

Christoph M. Schmidt, Jörg Ankel-Peters

Von der Glaubwürdigkeitsrevolution zur Komplexität

Wirtschaftswissenschaft in disruptiven Zeiten

Die „Glaubwürdigkeitsrevolution“ hat die empirische Wirtschaftsforschung durch designbasierte Forschung und den Fokus auf kausale Identifikation grundlegend verändert. Doch ihr Instrumentarium reicht für disruptive Transformationsfragen nicht aus. Neben verbleibenden Problemen der Robustheit und Übertragbarkeit stellt sich eine Relevanzfrage: Designbasierte Forschung setzt häufig hinreichend stabile Strukturen voraus, während Transformationen gerade deren Veränderung bedeuten. Wir plädieren deshalb dafür, die mit der Glaubwürdigkeitsrevolution verbundene Komplexitätsreduktion aufzubrechen und eine größere Bereitschaft zum Umgang mit Komplexität zu entwickeln. Dazu gehören komplexitätsbereitere Modellierung, die Öffnung gegenüber anderen epistemischen Traditionen und insbesondere ein systematischerer Umgang mit Expertise. Strukturierte Expertenerhebungen und institutionell organisierte Pluralität können hierzu beitragen.

Wissenschaft steht im Dienst der Gesellschaft, die sie finanziert. Für die Leibniz-Gemeinschaft stehen deshalb die gesellschaftliche Relevanz von Forschung und Politikberatung im Zentrum ihrer Mission (Leibniz-Gemeinschaft, 2019). In diesem Beitrag befassen wir uns mit der Frage, wie die Wirtschaftswissenschaft diesem Leitmotiv in Zeiten teils disruptiver Transformation gerecht werden kann. Eine trennscharfe Abgrenzung von Disruption ist dabei nicht unser Anliegen. Doch erscheint es offensichtlich, dass beispielsweise der demografische Wandel, die rasanten Entwicklungen der künstlichen Intelligenz und die unumgängliche Abkehr vom fossil dominierten Energiesystem die Gesellschaft tiefgreifender verändern als andere Entwicklungen der vergangenen Jahrzehnte.

Politik muss daher Entscheidungen über eine Zukunft treffen, deren Strukturen sich fundamental verändern. Zugleich ist der gesellschaftliche Anspruch an Politik gewachsen, langfristig zu gestalten und Verantwortung für Entwicklungen zu übernehmen, die teils weit in der Zukunft liegen. Dies zeigt sich am deutlichsten in der klimapolitischen Debatte, aber ebenso in Diskussionen um den

demografischen Wandel und den Themenkomplex Zuwanderung und Integration, geopolitische Herausforderungen oder die Prävention künftiger Krisen. Wie sollten gesellschaftsorientierte Forschung und wirtschaftspolitische Beratung vor diesem Hintergrund aussehen?

Dieser Frage gehen wir in diesem Beitrag nach, indem wir zunächst rekapitulieren, wo die empirischen Wirtschaftswissenschaften herkommen, dann die Defizite des Status quo identifizieren und schließlich Vorschläge für ein Update der wirtschaftswissenschaftlichen Forschung umreißen. Damit knüpfen wir an eine Reihe kritischer Beiträge innerhalb der Ökonomik an, die aus unterschiedlichen Perspektiven Grenzen der vorherrschenden empirischen Forschung für wirtschaftspolitische Schlussfolgerungen

Prof. Dr. Dr. h. c. Christoph M. Schmidt
ist Präsident der Leibniz-Gemeinschaft und
Professor an der Ruhr-Universität Bochum.

Prof. Dr. Jörg Ankel-Peters ist Ko-Leiter
des Kompetenzbereichs „Klima- und
Entwicklungspolitik“ am RWI – Leibniz-Institut
für Wirtschaftsforschung und Professor für
Volkswirtschaftslehre an der Universität Passau.

© Der/die Autor:in 2026. Open Access: Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht (creativecommons.org/licenses/by/4.0/deed.de).

Open Access wird durch die ZBW – Leibniz-Informationszentrum Wirtschaft gefördert.

aufgezeigt haben (Manski, 2011; Basu, 2014; Deaton & Cartwright, 2018; Rodrik, 2015; Ankel-Peters & Schmidt, 2025; Connolly et al., 2025).

Die Glaubwürdigkeitsrevolution und ihre Grenzen

In den vergangenen Jahrzehnten hat die Ökonomik eine Verschiebung hin zur empirischen Unterfütterung ihrer Aussagen erlebt (Hamermesh, 2025). In Deutschland wurde dies nicht zuletzt getragen von dem Neuaufbruch, den die großen außeruniversitären Institute – typischerweise Mitgliedsinstitute der Leibniz-Gemeinschaft – seit der Jahrhundertwende durchlaufen haben. Noch vor wenigen Jahrzehnten spielten in der deutschen wirtschaftspolitischen Beratung theorie- und ordnungspolitisch begründete Plausibilitätsüberlegungen eine wesentlich größere Rolle (Hien, 2024). Seither hat sich die Wirtschaftsforschung grundlegend verändert.

