

Research with Impact

Christoph M. Schmidt

Forschung und Politikberatung am RWI

RWI Position #54, 5. Dezember 2013

ZUSAMMENFASSUNG

Politische Entscheidungsträger wünschen sich oft klare Empfehlungen auf der Basis absoluter Gewissheit darüber, welche Auswirkungen eine neue Weichenstellung haben wird. Diese Sicherheit ist in einer unsicheren Welt aber nicht möglich. Dennoch müssen Entscheidungsträger auf wissenschaftlich gestützte Handlungsempfehlungen nicht verzichten. Denn die evidenzbasierte Politikberatung kann glaubwürdige Annäherungen an die tatsächlichen Folgen politischer Entscheidungen bereitstellen. Was genau ist mit „evidenzbasierter Politikberatung“ gemeint? Institute wie das RWI erarbeiten ihre Forschungsergebnisse durch die transparente Zusammenführung von theoretischen Annahmen und empirischen Belegen. Sie sprechen auf dieser Basis zwar klare Handlungsempfehlungen aus, dokumentieren aber auch die Beschränkungen ihrer Analysen.

Diese *RWI Position* zeigt Möglichkeiten und Grenzen der evidenzbasierten Politikberatung anhand zahlreicher Beispiele aus der Arbeit des RWI auf. Damit illustriert sie zugleich den neuen Claim des Instituts: „Research with Impact“. Der Zusatz „with Impact“ ist dabei in einem dreifachen Sinne zu verstehen: Erstens besitzen fast alle Arbeiten des RWI aufgrund ihrer praktischen Relevanz auch politisches Wirkungspotenzial („policy impact“). Zweitens genügt seine Forschung höchsten Standards und wird daher regelmäßig in sehr guten Journals (mit hohem „impact factor“) veröffentlicht. Drittens ist die empirische Wirkungsanalyse („impact evaluation“) eine bereichsübergreifende Kernkompetenz des Instituts und daher auch oft die zentrale Herausforderung in den Forschungs- und Beratungsprojekten des RWI.

AUTOR



Christoph M. Schmidt

Präsident des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung (RWI) in Essen, Vorsitzender des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung (SVR), Professor für Wirtschaftspolitik und angewandte Ökonometrie an der Ruhr-Universität Bochum

SUMMARY

Policy-makers often wish for clear advice on the basis of absolute certitude about the consequences of new policy measures. Yet, such certainty is not feasible in an uncertain world. Nevertheless, decision makers do not have to do without recommendations that are derived from sound scientific analysis. The tools of evidence-based policy advice provide credible assessments on the likely outcomes of political decisions. What exactly is meant by “evidence-based” policy advice? Institutes such as the RWI reach their findings through transparent combinations of theory-based assumptions and empirical data. On this basis, they derive clear policy recommendations but also document the restrictions of their analyses.

The present *RWI Position* policy paper highlights opportunities and restrictions of evidence-based policy advice by examples from different fields of research at RWI. Thus, this paper illustrates the institute’s new claim “Research with Impact”. The addition “with impact” is to be understood in a threefold sense: first, virtually all projects at RWI are of practical relevance and possess a high degree of potential policy impact. Second, research at RWI fulfills highest academic standards and research results are regularly published in renowned academic journals (i.e. with a high “impact factor”). Finally, the empirical establishment of cause-effect relationships, also called “impact evaluation”, is a trans-divisional core competency of RWI and, therefore, often the central challenge in research and policy consulting projects at the institute.

RÉSUMÉ

Les décideurs politiques demandent souvent des recommandations claires qui s'appuient sur une certitude absolue des futures conséquences d'une mesure politique. Mais dans un monde incertain, une telle certitude n'est pas possible. Néanmoins, les décideurs ne doivent pas se passer des recommandations basées sur l'évidence scientifique. Les études empiriques permettent aujourd'hui l'approximation crédible des conséquences réelles d'une mesure politique. Instituts comme le RWI obtiennent leur résultats par la combinaison des assumptions théoriques avec des données empiriques. Sur cette base, ils dérivent des recommandations d'actions, alors qu'ils soulignent en même temps les limitations de leurs analyses.

La présente *RWI Position* policy paper démontre les possibilités et, au même titre, les limitations de la recherche empirique par des exemples issus du travail au RWI. Ainsi, le nouveau slogan de l'institut – «Research with Impact» – est illustré du même coup. Le complément «with Impact» s'entend dans un triple sens: Primo, presque tous les projets du RWI possèdent une importance pratique et un haut niveau d'impact potentiel dans le domaine politique. Secundo, la recherche du RWI satisfait les plus hauts standards scientifiques et apparaît régulièrement dans les revues spécialisées prestigieuses. Finalement, l'établissement crédible des relations entre causes et effets, c'est-à-dire l'évaluation des impacts, est la compétence globale de l'institut à travers des spécialités et, par conséquent, souvent le défi central dans les projets de recherche et de conseil politique au sein du RWI.

IMPRESSUM

Herausgeber

Rheinisch-Westfälisches Institut
für Wirtschaftsforschung (RWI)

Hohenzollernstr. 1-3
45128 Essen
Fon: +49 (0) 201-8149-0

Büro Berlin

Invalidenstr. 112
10115 Berlin

ISBN 978-3-86788-513-3

Alle Rechte vorbehalten. Essen 2013

www.rwi-essen.de/positionen

Schriftleitung

Prof. Dr. Christoph M. Schmidt

Redaktion und Ansprechpartner

Nils aus dem Moore (verantwortlich)
Fon: +49 (0) 30-2 0215 98-15
nils.ausdemmoore@rwi-essen.de

Katja Fels
positionen@rwi-essen.de

Lektorat

Claudia Schmiedchen

Layout und Gestaltung

Julica Bracht & Daniela Schwindt

1. EVIDENZBASIERTE POLITIKBERATUNG¹

Die im Jahr 2008 ausgebrochene globale Wirtschaftskrise hat uns eine bittere Erkenntnis in Erinnerung gerufen: Wir leben in einer Welt voller Unsicherheiten, voller neuer Entwicklungen, voller Unwägbarkeiten. Politisches Handeln kann in dieser Welt niemals mit absoluter Sicherheit zum Erfolg führen. Somit kann ich in der politischen Sphäre einen Reformvorschlag oder eine wirtschaftspolitische Entscheidung nur dann als wohlbegründet ansehen, wenn die handelnde Politik mit dieser Unsicherheit offen umgeht, sich also zu ihrer begrenzten Wirkungsmacht bekennt.

Statt absolute Gewissheiten über die Effekte von Maßnahmen oder veränderter Weichenstellungen vorzutäuschen, muss eine gute Politik stets mögliche Erfolge und Risiken einer Maßnahme sorgfältig gegeneinander abwägen, insbesondere im Hinblick auf ihre – meist nur grob anzunähernden – Eintrittswahrscheinlichkeiten. Möglicherweise ergibt sich dabei sogar eine Situation, in der man zwar mit hoher Wahrscheinlichkeit erwarten kann, einen sehr großen Erfolg zu haben, in der aber die Risiken ebenfalls sehr bedeutend sind.

Statt absolute Gewissheiten über die Effekte einer veränderten Weichenstellung vorzutäuschen, muss eine gute Politik stets mögliche Erfolge und Risiken einer Maßnahme sorgfältig abwägen.

Ein klares, wenngleich extremes, Beispiel für eine derartige Situation aus der Sphäre der Energiepolitik ist die – für Deutschland bereits entschiedene – Frage des Ausstiegs aus der Nukleartechnologie. Hier ließe sich seitens der Befürworter der Kernenergie argumentieren, dass mit hoher Wahrscheinlichkeit ein Atomunfall ausgeschlossen werden kann – nicht jedoch mit absoluter Gewissheit. Das Kernargument der Gegner würde hingegen lauten, dass, wenn aber doch ein Atomunfall passiert, daraus eine riesige Katastrophe erwachsen würde, die in ihrer Dimension die Gesellschaft über alle Maßen belasten würde. Es ist offensichtlich, dass ein politischer Entscheidungsträger schlecht beraten wäre, seine Entscheidung lediglich auf den – geschätzten, die genauen Eintrittswahrscheinlichkeiten sind ja unbekannt – Mittelwert zu stützen.²

¹ Diese RWI Position basiert auf einem Vortrag vom 21. Mai 2013 in der gemeinsamen Vorlesungsreihe „Wissenschaftliche Politikberatung“ der Berlin-Brandenburgischen Akademie der Wissenschaften und der Leibniz-Gemeinschaft. Ich bedanke mich herzlich für konstruktive Kommentare bei Nils aus dem Moore, Philipp Breidenbach, Katja Fels, Wim Kösters und Lina Zwick. – Meine Ausführungen in diesem Abschnitt lehnen sich eng an die folgenden Beiträge an: Schmidt, Ch. M. (2007), *Policy Evaluation and Economic Policy Advice*, *Advances in Statistical Analysis* 91: 379-389, und Schmidt, Ch. M. (2009), *Wirtschaftswissenschaft und Politikberatung in Deutschland – Bedeutung, Möglichkeiten und Grenzen der Kausalanalyse*, in: Belke, A., H.-H. Kotz, S. Paul und Ch. M. Schmidt (Hrsg.), *Wirtschaftspolitik im Zeichen europäischer Integration: Festschrift für Wim Kösters*, *RWI Schriften* 83: 21-38.

² Im Juni 2011 beschloss die Bundesregierung, gestützt auf die Empfehlungen der ad hoc gegründeten Ethik-Kommission „Sichere Energieversorgung“, den endgültigen Ausstieg aus der Kernenergie bis zum Jahr 2022. Die Ethik-Kommission selbst sah ihren Ausgangspunkt in der Einsicht, dass es sich bei der Frage des Atomausstiegs um eine gesellschaftliche Wertentscheidung handelt und dass die vorangegangene Reaktorkatastrophe in Japan, wenngleich sie die mit der Nutzung der Kernenergie verbundenen Risiken nicht verändert hat, eine Neubewertung dieser Risiken erforderlich machte.

Eine aufgeklärte Gesellschaft kann und sollte es auf Dauer nicht akzeptieren, immer nur politische Kost vorgesetzt zu bekommen, die absolute Gewissheiten vortäuscht. Denn dann würden Reformvorschläge und politische Entscheidungen eher auf Glaubenssätzen beruhen als auf nüchterner Einsicht und belastbarer Evidenz.³ Doch oft genug ist es in der politischen Realität noch immer gängige Praxis. Denn wäre eine Evidenzbasierung wirklich bei allen Akteuren – Entscheidungsträgern wie Beratern – als die unabdingbare Grundlage politischer Entscheidungen anerkannt, dann wäre die politische Rhetorik durchweg weniger apodiktisch ausgeprägt, als sie es in der Praxis ist.⁴

Ein gutes Beispiel hierfür ist sicherlich die hoch emotionale Debatte um die Einführung eines flächendeckenden Mindestlohns, bei der Befürworter und Gegner grundlegend verschiedene Einschätzungen über die zu erwartenden Konsequenzen äußern. Das Spektrum der Erwartungen reicht dabei von hohen Zuwächsen an Kaufkraft und Wohlstand bis zu hunderttausendfachem Arbeitsplatzverlust und einem daraus resultierenden Desaster für die öffentlichen Finanzen. Da es für einen flächendeckenden Mindestlohn in unserer Volkswirtschaft keinen historischen Präzedenzfall gibt und man in der Analyse daher immer zu Analogieschlüssen aus ähnlichen Kontexten gezwungen ist, kann die empirische Basis eigentlich keine apodiktischen Aussagen stützen. Im Gegensatz zur politischen Debatte stehen ernsthafte wirtschaftswissenschaftliche Studien zu der mit ihren Schlussfolgerungen verbundenen Unsicherheit.

