

# Was funktioniert - und was nicht

Wirtschaftspolitische Programme sollten konsequenter überprüft werden, fordern Forscher.

► In den USA wird mehr und systematischer evaluiert.

► Bundesministerien versprechen Besserung.

**Hans Christian Müller**  
Düsseldorf

Wenn die Politik eine neue Subvention beschließt, dann stellen sich immer dieselben Fragen: Hilft die Förderung wirklich weiter? Werden die richtigen Unternehmen gefördert? Und brauchen sie die Hilfe wirklich - oder streichen sie das Geld einfach nur gerne ein? Viel zu oft gibt es viel zu wenige Antworten darauf. „Ich ärgere mich oft, wie wenige fundierte statistische Studien es eigentlich gibt, auf die man sich stützen kann“, sagt Monika Schnitzer, Ökonomin an der Uni München, die auch in der Politik als Beraterin geschätzt wird.

Gemeinsam mit Kollegen hat es sich Schnitzer zur Aufgabe gemacht, dafür zu werben, dass die Förderpolitik in Deutschland mehr als heute durch unabhängige Studien evaluiert wird. „Die Steuergelder sollen doch für Projekte ausgegeben werden, von denen man weiß, dass sie Erfolg versprechen“, meint Schnitzer.

Alle Fördermaßnahmen sollten grundsätzlich bewertet werden, wenn die Kosten einen Schwellenwert überschreiten, forderte jüngst der wissenschaftliche Beirat des Bundeswirtschaftsministeriums. „Wir müssen eine Evaluierungskultur aufbauen“, sagt Achim Wambach von der Uni Köln, der dem Gremium vorsteht. Andere Länder seien da weiter: In den USA etwa gibt das Arbeitsministerium bei jedem Projekt 0,5 Prozent der Budgetsumme für die Evaluation aus.

Aber auch in Deutschland tut sich etwas. Das Wirtschaftsministerium etwa hat 2010 einen unabhängigen Aufbaustab für das Fördercontrolling eingerichtet, der die Evaluation der milliardenschweren

Programme der Behörde koordiniert. Dort will man explizit dafür sorgen, dass die beauftragten Forscher bei ihrer Evaluation qualitative Standards beachten und unabhängig arbeiten können - und nicht etwa das schreiben, was gefällt.

Für die Politik können Evaluierungen durchaus ungemütlich werden. Das zeigte im vergangenen Jahr die groß angelegte Bewertung der deutschen Familienpolitik - die zum Totalverriss für die Regierung wurde. Ähnliches hatte es in der Vergangenheit im Bereich der Arbeitsmarktpolitik gegeben - in jenem Bereich also, in dem bereits konsequent kontrolliert wird. „Im Zuge der Hartz-Reformen wurde erstmals eine Evaluierungspflicht gesetzlich verankert“, sagt Marco Caliendo von der Uni Potsdam, der regelmäßig an entsprechenden Studien mitgearbeitet hat.

Gleich eine der ersten Untersu-

chungen hatte ein verblüffendes Ergebnis. Demnach bewirkten die in den 90er-Jahren so beliebten Arbeitsbeschaffungsmaßnahmen für Arbeitslose genau das Gegenteil von dem, was sie sollten: Die Teilnehmer fanden hinterher nicht etwa öfter eine normale Anstellung am ersten Arbeitsmarkt, sondern seltener. „Menschen mit eigentlich guten Chancen waren regelrecht in den Maßnahmen gefangen“, sagt Gesine Stephan vom Institut für Arbeitsmarkt- und Berufsforschung (IAB).

Doch manchmal gaben die Erfolgskontrollen der Politik auch Rückenwind - etwa bei den sogenannten Ich-AGs, also der staatlichen Förderung für Arbeitslose, die sich selbstständig machen wollen. „Die hatten einen echt schlechten Ruf“, sagt Wissenschaftler Caliendo - dabei seien sie durchaus erfolgreich gewesen, wie seine Forschung gezeigt habe.



**Vermittlungsgespräch für Jobsuchende:** Die Arbeitsmarktpolitik wird besonders intensiv kontrolliert.

”

Im Zuge der Hartz-Reformen wurde erstmals eine Evaluierungspflicht im Gesetz verankert.

**Marco Caliendo**  
Arbeitsmarktforscher, Uni Potsdam

Die Arbeitsmarktforscher haben einen großen Vorteil: Dank der Daten der Rentenversicherung können sie inzwischen einen riesigen Fundus von individuellen Lebensläufen auswerten, was ihre Ergebnisse valider macht.

**In anderen Bereichen ist die Datelage dagegen noch völlig unzureichend** - etwa wenn es darum geht, die Wirkung der Konjunkturpakete zu untersuchen, mit denen die Verwerfungen der Finanzkrise abgefedert werden sollten. „Eigentlich bräuchten wir genaue Daten, wo welches Geld bekommen und verwendet hat“, sagt Ökonomin Schnitzer. Doch zu oft gebe es nur Gesamtsummen. Dann sei die Evaluierung nur mit Hilfe starker Modellannahmen möglich - was sie angreifbar mache.

