

die Anbieter auf einer Plattform, zumeist aber nicht die Plattform selbst. Bisher wird diese Finanzierungsgrenze im deutschen *Crowdfunding*-Geschäft jedoch nur selten erreicht.

Auf der EU-Ebene sind bereits einige Direktiven auf den Weg gebracht oder in Planung, u. a. die Prospekttrichtlinie¹²⁹ oder auch die AIFM-Richtlinie. Bei deren nationaler Umsetzung ist darauf zu achten, dass sie die Entstehung einer *Crowdfunding*-Industrie in Deutschland insgesamt fördern und nicht einseitig auf den Anlegerschutz abzielen. Gleichzeitig ist es schwierig, mit den jeweiligen Direktiven allen unterschiedlichen *Crowdfunding*-Formen gleichermaßen gerecht zu werden. Auch eine explizite *Crowdfunding*-Direktive, ähnlich den Regelungen zu *funding portals* im Rahmen des *Jumpstart Our Business Start-ups (JOBS) Acts* 2012 in den USA, die eine rechtliche Grundlage für *Crowdfunding* in den Vereinigten Staaten geschaffen haben, wäre denkbar.¹³⁰ Bestandteil dieser Direktive ist unter anderem die Anhebung der Obergrenze der *Crowdfunding*-Finanzierung auf 1 Million US-Dollar, einem Vielfachen des Wertes, bei dem die Prospektspflicht in Europa einsetzt (100.000 Euro, s.o.).

Empfehlungen

- *Crowdfunding* bietet für *Start-ups* und KMU eine interessante Alternative oder Ergänzung zur Finanzierung durch Staat, Banken, Wagniskapitalgeber oder *Business Angels* und verfügt somit über das Potenzial, Innovationen zu fördern. Als sehr junges Phänomen ist *Crowdfunding* bisher allerdings kaum erforscht. Offene Fragen sind beispielsweise, wie die bisherige staatliche Förderung von *Start-ups* und KMU mit dieser neuen Finanzierungsform kombiniert werden kann und wie sich *Crowdfunding* auf die Akquise von Folgefinanzierung auswirkt. Es ist im Interesse aller Beteiligten, eine entsprechende Datengrundlage zu schaffen, um Transparenz bezüglich der Vor- und Nachteile von *Crowdfunding* zu ermöglichen.
- Wie das Beispiel anderer europäischer Länder zeigt, kann bei der Unternehmensgründung die Wahl der Rechtsform eine Eigenkapitalfinanzierung durch *Crowdfunding* behindern. So erfordern gerade *Crowdfunding*-Aktivitäten von Gründungen auf Basis eher privat ausgestalteter Rechtsformen in anderen europäischen Ländern oft notarielle

Formalien oder verbieten weitgestreutes Eigenkapital. Die bestehenden deutschen Regelungen zur stillen Beteiligung erweisen sich demgegenüber eher als Standortvorteil und werden daher aus Sicht der Expertenkommission positiv bewertet.

- Die neuen Regelungen des US-amerikanischen *JOBS Acts* haben die Rahmenbedingungen für *Crowdfunding*-Aktivitäten und damit für innovative KMU und Gründungen in den USA deutlich verbessert. Will Europa nicht hinter die USA zurückfallen, ist es umso wichtiger, auch auf europäischer Ebene eine Harmonisierung der Regulierung oder zumindest eine Abstimmung der *Crowdfunding*-relevanten, nationalen Regelungen anzustreben. Unter anderem betrifft dies die Grenzen der Prospektspflicht.
- Es ist zu klären, wie Anlegerschutz einerseits sowie der Schutz der Plattformbetreiber vor betrügerischen Anlegern andererseits zu bewerkstelligen ist, ohne dass staatliche Eingriffe die Wachstumspotenziale von *Crowdfunding*-Märkten in Deutschland und Europa behindern.¹³¹ Eine Stärkung des Anlegerschutzes könnte beispielsweise durch die Einführung einer Obergrenze für zulässige Investitionsbeträge einzelner, privater Geldgeber oder durch die explizite Beteiligung erfahrener und akkreditierter Anleger an einem Finanzierungsprojekt erreicht werden.¹³²

