

Titze, Mirko; Brachert, Matthias; Giebler, Alexander; Ehrenfeld, Wilfried

**Research Report**

**IWH-FuE-Mikrodatenbank. Teil 1: Daten, Datenherkunft und Datenqualität. Überarbeitung und Erweiterung der Fassung vom 14.10.2013**

IWH Technical Reports, No. 01/2016

**Provided in Cooperation with:**

Halle Institute for Economic Research (IWH) – Member of the Leibniz Association

*Suggested Citation:* Titze, Mirko; Brachert, Matthias; Giebler, Alexander; Ehrenfeld, Wilfried (2016) : IWH-FuE-Mikrodatenbank. Teil 1: Daten, Datenherkunft und Datenqualität. Überarbeitung und Erweiterung der Fassung vom 14.10.2013, IWH Technical Reports, No. 01/2016, Leibniz-Institut für Wirtschaftsforschung Halle (IWH), Halle (Saale)

This Version is available at:

<https://hdl.handle.net/10419/144722>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# IWH Technical Reports

Nr. 1

April 2016

## IWH-FuE-Mikrodatenbank



### Teil 1:

Daten, Datenherkunft und Datenqualität

Überarbeitung und Erweiterung der Fassung vom 14.10.2013

Mirko Titze, Matthias Brachert, Alexander Giebler, Wilfried Ehrenfeld

**Autoren:**

Mirko Titze, Tel +49 345 7753 861, [mirko.titze@iwh-halle.de](mailto:mirko.titze@iwh-halle.de)  
Matthias Brachert  
Alexander Giebler  
Wilfried Ehrenfeld

**Kontakt:**

Cornelia Lang  
Leiterin des IWH-Forschungsdatenzentrums  
Tel + 49 345 77 53 802  
Fax + 49 345 77 53 820  
[cornelia.lang@iwh-halle.de](mailto:cornelia.lang@iwh-halle.de)

**Herausgeber:**

Leibniz-Institut für  
Wirtschaftsforschung Halle (IWH)

**Geschäftsführender Vorstand:**

Prof. Reint E. Gropp, Ph.D.  
Prof. Dr. Oliver Holtemöller  
Dr. Tankred Schuhmann

**Hausanschrift:**

Kleine Märkerstraße 8  
D-06108 Halle (Saale)

**Postanschrift:**

Postfach 11 03 61  
D-06017 Halle (Saale)

Tel +49 345 7753 60  
Fax +49 345 7753 820

[www.iwh-halle.de](http://www.iwh-halle.de)

Alle Rechte vorbehalten

**Zitierhinweis:**

*Titze, Mirko; Brachert, Matthias; Giebler, Alexander; Ehrenfeld, Wilfried:* IWH-FuE-Mikrodatenbank – Teil 1: Daten, Datenherkunft und Datenqualität. Überarbeitung und Erweiterung der Fassung vom 14.10.2013. IWH Technical Reports 01/2016. Halle (Saale) 2016.

ISSN 2365-9076

# IWH-FuE-Mikrodatenbank

## Teil 1: Daten, Datenherkunft und Datenqualität

Überarbeitung und Erweiterung der Fassung vom 14.10.2013

*Mirko Titze, Matthias Brachert, Alexander Giebler, Wilfried Ehrenfeld*

### Zusammenfassung

Nahezu alle entwickelten Volkswirtschaften haben Programme zur Förderung von Forschung und Entwicklung (FuE) eingeführt. Allerdings ist sehr wenig darüber bekannt, welche Programme eigentlich genau zur Anwendung kommen, welche finanziellen Mittel dafür aufgebracht werden und ob die Programme in der Art und Weise wirken, wie sie ursprünglich intendiert waren. Evaluationsstudien, die auf kausalen Untersuchungsdesigns basieren, können einen wertvollen Beitrag zur Beantwortung der Frage leisten, ob ein Programm tatsächlich Wirkungen entfaltet und welcher der verschiedenen Ansätze am erfolgversprechendsten ist. Vor diesem Hintergrund haben Forscher am Leibniz-Institut für Wirtschaftsforschung Halle (IWH) damit begonnen, systematisch eine Vielzahl an verschiedenen Datensätzen der FuE-Projektförderung in Deutschland zu sammeln. Unter Verwendung von umfassenden und komplexen Methoden wurden diese Datensätze harmonisiert mit Firmendatenbanken zusammengespielt. Der vorliegende technische Report gibt einen Überblick über die in der IWH-FuE-Mikrodatenbank vorhandenen Programme und ihrer Variablen. Darüber hinaus geht der Report auch auf die Harmonisierungsroutinen ein.

## Inhaltsverzeichnis

|                                                                                                  |    |
|--------------------------------------------------------------------------------------------------|----|
| Zusammenfassung.....                                                                             | 1  |
| 1 Überblick.....                                                                                 | 3  |
| 2 Übersicht über die in der Datenbank enthaltenen Förderprogramme.....                           | 5  |
| 3 Datenanforderungen.....                                                                        | 7  |
| 3.1 Informationen in der Förderdatenbank und ihre Eignung für evidenzbasierte Evaluationen ..... | 7  |
| 3.2 Geeignete Datenquellen für evidenzbasierte Evaluationen .....                                | 8  |
| 3.2.1 Bureau van Dijk (BvD) .....                                                                | 8  |
| 3.2.2 Beschäftigtenstatistik der Bundesagentur für Arbeit (BA) .....                             | 9  |
| 3.2.3 Amtliche Firmendaten für Deutschland (AFiD) des Statistischen Bundesamts .....             | 9  |
| 3.2.4 Daten des Stifterbands für die Deutsche Wissenschaft .....                                 | 10 |
| 3.2.5 Einrichtungen der öffentlich finanzierten Wissenschaft.....                                | 10 |
| 3.2.6 Weitere einzelwirtschaftliche Indikatoren .....                                            | 10 |
| 4 Harmonisieren und Verknüpfen von Datensätzen über Record-Linkage-Techniken .....               | 11 |
| 4.1 Harmonisierung der IWH-FuE-Mikrodatenbank.....                                               | 11 |
| 4.2 Das Beispiel Sachsen.....                                                                    | 12 |
| 5 Weitere Bearbeitungsschritte am IWH .....                                                      | 14 |
| Literatur .....                                                                                  | 16 |

# IWH-FuE-Mikrodatenbank

## Teil 1: Daten, Datenherkunft und Datenqualität

Überarbeitung und Erweiterung der Fassung vom 14.10.2013

*Mirko Titze, Matthias Brachert, Alexander Giebler, Wilfried Ehrenfeld*

### 1 Überblick

Das IWH untersucht wirtschaftliche Aufholprozesse und die ökonomische Integration in Europa. Wirtschaftliche Aufhol- und Wachstumsprozesse hängen von der Effizienz der Allokation der Produktionsfaktoren und dem Produktivitätsfortschritt ab. Das IWH erforscht insbesondere, wie das Finanzsystem Kapital(re)-allokation, Strukturwandel, Innovationen, Produktivitätsfortschritt und somit eine effiziente und nachhaltige realwirtschaftliche Entwicklung gewährleisten kann. Ein wichtiger Aspekt in diesem Zusammenhang ist die Erforschung des Zusammenhangs zwischen Produktivität auf Unternehmensebene und Innovationen. Dies umfasst auch die systematische und umfassende Analyse der eingesetzten FuE-Förderprogramme.

