A 2 Kausalanalyse von Maßnahmeneffekten

iele der im Auftrag der Bundesregierung durchgeführten Evaluationsstudien zu Maßnahmen der Forschungs- und Innovationspolitik (F&I-Politik) lassen keine Rückschlüsse darauf zu, ob die beobachteten Entwicklungen tatsächlich auf die untersuchten Politikmaßnahmen zurückzuführen sind. Wesentlicher Grund hierfür ist, dass Evaluationsstudien häufig nicht den methodischen Anforderungen an eine Kausalanalyse genügen, nicht zuletzt, weil die Voraussetzungen für den sachgerechten Einsatz geeigneter Methoden nicht in jedem Fall erfüllt sind. Das fehlende Wissen über die Wirkung von Maßnahmen erschwert ein systematisches und evidenzbasiertes Politiklernen. Die Möglichkeiten, aus der Bewertung von und Erfahrung mit durchgeführten Maßnahmen zu lernen, um gegebenenfalls zukünftige Maßnahmen so anzupassen, dass deren Ziele besser erreicht werden, bleiben im Wesentlichen ungenutzt.

Die Bundesregierung ist daher aufgefordert, Kausalanalysen der Maßnahmeneffekte systematisch und umfassend in den Einsatz von Maßnahmen der F&I-Politik zu integrieren und so die Voraussetzungen für eine sach- und fachgerechte Durchführung dieser Analysen sowie für deren Nutzbarkeit beim Politiklernen zu gewährleisten. Hierzu ist es erforderlich, bei der Ausschreibung von Evaluationsstudien Kausalanalysen in die Leistungsbeschreibung aufzunehmen, die Datenverfügbarkeit für die evaluierenden Einrichtungen zu verbessern und alle in Auftrag gegebenen Evaluationsstudien zu veröffentlichen.

Kausalanalysen für lernende Forschungs- und Innovationspolitik erforderlich

Kausalanalysen der Effekte von Politikmaßnahmen sind notwendig, um die Effektivität und Effizienz dieser Maßnahmen bewerten zu können und dadurch Politiklernen zu ermöglichen. Politikmaßnahmen sind effektiv, wenn sie die verfolgten Ziele erreichen, und effizient, wenn es keine Möglichkeiten gibt, die verfolgten Ziele kostengünstiger zu realisieren. Politiklernen impliziert, dass Maßnahmen auf besserer Wissensbasis unverändert fortgeführt, nachjustiert oder gegebenenfalls aufgegeben werden.

Die Expertenkommission schätzt die Potenziale des Politiklernens in der F&I-Politik aus folgenden Gründen als hoch ein:

- Maßnahmen der F&I-Politik zielen auf ein breites Spektrum verschiedener F&I-Aktivitäten ab – von der Grundlagenforschung über die angewandte Forschung bis zur Anwendung. Die verschiedenen Arten von F&I-Aktivitäten unterscheiden sich hinsichtlich ihrer Ansatzpunkte sowie ihrer förderpolitischen Relevanz. Ihre Ergebnisse sind mit einem unterschiedlich hohen und sich zeitlich verändernden Grad der Unsicherheit verbunden.
- Inwieweit eine Maßnahme der F&I-Politik geeignet ist, bestimmte F&I-Aktivitäten zu unterstützen, hängt stark vom jeweiligen Kontext ab. So kann die Wirksamkeit einer Maßnahme je nach Branche oder Region ganz unterschiedlich ausfallen.
- Die F&I-Politik zielt zunehmend darauf ab, F&I-Aktivitäten in völlig neue Richtungen zu lenken und damit einen Beitrag zur Bewältigung der großen gesellschaftlichen Herausforderungen zu leisten. Dabei kann die F&I-Politik häufig nicht einfach auf etablierte Maßnahmen zurückgreifen, sondern muss diese anpassen oder gänzlich neue Wege beschreiten.

Vor diesem Hintergrund können Kausalanalysen der Maßnahmeneffekte erheblich dazu beitragen, Maßnahmen der F&I-Politik anzupassen, weiterzuentwickeln und somit deren Wirkung zu verbessern. Kausalanalysen sind Teil von Ex-post-Evaluationen, die Maßnahmen im Nachhinein untersuchen und bewerten. Neben solchen Ex-post-Evaluationen

sind häufig auch begleitende Evaluationen erforderlich, um bereits während der Laufzeit von Maßnahmen Erkenntnisse über deren Umsetzbarkeit zu gewinnen und gegebenenfalls frühzeitig nachsteuern zu können. Dies gilt insbesondere im Rahmen einer missionsorientierten und transformativen F&I-Politik, in der zahlreiche Maßnahmen parallel oder sequenziell zum Einsatz kommen, deren Zusammenwirken ex ante nur unvollständig abgeschätzt werden kann.