Dabei hindert die verbleibende Unsicherheit bei den Schlussfolgerungen der empirischen Analyse den Wissenschaftler keineswegs daran, in der wirtschaftspolitischen Debatte eine klare Position zu beziehen. So ist eine mögliche, ja sogar naheliegende Schlussfolgerung aus den bislang vorliegenden Studien, dass die Einführung eines flächendeckenden Mindestlohns mit hohen Risiken für die Arbeitnehmer verbunden ist. Man muss vom Ausbleiben negativer Beschäftigungseffekte demnach schon sehr überzeugt sein, um einen hohen Mindestlohn zu fordern.⁵ Die Anforderung an die wirtschaftspolitische Beratung ist also nicht, bei verbleibenden Unsicherheiten über die vermuteten Entwicklungen und

3 Herausragende Beiträge zur evidenzbasierten Politikberatung sind vor allem Manski, C.F. (1995), *Identification Problems in the Social Sciences*, Harvard University Press, und Manski, C.F. (2013), *Public Policy in an Uncertain World: Analysis and Decisions*, Harvard University Press. Grundlegend für Deutschland sind insbesondere Fitzenberger, B. und R. Hujer (2002), *Stand und Perspektiven der Evaluation der Aktiven Arbeitsmarktpolitik in Deutschland*, *Perspektiven der Wirtschaftspolitik* 3: 139-158, sowie Lechner, M. (2002), *Eine wirkungsorientierte aktive Arbeitsmarktpolitik in Deutschland und der Schweiz: Eine Vision – zwei Realitäten*, *Perspektiven der Wirtschaftspolitik* 2: 159-174.

4 Für die Stärkung der evidenzbasierten Politikberatung argumentierte erst kürzlich auch: Wissenschaftlicher Beirat beim BMWi (2013), *Evaluierung wirtschaftspolitischer Fördermaßnahmen als Element einer evidenzbasierten Wirtschaftspolitik*, Gutachten des Wissenschaftlichen Beirats beim Bundesministerium für Wirtschaft und Technologie, Berlin.

5 Ein Beispiel für eine ökonometrische Studie zu diesem Thema, welche zur Begründung des eigenen Studiendesigns die empirischen Fallstricke bei der Abschätzung der zu vermutenden Auswirkungen eines künftigen Mindestlohns diskutiert, ist Bauer, Th. K., J. Kluve, S. Schaffner und Ch. M. Schmidt (2009), *Fiscal Effects of Minimum Wages: An Analysis for Germany*, *German Economic Review* 10(2): 224-242. Das Thema wurde in diesem Jahr vom *German Economic Review* erneut aufgegriffen, mit ambivalenten Ergebnissen: Paloyo, A., S. Schaffner und Ch. M. Schmidt (2013), *Special Issue on the Economic Effects of Minimum Wages in Germany: Editorial*, *German Economic Review* 14 (3): 255-257.

Effekte zu schweigen. Stattdessen muss man immer offen sagen, was in der gegebenen Situation tatsächlich auf Wissen beruht und was lediglich vermutet oder nur aus Analogieschlüssen abgeleitet wird.

Wird im Sinne einer Evidenzbasierung die unsichere Welt anstelle absoluter Gewissheit als Rahmen für politisches Handeln akzeptiert, dann zieht dies unweigerlich auch hohe Anforderungen an die politischen Entscheidungsträger nach sich. Erstens müssen sie lernen damit zu leben, dass es bei wirtschaftspolitischen Fragestellungen nahezu unausweichlich eine gewisse Vielstimmigkeit der Analyse geben muss. Dass dies offenbar schwer fällt, sieht man schon allein an den Reaktionen zu Konjunkturprognosen, welche die Institute regelmäßig veröffentlichen: Politik und Presse fragen immer nach einer genauen Zahl und wollen nur sehr ungern eine Bandbreite akzeptieren, aber oft ist die Empörung groß, wenn die Vorhersagen nicht punktgenau eintreffen. Doch natürlich ist das überhaupt nicht zu leisten.⁶

Gerade für „große“ Fragen gibt es oft nur wenig Präzedenzfälle, so dass man für ihre Analyse bestenfalls Analogieschlüsse heranziehen kann und die verbleibende Unsicherheit groß ist.

Zweitens muss man einsehen, dass es gerade die „großen“ Fragen sind, für die es nur wenig Präzedenzfälle gibt, so dass man für ihre Analyse bestenfalls plausible Analogieschlüsse heranziehen kann und die verbleibende Unsicherheit über die Schlussfolgerungen groß ist. Derartige wirtschaftspolitische Weichenstellungen betreffen in der Regel entweder Eingriffe allumfassender Natur, weisen also eine erhebliche Reichweite auf, oder sie betreten konzeptionelles Neuland, so dass es schwer fallen muss, im bisherigen Erfahrungsschatz zum Funktionieren unserer Volkswirtschaft empirische Belege zu ihrer Wirkung zu sammeln. Je umfassender oder radikaler also die zu begründende Reform, umso mehr Sorgfalt muss die Politik auf die Diskussion der Grenzen von wissenschaftlicher Erkenntnis und eigener Wirkmächtigkeit legen.

Betrachten wir beispielsweise die Eurokrise. Niemand weiß mit absoluter Gewissheit, welche Konsequenzen die im Raum stehenden Vorschläge nach sich zögen, etwa der Schuldentilgungspakt des Sachverständigenrates⁷, würde die europäische Politik ihnen folgen. Man kann Vermutungen formulieren, man kann Plausibilitätsüberlegungen anstellen, man kann aus der Erkenntnis über Entscheidungssituationen in vergleichbaren Kontexten möglicherweise etwas lernen – aber einen Präzedenzfall auswerten, das kann

⁶ Eine Einführung in diese Diskussion geben Döhrn, R. und Ch. M. Schmidt (2005), *Kein Stein der Weisen – Prognosen erfordern eine intelligente Nutzung*, RWI Positionen #5.

⁷ Das Konzept des Schuldentilgungspakts wird ausführlich diskutiert in: Schmidt, Ch. M. und B. Weigert (2013), *Weathering the crisis and beyond: perspectives for the Euro area*, *International Tax and Public Finance* 20(4): 564-595.

man nicht. Daraus folgt insbesondere, dass die Politik ihre Entscheidungen so treffen muss, dass sie diese inhärente Unsicherheit berücksichtigt. Denn das Beste, was sie machen kann, ist die Wahrscheinlichkeit auf Erfolg zu erhöhen, nicht den garantierten Erfolg zu sichern.

Das bedeutet aber nicht, dass das, was wir als empirisch arbeitende Ökonomen und evidenzorientierte Politikberater tun, nutzlos ist. Das würde so mancher ab und an gerne schlussfolgern, frustriert über eine derartige Vielstimmigkeit und frustriert über das Ausbleiben einer größeren Gewissheit versprechenden Aussage eines wirtschaftspolitischen Beraters. Verlangen kann der zu Beratende sicherlich, dass sein Berater ihm deutlich sagt, was er an seiner Stelle tun würde, aber dann auch die mit dieser Entscheidung verbundenen Risiken nennt. Die Entscheidung bleibt aber immer noch die des politischen Entscheidungsträgers; wir können ihn nicht aus seiner demokratisch legitimierten Verantwortung befreien. Das ist aber keine nutzlose, sondern die einzig richtige Beratung, weil jeder andere Berater, der größere Gewissheit vortäuscht, diese wirklich nur vortäuschen kann.

Die evidenzbasierte Politikberatung zeigt offen auf, in welchem Ausmaß die Schlussfolgerungen durch empirische Belege abgesichert sind. Dabei wird transparent dargelegt, wie der vorliegende Erfahrungsschatz mit theoretischen Annahmen zusammengeführt wird.

Die evidenzbasierte Politikberatung ist somit eine Antwort auf das Problem, dass man in einer unsicheren Welt notwendigerweise Schwierigkeiten hat, Erkenntnisse darüber zu sammeln, wie etwas werden wird, wenn man nicht eingreift, bzw. wie es sich gestalten wird, wenn man eine bestimmte Maßnahme durchführt. Daher ist die evidenzbasierte Politikberatung weder ein Schlagwort, noch ist sie allein durch den intensiven Gebrauch von Daten gekennzeichnet. Es handelt sich vielmehr um ein Vorgehen, das offen aufzeigt, in welchem Ausmaß die Schlussfolgerungen, die man anbietet, durch empirische Belege abgesichert sind. Dabei wird ganz transparent darlegt, wie der vorliegende Erfahrungsschatz – also die Daten, auf die man in der Analyse zurückgreifen, und die Präzedenzfälle, auf die man im Analogieschluss verweisen kann – mit theoretischen Annahmen zusammengeführt wird, um diese Schlussfolgerungen zu ziehen.

Zu diesem Vorgehen gehört es insbesondere, offen darzulegen, welche Fehler bei dieser Zusammenführung möglicherweise passieren und welche Fehlschlüsse daraus erwachsen könnten. Eine solche Vorgehensweise ist umso erfolgreicher und gewinnt an Überzeugungskraft, wenn der Berater auf einen großen Erfahrungsschatz zurückgreifen kann. Viele Präzedenzfälle, die alle in die gleiche Richtung laufen, geben relativ viel Zuversicht,

dass die Schlussfolgerung in diesem Fall eine gute ist. Natürlich sind solche Analysen besonders überzeugend, bei denen man durch Variation der Art und Weise, in der man Daten und Theorien verknüpft, nachweist, dass die Sensitivität der grundsätzlichen Ergebnisse vergleichsweise klein ist: Wenn ich bei einer derartigen Sensitivitätsanalyse eine kleine Änderung in den Annahmen vollziehe, dann aber keine große Veränderungen in den Aussagen verzeichnen muss, so hilft dies natürlich, weitere Zuversicht in die Robustheit und die damit verbundene Güte der Schlussfolgerungen zu gewinnen.

2. AUFGABEN DER (ÖKONOMISCHEN) POLITIKBERATUNG

Beschäftigen sich Ökonomen eigentlich in hinreichender Art und Weise mit den „echten“ Problemen unserer Zeit? Ich bin fest davon überzeugt, dass die Antwort grundsätzlich „ja“ lautet. Da sind zum einen die „großen“ wirtschaftspolitischen Fragen unserer Zeit: Wie soll der gemeinsame Währungsraum in Europa ausgestaltet, wie die aktuelle Krise bewältigt werden?⁸ Welche energiepolitischen Weichenstellungen machen eine erfolgreiche Umsetzung des großen gesellschaftlichen Vorhabens „Energiewende“ überhaupt erst möglich?⁹ Was macht menschliche Wohlfahrt aus und wie kann man Wohlstand und Fortschritt sachgerecht statistisch erfassen?¹⁰ Welche Probleme wirft der anstehende demographische Wandel für das Wirtschaftswachstum und die Tragfähigkeit der sozialen Sicherungssysteme auf?¹¹

Die ökonomische Analyse kann darüber hinaus eine Fülle von Einsichten herausarbeiten, die ein weites Spektrum der schier unendlichen Facetten des menschlichen Lebens und der menschlichen Gemeinschaft betreffen, an die man aber beim Stichwort „Wirtschaftswissenschaften“ wohl nicht sofort denkt. Dazu gehören beispielsweise Themen wie ehrenamtliche Tätigkeiten, wie Fremdenfeindlichkeit und Antisemitismus oder die Bestimmungsgrößen des Umgangs mit Suchtmitteln.¹² Diese Beiträge sind der Ökonomik möglich, weil sie sich stets auf einige Grundprinzipien besinnt – die unvermeidliche

⁸ Siehe die vorangehende Fußnote 7.