EVALUATION INNOVATIONSPOLITISCHER MASSNAHMEN MIT HILFE RANDOMISierter EXPERIMENTE

A 6

Problemstellung

Zur Steigerung der FuE-Dynamik gibt es in Deutschland – genau wie in vielen anderen Ländern – eine Vielzahl an FuE-Fördermaßnahmen. Allerdings fehlt es bislang meist an einer systematischen und neuesten wissenschaftlichen Ansprüchen genügenden Wirkungsanalyse solcher Fördermaßnahmen. Es gibt noch zu wenig gesichertes Wissen über die tatsächliche Wirkung der unterschiedlichen Fördermaßnahmen und insbesondere über das Ausmaß von Mitnahmeeffekten im Vergleich zu ursächlichen Innovationssteigerungen, die auf diese Fördermaßnahmen zurückzuführen sind. Knappe staatliche Fördermittel sollten aber effizient und effektiv eingesetzt werden, so dass wissenschaftlich fundierte Untersuchungen der

Ursache-Wirkungsbeziehungen bei den staatlichen FuE-Fördermaßnahmen dringend geboten erscheinen. Obwohl in den vergangenen Jahren wissenschaftlich fundierte Analysen bereits zu einer verbesserten Evaluierung der FuE-Förderung beigetragen haben, kann von einer systematischen Wirkungsanalyse in der Innovationsforschung nach wie vor nicht die Rede sein. Und dies, obwohl in anderen Politikbereichen, wie zum Beispiel der Arbeitsmarktforschung, in den letzten Dekaden bedeutende Fortschritte in der Analyse der Wirksamkeit staatlicher Förderung verbucht wurden und große Ineffizienzen aufgedeckt und behoben werden konnten.¹³³ So gab es von 2000 bis 2005 für ca. 45 Milliarden Euro Maßnahmen der aktiven Arbeitsmarktpolitik,¹³⁴ von denen viele als ineffizient evaluiert und folgerichtig mit den Reformen im Jahr 2009 wieder zurückgenommen wurden. Die Expertenkommission empfiehlt vor diesem Hintergrund, auch in der staatlichen Innovationspolitik in Zukunft stärker auf eine evidenzbasierte Förderpolitik zu setzen und systematisch auf eine empirische Wirkungsforschung zu bauen.

Ein generelles Problem solcher Wirkungsanalysen ist, sei es in der Arbeitsmarkt-, der Gesundheits- oder der Innovationsförderung, dass nach Einführung einer Maßnahme die sogenannte kontrafaktische Situation nicht beobachtbar ist. Das heißt, man kann nie beobachten, wie sich die Geförderten verhalten hätten, wenn sie nicht gefördert worden wären. Weil sie gefördert wurden, ist es faktisch ausgeschlossen, dass man sie, die Förderempfänger, in ihrem nicht-geförderten Zustand als Vergleichsmaßstab heranzieht, um die Wirkung der Fördermaßnahme zu beurteilen. So kann man beispielsweise in der Innovationsförderung nicht beobachten, wie sich etwa die Patentierungsaktivitäten bei den geförderten Unternehmen im Falle einer Nicht-Förderung entwickelt hätten. Die Förderwirkung kann in diesem Fall also nur analysiert werden, indem man die Patente der geförderten Unternehmen (*Treatment*-Gruppe) anderen Unternehmen, nämlich denen einer geeigneten Vergleichsgruppe, die keine Förderung bekommen hat (der Kontrollgruppe), gegenüberstellt (ähnlich wie in experimentellen Designs bei der Einführung neuer Medikamente, vgl. Box 6).

Im Innovationsbereich ist die Schwierigkeit von Wirkungsanalysen mithilfe eines Vergleichs von *Treatment*- und Kontrollgruppen, exakt solche nicht-geförderten Kontrollgruppen-Unternehmen zu

