



STIFTERVERBAND

ÜBERBLICKSBERICHT

---

# NEUE DATENBASIS ZUR EVALUATION DER FUE-FÖRDERUNG VON UNTERNEHMEN IN DEUTSCHLAND

---

Ergebnisse und Perspektiven für die  
EVALDAT-Datenbasis (Teildatenbasis FuE)

Gefördert durch:



Bundesministerium  
für Forschung, Technologie  
und Raumfahrt

März 2026

---

# NEUE DATENBASIS ZUR EVALUATION DER FUE-FÖRDERUNG VON UNTERNEHMEN IN DEUTSCHLAND

---

Ergebnisse und Perspektiven für die EVALDAT-Datenbasis (Teildatenbasis FuE)

## 1. Zielsetzung der Datenverknüpfung im Rahmen des Projekts EVALDAT

In diesem Übersichtsartikel stellen wir die Arbeit des durch das BMFTR geförderten Projekts „Erstellung und Analyse einer Datenbasis für die Evaluierung von Outputs und Outcomes der FuE- und Innovationsförderung von Unternehmen – EVALDAT“ (Förderkennzeichen 16P4293) vor. Das Ziel des Projekts war die Konzipierung und Generierung eines Paneldatensatzes, der für Wirkungsanalysen zu Maßnahmen der Forschungs- und Innovationsförderung eingesetzt werden kann. Insbesondere standen dazu die Zusammenführung administrativer Förderdaten zu FuE- und Innovationsförderungen für deutsche Unternehmen sowie die Verknüpfung dieser Daten mit Informationen über FuE- und Innovationsaktivitäten sowie zusätzlichen relevanten Kennzahlen aus bestehenden Panelbefragungen im Vordergrund. Das Projekt wurde im Zeitraum 01.09.23-30.06.25 von der SV Wissenschaftsstatistik gGmbH (WISTAT) im Verbund mit dem ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH, Mannheim durchgeführt. Bei der WISTAT lag der Schwerpunkt auf der Verknüpfung der verschiedenen Förderdaten mit Daten der jährlich selbst durchgeführten FuE-Erhebung. Die in diesem Artikel getätigten Aussagen beziehen sich somit auf das FuE-Teilvorhaben der WISTAT beziehungsweise die dabei generierte FuE-Teildatenbasis.

Im Folgenden wird zunächst der Prozess der Datenzusammenführung erläutert und die resultierende Datenstruktur beschrieben. Anschließend werden die Ergebnisse der Datenzuordnung präsentiert sowie die damit verbundenen Herausforderungen vorgestellt. Beginnend mit einer strukturierten Darstellung der Variablenkategorien werden mögliche Ansätze für quantitative Wirkungsanalysen skizziert, die mit der neu geschaffenen Datenbasis potenziell umgesetzt werden können. Dabei stehen sowohl die Analyse von Einflussfaktoren auf die Empfängerauswahl als auch die Untersuchung der Wirkungen von Fördermaßnahmen im Fokus. Abschließend wird der Prozess der Zurverfügungstellung der Daten beschrieben und eine Perspektive für deren mögliche zukünftige Weiterentwicklung gegeben.

Die hier dargestellten Ergebnisse wurden auf der Tagung „Herausforderungen und innovative Ansätze der Wirkungsmessung“ des Arbeitskreises

- **Elsner, Svenja**  
Referentin Data Analysis und Datenmanagement
- **Hellmann, Marco (Dr.)**  
Wissenschaftlicher Referent

Forschungs-, Technologie- und Innovationspolitik der Deutschen Gesellschaft für Evaluation einem interessierten Fachpublikum präsentiert und gemeinsam reflektiert.

## 2. Datenbasis und Verknüpfungsergebnisse

Dieses Kapitel gibt einen Überblick über die zugrunde liegenden Datenquellen, den Prozess ihrer Verknüpfung sowie die daraus entstandene Datenbasis. Es werden ausgewählte Ergebnisse des Zuordnungsprozesses dargestellt und Perspektiven für die Weiterentwicklung der Datengrundlage skizziert.

### 2.1 Datenquellen, Verknüpfungsprozess und Aufbau der Panel-Datenbasis

Folgende (Förder-) Datenquellen wurden für das Vorhaben EVALDAT berücksichtigt:

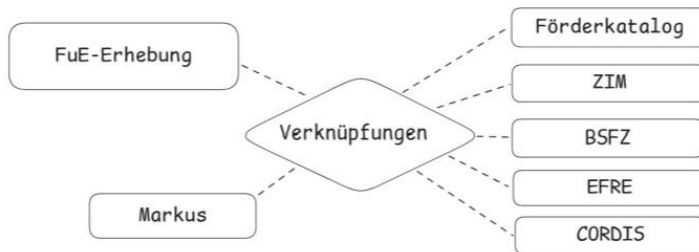
- Daten der regelmäßig durch die WISTAT durchgeführten Erhebung zu Forschung und Entwicklung der Unternehmen in Deutschland (folgend „FuE-Erhebung“) für die Erhebungsjahre 1995-2023
- Informationen der MARKUS-Datenbank für die Erleichterung und Vereinheitlichung der Zuordnung zwischen FuE- und Förderdaten
- Daten zur Projektförderung des Bundes (Förderkatalog) für die Jahre 2009 bis 2023
- Daten zu Förderungen im Zentralen Innovationsprogramm Mittelstand (ZIM) für die Jahre 2007 bis 2009, 2011, 2013, 2015 und 2017 bis 2023
- Daten der Bescheinigungsstelle Forschungszulage (BSFZ) für Anträge in den Jahren 2020 bis 2024
- Projektdaten der Bundesländer zu EFRE-kofinanzierten Förderungen für die EFRE-Förderperiode 2014-2020
- Förderdaten aus der CORDIS-Datenbank für die folgenden Forschungsrahmenprogramme: Forschungsrahmenprogramm 7, Horizon 2020, Horizon Europe

Die einbezogenen Datensätze unterscheiden sich hinsichtlich ihrer institutionellen Herkunft, Struktur, inhaltlichen Tiefe und Erfassungslogik. Vor diesem Hintergrund erwies sich die Erstellung eines einzelnen umfassenden Datensatzes im Projektverlauf als wenig zielführend. Stattdessen wurde eine flexible Datenbasis konzipiert, die eine einheitliche Grundstruktur bereitstellt und zugleich die spezifischen Inhalte der einzelnen Datenquellen erhält.

Zu diesem Zweck wurde eine eigene Datenbank eingerichtet, die die Verknüpfung der verschiedenen Datenquellen ermöglicht und so angelegt ist, dass zusätzliche Quellen integriert und bestehende Bestände erweitert werden können. Dadurch können je nach Analyse- oder Nutzeranfrage passgenaue Datenausschnitte generiert werden. Automatisierte

Sicherungen gewährleisten zudem die langfristige Verfügbarkeit des Datenbestands.

**Abbildung 1: Vereinfachte Darstellung der im EVALDAT-Projekt genutzten Datenbankstruktur**



Quelle: SV Wissenschaftsstatistik gGmbH (WISTAT). Eigene Darstellung.

Die Zuordnung von Einheiten zwischen den verschiedenen Datenquellen erfolgte im Wesentlichen über Namens- und Adressabgleiche sowie – soweit verfügbar – zusätzliche Identifikatoren. Dabei kamen sowohl automatisierte Verfahren als auch manuelle Prüfungen zum Einsatz; teilweise konnte zudem auf Vorarbeiten des Projektpartners oder eigene Vorarbeiten im Rahmen vergangener FuE-Erhebungen zurückgegriffen werden.

