



Christoph M. Schmidt

Politikberatung und Evaluationskultur in Deutschland

In:

Wissenschaftliche Politikberatung im Praxistest / hrsg. von Peter Weingart und Gert G. Wagner unter Mitarbeit von Ute Tintemann. – ISBN: 978-3-95832-046-8

Weilerswist: Velbrück Wissenschaft, 2015

S. 157-187

Persistent Identifier: [urn:nbn:de:kobv:b4-opus4-34595](https://nbn-resolving.org/urn:nbn:de:kobv:b4-opus4-34595)

Die vorliegende Datei wird Ihnen von der Berlin-Brandenburgischen Akademie der Wissenschaften unter einer Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (cc by-nc-sa 4.0) Licence zur Verfügung gestellt.



Christoph M. Schmidt

Politikberatung und Evaluationskultur in Deutschland

Das vorliegende Buch und dieser Beitrag¹ bieten aus der Sicht eines angewandt forschenden und intensiv in der Politikberatung tätigen Ökonomen eine ideale Gelegenheit, um zu erläutern, welche Ausrichtung die (wirtschaftswissenschaftliche) Politikberatung verfolgen sollte, wie sich unsere Tätigkeit ins Wissenschaftssystem insgesamt einfügt und welchen Stand sie in Deutschland zum gegenwärtigen Zeitpunkt erreicht hat. Besonders wichtig ist mir dabei die Diskussion kausalanalytischer Bemühungen: Nur wer Ursachen und Wirkungen verlässlich herausgearbeitet hat, kann Wirksamkeit versprechende wirtschaftspolitische Eingriffe entwerfen. Gleichzeitig ist das Erkennen von Ursachen und Wirkungen im nicht-experimentellen Kontext auch eine notwendige Bedingung, um die Wirksamkeit (wirtschafts-)politischer Maßnahmen im Nachhinein überzeugend zu bewerten.

Ich bin davon überzeugt, dass eine aufgeklärte Gesellschaft sich insgesamt glücklich schätzen könnte, wenn die Wirtschaftspolitik im Hinblick auf ihre tatsächlichen Wirkungen der regelmäßige Gegenstand kritischer und kompetenter Untersuchungen wäre. Ich werde hier argumentieren, dass es zum gegenwärtigen Zeitpunkt noch ein weiter Weg ist, um tatsächlich von einer derartigen »Evaluationskultur« in Deutschland sprechen zu können, wenngleich im vergangenen Jahrzehnt erhebliche Fortschritte bei ihrer Entwicklung zu verzeichnen waren. Dieser Weg muss konsequent weiter besritten werden, um das Potenzial der modernen Wirtschaftsforschung zum Wohle der Gesellschaft noch besser zu nutzen.

Bei meinen Ausführungen greife ich auf zwei verschiedene Quellen in besonders intensiver Weise zurück. Zum einen schöpfe ich aus dem Reservoir einer wachsenden Literatur zur evidenzbasierten Politikberatung, zu der ich in jüngerer Vergangenheit einige Gedanken aus deutscher

1 Überarbeitete Version des Vortrags vom 21. Mai 2013 in der gemeinsamen Vorlesungsreihe »Wissenschaftliche Politikberatung« der Berlin-Brandenburgischen Akademie der Wissenschaften und der Leibniz-Gemeinschaft. Der Vortragsduktus wurde beibehalten. Der Autor bedankt sich herzlich für konstruktive Kommentare bei Andrea Cassel, Nicole Kodat, Nils aus dem Moore, Philipp Breidenbach, Wim Kösters, Gert G. Wagner, Hartmut Westram und Lina Zwick.

Sicht beigetragen habe.² Herausragende Beiträge sind insbesondere Manski, C. F. (1995), *Identification Problems in the Social Sciences* und Manski, C. F. (2013), *Public Policy in an Uncertain World: Analysis and Decisions*. Ein deutschsprachiges Lehrbuch im Geist der evidenzbasierten Politikberatung ist Bauer, Th. K., M. Fertig und Ch. M. Schmidt (2009), *Empirische Wirtschaftsforschung*.

Zum anderen gehört der Anspruch, eine dezidiert evidenzbasierte Politikberatung zu betreiben und dabei die Evaluation wirtschaftspolitischer Eingriffe in den Mittelpunkt des eigenen Tuns zu stellen, zum Selbstverständnis desjenigen Leibniz-Instituts, dessen Präsident ich seit dem Jahr 2002 sein darf, nämlich des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung (RWI). Dass dieses Selbstverständnis unseres Instituts sehr viel mehr ist als lediglich die reine Willenserklärung, Daten sprechen zu lassen, wird – so hoffe ich – durch die Beispiele deutlich werden, die ich zur Illustration der konzeptionellen Argumente anführen werde.

Um die angestrebten Ziele des Vortrags zu erreichen, gliedern sich meine Ausführungen in fünf Abschnitte. Nach der Einführung werde ich im ersten Abschnitt kurz die Grundbausteine nennen, die zum Verständnis des Konzepts der evidenzbasierten Politikberatung nötig sind und sie damit gegenüber einer Politikberatung abgrenzen, die nur *vorgibt*, evidenzbasiert zu sein. Dass sie ein evidenzbasiertes Vorgehen verfolgten, möchten zwar mittlerweile sehr viele Akteure für sich reklamieren, aber dieses Konzept stellt strenge Anforderungen, die zu erfüllen sind, um seinem Anspruch gerecht zu werden.

Der Kern meiner Ausführungen im zweiten und dritten Abschnitt betrifft die konzeptionellen Aspekte der evidenzbasierten Politikberatung. Zunächst diskutiert der zweite Abschnitt die Themen der empirischen Wirtschaftsforschung und ihre Bedeutung für die wirtschaftspolitische Beratung. Ausgehend von einer Systematisierung dieser Aufgaben nach methodischen und organisatorischen Kriterien wird dabei insbesondere die wissenschaftspolitische Einbettung der außeruniversitären Forschungseinrichtungen in Deutschland sichtbar.

Der dritte Abschnitt wird die »Identifikationsproblematik«, das große konzeptionelle Kernproblem unseres Tuns, erläutern: Anders, als etwa bei den Technikwissenschaften und Naturwissenschaften üblich, können wir in den meisten Anwendungsfällen unsere Analysen nicht unter kontrollierten Studienbedingungen durchführen. Daraus erwachsen eine Reihe eigener, höchst anspruchsvoller Herausforderungen an unsere Arbeit. Illustriert anhand einiger jüngerer Studien aus der Arbeit des RWI werde

2 Meine Ausführungen hier entwickeln nicht zuletzt Schmidt (2007 und 2009) weiter. Der Beitrag von Schmidt (2014) entwickelt wiederum die hier diskutierten Argumente weiter.

ich im vierten Abschnitt die grundsätzlichen Möglichkeiten diskutieren, um sich diesen Herausforderungen bei der empirischen Erforschung von Ursachen und Wirkungen erfolgreich zu stellen.