identifizieren, die sich in keiner anderen wichtigen Variable von den geförderten Unternehmen unterscheiden als der Fördermaßnahme. Für beobachtbare Charakteristika wie Firmengröße, Alter, Branche mag dies noch vergleichsweise einfach sein, schwierig wird es bei nicht-beobachtbaren Charakteristika, die aber für die Wirkungsweise einer FuE-Fördermaßnahme eine große Bedeutung haben können, wie etwa Führungskompetenzen des FuE-Managements, Fähigkeiten und Motivation des FuE-Personals oder organisatorische Lernfähigkeit und Flexibilität eines Unternehmens. Wenn jedoch die Unternehmen in solchen wichtigen, aber nicht-beobachtbaren Variablen nicht vollkommen identisch sind, dann kann nicht mehr festgestellt werden, ob ein höherer Innovationsoutput eines geförderten Unternehmens im Vergleich zu einem nicht-geförderten Unternehmen auf die Wirkung der Fördermaßnahme zurückgeht. Es besteht die Möglichkeit, dass der höhere Innovationsoutput auf Unterschieden bei solchen nicht-beobachtbaren Charakteristika beruht und demnach mit der Maßnahme gar nichts zu tun hat. Fänden sich beispielsweise in der geförderten Unternehmensgruppe vorwiegend Unternehmen mit besser geführtem und motiviertem FuE-Personal, dann könnten Unterschiede im Innovationsoutput sogar alleine auf solche weichen Faktoren zurückgehen, ohne dass die FuE-Fördermaßnahme überhaupt eine Wirkung entfaltet hätte; sie würde gegebenenfalls einfach nur mitgenommen, weil besser geführtes und motiviertes FuE-Personal auch besser im Ausschöpfen von Fördermöglichkeiten ist. Als Konsequenz wäre die Wirkung der Fördermaßnahme bei den faktisch nicht-geförderten Unternehmen vollkommen anders als bei den geförderten Unternehmen: Wenn die Fördermaßnahme auf weitere Unternehmen ausgedehnt würde, wäre die Wirkung im einfachsten Fall gleich Null, im schlechtesten Fall sogar negativ. Die Identifikation einer geeigneten Kontrollgruppe ist also essenziell, um Fördermaßnahmen einer aussagekräftigen Evaluation unterziehen zu können und Rückschlüsse bezüglich der Wirkung einer Maßnahme zu erhalten.

Evaluationen mithilfe randomisierter Einteilung in Kontroll- und *Treatment*-Gruppen

Aus Sicht der Expertenkommission ist es bei manchen FuE-Fördermaßnahmen eine besonders vielversprechende Möglichkeit, geeignete Kontrollgruppen für eine qualitativ hochwertige Evaluation mithilfe

BOX 06

Randomisierte Evaluationen am Beispiel Medikamentenentwicklung

Evaluationen, die auf einer Zuteilung nach dem Zufallsprinzip („Randomisierung“) basieren, lassen sich mit einem medizinischen Test zur Wirksamkeit eines Medikaments vergleichen: Es gibt auf der einen Seite eine Patientengruppe, die ein Medikament erhält (*Treatment*-Gruppe), und auf der anderen Seite eine zweite Patientengruppe, die kein Medikament oder lediglich ein Placebo verabreicht bekommt (Kontrollgruppe). Wenn die Zuteilung der Patienten in beide Gruppen strikt zufällig erfolgt, kann davon ausgegangen werden, dass sie in beobachtbaren und nicht-beobachtbaren Charakteristika identisch sind. Daraus resultierend kann ein zwischen den beiden Gruppen unterschiedlicher Krankheitsverlauf mit großer Sicherheit auf die ursächliche Wirkung des Medikaments zurückgeführt werden. Unterschiede im Behandlungserfolg beruhen also nicht darauf, dass sich z. B. einfach nur Personen mit unterschiedlichem Ausgangsbefund für oder gegen eine Behandlung entschieden haben oder dass sich Personen mit unterschiedlichen finanziellen Ressourcen oder unterschiedlichem sozioökonomischen Status für eine Behandlung beworben haben. Dagegen lassen Unterschiede im Gesundheitszustand keinen Schluss auf die Wirkung des Medikaments zu, wenn sich beispielsweise nur Personen mit großen finanziellen Reserven und umfassendem Betreuungsumfeld für eine Behandlung entschieden hätten, alle anderen aber nicht. In diesem Fall wäre nicht zu unterscheiden, ob eine positivere Krankheitsentwicklung auf den Einsatz des Medikaments oder auf das bessere Betreuungsumfeld zurückzuführen ist. Es könnte sogar der gesamte Effekt auf das bessere Betreuungsumfeld zurückgehen und das Medikament wirkungslos sein. D. h. ein positiverer Krankheitsverlauf erlaubt bei einer solchen, nicht zufälligen Aufteilung der Patienten auf *Treatment*- und Kontrollgruppe keinerlei Aussage über die Wirkungsweise eines Medikaments.