Aussagekraft der Wirkungsevaluationen von Methodik abhängig

Das Ziel der Kausalanalysen von Maßnahmeneffekten ist es festzustellen, ob beobachtete Veränderungen in vorab definierten Zielgrößen kausal, d. h. ursächlich, auf die untersuchte Politikmaßnahme zurückzuführen sind. Die Aussagekraft solcher Kausalanalysen und damit die Möglichkeiten des Politiklernens hängen von der verwendeten Methodik ab. Diese hat sich in den letzten Jahrzehnten grundlegend weiterentwickelt. Experimentelle und quasi-experimentelle Verfahren der Kausalanalyse sind heute fest etablierter Standard in den empirischen Wirtschafts- und Sozialwissenschaften. ⁹³

Die Expertenkommission hat bereits in ihrem Jahresgutachten 2010 auf die Bedeutung aussagekräftiger Evaluationsstudien hingewiesen und seitdem immer wieder die Notwendigkeit betont, Kausalanalysen von Maßnahmeneffekten vorzunehmen. Harüber hinaus hat 2013 der wissenschaftliche Beirat des damaligen BMWi in einem seiner Gutachten die Bedeutung von Kausalanalysen für die Evaluation wirtschaftspolitischer Maßnahmen betont. Die von ihm formulierten methodischen Anforderungen sind weiterhin aktuell – auch für Maßnahmen der F&I-Politik. PS

Um aussagekräftige Kausalanalysen durchführen zu können, müssen die Voraussetzungen hierfür bereits bei der Konzeption von Politikmaßnahmen berücksichtigt und mit der Umsetzung der Maßnahmen geschaffen werden. Wesentlich dafür ist es, die Situation im Fall der durchgeführten Maßnahme mit der Situation zu vergleichen, die sich ohne Maßnahme ergeben hätte, die sogenannte kontrafaktische Situation (vgl. Box A 2-1). Dieser Vergleich wird über ein geeignetes Evaluationsdesign hergestellt. Zentrale Anforderungen sind, dass die von einer Maßnahme Erfassten, die Treatment-Gruppe,

mit einer geeigneten Kontrollgruppe Nichterfasster (vgl. Box A 2-1) verglichen werden und die Wirksamkeit der Maßnahme mit geeigneten experimentellen oder quasi-experimentellen Verfahren überprüft wird (vgl. Box A 2-1).

Bestmögliche Voraussetzungen für eine geeignete Zusammensetzung von Treatment- und Kontrollgruppe bietet eine echte Randomisierung der Maßnahmenteilnahme im Rahmen von Politikexperimenten (vgl. Box A 2-1). Für Politikexperimente gibt es international zahlreiche Vorbilder. 96 Echte Randomisierung bedeutet etwa bei einer Fördermaßnahme, dass die Auswahl der Geförderten zufällig erfolgt. Bei ausreichend großen Fallzahlen kann durch Randomisierung sichergestellt werden, dass sich die Gruppe der Geförderten und die Kontrollgruppe der Nichtgeförderten in allen beobachtbaren und nicht beobachtbaren Merkmalen mit Ausnahme des Merkmals "gefördert" bzw. "nicht gefördert" statistisch nicht unterscheiden. 97 Die für eine Kausalanalyse so wichtige kontrafaktische Situation wird auf diese Weise von vornherein in die Politikmaßnahme integriert.

Randomisierung kann häufig mit einer Vorauswahl nach inhaltlichen, mit der Maßnahme verbundenen förderrelevanten Kriterien kombiniert werden. So kann etwa zunächst die grundsätzliche Förderfähigkeit geprüft werden. Eine zufällige Auswahl erfolgt nur innerhalb der Gruppe der Förderfähigen, ggf. auch nur innerhalb einer Teilgruppe. 99

Randomisierte Politikexperimente sind vielfach nicht durchführbar, etwa weil die notwendigen Fallzahlen nicht erreicht werden. Ist dies absehbar, dann sollten von vornherein quasi-experimentelle Verfahren der Kausalanalyse in Betracht gezogen werden (vgl. Box A 2-1). Deren Anwendbarkeit und Aussagekraft kann durch ein geeignetes Design der Maßnahme und durch die Erhebung der erforderlichen Daten verbessert werden.