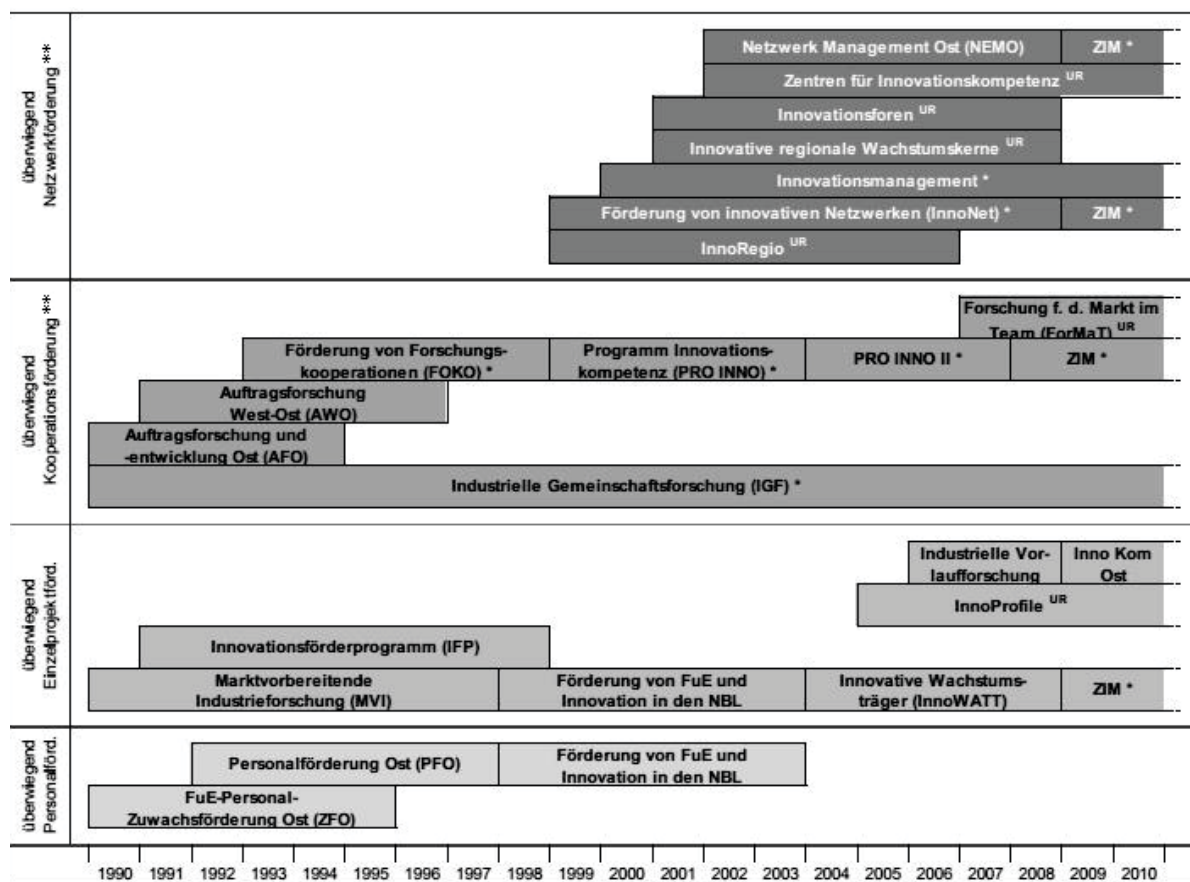
Alle modernen Industrienationen setzen heute in mehr oder weniger großem Umfang Instrumente zur Förderung von Forschung und Entwicklung (FuE) ein. Begründet wird die Anwendung von FuE-Förderprogrammen in der Regel damit, dass „Marktversagen“ den optimalen Umfang von FuE verhindert – Innovationen aber heute die treibende Kraft (regionalen) wirtschaftlichen Wachstums sind. Aus wissenschaftlicher Sicht ist der Nachweis eines Marktversagens allerdings nur unter äußerst restriktiven Bedingungen möglich, sodass schon allein deswegen eine wissenschaftliche Begleitung von staatlichen Fördermaßnahmen geboten ist.

Enger werdende Spielräume in den öffentlichen Haushalten erzwingen einen effizienten Einsatz der Fördermittel. Evidenzbasierte Evaluationen können diesbezüglich wertvolle Impulse für die Ausgestaltung der Förderrichtlinien geben. Das Ziel solcher Evaluationen muss sein, einen möglichst klaren Zusammenhang zwischen der Fördermaßnahme und ihrer Wirkung herzustellen. Zentral für eine Evaluierung sind neben klar definierten wirtschaftspolitischen Zielen auch Informationen über die Untersuchungseinheiten (z. B. Betriebe oder Regionen), und zwar nicht nur über die geförderten, sondern auch über die nicht geförderten. Diese Vorgehensweise stellt hohe Anforderungen an die Daten. Evidenzbasierte Evaluationen sind in der Regel mit den administrativen Förderdaten allein nicht durchführbar. Stattdessen ist ein Hinzuspielen weiterer Daten erforderlich, die nicht nur die geförderten Einheiten hinreichend genau beschreiben, sondern auch jene der Kontrollgruppe.

In der Praxis wird eine Vielzahl von Programmen zur Förderung von FuE und Innovationen angewendet, die zudem einem institutionellen Wandel (Zielgruppen, Förderquoten, Fördergegenstand usw.) unterliegen. Abbildung 1 verdeutlicht dies für (ausgewählte) Förderprogramme des Bundes, die in den Neuen Ländern seit der Wiedervereinigung eingesetzt wurden.

**Abbildung 1:**

Förderprogramme des Bundes (BMWi und BMBF) für FuE und Innovation in der ostdeutschen Wirtschaft (einschließlich Berlin) seit 1990



Anmerkungen: \* bundesweite Programme. – UR: Initiativen der Programmfamilie „Unternehmen Region“. – \*\* Netzwerkförderung beinhaltet in der Regel die Unterstützung des Netzwerkmanagements, während die Kooperationsförderung auf die Durchführung von Forschungsprojekten fokussiert. – Gründungsförderungsprogramme sind nicht enthalten.

Quelle: Günther, Jutta; Nulsch, Nicole; Urban-Thielicke, Dana; Wilde, Katja: 20 Jahre nach dem Mauerfall: Transformation und Erneuerung des ostdeutschen Innovationssystems. Expertenkommission Forschung und Innovation (Hrsg.), Studien zum deutschen Innovationssystem Nr. 16-2010, Berlin 2010, S. 32.

Ein wichtiger Aspekt ist, dass nicht nur der Bund Förderprogramme anwendet, sondern auch die Länder und die Europäische Union. Dem Antragsteller steht es frei, sich für eine der vielen Förderinstitutionen zu entscheiden. Damit aber wird unmittelbar deutlich, dass nur eine Kombination der Datensätze der verschiedenen Förderinstitutionen ein umfassendes Bild der FuE- und Innovationsförderung abgeben kann. Die **Einzigkeit** der IWH-FuE-Mikrodatenbank besteht darin, dass sie dieser Anforderung gerecht wird.

Im Mittelpunkt der Forschung mit FuE-Mikrodaten am IWH stehen derzeit zwei Aspekte:

- Analysen zum Ausmaß von innovativen Vorhaben verschiedener Akteure (Hochschulen, außeruniversitäre Forschungseinrichtungen, private Wirtschaft) sowie zur Verflechtung der Akteure im Rahmen der Kooperationsförderung und
- Analysen zur Wirkung der FuE-Förderung (in einer weiten Abgrenzung).<sup>1</sup>

<sup>1</sup> Im weiteren Sinne wird in diesem Dokument auch die Investitionsförderung als Bestandteil der Innovationsförderung verstanden. Dies rührt daher, da die Regularien bspw. nur die Anschaffung

Die Daten dienen neben der rein **wissenschaftlichen Verwertung** der zielgerichteten **wirtschaftspolitischen Beratung** auf regionaler, nationaler und internationaler Ebene.<sup>2</sup> Die IWH-FuE-Mikrodatenbank leistet Beiträge zur räumlichen Verortung von FuE-Aktivitäten und zur Analyse von Wirkungen der FuE-Förderung aus betrieblicher und regionaler Perspektive. Das Projekt wird inhaltlich durch das **Zentrum für evidenzbasierte Politikberatung am IWH (IWH-CEP)**<sup>3</sup> koordiniert.