Für die Datenquellen Förderkatalog und ZIM lagen bereits Zuordnungsroutinen aus früheren Erhebungen vor. Die bestehenden Verknüpfungen wurden im Rahmen des EVALDAT-Projekts auf ihre Anwendbarkeit für die vorliegende Datenbasis überprüft. Hinweise auf mögliche Inkonsistenzen wurden in der Datenbank durch entsprechende Flags dokumentiert, da rückwirkende Anpassungen der ursprünglichen Zuordnungen aufgrund der komplexen Datenstrukturen nicht realisierbar waren. Damit wurden Unsicherheiten transparent dokumentiert und zugleich die Grundlage für künftig optimierte Zuordnungsroutinen geschaffen.

Für die Datenquellen Forschungszulage, EFRE und CORDIS wurden neue Zuordnungen erstellt, da keine bestehende Zuordnungshistorie vorlag. Die Qualität der Zuordnungen wurde – wo möglich - bewertet und dokumentiert.

Über alle Datenquellen hinweg stellte die unterschiedliche zeitliche Erfassung identifizierender Merkmale eine methodische Herausforderung dar, da sie eine präzise Zuordnung historischer Datenpunkte erschwert – insbesondere bei Einheiten mit häufigen Änderungen in den Identifikatoren. Unsicherheiten wurden – wo möglich – dokumentiert.

Die Identifizierung der methodischen Herausforderungen bildet die Grundlage für die Optimierung zukünftiger Zuordnungsprozesse und die Reduktion von Zuordnungsunsicherheiten. Gleichzeitig wird die besondere Bedeutung verlässlicher Identifikatoren deutlich, da alleinige Merkmale wie Name und Adresse über Zeit und Quellen hinweg möglicherweise variieren oder nicht konsistent erfasst werden.

## 2.2 Zuordnungsergebnisse und Qualität der Verknüpfungen

Die Ergebnisse des Zuordnungsprozesses lassen sich aus zwei Perspektiven einordnen. Zum einen wird betrachtet, in welchem Umfang Einheiten aus unterschiedlichen Förderdatenquellen im FuE-Adressbestand abgebildet sind. Auf dieser Grundlage kann beurteilt werden, inwieweit ein aus Förderinformationen abgeleiteter FuE-Bezug mit einer entsprechenden Erfassung im Rahmen der FuE-Erhebung übereinstimmt und wo gegebenenfalls Abweichungen bestehen.

Zum anderen wird analysiert, für wie viele Einheiten der FuE-Erhebung ergänzende Informationen aus den Förderdaten vorliegen. Dies ermöglicht eine Einschätzung darüber, in welchem Umfang die Paneldaten der FuE-Erhebung grundsätzlich für Evaluationen unter Einbeziehung von Kontrollgruppenansätzen geeignet sind.

Tabelle 1 berichtet die Ergebnisse aus der ersten Perspektive. Ausgehend von den in den jeweiligen Förderdatenquellen enthaltenen Projekten und potenziellen Förderempfängern wird der jeweilige Anteil mit Zuordnung zu Einheiten der FuE-Erhebung wiedergegeben. Die Zuordnungsquoten variieren – sowohl bezogen auf Projekte als auch auf potenzielle Förderempfänger – deutlich zwischen den einzelnen Förderdatenquellen.

**Tabelle 1: Ergebnisse der Zuordnung von Förderdaten zu Unternehmen der FuE-Erhebung**

für die Gesamtzahl der jeweils vorhandenen Jahre bis 2023

Förderdatenquelle	Anzahl Projekte	... davon mit Verknüpfung (%)*	Anzahl (potenzieller) Förderempfänger	... davon mit Verknüpfung (%)*
Förderkatalog	125.680	39	22.389	79
ZIM	38.430	89	18.899	88
Forschungszulage	27.921	65	12.632	54
EFRE	16.298	42	6.645	45
CORDIS	12.684	30	4.750	47

Grundlage der Auswertung: FuE-Erhebungen der WISTAT der Jahre 2007 bis 2023; MARKUS-Datenbank von Creditreform (Auszüge); Einmalige Datenlieferung der BSFZ zur Forschungszulage eingeschränkt auf positive förderfähige Aufwendungen, EFRE-Daten für Periode 2014-2020 eingeschränkt auf Interventionsbereiche zu FuE/Innovation und Technologietransfer und positive förderfähige Ausgaben, CORDIS Daten zu Forschungsrahmenprogramm 7, Horizon 2020, Horizon Europe gefiltert nach Projektstatus und vergleichbaren Variablen, Förderkatalog Daten für die Jahre 2009-2023 gefiltert nach Vorhaben-Art und Leistungsplansystematik, ZIM-Daten für die Jahre 2007 bis 2009, 2011, 2013, 2015 und 2017 bis 2023; bei Forschungszulage, EFRE, CORDIS: Einschränkung auf Projekte mit Laufzeitbeginn vor 2024

\* Die angegebenen Prozentanteile beziehen sich auf Projekte bzw. (potenzielle) Förderempfänger, die für mindestens ein Jahr einem Unternehmen des Adressbestandes der FuE-Erhebung zugeordnet werden konnten. Dabei sind hier auch Unternehmen berücksichtigt, für die noch keine Fragebogen-Informationen vorliegen.

Diese Unterschiede sind unter anderem darauf zurückzuführen, dass ein Teil der Förderdatenquellen (ZIM und Förderkatalog) bereits vor Beginn des Projektzeitraums systematisch und regelmäßig zur Anreicherung des Adressbestandes herangezogen wurde, während andere Quellen erstmals im Rahmen dieses Projekts integriert wurden.

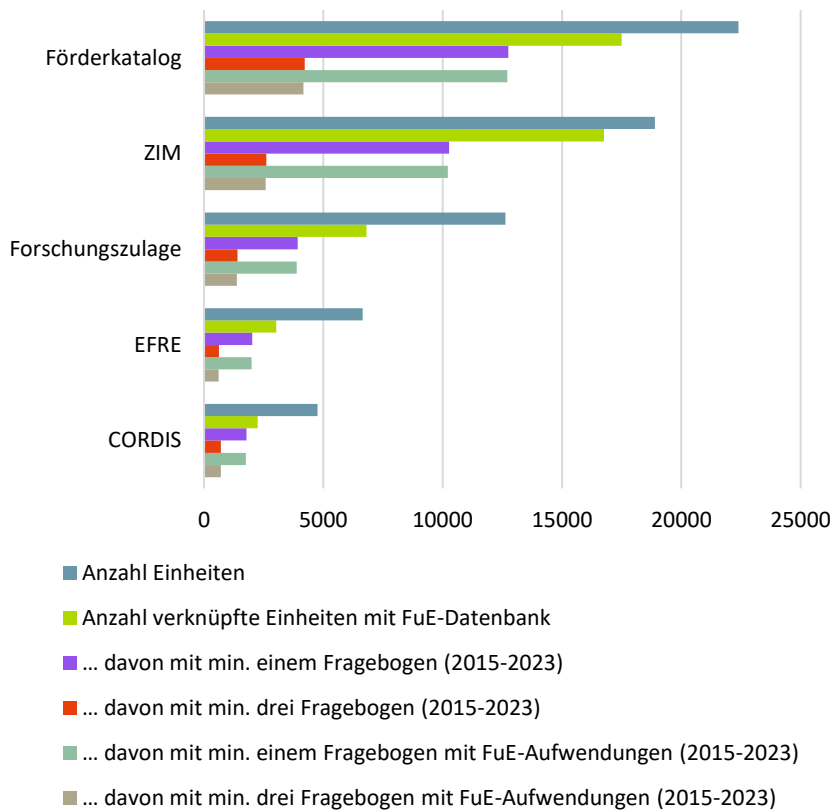
Darüber hinaus sind die Zuordnungsquoten vor dem Hintergrund struktureller Einschränkungen zu interpretieren. Eine vollständige Abdeckung ist unter anderem aufgrund fehlender oder nicht durchgängig verfügbarer Identifikatoren nicht optimal möglich. Unterschiede zwischen den Datenquellen ergeben sich zudem aus einem unterschiedlichen inhaltlichen Fokus als auch daraus, inwieweit relevante FuE-Projekte überhaupt mit den gegebenen Merkmalen identifiziert werden können. Im Fall der Daten der Forschungszulage ist beispielsweise zu beachten, dass das Vorhandensein eines Unternehmens in diesen Daten zunächst nur zeigt, dass ein Antrag für die Forschungszulage vorliegt, aber nichts darüber aussagt, ob das Projekt als FuE-Projekt bzw. das Unternehmen als FuE-aktiv einzustufen ist. Bei den EFRE-Daten ist die Eingrenzung grob auf die Förderbereiche FuE, Innovation und Technologietransfer erfolgt. Ein Indikator, der eine zuverlässige und differenzierte Zuordnung zu FuE-Projekten ermöglicht, liegt jedoch nicht vor.

