

Hanna Hottenrott, Bettina Peters

Evaluation staatlicher Fördermaßnahmen in Krisenzeiten: Zwischen Anspruch und Realität

Die deutsche Wirtschaftspolitik reagiert auf Krisensituationen mit einem bekannten Instrumentenmix aus direkten Förderungen, steuerlichen Anreizen und Liquiditätsunterstützungen. Trotz der gesetzlichen Verpflichtung zur Wirtschaftlichkeits- und Erfolgskontrolle bleibt die Evaluation dieser Maßnahmen häufig methodisch eingeschränkt. Neben der Herausforderung, klassische Evaluationsmethoden in Krisensituationen anzuwenden, erschweren gravierende Datenlücken die Evaluation von Maßnahmen zur Stützung der Wirtschaft nach wissenschaftlichen Standards. Die Verfügbarkeit, die Verknüpfbarkeit und die Qualität von Förder- und Unternehmensdaten sind unzureichend, was die Identifikation von Wirkungsketten behindert. Eine Verbesserung erfordert eine systematische Integration von Evaluationsansätzen bereits in die Programmdesignphase – inklusive klarer Zieldefinition, datenbasierten Monitorings und verpflichtender Datenlieferungen. So lässt sich eine evidenzbasierte Politikberatung auch unter Unsicherheit und zeitlichem Druck gewährleisten.

Den verschiedenen Krisen der jüngsten Zeit wurde mit umfangreichen staatlichen Rettungspaketen begegnet, die die Auswirkungen auf Unternehmen, Beschäftigte und Verbraucher abmildern sollten. Am Beispiel der Coronakrise hat sich zudem gezeigt, dass Bundes- und Landesregierungen schnell Entscheidungen treffen und in kürzester Zeit beträchtliche finanzielle Mittel auf den Weg bringen konnten. Spätestens seit Anfang der 2020er Jahre befindet sich die deutsche Wirtschaft jedoch auch in einer grundlegenden Wettbewerbsfähigkeitskrise. Insbesondere zunehmende Handelsrestriktionen auf Exportmärkten, abrupt steigende Energiepreise, technologische Pfadabhängigkeiten und der China-Schock 2.0 machen es deutschen Unternehmen aktuell schwer, im internationalen Wettbewerb zu bestehen. Die Reaktion der Bundesregierung darauf mit industriepolitischen Maßnahmen und der gezielten Förderung von strategisch als wichtig angesehenen Industrien und Technologien ist wenig überraschend.

Doch wo viel staatliches Geld fließt, besteht auch ein öffentliches Interesse an der Untersuchung der Wirksamkeit und Wirkung der Maßnahmen. Hätte ein gefördertes Unternehmen ohne die Förderung anders agiert? Haben

Kurzarbeitergeld oder Überbrückungshilfen tatsächlich Insolvenzen verhindert; haben steuerliche Investitionsanreize oder Subventionen Aktivitäten unterschiedlicher Art tatsächlich stabilisiert oder wurden lediglich Mitnahmeeffekte finanziert? Diese Fragen sind in Krisenzeiten besonders schwer zu beantworten, weil sowohl ein sich drastisch verändertes Umfeld als auch die ergriffenen Maßnahmen gleichzeitig auf die Akteure einwirken. Hinzu kommt, dass die Datenbasis vielfach unvollständig,

Prof. Dr. Hanna Hottenrott leitet den Forschungsbereich Innovationsökonomik und Unternehmensdynamik am ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung und ist Professorin für Innovationsökonomik an der School of Management der Technischen Universität München.

Prof. Dr. Bettina Peters ist stellvertretende Leiterin des ZEW-Forschungsbereichs Innovationsökonomik und Unternehmensdynamik sowie Honorarprofessorin für Innovation an der Fakultät für Rechts-, Wirtschafts- und Finanzwissenschaften der Universität Luxemburg.

© Der/die Autor:in 2026. Open Access: Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht (creativecommons.org/licenses/by/4.0/deed.de).

Open Access wird durch die ZBW – Leibniz-Informationszentrum Wirtschaft gefördert.

widersprüchlich oder schlichtweg nicht vorhanden ist. Gleichwohl verlangt die Politik nach Orientierung. So sieht sich die Wirtschaftsforschung insbesondere in Krisenzeiten mit einer Erwartungshaltung konfrontiert, die ihrem methodischen Selbstbild eigentlich widerspricht: schnelle, klare Aussagen zu treffen, obwohl die Datengrundlage dafür eigentlich noch nicht reif ist.