## 2 Übersicht über die in der Datenbank enthaltenen Förderprogramme

Die **IWH-FuE-Mikrodatenbank** umfasst zurzeit (Stand: 30.06.2015) acht Datensätze verschiedener Programme der direkten FuE-Projektförderung<sup>4</sup> des Bundes, der Länder und der Europäischen Union.

Die Daten stammen von den Projektträgern bzw. den zuständigen Ministerien. Der Zugriff ist bei einigen Datensätzen frei – andere wiederum können nur exklusiv über die Ministerien bezogen werden.<sup>5</sup>

Für die sächsischen Daten bspw. gilt, dass jedwede Veröffentlichung mit dem Leiter des Referats 37 (Innovationspolitik, Technologieförderung) im Sächsischen Staatsministerium für Wirtschaft, Arbeit und Verkehr (SMWA), Herrn Christoph Zimmer-Conrad, abgestimmt werden muss.

Die in der IWH-FuE-Mikrodatenbank kombinierten Programme sind in Tabelle 1 dargestellt. Sie gibt einen Überblick über die wichtigsten zu diesen Programmen erhobenen Merkmale. Die größte Abdeckung besteht zurzeit für die Länder Sachsen und Sachsen-Anhalt. Hier liegen Daten der direkten FuE-Projektförderung des Bundes, der Länder und der Europäischen Union sowie Daten zu den Kooperationspartnern der Antragsteller vor.

---

neuer Maschinen und Anlagen erlauben. Die Literatur nennt in diesem Zusammenhang sogenannte product-embodied R&D flows, also Innovationen, die über die Lieferkette weitergereicht werden (bspw. OECD 1996 und Papaconstantinou et al. 1996).

<sup>2</sup> Am IWH laufen aktuell zwei Forschungsprojekte zur evidenzbasierten Evaluierung der Förderung, erstens *Möglichkeiten einer wissenschaftlichen Evaluation der gewerblichen Förderung im Rahmen der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) in Sachsen-Anhalt* und zweitens *Evaluation der Förderinitiative „Innovative regionale Wachstumskerne“ im Rahmen der BMBF-Innovationsinitiative für die Neuen Länder „Unternehmen Region“*.

<sup>3</sup> Das Zentrum für evidenzbasierte Politikberatung am IWH (IWH-CEP) wurde im Jahr 2014 gegründet. Es handelt sich hierbei um eine Plattform, die Aktivitäten in Forschung, Lehre und Politikberatung bündelt und strukturiert mit dem Ziel, bessere Grundlagen für eine kausale Analyse wirtschaftspolitischer Maßnahmen in Deutschland zu schaffen. Das IWH-CEP ist als Serviceeinheit konzipiert und unterstützt die Aktivitäten in den Forschungsgruppen, indem der Zugang zu einem überregionalen Forschungs- und Politikberatungs-netzwerk sowie der Zugang zu Datensätzen für Kausalanalysen bereitgestellt wird. Weitere Informationen sind verfügbar unter dem folgenden Link: <http://www.iwh-halle.de/d/forschung/cep/start.asp>.

<sup>4</sup> Dieses Dokument verwendet eine weite Fassung des Begriffs Forschung und Entwicklung (FuE). In der weiten Abgrenzung zählen dazu auch Investitionsförderprogramme.

<sup>5</sup> Für alle Interessenten frei zugängliche Datensätze sind z. B. der Förderkatalog für geförderte Bundesprojekte ([www.foerderkatalog.de](http://www.foerderkatalog.de)) und die cordis-Datenbank für geförderte EU-Projekte (<http://cordis.europa.eu/search/index.cfm?fuseaction=search.simple>). Alle anderen Daten unterliegen strenger Vertraulichkeit – Veröffentlichungen dieser Daten müssen von den (Landes)Ministerien genehmigt werden.



**Tabelle 1:**  
Datensätze<sup>a</sup> der IWH-FuE-Mikrodatenbank (Stand: 30.06.2015)

| Nr. Datensatz/<br>Förderprogramm                                            | Förder-<br>einrichtung  | Art der<br>Projekt-<br>förderung | Förder-<br>gegenstand<br>im<br>IWH<br>seit | Daten decken<br>Zeitraum ab | enthaltene<br>Regionen                              | Bereit-<br>stellung<br>der Daten<br>durch | Zugriff                                                     | Anzahl<br>enthaltener<br>Projekte |
|-----------------------------------------------------------------------------|-------------------------|----------------------------------|--------------------------------------------|-----------------------------|-----------------------------------------------------|-------------------------------------------|-------------------------------------------------------------|-----------------------------------|
| 1 Förderkatalog                                                             | Bund                    | Einzel+<br>Verbund               | Innovation                                 | 2010                        | 1968-2014 (Oktober)                                 | BMBF, DLR                                 | frei verfügbar <sup>c</sup> / wichtige<br>Merkmale exklusiv | 162 910                           |
| 2 Förderdatenbank der SAB                                                   | Freistaat<br>Sachsen    | Einzel+<br>Verbund               | Innovation                                 | 2007                        | 1991-2012 (September)                               | SAB, SMWK                                 | exklusiv                                                    | 7 816                             |
| 3 Förderdatenbank                                                           | Land Sachsen-<br>Anhalt | Einzel+<br>Verbund               | Innovation                                 | 2012                        | 1998-2011                                           | MWW, IB                                   | exklusiv                                                    | 1 238                             |
| 4 Zentrales Innovationsprogramm Mittelstand<br>(ZIM), Koop+Solo             | Bund                    | Einzel+<br>Verbund               | Innovation                                 | 2012                        | 2004-2012 (Koop),<br>2009-2012 (Solo)               | BMWi, AiF,<br>EuroNorm                    | exklusiv                                                    | 4 899                             |
| 5 Pro Inno + Pro Inno II (Vorläufer von ZIM)                                | Bund                    | Verbund                          | Innovation                                 | 2013                        | 1999-2004 <sup>e</sup> , 2004-<br>2009 <sup>f</sup> | BMWi, AiF                                 | exklusiv                                                    | 4 406                             |
| 6 EU FP 7                                                                   | EU                      | Verbund                          | Innovation                                 | 2013                        | 2007-2013                                           | SMWK                                      | frei verfügbar <sup>d</sup> / wichtige<br>Merkmale exklusiv | 18 507                            |
| 7 EU FP 6                                                                   | EU                      | Verbund                          | Innovation                                 | 2015                        | 2000-2006                                           | SMWK,<br>BMBF                             | frei verfügbar <sup>d</sup> / wichtige<br>Merkmale exklusiv | 10 107                            |
| 8 Gemeinschaftsaufgabe „Verbesserung der<br>regionalen Wirtschaftsstruktur“ | Bund+ Länder            | Einzel                           | Investition                                | 2014                        | 2007-2013                                           | MWW, IB                                   | frei verfügbar <sup>g</sup> / wichtige<br>Merkmale exklusiv | 1 654                             |

Abkürzungen: AiF – Arbeitsgemeinschaft industrieller Forschungsvereinigungen, AiF Projekt GmbH, Projektträger des Bundesministeriums für Wirtschaft und Technologie (BMWi); BMBF – Bundesministerium für Bildung und Forschung; BMWi – Bundesministerium für Wirtschaft und Technologie; DLR – Projektträger im Deutschen Zentrum für Luft- und Raumfahrt; EU – Europäische Union; EuroNorm – EuroNorm Gesellschaft für Qualitätssicherung und Innovationsmanagement mbH, Projektträger des Bundes; IB – Investitionsbank Sachsen-Anhalt; MWW – Ministerium für Wissenschaft und Wirtschaft des Landes Sachsen-Anhalt; SAB – Sächsische Aufbaubank - Förderbank; SMWK – Sächsisches Staatsministerium für Wissenschaft und Kunst.