Die klassische Ordnungsökonomik wurde in Deutschland ebenso zurückgedrängt wie international das vorschnelle Schließen von Korrelationen auf Wirkungszusammenhänge. An seine Stelle ist zunehmend ein empirischer Ansatz getreten, der beansprucht, kausale Zusammenhänge nachvollziehbar zu identifizieren – die sogenannte Glaubwürdigkeitsrevolution (Angrist & Pischke, 2010). Diese designbasierte Forschung (Card, 2022) ist eng mit der Idee der evidenzbasierten Politik verbunden: Glaubwürdig identifizierte Kausalzusammenhänge sollten die Basis informierter politischer Entscheidungen sein (Boockmann et al., 2014).

Designbasierte Forschung ist eine wissenschaftliche Erfolgsgeschichte, deren methodischer Einfluss inzwischen weit über die Ökonomik hinausreicht (Torreblanca et al., 2026). Doch ihr Glaubwürdigkeitsversprechen ist nur teilweise eingelöst, denn es mehren sich die Anzeichen, dass auch die moderne empirische Forschung weniger belastbar ist als erhofft (Brodeur et al., 2020, 2026a; Ioannidis et al., 2017).

Das liegt paradoxerweise auch am Erfolg der empirischen Wende selbst. Die Glaubwürdigkeitsrevolution nahm ihren Ausgang in einer Welt relativer Datenknappheit. Geeignete Daten waren rar, überzeugende experimentelle oder quasi-experimentelle Situationen ebenso. Die empirischen Möglichkeiten haben sich seither fundamental verändert. Administrative und digitale Daten, immer neue Möglichkeiten ihrer Verknüpfung und erheblich gestiegene Rechenkapazitäten eröffnen heute einen ungleich größeren Raum möglicher empirischer Analysen.

Gerade dieser gewachsene Raum möglicher Analysen stellt neue Anforderungen an die Glaubwürdigkeit em-

pirischer Forschung. Der Logik des Hypothesentestens folgend sollte eine theoretisch begründete Hypothese formuliert werden, *bevor* sie mit den Daten konfrontiert wird. Werden Hypothesen hingegen erst in Kenntnis der empirischen Ergebnisse formuliert und anschließend anhand derselben Daten vermeintlich getestet, wird diese Logik verletzt. Die Daten dienen dann zugleich zur Formulierung der Hypothese und als Evidenz für ihre Gültigkeit. Ein solches Vorgehen wird als HARKing („hypothesizing after the results are known“) bezeichnet (Kerr, 1998).

Wo dies dennoch als Hypothesentest präsentiert wird, führt die empirische Analyse in die Irre: Sie suggeriert eine empirische Bewährung der Theorie, obwohl Theorie und Befund mit denselben Daten erzeugt wurden. Je größer der Raum möglicher Analysen – und in der modernen Datenwelt ist dieser Raum enorm –, desto größer ist der Spielraum für bewusstes oder unbewusstes HARKing. Umso entscheidender ist die Unterscheidung, ob eine Hypothese vor oder nach dem Blick auf die Daten formuliert wurde.

Die Präspezifizierung von Hypothesen und zentralen Analyseschritten sollte deshalb auch in der empirischen Wirtschaftsforschung weit stärker zur gängigen Praxis werden (Christensen & Miguel, 2018). Solange dies nicht der Fall ist, bleibt das Glaubwürdigkeitsversprechen der Glaubwürdigkeitsrevolution angreifbar. Tatsächlich sind die Transparenzstandards in der Profession zwar merklich gestiegen, nicht zuletzt durch inzwischen weit verbreitete Data-Sharing-Policies (Brodeur et al., 2026b), die systematische Robustheitsüberprüfung durch Replikation ermöglichen. Zweifel bleiben jedoch, ob die Kultur der Selbstreflexion und -kritik hinreichend ausgeprägt ist und bestehende Transparenzstandards selektivem Berichten und HARKing wirksam entgegenwirken (Ankel-Peters et al., 2023, 2025; Ofosu & Posner, 2023).

Die Glaubwürdigkeitsrevolution braucht daher zunächst ein Update aus sich selbst heraus. Präregistrierung, Replikation und eine klare Trennung zwischen Hypothesengenerierung und Hypothesenprüfung sollten selbstverständlicher Bestandteil empirischer Ausbildung und Forschung werden. Doch damit wäre lediglich das Glaubwürdigkeitsproblem adressiert.

Von der Glaubwürdigkeits- zur Relevanzfrage

Unser zweites Argument für ein Update der empirischen Wirtschaftsforschung setzt an einer anderen Stelle an: Selbst wenn die designbasierte Kausalforschung für sich genommen vollkommen glaubwürdige Ergebnisse liefert, ist der Ansatz nur begrenzt geeignet, die großen anstehenden Transformationsentscheidungen zu informieren.

Der Grund liegt in der Komplexitätsreduktion, auf der ihre besondere Glaubwürdigkeit wesentlich beruht. Für disruptive Transformationsfragen wird diese in zweierlei Hinsicht zum Problem.