Evaluationen als wissenschaftlicher Input in die Politik

Über die verschiedenen Krisen hinweg zeigt sich ein bemerkenswert konstanter Instrumentenmix. Die Politik greift immer wieder auf ähnliche Instrumente wie Zuschüsse, Förderkredite, Preisbremsen, Garantien und Steuervergünstigungen zurück. Die Unterschiede liegen jeweils eher im Umfang und in der konkreten Ausgestaltung der Maßnahmen. Umso wichtiger ist die Frage, welche Lehren aus früheren Krisen gezogen werden, inwieweit wissenschaftliche Evaluationen in die Gestaltung zukünftiger Maßnahmen einfließen und welchen Beitrag sie damit zu fundierten politischen Entscheidungen leisten können.

Die Evaluation von Maßnahmen der Bundesregierung ist klar geregelt. § 7 Abs. 2 der Bundeshaushaltsordnung (BHO) verpflichtet den Bund für alle finanzwirksamen Maßnahmen angemessene Wirtschaftlichkeitsuntersuchungen durchzuführen. Die Verwaltungsvorschriften zu § 7 BHO (VV-BHO) konkretisieren diese Vorgabe. Sie schreiben sowohl eine begleitende Erfolgskontrolle während der Durchführung als auch eine abschließende Erfolgskontrolle nach Beendigung einer finanzwirksamen Maßnahme vor. Dabei sollen drei Dinge geprüft werden. *Erstens*, ob und in welchem Ausmaß die angestrebten Ziele erreicht wurden (Zielerreichungskontrolle). *Zweitens*, ob die Maßnahme ursächlich für die Zielerreichung war und ob es unbeabsichtigte Nebenwirkungen gab (Wirkungskontrolle). Und *drittens*, ob die Maßnahme wirtschaftlich im Sinne eines angemessenen Ressourceneinsatzes umgesetzt wurde (Wirtschaftlichkeitskontrolle). Das entspricht im Kern genau dem Evaluationsdesign, das auch den höchsten wissenschaftlichen Standards genügt: Zielerreichung, Kausalität und eine Kosten-Nutzen-Abschätzung.

Doch wie sieht es in der Realität aus? Die Praxis bleibt oft hinter diesen Ansprüchen zurück. Der Bundesrechnungshof kommt in einer Querschnittsprüfung im Jahr 2023 zu dem Schluss, dass die Bundesverwaltung den rechtlichen Vorgaben zu § 7 BHO häufig nicht oder nur unzureichend nachkommt, unter anderem wegen mangelnder methodischer Vorgaben (Bundesrechnungshof, 2023). Ebenso kritisierte die Expertenkommission Forschung und Innovation (EFI) im Jahresgutachten 2024

die mangelnde wissenschaftliche Tiefe bei der Evaluation von deutschen Fördermaßnahmen. Sie untersuchte dazu 81 Evaluationsstudien forschungs- und innovationspolitischer Maßnahmen und Instrumente, die im Zeitraum 2009 bis 2023 vom Bundesministerium für Bildung und Forschung (BMBF) und Bundesministerium für Wirtschaft und Klimaschutz (BMWK) in Auftrag gegeben wurden. Die EFI kritisierte insbesondere, dass die Evaluationsstudien Aussagen über die Wirksamkeit treffen, ohne belastbare kausalanalytische Methoden zu verwenden. Häufig ließen sich die Effekte nicht eindeutig auf die Förderung zurückführen. Sie fordert daher eine deutliche Verbesserung der Evaluationsstandards durch fundiertere Kausalanalysen mit Kontrollgruppenansätzen, um die tatsächliche Wirksamkeit von Innovations- und Forschungspolitik belastbar messen zu können (EFI, 2024).

Evaluation von politischen Maßnahmen als Herausforderung

Doch warum ist das so? Im Grunde stehen einer Evaluation nach wissenschaftlichen Standards häufig drei Herausforderungen entgegen: *Erstens* fehlt vielfach eine geeignete Datengrundlage. *Zweitens* fallen nicht selten der Evaluationszeitraum und der Zeitpunkt, zu dem sich die Wirkungen entfalten, auseinander. Und *drittens* ist die Realität meist komplizierter, als es klassische Evaluationsmethoden abbilden können.