Erläuterungen: <sup>a</sup> Die originären Datensätze haben keine Panelstruktur, da die beobachtete Einheit das Projekt ist. Die Merkmale des Projektes (z. B. beteiligte Partner, bewilligte Summe, Branchenzugehörigkeit usw.) ändern sich während der Laufzeit des Projektes nicht. Die Datensätze können aber in eine Panelstruktur überführt werden, wenn die Zuwendungsempfänger (bzw. ausführenden Stellen) eindeutig, z. B. mit einer Identifikationsnummer, erfasst sind. Damit lässt sich im Zeitverlauf darstellen, wann ein Zuwendungsempfänger ein Projekt beantragt und bearbeitet hat. Über eindeutige Identifikationsnummern der Akteure verfügen die Datensätze der Nr. 2, 4, 5 und 6. Diese sind aber nur innerhalb der Datensätze und nicht zwischen den Datensätzen eindeutig (vgl. hierzu auch Abschnitt 3). – <sup>b</sup> Sowie Kooperationspartner außerhalb des Landes. – <sup>c</sup> <http://www.foerderkatalog.de>. – <sup>d</sup> [http://cordis.europa.eu/fp7/projects\\_de.html](http://cordis.europa.eu/fp7/projects_de.html) – <sup>e</sup> Pro Inno. – <sup>f</sup> Pro Inno II. – <sup>g</sup> Gemäß Vereinbarungen zur EU Strukturfondsförderung wird eine Liste der Begünstigten aus öffentlichen Mitteln für das Land Sachsen-Anhalt bei kofinanzierten Projekten erstellt.

Quelle: Eigene Darstellung.

### 3 Datenanforderungen

#### 3.1 Informationen in der Förderdatenbank und ihre Eignung für evidenzbasierte Evaluationen

Wirkungsanalysen sollen eine Verbindung zwischen den staatlichen Interventionen und vorgegebenen Zielen herstellen. Das fundamentale Evaluationsproblem besteht in der Beantwortung der Frage, was ohne eine staatliche Intervention geschehen wäre. Technisch gesprochen geht es um die Schaffung einer „kontrafaktischen“ Situation: wie hätte sich bspw. ein bestimmtes Unternehmen verhalten, wenn es keine Förderung bekommen hätte? Die Frage des „Was wäre gewesen, wenn ein Unternehmen nicht gefördert worden wäre?“ lässt sich nur durch den Kontrast zwischen geförderten und nicht-geförderten Unternehmen beantworten.

Grundsätzlich lassen sich die evidenzbasierten Evaluationsverfahren unterscheiden in Makro- und Mikro-Methoden. Eine Analyse der gesamtwirtschaftlichen Effekte von Fördermaßnahmen ist jedoch in der Regel schwierig (Bade et al. 2012), weshalb gängige Studien häufig auf die Mikroebene fokussieren. Derartige Analysen erfordern dann einzelwirtschaftliche Informationen (und zwar nicht nur für die Behandlungsgruppe, sondern auch für eine Kontrollgruppe) über:

- die geförderte Wirtschaftseinheit (vor allem Name, Region und Branche, Größe),
- die Förderung (Höhe der Förderung, Zeitpunkt der Förderung, rechtliche Regelungen) sowie
- ökonomische Zielgrößen (bspw. Beschäftigung, Löhne/Gehälter oder Umsatz).

Die genannten **Informationen** allerdings sind den in der IWH-FuE-Mikrodatenbank aufgeführten **Förderprogrammen nicht zu entnehmen**. Jeder der in Tabelle 1 aufgelisteten Datensätze enthält zunächst eine eindeutige Verschlüsselung der Projekte. Anhand einer Identifikationsnummer ist erkennbar, welche Projekte in Kooperation durchgeführt wurden und welche nicht. Je nach Programm sind Unternehmen, außeruniversitäre Forschungseinrichtungen und Hochschulen antragsberechtigt. Kooperationen finden daher in den verschiedenen Kombinationen aller drei genannten Institutionen statt, z. B. Hochschule-Unternehmen, Unternehmen-außeruniversitäre Forschungseinrichtung-Hochschule, Unternehmen-Unternehmen usw.

Jeder der Datensätze enthält darüber hinaus Informationen zum Namen des Antragstellers, seinem Regionalmerkmal (so genannte Allgemeine Gemeindekennziffer bzw. Ortsname mit PLZ), seiner Branche, der Projektlaufzeit (Datum zu Beginn und Ende) und der Bewilligungssumme. Der Datensatz der sächsischen FuE-Projektförderung (Nr. 2 in Tabelle 1) enthält weitere Merkmale, z. B. zum Projektvolumen, zu Förderquoten usw., die in den anderen Datensätzen nicht enthalten sind.

Da die Datensätze von unterschiedlichen Fördereinrichtungen bereitgestellt werden, sind die Variablenbezeichnungen für identische Merkmale unterschiedlich, bspw. die Bezeichnungen für die Fördersumme, den Beginn und das Ende der Finanzierung, den Zuwendungsempfänger und die ausführende Stelle usw. Am IWH wurde eine Konkordanz über die Variablennamen erstellt, die ein Zusammenspielen der Datensätze erlaubt. Des Weiteren beruhen die Brancheninformationen auf unterschiedlichen Klassifikationssystemen (z. B. Nace Rev. 1.1 versus Nace Rev. 2) und unterschiedlicher Gliederungstiefe (Abschnitte, Abteilungen, Gruppen, Klassen), die Regionalinformationen auf unterschiedlichen Gebietsständen. Auch diese Informationen werden zurzeit am IWH harmonisiert.

### 3.2 Geeignete Datenquellen für evidenzbasierte Evaluationen

Um dennoch evidenzbasierte Analysen durchzuführen, müssen Informationen über die geförderten Einheiten (und die Gruppe der nicht geförderten Akteure) aus anderen Datenquellen hinzugespielt werden. Grundsätzlich kommen für die Generierung fehlender ökonomischer Ziel- und Kontrollvariablen zwei Wege in Betracht: die Primärerhebung der Daten über einen Fragebogen sowie die Nutzung von Sekundärquellen (Abbildung 2).

Eine Primärerhebung verursacht einen hohen zeitlichen und finanziellen Aufwand – nicht nur bei der erhebenden Institution, sondern auch bei den befragten Akteuren. Die Teilnahme an einer solchen Befragung beruht in der Regel auf Freiwilligkeit. Hohe Rücklaufquoten sind nicht garantiert. Aufgrund dieser Nachteile geht heute vor allem die amtliche Statistik dazu über, Daten, die schon in irgendeiner Form erfasst worden sind, nicht erneut zu erheben. Vielmehr werden über geeignete Techniken der Datenverarbeitung (sogenannte Record-Linkage-Methoden) Informationen aus mehreren Datenquellen zusammengefügt.

**Abbildung 2:**  
Übersicht über geeignete Datenquellen



Quelle: In Anlehnung an IWH (2014).

Eine Zusammenfassung zu den Vor- und Nachteilen der im Folgenden näher erläuterten Datensätze findet sich in IWH (2014, Abschnitt 6.5).