einer zufallsgesteuerten Zuteilung auf Geförderte und Nicht-Geförderte sicherzustellen.¹³⁵ Teilt man nämlich eine ausreichend große Zahl an grundsätzlich förderungsfähigen Fällen (Unternehmen)¹³⁶ nach dem Zufallsprinzip in eine Kontrollgruppe und eine geförderte Gruppe (*Treatment*-Gruppe) ein, so sind diese beiden Gruppen in allen beobachtbaren

und nicht-beobachtbaren Charakteristika mit großer Sicherheit statistisch identisch – ausgenommen das Merkmal „gefördert“ bzw. „nicht-gefördert“. Wenn also nach Abschluss der Fördermaßnahme der Innovationsoutput unterschiedlich ausfällt, dann kann daraus geschlossen werden, dass dieser Unterschied tatsächlich auf die Maßnahme und nicht auf sonstige Unterschiede in den beiden Gruppen zurückzuführen ist. Eine zufallsgesteuerte Einteilung der Antragsteller in Kontrollgruppe und *Treatment*-Gruppe – analog zu medizinischen Wirkungsanalysen vor Einführung eines Medikaments – würde es demnach auch im Bereich der Innovationspolitik erlauben, die ursächliche Wirkung einer neuen Fördermaßnahme herauszuarbeiten. Nur wenn diese *ursächliche* Wirkung bekannt ist, kann festgestellt werden, wie effektiv und effizient eine Fördermaßnahme ist. Daraus lassen sich dann ökonomisch sinnvolle Schlussfolgerungen für die weitere Gestaltung von Förderprogrammen ableiten.

Zufallsgesteuerte Vergabeverfahren und Rationierung

Dabei ist eine zufallsgesteuerte Vergabe jedoch nicht gleichzusetzen mit einer unspezifischen Förderpolitik nach dem „Gießkannenprinzip“ oder mit einem generellen Verzicht auf Förder- und Antragskriterien bzw. qualifizierte, gutachterliche Auswahlprozesse.¹³⁷ Vielmehr wird die Gruppe der (geeigneten) Antragsteller einer Lotterie bzw. einer randomisierten Vergabe unterworfen werden. Da generell die verfügbaren Fördermittel beschränkt sind, konnten auch bei bisherigen Vergabeverfahren selten alle potenziell geeigneten Bewerber gefördert werden. Bei der randomisierten Vergabe gibt es also keine zusätzlichen Verlierer. Einziger Unterschied ist, dass diejenigen, die unter allen geeigneten Kandidaten nicht zum Zuge kommen, streng zufällig mithilfe eines Losverfahrens ausgewählt werden. Dadurch besitzen alle Kandidaten bei gleicher Qualität eine gleich hohe Wahrscheinlichkeit, von der Rationierung betroffen zu sein. Dabei sind Losverfahren in anderen mit Rationierungsproblemen behafteten Politikfeldern auch heute schon keineswegs unüblich. So wird beispielsweise bei Studienanfängern in der Medizin ein Teil der Studienplätze aufgrund eines Losverfahrens vergeben,¹³⁸ in Berlins Schulämtern werden knappe Plätze an Gymnasien und Sekundarschulen für Siebtklässler zugewiesen,¹³⁹ oder Finanzbehörden wählen kleine und mittelständische Unternehmen für ihre Routine-

betriebsprüfungen nach dem Zufallsprinzip aus.¹⁴⁰ Da im Fall der Innovationspolitik die Zufallsauswahl erfolgen soll, um belastbare Wirkungsanalysen der Fördermaßnahmen zu ermöglichen, würden längerfristig alle Unternehmen davon profitieren, also nicht nur die zufällig ausgewählten, denn die Vergabe eines begrenzten Topfs an Fördermitteln würde effektiver werden.¹⁴¹

