

## Frühjahrstreffen 2019 des AK FTI der DeGEval

### Methodische Herausforderungen und Lösungsansätze der Evaluation im Politikfeld FTI

*Karoline Rodriguez,<sup>1</sup> Jan Wessels,<sup>1</sup> Marianne Kulicke,<sup>2</sup>  
Sascha Ruhland<sup>3</sup>*

Am 18. Juni 2019 fand das diesjährige Frühjahrstreffen des AK FTI der DeGEval in Berlin statt. Es stellte methodische Herausforderungen und neue Herangehensweisen in den Mittelpunkt. Etwa 35 Teilnehmerinnen und Teilnehmer aus Evaluations-einrichtungen, Ministerien und Projektträgern diskutierten anhand aktueller Projektbeispiele.

#### Praxis der Wirtschaftlichkeitsanalyse (Vortrag: Karoline Rodriguez Rivera/iit)

Die Untersuchung der Wirtschaftlichkeit von Fördermaßnahmen ist durch den Grundsatz der Wirtschaftlichkeit der Bundeshaushaltsordnung eine zentrale Anforderung an Förderprogrammevaluationen. Wie in den letzten Jahren methodisch mit diesem Thema umgegangen wurde, zeigte der erste Beitrag des Frühjahrstreffens auf. Basierend auf einer Auswertung von 40 Evaluationen wurden gängige Ansätze zur Einschätzung der Vollzugs- und Maßnahmenwirtschaftlichkeit im Rahmen der Erfolgskontrolle präsentiert. Die Vollzugswirtschaftlichkeit, also die Wirtschaftlichkeit des Vollzugs der Maßnahme im Hinblick auf den Ressourcenverbrauch, wurde vor allen Dingen mithilfe der Errechnung des Verwaltungsaufwands, durch Prozessanalysen sowie durch Zuwendungsempfängerbefragungen und mit Soll-Ist-Vergleichen untersucht. Die Maßnahmenwirtschaftlichkeit („War die Maßnahme im Hinblick auf die übergeordneten Zielsetzungen wirtschaftlich?“) wurde oftmals anhand des Nutzens auf Seite der Fördernehmer eingeschätzt. Auch Hebel- und Mitnahmeeffekte wurden berechnet. Dabei wurde die Vollzugswirtschaftlichkeit insgesamt häufiger geprüft als die Maßnahmenwirtschaftlichkeit – auch deshalb, weil es nicht immer möglich ist, den Programmnutzen messbar zu machen. Auch in der anschließenden Dis-

1 Institut für Innovation und Technik (iit), Berlin

2 Fraunhofer-Institut für System- und Innovationsforschung, Karlsruhe

3 KMU Forschung Austria, Wien

kussion wurde deutlich, dass das Thema in den letzten Jahren zwar an Bedeutung gewonnen hat, es aber weiterhin Bedarf für fundierte Studien und eindeutige Handlungsrichtlinien gibt. Die Diskussion sollte deshalb vor allem auch mit Fachreferaten und Vertreter(inne)n des Bundesrechnungshofes weitergeführt werden.