Weiterhin können je nach Quelle auch potenzielle Förderempfänger enthalten sein, die nicht dem Berichtskreis der FuE-Erhebung zuzuordnen sind – dies ist insbesondere bei den Förderkatalog- und CORDIS-Daten der Fall. Insgesamt ist daher für alle Quellen davon auszugehen, dass die ausgewiesenen Quoten die tatsächliche Abdeckung relevanter Unternehmen eher unterschätzen. Eine präzisere Abgrenzung der sektoralen Zugehörigkeit der potenziellen Förderempfänger sowie des inhaltlichen Projektfokus durch die datenerfassenden Stellen würde die Aussagekraft künftiger Analysen weiter erhöhen.

Während sich die in Tabelle 1 ausgewiesenen Quoten auf das bloße Vorhandensein von Einheiten im FuE-Adressbestand beziehen – unabhängig davon, ob eine Rückmeldung des Unternehmens oder konkrete FuE-Angaben vorliegen –, veranschaulicht Abbildung 2, in welchem Umfang für die zugeordneten Förderdaten-Einheiten tatsächlich konkrete FuE-Informationen verfügbar sind.

Die Abbildung zeigt, dass für die Mehrheit, der dem FuE-Adressbestand zugeordneten Förderdaten-Einheiten im Zeitraum der letzten acht Jahre bis zum aktuellen Rand zumindest einmal konkrete FuE-Informationen vorliegen. Fehlende Angaben in einzelnen Jahren können unter anderem darauf zurückgehen, dass entsprechende Aktivitäten bereits vor dem betrachteten Zeitraum abgeschlossen wurden oder dass keine verwertbare Rückmeldung im Erhebungsprozess vorlag.

**Abbildung 2: Aufkommen der potenziellen Förderempfänger in den Daten der FuE-Erhebung**



Grundlage der Auswertung: FuE-Erhebungen der WISTAT der Jahre 2007 bis 2023; MARKUS-Datenbank von Creditreform (Auszüge); Einmalige Datenlieferung der BSFZ zur Forschungszulage eingeschränkt auf positive förderfähige Aufwendungen, EFRE-Daten für Periode 2014-2020 eingeschränkt auf Interventionsbereiche zu FuE/Innovation und Technologietransfer und positive förderfähige Ausgaben, CORDIS Daten zu Forschungsrahmenprogramm 7, Horizon 2020, Horizon Europe gefiltert nach Projektstatus und vergleichbaren Variablen, Förderkatalog Daten für die Jahre 2009-2023 gefiltert nach Vorhaben-Art und Leistungsplansystematik, ZIM-Daten für die Jahre 2007 bis 2009, 2011, 2013, 2015 und 2017 bis 2023; bei Forschungszulage, EFRE, CORDIS: Einschränkung auf Projekte mit Laufzeitbeginn vor 2024.

Hinweis: abweichende Definition bei der Anzahl verknüpfter Einheiten im Vergleich zur Auswertung für Tab 1 mit im Ergebnis möglichen geringfügigen Abweichungen.

Für etwa ein Fünftel der Förderdaten-Einheiten stehen im Zeitraum 2015 bis 2023 sogar Angaben für mindestens drei Jahre zur Verfügung. Durch die Anreicherung mit Förderdaten zu bislang nicht in den FuE-Daten erfassten Einheiten wird der Adressbestand der FuE-Erhebung schrittweise erweitert und systematisch geprüft, welche dieser Einheiten als FuE-relevant einzustufen sind. Für künftige Berichtsjahre ist daher von einer wachsenden Paneldatenbasis auszugehen – vorbehaltlich ausreichender Ressourcen für eine qualitätsgesicherte Verknüpfung und konsistente Zeitreihenpflege. Das Projekt hat hierfür die strukturellen Grundlagen geschaffen und dazu beigetragen, die Effizienz, Zielgenauigkeit und Verlässlichkeit der Zuordnungen sowie die analytische Nutzbarkeit der Daten zu verbessern.

Tabelle 2 stellt die Zuordnungsergebnisse aus der zweiten der oben genannten Perspektiven dar. Ausgehend von den in einem Berichtsjahr in der FuE-Erhebung erfassten Einheiten wird ausgewiesen, für wie viele Beobachtungen sowohl Informationen zu FuE-Aktivitäten als auch

mindestens ein laufendes Förderprojekt beziehungsweise eine entsprechende Zahlung im Berichtsjahr vorliegen.

**Tabelle 2: Zuordnungsergebnisse ausgehend von den Fragebogendaten der FuE-Erhebung für die Jahre 2019 bis 2023**

Variable/Jahr	2019	2020	2021	2022	2023
Anzahl Beobachtungen	18.832	1.911	20.340	2.070	25.036
Beobachtungen mit Crefo-Nr (%)	94	92	95	91	94
Davon mit FuE = Ja (%)	94	96	95	96	88
→ davon mit min. einer FuE-Förderung bzw. einem laufendem Förderprojekt in t (%)*	56	65	59	64	62

Grundlage der Auswertung: FuE-Erhebungen der WISTAT für die Berichtsjahre 2019 bis 2023; MARKUS-Datenbank von Creditreform (Auszüge); Einmalige Datenlieferung der BSFZ zur Forschungszulage (beantragte Projekte, die aktuell laufend sind im jeweils genannten Jahr) eingeschränkt auf positive förderfähige Aufwendungen, EFRE-Daten für Periode 2014-2020 eingeschränkt auf Interventionsbereiche zu FuE/Innovation und Technologietransfer und positive förderfähige Ausgaben (laufende Projekte im jeweiligen Jahr), CORDIS Daten zu Forschungsrahmenprogramm 7, Horizon 2020, Horizon Europe (laufende Projekte im jeweiligen Jahr) gefiltert nach Projektstatus und vergleichbaren Variablen, Förderkatalog Daten für die Jahre 2019 bis 2023 gefiltert nach Vorhaben-Art und Leistungsplansystematik, ZIM-Daten für die Jahre 2019 bis 2023; Hinweise: Da die FuE-Erhebung für gerade Berichtsjahre in Form einer Kurzerhebung durchgeführt wird, ist die Anzahl an Beobachtungen für diese Jahre wesentlich kleiner. Tatsächliche Mittelflüsse sind nur für ZIM und Förderkatalog-Daten bekannt. Für die übrigen Quellen wird von linear verteiltem Mittelfluss innerhalb der Projektlaufzeit ausgegangen.