Erstens sind die Methoden der designbasierten Forschung dort am stärksten, wo Interventionen randomisiert oder zumindest unter kontrollierten Bedingungen variiert werden können. Das ist in der Lebenswirklichkeit eher bei begrenzten Interventionen der Fall. Ein idealtypisches Beispiel bietet die umfangreiche Literatur zu Nudging und Verhaltensökonomik. Hier werden durchaus nützliche Erkenntnisse erzielt. Die großen Transformationsfragen lassen sich auf diese Weise jedoch kaum untersuchen (Connolly et al., 2025).

Geht es um weitreichendere Politikinterventionen, kommen in der designbasierten Forschung daher andere Identifikationsstrategien zum Einsatz, etwa Panelanalysen oder natürliche Experimente, die zeitliche, geografische oder anderweitig exogene Variation nutzen. Auch wenn sich bei diesen Ansätzen die diskutierten Glaubwürdigkeitsprobleme besonders stellen können und ihre kausale Interpretation teils starke Identifikationsannahmen erfordert (Haveresch et al., 2026; Mellon, 2025; de Chaisemartin & D'Haultfœuille, 2020; Goodman-Bacon, 2021), können sie wichtige Erkenntnisse liefern. Doch die für große Transformationsentscheidungen relevante Variation ist häufig gerade nicht beobachtbar.

Das zeigt sich exemplarisch in der Klimapolitik. Die designbasierte Forschung ist nicht in der Lage, die ökonomischen und gesellschaftlichen Auswirkungen einer Dekarbonisierung¹ der Volkswirtschaft zu untersuchen. Gleiches gilt für die tiefgreifenden Politikmaßnahmen, mit denen eine solche Transformation gestaltet werden soll, wie Technologiepolitik oder substanzielle CO₂-Bepreisung. Solche Politikmaßnahmen werden zwar bisweilen designbasiert untersucht, doch kommen dabei Varianten des randomisierten Lehrbuchbeispiels zum Einsatz, deren „Glaubwürdigkeit“ auf sehr speziellen Annahmen basiert. So muss etwa unterstellt werden, dass aus anderen Regionen, Ländern oder Zeiträumen ein valides kontrafaktisches Szenario konstruiert werden kann, das hinreichend genau abbildet, wie sich die behandelte Einheit ohne die Politikmaßnahme entwickelt hätte. Dabei stellen sich oftmals in besonderem Maße Belastbarkeitsfragen, die wir im vorangegangenen Kapitel umrissen haben (Ferman et al., 2020; Piseddu et al., 2026).

¹ Es geht eigentlich um eine „Defossilisierung“, also Transformation von einem überwiegend auf der Nutzung von fossilen Energieträgern beruhenden Energiesystem zu einem nicht-fossil basierten Energiesystem.

Der *zweite* Grund ist eine Variante der Kritik an der externen Validität empirischer Studien. Ist ein glaubwürdig identifizierter Zusammenhang aus der Vergangenheit auch relevant für eine Entscheidung über die Zukunft? Diese Frage ist keineswegs neu. Sie gewinnt jedoch eine andere Qualität, wenn nicht lediglich Parameter innerhalb bestehender Strukturen variieren, sondern diese Strukturen selbst aufbrechen. In disruptiven Transformationen verändern sich nicht nur relative Preise. Es kann sich verändern, ob ein Gut überhaupt knapp ist, ob es einen Markt und damit einen Preis gibt, oder ob aus einer ökonomischen Knappheit eine absolute Beschränkung wird. So werden demografischer Wandel und Zuwanderung heute wesentlich vor dem Hintergrund eines absehbar knapper werdenden Arbeitsangebots diskutiert. Eine disruptive KI-Transformation könnte jedoch Art und Umfang der Arbeitskräftenachfrage grundlegend verändern. Der Zusammenhang zwischen Demografie, Arbeitskräftebedarf und Zuwanderung müsste selbst neu bestimmt werden.

Die Grenzen der Aussagekraft designbasierter Forschung werden umso bedeutsamer, je länger die politischen Gestaltungshorizonte sind. Klimapolitik setzt sich Ziele für Zeiträume von 25 Jahren und mehr. Je weiter solche Ziele in die Zukunft reichen, desto größer ist die Möglichkeit, dass technologische Disruptionen die Bedingungen verändern, unter denen sie erreicht werden sollen. Wenn etwa grüner Wasserstoff oder Verfahren zur Entnahme von CO₂ aus der Atmosphäre in 15 Jahren zu geringen Kosten und in großem Maßstab verfügbar wären, hätte dies starke Auswirkungen darauf, welche klimapolitischen Maßnahmen wir heute für sinnvoll halten. Transformationsentscheidungen müssen dabei vielfach unter tiefer (sogenannter Knight'scher) Unsicherheit getroffen werden, d. h., für manche politikrelevante Szenarien lassen sich nicht einmal belastbare Wahrscheinlichkeiten angeben (Sunstein, 2021). Vergangenheitsbasierte Evidenz stößt hier an Grenzen (Heider et al., 2026).

Damit ist nicht gesagt, dass designbasierte Forschung keinerlei Bedeutung für die Politikberatung haben sollte. Wo relevante Variation beobachtbar ist und die untersuchten Strukturen hinreichend stabil sind, bleibt die mit designbasierter Forschung verbundene Komplexitätsreduktion sinnvoll und die Methode ein nützliches Instrument. Doch die Ökonomik muss ihren Horizont erweitern, um die anstehenden großen gesellschaftlichen Fragen zu adressieren.