⁹ Siehe beispielhaft Hübner, M., Ch. M. Schmidt und B. Weigert (2012), *Energiepolitik: Erfolgreiche Energiewende nur im europäischen Kontext*, *Perspektiven der Wirtschaftspolitik* 13(4): 286–307.

¹⁰ Mit diesem Thema befasste sich jüngst die Enquete-Kommission des Deutschen Bundestages „Wachstum, Wohlstand, Lebensqualität“. Ein aktueller Beitrag zu dieser in der ökonomisch-statistischen Literatur sehr lebendigen Debatte ist Schmidt, Ch. M. und N. aus dem Moore (2013), *Wohltandsmessung heute: Statistische Grundlagen einer umfassenden gesellschaftlichen Debatte*, in: Theurl, Th. (Hrsg.), *Akzeptanzprobleme der Marktwirtschaft: Ursachen und wirtschaftspolitische Konsequenzen*, Duncker & Humblot: 13–32.

¹¹Für eine aktuelle Übersicht zu diesem Thema siehe beispielhaft Schmidt, Ch. M. (2012), *Die demografische Herausforderung – ein aktueller Überblick*, in: Hüther, M. und G. Naeyege (Hrsg.), *Demografiepolitik – Herausforderungen und Handlungsfelder*. Springer: 71–95.

¹² Beispiele für derartige Beiträge sind Bauer, Th. K., J. Bredtmann und Ch. M. Schmidt (2013), *Time vs. Money – The Supply of Voluntary Labor and Charitable Donations Across Europe*, *European Journal of Political Economy* 32: 80–94; Fertig, M. und Ch. M. Schmidt (2011), *Attitudes towards Foreigners and Jews in Germany: Identifying the Determinants of Xenophobia in a Large Opinion Survey*, *Review of Economics of the Household* 9: 99–128; Tauchmann, H., S. Lenz, T. Requate und Ch. M. Schmidt (2013), *Tobacco and Alcohol: Complements or Substitutes? A Structural Model Approach to Insufficient Price Variation in Individual-Level Data*, *Empirical Economics* 45(1): 539–566.

Existenz von Knappheiten („Budgetrestriktionen“) und die Einsicht, dass ein gegebener Zustand tendenziell eher eine Balance (ein „Gleichgewicht“) zwischen näher zu analysierenden widerstreitenden Kräften und kein Zufallsprodukt darstellt.

Doch auch die Einsicht, dass das menschliche Leben nicht im Hier und Jetzt entspringt, sondern nur durch eine Rückbesinnung auf seine historischen Entwicklungen verstanden werden kann, beschäftigt die Ökonomik. Daher zählen zu unserem Arbeitsspektrum natürlich auch einerseits wirtschaftshistorische und andererseits ideengeschichtliche Beiträge.¹³ Dabei leitet den Forscher der Dreiklang „Erkennen, Erwarten, Ergründen“.

So ist es häufig schwer genug, überhaupt einen Überblick über die aktuelle Lage zu gewinnen. Die Verdichtung der Fülle an vorhandenen Fakten, die Informationsreduktion, ist eine Hauptaufgabe der Statistik. Eine „Erbsenzählerei“, wie sie häufig im Rahmen von Sportreportagen transportiert wird („Für die Statistiker unter Ihnen: Das Tor fiel in der 63. Minute.“), liefert hingegen ein falsches Bild der Statistik. Denn der einzelne Datenpunkt interessiert den Statistiker gar nicht. Was zählt, ist etwa die Frage, ob typischerweise gegen eine bestimmte Mannschaft in der zweiten Halbzeit mehr Tore fallen als in der ersten, da man dann Rückschlüsse auf deren Trainingszustand ziehen kann.

Ein ökonomisches Beispiel der Informationsreduktion liefern die diversen Berichte des RWI zu den Innovationsanstrengungen im Land Nordrhein-Westfalen im Vergleich zu Bayern und Baden-Württemberg. Dieser Quervergleich ist für die Forschungs- und Innovationspolitik von großer Bedeutung, denn nur im Kontrast lässt sich gegebenenfalls ein Handlungsbedarf erkennen.¹⁴ Aber wenn man diesen Kontrast nicht auf der bereits durch Informationsreduktion verdichteten Ebene von Ländern oder Regionen suchte, sondern durch die Einzelbetrachtung forschender Unternehmen, dann würde die Analyse in der Fülle der Detailinformationen erstickt werden.

Die zweite Aufgabe empirischer Wirtschaftsforschung besteht darin, darzulegen, welche künftigen Entwicklungen in verschiedenen Wirtschaftsbereichen zu erwarten sind. Das klassische Beispiel für diese Leistung der empirischen Wirtschaftsforschung sind die regelmäßigen Konjunkturprognosen der Leibniz-Institute im Rahmen der sogenannten Gemeinschaftsdiagnose.¹⁵ Prognosen sind in der Praxis weit mehr als akribische Fortschreibung am „aktuellen Rand“ des Geschehens: Prognosen sind, so trivial das klingen mag, zum einen deswegen nötig, weil die Zukunft nicht so ist wie die Vergangenheit; sie

¹³ Beispielhaft für diese Art von Beiträgen siehe: Guinnane, T.W., J. Streb, M. Frondel und Ch. M. Schmidt (2011), *Krank oder Simulant? Das Identifikationsproblem der Krankenkassen aus wirtschaftshistorischer Perspektive*, *Perspektiven der Wirtschaftspolitik* 12(4): 413-429; aus dem Moore, N. und Ch. M. Schmidt (2013), *On the Shoulders of Giants: Tracing Back the Intellectual Sources of the Current Debate on “GDP and Beyond” to the 19th Century*, *Jahrbücher für Nationalökonomie und Statistik* 233(3): 266-290.

¹⁴ RWI und Stifterverband Wissenschaftsstatistik (2010), *Innovationsbericht 2009 – Zur Leistungsfähigkeit des Landes Nordrhein-Westfalen in Wissenschaft, Forschung und Technologie*, *RWI Projektberichte*.

¹⁵ Siehe beispielhaft Döhrn, R. (2013), *Deutsche Konjunktur erholt sich – Ergebnisse der Gemeinschaftsdiagnose vom Frühjahr 2013*, *Wirtschaftsdienst* 93(5): 323-327.

sind zum anderen möglich, weil sich meist keine sprunghaften Unterschiede zwischen Vergangenheit, Gegenwart und Zukunft ergeben. Die Kunst der Prognose liegt also darin, diejenigen Elemente der Stabilität zu identifizieren, etwa stetige Entwicklungspfade zentraler Bestimmungsfaktoren, die es erlauben, Aussagen über künftige Entwicklungen abzuleiten.¹⁶

Gewissermaßen die Krönung unseres Tuns sind, drittens, Kausalanalysen. So steht insbesondere im Fokus, ob eine gegebene wirtschaftspolitische Weichenstellung tatsächlich die angestrebte Veränderung der Umstände bewegt hat oder ob es für die beobachteten Entwicklungen andere Ursachen gegeben hat: Warum, so ist zu fragen, hat sich hier ein Erfolg oder ein Misserfolg ergeben? War dies der Fall, obwohl oder etwa weil der Staat eingegriffen hat?

Die große Herausforderung bei der Frage nach Ursache und Wirkung in unserem Tätigkeitsfeld ist – und das ist zugleich das Schöne an der empirischen Wirtschaftsforschung – dass das Objekt unseres Erkenntnisinteresses einen eigenen Willen besitzt: Wir untersuchen die Entscheidungen und Handlungen von Menschen, also nicht die Geschehnisse von willenlosen Zellen oder landwirtschaftlich beackerten Feldern. Diesen eigenen Willen können unsere Studienobjekte glücklicherweise in unserer freiheitlichen Gesellschaft ausleben, indem sie selbst großen Einfluss darauf nehmen, ihre Geschehnisse zu steuern, allen wirtschaftspolitischen Einfluss- und Steuerungsversuchen zum Trotz. Aus diesem Grunde ist in der Praxis häufig die Erkenntnis einer Korrelation, die tendenzielle Gleichzeitigkeit von beobachteten Phänomenen, nicht notwendigerweise mit dem Nachweis von Kausalität, also einer Ursache-Wirkung-Beziehung, gleichzusetzen.¹⁷

Das übergreifende Ziel unserer Bemühungen als empirische Wirtschaftsforscher und wirtschaftspolitische Berater ist es natürlich, den Entscheidungsträgern in der Wirtschaftspolitik zu helfen, effektive und effiziente Weichenstellungen zu beschließen, also „gute“ Wirtschaftspolitik zu betreiben. Es geht dabei nicht allein darum, den Akteuren in Politik und Verwaltung – oder gar allein der jeweils aktuellen Regierung – Hinweise zu geben. Sondern unsere Aufgabe ist es auch, die breite Öffentlichkeit über die für sie nach dem genannten Dreiklang, „Erkennen, Erwarten, Ergründen“ relevanten Erkenntnisse der empirischen Wirtschaftsforschung aufzuklären und dadurch den öffentlichen Diskurs zu beflügeln.

¹⁶ Diese Argumentation wird näher ausgeführt in Fertig, M. und Ch. M. Schmidt (2002), *Reading Tea Leaves or Science? Forecasting for Beginners*, *AStA – Allgemeines Statistisches Archiv* 86: 459-467.

¹⁷ Ein Beispiel für die in den vergangenen Jahren nahezu explosionsartig gewachsene Literatur über unterschiedliche konzeptionelle Ansätze zur Evaluation von wirtschaftspolitischen Eingriffen und deren jeweilige Anforderungen ist, hier speziell auf die Umweltökonomik zugeschnitten, Frondel, M. und Ch. M. Schmidt (2005), *Evaluating Environmental Programs: The Perspective of Modern Evaluation Research*, *Ecological Economics* 55(4): 515-526.

Die dezidiert „angewandte“ Forschung am RWI zielt darauf ab, Beiträge dazu zu leisten, die Geschicke der Menschen und der gesamten Gesellschaft zu verbessern. Sie ist notwendigerweise empirisch ausgerichtet, da die Ökonomik allein auf dem Reißbrett nicht funktionieren kann. Insbesondere für Leibniz-Institute wie das RWI ist die Frage der Relevanz und des Anwendungsbezugs daher maßgeblich für die Auswahl der Forschungsfragen. Theoriebildung oder abstrakte Beiträge zur Verbesserung des empirischen Methodenkastens gehören nicht zu unseren vorrangigen Aufgaben. Nur wenn der Analyse der eigentlichen inhaltlich-ökonomischen Sachfrage eine Datensituation zugrunde liegt, die so komplex ist, dass sie einen eigenständig auf sie zugeschnittenen methodischen Beitrag verlangt, sollten wir uns dieser Herausforderung stellen.