Für die Mehrheit der erfassten Einheiten sind Förderinformationen verfügbar. Der Anteil der FuE-Unternehmen mit Angaben zu gleichzeitigen Förderungen liegt in den betrachteten Jahren zwischen etwa der Hälfte und zwei Dritteln. In geraden Berichtsjahren fällt dieser Anteil höher aus, da in diesen Jahren eine Kurzerhebung auf Basis einer kleineren Stichprobe durchgeführt wird.

Von allen Einheiten mit FuE-Informationen im Zeitraum 2019 bis 2023 liegen für rund 15 Prozent in der gesamten Datenbasis (also auch vor 2019) keine Förderinformationen vor, was auf eine breite Streuung der Fördermittel hindeutet und – vorbehaltlich vertiefter fallbezogener Prüfungen – grundsätzlich die Möglichkeit eröffnet, Kontrollgruppen für weiterführende Analysen zu bilden.

Die Eignung der Zeitreihendaten für ein konkretes Analyseziel ist jeweils gesondert zu prüfen. Dabei sind mögliche Unterschiede in der Validität der Zeitreihen zu berücksichtigen, die aus Einschränkungen bei der nachträglichen Zuordnung von Förderinformationen zu historisch gewachsenen Datenstrukturen der FuE-Erhebung resultieren können.

Während die FuE-Erhebung bislang vor allem auf Aggregatdaten einzelner Berichtsjahre ausgerichtet war, erfordert die konsistente Analyse auf Mikroebene eine über die Zeit stabile und eindeutig nachvollziehbare

Einheitenzuordnung, die für vergangene Zeiträume nicht durchgängig gewährleistet ist. Für zukünftige Berichtsjahre ist jedoch zu erwarten, dass entsprechende Fälle aufgrund verbesserter Prozesse seltener auftreten. Zur besseren Einordnung im Rahmen von Datenanalysen mit der im Projekt erstellten Paneldatenbasis wurden im Projekt methodische Hinweise in Form von Daten-Flags erarbeitet.

Aktuell wird über die gesamte Datenbasis hinweg davon ausgegangen, dass rund 75 Prozent der Einheiten nicht von den identifizierten Einschränkungen betroffen sind. Ergänzend ist zu berücksichtigen, dass sich die FuE-Zeitreihen nicht ausschließlich aus originär erhobenen Werten zusammensetzen, sondern auch imputierte und geschätzte Werte enthalten. In der Erhebung 2023 stammen beispielsweise 36 Prozent der Werte aus der aktuellen Erhebung, 39 Prozent wurden aus vorherigen Erhebungen imputiert, und weitere 25 Prozent beruhen auf Schätzungen.

Zukünftig sollen zentrale Prozesse der FuE-Erhebung weiter optimiert und teilweise automatisiert werden, sodass perspektivisch mit einer weiteren Verbesserung der Qualität der bei EVALDAT zusammengestellten Paneldatenbasis zu rechnen ist, sollte eine Weiterentwicklung dieser Datenbasis realisierbar sein.

### 3. Indikatorenübersicht und Potentiale für Wirkungsanalysen

Nachfolgend werden Analysemöglichkeiten dargestellt, die durch die oben beschriebenen Maßnahmen der Verknüpfung potenziell eröffnet werden. Kontrollgruppenansätze, die geförderte Unternehmen mit nicht geförderten vergleichen, gelten hier als wissenschaftlicher Standard (vgl. Cerulli 2010, Zúñiga-Vicente et al. 2014). Voraussetzung für solche Wirkungsanalysen mit Kontrollgruppenansätzen sind Mikrodaten, die beide Gruppen abbilden und ein breites Set an Input- und Outputvariablen über einen ausreichend langen Zeitraum enthalten.

Zunächst wird ein inhaltlicher Überblick über die Variablen der Förderquellen sowie der FuE-Erhebung gegeben. Darauf aufbauend werden die Analysemöglichkeiten anhand beispielhafter Ansätze veranschaulicht. Diese Veranschaulichung dient in erster Linie einer grundlegenden Orientierung zum Umgang mit den verknüpften Datensätzen sowie einer zusätzlichen Erläuterung etwaiger Besonderheiten. Zwei zentrale Fragestellungen dienen hier als Ausgangspunkt: Welche Faktoren beeinflussen die Förderauswahl beziehungsweise den Fördererhalt? Welche Wirkungen des Fördererhalts lassen sich beobachten?

Es ist stets zu beachten, dass es sich vorrangig um beispielhafte Darstellungen handelt, also keine tatsächlich durchgeführten Wirkungsanalysen. Im Vordergrund steht also ein Überblick über Indikatoren und Vorgehensweisen für interessierte Forschende. Bei realen Analysen sind die zuvor beschriebenen Einschränkungen zu beachten.

### 3.1 Übersicht zu Förder- und FuE-Daten

Die berücksichtigten Förderdatenquellen enthalten strukturierte Informationen zu geförderten Projekten und deren Zuwendungsempfängern. Grundsätzlich lassen sich mehrere übergreifende Variablenkategorien identifizieren, die in den meisten Datensätzen – wenn auch teilweise mit unterschiedlichem Detaillierungsgrad – vorhanden sind. Dazu zählen insbesondere Informationen zu den Antragstellern beziehungsweise Zuwendungsempfängern, zu den geförderten Projekten, zu finanziellen Förderstrukturen, zu regionalen Zuordnungen sowie zu zeitlichen Merkmalen der Förderung. Teilweise sind darüber hinaus auch Angaben zu Projektinhalten oder Projektergebnissen verfügbar.

Die genaue Ausprägung und Verfügbarkeit dieser Variablen unterscheidet sich jedoch zwischen den einzelnen Datenquellen. Eine Übersicht zu den identifizierten Kategorien sowie beispielhaften Variablen bietet die folgende Tabelle.