Komplexitätsbereitschaft

Mit der seit den 1980er und 1990er Jahren einsetzenden Hinwendung weiter Teile der empirischen Ökonomik zu

experimentellen und quasi-experimentellen Forschungsdesigns verlor eine andere epistemische Tradition relativ an Bedeutung: die strukturelle Modellierung ökonomischer Zusammenhänge (Panhans & Singleton, 2017). Diese Entwicklung war keineswegs unumstritten. Vertreter der Strukturmodellierung warnten – weitgehend ohne durchzudringen – davor, dass der Gewinn an Glaubwürdigkeit mit einer Beschränkung der Fragen erkaufte werden könnte, die sich empirisch beantworten lassen (Heckman & Vytlacil, 2005; Keane, 2010; Rust, 2010).

Aufschlussreich ist es, etwa Levitt (2025) folgend, den Versuch, komplexe wirtschaftliche Strukturen explizit zu modellieren, mit der Klimawissenschaft zu vergleichen. Diese setzt zwangsläufig auf die Modellierung des überaus komplexen Weltklimas, weil ein klassisch-experimenteller Zugang offensichtlich nicht möglich ist. Levitt argumentiert, dass die Ökonomik vor rund 50 Jahren ganz ähnliche Ambitionen verfolgt habe: komplexe strukturelle Modelle mit zahlreichen Gleichungen, mit denen gesamtwirtschaftliche Gleichgewichte abgebildet und alternative Szenarien simuliert werden sollten. Die Profession habe diesen Weg jedoch weitgehend aufgegeben, weil die Modelle aus damaliger Sicht menschliches Verhalten und wirtschaftliche Entwicklungen nicht hinreichend gut vorhergesagt hätten.

In der Tat weisen Klima und Volkswirtschaft trotz aller Unterschiede eine wichtige Gemeinsamkeit auf: Beide sind nicht lediglich komplizierte, sondern komplexe Systeme (Arthur, 2021; Curry, 2023). Ihre Entwicklung entsteht aus zahlreichen simultanen Beziehungen, Rückkopplungen und nichtlinearen Prozessen und lässt sich daher nicht deterministisch aus dem Verhalten einzelner Komponenten ableiten (Farmer & Foley, 2009). Die Klimawissenschaft hat gerade deshalb eine modellbasierte epistemische Praxis entwickelt, die nicht auf punktgenaue Vorhersagen zielt, sondern Szenarien, Sensitivitäten und Bandbreiten untersucht (Stainforth, 2023).

Hat die Ökonomik das Potenzial komplexitätsbereiter Modellierung möglicherweise zu früh aufgegeben? Diese Frage ist nicht nur ideengeschichtlich relevant. Denn gerade angesichts rapide gewachsener Rechenkapazitäten könnten solche Formen der Modellierung für die anstehenden langfristigen Transformationsfragen wieder stärker an Bedeutung gewinnen.

Komplexitätsbereitschaft muss sich jedoch nicht auf komplexere Modelle beschränken. Auch von anderen Disziplinen lässt sich lernen. Historische und qualitative Forschungstraditionen sind stärker darauf ausgerichtet, komplexe Kausalzusammenhänge, unterschiedliche kausale Pfade und die Einbettung von Entwicklungen in

ihrem jeweiligen Kontext zu untersuchen (Mahoney & Goertz, 2006; Röth, 2023; Connolly et al., 2025). Der Wert all dieser Ansätze liegt dabei vor allem darin, dass sie ein tieferes Verständnis ihres Gegenstands ermöglichen, indem sie komplexe kausale Prozesse und Zusammenhänge in den Blick nehmen. Damit führt die Frage nach komplexitätsbereiter Forschung unmittelbar zu einer weiteren Wissensquelle, deren Rolle in der modernen Ökonomik bislang wenig systematisch untersucht wird: der Expertise.

Expertise

Expertise ist eine zentrale epistemische Ressource für eine komplexitätsbereitere Wirtschaftswissenschaft. In diesem Abschnitt fragen wir deshalb, was Expertise überhaupt ist und wie sie identifiziert werden kann, ehe wir uns im nächsten Abschnitt ihrer Nutzbarmachung widmen. Bei aller Kritik an den empirischen Wirtschaftswissenschaften muss man konstatieren: In der Politikberatung und in wirtschaftspolitischen Gremien haben Ökonom:innen sich komplexen Fragen schon immer auf Grundlage ihrer Expertise gestellt. Diese Ressource sollte auch die akademische Wirtschaftsforschung stärker und systematischer nutzen.

Was aber ist Expertise? Expertise ist weder gleichbedeutend mit Evidenz noch schlicht mit Expertenmeinung. Collins und Evans (2002) unterscheiden unterschiedliche Formen von Expertise. Für unsere Zwecke ist zunächst vor allem die Abgrenzung zwischen über längere Zeit gewachsener Expertise und kurzfristig erwerbbaarem Wissen relevant. Wer Fakten und Forschungsergebnisse eines Feldes kennt, verfügt noch lange nicht über diesbezügliche Expertise. Expertise entsteht durch tiefe Auseinandersetzung mit einem Wissensgebiet und, ganz zentral, durch Sozialisation in dessen Fachgemeinschaft sowie durch praktische Erfahrung in Forschung, Anwendung oder Politikberatung.