Abschließen werde ich im fünften Abschnitt mit einer Einordnung der Bemühungen der unabhängigen empirischen Wirtschaftsforschung in den Wettbewerb um wirtschaftspolitische Beratungsprojekte und um wirtschaftspolitisches Gehör. Denn wissenschaftliche Methoden anwenden, um Politikberatung zu betreiben, kann grundsätzlich jeder, insbesondere natürlich auch Einrichtungen mit kommerziellen oder ideologischen Interessen. Aber, so lautet die eindeutige Schlussfolgerung aus der hier vorgelegten Diskussion zur evidenzbasierten Politikberatung, lediglich die Ergebnisse von Regressionsanalysen oder anderer statistischer Auswertungen vorzulegen, rechtfertigt noch nicht den Anspruch, evidenzbasiert politikberatend unterwegs zu sein. Häufig nutzen, so meine Einschätzung auf Basis meiner langjährigen Erfahrungen, manche Anbieter wirtschaftspolitischer Ratschläge sogar derartige Methoden, um einen eigentlich vorher feststehenden (oder ihnen vom Auftraggeber vorgegebenen) Standpunkt zu untermauern und sich dabei den Anstrich der Wissenschaftlichkeit zu geben.

Nicht nur im Wettstreit untereinander, sondern auch in Konkurrenz mit derartigen Akteuren müssen wir als unabhängige Wissenschaftler in der täglichen Praxis operieren und uns auf dem Markt für wirtschaftspolitische Beratung behaupten. Je mehr die zu Beratenden, also die Nachfrager wirtschaftspolitischer Beratung, von den konzeptionellen Herausforderungen verstehen, mit denen wir konfrontiert sind, umso eher wird es ihnen gelingen, die Spreu vom Weizen der Beratungsangebote zu trennen – aber nur dann, wenn sie selbst das auch wollen. Meine folgenden Ausführungen bereiten insofern auch die Basis dafür, diese »Nachfragekompetenz« aktiv zu erwerben.

1 Der Schlüssel zu besserer Politik: evidenzbasierte Politikberatung

Zunächst gilt es, das Konzept der evidenzbasierten Politikberatung zu erläutern. Dies kann in recht einfacher Weise gelingen, indem man sich eine klare und deutliche Botschaft aus der aktuellen großen Wirtschaftskrise vor Augen hält. Denn diese hat uns eine bittere Erkenntnis eindringlich 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 vorzutauschen, 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.

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. Im Juni 2011 beschloss jedenfalls 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 dabei um eine gesellschaftliche Wertentscheidung handeln muss und dass die ihrer Einsetzung vorangegangene Reaktorkatastrophe in Japan, wenngleich sie diese Risiken nicht verändert hat, eine Neubewertung der mit der Nutzung der Kernenergie verbundenen Risiken erforderlich machte.

Eine aufgeklärte Gesellschaft kann und sollte es auf Dauer nicht akzeptieren, politische Kost vorgesetzt zu bekommen, die unzutreffenderweise absolute Gewissheiten vortauscht. Denn dann würden Reformvorschläge und politische Entscheidungen eher auf Glaubenssätzen beruhen als auf nüchterner Einsicht und belastbarer Evidenz. Doch oft genug ist dies 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.

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. Doch 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 (siehe beispielsweise Bauer, Kluge et al. 2009 oder das entsprechende Symposium im *German Economic Review*: Paloyo et al. 2013).

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 Beschäftigungsrisiken für die Arbeitnehmer verbunden ist. Man muss vom Ausbleiben negativer Beschäftigungseffekte demnach schon sehr überzeugt sein, um einen hohen Mindestlohn zu fordern. Die Anforderung an die wirtschaftspolitische Beratung ist also nicht, bei verbleibenden Unsicherheiten über die vermuteten Entwicklungen und 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 schwerfällt, sieht man schon allein an den Reaktionen zu Konjunkturprognosen, welche die Institute regelmäßig veröffentlichen: Politik und Presse fragen uns immer nach einer genauen Zahl und wollen nur sehr ungern eine Bandbreite akzeptieren, aber oft ist die Empörung groß, wenn dann die Vorhersagen nicht punktgenau eintreffen. Doch natürlich ist das überhaupt nicht zu leisten. Eine Einführung in diese Diskussion geben Döhrn und Schmidt (2005).

Zweitens muss man einsehen, dass es gerade die »großen« Fragen sind, für die es nur wenig Präzedenzfälle gibt, sodass 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, sodass es schwerfallen muss, aus den bisherigen Erfahrungen über das Funktionieren der Volkswirtschaft empirische Belege zu ihrer Wirkung abzuleiten. 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 im Raum stehende Vorschläge nach sich zögen, etwa der Schuldentilgungspakt des Sachverständigenrates (vgl. Schmidt und Weigert 2013), 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 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. Nun, verlangen kann der zu Beratende sicherlich, dass ich ihm deutlich sage, was ich an seiner Stelle tun würde – ihn aber auch auf die mit dieser Entscheidung verbundenen Risiken hinweise. Die Entscheidung bleibt dann 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 Position, weil natürlich jeder andere Berater, der größere Gewissheit vortäuscht, diese in der Tat nur vortäuschen kann.

Die evidenzbasierte Politikberatung ist somit eine Antwort auf das Problem, dass man in einer unsicheren Welt notwendigerweise Schwierigkeiten hat, darüber Erkenntnisse zu sammeln, wie etwas werden wird, wenn man entweder nicht eingreift oder 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 dargelegt, 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 zusammgeführt 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 ich auf einen großen Erfahrungsschatz zurückgreifen kann. Viele Präzedenzfälle, die alle in die gleiche Richtung laufen, geben mir relativ viel Zuversicht, dass die Schlussfolgerung in diesem Fall auch 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 großen Veränderungen in den Aussagen verzeichnen muss, so hilft dies natürlich, weitere Zuversicht in die Robustheit und die damit verbundene Güte der Schlussfolgerungen zu gewinnen.

Aus diesen Ausführungen ergibt sich eine dritte Botschaft für politische Entscheidungsträger: Man sollte gerade diejenigen wirtschaftspolitischen Beratern am wenigsten trauen, die sich diesem Mangel an absoluten Gewissheiten und den Schwierigkeiten bei der Suche nach überzeugenden empirischen Belegen nicht offen stellen. Eine »echte« evidenzbasierte Politikberatung hat sehr wohl den Anspruch, klare Empfehlungen dazu abzugeben, welche Weichenstellungen man in dieser unsicheren Welt verfolgen sollte, wohlwissend, was alles schiefgehen kann. Sie ist sich aber auch ihrer Grenzen bewusst und kommuniziert diese in transparenter Form. Sie findet idealerweise ihr Gegenstück in einer Ausgestaltung von politischen Entscheidungsmechanismen, die dieses Fehlen von absoluten Gewissheiten und die begrenzte Wirkmächtigkeit von Politik anerkennen und als lernende Systeme eine stetige Nachjustierung der Entscheidungen erlauben.