Die heute vorherrschende Evaluationspraxis in der deutschen wie in der europäischen Innovationspolitik stützt sich häufig lediglich auf die Beschreibung statistisch beobachtbarer Unterschiede zwischen Geförderten und Nicht-Geförderten oder analysiert die Zusammenhänge auf Basis ökonometrischer Verfahren und unter Berücksichtigung einer Vielzahl an Kontrollvariablen.¹⁴² Allerdings besteht die Gefahr, dass die beobachteten Unterschiede oft nicht auf die Förderung zurückgehen, sondern darauf, dass sich von vornherein andersartige Unternehmen für die Förderung beworben bzw. nicht beworben haben. Die Einbeziehung beobachtbarer Kontrollvariablen mag die Probleme zwar reduzieren, schafft aber keine Abhilfe bezüglich der Tatsache, dass die wesentlichen Unterschiede in unbeobachtbaren Variablen bestehen könnten, wie beispielsweise bessere Führungskompetenzen des FuE-Managements oder organisatorische Lernfähigkeit. Diese Kriterien lassen sich unter realistischen Annahmen selbst bei aufwendigen Datenerhebungen nicht erfassen. Eine verbesserte Innovationstätigkeit bei geförderten Unternehmen darf dann aber nicht ursächlich auf die Förderung zurückgeführt werden.

Neue, elaboriertere ökonometrische Methoden¹⁴³ mögen dem Kausalitätsproblem besser gerecht werden, sind aber aufgrund der hohen Anforderungen an die Daten ebenfalls begrenzt und mit hoher statistischer Unsicherheit behaftet. Nicht verwunderlich ist es deshalb, dass Politikmaßnahmen in der Evaluationsliteratur oftmals sehr unterschiedlich bewertet werden, je nachdem, welche Methode der Evaluation zu Grunde liegt.¹⁴⁴

Im Vergleich dazu ist die Überzeugungs- und Aussagekraft einer Evaluation, die auf einer randomisierten Vergabe beruht, hoch, denn die Resultate sind einfach zu verstehen (vgl. Box 7). Die Wirkung der Fördermaßnahme („*Treatment*-Effekt“) ist der Unterschied im Innovationsoutput zweier Gruppen: der Gruppe der Geförderten einerseits und der Kontrollgruppe

andererseits. Auf weitere, methodisch hochkomplexe Verfahren und Annahmen, die eine einfache Interpretation erschweren, die aber in der ökonometrisch anspruchsvollen Evaluationspraxis üblich und notwendig sind, kann damit verzichtet werden. Wichtig ist lediglich, dass die *Treatment*- und die Kontrollgruppe streng nach dem Zufallsprinzip zusammengestellt werden. Außerdem müssen beide Gruppen ausreichend groß sein, damit man statistisch zuverlässige Aussagen über die Wirksamkeit einer Maßnahme treffen kann.

Je nach Ausgangslage bei einer zu evaluierenden Fördermaßnahme kann es sich anbieten, entweder alle Bewerber zufällig in eine Kontroll- und eine Fördergruppe einzuteilen, oder nur einen Teil in das Losverfahren einzubeziehen, wie etwa bei der Studienplatzvergabe in der Medizin. Bei randomisierten Evaluationen im Innovationsbereich könnte man beispielsweise entscheiden, die Randomisierung nur auf einen Teil zu beschränken, wenn es um Maßnahmen geht, bei denen man im Vorfeld sicher ist, dass diese Maßnahme für eine Top-Gruppe von eindeutig förderungswürdigen Bewerbern die größtmögliche Wirkung erzielt, aber das Gleiche möglicherweise nicht für alle gilt. Die erstgenannte kleine Gruppe würde komplett ausgewählt und die Randomisierung würde beschränkt auf die größere Restgruppe oder auf eine Gruppe an Fällen an einer unscharfen Grenze. Zu beachten ist bei einer solchen gemischten Vorgehensweise allerdings, dass statistisch gesicherte Aussagen über die Wirkungsfähigkeit des Instruments nur für diese Grenzfälle möglich sind und vor allem kein Vergleich mit der Top-Gruppe erfolgen kann. Letzteres würde im Sinne einer bestmöglichen Evaluation eher für eine vollständige Randomisierung aller Fälle für einen begrenzten Zeitraum sprechen. Eine vollständige Randomisierung bietet sich ebenfalls an, wenn es keine oder nur wenig zuverlässige Kriterien für eine Vorauswahl gibt. Evaluationen, die auf der randomisierten Einführung einer Fördermaßnahme beruhen, bieten zusätzlich den Vorteil, dass nicht nur verglichen werden kann, ob eine Förderung besser ist als eine Nicht-Förderung. Sondern es können – unter der Voraussetzung ausreichend großer Fallzahlen – auch verschiedene Ausgestaltungen eines Förderinstruments gegenübergestellt werden, um so die relativ kostengünstigeren oder wirkungsvolleren Maßnahmen identifizieren zu können. Im politischen Entscheidungsprozess kann dann entschieden werden, welches Kriterium

gegebenenfalls höher zu gewichten ist. Insofern können Evaluationen, basierend auf einer randomisierten Einführung von Fördermaßnahmen (für bestimmte Förderbereiche), ein breites Spektrum an Informationen zur Verbesserung staatlicher FuE-Politik liefern.