#### 3.2.1 Bureau van Dijk (BvD)

Das Unternehmen Bureau van Dijk (BvD) bietet in Zusammenarbeit mit Creditreform e. V. in der sogenannten dafne-Datenbank Geschäftsinformationen von ca. 1,3 Millionen deutschen Unternehmen an. Diese Datenbank enthält wichtige Unternehmensmerkmale (z. B. Rechtsform,

Branchenschlüssel, Gemeindekennziffer, Anzahl Mitarbeiter, Umsatz, Kapital usw.). Die Daten stammen nicht aus der amtlichen Statistik, sondern beruhen im Wesentlichen auf veröffentlichungspflichtigen Angaben der Kapitalgesellschaften sowie auf Selbstauskünften der Unternehmen. Der Datensatz weist mitunter Lücken bei wichtigen Kennziffern wie der Anzahl der Mitarbeiter, beim Umsatz oder dem Kapital auf.

Ein Zusammenführen mit der Förderstatistik lässt sich anhand des Betriebsnamens und eines Regionalmerkmals unter Verwendung von Record-Linkage-Techniken vornehmen. BvD-Daten liegen im Wesentlichen für den Zeitraum 2003-2012 im IWH vor.

### 3.2.2 Beschäftigtenstatistik der Bundesagentur für Arbeit (BA)

Die Verwendung der Beschäftigtenstatistik der BA bringt eine Reihe von Vorteilen mit sich (Bade und Alm 2010, S. 2 ff.): gleichmäßige und zuverlässige Erfassung, Informationen über zusätzliche Merkmale (z. B. Arbeitseinkommen, Qualifikation der Beschäftigten, ausgeübte Tätigkeit) sowie ein hoher Erfassungsgrad. Diese zusätzlichen Merkmale erlauben somit eine Evaluation der Effekte der Förderung auf Quantität und Qualität der geschaffenen Arbeitsplätze.

Ein weiterer Vorteil besteht darin, dass sich die Förderfälle vergleichsweise unkompliziert mit der Beschäftigtenstatistik über die Betriebsnummer verknüpfen lassen. Die Betriebsnummer wird in der Förderstatistik mit erhoben. Gewisse Einschränkungen sind durch die Fokussierung auf sozialversicherungspflichtig Beschäftigte gegeben. Gleichwohl dürfte der Abdeckungsgrad dennoch rund 70-80% aller Erwerbstätigen betragen. Ein Nachteil besteht darin, dass diese Statistik keine Angaben über die Beschäftigung hinaus erfasst (z. B. Umsatz, Produktpalette oder Investitionen). Ein Zugang zu dieser Datenquelle ist nur über einen Antrag an die BA bzw. das Bundesministerium für Arbeit und Soziales möglich. Zwischen dem IWH und dem Regionalen Forschungsnetzwerk am IAB (Standorte Chemnitz und Halle) besteht seit Dezember 2014 ein Kooperationsvertrag. Der Datensatz Nr. 8 aus Tabelle 1 wurde bereits mit Betriebsnummern verknüpft.

### 3.2.3 Amtliche Firmendaten für Deutschland (AFiD) des Statistischen Bundesamts

Eine weitere Quelle, die Angaben über die Grundgesamtheit und geeignete ökonomische Ziel- und Kontrollgrößen enthält, sind die amtlichen Firmendaten für Deutschland (AFiD). Diese Daten werden von den Statistischen Ämtern der Länder bereitgestellt. Über das Unternehmensregister verknüpfen sie alle Mikrodaten der Wirtschafts- und Umweltstatistiken. Zudem wurden die Daten in eine Panelstruktur überführt. Integrierte Paneldaten stehen zurzeit für die Bereiche Landwirtschaft, Dienstleistungen, Verarbeitendes Gewerbe und Energie zur Verfügung. Vor allem das Verarbeitende Gewerbe spielt in der FuE-Förderung eine große Rolle. Für eine Analyse genau dieser Branchen eignen sich die AFiD-Panel Industriebetriebe und Industrieunternehmen. Es handelt sich um eine Totalerhebung von Einheiten mit 20 oder mehr Beschäftigten. Zum jährlichen Berichtskreis gehören 68.000 Unternehmen.

Gegenüber der Beschäftigtenstatistik der BA weist dieser Datensatz den Vorteil auf, dass er neben der Beschäftigung auch zusätzliche ökonomische Zielgrößen, wie z. B. Umsatz und Investitionen enthält. Damit können Aussagen über die Produktivität – d. h. über die Verbesserung der Wettbewerbsfähigkeit der Unternehmen – getroffen werden.

Der genannte Datensatz wird im Forschungsdatenzentrum (FDZ) des Statistischen Landesamts Sachsen-Anhalts in Halle (Saale) gepflegt. Ein Hinzuspielen der Förderinformationen über Record-Linkage-Techniken ist möglich und würde von den Mitarbeitern des FDZ vorgenommen. Praktisch wird dies anhand der Handelsregisternummern, der Steuernummern sowie der Nummern von Bureau van Dijk vorgenommen. Diese Angaben sind in den administrativen Daten jedoch nicht enthalten, weshalb in einem ersten Schritt eine Verknüpfung mit dem BvD-Datensatz (siehe Abschnitt 3.2.1) vorgenommen werden muss. Die rechtliche Anforderung für

das Zusammenführen der Daten ergibt sich aus §13a des Bundesstatistikgesetzes. Danach müssen die hinzuzuspielenden Daten allgemein zugänglichen Quellen entstammen.

Für den Dienstleistungsbereich liegt ebenfalls ein Panel vor – allerdings handelt es sich um eine Stichprobenziehung. In dieser sind rund 15% aller Dienstleistungsunternehmen enthalten. Für die Evaluation von geförderten Betrieben der Dienstleistungsbranchen scheint dieser Datensatz weniger geeignet.

Seit Juni 2015 besteht zwischen dem IWH und dem Statistischen Landesamt ein Vertrag zur Nutzung der AFiD Daten. Ein Matching mit den Förderdatensätzen Nr. 1 und 8 läuft aktuell – es geht darum, die Trefferquoten zu verbessern.

### 3.2.4 Daten des Stifterbands für die Deutsche Wissenschaft

Alle bislang beschriebenen Datensätze enthielten keine (oder nur geringe) Informationen über die FuE-Aktivitäten der Unternehmen. Diese werden von der Wissenschaftsstatistik GmbH im Stifterverband für die Deutsche Wissenschaft im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) erhoben. Wesentliche Kenngrößen sind hier die FuE-Aufwendungen sowie das FuE-Personal. Ein Verknüpfen der Datensätze kann über das Förderkennzeichen (bei der Bundesförderung) und darüber hinaus über die Bureau van Dijk Nummer vorgenommen werden. Ein Kooperationsvertrag zwischen dem IWH und dem Stifterverband befindet sich zurzeit in Abstimmung.

### 3.2.5 Einrichtungen der öffentlich finanzierten Wissenschaft

Die in Tabelle 1 genannten Förderprogramme richten sich nicht nur an Akteure der Wirtschaft. Vielmehr legt die Politik großen Wert darauf, dass auch Akteure der öffentlich finanzierten Wissenschaft in die Innovationsprozesse eingebunden sind. Ganze Programmlinien zielen darauf ab, dass sich Unternehmen mit Hochschulen und außeruniversitären Forschungseinrichtungen vernetzen. Die Grundgesamtheit der öffentlich finanzierten wissenschaftlichen Einrichtungen wird im sogenannten Research Explorer aufgeführt. Diese Datenquelle wird als Online-Datenbank unter der Adresse [http://research-explorer.dfg.de/research\\_explorer.de.html](http://research-explorer.dfg.de/research_explorer.de.html) kostenlos bereitgestellt. Dieses Verzeichnis umfasst ca. 23.000 Einträge zu Instituten an deutschen Hochschulen und außeruniversitären Forschungseinrichtungen. Bei Universitäten ist die Untergliederung dabei bis auf Lehrstuhlebene abgebildet. Die Angaben sind nach geografischen, fachlichen und strukturellen Kriterien geordnet. Ein Zusammenführen mit der Förderstatistik lässt sich anhand des Namens und eines Regionalmerkmals unter Verwendung von Record-Linkage-Techniken vornehmen. Da die Daten des Research Explorer nicht eindeutig verschlüsselt sind, findet am IWH gegenwärtig eine systematische Aufbereitung statt.