## Fokus auf die ‚Besten‘ und Kontrollgruppenansätze (Vortrag: Christian Rammer/ZEW)

Der zweite Vortrag präsentierte einen Kontrollgruppenansatz, um das österreichische Förderprogramm Frontrunner zu evaluieren. Hier lag die große Herausforderung darin, eine geeignete Kontrollgruppe zu finden, da sich das Programm bewusst an die ‚besten‘ Unternehmen des Landes richtet und diese deshalb schlecht mit anderen Firmen in Österreich vergleichbar sind. Um diese Beschränkung zu umgehen, wurde eine Kontrollgruppe von deutschen Unternehmen mithilfe des Mannheimer Innovationspanels konstruiert, die in ihren Eigenschaften den österreichischen Frontrunner-Unternehmen stark ähneln (hoher Marktanteil, hohe Exportquote, Exportquote außerhalb Deutschlands und interne FuE-Aktivitäten). Für einen Zeitraum von 2010-2017 wurden mit einem Conditional-Difference-in-Differences-Ansatz Unternehmen aus Kontroll- und Treatmentgruppe hinsichtlich ihrer Branche und Größenklasse gematched und die vorhandenen Paneldaten mit Hinblick auf potenzielle Effekte der Frontrunner-Förderung auf relevante Performance-Indikatoren analysiert. Auch wenn die Ergebnisse dafür sprechen, dass das Frontrunner-Programm einen positiven Einfluss auf Wachstumsaspekte und FuE-Aktivitäten hatte, so stellte sich heraus, dass die ausgewählten deutschen Unternehmen aufgrund von *Country Fixed Effects* keine geeignete Kontrollgruppe darstellten. Das österreichische FuE-System ist im Vergleich zu Deutschland dynamischer und die FuE-Intensität ist in Österreich generell viel höher. Dies wird vor allen Dingen dann sichtbar, wenn die Sektorstruktur beim Vergleich von FuE-Ausgaben von Unternehmen beider Länder berücksichtigt wird. Auch die Kontrolle dieser systematischen Unterschiede über Innovationssystem-Indikatoren wie z.B. die Veränderung der FuE-Intensität im Unternehmenssektor, reichte aufgrund der geringen Varianz über die Zeit und der hohen Multikollinearität der unabhängigen Variablen auf Länderebene nicht aus, um Rückschlüsse ziehen zu können. Bei der Beschränkung der Treatmentgruppe auf Unternehmen, die Investitionszuschüsse zum ERP-Kredit erhielten, trat zusätzlich das Problem des *Selection Bias* auf, da dieser Programmzweig nur von Unternehmen genutzt wird, die expandieren möchten. Zusätzlich wurde eine alternative Kontrollgruppe aus abgelehnten Förderantragsstellern konstruiert. Hier deuteten die Ergebnisse auf keine Effekte der Frontrunner-Förderung hin; auch deshalb, weil fast alle abgelehnten Unternehmen andere FFG-Förderung erhielten. Ein Fazit der anschließenden Diskussion war, dass ein länderübergreifender Kontrollgruppenansatz eine interessante Alternative bietet, um die Förderung der ‚Besten‘ zu evaluieren, aber an seine Grenzen stößt, wenn sich Länder systematisch unterscheiden. Ein Methodenmix aus Fallstudien, Expertengesprächen und Szenarien ist in diesem Fall angebracht.

## Anwendung analytisch-quantitativer Methoden der empirischen Sozialforschung im Rahmen dreier Programmevaluationen (Vortrag: Peter Kaufmann/KMU Forschung Austria)

In diesem Vortrag wurden drei Kontrollgruppenvergleiche vorgestellt, die für die Evaluation von Förderprogrammen aus der Schweiz, Österreich und Deutschland angestellt wurden. Das erste Beispiel befasste sich mit der Evaluation der CTI Entrepreneurship Sensibilisierungs- und Trainingsprogramme, die in den Jahren 2016 und 2017 an ausgewählten Schweizer Hochschulen implementiert wurden. Dabei stützte sich die Evaluation auf die *Theory of Planned Behavior* (TPB), ein Modell aus der Sozialpsychologie zur Untersuchung von Faktoren, die das unternehmerische Verhalten beeinflussen. Während mithilfe von Programmteilmehrfbefragungen die Hypothesen der *Theory of Planned Behavior* bestätigt werden konnten, war es aufgrund von signifikanten Unterschieden zwischen Treatment- und Kontrollgruppe (Student(inn)en einer Schweizer Hochschule ohne CTI Programm) nicht möglich, die beobachtete Wirkung auf die Kursteilnahme zurückzuführen. Als Positivbeispiel wurde die Evaluation des EFRE-Programms in Österreich vorgestellt. Mithilfe einer großen Menge an Bilanzdaten für geförderte und nicht geförderte Unternehmen konnte eine vergleichbare Kontrollgruppe identifiziert werden, die für die Wirkungsanalyse hinzugezogen werden konnte. Zum Abschluss wurden vorläufige Ergebnisse aus dem Kontrollgruppenansatz im Rahmen der Evaluation des deutschen ZIM-Programms vorgestellt. Mithilfe von Daten des Mannheimer Innovationspanels sowie Befragungen von geförderten und nicht geförderten Unternehmen konnte eine geeignete Kontrollgruppe konstruiert werden. Allerdings machte dieses Beispiel auch deutlich, dass ein quasi-experimentelles Design oftmals dazu führt, dass durch den sehr hohen Vergleichbarkeitsanspruch an die Daten viele Unternehmen aus der Beobachtungsmenge fallen und die Ergebnisse somit an Aussagekraft verlieren.