Klassische Evaluationsmethoden – etwa Differenz-von-Differenzen-Schätzungen, Propensity-Score-Matching und randomisierte Kontrollgruppenstudien – beruhen auf einem einfachen Prinzip. Die Wirkung einer Maßnahme lässt sich dadurch kausal bestimmen, dass eine geförderte Gruppe einer möglichst ähnlichen, aber nicht geförderten Kontrollgruppe gegenübergestellt wird. In einer Krise wie beispielsweise der Coronapandemie war diese Voraussetzung aber häufig nicht erfüllt: Nahezu alle Unternehmen einer Branche – etwa in der Gastronomie oder im Einzelhandel – waren von den Schließungsanordnungen betroffen und wurden daher gleichzeitig gefördert. Eine vergleichbare, aber unbeeinflusste Kontrollgruppe existierte damit oft nicht. Hinzu kommen Übertragungseffekte zwischen Unternehmen entlang der Wertschöpfungsketten. Die Evaluation steht damit vor dem Problem, dass sich Förder- und Krisenwirkung kaum trennen lassen. Hat ein Betrieb nach Erhalt von Coronahilfen überlebt, bleibt unklar, ob er auch ohne Förderung überlebt hätte oder ob die staatliche Förderung den Ausschlag gegeben hat.

Eine weitere Herausforderung für Evaluationen besteht darin, dass in Krisenzeiten Förderprogramme häufig noch während ihrer Laufzeit angepasst werden. Fördersätze,

Antragsvoraussetzungen oder Zielgruppen können sich innerhalb weniger Monate mehrfach ändern. Für die Evaluation bedeutet dies, dass die zu evaluierende Maßnahme selbst nicht stabil ist. Klassische Ansätze verlieren an Aussagekraft, weil unklar bleibt, welche konkrete Ausgestaltung des Programms für die beobachteten Effekte verantwortlich ist.

Ohne geeignete Daten keine zuverlässige Evaluation

Die Erfahrungen am ZEW mit Evaluationen in verschiedensten Politikbereichen und für unterschiedliche Maßnahmentypen zeigen vor allem eines: Das Hauptproblem für eine Evaluation nach wissenschaftlichen Standards sind die eingeschränkte Verfügbarkeit und Qualität von Daten. Für eine solide Untersuchung von Zielerreichung, Wirksamkeit und Wirtschaftlichkeit sind umfangreiche Informationen erforderlich, die in der Praxis nur sehr selten vorliegen. Die Verantwortung für die Bereitstellung solcher Informationen liegt primär beim Staat als Auftraggeber. Mangels geeigneter Daten behelfen sich Evaluationen daher oft mit dem Verfügbaren. Dies führt zwangsläufig zu Einschränkungen in der Aussagekraft der Analysen, die in guten Studien auch benannt werden. Besser als keine Evaluation, könnte man meinen. Aber ist es tatsächlich besser? Nein.

Die Konsequenz daraus muss lauten: Die Evaluation muss bei der Ausgestaltung von Förderprogrammen von Anfang an mitgedacht werden. Dies umfasst sowohl die wissenschaftliche Methode als auch die dafür notwendigen Daten. Idealerweise würden Programme schon so gestaltet, dass eine Evaluation mit wissenschaftlichen Methoden möglich ist. Das müssen nicht gleich randomisierte Vergabe sein, etwa Losverfahren, um Förderungen an Förderberechtigte zuzuteilen. Das ist bei vielen Maßnahmen auch nur schwer vermittelbar, denn es besteht in der Regel ein Interesse an der Zielgerichtetheit der Maßnahmen. Diese Ziele müssen jedoch klar formuliert und messbar sein.

Als absolutes Minimum müssen Informationen dazu vorliegen, wie geförderte Projekte oder Akteure ausgewählt bzw. bewertet werden. Die Evaluatoren müssen wissen, wer sich beworben hat oder grundsätzlich förderfähig gewesen wäre, aber sich nicht beworben hat und wer letztlich gefördert und nicht gefördert wurde. Wenn keine ausreichenden Informationen zu den Förderempfängern vorliegen, dann muss zumindest die Verknüpfung mit anderen Datenquellen möglich sein. Es fehlen insbesondere häufig Daten auch nach Ende der Förderung. Für die nicht geförderte Kontrollgruppe gibt es in den Förderdaten typischerweise überhaupt keine (oder nur sehr begrenzte) Informationen, sodass eine Verknüpfung mit anderen Datensätzen häufig unumgänglich für eine Evaluation ist.