### 3.2.6 Weitere einzelwirtschaftliche Indikatoren

Schließlich gibt es Datenbanken, die weitere wichtige Indikatoren des Innovationsprozesses bereitstellen. Hierbei handelt es sich um Patentdatenbanken (PATSTAT, RegPat, DPMA) sowie Publikationsdaten (Web of Knowledge Datenbank des Anbieters Thomson Reuters). Erfasste Einheiten in der erstgenannten Quelle sind Patente, bei denen im Wesentlichen Angaben zu Anmeldern, Erfindern, Patentklassen (Technologiefelder) sowie Patentzitationen aufgeführt sind. Zu Anmeldern und Erfindern liegt in der Regel auch ein Regionalmerkmal vor. Allerdings besteht bei den Patentdatenbanken häufig das Phänomen, dass Erfinder- und Anmelderort auseinanderfallen. Erfinder sind in einen großen Umfang auch Einzelpersonen, was eine Zuordnung zu einer Institution (Unternehmen oder wissenschaftliche Einrichtung) nahezu unmöglich macht, da die Affiliation in den Datenbanken nicht bereitgestellt wird. Unter Anwendung von Record-Linkage-Techniken wurden am IWH die Anmelder europäischer Patente (Datenbank RegPat) in der Bundesrepublik Deutschland harmonisiert und mit der

Datenbank des Anbieters Bureau van Dijk verknüpft (siehe dazu auch den Abschnitt 4 in diesem Dokument).

Die Angaben in Publikationsdatenbanken beruhen im Wesentlichen auf Veröffentlichungen in gerankten internationalen Fachzeitschriften. Hier gibt es in der Regel zu jedem Autor die Angabe einer Affiliation, die auch ein Regionalmerkmal beinhaltet. Allerdings unterscheiden sich die Schreibweisen der Autorennamen und der Affiliationen teilweise gravierend, so dass umfangreiche Arbeiten zur Harmonisierung der Daten notwendig sind. Über entsprechende Erfahrungen auf dem Gebiet der Publikationsdaten verfügt in der Bundesrepublik Deutschland das Kompetenzzentrum Bibliometrie in Bielefeld.<sup>6</sup>

## 4 Harmonisieren und Verknüpfen von Datensätzen über Record-Linkage-Techniken

### 4.1 Harmonisierung der IWH-FuE-Mikrodatenbank

Am IWH hat sich folgende Vorgehensweise für die Bereitstellung von Datensätzen für evidenzbasierte Evaluationen als geeignet erwiesen. In einem *ersten Schritt* werden die in Tabelle 1 genannten Förderdatensätze zusammengeführt, zunächst erfolgt anhand einer am IWH erstellten Konkordanz eine Vereinheitlichung der Variablennamen und der Variablenformate (z. B. beim Datumsformat).

Sodann werden die in der Datenbank genannten Akteure harmonisiert.<sup>7</sup> Da die Datensätze durch unterschiedliche Institutionen bereitgestellt werden, sind die Namen der Akteure nicht eindeutig erfasst. Hinzu kommen Schreibfehler bei der Erfassung der Namen. Von der Situation leicht abweichender Schreibweisen zu unterscheiden sind im technischen Sinne deutlich andere Bezeichnungen für dieselbe Institution. Als Beispiel hierfür seien die Technischen Universitäten genannt (mit ihre oft verwendeten Kurzform „TU“) sowie (als „Klassiker“) die Rheinisch-Westfälische Technische Hochschule Aachen (kurz: RWTH Aachen).

Hierzu ist eine systematische Harmonisierung der Akteure unter Nutzung von Record-Linkage- bzw. Data-Matching Techniken nötig (siehe hierzu beispielsweise Christen 2009 sowie Magerman et al. 2006). Der Begriff „Record Linkage“ bezeichnet dabei das Zusammenführen von Informationen zweier Datensätze, von denen angenommen wird, dass sie sich auf dieselbe Einheit/Entität beziehen (Herzog et al. 2007, S. 81).

Vor der eigentlichen Record-Linkage-Prozedur hat sich eine Vereinheitlichung der Regionalmerkmale bewährt, und zwar von Postleitzahlen auf den Amtlichen Gemeindeschlüssel (AGS 8). Zusätzlich zu berücksichtigen sind Gemeindegebietsreformen, bspw. in Sachsen 2006, in Sachsen-Anhalt 2008, in Mecklenburg-Vorpommern 2011. Nach der Anpassung der Regionalmerkmale ist es noch sinnvoll, „störende“ Zusatzinformationen aus den Akteursnamen entfernen (z. B. „Abteilung“, „Bereich“, „Fakultät“, „Lehrstuhl“).

Danach erfolgt ein sogenanntes „Precleaning“ (vgl. Magerman et al. 2006). Hierbei handelt es sich um eine Zeichenbereinigung. Dazu wird der Akteursname komplett in Großbuchstaben

---

<sup>6</sup> <http://www.forschungsinfo.de/Bibliometrie/index.php?id=home>.

<sup>7</sup> Hierbei ist zu beachten, dass ein Fördervorgang zwei Typen von Akteuren unterscheidet, den Zuwendungsempfänger und die ausführende Stelle. Erstgenannter erhält die Zuwendung, der Zweitgenannte bearbeitet das Projekt tatsächlich. In den meisten Fällen sind beide Akteurstypen identisch. In einigen Fällen aber treten gravierende Unterschiede auf. Der typische Fall sind die Institute und Anwendungszentren der Fraunhofer-Gesellschaft – Zuwendungsempfänger ist immer die Zentrale am Standort München. Tatsächlich bearbeitet aber werden die Projekte an den jeweiligen Institutsstandorten.



umgewandelt. Im Zuge dessen werden auch deutsche Umlaute oder Zeichen mit Akzenten ersetzt. Doppelte Leerzeichen im Namen sowie Leerzeichen am Beginn oder Ende des Namens werden entfernt. Anschließend werden Klammersymbole und Schreibweisen für „und“ vereinheitlicht sowie vorhandene Klammersymbole extrahiert. Anschließend werden die Gesellschaftsformen von Unternehmen identifiziert (und in einer separaten Variablen abgespeichert). Dies geschieht mittels einer Identifikationstabelle, die zum aktuellen Zeitpunkt über 600 Schreibweisen für verschiedene Gesellschaftsformen umfasst. Die Originalschreibweisen der Gesellschaftsform werden anschließend aus dem Unternehmensnamen entfernt. Schließlich werden einige Schreibweisen von häufig verwendeten, aber unterschiedlich geschriebenen Begriffen harmonisiert. Zuletzt werden alle Leerzeichen aus dem Ausdruck entfernt (Kondensierung). Diese Prozeduren eignen sich sehr gut um eine Zuordnung bei leicht abweichenden Schreibweisen für denselben Akteursnamen zu gewährleisten. Da Teile dieser Daten manuell erfasst wurden oder aus Scans von Dokumenten in Papierform stammen, sind auch einzelne Buchstabenfehler in den Datensätzen vorhanden. Hierzu gehören auch Unterschiede in der Verwendung von Trennstrichen oder Leerzeichen, die eine direkte Zuordnung verhindern. Versäumnisse in dieser Phase können selbst durch die Verwendung „unscharfer“ Zuordnungsalgorithmen kaum ausgeglichen werden. Die hier geleistete Arbeit ist wichtig und eine gute Investition in eine sichere Zuordnung.