## Übergreifendes, vergleichendes Programm-Monitoring mit gemeinsamen Erhebungsinstrumenten über mehrere Förderprogramme hinweg (Vortrag: Stefan Schröder/PTJ, Jan Wessels/VDI/VDE-IT)

Dieser Beitrag präsentierte erste Ergebnisse aus einem laufenden förderprogrammübergreifenden Maßnahmenmonitoring, das durch die beiden Projektträger PTJ und VDI/VDE-IT durchgeführt wird. Hier lag neben der Einordnung und der Beschreibung des Maßnahmenmonitorings ein Schwerpunkt auch auf den Vergleichen von Prozess- und Befragungsdaten der drei Förderprogramme, die durch die Standardisierung der Befragungsdaten möglich wurden. Einen weiteren Schwerpunkt bildeten die Chancen, die ein kontinuierliches Monitoring von Förderprogrammen den Programmteilnehmenden bietet. Wie auch bei Evaluationen werden hier Zuwendungsempfängerbefragungen durchgeführt, allerdings ist es durch den frühen Zeitpunkt

der Befragungen und die Nähe zum Projektträger möglich, Förderprogrammaktivitäten und Prozesse während der Laufzeit den Befragungsergebnissen anzupassen (z.B. Zusatzangebote für Fördernehmer). So kommt es durch das direkte Feedback zu einer höheren Kundenorientierung der Projektträgerarbeit. Weiterhin bietet das Maßnahmenmonitoring einen besseren Einblick in die mittelfristige Wirkung der Förderung, da auch mit deutlichem Abstand zum Projektende (zwei sowie fünf Jahre) befragt wird. Eine zusätzliche Chance bietet sich für externe Evaluator(inn)en, die die Ergebnisse des Maßnahmenmonitorings nutzen können und so möglicherweise ihre Kapazitäten für andere, innovative Auswertungsmethoden und -aspekte verwenden können. Ebenfalls wurden mögliche Probleme diskutiert, die das Maßnahmenmonitoring mit sich bringt. So sind die Befragungen im Rahmen des Monitorings eine zusätzliche Belastung der Fördernehmer. Es ist daher wichtig, deutlich zu machen, welchen Nutzen, die Befragung auch für die Teilnehmer(innen) haben kann. Mögliche Rollenkonflikte, ein erhöhter Aufwand beim Projektträger und mögliche Probleme der Datenverfügbarkeit nach Abschluss der Förderung sind weitere Punkte, die es zu beachten gilt.

### **Vernetzungsanalysen im Rahmen von Evaluationen in den Bereichen FTI (Vortrag: Martina Kauffeld-Monz/iit)**

Eine Einführung in das Thema Vernetzungsanalysen bot der Vortrag von Frau Kauffeld-Monz. Hier wurde aufgezeigt, welchen Mehrwert das Hinzuziehen von Netzwerkanalysen für Evaluationen bietet. Mithilfe von Vernetzungsanalysen können Forschungsnetzwerke, Strukturen der Technologieentwicklung und Innovationsnetzwerke bzw. Cluster visualisiert und analysiert werden. Hier kann zum einen das System selbst betrachtet werden, aber auch (Austausch-)Beziehungen, Kooperationen, Ko-Publikationen/-Patente, Wissensflüsse und Kommunikationsstrukturen für verschiedene Teilbereiche und auf unterschiedlicher Akteursebene dargestellt werden. Die Multiplexität bezeichnet dabei das gleichzeitige Auftreten verschiedener Beziehungsformen zwischen zwei Knoten (Akteur(inn)en, d.h. Personen oder Organisationen). Bei der Analyse eines Netzwerkes gestaltet sich das Netzwerk je nach Auswahl einer bestimmten Beziehungsform entsprechend unterschiedlich. Netzwerkgrafiken ermöglichen es, verschiedene Informationen strukturiert zu präsentieren, damit diese vom Betrachter/von der Betrachterin ‚ökonomischer‘ (kognitiv) verarbeitet und beurteilt werden können. Die Analyseverfahren geht hierbei weit über die reine Visualisierung von strukturellen Zusammenhängen hinaus. So werden basierend auf quantitativen Maßzahlen qualitative Einschätzungen getroffen. Aufbauend auf theoretisch-konzeptionellen Überlegungen können Zusammenhänge zwischen den strukturellen Merkmalen und dem Output oder Outcome einer bestimmten Struktur erklärt werden. Auch bieten sich Vernetzungsanalysen dann an, wenn es bei einer zu evaluierenden Intervention explizit um Fragen der Veränderung von Vernetzung und Austausch geht. Die Netzwerk-SWOT-Analyse ist ein hilfreicher Rahmen zur umfassenden Auswertung.