2 Aufgaben der (ökonomischen) Politikberatung

Es wird immer wieder gefragt: »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. Nun, das gilt sicherlich nicht für alles, was wir tun. Aber Forschung soll natürlich auch Spaß machen, und da ist es schon legitim, ökonomische oder ökonometrische Ansätze zu nutzen, um etwa über historisches Schlachtenglück oder den gefühlten Wert von Weihnachtsgeschenken zu arbeiten oder zu der Frage, ob bei Vollmond mehr Kinder geboren werden, wie es der Folklore nach die Hebammen sagen. Entsprechend bekenne ich mich schuldig im Sinne der Anklage:

Vgl. Rotte und Schmidt (2003); Bauer und Schmidt (2012) sowie Bauer et al. (2012b).

Doch wir Ökonomen wissen auch, wenn es um Emotionen, um Liebe und Hass, um Freundschaften geht, braucht man natürlich nicht nur einen Ökonomen, der darauf abhebt, welche materiellen Vorteile man aus einer Freundschaft oder einer Partnerschaft gewinnt. Die Lebenswirklichkeit so auf ein wirtschaftliches Kalkül zuzuspitzen, wäre wirklich albern, selbst wenn man manchmal den Eindruck gewinnen mag, dass einige Ökonomen mit einem Anspruch auf Allgemeingültigkeit ihrer Weltansicht auftreten. In unserem »Kerngeschäft« sind wir nichtsdestoweniger für viele Aspekte des wirtschaftspolitischen Handelns und menschlichen Lebens zuständig, weit über das regelmäßige Anfertigen von Konjunkturprognosen hinaus.

Da sind zum einen die »großen« wirtschaftspolitischen Fragen unserer Zeit:

- Wie – ich hatte es im vorangegangenen Abschnitt bereits angesprochen – soll der gemeinsame Währungsraum in Europa ausgestaltet, wie die aktuelle Krise bewältigt werden? Siehe dazu z. B. Schmidt und Weigert (2013).
- Welche energiepolitischen Weichenstellungen machen eine erfolgreiche Umsetzung des großen gesellschaftlichen Vorhabens »Energiewende« überhaupt erst möglich? Siehe beispielhaft dazu Hübner et al. (2012).
- Was macht menschliche Wohlfahrt aus und wie kann man Wohlstand und Fortschritt sachgerecht statistisch erfassen? Mit diesem Thema befasste sich jüngst die Enquete-Kommission des Deutschen Bundestages »Wachstum, Wohlstand, Lebensqualität«. Ein aktueller Beitrag zu dieser Debatte ist Schmidt und aus dem Moore (2013).
- Welche Probleme wirft der anstehende demographische Wandel für das Wirtschaftswachstum und die Tragfähigkeit der sozialen Sicherungssysteme auf? Eine aktuelle Übersicht zu diesen Fragen bietet beispielsweise Schmidt (2012).

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 (vgl. beispielhaft Bauer et al. 2013; Fertig und Schmidt 2011; Tauchmann et al. 2013). Diese Beiträge sind der Ökonomik möglich, weil sie sich stets auf einige Grundprinzipien besinnt – die unvermeidliche Existenz von Knappheiten (»Budgetrestriktionen«) und die Einsicht, dass ein gegebener Zustand

tendenziell eher eine Balance (ein »Gleichgewicht«) zwischen widerstrebenden Kräften und kein Zufallsprodukt darstellt.

Und 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 (beispielhaft siehe Guinnane et al. 2011 bzw. aus dem Moore und Schmidt 2013). Bei all diesem Tun haben wir im Grunde genommen drei Aufgaben zu erfüllen, am besten im Dreiklang »Erkennen, Erwarten, Ergründen« zu fassen. Dabei erweisen sich alle drei Aufgaben als gleichermaßen wichtig. Es kommt daher nicht von ungefähr, dass alle wirtschaftswissenschaftlichen Institute der Leibniz-Gemeinschaft diesen Dreiklang in ihrem Leistungs-Portfolio aufweisen.

So ist es, erstens, 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 doch letztendlich 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 (vgl. RWI und Stifterverband Wissenschaftsstatistik 2010). Im Übrigen geht es dabei häufig um weit mehr als nur um Mittelwertvergleiche, insbesondere dann, wenn es sich bei den vorliegenden Daten nicht um Zufallsstichproben handelt, etwa aus Gründen der Antwortverweigerung. In diesem Falle muss eine entsprechende statistische Korrektur zum Einsatz kommen, um sich der Zufallsstichprobe – und damit tatsächlicher Vergleichbarkeit – anzunähern. Dies ist keinesfalls eine intellektuell gering einzuschätzende Herausforderung.

Zweitens besteht unsere Aufgabe häufig darin, darzulegen, welche künftigen Entwicklungen in verschiedenen Wirtschaftsbereichen zu erwarten sind. Das klassische Beispiel für diese Leistung der empirischen Wirtschaftsforschung sind natürlich die regelmäßigen Konjunkturprognosen der Leibniz-Institute im Rahmen der sogenannten Gemeinschaftsdiagnose (vgl. beispielhaft Döhrn 2013). Prognosen sind in der Praxis

allerdings weit mehr als lediglich die akribische Fortschreibung am »aktuellen Rand« des Geschehens: Prognosen sind, so trivial das klingen mag, zum einen nötig, weil die Zukunft nicht so ist wie die Vergangenheit; sie sind zum anderen aber auch 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 (vgl. Fertig und Schmidt 2002).

Gewissermaßen die Krönung unseres Tuns sind, drittens, Kausalanalysen. Es gehört sicherlich ohnehin zum menschlichen Bestreben, Ursachen zu ergründen, also herauszufinden, warum die Dinge so sind, wie sie sind. Im Kontext der Wirtschaftspolitik ist insbesondere zu fragen, ob eine gegebene wirtschaftspolitische Weichenstellung nun 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 jetzt hier ein Erfolg oder ein Misserfolg ergeben, war das der Fall, obwohl, oder sogar, weil der Staat eingegriffen hat? Die (vermeintliche) Einsicht in das Zusammenspiel von Ursache und Wirkung stellt nicht zuletzt die intellektuelle Unterfütterung all unserer wirtschaftspolitischen Empfehlungen dar.

Die große Herausforderung bei der Frage nach Ursache und Wirkung im Tätigkeitsfeld von Ökonomen – und das ist auch zugleich das Schöne an der empirischen Wirtschaftsforschung – ist, 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. Und diesen eigenen Willen können unsere Studienobjekte glücklicherweise in unserer freiheitlichen Gesellschaft auch 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 nicht notwendigerweise mit dem Nachweis von Kausalität gleichzusetzen.

Der beste Weg, um sich der grundlegenden Frage von Ursache und Wirkung überzeugend zu nähern, ist die Durchführung eines kontrollierten Experiments. Meist besteht aber dazu keine Gelegenheit. Dann muss ich danach streben, die vorliegende nicht-experimentelle Situation so auszunutzen, dass ich dem erwünschten, aber leider nicht durchführbaren Experiment durch mein Studiendesign möglichst nahe komme. Die Einsicht, wie mit diesem Problem in der (nicht-experimentellen) Wirtschaftsforschung umzugehen ist, stellt nach meiner Einschätzung einen der größten Fortschritte unserer Disziplin in den vergangenen Jahrzehnten dar. Ein Beispiel für die in den vergangenen Jahren nahezu

explosionsartig gewachsene Literatur zur Evaluation von wirtschaftspolitischen Eingriffen und deren jeweilige Anforderungen ist, hier speziell auf die Umweltökonomik zugeschnitten, Frondel und Schmidt (2005).