Bedenken gegenüber randomisierten Evaluationsverfahren

Unabhängig von den vielfältigen methodischen Vorteilen werden der Einführung zufallsgesteuerter Evaluationsverfahren in der Politikpraxis häufig vergaberechtliche oder ethische Bedenken entgegengebracht. So werden Einwände vorgetragen, dass eine zufällige Aufteilung auf eine geförderte und eine nicht-geförderte Gruppe nicht mit vorherrschendem Vergaberecht vereinbar ist. Allerdings wird vernachlässigt, dass bei einer Überzeichnung eines Programms ebenfalls zusätzliche Auswahlkriterien (wie beispielsweise Windhundverfahren oder regionale Verteilungskriterien) herangezogen werden, deren Wirkungen einer gewissen Beliebigkeit nicht entbehren. Mitunter werden ethische Bedenken gegen ein Losverfahren vorgetragen, da dies Ungerechtigkeiten schüre. Dabei wird außer Acht gelassen, dass nicht das Losverfahren, sondern die Rationierung an sich dazu führt, dass einzelne Personen oder Unternehmen leer ausgehen. Das Losverfahren stellt im Gegenteil sogar sicher, dass alle Bewerber mit gleich großer Wahrscheinlichkeit leer ausgehen und die Ausgangswahrscheinlichkeit nicht für manche höher ist – und dies oft auf der Basis von inhaltlich oder statistisch fragwürdigen Kriterien. Solche Probleme und Bedenken lassen sich – wie sich im Ausland (bspw. in den Niederlanden) zeigt – durch eine sorgfältige Implementierung des randomisierenden Verfahrens und eine geeignete Kommunikation weitgehend ausräumen.¹⁴⁵

Vor Kurzem sind auch mehrere europäische Länder in der Wissenschafts- und Innovationspolitik dazu übergegangen, bei einzelnen Fördermaßnahmen ein solches randomisiertes Vergabeverfahren einzusetzen. Dazu gehört unter anderem die Ausgabe von Innovationsgutscheinen an KMU in den Niederlanden¹⁴⁹ oder in verschiedenen Pilotregionen Großbritanniens. Die Evaluation dieser neuen Maßnahmen, die auf den Technologietransfer zwischen Wirtschaft und Wissenschaft abzielten, ergab, dass die Anzahl vertraglicher Kooperationen kurzfristig gesteigert wurde. Zugleich quantifizieren die Ergebnisse prä-

Erfahrungen mit randomisierten Evaluationen in der US-amerikanischen Bildungspolitik

BOX 07

In der US-amerikanischen Bildungspolitik sind randomisierte Evaluationsverfahren bereits seit längerer Zeit üblich.¹⁴⁶ So untersuchte beispielsweise eine viel beachtete Studie im Bundesstaat Tennessee¹⁴⁷ die langfristige Wirkung der Einführung kleiner Unterrichtsklassen und zusätzlicher Hilfslehrer auf den Bildungserfolg. Mit Einführung der Maßnahme wurden die Schüler vorab auf zufälliger Basis unterschiedlich großen Klassen zugeordnet bzw. Klassen mit einem oder ohne einen zusätzlichen Hilfslehrer. Die Studie kommt zu dem Ergebnis, dass die Schüler in den kleinen Klassen (auch langfristig) durchschnittlich deutlich besser abschneiden, dass aber die Unterstützung durch einen Hilfslehrer keine Auswirkungen hat. Diese verlässlichen und nachvollziehbaren Evaluationsergebnisse trugen im Zuge der politischen Debatten maßgeblich dazu bei, dass in der Folge in mehr als zwölf US-Bundesstaaten die Klassengrößen reduziert und zugleich auf den kostspieligen Einsatz von Hilfslehrern verzichtet wurde. Damit konnte eine größere Effizienz der eingesetzten Mittel sichergestellt werden.¹⁴⁸

zise das Ausmaß der Mitnahmeeffekte: Es zeigt sich, dass etwa jedes neunte Projekt auch ohne den Gut-schein durchgeführt worden wäre.¹⁵⁰ In den Niederlanden wurde außerdem eine randomisierte Vergabepaxis im Kontext von Innovationskredit und Stipendienprogrammen für Nachwuchswissenschaftler eingeführt, so dass zukünftig auch zuverlässige Aussagen über die Wirkung solcher Programme möglich sein werden.