Es geht aber nicht nur um das Vorliegen geeigneter Daten, sondern auch um die Verfügbarkeit bzw. Nutzbarkeit. Die übermäßig strikte Interpretation der Datenschutz-Grundverordnung (DSGVO) in Deutschland führt dazu, dass vorhandene Datenbestände, insbesondere die amtlichen Statistiken, nicht in ausreichendem Maße verwendet werden können. Für eine Evaluation sind häufig Verknüpfungen von Datenquellen erforderlich, was durch die strenge Interpretation verhindert wird. Eine bessere Abstimmung zwischen den involvierten Einheiten könnte zu einer deutlichen Verbesserung der Datenlage und damit der Qualität der Evaluationen führen.

Eine der größten Herausforderungen bei der Evaluation von Einzelmaßnahmen besteht darin, ihre Wirkungen von denen anderer Maßnahmen zu trennen. Die Identifizierung von Wirkungen einer bestimmten Maßnahme erfordert notwendigerweise, dass Informationen zur Teilnahme an anderen relevanten Fördermaßnahmen vorliegen. In vielen Politikbereichen zeichnen sich Fördermaßnahmen oft über Portfolioansätze über verschiedene Akteursgruppen hinweg aus. Unternehmen profitieren oft zeitgleich oder zeitlich versetzt von unterschiedlichen Förderungen. Ohne eine umfassende und transparente Förderdatenbank ist die Identifizierung der Wirksamkeit einzelner Maßnahmen immer ungenau. Dazu braucht es dringend eine geeignete Dateninfrastruktur, die es erlaubt, Wirkungen zu isolieren oder explizit in Kombination zu evaluieren. Das Projekt „Erstellung und Analyse einer Datenbasis für die Evaluierung von Outputs und Outcomes der FuE- und Innovationsförderung von Unternehmen (EVALDAT)“, an dem das ZEW beteiligt ist, geht mit Erfolg in diese Richtung. Weitere Initiativen für andere Förderkontexte erscheinen daher sehr sinnvoll.

Ein weiteres Problem ist die Wirkgeschwindigkeit. Bei manchen Maßnahmen kann eine Evaluation direkt nach Beendigung oder vielleicht sogar noch während der Laufzeit sinnvoll sein. Bei anderen Maßnahmen, insbesondere wenn es um die Förderung von längerfristigen Aktivitäten wie Forschung und Entwicklung oder Technologieadoption geht, ist es oft direkt nach Ende einer Maßnahme noch zu früh, um Ergebnisse sehen zu können. Wann ist also der richtige Zeitpunkt zur Messung der Wirkung?

In diesem Zusammenhang sei jedoch auch gesagt, dass die Wirkungsanalyse einzelner Maßnahmen nur ein Teil einer guten Evaluation ist. Daneben gilt es auch, das Zusammenspiel mit anderen Fördermaßnahmen und die Implementierungseffizienz, d. h. die Frage „welche Programme mit welchen Zielen gibt es und wie passen sie zusammen?“ explizit zu analysieren. Hier stoßen auch einzelne experimentelle Designs an die Grenzen der Interpretier- und Generalisierbarkeit. Es scheint daher in der

Tat wichtig und richtig, sich nicht nur auf sehr gut identifizierbare Wirkungen zu beschränken. Kontextfaktoren und Rahmenbedingungen, die sich förderlich oder hinderlich auf die Zielerreichung von staatlichen Interventionen und die Wirkungsentfaltung auswirken können, sind ebenfalls wichtige Bausteine einer guten Evaluation, auch wenn hier kausale Aussagen grundsätzlich schwieriger sind. Der Einsatz qualitativer Methoden kann hier die quantitativen Ansätze sinnvoll ergänzen.

Es ist allerdings auch wichtig zu betonen, dass selbst bei idealer Datenverfügbarkeit alle Evaluationsmethoden auf Annahmen beruhen, die sich nicht immer überprüfen lassen. Ziel kann es also nicht sein, nur noch solche Programme zu implementieren, bei denen eine kausale Wirkungsanalyse nach höchsten wissenschaftlichen Standards ohne Einschränkungen möglich ist. Das würde in vielen Fällen zur politischen Handlungsunfähigkeit führen. Aber die Programme, die eingeführt werden, sollten zumindest grundsätzlich evaluierbar sein.