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.

Neben eigens zu diesem Zweck eingerichteten Beratungsgremien, wie etwa dem Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (SVR) und den jeweiligen Wissenschaftlichen Beiräten des Bundesministeriums für Wirtschaft und Technologie (BMWi) sowie des Bundesministeriums der Finanzen (BMF), ist insbesondere die Arbeit der wirtschaftswissenschaftlichen Institute der Leibniz-Gemeinschaft (WGL) mit ihrer Mission einer »angewandten Grundlagenforschung« darauf ausgerichtet, diese Leistungen zu erbringen. Die Leibniz-Gemeinschaft ist daher genau die richtige Heimat für die empirische Wirtschaftsforschung und die eng damit verbundene wissenschaftliche Beratung der Politik. Es ist somit kein Zufall, dass alle der oft als »große« Wirtschaftsforschungsinstitute bezeichneten Institute in der Leibniz-Gemeinschaft eingeordnet sind. Diese dezidiert »angewandte« Forschung zielt, in die Lebenswirklichkeit eingebettet, 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 eigentlich 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 daher 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 – aber dann auch kompetent genug sein, dies zu leisten. Uns gibt also nicht die Ästhetik eines mathematisch anspruchsvollen Ansatzes den Anlass für methodische Beiträge, sondern das inhaltliche Interesse. Die Methoden bilden somit lediglich das Rüstzeug für unsere Arbeit.

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 auch schwer vorstellbar ist.

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 in der Wissenschaft hoch angesehen sind (im deutschen Kontext bedeutet dies meist, dass eine Zeitschrift im sogenannten »Handelsblatt-Ranking« vertreten ist). Rein deskriptive Studien – wie der oben angesprochene »Innovationsbericht« – gehören trotz ihrer hohen wirtschaftspolitischen Bedeutung sicherlich nicht dazu.

In der Konsequenz ist die strikte institutionelle 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, Beschreibung und Fortschreibungen am aktuellen Rand machen die fleißigen, aber konzeptionell schlichten angewandten Forscher in den außeruniversitären Instituten, doch die »richtige Forschung« findet in 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 auch als Serviceeinrichtung für externe Forscher derartige Daten bereitstellt. Und unsere Forschung muss stets »richtige Forschung« sein,

gleich, ob es sich um Deskription und Prognose oder Kausalanalyse handelt.

Insbesondere sind Evaluationsstudien mittlerweile ein Markenzeichen der außeruniversitären Forschung in Deutschland geworden, nicht zuletzt deshalb, weil dabei die eigene, auf die konkrete Anwendungsfrage maßgeschneiderte Datensammlung, die in den Universitäten nur schwer zu leisten ist, nahezu unabdingbar geworden ist. Warum dies so ist, werde ich im nächsten Abschnitt darzulegen versuchen. Aber ein weiterer Grund ist mindestens genauso wichtig: Nur in einem außeruniversitären Forschungsinstitut werden sich typischerweise Gruppen von empirisch arbeitenden Ökonomen in einer Anzahl zusammenfinden, welche die Anforderungen einer »kritischen Masse« erfüllt. Dies macht eine umfassende 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 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. Wenn Sie vor dieser Aufgabe stehen, haben Sie eigentlich immer einen gewissen Satz an Minimalvoraussetzungen zu erfüllen. Um die Diskussion darüber, welche Aspekte des Phänomens Sie als gegeben annehmen können und welche noch zu erobern, zu errahnen, zu erarbeiten sind, kommt niemand herum. Diese Fragen stehen unweigerlich am Anfang einer jeden empirischen Studie.

Denn nur das, von dem man wenigstens errahnt, dass man darüber Erkenntnisse sammeln kann, lässt sich überhaupt empirisch erforschen. Und über das, von dem ich nicht einmal im Ansatz weiß, dass ich es nicht kenne, kann ich als empirischer Forscher auch keine Erkenntnisse gewinnen. 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 errahne, zu erkennen.

Jeder, der einmal an einem Statistikkurs teilgenommen hat, wird leicht nachvollziehen können, dass es bei den intellektuellen Herausforderungen der statistischen Arbeit im Grunde genommen drei Perspektiven gibt, von denen jedoch in einem der üblichen Statistikkurse nur zwei tiefer

diskutiert werden. Die erste Perspektive des statistischen Denkens betrifft die Wahrscheinlichkeitstheorie. Der 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 ich jetzt beispielsweise eine Stichprobe zusammenstelle, indem ich mit Zurücklegen zehnmal eine Kugel aus der Urne hole und mir die Farbe der jeweils gezogenen Kugel merke: Welchen Anteil an roten Kugeln dürften wir in der Stichprobe erwarten? Nun, das Ergebnis ist unter den genannten Bedingungen, dass man erwarten darf, dass die Hälfte der gezogenen Kugel rot sein wird.

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 die Existenz von schwarzen Kugeln ausgeschlossen werden kann, lässt sich der oben geschilderte Analogieschluss ziehen. Treffe ich eingangs meiner statistischen Analyse die (Identifikations-)Annahme, dass es in der Grundgesamtheit keine schwarzen Kugeln gibt, dann bin ich auf einmal in der Rolle des Statistikers, dann greifen die bewährten Einsichten der schließenden Statistik.

In der Praxis der empirischen Wirtschaftsforschung ist aber genau diese Anforderung entscheidend: Ohne Identifikationsannahmen geht es nicht, ohne sie können keine Erkenntnisse gesammelt, keine Schlussfolgerungen abgeleitet werden! Nun ist das gewählte Beispiel sehr stilisiert, aber ich hoffe, dass es die grundsätzliche Idee dennoch plastisch transportieren kann. 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 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 beantwortet, schon von vornherein völlig falschliegen, völlig unabhängig davon, wie viel statistisch-ökonomisches »Muskelspiel« in Form fortgeschrittener statistischer Methoden 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 ökonomische Rüstzeug als intellektuelle Grundausstattung eines modernen Ökonomen unbedingt dazugehört, 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: Nahezu alle empirischen 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 (vgl. Döhrn et al. 2009). 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.

Vor drei Jahrzehnten, als ich selbst die empirische Wirtschaftsforschung als studentische Hilfskraft beim unvergessenen Heinz König am Lehrstuhl für Ökonometrie in Mannheim kennenlernen durfte, war die Ökonometrie in Deutschland noch ein relativ neuer Forschungszweig. Die angewandte Wirtschaftsforschung hat sich damals natürlich mit viel primitiveren Methoden, weniger leistungsfähigen Rechnern und mit viel größeren Daten begnügen müssen, als es heute der Fall ist. Damals glaubte man noch, der größte zu erwartende Fortschritt in unserem Fach würde künftig darin liegen, dass wir die Systeme, mit denen wir umgehen, komplexer werden lassen, wenn unsere Rechner schneller und die Datensätze größer werden. Man erwartete, künftig immer mehr in Richtung großer Vielgleichungssysteme mit viel Nichtlinearität zu gehen.