Dabei wird auch Wissen erworben, das sich nicht ohne Weiteres kodifizieren und weitergeben lässt – tacit knowledge. Für die Wirtschaftsforschung umfasst dies etwa die Fähigkeit einzuschätzen, welche Annahmen plausibel, welche Befunde belastbar und welche Erkenntnisse auf einen konkreten Fall übertragbar sind. Evidenz lässt sich dokumentieren und weitergeben; Expertise umfasst zusätzlich die Kompetenz, ihre Bedeutung und Reichweite zu beurteilen. Sie ist damit das, was einer Expertenmeinung epistemisches Gewicht verleihen kann.

Collins und Evans (2002) gehen jedoch darüber hinaus. Sie unterscheiden zwischen zwei Arten von Expertise, contributory expertise und interactional expertise. Wer heute als wissenschaftlicher Experte gilt, verfügt typi-

scherweise mindestens über *contributory expertise*: Man hat Forschung in einem bestimmten Feld nicht nur eigenhändig durchgeführt, sondern ist durch langjährige wissenschaftliche Praxis in diesem Feld sozialisiert und verfügt somit nicht nur über das kodifizierte Wissen, sondern auch das *tacit knowledge*. Bereits diese klassische Form von Expertise ist von außen keineswegs leicht zu identifizieren. Wissenschaftliche Publikationen und Reputation dienen als Signale, messen die zugrunde liegende Expertise aber nur unvollkommen (Heckman & Moktan, 2020). Hinzu kommen die diskutierten Grenzen des wissenschaftlichen Qualitätssicherungsprozesses. Wer in einem Forschungsfeld sichtbar erfolgreich ist, muss deshalb noch nicht über genau jene *contributory expertise* verfügen, die für die Beurteilung einer konkreten wirtschaftspolitischen Frage benötigt wird.

Noch weniger sichtbar ist eine zweite Form von Expertise, die für komplexe transformative Politikberatung unseres Erachtens unerlässlich ist. Collins und Evans bezeichnen sie als *interactional expertise*: die Fähigkeit, die Sprache, Wissensbestände und Problemwahrnehmungen einer anderen Expertengemeinschaft so weit zu beherrschen, dass deren Wissen verstanden und mit dem eigenen in Beziehung gesetzt werden kann. *Interactional expertise* ist zentral für das vorgeschlagene Update der Wirtschaftswissenschaften. Denn die für die Transformation relevanten Wissensbestände liegen typischerweise nicht innerhalb eines einzelnen Forschungsfeldes – und oftmals nicht einmal innerhalb einer Disziplin. Wer etwa die Folgen künstlicher Intelligenz für Arbeitsmarkt, Energiebedarf und gesellschaftliche Teilhabe beurteilen will, muss Erkenntnisse unterschiedlicher wissenschaftlicher und praktischer Wissensgemeinschaften zusammenführen können.

Während *contributory expertise* – optimistisch betrachtet – über das etablierte wissenschaftliche Reputationssystem zumindest näherungsweise identifiziert werden kann, fehlen für *interactional expertise* vergleichbare Mechanismen. Ihr Erwerb erfolgt eher beiläufig und wird wissenschaftlich kaum systematisch honoriert. Das sollte sich ändern. Wirtschaftsforschung und wirtschaftspolitische Beratung in disruptiven Zeiten brauchen eine Vielfalt spezialisierter Expertisen – und Menschen, die in der Lage sind, diese über disziplinäre und institutionelle Grenzen hinweg miteinander in Beziehung zu setzen.

Interactional expertise entsteht nicht von selbst. Sie entsteht nicht zuletzt durch inter- und transdisziplinäre Erfahrung. Interdisziplinäre Forschung ist deshalb mehr als die Kombination unterschiedlicher Methoden oder die arbeitsteilige Bearbeitung eines gemeinsamen Projekts. Im besten Fall sozialisiert sie Forschende in die Sprache,

Problemwahrnehmungen und Wissensbestände anderer Disziplinen. Damit schafft sie ebenso *interactional expertise* wie transdisziplinäre Forschung, also die dauerhafte Interaktion zwischen Wissenschaft, Politik und Praxis. Solches Wissen, etwa über institutionelle Restriktionen, die Umsetzbarkeit politischer Maßnahmen oder die Funktionsweise konkreter Märkte und Organisationen, erschließt sich nicht allein aus der Lektüre wissenschaftlicher Publikationen.

Gerade hierin liegt eine besondere Stärke der Leibniz-Gemeinschaft: Sie vereint unterschiedliche Disziplinen, Forschungsinfrastrukturen und institutionelle Schnittstellen zur Praxis. Diese Vielfalt ist für Politikberatung nicht nur deshalb wertvoll, weil sie viele Stimmen hervorbringt. Sie kann die Voraussetzungen dafür schaffen, dass unterschiedliche Expertisen einander verstehen und zu einem Urteil über komplexe Transformationsfragen zusammengeführt werden (Fuchs-Schündeln et al., 2026).