**Tabelle 3: Übersicht über die Variablen aus unterschiedlichen Förderdatenquellen**

Förderdatenquelle	Förderkatalog	ZIM	EFRE	CORDIS	Forschungszulage
<b>Empfängerdetails</b>	z. B. Trennung nach Zuwendungsempfänger und ausführender Stelle, Einordnung als KMU, Ort, PLZ	z.B. Anzahl Beschäftigte, Beschäftigte in FuE, Umsatz, WZ, Ort, PLZ	z.B. Sektor, Ort, PLZ (Ort hier teilweise nur projektbezogen)	z.B. Einordnung als KMU, Organisationstyp, Rolle im Projekt, Ort, PLZ	z.B. Anzahl Beschäftigte, Beschäftigte in FuE, FuE-Aufwendungen z.B. Umsatz, WZ, Rechtsform, Ort, PLZ
<b>Projektetails</b>	z.B. Titel bzw. Thema	z.B. Programmart	z.B. Interventionskategorien, Vorhabenbeschreibung	z.B. Titel, Beschreibung, Ziele, Keywords	-
<b>Fördermittel</b>	z.B. geflossene Mittel nach Jahren, Projekt-Gesamtkosten, Anteile Bundesmittel	z.B. geflossene Mittel nach Jahren, Gesamtkosten des Projekts	Förderfähige Ausgaben	z.B. Gesamtkosten des Projekts, Beitrag einzelner Projektteilnehmer zu Gesamtkosten (inklusive EU Förderung), EU Förderung Gesamt und nach Projektteilnehmern	z.B. FuE-Auftrags-/Personalkosten, Eigenleistung in Stunden z.B. Prüfergebnis Antrag, Antrag zurückgezogen?, Erhalt anderer Fördermittel
<b>Klassifizierung innerhalb des Katalogs</b>	z.B. Art des Vorhabens, Art der FuE, Rahmenprogramm, Projektträger, Ministerium	z.B. Förderprogramm	z.B. Veröffentlichendes Bundesland	z.B. EU Framework Programmes, Thematischer Schwerpunkt, Type of Action z.B. "Call" und "Subcall"	-
<b>zusätzliche zeitliche Angaben</b>	-	z.B. Antrags- und Bewilligungsdatum	-	z.B. Ende der Projekt-partizipation einzelner Projekt-partner	z.B. Bescheidversand
<b>Projektoutcome</b>	-	-	-	z.B. "Intellectual Property Rights", Publikationen	-

Quelle: SV Wissenschaftsstatistik gGmbH (WISTAT). Eigene Darstellung.

Insgesamt zeigt sich, dass trotz teils divergierender Datenstruktur eine Reihe von inhaltlichen und strukturellen Überschneidungen besteht, die eine vergleichende Auswertung und Zusammenführung der Datensätze grundsätzlich ermöglichen. Unterschiedliche Detaillierungstiefen und

Erhebungsperspektiven machen jedoch eine sorgfältige Variablenharmonisierung sowie quellenkritische Interpretation unabdingbar.

Die Verknüpfung der Förderdaten mit den FuE-Daten der WISTAT liefert die nachfolgenden Ergänzungen. Die FuE-Erhebung wird jährlich im Auftrag des Bundes durchgeführt und basiert aktuell auf der EU-Verordnung 2019/2152 als rechtlicher Grundlage. In ungeraden Berichtsjahren erfolgen Vollerhebungen, während in geraden Jahren Kurzerhebungen durchgeführt werden. Letztere umfassen eine Stichprobe der zu betrachtenden Unternehmen sowie einen gekürzten Fragebogen. Im Rahmen der Vollerhebungen werden sämtliche Unternehmen mit mindestens einem Beschäftigten befragt, die als FuE-aktiv identifiziert wurden oder bei denen ein entsprechender Verdacht besteht. Der derzeit aktive Adressbestand umfasst rund 34.000 Unternehmen. In der letzten Vollerhebung für das Berichtsjahr 2023 wurden insgesamt 25.036 Fälle erfasst, einschließlich der Unternehmen, die FuE-Inaktivität gemeldet haben. Zeitreiheninformationen liegen derzeit für die Berichtsjahre 1979, 1981, 1983, 1985 sowie durchgehend bis 2023 vor; die Erhebung für das Berichtsjahr 2024 ist abgeschlossen und die Vollerhebung für das Berichtsjahr 2025 steht bevor.

**Tabelle 4: Übersicht über die FuE-Daten der WISTAT**

	<b>Vollerhebung (zuletzt 2023)</b>	<b>Teilerhebung (zuletzt 2024)</b>
<b>FuE-Aktivitäten</b>	z.B. FuE-Aufwendungen insgesamt, interne und externe Aufwendungen, Mittelquellen	z.B. FuE-Aufwendungen insgesamt, interne und externe Aufwendungen, (weniger detailliert)
<b>FuE-Personal</b>	z.B. wissenschaftliches, technisches und sonstiges FuE-Personal (Head Count/Vollzeitäquivalente)	z.B. FuE-Personal insgesamt, wissenschaftliches FuE-Personal (Vollzeitäquivalente)
<b>Unternehmensdetails</b>	z.B. Konzernstruktur, Forschungsstandorte, Umsatz, Anzahl Beschäftigter, Wirtschaftszweig	siehe Vollerhebung (weniger detailliert)
<b>Regionale Zuordnungen</b>	z.B. Ort, PLZ, AGS	siehe Vollerhebung
<b>Klassifikation der FuE-Aktivitäten</b>	z.B. Art der FuE (Grundlagenforschung, ...), Zweck der FuE (Entwicklung neuer Verfahren, ...)	-
<b>FuE-Outcome</b>	z.B. Anteil des durch FuE generierten Outputs am Umsatz	-
<b>Jährliche wechselnde Vertiefungsfragen</b>	z.B. Forschungsfelder, Produktgruppen, Fachliche Qualifikationen des FuE-Personals, Gründe für externe FuE, Nutzung der Forschungszulage, Gründe für die Nichtnutzung der Forschungszulage	z.B. Nutzung der Forschungszulage, Gründe für die Nichtnutzung der Forschungszulage

Quelle: SV Wissenschaftsstatistik gGmbH (WISTAT)

Seit dem Berichtsjahr 2020 nimmt die Forschungszulage einen zusätzlichen Stellenwert in der FuE-Erhebung ein. Neben jährlich wechselnden Vertiefungsfragen werden stets die Nutzung und Gründe für die Nicht-

Nutzung der Forschungszulage abgefragt. Weitere Details können bei Bedarf dem aktuellen Methodenbericht zur FuE-Erhebung<sup>1</sup> entnommen werden.

### 3.2 Analyse der Förderauswahl und des Fördererhalts

Die systematische Untersuchung von Einflussfaktoren, die den Erhalt öffentlicher Innovationsförderungen auf Seiten von Unternehmen determinieren, ist sowohl aus ökonomischer Perspektive als auch im Kontext evidenzbasierter Förderpolitik von erheblicher Bedeutung.

#### 3.2.1 Allokationseffizienz

Wenn bestimmte Unternehmensmerkmale – etwa Größe, Standort, Branchensegment oder FuE-Kapazität – systematisch mit der Förderwahrscheinlichkeit korrelieren, stellt sich die Frage nach der Allokationseffizienz öffentlicher Mittel. Diese Vergabe ist in der Regel kompetitiv, wobei die bereits vorhandenen Innovationskapazitäten und technologischen Fähigkeiten sowie vergangene Innovationserfolge der Unternehmen eine nicht unerhebliche Rolle bei der Förderauswahl spielen können (siehe zum Beispiel Falk & Svensson, 2020). Die Analyse solcher zugrundeliegender Einflussfaktoren kann dabei helfen, potenzielle Fehlsteuerungen, Zugangshürden oder ungewollte Selektionsmechanismen im Förderprozess zu identifizieren. Die von uns erzeugte Paneldatenbasis kann als Grundlage dienen diese Faktoren gezielt für einzelne Förderquellen zu analysieren und Förderquellen miteinander zu vergleichen.

Indikatoren, die den Erhalt einer Förderung und die Zielvariablen von Förderungen beeinflussen können, liegen teilweise bereits im Datenbestand der FuE-Erhebung vor. Durch die Verknüpfung der Förderdaten mit den FuE-Daten können für alle zugeordneten Förderempfänger relevante Empfängerdetails auf Basis derselben Erfassungslogik bereitgestellt werden. Zu diesen zählen etwa die Unternehmensdetails (zum Beispiel Unternehmensgröße, Konzernstruktur, Umsatz, Wirtschaftszweig), die FuE-Aktivitäten (zum Beispiel FuE-Aufwendungen) oder das FuE-Personal (zum Beispiel Anzahl wissenschaftliches Personal). Sofern es aus Sicht des Forschenden als plausibel anzunehmen ist, dass Merkmale wie die Qualifikationsstruktur, Produktgruppen oder Forschungsfelder im Zeitverlauf recht stabil bleiben, können auch diese als relevante Einflussfaktoren analysiert werden, obwohl diese nur für ausgewählte Jahre vorliegen.