Das ist interessanterweise genau nicht das, was tatsächlich passiert ist. Mittlerweile dominiert stattdessen die Einsicht, dass am Anfang die Frage steht, was man bei einer begrenzten Informationslage überhaupt verlässlich lernen kann: Wo sind meine Möglichkeiten, aber vor allem, wo sind auch die Grenzen meines Erkenntnisgewinns? Das, was man sich ehemals 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.

Die schließende Statistik – so habe ich sie ja bereits in gewisser Weise charakterisiert – 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 ökonomischen 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 restriktive Identifikationsannahmen empirische »Freiheitsgrade« freischaufle, wächst die ausgewiesene Präzision der geschätzten Modellparameter – aber auch das Risiko, damit völlig falschzuliegen.

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. Aber es wird immer einen Grundkanon an Identifikationsannahmen geben, ohne den ich überhaupt kein Ergebnis ableiten kann. Im obigen Beispiel ist dies etwa die Nicht-Existenz von schwarzen Kugeln. Diese Annahmen können 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, erst gar nicht einfallen, auf Basis der Stichprobe meine identifizierende Grundannahme zu testen, dass es in der Grundgesamtheit keine schwarzen Kugeln gibt. Denn meine Aussagen über die bestehenden Zusammenhänge kann ich nur aus der vorliegenden Stichprobe ableiten, in der schwarze Kugeln nicht enthalten sind.

Da die unbedingt nötigen Identifikationsannahmen niemals getestet werden können, gibt es am Anfang jeder empirischen Studie ein Abwägungsproblem, über das wir im normalen Statistikkurs nie etwas Näheres lernen: 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 daraus 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 falschliegen, aber den oberflächlichen Betrachter mag die hohe Trennschärfe der Aussage dennoch beeindrucken. Es dürfte aber für die Wirtschaftspolitik nicht zielführend sein, ganz präzise völlig falschzuliegen. Daher ist in der Praxis besonders dort Vorsicht geboten, wo die Anbieter von wirtschaftspolitischen Beratungsleistungen mit unangemessen geringen Zweifeln an der Belastbarkeit ihrer empirisch gestützten Schlussfolgerungen auftreten.

4 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.

Man sagt ja so schön, man steige nie zweimal in denselben Fluss. Genau dies charakterisiert die Beobachtungslage im Falle der Kausalanalyse. Denn 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. Wie man in der praktischen Arbeit die

Evaluierung von arbeitsmarktpolitischen Maßnahmen konstruktiv angehen kann, zeigen anschaulich Fertig et al. (2004). Beispielsweise könnte man sich vorstellen, dass man die Wirkungen eines Weiterbildungsprogramms für Langzeitarbeitslose evaluieren möchte. In einer 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 ja 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 Vergleichsbeobachtungen bilden, einige möglicherweise gar nicht den Lebensmut oder das nötige Durchhaltevermögen aufweisen, um dies zu schaffen. Insgesamt dürften diese also keine zwei völlig vergleichbaren Gruppen darstellen.

Aus diesem Reservoir muss ich mir in einer empirischen Kausalanalyse jetzt diejenigen Beobachtungen herausuchen, die – hoffentlich, wenn ich meine Arbeit richtig mache – 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, für den empirischen Wirtschaftsforscher in der Regel nicht beobachtbar sind. 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.

Aus diesem Dilemma kann man nicht ausbrechen: Empirische Forschung lebt immer davon, sich von der Anekdote zu lösen und durch Informationsverdichtung gemeinsame Muster herauszuarbeiten. Daher 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.

In einem kontrollierten Zufalls-Experiment kann ich diese Problematik 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 verrät, 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. Diese wird selbst dann noch da sein, wenn man als näherungsweise Ansatz Indikatoren aus der psychologischen Forschung in die Analyse einbeziehen kann, z. B. die Risikoneigung. 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-beobachtbaren Eigenschaften – so die Identifikationsannahme. Ein Beispiel für eine derartige Suche nach »statistischen Zwillingen«, von denen der eine die zu untersuchende Maßnahme erfährt, während der andere aus dem Reservoir der Nicht-Teilnehmer stammt, ist bei Kluge, Lehmann und Schmidt (2008) zu finden. Dort entscheiden insbesondere die jeweiligen früheren Erfahrungen am Arbeitsmarkt über die Zuordnung zu einem »Zwillingspaar«. Auf weitere illustrative Beispiele für die nicht-experimentelle Kausalanalyse komme ich im Folgenden noch zu sprechen.

Die wichtigste Frage kommt allerdings bei jeder Studie zuerst. So trivial dies auch klingen mag, aber bereits beim Einstieg in die Analyse wird in der Praxis leider viel falsch gemacht: Es kommt erst einmal darauf an, die richtige Frage zu stellen – wo ist das informative Datenmaterial, was sind die richtigen Beobachtungseinheiten, welche kontrafaktische Situation möchte ich eigentlich konstruieren? Erst auf dieser Basis kann ich daran gehen, den Kontrast zwischen Teilnehmer und Vergleichsbeobachtung zu konstruieren, der mir die Antwort auf die Frage nach der Wirkung der Maßnahme gibt. Und erst dann ist die Frage nach dem überzeugenden Studiendesign möglich, das mir diesen Kontrast liefert.

Nun, wir haben gesehen, dass – gegeben, dass ich diese ersten Schritte erfolgreich bewältigt habe – bei der Identifikation kausaler Effekte die

Konstruktion der kontrafaktischen Situation das Entscheidende ist. Und wir haben zwei grundsätzliche Vorgehensweisen angesprochen, das sozialwissenschaftliche Experiment einerseits und nicht-experimentelle Methoden andererseits. Beide haben ihre eigenen Fallstricke und bringen ihre eigenen Herausforderungen mit sich.

Ein Beispiel, an dem ich das Vorgehen bei experimentellen Studien erläutern möchte, betrifft die Gesundheitsökonomik. Es handelt sich dabei um ein sozialwissenschaftliches Feldexperiment zum Thema »Fettleibigkeit«, das vom Pakt für Forschung und Innovation finanziert worden ist (vgl. Augurzky et al. 2012). Dabei wurde Patienten, die eine Reha-Klinik verlassen haben und gleichzeitig einen Body-Mass-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 wenigen 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 die Fähigkeit zur überzeugenden empirischen Analyse beeinflusst.