Gute wirtschaftspolitische Beratung lässt sich vom aktuellen Stand der wissenschaftlichen Erkenntnis nicht entkoppeln. Veröffentlichungen in referierten Zeitschriften sind als Qualitätskontrolle und -nachweis unerlässlich.

In unserer täglichen Arbeit sind dabei natürlich die beiden eng miteinander verwobenen Aufgaben der empirischen Wirtschaftsforschung und der wirtschaftspolitischen Beratung in einem ständigen Balanceakt miteinander in Einklang zu bringen. Aus drei Gründen müssen wir erstklassige Forschungsleistungen erbringen und können nicht nur Beratungsleistungen anbieten: Gute wirtschaftspolitische Beratung lässt sich, erstens, vom aktuellen Stand der wissenschaftlichen Erkenntnis nicht entkoppeln. Für uns sind, zweitens, Veröffentlichungen in referierten Fachzeitschriften als Qualitätskontrolle und -nachweis unerlässlich, weil eine vergleichbare Qualitätskontrolle im Bereich der wirtschaftspolitischen Beratung nicht existiert.

Und wir müssen, drittens, für talentierte Nachwuchsforscher(innen) attraktiv bleiben. Denn nicht nur müssen wir das eigene Talentreservoir pflegen, um selbst leistungsfähig zu bleiben. Wir haben auch die wichtige Aufgabe, komplementär zur eher theoretisch ausgeprägten Ausbildung der Universitäten talentierte Nachwuchskräfte durch angewandte Arbeit am Objekt, also bei der praktischen empirischen Wirtschaftsforschung, zu fertigen Forschern auszubilden.

Bei aller Konkurrenz zur universitären Forschung, die wir somit entfalten müssen, macht allein das Beispiel der deskriptiven Forschung deutlich, wie sinnvoll es ist, außeruniversitäre Forschungseinrichtungen wie die Leibniz-Gemeinschaft öffentlich zu fördern: Wer würde sich in der Forschung, wäre diese rein universitär organisiert, überhaupt deskriptiven Forschungsfragen widmen? Wohl fast niemand, denn junge Wirtschaftswissenschaftler wissen genau, dass sie ihre Berufungschancen nur dadurch erhöhen können, dass sie Publikationen vorzuweisen haben, die im einschlägigen Handelsblatt-Ranking

hoch eingeordnet sind. Rein deskriptive Studien – wie der oben angesprochene „Innovationsbericht“ – gehören trotz ihrer hohen wirtschaftspolitischen Bedeutung sicherlich nicht dazu.

In der Konsequenz ist die strikte Arbeitsteilung verschwunden, die es früher sicherlich gegeben hat. Eine zugespitzte Karikatur der Vergangenheit sah folgendermaßen aus: Die Datensammlung wird von den „Erbsenzählern“ in den statistischen Ämtern durchgeführt, Deskription und Fortschreibungen am aktuellen Rand machen die fleißigen, aber konzeptionell schlichten angewandten Forscher in den außeruniversitären Instituten, aber die „richtige Forschung“ findet an den Universitäten statt. Die aktuelle Szene sieht stattdessen so aus, dass die Sammlung und Bereitstellung von einzelbasierten Individualdaten von angewandten Forschern häufig auch direkt durchgeführt wird. So haben wir z.B. am RWI ein „Forschungsdatenzentrum Ruhr“, das nicht zuletzt als Serviceeinrichtung für externe Forscher derartige Daten bereitstellt.

Insbesondere sind Evaluationsstudien mittlerweile ein Markenzeichen der außeruniversitären Forschung geworden, unter anderem sicherlich deshalb, weil dabei die eigene, auf die konkrete Anwendungsfrage maßgeschneiderte Datensammlung nahezu unabdingbar geworden ist. Aber ein weiterer Grund ist mindestens genauso wichtig: Nur in einem außeruniversitären Forschungsinstitut werden sich Gruppen von empirisch arbeitenden Ökonomen in einer Anzahl zusammenfinden, welche die Anforderungen einer „kritischen Masse“ erfüllt. Dies macht „echte“ Teamarbeit möglich, die den ständigen Austausch von Ideen erlaubt und in der Zusammenarbeit von jungen und erfahrenen Forschern die Weitergabe von Problemlösungen fördert. In den wirtschaftswissenschaftlichen Fakultäten deutscher Universitäten gibt es hingegen nur selten mehr als ein Lehrstuhlteam für (angewandte) Ökonometrie.

3. GRUNDLAGEN ÖKONOMISCHER POLITIKBERATUNG: ANNAHMEN UND IDENTIFIKATIONSPROBLEME

Empirische Forschung besteht grundsätzlich in einem Versuch, zunächst durch die Beobachtung und dann durch die – in geeigneter Form informationsverdichtende – Auswertung einer Stichprobe einen Erkenntnisgewinn über die Eigenschaften der jeweiligen Grundgesamtheit zu erlangen. Dabei gehört die Diskussion darüber, welche Aspekte des Phänomens sie als gegeben annehmen können und welche noch zu erobern, zu erahnen, zu erarbeiten sind, unweigerlich an den Anfang einer jeden empirischen Studie.

Denn nur das, von dem man wenigstens erahnt, dass man darüber Erkenntnisse sammeln kann, lässt sich überhaupt empirisch erforschen. Die grundlegenden Annahmen darüber, was ich im Laufe der empirischen Arbeit nicht mehr in Frage stelle, werden in der

modernen Forschung üblicherweise als „Identifikationsannahmen“ bezeichnet, denn erst sie erlauben es, bislang unbekannte Eigenschaften der Grundgesamtheit, deren Existenz ich erahne, zu erkennen.

Eine erste Perspektive des statistischen Denkens betrifft die Wahrscheinlichkeitstheorie. Ein Wahrscheinlichkeitstheoretiker weiß genau, wie das Wahrscheinlichkeitsgesetz aussieht, das die vorliegende Datensituation prägt. Aus dieser Kenntnis heraus kann es ihm gelingen, Eigenschaften von repräsentativen Stichproben aus der Grundgesamtheit zu charakterisieren, beispielsweise Erwartungswerte einschlägiger Anteilsgrößen.

Gedanklich steht vor ihm, um es ganz plastisch zu machen, eine Urne, die mit Kugeln gleicher Größe gefüllt ist. Er weiß nun, woher auch immer, dass die eine Hälfte der Kugeln rot und die andere Hälfte blau ist. Die Fragen, zu deren Beantwortung er jetzt befähigt ist, beziehen sich auf die Eigenschaften einer jeden aus dieser Grundgesamtheit zu ziehenden Stichprobe. Wenn er jetzt beispielsweise eine Stichprobe zusammenstellt, indem er mit Zurücklegen zehnmal eine Kugel aus der Urne holt und die Farbe der jeweils gezogenen Kugel notiert: Welchen Anteil an roten Kugeln dürften wir in der Stichprobe erwarten? Unter den genannten Bedingungen darf man erwarten, dass die Hälfte der gezogenen Kugel rot sein wird.

Grundlegende Annahmen, die ich im Verlauf der empirischen Arbeit nicht mehr in Frage stelle, werden als „Identifikationsannahmen“ bezeichnet.

Aus der Sicht eines Außenstehenden, der sich fragt, wie die Grundgesamtheit beschaffen ist, weiß der Wahrscheinlichkeitstheoretiker also bereits alles, was es zu erkennen gilt. Ein (schließender) Statistiker nimmt hingegen eine zweite, gerade entgegengesetzte Perspektive ein. Er kennt gerade nicht die Grundgesamtheit, sondern ihm liegt lediglich eine daraus gezogene Stichprobe vor. In unserem Beispiel mag die eine Hälfte der Kugeln in seiner Stichprobe rot, die andere Hälfte blau sein. Typischerweise ist der Statistiker nun aufgefordert, eine Aussage darüber zu treffen, wie wohl die Grundgesamtheit der Kugeln in der Urne beschaffen ist, aus der diese Stichprobe gezogen wurde.

Es liegt nahe, dass der Analogieschluss von der Stichprobe zur Grundgesamtheit eine ebenso gute empirische Strategie darstellt, wie es aus der oben diskutierten ersten Perspektive der Analogieschluss von der Grundgesamtheit hin zur Stichprobe war: Als vermutlichen Anteil der roten Kugeln in der Grundgesamtheit dürfte der Statistiker hier die Hälfte nennen und damit ziemlich richtig liegen. Aber er hat – vielleicht sogar, ohne darüber weiter nachzudenken – bei dieser Aussage implizit unterstellt, dass es in der Grundgesamtheit in der Tat lediglich zwei verschiedene Farben gibt.

Der Ökonometriker muss eine dritte Perspektive einnehmen, die ihn in eine viel heiklere Situation bringt: Weder kennt er das seinen Beobachtungen, seinem „Erfahrungsschatz“, zugrundeliegende Wahrscheinlichkeitsgesetz. Noch weiß er verlässlich, ob die ihm zur Verfügung stehende Stichprobe die Grundgesamtheit wirklich repräsentiert. Im Beispiel weiß er insbesondere nicht, ob es neben den roten und den blauen nicht vielleicht auch noch schwarze Kugeln gibt, die aber nur sehr selten oder nur unter bestimmten Bedingungen Einzug in die Stichprobe finden können. Erst wenn dies ausgeschlossen werden kann, lässt sich der oben geschilderte Analogieschluss ziehen. Trifft der Ökonometriker eingangs seiner statistischen Analyse die (Identifikations-)Annahme, dass es in der Grundgesamtheit keine schwarzen Kugeln gibt, dann ist er auf einmal in der Rolle des Statistikers. Es greifen die bewährten Einsichten der schließenden Statistik.

In der Praxis der empirischen Wirtschaftsforschung ist aber genau diese Anforderung entscheidend: Ohne Identifikationsannahmen geht es nicht, können keine Erkenntnisse gesammelt, keine Schlussfolgerungen abgeleitet werden! Wenn alles im Fluss ist und meine Stichprobe somit keine stabile Grundgesamtheit repräsentieren kann, dann kann ich auch deren Eigenschaften nicht durch empirische Arbeit – Schlussfolgerungen, die aus der Stichprobe abgeleitet werden – erhellen. Dies gilt für die Beschreibung des Ist genauso wie für Prognosen und Kausalanalysen: Ohne einen Grundkanon an – auf die jeweilige Situation zugeschnittenen – Identifikationsannahmen kann es keine empirische Analyse geben, in keinem der drei oben geschilderten Aufgabengebiete.

In der angewandten Wirtschaftsforschung und der damit verbundenen wirtschaftspolitischen Beratung ist dies natürlich derjenige Aspekt der Forschungsleistung, bei dem die Kompetenz des entsprechenden Forschers und die Güte des von ihm entworfenen Studiendesigns die größte Wirkung aufweisen. Insbesondere mag jemand, der die Frage nach den angemessenen Identifikationsannahmen am Anfang leichtfertig stellt, schon von vornherein völlig falsch liegen, völlig unabhängig davon, wie viel statistisch-ökonomisches „Muskelspiel“ er danach noch zum Einsatz bringt.

Somit entscheidet nicht die mathematische Eleganz des Vorgehens oder der vom Computer geleistete Rechenaufwand, sondern das Verknüpfen von institutionellen und historischen Kenntnissen mit ökonomischem Sachverstand über die Relevanz und Überzeugungskraft der abgeleiteten Schlussfolgerungen. Dass das mathematische und ökonometrische Rüstzeug als intellektuelle Grundausstattung eines modernen Ökonomen unbedingt dazu gehört, ist dadurch in keiner Weise in Frage gestellt. Es reicht aber nicht aus, wenn man gute wirtschaftspolitische Beratung anbieten will.