Mit einer rein deterministischen Zuordnung der Originaldatensätze oder „unscharfer“ Methoden alleine ist es aber in Fällen sehr unterschiedlicher Schreibweisen (z. B. „TU“ für „Technische Universität“) kaum möglich, eine sichere Zuordnung zu gewährleisten. Diese Fälle können zur Standardisierung mittels automatisierter Ersetzungsregeln bzw. mit einer zusätzlichen Tabelle für unterschiedliche Schreibweisen derselben Institution erfasst werden.

Die eigentliche Record-Linkage-Prozedur greift auf die Software MergeToolBox zurück, welche am German Record Linkage Center<sup>8</sup> entwickelt und für die Wissenschaft zum Download bereitgestellt wurde (Schnell et al. 2005). Über einen Algorithmus (sogenannte Bi-Gramme) werden ähnliche Akteursnamen innerhalb einer gleichen Region (AGS 3-Steller) anhand eines eindeutigen Identifikationsmerkmals einander zugeordnet.

Der *zweite Schritt* beinhaltet schließlich dieselbe Vorgehensweise für das Anspielen von Sekundärdaten an die Förderdatensätze. Technisch gesprochen wird dem im IWH erzeugten Primärschlüssel für den Akteur der Primärschlüssel der externen Datenquelle zugeordnet.<sup>9</sup>

## 4.2 Das Beispiel Sachsen

Wie eingangs bereits erwähnt, besteht die größte Abdeckung an Förderprogrammen gegenwärtig für die Länder Sachsen und Sachsen-Anhalt. Exemplarisch erläutert dieser Abschnitt die Harmonisierung der Datensätze in der IWH-FuE-Förderdatenbank am Beispiel des Freistaats Sachsen. Aus der Tabelle 1 gehen die Datensätze Nr. 1 und 2 sowie 4 bis 7 in die Auswertung ein, und zwar in der Gruppierung nach Landesprogrammen (Nr. 2), Bundesprogrammen (Förderkatalog sowie ZIM und Vorläuferprogramme, d. h. Nr. 1, 4 und 5) und den EU Forschungsrahmenprogrammen (Nr. 5 und 6). Üblich ist zudem eine Unterteilung nach den Förderperioden der EU. Im konkreten Fall umfasst die Analyse die Perioden 2000-2006 sowie 2007-2013. Die Einleitung betonte bereits, dass die Fördermittellempfänger aus den

<sup>8</sup> <http://www.record-linkage.de>.

<sup>9</sup> Die hier beschriebene Prozedur wurde in dem vom BMBF-geförderten Projekt „Hochschulstrategien für Beiträge zur Regionalentwicklung unter Bedingungen demografischen Wandels“ erfolgreich angewendet. Daten der direkten FuE-Projektförderung des Bundes (Nr. 1 in Tabelle 1) wurden (für ausgewählte Untersuchungsregionen) mit Patent- und Bibliometriedaten zusammengespielt. Konkret ging es darum, zu messen, in welchem Ausmaß Akteure, die in den Genuss von FuE-Förderung gekommen sind, in unterschiedliche Bereiche der Wissensproduktion und des Wissenstransfers eingebunden sind (Titze et al. im Erscheinen).

verschiedenen Förderangeboten wählen können, da die Programmlinien in weiten Teilen Substitute darstellen.

Zweifelsohne bestehen zwischen ihnen marginale Unterschiede – im Großen und Ganzen jedoch sind sie ident:

- Es handelt sich in allen Fällen um eine Zuschussförderung (sogenannte nicht-rückzahlbare Zuschüsse).<sup>10</sup>
- Der Zuschuss bemisst sich nach einem Anteil an förderfähigen Kosten (im Wesentlichen Personalkosten für die am Projekt beteiligten Wissenschaftler). Die Abgrenzung der förderfähigen Kosten erfolgt letztendlich nach den Richtlinien der EU.
- Es gelten für alle drei Programmlinien dieselben Förderhöchstgrenzen (unterschieden nach Wissenschaft und Wirtschaft, gestaffelt nach Betriebsgrößenabgrenzungen gemäß der EU Richtlinien).

Aus diesen Konstellationen sind insgesamt sieben Kombinationen möglich, und zwar die drei Einzelformen sowie Mischformen. Tabelle 2 zeigt die Inanspruchnahme der verschiedenen Kombinationsmöglichkeiten durch die sächsischen Akteure. Insgesamt erhielten im Zeitraum von 2000 bis 2013 4230 Akteure Bewilligungen. 2120 (rund 50%) von ihnen wählten ausschließlich Bundesprogramme – der größte Anteil davon entfällt auf die ZIM-Förderung und ihre Vorläuferprogramme (in der Tabelle nicht explizit dargestellt). 1115 (rund 26%) der Akteure entschieden sich ausschließlich für die sächsische Landesförderung.

**Tabelle 2:**

### Inanspruchnahme von Förderprogrammen durch sächsische Akteure

Zeitraum 2000-2013

| Nr. | Kombination aus Förderprogrammen | Frequenz | in Prozent |
|-----|----------------------------------|----------|------------|
| 1   | nur Land                         | 1.115    | 26,4       |
| 2   | nur Bund                         | 2.120    | 50,1       |
| 3   | nur EU                           | 89       | 2,1        |
| 4   | Land und Bund                    | 741      | 17,5       |
| 5   | Land und EU                      | 8        | 0,2        |
| 6   | Bund und EU                      | 73       | 1,7        |
| 7   | Land, Bund und EU                | 84       | 2,0        |
|     | insgesamt                        | 4.230    | 100,0      |

Quelle: Rohdaten, IWH-FuE-Mikrodatenbank; Berechnungen des IWH.

Tabelle 2 verdeutlicht, dass die Evaluation eines einzelnen Programms ohne die generelle Einbettung in die gesamtdeutsche Förderlandschaft nicht möglich ist. Wird beispielsweise das Programm ZIM evaluiert, dann darf die Kontrollgruppe nicht aus den nicht-ZIM geförderten Akteuren bestehen. Es gäbe in der Gruppe der nicht ZIM-geförderten Akteure diejenigen, die ein Substitut von ZIM gewählt haben, beispielsweise die sächsische Landesförderung. Gleichwohl offenbart Tabelle 2 aber Evaluationsdesigns innerhalb der Gruppe von FuE geförderten Akteuren. Beispielsweise kann man für die Akteure, die die Landesförderung in Anspruch genommen haben den additionalen Effekt der Bundesförderung untersuchen (Zeilen 1 und 4).

<sup>10</sup> Andere Förderarten sind Darlehen, Zulagen, Bürgschaften.



Tabelle 3 stellt das Antragsverhalten im Zeitverlauf dar. Von den betrachteten 4230 Akteuren waren 1485 (rund 35%) nur im Zeitraum 2000-2006 in der Förderung, 1668 (rund 40%) sind dagegen in der Periode 2007-2013 neu dazugekommen. Insgesamt zeigt sich, dass viele Akteure (rund 60%) kontinuierlich in irgendeiner Art von FuE-Förderung verbleiben.