Einsatzmöglichkeiten randomisierter Evaluationsverfahren und ihre Grenzen

Randomisierte Evaluationsverfahren haben sehr breite Einsatzmöglichkeiten, eignen sich aber nicht für alle Politikbereiche gleichermaßen.¹⁵¹ Grundsätzlich kommen sie eher für solche Förderprogramme der Wissenschafts- und Innovationspolitik in Frage, die eine größere Zahl von Förderempfängern bedient, da nur so eine ausreichende statistische Trennschärfe garantiert werden kann. Es ist eine ausreichend große Zahl an Geförderten und Nicht-Geförderten erforderlich, um bei der anschließenden Analyse der potenziel-

len Unterschiede zwischen Kontroll- und *Treatment*-Gruppe statistisch gesicherte Ergebnisse zu erhalten und damit zuverlässig auf die Wirkung der Maßnahme zu schließen. Für bestimmte Programme, insbesondere zur Förderung größerer Technologieprojekte, werden sich daher randomisierte Evaluationsverfahren auch in Zukunft nicht eignen. Für Programme mit vielen Antragstellern sollte jedoch eine randomisierte Einführung in Betracht gezogen werden. Zugleich empfehlen sich gemäß ethischer und politischer Akzeptanzüberlegungen vor allem solche Förderaktivitäten, die typischerweise „überzeichnet“ sind, d. h. bei denen die Zahl der grundsätzlich geeigneten Anträge das vorgesehene Gesamtbudget übersteigt und es insofern eine unvermeidliche Begleiterscheinung des Programms ist, dass manche Antragsteller in jedem Falle leer ausgehen. Dadurch werden vor allem die ethischen Bedenken gegenüber einer zufallsgesteuerten Vergabe reduziert. Ein letztes wichtiges Kriterium bezüglich der Eignung einer Fördermaßnahme für das randomisierte Evaluationsverfahren ist, dass die Ziele des Programms explizit und konkret definiert sein müssen, denn nur so ist auch ein Erfolg eindeutig und unbestritten messbar. Eine randomisierte Vergabe setzt also von Seiten der Politik voraus, dass diese frühzeitig klar abgesteckte und messbare Ziele für eine Maßnahme formuliert und damit festlegt, unter welchen Bedingungen diese als Erfolg zu werten ist.

Vor diesem Hintergrund kommen im deutschen Politikkontext für das randomisierte Evaluationsverfahren generell einige aktuelle Förderprogramme zur Innovationstätigkeit im Mittelstand in Frage, wie beispielsweise das ZIM-Programm oder KMU-innovativ. Ähnliches gilt für die Gründerförderung oder die Vergabe von Forschungsstipendien an einzelne Wissenschaftler, wie beispielsweise das EXIST-Gründerstipendium oder die Auslandsstipendienprogramme der DFG für Nachwuchswissenschaftler. Hier könnten nicht nur wichtige Erkenntnisse über die Wirksamkeit der konkreten Maßnahme, sondern auch wertvolle Erfahrungen für den Umgang mit randomisierten Evaluationsverfahren gewonnen werden. Besonders lohnenswert im Hinblick auf den ökonomischen Wert der Ergebnisse wäre eine randomisierte Evaluation der Mittelstandsförderung des Bundes, da diese ein hohes Finanzvolumen aufweist. Allein im Jahr 2013, und aller Voraussicht nach auch für das Jahr 2014, sieht der Bundeshaushalt für das ZIM-Programm jeweils Mittel in Höhe von rund 500 Millionen Euro

vor.¹⁵² Selbst ein nur leicht verbesserter Mitteleinsatz könnte hier also eine erhebliche Wirkung entfalten.

