Management Consulting (2022): Begleitende Evaluation der Förderinitiative „Forschungscampus – öffentlich-private Partnerschaft für Innovationen“. Beauftragt vom Bundesministerium für Bildung und Forschung. Hamburg: Ramboll Management Consulting.
- Rammer, C.; Gottschalk, S.; Fünér, L. (2020): Studie zur Evaluation des KfW-Förderprogramms „ERP-Digitalisierungs- und Innovationskredit“. Abschlussbericht. Mannheim: ZEW.
- Rammer, C.; Aschoff, B.; Crass, D.; Eckert, T.; Licht, G.; Astor, M.; Heinrich, S.; Riesenber, D.; Woywode, M.; Strohmeyer, R.; Tonoyan, V.; Ruffer, N. (2011): Systemevaluierung „KMU-innovativ“. Begleit- und Wirkungsforschung zur Hightech-Strategie. Abschlussbericht. Mannheim, Berlin: ZEW; Prognos AG; ifm.
- Ratschinski, G.; Sommer, J.; Eckhardt, C.; Struck, P. (2017): Evaluation des BMBF-Programms zur „Förderung der Berufsorientierung in überbetrieblichen und vergleichbaren Berufsbildungsstätten“.

Endbericht. Gefördert vom Bundesministerium für Bildung und Forschung. Berlin, Hannover, Duisburg: InterVal; Leibniz Universität Hannover; qualiNETZ.

Reineck, C.; Suerkemper, F.; Vondung, F.; Wörten, C.; Thomas, S. (2022): Abschlussbericht zur Evaluation der Richtlinie über die Förderung der Heizungsoptimierung durch hocheffiziente Pumpen und hydraulischen Abgleich. Wuppertal, Berlin: Arepo GmbH; Wuppertal Institut.

Ribka, L. (2022): Evaluation der IW JUNIOR-Programme. Endbericht. Auftraggeber: Bundesministerium für Wirtschaft und Klimaschutz (BMWK). Eschborn: BAFA.

Rudiger, G.; Beauducel, A.; Engin-Stock, T.; Haverkamp, N.; Krüger, T.; Mehlis, K. (2015): Evaluation nach § 15 Stipendienprogramm-Gesetz. Abschlussbericht. Gefördert vom Bundesministerium für Bildung und Forschung. Bonn: uzbonn GmbH; ZEM.

Schmidt, U.; Besch, C.; Schulze, K.; Heinzemann, S.; Andersson, M. (2020): Evaluation des Bund-Länder-Programms für bessere Studienbedingungen und mehr Qualität in der Lehre (Qualitätspakt Lehre). Abschlussbericht über den gesamten Förderzeitraum 2011-2020. Berlin, Mainz: ZQ Uni Mainz; Prognos AG.

Schwierting, T.; Zitzmann, I.; Greif, A.; Hoffmann, D.; Jossek, R.; Kating, P.; Rauschen, S.; Zimmermann, R. (2020): Ex-ante-Evaluation für ein Förderprogramm „Industrielle Bioökonomie“. Auftrag des Bundesministeriums für Wirtschaft und Energie. Jülich: PTJ.

Sommer, J.; Rennert, C. (2020): Endbericht der wissenschaftlichen Begleitung zur Interventionsstudie Potenzialanalyse (ISPA). Berlin: InterVal.

Stehnken, T.; Malik, F.; Danneil, T.; Reichert, J.; Wichmann, V. (2022): Evaluation der Bundeszuwendungen an die Germany Trade and Invest GmbH (GTAI). Zusammenfassung. Im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz. Berlin: Prognos AG.

Stehnken, T.; Neumann, M.; Talamo, J.A.; Astor, M.; Ploder, M.; Breitfuss-Loidl, M.; Deixelberger, D.; Linshalm, E.; Forchert, C.-E.; Blume, A.; Viebranz, T. (2022): Evaluation des Rahmenprogramms „Mikroelektronik aus Deutschland – Innovationstreiber der Digitalisierung“. Erstellt im Auftrag des Bundesministeriums für Bildung und Forschung. Berlin: Prognos AG.

Stehnken, T.; Astor, M.; Neumann, M.; Talamo, J.A.; Ploder, M.; Breitfuss-Loidl, M.; Rosenball, R.; Rammer, C.; Gottschalk, S.; Peters, B. (2021): Evaluation des BMBF-Förderprogramms „Photonik Forschung Deutschland“. Abschlussbericht. Kurzfassung. Im Auftrag des Bundesministeriums für Bildung und Forschung. Berlin: Prognos AG; Joanneum Research; ZEW.

Wangler, L.; Kerlen, C.; Bräuninger, M.; Meyer, S.; Teuber, M.-O. (2019): Evaluation des Luftfahrtforschungsprogramms (LuFo). Zusammenfassung zentraler Ergebnisse und Handlungsempfehlungen. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Berlin: iit.

Weber, A. (2018): Evaluation des Kompetenzzentrums Kultur- und Kreativwirtschaft. Endbericht. Auftraggeber: Bundesministerium für Wirtschaft und Energie. Eschborn: BAFA.

Wessels, J.; Brylla, K.; Kerlen, C.; Kind, S.; Neumann, K.; Wangler, L.; Schließer, R.; Pieper, O. (2017): Evaluation des Förderprogramms „ELEKTRO POWER“. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Berlin: iit.

Winterhager, N.; Birner, N.; Krabel, S.; Mozhova, A.; Shajek, A.; Breitbach, M.; Lüthje, J. (2017): Untersuchungen der Auswirkungen des Hochschulpakts 2020. Abschlussbericht/Kurzfassung. Berlin: iit.

Zech, D.; Ullrich, S.; Wühlbeck, H.-F.; Stuible, A.; Wapler, J.; Valenbreder, P.; Meyer, R.; Miara, M.; Hartmann, H.; Reisinger, K.; Werner, F.; Orozaliev, J.; Vajen, K.; Schuhmann, E.; Erler, R.; Heinrich, P.; Schröder, G. (2019): Evaluation des Marktanreizprogramms zur Förderung von Maßnahmen zur

Nutzung erneuerbarer Energien im Wärmemarkt im Förderzeitraum 2015 bis 2018. Evaluation des Förderjahres 2018. Ausarbeitung im Auftrag des Bundesministeriums für Wirtschaft und Energie. Stuttgart: Fichtner.