**Tabelle 3:**  
**Inanspruchnahme von Förderprogrammen durch sächsische Akteure im Zeitverlauf**

Zeitraum 2000-2006 und 2007-2013

| 2000-2006 \ 2007-2013 |                 |          |          |        |               |             |             |                   |           |
|-----------------------|-----------------|----------|----------|--------|---------------|-------------|-------------|-------------------|-----------|
|                       | keine Förderung | nur Land | nur Bund | nur EU | Land und Bund | Land und EU | Bund und EU | Land, Bund und EU | insgesamt |
| keine Förderung       | 0               | 427      | 1,044    | 50     | 118           | 3           | 11          | 15                | 1,668     |
| nur Land              | 561             | 127      | 74       | 0      | 81            | 1           | 0           | 6                 | 850       |
| nur Bund              | 740             | 23       | 336      | 5      | 67            | 1           | 27          | 4                 | 1,203     |
| nur EU                | 36              | 0        | 5        | 3      | 1             | 0           | 0           | 0                 | 45        |
| Land und Bund         | 112             | 36       | 78       | 1      | 152           | 1           | 1           | 15                | 396       |
| Land und EU           | 2               | 1        | 2        | 1      | 1             | 0           | 0           | 0                 | 7         |
| Bund und EU           | 7               | 2        | 8        | 1      | 2             | 0           | 9           | 0                 | 29        |
| Land, Bund und EU     | 0               | 1        | 3        | 1      | 14            | 1           | 1           | 11                | 32        |
| insgesamt             | 1,458           | 617      | 1,550    | 62     | 436           | 7           | 49          | 51                | 4,230     |

Quelle: Rohdaten, IWH-FuE-Mikrodatenbank; Berechnungen des IWH.

Der weitere Weg der Datenaufbereitung besteht nun darin, eine Verknüpfung mit den Daten von Bureau van Dijk vorzunehmen. In einem ersten Schritt konnten von den 4230 Akteuren 2532 (rund 60%) in der genannten Datenbank gefunden werden. Da unter den Fördermittelpfängern auch Akteure der Wissenschaft sind (welche in den Daten von Bureau van Dijk nicht aufgeführt werden), ist in einem zweiten Schritt die Verknüpfung mit dem Research Explorer notwendig (vgl. auch Abschnitt 3.2.5). Die dann noch verbleibenden Akteure müssten manuell nachrecherchiert werden. Stichproben unter den bislang nicht in Bureau van Dijk gefundenen Akteuren haben ergeben, dass diese im Onlineportal firmenwissen.de mit einer entsprechenden BvD-Nummer aufgeführt sind.

Das Matching mit den Bureau van Dijk Daten stellt ferner die Schlüsselbrücke her zu den Patentdaten, die zweifelsohne eine interessante Ergebnisgröße für die FuE-Förderung darstellen (vgl. Abschnitt 3.2.6).

## 5 Weitere Bearbeitungsschritte am IWH

Zukünftig wird sich das IWH verstärkt darauf konzentrieren, das bereits angesammelte Wissen zu den Programmen der FuE-Förderung in der Bundesrepublik Deutschland systematisch und umfassend zu erweitern. Dies geschieht im Wesentlichen entlang der folgenden Punkte:

- In der Endversion soll das Dokument „IWH-FuE-Mikrodatenbank“ aus drei Teilen bestehen. Der vorliegende Teil 1: **Daten, Datenherkunft und Datenqualität** wird ergänzt um zwei weitere Teile. Teil 2 wird eine **Übersicht zu den institutionellen Arrangements** der FuE-Förderpolitik enthalten, Teil 3 eine **Übersicht über Forschungsfragen und geeignete Methoden**, die mit den Daten der FuE-Projektförderung beantwortet bzw. analysiert werden können.

- Abbildung 1 verdeutlicht, dass **Förderprogramme einem Wandel unterliegen**. Beispielsweise können sich Förderquoten, der Umfang förderfähiger Kosten oder der Umfang förderfähiger Branchen ändern. Auch ist zu recherchieren, **welche FuE-Förderprogramme welche Bedeutung** (gemessen an der Anzahl der Bewilligungen sowie der bewilligten Summe) haben bzw. hatten. Diese Informationen sind notwendig, um abzuschätzen, zu welchem Grad die Datenbank die tatsächliche FuE-Projektförderung abdeckt. Entsprechende Übersichten liegen zurzeit – soweit bekannt – nicht vor und müssen umfassend erstellt werden.<sup>11</sup>
- Die Daten der IWH-FuE-Mikrodatenbank sollen zukünftig regelmäßig aktualisiert werden. Zudem wird eine Ergänzung, vor allem um weitere Landesdaten, angestrebt. Vor allem in den Neuen Ländern werden und wurden viele Instrumente der direkten FuE-Projektförderung von den Ländern eingesetzt.

---

<sup>11</sup> Frau Dr. Gisela Hillmann vom DLR ([Gisela.Hillmann@dlr.de](mailto:Gisela.Hillmann@dlr.de)) war mit dem Aufbau der Datenbank „PROMO“ beschäftigt, die sämtliche in Deutschland praktizierte Förderung (Bund, Länder, EU) erfassen sollte. Mit Frau Hillmann gab es am IWH bereits Arbeitstreffen, der Kontakt besteht also bereits. Soweit bekannt, liegen Angaben zur betrieblichen Förderung für die Förderperiode 2007-2013 nahezu vollständig vor. Für die Perioden davor dürften größere Lücken vorhanden sein, da viele Informationen nicht mehr verfügbar waren.

## Literatur

- Bade, F.-J.; Bornemann, H.; Breuer, A.; Rautenberg, R.:* Ansätze für ein besseres Monitoring und eine verbesserte Erfolgskontrolle der Infrastrukturförderung innerhalb der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“. Endbericht. Gutachten im Auftrag des Thüringer Ministerium für Wirtschaft, Arbeit und Technologie. Berlin, Bremen, Dortmund 2012.
- Bade, F.-J.; Alm, B.:* Evaluierung der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) durch einzelbetriebliche Erfolgskontrolle für den Förderzeitraum 1999-2008 und Schaffung eines Systems für ein gleitendes Monitoring. Dortmund 2010.
- Christen, P.:* Data Matching – Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. Springer: Berlin, Heidelberg 2012.
- Herzog, T. N.; Scheuren, F. J.; Winkler, W. E.:* Data Quality and Record Linkage Techniques. Springer Science + Business: New York 2007.
- Institut für Wirtschaftsforschung Halle:* Möglichkeiten einer wissenschaftlichen Evaluation der gewerblichen Förderung im Rahmen der Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ (GRW) in Sachsen-Anhalt, Stand 19.02.2014, mimeo.
- Magerman, T.; Van Looy, B.; Song, X.:* Data production methods for harmonized patent statistics: Patentee name harmonization. 2006
- OECD:* Technology and industrial performance, OECD: Paris 1996.
- Papaconstantinou, G.; Sakurai, N.; Wyckoff, A.:* Domestic and international product-embodied R&D diffusion, in: Research Policy, 27, 1996, 301-314.
- Schnell, R.; Bachteler, T.; Reiher, J.:* MTB: Ein Record-Linkage-Programm für die empirische Sozialforschung. 2005.
- Titze, M.; Ehrenfeld, W.; Piontek, M.; Pippel, G.:* Netzwerke zwischen Hochschulen und Wirtschaft: Ein Mehrebenenansatz, in: Fritsch, Michael; Pasternack, Peer; Titze, Mirko (Hrsg.): Schrumpfende Regionen – dynamische Hochschulen. Hochschulstrategien im demografischen Wandel. Springer Fachmedien, Wiesbaden 2015, 213-235.



Leibniz-Institut für  
Wirtschaftsforschung Halle (IWH)

Kleine Märkerstraße 8  
D-06108 Halle (Saale)

Postfach 11 03 61  
D-06017 Halle (Saale)

Tel +49 345 7753 60  
Fax +49 345 7753 820  
[www.iwh-halle.de](http://www.iwh-halle.de)

ISSN: 2365-9076