Methodisches Potenzial in Evaluationspraxis nicht ausgeschöpft

Die Bundesregierung hat in den vergangenen Jahren eine Vielzahl von Studien zur Evaluation von Maßnahmen der F&I-Politik in Auftrag gegeben. ¹⁰⁰ Die Expertenkommission ist der Frage nachgegangen, wie viele dieser Studien den Anspruch erheben, kausale Effekte der untersuchten Maßnahmen nachzu-

Box A 2-1 Erläuterung kausalanalytischer Grundbegriffe

Kausalanalyse: Unter Kausalanalyse werden statistische Verfahren zur Ermittlung von Ursache-Wirkungs-Beziehungen zwischen verschiedenen Größen verstanden. In der Politikanalyse wird die Kausalanalyse verwendet, um festzustellen, ob und in welchem Ausmaß eine bestimmte Intervention wie etwa eine Maßnahme der F&I-Politik (die unabhängige Größe) einen direkten Effekt auf das gewünschte Ergebnis (die abhängige Größe) hat.

Kontrafaktischer Zustand: Der kontrafaktische Zustand bezeichnet die hypothetische Situation, in der sich eine von einer Intervention (dem sogenannten Treatment) betroffene Einheit (z.B. eine Person, eine Gruppe, ein Unternehmen, eine Region) befände, wenn sie nicht der Intervention ausgesetzt gewesen wäre.

Randomisierung: Randomisierung bedeutet, dass die Maßnahmenteilnehmerinnen und -teilnehmer zufällig zwei Gruppen zugeordnet werden:

- Treatment-Gruppe: Diese Gruppe unterliegt der Intervention, deren Effekt untersucht werden soll.
- Kontrollgruppe: Diese Gruppe unterliegt entweder keiner Intervention oder einer anderen Intervention als der untersuchten. Aus den Ergebnissen der Kontrollgruppe wird auf den kontrafaktischen Zustand der Treatment-Gruppe geschlossen, um die Wirkung der Intervention zu messen bzw. die Wirkung der verschiedenen Interventionen zu vergleichen.

Die Randomisierung soll sicherstellen, dass es vor der Intervention keine systematischen Unterschiede zwischen den Gruppen gibt. Das bedeutet, dass alle bekannten und unbekannten Größen, die das Ergebnis beeinflussen könnten, gleichmäßig über beide Gruppen verteilt sind.

Experimentelle Verfahren: Randomisierte Experimente, oft auch als randomisierte kontrollierte Studien (Randomized Controlled Trials, RCTs) bezeichnet, werden in der Forschung verwendet, um die Wirksamkeit einer Intervention unter kontrollierten Umweltbedingungen zu testen. In geeigneter Weise konzipiert und durchgeführt, liefern randomisierte Experimente die stärkste Evidenz für kausale Effekte einer Intervention.

Quasi-experimentelle Verfahren: Quasi-experimentelle Verfahren basieren im Gegensatz zu experimentellen Verfahren nicht auf einer echten Randomisierung. Stattdessen werden Konstellationen untersucht, in denen die Teilnahme an einer Intervention "quasi" zufällig erfolgte und für die Beteiligten kaum zu beeinflussen war. Beispielsweise ergibt sich bei Interventionen, die auf einzelne Regionen begrenzt waren, die Möglichkeit der Teilnahme quasi-zufällig durch die Zugehörigkeit zur Region. Daher können vergleichbare Einheiten aus anderen Regionen als Kontrollgruppe eingesetzt werden. Unter bestimmten Bedingungen, die sich je nach Verfahren unterscheiden, können auf diese Weise kausale Effekte von Interventionen geschätzt werden. 101

Matching-Verfahren: Matching-Verfahren¹⁰² sind statistische Techniken, bei denen für jede Beobachtungseinheit in der Treatment-Gruppe eine oder mehrere Einheiten in der Kontrollgruppe gefunden werden, die in Bezug auf bestimmte beobachtete Größen ähnlich sind. Matching-Verfahren können dazu beitragen, Verzerrungen zu reduzieren, die durch die nicht-zufällige Teilnahme an Maßnahmen entstehen. Im Rahmen von Kausalanalysen werden sie in Ergänzung zu quasiexperimentellen Verfahren eingesetzt. Ein alleiniges Matching löst nicht das Problem, dass die Maßnahmenteilnahme durch nicht-beobachtbare Größen beeinflusst sein kann.