Ein weiteres Beispiel, an dem ich nunmehr 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, darüber muss man erst gar nicht diskutieren. In unserer Studie, die ich hier kurz vorstellen werde, konnten wir auf einen großen Datenschatz zugreifen, der uns retrospektiv die Lebensleistung von Arbeitnehmern darlegt – und jeweils einen Hinweis darauf gibt, ob sie als junge Männer in der Bundeswehr gedient hatten oder nicht (Bauer et al. 2012a). Dies ist eine ganz klar nicht-experimentelle Situation, denn ich weiß jetzt zwar im Nachhinein relativ gut, ob

sie ihr Arbeitsleben im Endeffekt erfolgreich oder nicht erfolgreich bestritten haben, aber die Zuweisung zum Wehrdienst kann ich zum Zwecke des Studiendesigns nicht mehr beeinflussen.

Damit kann ich leider auch nicht mehr verhindern, dass sich die Teilnehmer am Wehrdienst und die Nicht-Teilnehmer systematisch in im Datensatz nicht aufgeführten Eigenschaften unterscheiden, etwa in ihrer physischen Robustheit. Allerdings kann ich, 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 kann ich immer dann in die Tat umsetzen, wenn ich äußere – aus Sicht des Analytikers »exogene« – Umstände finde, welche die Zuweisung zu der zu bewertenden Maßnahme beeinflusst haben, ohne dass sie von den Betroffenen beeinflusst werden konnten und ohne dass sie dann wiederum selbst über die Zuweisung hinaus einen Einfluss auf deren Erfolg am Arbeitsmarkt aufweisen konnten. 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, meines Datensatzes gelegen haben. Im vorliegenden Fall ist es das konkrete Geburtsdatum, das zumindest für eine kleine Gruppe von Arbeitnehmern, nämlich diejenigen, die um den 1. Juli 1937 geboren wurden, über ihre Wehrpflicht entscheidet: Wer nach dem 1. Januar 1929 und vor dem 30. Juni 1937 geboren wurde, gehörte zu den sogenannten »weißen Jahrgängen«; 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 ich die Variation im Erfolg, die wir an dieser Sprungstelle sehen, wie in einem Experiment zur Ableitung einer Aussage über die Langzeitwirkungen des Wehrdienstes nutzen. Diese fallen im Übrigen vernachlässigbar aus; das kann man aber erst sehen, wenn man unserem Studiendesign folgt.

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. Ich denke, dass die entsprechende Identifikationsannahme einer quasi-randomisierten Zuweisung der Wehrpflicht für um dieses Datum herum Geborene im vorliegenden Falle recht überzeugend sein dürfte.

Denn es ist ja wohl 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 Gedankengang 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 aber leider wahrlich nicht jede empirische Studie vorweisen.

5 Wettbewerb und Weiterentwicklung 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 ich Neuland betrete, ist es notgedrungen sehr schwer, empirische Evidenz für wirtschaftspolitischen Rat anzubieten.

Aber so ernüchternd das ist: Wenn ich alles bereits wissen würde, dann gäbe es erst gar kein empirisches Problem, und wenn ich akzeptierte, dass ich überhaupt nichts weiß, könnte ich keinerlei Schlussfolgerungen ziehen. Die angemessenen Identifikationsannahmen muss ich also, es bleibt nichts anderes übrig, intelligent und mit dem Wissen um die bei dieser Wahl auftretenden Fallstricke treffen. Die Daten allein werden mir keine ultimative Sicherheit über die Qualität meines Vorgehens liefern können. Denn da ich zunächst immer zumindest einige grundlegende Annahmen treffen muss, um aus der Stichprobe überhaupt erst Schlussfolgerungen ableiten zu können, kann ich diese als gegeben angenommenen Aspekte nicht mehr hinterfragen. Jede Identifikationsstrategie muss daher argumentativ gut unterfüttert werden.

Es wäre aber – das hatte ich eingangs bereits sehr deutlich angesprochen – nicht hilfreich für die Politik, wenn wir nur einfach die Hände in die Luft werfen und sagen würden: »Na ja, das wissen wir halt alles nicht so genau, wir hoffen auf weitere Daten.« Die (wirtschaftswissenschaftliche)

Politikberatung ist vielmehr, das ist meine feste Überzeugung, 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, wengleich 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 jahrelangen Praxis eines wirtschaftspolitischen Beraters sagen, ist es zumindest häufig noch immer so, dass dieser Wille nicht vorhanden ist. Häufig werden schöne bunte Charts und klare Aussagen vorgezogen, vielleicht nicht zuletzt deshalb, weil sie helfen können, eine Politik zu »verkaufen«. Doch wahrhaft gute (Wirtschafts-)Politik sollte stattdessen den Stand der aktuellen Erkenntnis, die begrenzte eigene Wirkungsmacht und die mit jeder Handlungsoption verbundenen Risiken berücksichtigen.

Als empirische Wirtschaftsforscher müssen wir wohl oder übel damit leben, dass unsere wirtschaftspolitische Beratung manchmal ungenert gehört wird, weil sie dem Vortäuschen von ultimativen Gewissheiten entsagt. Für uns gilt es aus meiner Sicht, unsere Linie durchzuhalten und insbesondere immer darauf aufmerksam zu machen, wenn man keine apodiktische Aussage treffen kann. Wir müssen zudem anerkennen: 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 all 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 in Deutschland realisierte 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. Dass dies nicht jedem gelingt, der dies für sich in Anspruch nimmt, ändert nichts daran, sondern verstärkt diese Forderung nur. Dabei ist offensichtlich, dass man sich mit klaren wirtschaftspolitischen Stellungnahmen, die den inhaltlichen Positionen von starken Interessengruppen, Parteien oder Regierungen widersprechen, nicht unbedingt Freunde macht. Beispielhaft seien hier genannt: Frondel et al. (2006); Frondel et al. (2010); aus dem Moore et al. (2009) sowie **Gebhardt et al. (2011)**. Interessant an diesen Beispielen aus zwei Bereichen, der Energie- und der Finanzpolitik, dürfte insbesondere sein, dass sich jeweils im Diskurs völlig entgegengesetzt einzuordnende Gruppen von diesen Stellungnahmen erkennbar angegriffen gefühlt haben.

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 ist aus meiner Sicht durch das Modell einer monopolistischen Konkurrenz sehr gut beschrieben. 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. Und natürlich versuchen wir alle, auch gute Konjunkturprognosen anzubieten.

Gleichzeitig müssen wir uns aber auch gewisse Alleinstellungsmerkmale erarbeiten. Ob man dies tatsächlich schafft, wird gerade in den Evaluationsrunden der Leibniz-Gemeinschaft kritisch geprüft, die alle sieben Jahre unsere Existenz als Mitglied der Leibniz-Gemeinschaft aufs Neue in Frage stellen. Im Falle des RWI sind diese Alleinstellungsmerkmale nachweislich vor allem mit der mikroökonomischen Evaluation wirtschaftspolitischer Maßnahmen verknüpft und betreffen beispielweise die Umwelt- und Energieökonomik oder auch die Gesundheitsökonomik. Und auch die anderen wirtschaftswissenschaftlichen Institute der Leibniz-Gemeinschaft haben sicherlich derartige Alleinstellungsmerkmale vorzuweisen, sonst verlören sie ihre Existenzgrundlage.