Expertise nutzbar machen: Abduktion und Expertensurveys

Wie aber lässt sich Expertise systematischer für Forschung und Politikberatung nutzbar machen? Die wirtschaftswissenschaftliche Forschung sollte Ansätze entwickeln und testen, um Expertenwissen systematisch zu erheben und auszuwerten. Besonders erkenntnisreich sind Expertenurteile dann, wenn der Erhebungsansatz Expert:innen dazu bringt, unterschiedliche Wissensquellen auf die konkrete Fragestellung anzuwenden und miteinander zu verbinden – eine Form abduktiver Inferenz (Heckman & Singer, 2017). Expertise kann so empirische Evidenz mit theoretischem Wissen, Erfahrungen aus Politik und Praxis sowie *tacit knowledge* über Institutionen und Zusammenhänge verbinden. Bei komplexen und neuartigen Fragestellungen kann ein solches Urteil Erkenntnisse ermöglichen, die sich aus den verfügbaren Daten allein nicht gewinnen lassen.

Eine solche epistemische Praxis findet natürlich in Teilen längst statt. Sie bleibt jedoch meist implizit und wird wissenschaftlich kaum systematisch erschlossen. Einzelne Expert:innen werden in Gremien berufen, von Ministerien konsultiert oder um öffentliche Einschätzungen gebeten. Doch Expertise bedeutet nicht, dass einzelne Expert:innen über das gesamte relevante Wissen verfügen. Auch unter Expert:innen unterscheiden sich Erfahrungen, Wissensbestände und Einschätzungen – und ihre Urteile können durch persönliche Interessen, Präferenzen oder normative Überzeugungen geprägt sein. Gerade deshalb liegt es nahe, Expertise systematisch und in größerem Maßstab zu erschließen.

Dieser Vorschlag beginnt keineswegs bei null. Unter Begriffen wie „structured expert elicitation“ oder „structured expert judgment“ existiert eine etablierte Literatur dazu, wie Expert:innen ausgewählt, ihre Einschätzungen systematisch erhoben, Unsicherheiten quantifiziert und unterschiedliche Urteile zusammengeführt werden können (Cooke, 1991; O’Hagan et al., 2006; Aspinall, 2010; Morgan, 2014; Hemming et al., 2018; Bamber et al., 2019). Wir plädieren dafür, solche Ansätze wesentlich breiter und skalierbarer in der wirtschafts- und sozialwissenschaftlichen Forschung einzusetzen. Expert Surveys könnten nicht einzelne Expert:innen oder kleine Gremien, sondern Dutzende oder Hunderte sorgfältig ausgewählte Expert:innen mit denselben Fragen konfrontieren.

Ziel wäre dabei nicht, wissenschaftliche Wahrheit durch Mehrheitsentscheid zu erzeugen. Vielmehr ließe sich das verteilte Wissen einer Expertengemeinschaft erschließen, robuste Gemeinsamkeiten von idiosynkratischen Einschätzungen unterscheiden und zugleich sichtbar machen, wo auch unter Expert:innen erhebliche Unsicherheit oder Dissens besteht. Dafür gibt es Vorbilder sowohl in einzelnen Forschungsprojekten als auch bei der systematischen Erhebung des Wissensstands.

Schluss

Die Errungenschaften der Glaubwürdigkeitsrevolution werden oft mit der naiven Korrelationsforschung kontrastiert, wie sie in den Jahrzehnten zuvor auch in den Wirtschaftswissenschaften weit verbreitet war (Leamer, 1983). Richtig ist aber auch, dass die Hinwendung zu designbasierter Forschung eine implizite epistemische Komplexitätsreduktion mit sich gebracht hat. In diesem Beitrag haben wir dafür geworben, sich der Komplexität wieder stärker zuzuwenden – auch wenn dies bedeutet, einen Teil jener Einfachheit und Nachprüfbarkeit aufzugeben, auf der der besondere Rigoritätsanspruch der modernen empirischen Ökonomik beruht.

Für die anstehenden tiefgreifenden Transformationen ist eine Wissenschaft, die nur dort spricht, wo sie unter vergleichsweise einfachen Annahmen besonders glaubwürdige Aussagen treffen kann, zu wenig. Mit dieser Öffnung für größere epistemische Ambiguität muss die Wirtschaftswissenschaft aber auch ihren Deutungshoheitsanspruch zurücknehmen. Erkenntnisse einzelner Forschender sind dann stärker als Beiträge zu einer gesellschaftlichen Deliberation zu verstehen, in die auch andere Disziplinen und eine größere Vielfalt ökonomischer Perspektiven einfließen.

Mittelfristig könnten moderne und weiter wachsende Rechenkapazitäten neue Formen der Simulation komplexer

ökonomischer Systeme ermöglichen, orientiert an der Modellierungspraxis der Klimawissenschaft. Auch sollte die Ökonomik ihre methodische Sonderstellung unter den Sozialwissenschaften hinterfragen und sich stärker für komplexitätsorientierte Herangehensweisen ihrer Nachbardisziplinen öffnen (Fourcade et al., 2015). Eine solche Öffnung rückt zugleich die Rolle von Expertise stärker in den Vordergrund.