So mag bereits die ganz primitive Frage, ob ich davon ausgehen kann, dass die vorliegende Stichprobe nach wie vor die relevante Grundgesamtheit repräsentiert, die Güte der empirisch abgeleiteten Schlussfolgerungen vorausbestimmen. Ein aktuelles Beispiel: Nicht wenige empirische Arbeitsmarktforscher hatten zu Beginn der großen Wirtschaftskrise unter dem Eindruck der massiv eingebrochenen Wirtschaftsleistung prognostiziert,

dass die Arbeitslosigkeit in Deutschland in der Folge massiv ansteigen dürfte.¹⁸ Was sie dabei nicht bedacht hatten, aber auch nicht aus empirischen Beobachtungen hätten ableiten können: Nicht zuletzt aufgrund der vielfältigen Reformen hatte sich im vergangenen Jahrzehnt die Art und Weise, wie Angebot und Nachfrage auf dem Arbeitsmarkt zusammengeführt werden, verändert; und als ein großer Nachfrageschock die deutsche Wirtschaft traf, nutzten die deutschen Unternehmen in der Annahme, die Erholung stünde bald bevor, jede ihnen gebotene Möglichkeit, an ihren Belegschaften festzuhalten, der deutsche Arbeitsmarkt blieb erstaunlich (sic!) stabil.

Das, was man sich in den Anfängen der ökonometrischen Forschung in Deutschland nicht klargemacht hatte ist, dass man diese ganzen Systeme mit vielen nichtlinearen Gleichungen und den vielen sie charakterisierenden Parametern ja irgendwie trennscharf aus dem Erfahrungsschatz ableiten muss. Dieses Vorhaben ist aber, so stellte es sich heraus, unheimlich schwer umzusetzen: Denn als empirischer Forscher habe ich eigentlich immer recht wenig Daten im Vergleich zur Vielfalt der Fragen, die ich gerne beantworten möchte.

Nicht mathematische Eleganz oder der von Computern geleistete Rechenaufwand entscheidet über die Relevanz und Überzeugungskraft von Schlussfolgerungen, sondern das Verknüpfen von institutionellen und historischen Kenntnissen mit ökonomischem Sachverstand.

Die schließende Statistik ist dieser Problematik bereits einen wichtigen Schritt entrückt: Die Antwort auf die Frage nach dem konkreten Wert der unbekannt Parameter einer ansonsten bekannten Wahrscheinlichkeitsverteilung kann ich aus einer Stichprobe relativ problemlos ableiten. Dabei wächst die Trennschärfe, die „Präzision“, meiner Aussagen mit der Anzahl der Beobachtungen. Und bei steigender Trennschärfe kann ich zudem Hypothesen besser testen, denn ich kann dann umso besser die Frage stellen, ob die Ausprägungen der Stichprobe weitgehend im Einklang mit diesen Hypothesen sind.

Bei gegebener Anzahl der Beobachtungen kann ich wiederum an Trennschärfe gewinnen, wenn ich striktere Identifikationsannahmen treffe. In der ökonometrischen Praxis geht es dabei häufig um die Frage, ob ich gewillt bin, die Restriktionen eines theoretischen Strukturmodells als gegeben anzunehmen oder nicht. Denn wenn ich durch eine erhöhte Komplexität der Identifikationsannahmen „Freiheitsgrade“ freischaufle, wächst die ausgewiesene Präzision der geschätzten Modellparameter – aber auch das Risiko, damit völlig falsch zu liegen.

¹⁸ Döhrn, R., G. Barabas, H. Gebhart, T. Kitlinski, T. Schmidt und S. Vosen (2009), *Die wirtschaftliche Entwicklung im Inland: Weltweite Finanzkrise verursacht historischen Konjunkturbruch*, RWI Konjunkturberichte 60(1): 35-89.

Bis zu einem gewissen Grad werde ich in der praktischen Arbeit den Restriktionsgrad der Identifikationsannahmen durch statistische Tests auf den Prüfstand stellen können. Doch auf ihre Plausibilität testen kann ich nur solche Identifikationsannahmen, die ich bei der Analyse zum Zwecke der Gewinnung zusätzlicher Freiheitsgrade getroffen habe. Dabei wird es immer einen Grundkanon an Identifikationsannahmen geben, ohne den ich überhaupt kein Ergebnis ableiten kann. Diese Annahmen können dann naturgemäß nicht getestet werden.

Ich komme also unweigerlich in jedem Falle wieder auf die Grundeinsicht zurück: Ich muss die Fragen kennen, wenn ich aus den Daten konkrete Antworten ableiten will. Also wird es mir insbesondere, wenn man bei dem Beispiel mit roten, blauen und schwarzen Kugeln bleiben will, nie einfallen, aufgrund der Stichprobe zu testen, ob es in der Grundgesamtheit nicht doch einen gewissen Anteil von schwarzen Kugeln gibt, wenn ich meine Kenntnis über die Zusammenhänge nur aus der vorliegenden Stichprobe ableite, weil schwarze Kugeln in meiner Stichprobe nicht enthalten sind.

Da demnach (zumindest die unbedingt nötigen) Identifikationsannahmen niemals getestet werden können, gibt es am Anfang jeder empirischen Studie ein Abwägungsproblem: Zum einen kann ein empirischer Wirtschaftsforscher möglichst restriktive Modelle anbieten und dann unter der Annahme, diese Modelle wären in der Tat zutreffend, vergleichsweise präzise Schätzungen und damit verbunden klare Schlussfolgerungen anbieten. Allerdings besteht dabei die Gefahr aufgrund zu restriktiver Annahmen komplett falsche Schlüsse aus diesen Modellen zu ziehen. Andererseits gibt es immer die Möglichkeit, dem Modell diese Restriktionen nicht in dieser Schärfe aufzuerlegen und damit zu riskieren, dass die Parameter, die man zusätzlich noch aus der Stichprobe ableiten muss, relativ grobe Annäherungen an die tatsächlichen Werte darstellen. Dann besteht jedoch die Gefahr, dass man aus der Analyse keine brauchbaren Schlussfolgerungen ableiten kann.

Um es sehr zuzuspitzen: Wenn ich sage, die Antwort auf meine Frage nach dem Anteil der roten Kugeln in der Grundgesamtheit ist immer „ein Viertel rot“, dann leite ich damit eine unglaublich trennscharfe, wenngleich recht nutzlose Aussage ab. Denn angesichts jeder erdenklichen Stichprobe, selbst wenn drei Viertel oder mehr der gezogenen Kugeln rot sind, würde ich „ein Viertel rot“ schlussfolgern. Mit dieser Strategie wird man in der Regel falsch liegen, aber den oberflächlichen Betrachter mag die hohe Trennschärfe der Aussage dennoch beeindrucken. Es dürfte aber für die Wirtschaftspolitik nicht zielführend sein, ganz präzise völlig falsch zu liegen.

4. PROBLEMLÖSUNG: IDENTIFIKATIONSSTRATEGIEN DER KAUSALANALYSE

Dass erhebliche Identifikationsprobleme zu bewältigen sind, gilt vor allem bei der für gute Wirtschaftspolitik entscheidenden Frage nach Ursache und Wirkung, denn diese Frage ist unweigerlich eine Frage nach dem Nicht-Beobachtbaren, dem „was wäre gewesen, wenn“. Die Fakten zu beobachten ist immer vergleichsweise einfach. Wir mögen

beispielsweise einen Mindestlohn der Höhe W und eine mit diesem einhergehende Beschäftigung der Höhe L beobachten. Die entscheidende Frage ist dann aber, wie hoch die Beschäftigung gewesen wäre, wenn wir diesen Mindestlohn nicht gesetzt hätten. Diese sogenannte „kontrafaktische“ Situation, also jenes Ergebnis, das sich ergeben hätte, wenn die wirtschaftspolitische Weichenstellung anders erfolgt wäre, als es tatsächlich der Fall war, gilt es zu beschreiben. Doch diese kontrafaktische Situation ist niemals zu beobachten, handelt es sich doch um eine hypothetische Situation, die gerade nicht „faktisch“ vorliegt.

Für die Identifikation kausaler Effekte ist die Konstruktion der „kontrafaktischen“ Situation der entscheidende Schritt.

Jedes Mal, wenn ich Vergleichssituationen zum Kontrast heranziehe, um die nicht-beobachtbare kontrafaktische Situation anzunähern, muss ich mich fragen, ob es sich wirklich um eine überzeugende Vergleichssituation handelt. Im Idealfall ist hier alles genau so wie in der faktischen Situation, bis auf den einen Umstand, dass diesmal die auf ihre Wirkung zu untersuchende wirtschaftspolitische Maßnahme, also etwa der Mindestlohn, nicht eingeführt wird. Wenn diese Bedingung tatsächlich erfüllt ist, dann kann ich den Kontrast der Beschäftigungsvolumina nehmen und diese Differenz als Effekt der Maßnahme deuten. Wenn ich diesen Nachweis der Vergleichbarkeit aber nicht führen kann, dann kann ich auch keine kausalanalytischen Schlussfolgerungen ziehen.

In der Praxis der empirischen Wirtschaftsforschung muss ich diese Vergleichssituation immer aus dem Reservoir meiner vorliegenden Beobachtungen konstruieren. So könnte man sich beispielsweise vorstellen, dass man die Wirkungen eines Weiterbildungsprogramms für Langzeitarbeitslose evaluieren möchte.¹⁹ In der Stichprobe kann man Langzeitarbeitslose beobachten, die ein bestimmtes Trainingsprogramm wahrnehmen, und andere Langzeitarbeitslose, die dieses Programm nicht in Anspruch nehmen. Nun mögen diejenigen Langzeitarbeitslosen, die dieses Trainingsprogramm durchlaufen haben, vielleicht besonders erfolgreich bei den folgenden Bemühungen um eine Wiederaufnahme von Beschäftigung sein, während die anderen im Durchschnitt nicht so erfolgreich sind.

Doch diejenigen Langzeitarbeitslosen, die das Trainingsprogramm nicht in Anspruch genommen haben, sind lediglich das Reservoir an möglichen Vergleichsbeobachtungen. Würde ich einfach den Mittelwert ihres Erfolgs zum Maßstab nehmen, dann würde ich unterschlagen, dass die Willigen, die Erfolgshungrigen, also die Langzeitarbeitslosen, die es einfach schaffen wollen, aus ihrer misslichen Lage herauszukommen, eher die Maßnahme ergreifen werden, während bei denjenigen, die das Reservoir der Vergleichsbe-

¹⁹ *Wie man in der praktischen Arbeit die Evaluierung von arbeitsmarktpolitischen Maßnahmen so angehen kann, dass die arbeitsmarktpolitische Entscheidungsfindung konstruktiv unterstützt wird, zeigen anschaulich Fertig, M., J. Kluve, Ch. M. Schmidt, H. Apel, W. Friedrich und H. Haegele (2004), Die Hartz-Gesetze zur Arbeitsmarktpolitik – Ein umfassendes Evaluationskonzept, RWI Schriften 74, Duncker & Humblot.*

obachtungen bilden, einige möglicherweise gar nicht den Lebensmut oder das nötige Durchhaltevermögen aufweisen, um dies zu schaffen. Insgesamt dürften diese also ohne weiteres keine zwei völlig vergleichbaren Gruppen darstellen.