Es gibt in der Tat – wie oben bereits ausgeführt – gute Gründe für die beratungsorientierte Wirtschaftsforschung, Spitzenforschung anzustreben, und gute Gründe für unsere wissenschaftlichen Aufsichtsgremien, dies auch zu verlangen. In der universitären Forschung ist 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 Interessen ohne weiteres jederzeit neu justieren können. Wenngleich Forschungsinstitute typischerweise eine Abstimmung der in ihrem Team verfolgten Forschungsfragen vornehmen, 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 – nichtsdestoweniger dazu, eine gewisse Breite der Themen anzubieten (vgl. Schmidt 2006). 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 Berater 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. Denn was man vergleichsweise leicht messen 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. Doch ohne das ernsthafte Bemühen, gute wirtschaftspolitische Beratung auch als solche zu würdigen, dürfte es schwer werden, langfristig ihre hohe fachliche Qualität sicherzustellen. Meine persönliche Einschätzung der aktuellen Situation findet sich in Schmidt et al. (2013b).

An dieser Stelle ist es keineswegs abwegig, eine sprichwörtliche Lanze für die Leibniz-Gemeinschaft zu brechen: Im vergangenen Jahrzehnt, in dem ich das Evaluationsverfahren aus nächster Nähe beobachten konnte, hat sich die Evaluierung der Institute, ich spitze es ein bisschen zu, von einem schlichten Abzählreim (»Wie viele hochwertige Publikationen gibt es denn?«) mehr und mehr zu einer Suche nach einem runderen Bild gewandelt. Es wird nach wie vor sehr kritisch untersucht, was die Institute wissenschaftlich leisten, aber es wird zumindest im Ansatz gewürdigt, welche Qualität der wirtschaftspolitischen Beratung (und der Nachwuchsförderung sowie der Bereitstellung von Forschungsinfrastruktur)

sie anbieten. Man bemüht sich also durchaus, nicht nur unter der sprichwörtlichen Laterne zu suchen, die durch die einschlägigen Publikationsrankings bereitgestellt wird, sondern um eine balanciertere Einschätzung der Qualität von Forschung und Beratung.

Was nützte jedoch selbst eine hervorragend aufgestellte evidenzbasierte Politikberatung, wenn die Nachfrageseite für ihre Leistungen – und vor allem für ihre dabei gezeigte Ablehnung eines Vortäuschens 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.

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

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

Lediglich darum zu wissen, was eigentlich gute, also evidenzbasierte wirtschaftspolitische Beratung ausmacht, wird kaum hinreichend sein. Es muss stattdessen für die Nachfrager in Politik und Verwaltung auch zunehmend unausweichlich werden, ihr Gehör zu schenken. Dies kann

aber nur dadurch gesichert werden, dass der öffentliche Druck entsprechend wächst, die Verausgabung 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 rechtfertigen. Dies bedeutet aber, dass auch der Journalismus und letztlich auch der mündige Bürger ohne ein gewisses Verständnis für die Herausforderungen der empirischen Wirtschaftsforschung und der evidenzbasierten Politikberatung ihre Kontrollfunktionen nur sehr eingeschränkt werden wahrnehmen können. Hier ist sicherlich noch viel zu tun.

Aber manchmal hilft auch ein Missverständnis: Wer sich an den Skandal um die damalige Bundesanstalt für Arbeit zurückerinnert, der sich kurz nach der Jahrtausendwende abspielte, dem fallen sicherlich die, sagen wir es mal vorsichtig, geschönten Vermittlungsstatistiken ein. Aber 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 Abruf der vorgesehenen Mittel wurde bis dahin als großer Erfolg gewertet – tatsächlich bewirkten.

Um diesem öffentlichen Druck zu reagieren, wurde in die Hartz-Gesetze zur Reform der aktiven Arbeitsmarktpolitik die Forderung aufgenommen, dass eine wissenschaftliche Evaluation durch unabhängige Dritte in Auftrag zu geben sei. Jahrelange Mahnungen von Arbeitsökonominnen hatten das nicht bewirken können, die öffentliche Empörung schaffte es sofort. Und – etwa zehn Jahre später – wurde in das im Jahr 2014 verabschiedete »Mindestlohngesetz« die Notwendigkeit einer Evaluation seiner Wirkungen ohne große Diskussionen aufgenommen. Allerdings darf man gespannt sein, wie sehr sich die Umsetzung dieses Vorhabens dann in der Praxis an den hier ausführlich skizzierten Standards orientieren wird.

Meine Ausführungen zum Thema »Politikberatung und Evaluationskultur« sollen nicht enden, ohne kurz aus ihrem dezidiert deutschen Fokus auszubrechen. Schließlich gilt für dieses Thema das Gleiche wie für viele andere wirtschaftspolitische Themen: Unsere Herausforderungen liegen längst nicht mehr allein auf der nationalen Ebene und daher lassen sich dort auch die Lösungen für viele unserer Probleme nicht mehr finden. Auf Dauer wird unser Tun als empirische Wirtschaftsforscher und wirtschaftspolitische Berater daher mehr und mehr eine Internationalisierung, vor allem eine Europäisierung benötigen. Das wird ein Schritt sein, den Angebot und Nachfrage gemeinsam werden gehen müssen. Hoffen wir, dass die »Evaluationskultur« in Deutschland bis dahin

noch viel besser Fuß gefasst hat. Dann wird dieser Schritt umso eher gelingen können.

Literatur

- Augurzyk, Boris; Bauer, Thomas K.; Reichert, Arndt R.; Schmidt, Christoph M. und Tauchmann, Harald (2012): Does Money Burn Fat? – Evidence from a Randomized Experiment. *Ruhr Economic Papers* 368.
- aus dem Moore, Nils; Kambeck, Rainer und Kasten, Tanja (2009): Auswirkungen der Steuerprogramme zur Bundestagswahl 2009 – Eine mikrodatenbasierte Analyse der Reformvorschläge von CDU/CSU, SPD und FDP. *RWI Materialien* 55.
- aus dem Moore, Nils und Schmidt, Christoph M. (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), S. 266–290.
- Bauer, Thomas K. und Schmidt, Christoph M. (2012): WTP vs. WTA: Christmas Presents and the Endowment Effect. *Jahrbücher für Nationalökonomie und Statistik* 232(1), S. 4–11.
- Bauer, Thomas K.; Fertig, Michael und Schmidt, Christoph M. (2009): *Empirische Wirtschaftsforschung. Eine Einführung*. Heidelberg: Springer.
- Bauer, Thomas K.; Kluge, Jochen; Schaffner, Sandra und Schmidt, Christoph M. (2009): Fiscal Effects of Minimum Wages: An Analysis for Germany. *German Economic Review* 10(2), S. 224–242.
- Bauer, Thomas K.; Bender, Stefan; Paloyo, Alfredo und Schmidt, Christoph M. (2012a): Evaluating the Labor-market Effects of Compulsory Military Service. *European Economic Review* 56(4), S. 814–829.
- Bauer, Thomas K.; Bender, Stefan; Heining, Jörg und Schmidt, Christoph M. (2012b): The Lunar Cycle, Sunspots and the Frequency of Births in Germany, 1920–1989. *Economics & Human Biology* 11(4), S. 545–550.
- Bauer, Thomas K.; Bredtmann, Julia und Schmidt, Christoph M. (2013): Time vs. Money – The Supply of Voluntary Labor and Charitable Donations Across Europe. *European Journal of Political Economy* 32, S. 80–94.
- Döhrn, Roland und Schmidt, Christoph M. (2005): Kein Stein der Weisen – Prognosen erfordern eine intelligente Nutzung. *RWI Positionen* 5.
- Döhrn, Roland; Barabas, Györgi; Gebhart, Heinz; Kitlinski, Tobias und Vossen, Simeon (2009): Die wirtschaftliche Entwicklung im Inland: Weltweite Finanzkrise verursacht historischen Konjunkturreinbruch. *RWI Konjunkturberichte* 60(1), S. 35–89.
- Döhrn, Roland (2013): Deutsche Konjunktur erholt sich – Ergebnisse der Gemeinschaftsdiagnose vom Frühjahr 2013. *Wirtschaftsdienst* 93(5), S. 323–327.
- Fertig, Michael und Schmidt Christoph M. (2002): Reading Tea Leaves or Science? Forecasting for Beginners. *ASTA – Allgemeines Statistisches Archiv* 86(4); S. 459–467.