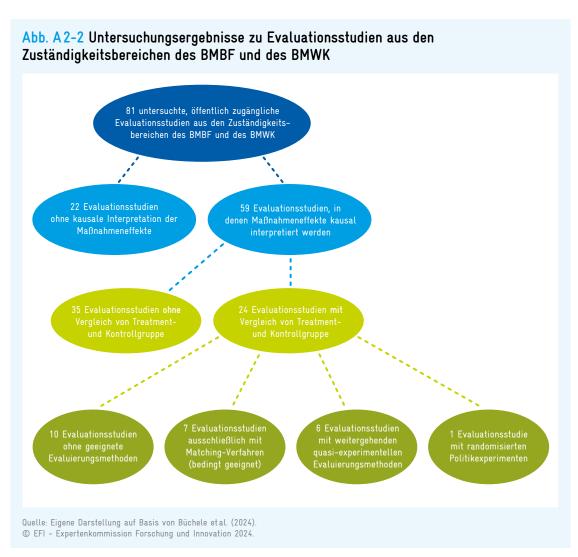
weisen und inwieweit dieser Anspruch eingelöst wird. Dazu wurden 81 zwischen 2009 und 2023 verfasste und öffentlich zugängliche Evaluationsstudien aus den Zuständigkeitsbereichen von BMBF und BMWK nach einem einheitlichen Analyseschema untersucht. Die Ergebnisse dieser Untersuchung zeigen, dass eine belastbare Kausalanalyse auch in denjenigen Evaluationsstudien, die Aussa-

gen zur Wirkung von Maßnahmen treffen, in der Vergangenheit häufig nicht durchgeführt wurde. Zudem deuten die Ergebnisse darauf hin, dass die Voraussetzungen, etwa hinsichtlich der Datenverfügbarkeit, für Kausalanalysen im Rahmen der Evaluation von Politikmaßnahmen verbessert werden müssen.¹⁰⁴

In 59 der 81 untersuchten Evaluationsstudien wurden die Ergebnisse kausal interpretiert - unabhängig von der angewandten Methode (vgl. Abbildung A 2-2). In 24 dieser 59 Studien wurden Treatmentund Kontrollgruppen verglichen. Nur in einer dieser Studien wurden randomisierte Politikexperimente durchgeführt. In sechs Studien kamen quasi-experimentelle Methoden zum Einsatz. In sieben Studien wurden ausschließlich Matching-Verfahren (vgl. Box A 2-1) angewandt, die nur bedingt für eine Kausalanalyse geeignet sind. In den übrigen zehn der 24 Studien wurden keine Methoden verwendet, die eine Kausalanalyse der Maßnahmeneffekte erlauben. Das heißt, nur in sieben der 81 untersuchten Evaluationsstudien konnten belastbare Aussagen darüber getroffen werden, ob die evaluierten Maßnahmen wirksam waren. In 45 der 81 Evaluationsstudien wurden Entwicklungen jedoch als kausale Maßnahmenwirkungen interpretiert, obwohl die verwendete Methodik diese Schlussfolgerung nicht zulässt.

Insgesamt 40 der 59 Studien mit kausaler Ergebnisinterpretation attestieren den evaluierten Maßnahmen der F&I-Politik eine positive und 19 der Studien eine zumindest teilweise positive Wirkung. Im Ergebnis finden alle betrachteten Studien mit einer Wirkungsaussage auch eine positive Wirkung. Keine Studie kommt zu dem Schluss, dass Wirkungen der betrachteten Maßnahmen nicht nachweisbar sind, oder findet gar negative Wirkungen. Dieses Muster unterscheidet sich deutlich von den Ergebnissen der in der wissenschaftlichen Literatur zu findenden Wirkungsanalysen. Diese kommen häufig zu dem Ergebnis, dass eine Wirksamkeit der untersuchten Maßnahmen nicht nachgewiesen werden kann. 106

Die Expertenkommission stellt fest, dass die Mehrzahl der betrachteten Evaluationsstudien den Anforderungen an eine aussagekräftige Kausalanalyse nicht genügt. Randomisierte Politikexperimente wurden in der Vergangenheit so gut wie nie durch-



Download der Abbildung und Daten geführt. Auch Verfahren der quasi-experimentellen Kausalanalyse, die über ein bloßes Matching hinausgehen, wurden – auch wenn hier ein leicht positiver Trend zu beobachten ist – nur in Einzelfällen eingesetzt. ¹⁰⁷ Damit fehlt der Politik die Grundlage, um aus der (Nicht-)Wirksamkeit früherer Maßnahmen zu lernen und die Wirksamkeit zukünftiger Maßnahmen der F&I-Politik zu erhöhen.