Darüber hinaus bietet die Verknüpfung mit den Förderdaten die Möglichkeit weitere Indikatoren zu bilden. Für die in der Datenbasis erfassten Unternehmen liegen im Zeitverlauf Informationen zu bereits erhaltenen Fördermitteln und beantragten beziehungsweise durchgeführten Förderprojekten vor. Durch deren Aggregation kann die „Fördererfahrung“ abgebildet werden. Neben der Vielfalt der genutzten Förderprogramme kann auch die Vielfalt der durchgeführten Förderprojekte und thematischen Schwerpunkte untersucht werden. Dies wird durch die verschiedenen Projektdetails teilweise ermöglicht. Die „Diversität der Forschung“ könnte

---

<sup>1</sup> [https://www.stifterverband.org/sites/default/files/2025-11/fue-erhebung\\_2023\\_methodenbericht.pdf](https://www.stifterverband.org/sites/default/files/2025-11/fue-erhebung_2023_methodenbericht.pdf)

sowohl im Sinne einer technischen Kompetenz als auch im Sinne der Fördererfahrung interpretiert werden.

Eine Zuordnung mehrerer Unternehmen zum selben Förderprojekt ist in den einzelnen Förderdaten teilweise bereits zugänglich, insbesondere in den CORDIS-Daten. Eine eindeutige Identifikation der Unternehmen über alle Datensätze hinweg ermöglicht für Daten dieser Quelle die Identifikation von Verbundpartnern und erlaubt damit potenziell die übergreifende Analyse von „Kooperationserfahrungen“. Berücksichtigt man dazu zusätzlich die Merkmale der einzelnen Verbundpartner kann dazu die Erfahrung mit homogenen und heterogenen Akteuren untersucht werden. Entstehen solche Verbundprojekte zudem häufiger, können daraus auch Annahmen über mögliche Netzwerke getroffen werden.

Ein mögliches Vorgehen für die oben beschriebene Analyse basiert wesentlich auf zwei Maßnahmen zur Datenaggregation. Erstens müssen für alle Unternehmen die tatsächlich geflossenen Mittel absolut bestimmt oder der Fördererhalt binär kodiert werden. Zweitens müssen die ausgewählten Indikatoren gebildet und gegebenenfalls gewichtet werden. Beispielsweise sind in der Regel nicht die absoluten Förderaufwendungen von Bedeutung, sondern die Förderaufwendungen in Relation zur Unternehmensgröße. Sollten neben der Unterscheidung nach Förderquellen auch darunterfallende Förderprogramme und Förderlinien verglichen werden, können dazu die aufgelisteten Klassifizierungsvariablen innerhalb der jeweiligen Förderquellen herangezogen werden.

### 3.2.2 Förderintentionen

Aus Sicht der Fördergeber besteht neben der Analyse von Verteilungen der Fördermittel auch Potenzial, die eigenen Förderintentionen detaillierter als zuvor zu untersuchen. Förderprogramme sollen nicht nur Innovationen generieren, sondern auch Innovationskapazitäten langfristig stärken und strukturelle Entwicklungspfade positiv beeinflussen. Dazu kann es von Interesse sein, ob Unternehmen mit bestimmten Forschungsaktivitäten oder Unternehmen mit hohem FuE-Output erreicht werden.

Indikatoren, die ausdrücken welche Art von FuE-Unternehmen adressiert beziehungsweise vorrangig erreicht werden, können sich an den oben genannten Indikatoren zur Verteilung von Fördermitteln durchaus orientieren. Weitere relevante Indikatoren sind zudem die Art und der Zweck der FuE. So kann beispielsweise untersucht werden, ob Unternehmen mit Schwerpunkt „Grundlagenforschung“ stärker erreicht werden als Unternehmen mit dem Schwerpunkt „angewandte Forschung“. Dasselbe gilt für die Gegenüberstellung verschiedener Zwecke hinsichtlich verschiedener Produktgruppen wie Software, Maschinen oder chemische Erzeugnisse. Besonders interessant ist dabei die Hinzunahme einer Output-Bewertung. Die FuE-Daten liefern hier den „Anteil des durch FuE generierten Outputs am Umsatz“ als zentralen Indikator. Damit eröffnen sich Möglichkeiten, um zu untersuchen, ob Unternehmen mit einem hohen FuE-Output stärker erreicht werden und Fördermittel so effektiver eingesetzt werden. Wichtig ist hier anzumerken, dass es nicht darum geht, die Art der Projekte zu bewerten, sondern Fördermittelempfänger möglichst

stark nach ihren FuE-Aktivitäten und „forschungsstrategischen Ausrichtungen“ zu differenzieren.

Inwiefern auch die Analyse von Förderanträgen möglich ist, hängt von der Verfügbarkeit von nicht bewilligten Anträgen in den einzelnen Förderdaten ab. Interessant wäre hier auch die Analyse von Empfängermerkmalen hinsichtlich der Zeit zwischen Förderantrag und Förderbewilligung. So bieten etwa die ZIM-Daten neben einem Antragsdatum auch ein Bewilligungsdatum. Ähnlich verhält es sich mit den Daten zur Forschungszulage. Hier können zudem die Hinweise aus der FuE-Erhebung zur Bewertung der Forschungszulage zusätzlich Hinweise auf mögliche Ursachen für Unterschiede bei der Reichweite der Fördermaßnahme geben.

### 3.3 Analyse der Fördermittelwirkung

Während zuvor die Determinanten der Förderauswahl und des Fördererhalts im Vordergrund standen, richtet sich der Blick nun stärker auf die Wirkungen der gewährten Fördermittel. Im Zentrum stehen dabei insbesondere outputbezogene Indikatoren sowie deren Entwicklung im Zeitverlauf. Eine trennscharfe Abgrenzung zwischen Auswahl- und Wirkungsanalyse ist dabei nicht immer möglich, da viele Indikatoren sowohl Selektionsmechanismen als auch Wirkungseffekte widerspiegeln können.

Grundsätzlich lassen sich Wirkungsanalysen entlang dreier zentraler Konzepte strukturieren: Inputadditionalität, Outputadditionalität und Verhaltensadditionalität (Peters et al., 2012). Diese Konzepte zielen jeweils auf unterschiedliche Dimensionen möglicher Förderwirkungen ab und können mithilfe der verknüpften Förder- und FuE-Daten differenziert untersucht werden. Voraussetzung für belastbare Aussagen ist dabei stets die Berücksichtigung potenzieller Auswahleffekte sowie die Einbeziehung geeigneter Kontrollgruppen, bestehend aus Unternehmen ohne Fördererhalt.

#### 3.3.1 Inputadditionalität

Die Analyse der Inputadditionalität untersucht, ob öffentliche Fördermittel zu einer Ausweitung der unternehmerischen FuE-Aktivitäten führen oder lediglich private Mittel substituieren. Im Fokus stehen dabei Veränderungen der FuE-Kapazität vor und nach dem Fördererhalt. Relevante Indikatoren sind insbesondere interne und externe FuE-Aufwendungen, die jeweilige Finanzierungsstruktur der FuE sowie die Anzahl der FuE-Beschäftigten.