Gute Politikberatung braucht Zusammenarbeit zwischen Wissenschaft und Politik

Für die Evaluation spielt die Einbindung der Entscheidungsträger und ausführenden Einheiten ebenfalls eine wichtige Rolle. So schwierig es auch erscheinen mag, die für Konzeption und Umsetzung verantwortlichen Akteure in Politik und Verwaltung in den Evaluationsprozess einzubinden – so unverzichtbar ist es dennoch. Externen Evaluatoren fehlt häufig das nötige Prozesswissen, um beispielsweise sinnvolle Kosten-Nutzen-Analysen durchführen zu können.

Die Defizitdiagnose ist eine Handlungsaufforderung an die Wissenschaft und die Politik zugleich. Wie können wir unsere Ansätze weiterentwickeln, sodass sie einer Realität von komplexen, gleichzeitigen, sich überlappenden oder sich gegenseitig beeinträchtigenden Maßnahmen gerecht werden? Wie können Verknüpfungen zu Datenquellen rechtssicher hergestellt werden? Wie können Behörden verpflichtet werden, Daten so zur Verfügung zu stellen, dass sie für Evaluationen genutzt werden können? Fortschritte wären hier bereits eine saubere Dokumentation von abgelehnten und bewilligten Anträgen, Protokolle von Beschlüssen, etc. und die Speicherung von Informationen auf eine Art und Weise, die eine lückenlose Nachvollziehbarkeit erlaubt. Ein Beispiel für eine Planung, die eine bessere Evaluation erlaubt: Wenn Förderprogramme aus administrativen oder finanziellen Gründen nicht auf einen Schlag, sondern in zeitlich oder regional gestaffelten Wellen eingeführt werden, entstehen dadurch Vergleichsgruppen: Wer zuerst gefördert wurde, kann mit noch nicht geförderten, aber später begünstigten Betrieben vergli-

chen werden. Solche „Staggered Rollout“-Designs erlauben kausale Schätzungen, obwohl keine klassische Kontrollgruppe absichtlich zurückgehalten wurde.

Datenverfügbarkeit ist aber auch für jede Methodenweiterentwicklung ein zentraler Faktor: Die amtlichen Statistiken liefern Zahlen mit einer Verzögerung von Monaten oder Jahren, während sich die reale Wirtschaft gerade in Krisen schnell verändert und Programme eingeführt, ersetzt oder beendet werden. Die klassische Antwort der Wirtschaftsforschung – abzuwarten, bis belastbare Daten vorliegen, um Förderwirkungen sauber identifizieren zu können – ist in solchen Situationen keine Option. Denn bis dahin hat sich die Welt meist schon derart geändert, dass der Erkenntnisgewinn der Realität hinterherhinkt. Hinzu kommt ein methodisches Problem, das in ruhigen Zeiten kaum ins Gewicht fällt, in der Krise aber gravierend wird: die nachträgliche Revision amtlicher Daten. Gerade dieser Punkt zeigt aber, dass eine umfangreiche und auf verschiedenen Quellen beruhende kontinuierliche Infrastruktur zur Datenbereitstellung zur Evaluation staatlicher Fördermaßnahmen auch bei der methodischen Weiterentwicklung der Generierung und Auswertung von Daten helfen kann. Gleichzeitig kann wirtschaftspolitische Beratung auch früher ansetzen. Statt nur ex-post Evaluationen auszuschreiben, sollte auch das Design von Maßnahmen ex-ante verstärkt wissenschaftlich begleitet werden.

Die Frage, die sich aus diesen Ausführungen ergibt, berührt das Selbstverständnis der Wirtschaftsforschungsinstitute als Anbieter evidenzbasierter wirtschaftspolitischer Beratung: Wie lässt sich der Evidenzanspruch erfüllen, wenn die Realität eine andere ist. Ist evidenzbasierte Beratung in der Krise überhaupt noch möglich? Die Antwort ist, dass Pragmatismus gefordert ist, der allerdings die Grundstandards nicht über den Haufen wirft. Dabei ist neben einer engeren Zusammenarbeit mit den ausführenden Organisationen (z. B. mit den entsprechenden Einheiten in den zuständigen Ministerien) auch die Kommunikation gegenüber Politik und Öffentlichkeit wichtig. Evidenzbasierte Politikberatung bedeutet nicht, Sicherheit zu simulieren, sondern Unsicherheit explizit zu machen. Konfidenzintervalle und die Sensitivität der Ergebnisse gegenüber einzelnen Annahmen oder Methoden sollten fester Bestandteil jeder Maßnahmenevaluation sein.