Aus diesem Reservoir muss ich mir in einer empirischen Kausalanalyse jetzt diejenigen Beobachtungen heraussuchen, die so beschaffen sind, dass man sie tatsächlich als Vergleichsgruppe verwenden kann. Das ist keineswegs einfach, weil viele von den Dingen, die wir dabei zur Auswahl aus dem Reservoir eigentlich gerne heranziehen wollen, wie etwa der persönliche Erfolgshunger, nicht für den empirischen Wirtschaftsforscher beobachtbar sind.²⁰ Somit ist die Frage, ob entweder die Teilnahme oder der Erfolg der Maßnahme durch in meinem Datenreservoir nicht zu beobachtende Faktoren beeinflusst ist, entscheidend für die Glaubwürdigkeit jeder nicht-experimentellen Evaluationsstudie.

Ein großer Teil der Anstrengungen in der empirischen Wirtschaftsforschung ist darauf ausgerichtet, in der Realität jene Situationen zu erkennen, die den Anforderungen eines „natürlichen Experiments“ möglichst gut entsprechen.

In einem kontrollierten Zufalls-Experiment kann ich dieses Problem recht gut in den Griff bekommen. Denn dann weise ich die betreffende Maßnahme in einem randomisierten Verfahren ganz zufällig zu. Das klassische Beispiel wäre die klinische Studie, bei der man den Probanden ohne weitere Anschauung der Person zufällig ausgewählt entweder Aspirin oder Kohletabletten verabreicht, die beide gleich aussehen, ihnen aber nicht ver-rät, was sie einnehmen. Dann sagen uns die Kontraste im durchschnittlichen Kopfschmerz nach der Einnahme, ob das Aspirin wirklich Wirkung zeigt.

Aber so eine Situation liegt in der empirischen Wirtschaftsforschung nur in ganz seltenen Fällen vor, denn bei unseren Fragen spielen solche Dinge wie Motivation oder Durchhaltevermögen meist eine große Rolle, also Aspekte, die man aus analytischer Sicht als „unbeobachtete Heterogenität“ bezeichnet. Wie geht man damit in der praktischen Arbeit um? Ein mittlerweile häufig eingesetzter Kniff, mit dem sich empirische Arbeitsmarktforscher in ihrer praktischen Arbeit über dieses Problem fehlender Beobachtungen hinwegzuhelfen suchen, ist der Rückgriff auf die persönlichen Biographien der Langzeitarbeitslosen: Ähneln sie einander darin, dann ähneln sie sich vermutlich auch in ihren nicht-beobacht-

²⁰ Da empirische Forschung immer davon lebt, sich von der Anekdote zu lösen und durch Informationsverdichtung gemeinsame Muster herauszuarbeiten, kann sie nur auf quantitativ aufbereitete Datensätze zugreifen, die den Einzelfall standardisiert abbilden. Die Ergebnisse im praktischen Wirtschaftsleben, die diese Forschung untersucht, sind jedoch das Ergebnis individuellen Austauschs zwischen den betreffenden Akteuren: Ein potenzieller Arbeitgeber mag hingegen vielleicht schon erkennen, ob er von Angesicht zu Angesicht einem leistungsbereiten Kandidaten gegenüber sitzt oder nicht. Doch dies kann kein Anlass sein, Evaluation per Fallstudie zu betreiben, denn diese weisen noch größere Schwächen auf: die Überbetonung des zufällig herausgegriffenen Einzelfalls und den fehlenden Maßstab für „typische“ Ausprägungen.

baren Eigenschaften – so die Identifikationsannahme.²¹ Nun haben wir gesehen, dass bei der Identifikation kausaler Effekte die Konstruktion der kontrafaktischen Situation das Entscheidende ist. Beide grundsätzlichen Vorgehensweisen, das sozialwissenschaftliche Experiment und nicht-experimentelle Methoden, haben ihre eigenen Fallstricke und werfen ihre eigenen Herausforderungen auf.

Ein Beispiel, an dem das Vorgehen bei experimentellen Studien erläutert werden kann, betrifft die Gesundheitsökonomik. Es handelt sich dabei um ein sozialwissenschaftliches Feldexperiment zum Thema „Fettleibigkeit“, das vom Pakt für Forschung und Innovation finanziert worden ist.²² Dabei wurde Patienten, die eine Reha-Klinik verlassen haben und gleichzeitig einen Bodymass-Index (BMI) über dem Wert 30 aufwiesen, eine Prämie für die Leistung ausgelobt, ihr Gewicht in einem vorgegebenen Zeitintervall auf einen erheblich reduzierten Wert zu führen. Doch nur ein – zufällig ausgewählter – Teil dieser Patienten erhielt ein derartiges Angebot von einigen Hundert Euro, ein anderer Teil hörte sich lediglich den mit dem Verlassen der Reha-Klinik fälligen Appell an. Im Ergebnis haben diejenigen Patienten, die einen zusätzlichen Anreiz durch die Prämie erfahren haben, tatsächlich einen größeren Rückgang in ihrem BMI zu verzeichnen.

Hier ist die Kausalitätsaussage über die Wirkung der angebotenen Prämie sehr überzeugend, denn die Patienten haben nicht selbst gesagt „Ich will ein Abnehmprogramm durchführen, gebt mir deswegen mal bitte 300 Euro!“, sondern dies war ein Anreiz, der ohne ihr Zutun, in der Sprache der empirischen Wirtschaftsforschung „exogen“, gesetzt worden ist. Doch nichts im Leben ist ganz umsonst: Die Überzeugungskraft der Schlussfolgerungen der Studie lebt davon, dass es uns gelungen ist, mit unseren statistischen Anpassungsmaßnahmen zu verhindern, dass die Austritte aus der Studie von denjenigen Patienten, die keine Prämie in Aussicht gestellt bekamen, den ursprünglich durch die Randomisierung erzeugten Idealcharakter der Vergleichsgruppe wieder zu verwischen drohten. Daran sieht man wieder einmal eindrucksvoll, wie sehr der oben angesprochene „eigene Wille“ unserer Studienobjekte unsere Fähigkeiten zur überzeugenden empirischen Analyse beeinflusst.

Ein Beispiel, an dem ich die Vorgehensweise bei nicht-experimentellen Studien erläutern will, beschäftigt sich mit der Frage, wie sich das Ableisten des Wehrdiensts langfristig auf den Arbeitsmarkterfolg auswirkt. In diesem Falle ist ein experimentelles Vorgehen von vornherein ausgeschlossen. In unserer Studie konnten wir auf einen großen Datenschatz zugreifen, der uns retrospektiv die Lebensleistung von Arbeitnehmern darlegt –

²¹ Ein Beispiel für die Suche nach „statistischen Zwillingen“, von denen der eine die zu untersuchende Maßnahme erfährt, während sich der andere im Reservoir der Nicht-Teilnehmer befindet (die „Zwillingspaare“ wurden gekoppelt über ihre jeweiligen früheren Erfahrungen am Arbeitsmarkt), ist: Kluve, J., H. Lehmann und Ch. M. Schmidt (2008), *Disentangling Treatment Effects of Active Labor Market Policies: The Role of Labor Force Status Sequences*, *Labour Economics* 15(6): 1270-1295.

²² Augurzky, B., Th. K. Bauer, A. R. Reichert, Ch. M. Schmidt und H. Tauchmann (2012), *Does Money Burn Fat? – Evidence from a Randomized Experiment*, *Ruhr Economic Papers* #368.

und jeweils einen Hinweis darauf gibt, ob sie als junge Männer in der Bundeswehr gedient hatten oder nicht.²³ Dies ist eine ganz klar nicht-experimentelle Situation, denn im Nachhinein wissen wir relativ gut, ob die jungen Männer ihr Arbeitsleben erfolgreich oder nicht erfolgreich bestritten haben. Eine Zuweisung zum Wehrdienst kann zum Zwecke des Studiendesigns jedoch nicht mehr beeinflusst werden.

Damit können wir als Forscher leider auch nicht mehr verhindern, dass sich die Teilnehmer am Wehrdienst und die Nicht-Teilnehmer systematisch durch – im Datensatz nicht aufgeführte – Eigenschaften unterscheiden, etwa in ihrer physischen Robustheit. Allerdings können wir, und das ist der Schlüssel zu unserem Beitrag, unter bestimmten Bedingungen die nicht-experimentellen Daten so auswerten, als handele es sich bei ihnen um das Resultat eines experimentellen Studiendesigns. Dieses Vorhaben ist immer dann in die Tat umsetzbar, wenn man äußere – aus Sicht des Analytikers „exogene“ – Umstände findet, welche die Zuweisung zu der zu bewertenden Maßnahme beeinflusst haben. Diese Umstände charakterisieren sich dadurch, dass sie von den Betroffenen nicht beeinflusst werden konnten aber einen Einfluss auf deren Erfolg am Arbeitsmarkt aufweisen. In solchen Fällen spricht man von einem „natürlichen Experiment“. Mittlerweile ist ein großer Teil unserer intellektuellen Anstrengungen als empirische Wirtschaftsforscher darauf ausgerichtet, derartige Situationen zu erkennen.

Ein Quell derartiger „exogener“ Variationen im Zugang zu Maßnahmen liegt häufig in institutionellen Reformen – davon betroffen zu sein, mag unter Umständen nicht im Einflussbereich der einzelnen „Beobachtungseinheiten“, hier Arbeitnehmer, gelegen haben. Im vorliegenden Fall ist es das konkrete Geburtsdatum, das zumindest für eine kleine Gruppe von Arbeitnehmern: Wer vor dem 30. Juni 1937 geboren wurde, gehörte zu den sogenannten „weißen Jahrgängen“ und wer ab dem 1. Juli 1937 geboren wurde, war im Prinzip in der in den 1950er Jahren gegründeten Bundeswehr wehrpflichtig. Da uns ein Datensatz vorliegt, der selbst bei einer Beschränkung der Analyse auf kurz vor und kurz nach der Schwelle zwischen dem 30. Juni und dem 1. Juli 1937 Geborene noch viele Beobachtungen zur Lebensleistung am Arbeitsmarkt liefert, kann die Variation im Erfolg, die wir an dieser Sprungstelle sehen, wie in einem Experiment zur Ableitung einer Aussage über die Langzeitwirkungen des Wehrdienstes genutzt werden (diese Langzeitwirkung, so hat die Studie ergeben, ist vernachlässigbar). Wie bei allen Studien der empirischen Wirtschaftsforschung gilt auch hier, dass die Überzeugungskraft der Identifikationsannahmen darüber entscheidet, wie sehr man den Schlussfolgerungen der Studie folgen kann. Die entsprechende Identifikationsannahme einer quasi-randomisierten Zuweisung der Wehrpflicht für um dieses Datum herum Geborene dürfte im vorliegenden Falle recht überzeugend sein: Es ist sicherlich ausgeschlossen, dass sich die werdenden Eltern am 30. Juni 1937 gesagt haben „Besser noch ein bisschen mit der Niederkunft warten, weil der Bub dann in 18 Jahren in die Bundeswehr kann.“ Ebenso wird niemand dem Gedan-

²³ Bauer, Th. K., S. Bender, A. Paloyo und Ch. M. Schmidt (2012), *Evaluating the Labor-market Effects of Compulsory Military Service*, *European Economic Review* 56(4): 814-829.

kengang gefolgt sein: „Bloß alles daran setzen, dass die Geburt vor Mitternacht erfolgt, dann bleibt meinem Kind in 18 Jahren die Wehrpflicht erspart.“ Eine derartig überzeugende Begründung für ihre Identifikationsannahmen kann jedoch nicht jede empirische Studie vorweisen.