Und so geht es den Forschern gar nicht so sehr darum, dass mehr Mittel zur Verfügung gestellt werden, sondern dass die Grundlagen besser werden: „Wir brauchen umfassendere Mikrodaten, die allen Wissenschaftlern zur Verfügung stehen“, sagt Beiratschef Wambach. Nur so könnten hinterher auch andere Forscher die Ergebnisse der Studien überprüfen - und in einen wissenschaftlichen Wettbewerb eintreten. Um prüfen zu können, ob eine Maßnahme wirklich die Ursache einer Entwicklung ist und nicht nur eine Begleiterscheinung, empfiehlt der Beirat, mehr mit Experimenten zu forschen - was bisher in Deutschland nicht üblich ist (siehe unten).

Anderswo hat man da weniger Berührungängste. Ohnehin holen sich die Forscher viele Ideen aus dem Ausland: „In Großbritannien werden Fördermaßnahmen erst einmal befristet“, sagt Schnitzer - und nur dann fortgeführt, wenn ihre Sinnhaftigkeit belegt sei. Und in den USA gibt es zahlreiche Websites, die schlichtweg alle Erkenntnisse zu einem Bereich zusammenstellen und aufbereiten. So können sich Interessierte schnell einen Überblick verschaffen. Der schlichte Titel der Seiten: What works. Zu Deutsch: Was funktioniert.

## EVALUIERUNG

# Experiment am arbeitenden Menschen

Moderne Forschungsmethoden treffen in Deutschland auf große Skepsis - anders als in anderen Ländern.

Wenn Mediziner wissen möchten, ob ein Medikament hilft, dann machen sie ein Experiment: Sie teilen eine Gruppe Probanden per Zufall in zwei Gruppen, geben der einen Gruppe Placebos und der anderen richtige Pillen. Geht es der zweiten Gruppe hinterher im Schnitt deutlich besser, dann wirkt das Medikament wahrscheinlich - sonst nicht.

Um zweifelsfrei feststellen zu können, ob eine Fördermaßnahme den gewünschten Zweck erfüllt, müssten Wirtschaftsforscher genauso vorgehen. „Doch in Deutschland tut man sich noch schwer damit“, meint Arbeitsmarktforscher Marco Caliendo. Hier gelte es weithin als unethisch,

einem Teil der Menschen zufällig eine Förderung zuzugestehen und anderen nicht, nur um die Wirkung sauber studieren zu können.

Allerdings: Wer einfach vergleicht, ob die Teilnehmer einer Maßnahme später mehr verdienen als die Nicht-Teilnehmer, macht es sich zu leicht - denn es könnte ja sein, dass einfach die Schlauderen gefördert wurden. Dann aber wäre die Maßnahme nicht der wahre Grund für den Erfolg.

Um trotzdem eine Aussage treffen zu können, greifen Forscher wie Caliendo auf einen Trick zurück. Sie suchen in ihren Daten sogenannte statistische Zwillinge, die sich - was Alter, Bildung oder Familienstand

angeht - möglichst ähnlich sind, aber von denen nur einer gefördert wurde. Vergleiche solcher Paare bringen dann auch gute Resultate.

Anderswo sind Feldexperimente bereits etabliert: In Frankreich etwa untersuchten Forscher, ob private Arbeitsvermittler besser arbeiten als staatliche. Dafür teilten sie 200 000 Arbeitslose per Los zu - und prüften, wer ein Jahr später häufiger einen richtigen Job hatte. Das Ergebnis war überraschend: Die staatlichen Arbeitsvermittler waren doppelt so erfolgreich.

In Dänemark zeigte ein ähnliches Experiment, dass Arbeitslose häufiger neue Jobs finden, wenn sich die Arbeitsvermittler regelmäßig inten-

”

Wirklich valide ist eine Erkenntnis erst, wenn sie mit verschiedenen Daten und Methoden bestätigt werden kann.

**Katharina Spieß**  
Familienforscherin, DIW

siv mit ihnen zusammensetzen - und sie nicht in Gruppensitzungen betreuen. Die höheren Kosten lohnen sich, so das Fazit. Und auch in Deutschland gab es inzwischen erste Versuche mit Feldexperimenten - etwa im Bereich der Pflege oder der Arbeitsvermittlung.

Letztlich sind aber auch Feldexperimente nicht der alleinige Erkenntnisbringer. „Man darf sich nicht nur auf eine Methode verlassen“, findet Katharina Spieß, Familienforscherin beim DIW Berlin. Wirklich valide sei eine Erkenntnis erst, wenn verschiedene Studien mit Hilfe verschiedener Daten und Methoden zum selben Ergebnis kämen. Hans Christian Müller