Gründe für mangelnde Aussagekraft von Evaluationsstudien vielfältig

Es gibt verschiedene Gründe dafür, dass das Forschungsdesign einer Evaluationsstudie den Anforderungen an eine aussagekräftige Wirkungsmessung durch Kausalanalyse nicht oder nicht vollständig genügt. In 32 der 81 von der Expertenkommission betrachteten Evaluationsstudien werden solche Gründe angeführt. Zu diesen Gründen zählen:

- Unzureichende Datenverfügbarkeit und -qualität erschweren oder verhindern die Identifikation geeigneter Kontrollgruppen und die Durchführung aussagekräftiger Kausalanalysen. Nicht in allen Fällen stehen Daten, die zu Beginn der Maßnahme prinzipiell verfügbar wären, für eine Kausalanalyse zur Verfügung.
- Kausalanalysen können an methodische Grenzen stoßen. So ist ein Vergleich von Treatment- und Kontrollgruppe bei zu geringen Fallzahlen nicht aussagekräftig. Zudem ist die Identifikation einer Kontrollgruppe gar nicht möglich, wenn alle potenziellen Adressatinnen und Adressaten einer Fördermaßnahme gefördert werden. Auch können Maßnahmeneffekte nicht zeitnah identifiziert werden, wenn Effekte der zu evaluierenden Politikmaßnahme erst langfristig zu erwarten sind.
- Die Effektivität von Maßnahmen wird anhand definierter Zielgrößen bewertet. Voraussetzung dafür ist, dass sich diese Zielgrößen messen lassen oder man sich ihnen zumindest annähern kann. Auch nicht-intendierte Effekte positive wie negative sind nicht immer messbar oder können übersehen werden.

Die Bewertung der Effizienz einer Maßnahme ist mit besonders großen Anforderungen an die Daten verbunden, da hierbei in der Regel erforderlich ist, dass die Effekte verschiedener Maßnahmen gemessen und miteinander verglichen werden. Auf Basis einer Kausalanalyse lässt sich jedoch in vielen Fällen zumindest eine Kosten-Nutzen-Abschätzung vornehmen.

Handlungsempfehlungen

Die Bundesregierung verfolgt im Bereich der F&I-Politik ambitionierte Ziele und unterliegt gleichzeitig zunehmenden Sparzwängen. Eine aussagekräftige Kausalanalyse der Maßnahmeneffekte als Teil der Evaluation von Politikmaßnahmen schafft die Voraussetzungen für eine lernende F&I-Politik. Die Expertenkommission spricht sich deshalb dafür aus, die Kausalanalyse von Maßnahmen der F&I-Politik als Grundlage einer lernenden F&I-Politik zu stärken, und knüpft dabei an Handlungsempfehlungen früherer Jahresgutachten an.

Kausalanalysen systematisch in Maßnahmen der F&I-Politik integrieren

- Schon bei der Planung von Maßnahmen sind Bedingungen für eine Kausalanalyse der Maßnahmeneffekte mitzudenken und die Voraussetzungen für die Erhebung der erforderlichen Daten zu schaffen. Hierfür sind konkrete Ziele festzulegen und nach Möglichkeit geeignete Indikatoren zur Überprüfung der Zielerreichung zu benennen. 108 Bereits in dieser Phase sollten in der Methodik von Kausalanalysen ausgewiesene interne und externe Expertinnen und Experten einbezogen werden.
- Bei der Ausschreibung von Evaluationsstudien sollte die Leistungsbeschreibung eine Kausalanalyse sowie diesbezügliche Mindestanforderungen an das Evaluierungskonzept enthalten. Dabei sollte auch eine grundsätzliche Bewertung der Effizienz der Maßnahme, z. B. durch eine Kosten-Nutzen-Analyse, angestrebt werden. Der Zeitpunkt der Evaluation ist so zu wählen, dass die relevanten Effekte der jeweiligen Maßnahme realistischerweise beobachtet werden können.