5. WETTBEWERB IN DER (ÖKONOMISCHEN) POLITIKBERATUNG

Wie sich aus der konzeptionellen Diskussion und den illustrierenden Beispielen erkennen lässt, ist eine empirische Wirtschaftsforschung, welche die allumfassenden Identifikationsprobleme erfolgreich bewältigt und somit als belastbare Basis für die wirtschaftspolitische Beratung dienen kann, eine große Herausforderung. Insbesondere sind es ganz offenbar häufig die „großen“ Fragen, die sich uns stellen, für deren Beantwortung uns nur ein kleiner Erfahrungsschatz – und damit ein dürftiges Reservoir für überzeugende Identifikationsannahmen – zur Verfügung steht. Wenn wir eine üppige Variation in bestimmten Weichenstellungen sehen, aber ansonsten vergleichbare Umstände vorliegen, dann können wir aus dem mit dieser Variation verbundenen Kontrast der Ergebnisse die gesuchten Schlussfolgerungen ziehen. Doch da, wo wir Neuland betreten, wie beispielsweise bei der aktuellen Krise im Euro-Raum, ist es notgedrungen sehr schwer, empirische Evidenz für wirtschaftspolitischen Rat anzubieten.

Aber so ernüchternd dies ist: Wenn wir alles bereits wissen würden, dann gäbe es erst gar kein empirisches Problem, und wenn wir akzeptierten, dass wir überhaupt nichts wissen, könnten wir keinerlei Schlussfolgerungen ziehen. Die angemessenen Identifikationsannahmen müssen also mit dem Wissen um die bei dieser Wahl auftretenden Fallstricke gewählt werden. Die Daten allein werden keine ultimative Sicherheit über die Qualität meines Vorgehens liefern können. Da ich als empirischer Forscher zunächst immer einige grundlegende Annahmen treffen muss, um aus der Stichprobe überhaupt Schlussfolgerungen ableiten zu können, kann ich diese als gegeben angenommenen Aspekte auch nicht mehr hinterfragen. Jede Identifikationsstrategie muss daher argumentativ gut unterfüttert werden.

Wo wir Neuland betreten, wie bei der aktuellen Krise in der Eurozone, ist es sehr schwierig, wirtschaftspolitischen Rat auf der Basis solider empirischer Evidenz anzubieten.

Es wäre aber nicht hilfreich für die Politik, wenn wir nur einfach die Hände in die Luft werfen würden und sagen würden „Na ja, das wissen wir halt alles nicht so genau, wir hoffen auf weitere Daten.“ Die (wirtschaftswissenschaftliche) Politikberatung ist vielmehr aufgerufen, Farbe zu bekennen und Empfehlungen für wirtschaftspolitisches Handeln auszusprechen. Dies darf jedoch nicht verwechselt werden mit dem Vortäuschen einer

Gewissheit darüber, was als Konsequenz dieses Handelns passieren wird, wenngleich viele, die sich in dieser Szene tummeln, vermeintlich solche Gewissheiten anzubieten haben.

Aus meiner Sicht ist es daher ganz entscheidend, dass die Politik sich klar zu der Frage bekennt, ob sie der ernsthaften empirischen Wirtschaftswissenschaft mehr Gehör schenken will, nicht obwohl, sondern weil sie sich offen zur Unsicherheit ihrer Aussagen bekennt. In der wirtschaftspolitischen Realität, das kann ich aus der Praxis eines jahrelangen wirtschaftspolitischen Beraters sagen, ist es zumindest häufig noch so, dass dieser Wille nicht vorhanden ist. Oft werden schöne bunte Charts und klare Aussagen vorgezogen, vielleicht nicht zuletzt deshalb, weil sie helfen können, eine Politik zu „verkaufen“. Doch wahrhaft gute Wirtschaftspolitik sollte stattdessen den Stand der aktuellen Erkenntnis, die begrenzte eigene Wirkungsmacht und die mit jeder Handlungsoption verbundenen Risiken berücksichtigen.

Ebenso, wie wir als Verfechter einer evidenzbasierten Politikberatung im politischen und öffentlichen Diskurs häufig einen schweren Stand haben, so geht es auch denjenigen Politikern, die das Richtige tun und meiner Forderung nach Anerkennung von unsicherer Erkenntnis, begrenzter Wirkungsmacht und verbleibenden Risiken folgen, nicht viel anders. Es ist für die Politik nicht gerade leicht, Populismus mit guten Sachargumenten zu kontern, insbesondere im heute üblichen Format einer Talkshow, in der man vielleicht nur 20 Sekunden Zeit hat, um zu antworten.

Was kann die empirische Wirtschaftsforschung vor diesem Hintergrund tun, um ihren Argumenten und Schlussfolgerungen trotz der beschriebenen und unvermeidbaren Einschränkungen Gehör zu verschaffen? Aus meiner Sicht stehen dabei auf der Angebotsseite des Marktes für wirtschaftspolitische Beratung drei Aspekte im Vordergrund: die Unabhängigkeit der wirtschaftspolitischen Beratung, die Balance zwischen Fokus und Breite im jeweiligen Forschungsprogramm und die Sicherstellung einer hohen fachlichen Qualität.

Die Unabhängigkeit der universitären und außeruniversitären Wirtschaftsforschung von Partikularinteressen ist ein hohes Gut, das die Organisation der wirtschaftspolitischen Beratung in unserem Land in besonderem Maße auszeichnet. Nicht umsonst lautet die korrekte Bezeichnung des Sachverständigenrates „Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung“, nicht „zur Beratung“. Natürlich ist die rein institutionelle Unabhängigkeit nur eine notwendige, keine hinreichende Bedingung für tatsächliche Unabhängigkeit. Die Reputation, wirklich unabhängig zu sein, muss man

sich durch seine Taten immer wieder neu erarbeiten.²⁴ Dass dies nicht jedem gelingt, der Unabhängigkeit für sich in Anspruch nimmt, ändert nichts daran, sondern verstärkt diese Forderung nur.

Im Hinblick auf die Ausgestaltung des wirtschaftspolitischen Beratungsangebots lässt sich zumindest für den Reigen der Leibniz-Institute eine gewisse Vorstellung über das anzustrebende Optimum entwickeln, bei dem Fokus und Breite in geeigneter Form miteinander in Einklang gebracht werden: Der Wettbewerb dieser Einrichtungen kann aus meiner Sicht durch das Modell einer monopolistischen Konkurrenz sehr gut beschrieben werden. Dabei gibt es in umkämpften Marktfeldern einen sehr harten Wettbewerb untereinander. So wird es sich mittlerweile wohl kaum ein Wirtschaftsforschungsinstitut, das etwas auf sich hält, nehmen lassen, im Markt für die Evaluation von arbeitsmarktpolitischen Maßnahmen präsent zu sein. Ebenso versuchen alle, gute Konjunkturprognosen anzubieten.

Die Unabhängigkeit der Wirtschaftsforschung ist ein hohes Gut. Die Reputation, wirklich unabhängig zu sein, muss man sich aber durch Taten immer wieder neu erarbeiten.

Gleichzeitig müssen sich die Institute gewisse Alleinstellungsmerkmale erarbeiten. Ob man dies tatsächlich schafft, wird in den Evaluationsrunden der Leibniz-Gemeinschaft kritisch geprüft, die alle sieben Jahre die Existenz als Mitglied der Leibniz-Gemeinschaft aufs Neue in Frage stellen. Im Falle des RWI sind diese Alleinstellungsmerkmale beispielsweise nachweislich umweltökonomische Fragestellungen, die Evaluation von energiepolitischen Maßnahmen und auch die Gesundheitsökonomik.

Es gibt in der Tat gute Gründe für wirtschaftswissenschaftliche Politikberater, Spitzenforschung anzustreben, und gute Gründe für unsere wissenschaftlichen Aufsichtsgremien, dies zu verlangen. In der universitären Forschung wäre die Antwort auf die Frage nach den idealen Arbeitsbedingungen für ein derartiges Ziel ganz klar: Spezialisierung ist der Schlüssel zum Erfolg, ebenso wie es wichtig ist, die Voraussetzungen dafür zu schaffen, dass Kreativität sich spontan entfalten kann, und zuzulassen, dass Forscher ihre Interes-

²⁴ Es ist offensichtlich, dass man sich mit klaren wirtschaftspolitischen Stellungnahmen, welche den inhaltlichen Positionen von starken Interessengruppen, Parteien oder Regierungen widersprechen, nicht unbedingt Freunde macht. Beispielhaft seien hier genannt: Frondel, M., R. Kambeck und Ch. M. Schmidt (2006), Kohlesubventionen um jeden Preis? – Eine Streitschrift zu den Argumentationslinien des Gesamtverbandes des deutschen Steinkohlenbergbaus, *RWI Materialien* 25; Frondel, M., N. Ritter, Ch. M. Schmidt und C. Vance (2010), Die ökonomischen Wirkungen der Förderung Erneuerbarer Energien: Erfahrungen aus Deutschland, *Zeitschrift für Wirtschaftspolitik* 59(2): 107-133; aus dem Moore, N., R. Kambeck und T. Kasten (2009), Auswirkungen der Steuerprogramme zur Bundestagswahl 2009 – Eine mikrodatenbasierte Analyse der Reformvorschläge von CDU/CSU, SPD und FDP. *RWI Materialien* 55; Gebhardt, H., R. Kambeck und F. Matz (2011), Der NRW-Haushalt für das Jahr 2011: Scheinerfolge bei der Konsolidierung, *RWI Positionen* #47. Interessant an diesen Beispielen aus zwei Bereichen ist insbesondere, dass sich von diesen Stellungnahmen jeweils im Diskurs völlig entgegengesetzt einzuordnende Gruppen erkennbar angegriffen gefühlt haben.

sen ohne weiteres jederzeit neu justieren können. Wenngleich Forschungsinstitute typischerweise eine Abstimmung der in ihrem Team verfolgten Forschungsfragen vornehmen, dann können sie ähnlich hohe Forschungsleistungen nur dann erzielen, wenn das dabei leitende Forschungsprogramm den einzelnen Forschern immer noch hinreichende Freiheiten für die Entfaltung von Kreativität lässt.

Der Auftrag der empirischen Wirtschaftsforschung und der damit verbundenen wirtschaftspolitischen Beratung verpflichtet uns aber – im Gegensatz zur universitär verfassten Forschung – dazu, eine gewisse Breite der Themen anzubieten.²⁵ Wie sollte man etwa zum Spektrum der arbeitsmarktpolitischen Reformen ein ganzheitliches Urteil abgeben wenn man sich schlicht weigerte, für mehr als einen kleinen Ausschnitt davon Interesse zu entwickeln? Somit müssen wir es schaffen, fachlich für ein breites Spektrum von Themen kompetent zu sein. Dementsprechend sind Forschungsinstitute in der Tat gut beraten, durch eine Abstimmung der in ihrem Team verfolgten Forschungsfragen dafür zu sorgen, dass aus dem Spektrum der insgesamt bearbeiteten Themen ein Mehrwert entsteht.