Eine sinnvolle Auswahl an Wirkungsindikatoren ist ebenfalls zentral. Wenn die eigentlichen Zielgrößen – etwa langfristige Innovationsleistung oder Produktivitätswachstum – erst mit erheblicher Verzögerung messbar sind, kann die Evaluation auf Zwischenindikatoren ausweichen: Wurde die Liquidität der Unternehmen gesichert? Wurden Entlassungen vermieden oder neue Personen eingestellt? Wurden Forschungsabteilungen

personell erweitert, verlagert oder geschlossen? Diese Indikatoren sind prinzipiell früher verfügbar und erlauben zumindest vorläufige Aussagen zur Plausibilität eines Programms, bevor belastbare Langfristevaluationen vorliegen. Allerdings setzt dies auch wieder die Kombination von Datenquellen voraus. In diesem Beispiel das Zusammenspiel von Förderinformationen, Unternehmens- und Beschäftigtenstatistiken.

Wenn belastbare kausale Schätzungen kurzfristig nicht möglich sind, kann Politikberatung sich auf die Umsetzungsqualität konzentrieren: Wie hoch ist die Antragsbearbeitungsdauer? Wie viele Anträge werden abgelehnt und warum? Wird die Zielgruppe überhaupt erreicht? Solche Prozessindikatoren sind kein Ersatz für Wirkungsnachweise, liefern der Politik aber schnell verwertbare Stellschrauben, um ein Programm noch während der Krise nachzusteuern.

Da valide Wirkungsanalysen oft erst Jahre nach einer Intervention möglich sind, sollte bereits bei der Programmgestaltung der spätere Evaluierungsansatz mitgedacht werden – etwa durch Dokumentation abgelehnter Anträge als potenzielle Vergleichsgruppe, durch Zufallskomponenten bei der Antragsbearbeitung dort, wo rechtlich möglich, durch die Erfassung der wirtschaftlichen Situation und Entwicklung von Antragstellern vor Maßnahmenbeginn oder durch verpflichtende Datenlieferungen der geförderten Unternehmen über die Förderlaufzeit hinaus.

Wirtschaftsforschungsinstitute können hier eine wichtige beratende Funktion übernehmen, indem sie evaluationsfreundliches Programmdesign von Anfang an einfordern, statt erst im Nachhinein zu beklagen, dass valide Aussagen nicht möglich sind.

Für wissenschaftliche Evaluationen ergibt sich daraus eine wichtige Konsequenz: Die Frage „Wirkt Maßnahme X?“ lässt sich in der Realität selten mit Ja oder Nein beantworten. Seriöse Evaluationen sollten stattdessen klar differenzieren zwischen dem, was kurzfristig über Reichweite und Prozessqualität gesagt werden kann, und dem, was erst mittel- oder langfristig über kausale Wirkung ausgesagt werden kann. Diese Unterscheidung sollte bei Ausschreibungen für Evaluationsstudien besser berücksichtigt werden. Dies gilt umso mehr in unruhigen Zeiten, in denen pragmatische Ansätze zur Bewertung von Maßnahmen zur Milderung der negativen Auswirkungen von Schocks erforderlich sind.

Literatur

- Bundesrechnungshof. (2023). *Bericht nach § 88 Absatz 2 BHO an den Haushaltsausschuss des Deutschen Bundestages. Erfolgskontrollen als Voraussetzung für eine wirkungsorientierte Haushaltsführung.*
- EFI – Expertenkommission Forschung und Innovation. (2024). *Gutachten zu Forschung, Innovation und technologischer Leistungsfähigkeit Deutschlands 2024.*

Title: *Evaluating government support measures in times of crisis: between ambition and reality*

Abstract: *German economic policy responds to crisis situations with a familiar mix of instruments comprising direct subsidies, tax incentives, and liquidity support. Despite a legal obligation to monitor cost-effectiveness and outcomes, the evaluation of these measures is often limited by methodological constraints. Causal analyses are often not feasible either because the prerequisites for traditional evaluation designs are not met in crisis situations or because there are significant data gaps. The availability, linkability and quality of support and company data are often inadequate, which makes it difficult to identify causal chains. Improvement requires the systematic integration of evaluation approaches right from the program design phase – including a clear definition of objectives, data-driven monitoring, and mandatory data provision. This ensures that evidence-based policy advice can be provided even in the face of uncertainty and time pressure.*