Trotz aller Vorteile sind auch dem Einsatz von Evaluationsergebnissen auf Grundlage einer randomisierten Vergabe Grenzen gesetzt. Eine der wichtigsten Einschränkungen in der Verwendung der Evaluationsergebnisse besteht darin, dass diese sich nicht zwangsläufig verallgemeinern lassen, so dass jede neue Ausgestaltung eines Instruments oder die Anwendung in einem anderen Kontext idealerweise auf einer eigenen randomisierten Einführung basiert. So lässt sich die Wirkung eines Instruments auf ausgewählte antragstellende Unternehmen – selbst wenn sie sorgfältig ermittelt wurden – nicht oder nur sehr eingeschränkt auf diejenigen Unternehmen übertragen, die bisher noch nie oder allenfalls an anderer Stelle einen Förderantrag eingereicht haben. Allerdings besteht dieses Problem nicht nur bei randomisierten Evaluationen, sondern letztlich bei allen Evaluationsverfahren. Dies spricht somit nicht grundsätzlich gegen randomisierte Evaluationen. Als ein weiterer Nachteil von Evaluationen mit *Treatment*- und Kontrollgruppen wird häufig angeführt, dass Wissensspillover zwischen geförderten und nicht-geförderten Unternehmen auftreten können, so dass beim Vergleich der beiden Gruppen die Wirkung unterschätzt werden könnte. Aber auch dieses Problem kann in ähnlicher Weise bei herkömmlichen Evaluationen auftreten. Außerdem erfordern randomisierte Verfahren – wie auch alle anderen Evaluationsverfahren – eine sehr gute und langfristige Datenerhebung, da die Validität der Ergebnisse essenziell von der Qualität der Ergebnismessungen und von den kurz- und langfristigen Vergleichsmöglichkeiten abhängt. Zusätzlich muss im Rahmen einer sorgfältigen Vorbereitung der Evaluation eine qualitativ hochwertige Datenbasis über die Grundgesamtheit der potenziellen Förderempfänger zur Verfügung stehen, damit das Design des Losverfahrens und der Zuteilung in Kontroll- und *Treatment*-Gruppe statistisch angemessen festgelegt werden kann. Nach dem offiziellen Ablauf der Förderperiode müssen die entsprechenden Daten nicht nur bei den geförderten, sondern auch bei den nicht-geförderten Antragstellern gesammelt und bereitgestellt werden.

Politikempfehlungen

Die Expertenkommission empfiehlt, randomisierte Evaluationen als eines der Standardinstrumente in das Evaluationsportfolio staatlicher FuE-Förderung aufzunehmen. Eine randomisierte Einführung von Fördermaßnahmen ist insbesondere in Bereichen sinnvoll, in denen mit einer größeren Zahl an Antragstellern zu rechnen ist und aufgrund begrenzter Budgets von einer Überzeichnung der Fördermaßnahme ausgegangen werden kann. Dieses Verfahren bietet die Möglichkeit, wertvolle Informationen über die Wirksamkeit und die Ausgestaltungsmöglichkeiten einer effizienten Fördermaßnahme zu erhalten, die wiederum die Entscheidung darüber unterstützen, ob ein Förderinstrument verstetigt oder auf weitere Bereiche ausgedehnt werden sollte. Randomisierte Evaluationsverfahren können so zu Effizienzgewinnen und zu einem deutlich verbesserten Einsatz der Fördermittel führen – sofern die Erkenntnisse auch Eingang in politische Entscheidungsprozesse finden.

Die Expertenkommission empfiehlt für das auf den innovativen Mittelstand ausgerichtete Instrument ZIM, eine Evaluation auf Basis einer randomisierten Mittelvergabe zu starten. Damit soll einerseits Wissen generiert werden, wie die Effizienz der genannten Fördermaßnahme gesteigert werden kann, und andererseits sollen wertvolle Erkenntnisse und Erfahrungen im Umgang mit randomisierten Evaluationen gesammelt werden, die später systematisch auf weitere Anwendungsfelder übertragen werden können.

Da in Europa die randomisierte Einführung von Fördermaßnahmen im Bereich der Innovationspolitik noch wenig verbreitet ist, eröffnet sich hier für die deutsche Politik die Chance, mit einer intelligenten, evidenzbasierten Forschungs- und Innovationspolitik eine Vorreiterrolle einzunehmen.