Schwierig ist sicherlich, gute wirtschaftspolitische Beratung zu würdigen und fachliche Qualität als solche trennscharf zu erkennen.²⁶ Denn was vergleichsweise leicht gemessen werden kann, lässt sich auch leichter würdigen, beispielsweise Publikationen in einschlägigen Fachzeitschriften. Die evidenzbasierte Politikberatung entzieht sich jedoch zu einem gewissen Grad der quantitativen Messbarkeit, gerade deswegen, weil sie sich häufig durch die Transparenz über verbleibende Restunsicherheiten und ihre Demut gegenüber ihren Grenzen auszeichnet. Ohne das ernsthafte Bemühen, gute wirtschaftspolitische Beratung auch als solche zu würdigen, dürfte es jedoch schwer werden, langfristig deren hohe fachliche Qualität sicherzustellen.

Hier ist es keineswegs abwegig, eine sprichwörtliche Lanze für die Leibniz-Gemeinschaft zu brechen: Im vergangenen Jahrzehnt hat sich die Evaluierung der Institute von einem Abzählreim (zugespitzt: „Wie viele hochwertige Publikationen gibt es denn?“) mehr und mehr zu einer Suche nach einem runderen Bild gewandelt. Es wird nach wie vor kritisch untersucht, was die Institute wissenschaftlich leisten, aber es wird zumindest im Ansatz gewürdigt, welche Qualität der wirtschaftspolitischen Beratung und der Nachwuchsförderung sie anbieten. Man bemüht sich also durchaus, nicht nur unter der sprichwörtlichen Laterne zu suchen, die durch die einschlägigen Publikationsrankings bereitgestellt wird, sondern versucht eine balanciertere Einschätzung der Qualität von Beratung und Forschung.

²⁵ Schmidt, Ch. M. (2006), *Fokus, Fokus, Fokus? Zur künftigen Rolle der außeruniversitären Wirtschaftsforschungsinstitute*, *AStA – Allgemeines Statistisches Archiv* 90: 617-622.

²⁶ *Meine persönliche Einschätzung der aktuellen Situation findet sich in Schmidt, Ch. M., N. aus dem Moore und M. Themann (2013), Mission Impossible? Zur Verbindung von Politikberatung und „Spitzenforschung“*, *Wirtschaftsdienst* 93(8): 511-515, als Beitrag eines „Zeitgesprächs“ zum Thema „Entwickeln sich wirtschaftswissenschaftliche Forschung und Politikberatung auseinander?“

Was nützte jedoch selbst eine hervorragend aufgestellte evidenzbasierte Politikberatung, wenn die Nachfrageseite für ihre Leistungen – und vor allem für ihre dabei gezeigte Zurückhaltung gegenüber dem Vortäuschen von vermeintlichen Gewissheiten – entweder kein Verständnis entwickelte oder diese Leistungen bewusst nicht abrufen wollte? Was die Frage des Verständnisses für unser Tun in Politik und Verwaltung angeht, so ist meine persönliche Einschätzung, dass einerseits die Komplexität der Fragestellung und die Geschwindigkeit, mit der sich unser Umfeld verändert, in den vergangenen Jahren deutlich gewachsen ist, und dass schritthaltend damit das Verlangen seitens der Nachfrager wirtschaftspolitischer Beratung zugenommen hat, rasche – am besten umgehende – Antworten auf ihre Beratungsfragen zu erhalten.

Dass dies nicht immer ohne Kompromisse bei der Qualität funktionieren kann und dass es gerade für die Aufrechterhaltung der Qualität wichtig ist, den wirtschaftspolitischen Beratern hinreichend Zeit für das Erarbeiten ihrer eigenen Forschungsergebnisse, deren Diskussion mit ihren (internationalen) Fachkollegen und die Publikationen in einschlägigen Fachzeitschriften zu lassen, gerät dabei leicht in Vergessenheit. Angesichts der Vielfalt der wirtschaftspolitischen Probleme und der durch sie aufgeworfenen Forschungsfragen wirkt die in der öffentlichen Diskussion nicht selten gehörte Einschätzung, es gäbe in Deutschland mehr als genug empirische Wirtschaftsforscher und Wirtschaftsforschungsinstitute, absurd.

Andererseits, und das ist die gute Botschaft, hat sich in den vergangenen Jahren durchaus ein aufkeimendes Verständnis für die Notwendigkeit ergeben, empirische Belege zu sammeln, um rationale wirtschaftspolitische Entscheidungen zu treffen. Es entwickelt

Auch der Journalismus – letztlich: jeder mündige Bürger – ist auf ein gewisses Verständnis für die Herausforderungen der evidenzbasierten Politikberatung angewiesen, um seine demokratische Kontrollfunktion wahrnehmen zu können.

sich darüber hinaus sogar ein Verständnis für die Möglichkeiten und Grenzen, die unsere technischen Ansätze und unsere Datensituation mit sich bringen, insbesondere bei der Frage nach Ursachen und Wirkung. Schließlich hat sich auch mehr und mehr ein offener Dialog zwischen Ministerialbürokratie, Politik und Wissenschaft entwickelt, wenn auch noch nicht so, wie ich es mir im Ideal vorstellen würde. Die Nachfrageseite befindet sich aber auf dem Weg, selbst den Qualitätssprüngen nachzueifern, die in den vergangenen Jahren auf der Angebotsseite zu verzeichnen waren.

Allein, es mag nicht hinreichend sein, darum zu wissen, was eigentlich gute, also evidenzbasierte wirtschaftspolitische Beratung ausmacht. Es muss für die Nachfrager in Politik und Verwaltung auch unausweichlich sein, ihr Gehör zu schenken. Dies kann aber

nur dadurch gesichert werden, dass der öffentliche Druck entsprechend wächst, die Ver-
ausgabung von öffentlichen Mitteln und das Stellen von regulierenden Weichen durch
den Nachweis der angestrebten Wirkungen – durch entsprechende empirische Belege,
die den Standards der evidenzbasierten Politikberatung tatsächlich genügen – zu recht-
fertigen. Dies bedeutet aber, dass auch der Journalismus und letztlich auch der mündige
Bürger ohne ein gewisses Verständnis für die Herausforderungen der empirischen Wirt-
schaftsforschung und der evidenzbasierten Politikberatung ihre demokratischen Kont-
rollfunktionen nur sehr eingeschränkt wahrnehmen können. Hier ist sicherlich
noch viel zu tun.²⁷

Aber manchmal hilft auch ein Missverständnis: Wer sich an den Skandal um die damalige
Bundesanstalt für Arbeit zurückerinnert, der sich kurz nach der Jahrtausendwende ab-
spielte, dem fallen sicherlich die, sagen wir es mal vorsichtig, geschönten Vermittlungs-
statistiken ein. Im Endeffekt hat dieser Skandal viel mehr bewegt als nur diese Statistiken
in Frage zu stellen: Aus dieser empörten Diskussion heraus wurde, eigentlich, ohne dass
ein tatsächlicher Zusammenhang damit bestand, in Frage gestellt, was denn die bis dahin
ebenfalls sehr positiv kommentierten Maßnahmen der aktiven Arbeitsmarktpolitik – ein
nahezu vollständiger Mittelabruf wurde bis dahin als großer Erfolg gewertet – tatsächlich
bewirkten. Um diesem öffentlichen Druck zu entsprechen, wurde in die Hartz-Gesetze
zur Reform der aktiven Arbeitsmarktpolitik die Forderung aufgenommen, dass eine wis-
senschaftliche Evaluation durch unabhängige Dritte in Auftrag zu geben sei. Jahrelange
Mahnungen durch ArbeitsökonomInnen hatten das nicht bewirken können, die öffentliche
Empörung schaffte es sofort.

Meine Ausführungen zum Thema „Politikberatung und Evaluationskultur“ hatten insge-
samt einen dezidiert deutschen Fokus. Doch für dieses Thema gilt das Gleiche wie für
viele andere wirtschaftspolitische Themen: Unsere Herausforderungen liegen längst nicht
mehr allein auf der nationalen Ebene und auch die Lösungen für viele unserer Probleme
lassen sich nicht mehr nur dort finden. Auf Dauer wird unser Tun als empirische Wirt-
schaftsforscher und wirtschaftspolitische Berater mehr und mehr eine Internationalisie-
rung, vor allem eine Europäisierung benötigen. Das wird ein Schritt sein, den Angebot
und Nachfrage gemeinsam werden gehen müssen.

²⁷ Ein Lehrbuch, das die statistischen und ökonometrischen Grundlagen der evidenzbasierten Politikberatung vermittelt,
ist Bauer, Th. K., M. Fertig und Ch. M. Schmidt (2009), *Empirische Wirtschaftsforschung. Eine Einführung*, Springer.

ZULETZT ERSCHIENENE RWI POSITIONEN

- #53 Was ist der optimale Mindestlohn?
- #52 Improved Cooking Stoves that End up in Smoke?
- #51 Im Zweifel für die Freiheit: Tarifpluralität ohne Chaos
- #50 Der Markt macht's: Hohe Benzinpreise sind kein Grund für politischen Aktionismus
- #49 Ernsthafte Konsolidierung muss Priorität der neuen NRW-Landesregierung werden
- #48 Das GKV-Versorgungsstrukturgesetz: Richtung richtig, Umsetzung unklar
- #47 Der NRW-Haushalt für das Jahr 2011: Scheinerfolge bei der Konsolidierung
- #46 Der Weg zu nachhaltigen Finanzen: Weniger Soziales, mehr Investitionen
- #45 Die Kosten des Klimaschutzes am Beispiel der Strompreise
- #44 Perspektiven des Gesundheitssektors: Wachstumsmotor oder Milliardengrab?
- #43 Wer trägt den Staat?
- #42 Wirtschaftsleistung, Wertschöpfung und Wachstumspotenziale in Deutschland
- #41 Eine Wirtschaftsregierung für Europa?

www.rwi-essen.de/positionen

RWI – Rheinisch-Westfälisches Institut für Wirtschaftsforschung

Das Rheinisch-Westfälische Institut für Wirtschaftsforschung e.V. (RWI) ist eines der führenden Zentren für wissenschaftliche Forschung und evidenzbasierte Politikberatung in Deutschland. Das Institut ist Mitglied der Leibniz-Gemeinschaft. Es wurde 1926 gegründet und arbeitet seit 1943 in rechtlicher Selbständigkeit. Das RWI stützt seine Arbeiten auf neueste theoretische Konzepte und aktuelle empirische Methoden. In fünf Kompetenzbereichen werden ökonomische Zusammenhänge auf allen Ebenen erforscht – vom Individuum bis zur Weltwirtschaft. Die individuelle Prosperität steht in den Kompetenzbereichen „Arbeitsmärkte, Bildung, Bevölkerung“ sowie „Gesundheit“ im Vordergrund. Unternehmen und Märkte werden in „Unternehmen und Innovation“ sowie „Umwelt und Ressourcen“ untersucht. Der Kompetenzbereich „Wachstum, Konjunktur, Öffentliche Finanzen“ analysiert gesamtwirtschaftliche Fragestellungen. Das „Forschungsdatenzentrum Ruhr am RWI“ (FDZ Ruhr) versorgt die Wissenschaftler mit aktuellsten Methoden und Zahlen. Das RWI veröffentlicht Forschungsergebnisse und Beiträge zur Politikberatung in verschiedenen Publikationsreihen. Weitere Informationen im Internet unter: www.rwi-essen.de