- Fertig, Michael und Schmidt, Christoph M. (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(1); S. 99–128.
- Fertig, Michael; Kluge, Jochen; Schmidt, Christoph M.; Apel, Helmut; Friedrich, Werner und Haegele, Heinrich (2004): Die Hartz-Gesetze zur Arbeitsmarktpolitik – Ein umfassendes Evaluationskonzept. *RWI Schriften* 74. Berlin: Duncker & Humblot.
- Frondel, Manuel und Schmidt, Christoph M. (2005): Evaluating Environmental Programs: The Perspective of Modern Evaluation Research. *Ecological Economics* 55(4), S. 515–526.
- Frondel, Manuel; Kambeck, Rainer und Schmidt, Christoph M. (2006): Kohlesubventionen um jeden Preis? – Eine Streitschrift zu den Argumentationslinien des Gesamtverbandes des deutschen Steinkohlenbergbaus. *RWI Materialien* 25.
- Frondel, Manuel; Ritter, Nolan; Schmidt, Christoph M. und Vance, Colin (2010): Die ökonomischen Wirkungen der Förderung Erneuerbarer Energien: Erfahrungen aus Deutschland. *Zeitschrift für Wirtschaftspolitik* 59(2), S. 107–133.
- Gebhardt, Heinz; Kambeck, Rainer und Matz, Florian (2011): Der NRW-Haushalt für das Jahr 2011: Scheinerfolge bei der Konsolidierung. *RWI Positionen* 47.
- Guinnane, Timothy W.; Streb, Jochen; Frondel, Manuel und Schmidt, Christoph M. (2011): Krank oder Simulant? Das Identifikationsproblem der Krankenkassen aus wirtschaftshistorischer Perspektive. *Perspektiven der Wirtschaftspolitik* 12(4), S. 413–429.
- Hübner, Malte; Schmidt, Christoph M. und Weigert, Benjamin (2012): Energiepolitik: Erfolgreiche Energiewende nur im europäischen Kontext. *Perspektiven der Wirtschaftspolitik* 13(4), S. 286–307.
- Kluge, Jochen; Lehmann, Hartmut und Schmidt, Christoph M. (2008): Disentangling Treatment Effects of Active Labor Market Policies: The Role of Labor Force Status Sequences. *Labour Economics* 15(6), S. 1270–1295.
- Manski, Charles F. (1995): *Identification Problems in the Social Sciences*. Cambridge: Harvard University Press.
- Manski, Charles F. (2013): *Public Policy in an Uncertain World: Analysis and Decisions*. Cambridge: Harvard University Press.
- Paloyo, Alfredo; Schaffner, Sandra und Schmidt, Christoph M. (2013): Special Issue on the Economic Effects of Minimum Wages in Germany: Editorial. *German Economic Review* 14(3), S. 255–257.
- Rotte, Ralph und Schmidt, Christoph M. (2003): On the Production of Victory: Empirical Determinants of Battlefield Success in Modern War. *Defence and Peace Economics* 14(3), S. 175–192.
- RWI und Stifterverband Wissenschaftsstatistik (2010): Innovationsbericht 2009 – Zur Leistungsfähigkeit des Landes Nordrhein-Westfalen in Wissenschaft, Forschung und Technologie, *RWI Projektberichte*.
- Schmidt, Christoph M. und Weigert, Benjamin (2013): Weathering the crisis

- and beyond: Perspectives for the Euro area. *International Tax and Public Finance* 20(4), S. 564–595.
- Schmidt, Christoph M. (2006): Fokus, Fokus, Fokus? Zur künftigen Rolle der außeruniversitären Wirtschaftsforschungsinstitute. *AStA – Allgemeines Statistisches Archiv* 90(4), S. 617–622.
- Schmidt, Christoph M. (2007): Policy Evaluation and Economic Policy Advice. *Advances in Statistical Analysis* 91(4), S. 379–389.
- Schmidt, Christoph M. (2009): Wirtschaftswissenschaft und Politikberatung in Deutschland – Bedeutung, Möglichkeiten und Grenzen der Kausalanalyse, In: Belke, Ansgar; Kotz, Hans-Helmut; Paul, Stephan und Schmidt, Christoph M. (Hrsg.): *Wirtschaftspolitik im Zeichen europäischer Integration: Festschrift für Wim Kösters*. RWI: Schriften 83, S. 19–36.
- Schmidt, Christoph M. (2012): Die demografische Herausforderung – ein aktueller Überblick. In: Hüther, Michael und Naeyele, Gerhard (Hrsg.): *Demografiepolitik – Herausforderungen und Handlungsfelder*. Heidelberg: Springer, S. 71–95.
- Schmidt, Christoph M. (2014): Wirkungstreffer erzielen – Die Rolle der evidenzbasierten Politikberatung in einer aufgeklärten Gesellschaft. *Perspektiven der Wirtschaftspolitik* 15(3): S. 219–233.
- Schmidt, Christoph M. und aus dem Moore, Nils (2013): Wohlstandsmessung heute: Statistische Grundlagen einer umfassenden gesellschaftlichen Debatte. In: Theurl, Theresia (Hrsg.): *Akzeptanzprobleme der Marktwirtschaft: Ursachen und wirtschaftspolitische Konsequenzen*. Berlin: Duncker & Humblot, S. 13–32.
- Schmidt, Christoph M.; aus dem Moore, Nils und Themann, Michael (2013b): Mission Impossible? Zur Verbindung von Politikberatung und »Spitzenforschung«. *Wirtschaftsdienst* 93(8), S. 511–515.
- Tauchmann, Harald; Lenz, Silja; Requate, Till und Schmidt, Christoph M. (2013): Tobacco and Alcohol: Complements or Substitutes? A Structural Model Approach to Insufficient Price Variation in Individual-Level Data. *Empirical Economics* 45(1), S. 539–566.