## Innovationsfähigkeitsindikator

(Vortrag: Leo Wangler/iit, Christiane Kerlen/Kerlen Evaluation)

Zahlreiche Veröffentlichungen bestätigen, dass Innovationsfähigkeit eine wichtige Schlüsselkompetenz für Unternehmen ist. Doch wie lässt sich Innovationsfähigkeit messen und welche Rolle spielt sie für den Erfolg von Förderprogrammen? In diesem Vortrag wurde das Konzept des Innovationsfähigkeitsindikators präsentiert und die Wirkung der dadurch messbaren Innovationsfähigkeit von Fördernehmern mithilfe von Daten aus der Evaluation des Luftfahrtprogramms (LuFo) auf verschiedene Innovations-Output-Variablen geprüft. Der Innovationsfähigkeitsindikator setzt sich dabei aus einer Reihe von Indikatoren für Human-, Struktur- und Beziehungskapital zusammen. Das Humankapital umfasst Faktoren, die beeinflussen, inwiefern wichtige externe Entwicklungen in Wissenschaft und Wirtschaft überhaupt wahrgenommen und verarbeitet werden können (Absorptionsfähigkeit). Das Strukturkapital wird an den Strukturen wie der Aufbauorganisation und den technischen Infrastrukturen sowie Arbeits- und Kommunikationsprozessen im Unternehmen gemessen. Das Beziehungskapital ergibt sich u.a. aus Beziehungen zu externen Partnern entlang der Wertschöpfungskette, der eigenen Branche, Forschungseinrichtungen und Bildungsanbietern. In einem ersten Schritt wurde an einem Wirkmodell der FuE-Förderung erläutert, wie eine erfolgreiche Förderprogrammteilnahme zur Steigerung der Innovationsfähigkeit führen kann. Diese Annahme konnte unter anderem mithilfe eines Vergleichs von Eingangs- und Abschlussbefragungsdaten der Fördernehmer des Luftfahrtprogramms bestätigt werden. Im zweiten Teil der Präsentation wurde anhand verschiedener Regressionsmodelle der Einfluss der anfänglichen Werte für den Innovationsindikator und dessen Teilaspekte auf die Anzahl der entstandenen Patente und Publikationen bei Unternehmen, auf die Ausbildung wissenschaftlicher Nachwuchskräfte, auf Produkt- und Prozessinnovationen und die Geschwindigkeit der Technologieentwicklung (TRL-Sprünge) geprüft. Für Patente, Ausbildungszahlen und Produkt- und Prozessinnovationen konnten die Hypothesen bestätigt werden, dass eine hohe Innovationsfähigkeit zu Beginn des Förderprojekts zu besseren Resultaten führt. Im Gegensatz dazu konnte die Anzahl der Publikationen und die Veränderung des TRL-Levels nicht mit der anfänglichen Innovationsfähigkeit erklärt werden. Dies wohl auch deshalb, weil die Entwicklungszyklen in der Luftfahrtbranche besonders lang sind.

## Evaluation der österreichischen Weltraumstrategie

(Vortrag: Peter Kaufmann/KMU Forschung Austria, Guido Zinke/iit)

Der letzte Vortrag bot einen spannenden Einblick in die laufende Evaluation der österreichischen Weltraumstrategie, die zum einen das BMVIT-Weltraumengagement im internationalen Kontext sowie auch das nationale Weltraumprogramm ASAP umfasst. Neben der Design-, Prozess- und Wirkungsanalyse ist auch die Ausarbeitung von konkreten Vorschlägen für die kommende Weltraumstrategie Österreichs ein

zentraler Bestandteil des Evaluationsauftrags. Über eine Literaturrecherche und Experteninterviews wurden in einem ersten Schritt die Strategien wichtiger und auch vergleichbarer Player auf europäischer Ebene sowie die Strategien der ESA und der EU analysiert. Die Erkenntnisse aus diesem Strategiescreening flossen dann in zentrale Thesen zu der Entwicklung und den Potenzialen Österreichs in den Bereichen Wirtschaft, Technologie und Politik ein. Die Thesen bildeten die Grundlage für eine zweitstufige Delphi-Befragung, die sich momentan in der zweiten Phase befindet. Nachdem eine hohe Zahl an internationalen Weltraumexpert(inn)en aus ganz unterschiedlichen Bereichen die Potenziale für Österreich eingeschätzt hat, werden zentrale Befragungsergebnisse in einem zweiten Schritt einer Auswahl an Expert(inn)en zurückgespielt, um fundierte Schlussfolgerungen ziehen zu können. Die Befragung strukturiert sich dabei entlang der relevanten Technologie-, Anwendungs- und Forschungsbereiche des Weltraumsektors differenziert nach Up-, Mid- und Downstream-Sektoren. Das breite Spektrum an Methoden umfasst zudem bibliometrische Analysen und eine Auswertung der Förderdaten.