Bei der Vergabe von Evaluationsstudien ist insbesondere auf die methodische Kompetenz der Anbieter zu achten. Evaluationen sollten nicht von Organisationen durchgeführt werden, die selbst Träger der zu untersuchenden Maßnahme sind. Durch geeignete Ausschreibungsmodalitäten sollten Anreize für neue qualifizierte Anbieter geschaffen werden.¹⁰⁹

Kausalanalysen bestmöglich auf Evaluationskontext abstimmen

- Randomisierte Politikexperimente bieten ideale Voraussetzungen für aussagekräftige Kausalanalysen. Sie sollten stärker als bisher in der F&I-Politik eingesetzt werden. Eine Randomisierung bietet sich insbesondere dann an, wenn mit einer Überzeichnung von Fördermaßnahmen zu rechnen ist.
- Reallabore stellen eine besondere Kategorie von F&I-politischen Maßnahmen dar, die Politiklernen ermöglichen sollen. Bei der Einrichtung von Reallaboren sollte daher auch die spätere Kausalanalyse mitgedacht werden, z. B. indem sichergestellt wird, dass es aussagekräftige Vergleichseinheiten zum Reallabor gibt, um so dessen Gesamtwirkung abschätzen zu können.
- Um das Zusammenwirken verschiedener Maßnahmen im Kontext missionsorientierter F&I-Politik zu analysieren, kann es sinnvoll sein, die Erreichung der Missionsziele durch das Zusammenwirken der Maßnahmen aggregiert ggf. auch im internationalen Vergleich zu untersuchen.
- Der Verzicht auf eine Kausalanalyse der Maßnahmeneffekte kann im Einzelfall begründet sein, nicht zuletzt, weil die Voraussetzungen hierfür nicht in jedem Fall geschaffen werden können. Ein solcher Verzicht muss jedoch mit den Besonderheiten des jeweiligen Maßnahmenkontextes begründet werden.
- Wo die Voraussetzungen belastbarer Kausalanalysen nicht erfüllt sind, kann die Verwendung einfacherer Verfahren wie etwa Matching-Verfahren Hinweise auf die Effektivität von Maßnahmen geben. Jedoch sollten in solchen Fällen keine Aussagen über die Maßnah-

menwirksamkeit getroffen werden, die nicht durch die verwendeten Verfahren gedeckt sind.

Know-how erweitern und Transparenz schaffen

- Zentrale Voraussetzung für Politiklernen mittels Kausalanalysen von Maßnahmeneffekten ist, dass sie ergebnisoffen durchgeführt und positive wie negative Ergebnisse gleichermaßen als Erkenntnisfortschritt geschätzt werden.¹¹¹
- Unabhängig von den Ergebnissen sollten alle von den Ministerien beauftragten Analysen von Maßnahmeneffekten veröffentlicht werden. Dabei sind auch die verwendeten Daten und Programmiercodes offenzulegen oder interessierten Forscherinnen und Forschern auf Anfrage zugänglich zu machen. Hierdurch wird eine Replikation der Ergebnisse, ggf. auch mit später entwickelten verbesserten Analysemethoden, ermöglicht.
- Die institutionelle Verankerung der Evaluationspraxis in den Ministerien sollte weiter gestärkt und dabei ein besonderes Augenmerk auf die Weiterbildung der Beschäftigten in den Fachreferaten gelegt werden.¹¹²

Verfügbarkeit und Qualität von Daten erhöhen

- Die für die Durchführung einer Maßnahme verantwortlichen Referate oder Projektträger sollten den Evaluationsteams alle mit einer Maßnahme verbundenen Dokumente systematisch zur Verfügung stellen. Hierzu zählen etwa Antragsskizzen, Informationen zu allen und nicht nur zu den letztendlich geförderten Antragstellern sowie die Vorbereitungsdokumente und Protokolle der Auswahl- und Fördergremien.
- Die Datenlabore in den Ministerien sollten zu Forschungsdatenzentren weiterentwickelt werden, in denen die Daten aller Fördermaßnahmen der F&I-Politik gesammelt, aufbereitet und für die Forschung vorgehalten werden. Ressortübergreifend sollten einheitliche Definitionen und, soweit möglich, Indikatoren entwickelt und genutzt werden, um die Vergleichbarkeit der Daten über verschiedene Maßnahmen hinweg zu gewährleisten.

 Für Evaluationsstudien sind Daten, die ohnehin administrativ erhoben werden, oft von hoher Bedeutung. Der Zugang zu diesen Daten zu Zwecken der Evaluation von Politikmaßnahmen sollte niedrigschwellig geregelt und umfassend, d.h. auch auf Ebenen geringerer Datenaggregation, ermöglicht werden.¹¹⁴