Ein mögliches Vorgehen besteht zunächst in der Analyse der Mittelflüsse. Hierbei kann entweder die absolute Höhe der tatsächlich geflossenen Fördermittel berücksichtigt oder der Fördererhalt binär kodiert werden. Aufbauend darauf können Veränderungen zentraler FuE-Inputgrößen im Zeitverlauf betrachtet werden.

Als Indikatoren für diese Art der Analyse sind insbesondere die verknüpften Förderdaten in Kombination mit den FuE-Erhebungsdaten geeignet. Die Förderdaten liefern Informationen zu Bewilligungszeitpunkten, Förderhöhen und Förderkategorien, während die FuE-Daten standardisierte

Angaben zu Aufwendungen, Personal und Finanzierungsquellen bereitstellen. Die Verknüpfung beider Datenquellen bildet eine Grundlage, um Förderintensitäten in Relation zur Unternehmensgröße oder zur bisherigen FuE-Kapazität zu setzen und deren Entwicklung über mehrere Jahre hinweg zu analysieren.

### 3.3.2 Outputadditionalität

Die Analyse der Outputadditionalität richtet den Fokus auf die Frage, ob Fördermittel zu messbaren Verbesserungen der Innovations- und Marktergebnisse führen. Als zentrale Zielvariablen bieten sich hierbei Umsatzentwicklungen sowie insbesondere der Anteil des durch FuE generierten Outputs am Gesamtumsatz an.

Ein mögliches Vorgehen besteht darin, die Entwicklung dieser Outputindikatoren vor und nach dem Fördererhalt zu analysieren und mit der Entwicklung nicht geförderter Unternehmen zu vergleichen. Neben absoluten Umsatzgrößen sind dabei relative Kennzahlen von besonderer Bedeutung, um strukturelle Unterschiede zwischen Unternehmen angemessen zu berücksichtigen.

Als relevante Indikatoren liefern die FuE-Daten Angaben zum innovationsbezogenen Umsatz, während die Förderdaten eine zeitliche Zuordnung der Förderereignisse ermöglichen. Durch die Kombination beider Datenquellen können Förderzeitpunkte potentiell mit nachgelagerten Outputentwicklungen in Beziehung gesetzt und über längere Beobachtungszeiträume hinweg analysiert werden. Auch hier ist die Anwendung ökonomischer Verfahren – insbesondere Panelregressionen oder Differenz-von-Differenzen-Modelle – denkbar, um beobachtete Veränderungen von allgemeinen Konjunktur- oder Brancheneffekten abzugrenzen.

### 3.3.3 Verhaltensadditionalität

Neben Input- und Outputeffekten kann Förderung auch zu Veränderungen im Innovationsverhalten von Unternehmen führen. Unter dem Begriff der Verhaltensadditionalität wird untersucht, ob Fördermittel strategische Entscheidungen beeinflussen, etwa die Aufnahme oder Intensivierung von FuE-Aktivitäten in bestimmten Technologiefeldern oder Branchen.

Geeignete Indikatoren sind hier beispielsweise die Anzahl FuE-aktiver Unternehmen in bestimmten Wirtschaftszweigen oder Technologiefeldern, Veränderungen in der thematischen Ausrichtung der Forschung oder die Aufnahme neuer Forschungsfelder. Auch die Entwicklung der FuE-Aufwendungen in bislang wenig bearbeiteten Bereichen kann Hinweise auf Verhaltensänderungen liefern.

Die verknüpften Datensätze können dazu beitragen, solche Entwicklungen sowohl auf Unternehmensebene als auch aggregiert nach Branchen, Förderlinien oder thematischen Schwerpunkten zu analysieren. Für alle Wirkungsanalysen gilt zudem, dass neben der Differenzierung nach Förderquellen auch eine Unterscheidung nach Förderprogrammen, thematischen Schwerpunkten einzelner Förderlinien sowie regionalen Faktoren

vorgenommen werden kann. Dies kann zu einer systematischeren Untersuchung von eventuellen Unterschieden in den Wirkungen verschiedener Förderinstrumente beitragen.

### 3.3 Methodische Hinweise und Erweiterungsperspektiven

Die Analyse der Förderauswahl, des Fördererhalts sowie der Fördermittelwirkung erfordert ein methodisch differenziertes Vorgehen. Grundsätzlich kommen sowohl deskriptive als auch inferenzstatistische Verfahren in Betracht, deren Einsatz sich nach der jeweiligen Fragestellung und Datenverfügbarkeit richtet.

Auf deskriptiver Ebene können zunächst Verteilungen zentraler Merkmale analysiert werden. Hierzu zählen insbesondere die Anzahl der Unternehmen in verschiedenen Kategorien sowie Kreuztabellen zu Unternehmensmerkmalen, FuE-Kapazitäten oder Förderkategorien. So lassen sich beispielsweise der Anteil kleiner und mittlerer Unternehmen oder der Anteil forschungsintensiver Unternehmen in einzelnen Förderprogrammen bestimmen. Auch Vorher-Nachher-Vergleiche zentraler Input- und Output-Indikatoren liefern erste Hinweise auf mögliche Förderwirkungen.

Für vertiefende Analysen bieten sich Regressionsmodelle an. Im Kontext der Förderauswahl können (log-)lineare Modelle eingesetzt werden, um die Beiträge einzelner Indikatoren zum Fördererhalt zu untersuchen, indem geförderte und nicht geförderte Unternehmen systematisch miteinander verglichen werden. Auf diese Weise lassen sich Zusammenhänge zwischen Unternehmensmerkmalen, FuE-Strukturen und der Wahrscheinlichkeit eines Fördererhalts statistisch modellieren.

Im Rahmen der Wirkungsanalyse kommt darüber hinaus dem Einsatz quasi-experimenteller Verfahren besondere Bedeutung zu. Insbesondere Differenz-von-Differenzen-Ansätze können ermöglichen, geförderte Unternehmen mit strukturell vergleichbaren nicht geförderten Unternehmen zu kontrastieren, um kausale Effekte des Fördererhalts näherungsweise zu identifizieren. Die Auswahl geeigneter Kontrollgruppen kann sowohl auf Basis allgemeiner Unternehmensmerkmale als auch anhand FuE-spezifischer Indikatoren erfolgen. Voraussetzung für belastbare Ergebnisse ist dabei die sorgfältige Berücksichtigung potenzieller Auswahlwirkungen.

Unabhängig vom gewählten methodischen Ansatz gilt, dass für tragfähige Wirkungsanalysen sowohl fundierte Kenntnisse über Selektionsmechanismen als auch über die Zielgrößen der Förderung erforderlich sind. Die im Bericht erwähnten sowie im Methodenbericht<sup>2</sup> detaillierter beschriebenen Einschränkungen der bisherigen Datenverknüpfungen – insbesondere mögliche Fehlverknüpfungen oder unvollständige Zeitreihen – sind bei der Interpretation empirischer Ergebnisse zwingend zu berücksichtigen,

---

<sup>2</sup> Der Methodenbericht zum EVALDAT-Projekt kann auf Anfrage interessierten Forschenden zur Verfügung gestellt werden.

da sie die Durchführbarkeit und Güte einzelner statistischer Verfahren beeinflussen können.