Hier sollte die Disziplin ein belastbares Verständnis davon entwickeln, was Expertise ist, und Verfahren etablieren, um sie systematischer für Forschung und Politikberatung nutzbar zu machen. Expertise ist dabei kein Ersatz für Evidenz. Sie hilft, Evidenz zu interpretieren und mit anderen Wissensquellen zu verbinden. Umgekehrt kann Expertise selbst zum Gegenstand der Evidenzgewinnung werden – indem wir empirisch untersuchen, wer über Expertise verfügt und wie sich diese systematisch erschließen lässt.

Wenn komplexe und disruptive Probleme unterschiedliche Formen von Expertise verlangen, brauchen wirtschaftswissenschaftliche Forschung und Politikberatung zudem institutionell organisierte Pluralität. Die Zusammenarbeit über disziplinäre und institutionelle Grenzen hinweg ist dafür keine Zugabe, sondern Voraussetzung. Die Leibniz-Gemeinschaft erscheint prädestiniert dafür, Teile dieser Entwicklung in Deutschland voranzutreiben.

Literatur

- Angrist, J. D. & Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, 24(2), 3–30.
- Ankel-Peters, J., Fiala, N. & Neubauer, F. (2023). Do Economists Replicate? *Journal of Economic Behavior & Organization*, 212, 219–232.
- Ankel-Peters, J. & Schmidt, C. M. (2025). Rural Electrification, the Credibility Revolution, and the Limits of Evidence-Based Policy. *Environment and Development Economics*, 30(1), 70–81.
- Ankel-Peters, J., Fiala, N. & Neubauer, F. (2025). Is Economics Self-Correcting? Replications in the American Economic Review. *Economic Inquiry*, 63(2), 463–485.
- Arthur, W. B. (2021). Foundations of Complexity Economics. *Nature Reviews Physics*, 3, 136–145.
- Aspinall, W. (2010). A route to more tractable expert advice. *Nature*, 463, 294–295.
- Bamber, J. L., Oppenheimer, M., Kopp, R. E., Aspinall, W. P. & Cooke, R. M. (2019). Ice sheet contributions to future sea-level rise from structured expert judgment. *Proceedings of the National Academy of Sciences*, 116(23), 11195–11200.
- Basu, K. (2014). Randomisation, Causality and the Role of Reasoned Intuition. *Oxford Development Studies*, 42(4), 455–472.
- Boockmann, B., Buch, C. M. & Schnitzer, M. (2014). Evidenzbasierte Wirtschaftspolitik in Deutschland: Defizite und Potentiale. *Perspektiven der Wirtschaftspolitik*, 15(4), 307–323.
- Brodeur, A., Cook, N. & Heyes, A. (2020). Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics. *American Economic Review*, 110(11), 3634–3660.
- Brodeur, A., Bruns, S. B., Herwartz, H. & Islam, C. G. (2026b). How Did Empirical Economic Research Improve over the Last Decades: Re-

- cent Trends and Prospects. *European Economic Review*, erscheint demnächst.
- Brodeur, A., Mikola, D., Cook, N., Fiala, L., Brailey, T., Briggs, R., de Grendre, A., Dupraz, Y., Gabani, J., Gauriot, R., Haddad, J., Lima, G., Ankel-Peters, J., Dreber, A., Campbell, D., Kattan, L., Marino Fages, D., Mierisch, F., Sun, P. [...] & Zhong, Y. (2026a). Reproducibility and Robustness of Economics and Political Science Research. *Nature*, 652, 151–156.
- Card, D. (2022). Design-Based Research in Empirical Microeconomics. *American Economic Review*, 112(6), 1773–1781.
- de Chaisemartin, C. & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9), 2964–2996.
- Christensen, G. & Miguel, E. (2018). Transparency, Reproducibility, and the Credibility of Economics Research. *Journal of Economic Literature*, 56(3), 920–980.
- Collins, H. M. & Evans, R. (2002). The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science*, 32(2), 235–296.
- Connolly, D. J., Loewenstein, G. & Chater, N. (2025). An S-Frame Agenda for Behavioral Public Policy Research. *Behavioural Public Policy*, 9(3), 593–613.
- Cooke, R. M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Oxford University Press.
- Curry, J. A. (2023). *Climate Uncertainty and Risk: Rethinking Our Response*. Anthem Press.
- Deaton, A. & Cartwright, N. (2018). Understanding and Misunderstanding Randomized Controlled Trials. *Social Science & Medicine*, 210, 2–21.
- Farmer, J. D. & Foley, D. (2009). The economy needs agent-based modelling. *Nature*, 460(7256), 685–686.
- Ferman, B., Pinto, C. & Possebom, V. (2020). Cherry Picking with Synthetic Controls. *Journal of Policy Analysis and Management*, 39(2), 510–532.
- Fourcade, M., Ollion, E. & Algan, Y. (2015). The Superiority of Economists. *Journal of Economic Perspectives*, 29(1), 89–114.
- Fuchs-Schündeln, N., Kübler, D. & Palermo Kuss, A. H. (2026). Blinde Flecken in der Polykrise: Mehr Diversität führt zu einer soliden Wirtschaftspolitik. *Wirtschaftsdienst*, 106(8).
- Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*, 225(2), 254–277.
- Hamermesh, D. S. (2025). The Demographic and Research Styles of Economics Writing. *SSRN Electronic Journal*.
- Haveresch, N., Ankel-Peters, J. & Bensch, G. (2026). A Slippery Slope: Topographic Variation as an Instrument. *USAAE Working Paper*, No. 26-676.
- Heckman, J. J. & Moktan, S. (2020). Publishing and Promotion in Economics: The Tyranny of the Top Five. *Journal of Economic Literature*, 58(2), 419–470.
- Heckman, J. J. & Singer, B. (2017). Abducting Economics. *American Economic Review: Papers & Proceedings*, 107(5), 298–302.
- Heckman, J. J. & Vytlacil, E. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica*, 73(3), 669–738.
- Heider, F., Lindner, V. & Schaffranka, C. (2026). Wie evidenzbasierte Politikberatung in Krisenzeiten trotz unvollständiger Daten gelingt. *Wirtschaftsdienst*, 106(8).
- Hemming, V., Burgman, M. A., Hanea, A. M., McBride, M. F. & Wintle, B. C. (2018). A Practical Guide to Structured Expert Elicitation Using the IDEA Protocol. *Methods in Ecology and Evolution*, 9(1), 169–180.
- Hien, J. (2024). The rise and fall of ordoliberalism. *Socio-Economic Review*, 22(4), 1947–1966.
- Ioannidis, J. P. A., Stanley, T. D. & Doucouliagos, H. (2017). The Power of Bias in Economics Research. *Economic Journal*, 127(605), F236–F265.
- Keane, M. P. (2010). A Structural Perspective on the Experimentalist School. *Journal of Economic Perspectives*, 24(2), 47–58.
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and social psychology review*, 2(3), 196–217.
- Leamer, E. E. (1983). Let's Take the Con out of Econometrics. *American Economic Review*, 73(1), 31–43.
- Leibniz-Gemeinschaft. (2019). *Leitbild Leibniz-Transfer*.
- Levitt, S. D. (2025, 18. Juli). *Will We Solve the Climate Problem?* Gespräch mit Kate Marvel, People I (Mostly) Admire, Episode 162. Freakonomics Radio Network.
- Mahoney, J. & Goertz, G. (2006). A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research. *Political Analysis*, 14(3), 227–249.
- Manski, C. F. (2011). Policy Analysis with Incredible Certitude. *Economic Journal*, 121(554), F261–F289.
- Mellon, J. (2025). Rain, Rain, Go Away: 194 Potential Exclusion-Restriction Violations for Studies Using Weather as an Instrumental Variable. *American Journal of Political Science*, 69(3), 881–898.
- Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences*, 111(20), 7176–7184.
- Ofosu, G. K. & Posner, D. N. (2023). Pre-analysis plans: An early stocktaking. *Perspectives on Politics*, 21(1), 174–190.
- O'Hagan, A., Buck, C. E., Daneshkhan, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, J. D., Oakley, J. E. & Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley & Sons.
- Panhans, M. T. & Singleton, J. D. (2017). The Empirical Economist's Toolkit: From Models to Methods. *History of Political Economy*, 49(S1), 127–157.
- Piseddu, E., Brodeur, A., Rose, J., Sievert, M. & Ankel-Peters, J. (2026). The Power of Carbon Pricing – A Comment on Döbbling-Hildebrandt et al. (2024) and Its Press Release. *I4R Discussion Paper Series*, 283.
- Röth, L. (2023). Pathway Analysis, Causal Mediation, and the Identification of Causal Mechanisms. In A. Damonte & F. Negri (Hrsg.), *Causation in Policy Studies: A Pluralist Toolbox* (S. 111–135). Springer.
- Rodrik, D. (2015). *Economics Rules: The Rights and Wrongs of the Dismal Science*. W. W. Norton.
- Rust, J. (2010). Comments on “Structural vs. Atheoretic Approaches to Econometrics” by Michael Keane. *Journal of Econometrics*, 156(1), 21–24.
- Stainforth, D. A. (2023). *Predicting Our Climate Future: What We Know, What We Don't Know, and What We Can't Know*. Oxford University Press.
- Sunstein, C. R. (2021). *Averting Catastrophe: Decision Theory for COVID-19, Climate Change, and Potential Disasters of All Kinds*. New York University Press.
- Torreblanca, C., Dinneen, W., Grossman, G. & Xu, Y. (2026). The credibility revolution in political science. *arXiv preprint arXiv:2601.11542*.

Title: From the credibility revolution to complexity – economics in disruptive times

Abstract: The “credibility revolution” has fundamentally transformed empirical economic research through design-based research and a focus on causal identification. However, its toolkit is insufficient for addressing disruptive transformation issues. In addition to lingering problems of robustness and generalisability, there is a question of relevance: design-based research often presupposes sufficiently stable structures, whereas transformations involve precisely the alteration of such structures. We therefore advocate moving beyond the reduction in complexity associated with the credibility revolution and developing a greater willingness to engage with complexity. This includes modelling that is better equipped to handle complexity, openness to other epistemic traditions and, in particular, a more systematic approach to expertise. Structured expert surveys and institutionally organised plurality can contribute to this.