Eine kontinuierliche Weiterentwicklung der Verknüpfung der bestehenden Förder- und FuE-Daten ist perspektivisch vorgesehen. Im Rahmen des Projekts wurden hierfür die konzeptionellen und technischen Grundlagen geschaffen, die künftig eine Bereinigung bislang unpräziser Verknüpfungen sowie eine Anreicherung mit zusätzlichen Datenquellen ermöglichen können.

Perspektivisch besteht insbesondere die Möglichkeit, weitere innovationsrelevante Indikatoren zu integrieren, etwa Patentinformationen, regionalstatistische Kennzahlen oder strukturierte Informationen aus Webrecherchen. Liegen externe Datenquellen mit eindeutigen Identifikatoren – insbesondere der Crefonummer – vor, können diese effizient mit mehreren bestehenden Datenressourcen verknüpft werden.

#### **4. Fazit und Ausblick**

Mit dem Projekt EVALDAT wurde ein zentraler methodischer und infrastruktureller Grundstein für evidenzbasierte Analysen der Forschungs- und Innovationsförderung von Unternehmen in Deutschland gelegt. Durch die Zusammenführung heterogener administrativer Förderdaten mit den FuE-Daten der WISTAT ist erstmals eine integrierte Paneldatenbasis entstanden, die für Analysen zur Förderauswahl, zu Förderintentionen sowie zu Förderwirkungen herangezogen werden kann.

Die Arbeiten zeigen zugleich, dass die Leistung des Projekts nicht allein in der verbesserten Datenverfügbarkeit durch eine breitere, verknüpfte Datenbasis liegt, sondern auch in der Analyse und Einordnung historisch gewachsener Datenstrukturen sowie in der systematischen Identifikation und Aufarbeitung der damit verbundenen Herausforderungen. Daraus konnten Ansatzpunkte für zukünftige Prozesse abgeleitet werden. Trotz verbleibender Einschränkungen – etwa infolge unvollständiger Identifikatoren, imputierter Werte oder gekennzeichnete Unsicherheiten – stellt der aufgebaute Datenbestand eine wertvolle Grundlage für weiterführende mikroökonomische Wirkungsanalysen dar.

Die im Rahmen des EVALDAT-Projekts zusammengestellten Daten können interessierten Forschenden unter bestimmten Bedingungen über das Forschungsdatenzentrum der WISTAT (FDZ) bereitgestellt werden. Voraussetzung ist zunächst der Nachweis eines beauftragten Evaluations- oder wissenschaftlichen Forschungsprojekts. Je nach Umfang der gewünschten Daten ist zudem die individuelle Prüfung und Zustimmung der ursprünglichen datengebenden Stellen erforderlich.

Interessierte Forschende stellen einen Antrag, in dem sie unter anderem ihr Forschungsvorhaben und die konkrete Datenanfrage formulieren – gegebenenfalls nach einem Beratungsgespräch mit dem FDZ. Auf dieser Grundlage erstellt das FDZ ein Angebot zur Datennutzung. Die Daten können vor Ort im FDZ genutzt werden und perspektivisch auch über einen

Remote-Zugang. Weiterführende Informationen zu Zugang und Nutzungsbedingungen sind auf den Seiten des FDZ verfügbar.<sup>3</sup>

Die Bereitstellung des bisherigen Datenstandes über das FDZ der WISTAT schafft die institutionellen Voraussetzungen für eine nachhaltige wissenschaftliche Nutzung und Weiterentwicklung des Datensatzes. Die bisherigen Arbeiten zeigen, dass das Potenzial der aufgebauten Datenbasis noch nicht ausgeschöpft ist und dass die im Projekt gewonnenen Erkenntnisse sowie identifizierten Herausforderungen zentrale Ansatzpunkte für eine Weiterentwicklung der Datenbasis liefern. Technische und methodische Fortschritte, vertiefte Analysen bestehender Inhalte, zusätzliche Berichtsjahre und Unternehmenseinheiten sowie die Einbindung weiterer Datenquellen können die Nutzbarkeit und Differenzierbarkeit der Daten weiter erhöhen. Aufbauend auf der im EVALDAT-Projekt geschaffenen Grundlage eröffnet dies die Möglichkeit, den Datensatz in zukünftigen Projekten schrittweise zu erweitern und seine analytische Aussagekraft weiter zu stärken.

---

<sup>3</sup> <https://www.fdz-wissenschaftsstatistik.de/>

## Quellen

Cerulli, G. (2010), Modelling and Measuring the Effect of Public Subsidies on Business R&D: A Critical Review of the Econometric Literature. *Economic Record*, 86, S. 421-449. <https://doi.org/10.1111/j.1475-4932.2009.00615.x>

Falk, M.Th. & Svensson, R. (2020): Evaluation criteria versus firm characteristics as determinants of public R&D funding. In: *Science and Public Policy*, Volume 47, Issue 4, August 2020, Pages 525–535, <https://doi.org/10.1093/scipol/scaa032>

Peters, B., Hud, M., Köhler, Ch. & Licht, G. (2012): Ökonomische Bewertung von staatlichen Investitionen in Forschung und Innovation. *Studien zum deutschen Innovationssystem Nr. 15-2012*, Zentrum für Europäische Wirtschaftsforschung (ZEW)

Zúñiga-Vicente, J.Á., Alonso-Borrego, C., Forcadell, F.J. and Galán, J.I. (2014): Assessing the effect of public subsidies on firm R&D investment: a survey. *Journal of Economic Surveys*, 28, S. 36-67. <https://doi.org/10.1111/j.1467-6419.2012.00738.x>

## Datenquellen

Publications Office Of The European Union (2015): *CORDIS - EU research projects under FP7 (2007-2013)* [Data set]. Publications Office of the European Union. <https://doi.org/10.2906/112117098108/11>

Publications Office (2015): *CORDIS - EU research projects under Horizon 2020 (2014-2020)* [Data set]. Publications Office of the European Union. <https://doi.org/10.2906/112117098108/12>

Publications Office (2022): *CORDIS - EU research projects under HORIZON EUROPE (2021-2027)* [Data set]. Publications Office. <https://doi.org/10.2906/112117098108/20>

Bureau van Dijk (2025). *MARKUS-Datenbank* [unveröffentlichter Datensatz].

SV Wissenschaftsstatistik (2025): *Mikrodaten aus der Erhebung zu Forschung und Entwicklung im Wirtschaftssektor* [unveröffentlichter Datensatz].

Bundesministerium für Forschung, Technologie und Raumfahrt (2024): *Förderkatalog* [unveröffentlichter Datensatz]. Öffentlich zugängliche Version der Daten zu finden unter: <https://foerderportal.bund.de/foekat/>

Bundesministerium für Wirtschaft und Energie (2024): *Zentrales Innovationsprogramm Mittelstand (ZIM)* [unveröffentlichter Datensatz].

Bescheinigungsstelle Forschungszulage (2024): Mikrodatensatz mit Anträgen zur Forschungszulage [unveröffentlichter Datensatz].

ZEW (2024): Aufbereiteter Datensatz zur EFRE-Förderung [unveröffentlichter Datensatz].

---

### Impressum

Essen, 2026

DOI: 10.60598/pub.evaldat.2026

### Herausgeber

Stifterverband für die Deutsche Wissenschaft e.V.

Baedekerstraße 1 · 45128 Essen

T 0201 8401-0 · mail@stifterverband.de

www.stifterverband.org

### Redaktion

Simone Höfer

### Zitationshinweis

Elsner, S. & Hellmann, M. (2026). Neue Datenbasis zur Evaluation der FuE-Förderung von Unternehmen in Deutschland. Essen, Stifterverband.

---