DOI: 10.2478/wd-2023-0206

Zeitgespräch

sciendo

Wirtschaftsdienst, 2023, 103(11), 741-745

JEL: E00, K39, C80

Rüdiger Bachmann, Almut Balleer, Zeno Enders, Niklas Garnadt, Heiko Stüber*

Makrodaten: Die amtliche Statistik stärken

Exzellente empirische Wirtschaftsforschung, die Evaluation wirtschafts-, finanz- und arbeitsmarktpolitischer Maßnahmen sowie evidenzbasierte Politikberatung sind angewiesen auf die Verfügbarkeit hochqualitativer Daten. Eine im Januar 2023 durchgeführte Erhebung unter den Mitgliedern des Vereins für Socialpolitik (VfS) ergab, dass viele Forschende die Möglichkeiten, deutsche Daten für die Forschung zu nutzen, für unzureichend halten. Die Mitglieder stimmen darin überein, dass fehlende Datenverknüpfungen einen der größten Mängel in der deutschen Dateninfrastruktur darstellen, und sind insgesamt eher unzufrieden mit dem Datenzugang in Deutschland. Sie forschen daher oft mit ausländischen Daten. Deshalb leide, nach Meinung der Mitglieder, auch die aktuelle Politikberatung in Deutschland unter fehlenden sowie unzureichenden Forschungsergebnissen mit deutschen Daten. Zwar gab es in den vergangenen zwei Jahrzehnten Verbesserungen beim Forschungsdatenzugang, die heutigen Datenbedarfe gehen jedoch deutlich über die existierende Infrastruktur und den gesetzlichen Rahmen hinaus: Die Gesetzeslage selbst muss angepasst werden, um leistungsstarke Forschung zu ermöglichen.

Forschungsdateninfrastruktur im Bereich Makroökonomik – der Status quo

Unzufriedenheit mit der Dateninfrastruktur

Wie die Mitglieder des VfS insgesamt bemängeln auch die Makroökonom:innen die unzureichende Dateninfrastruktur. Beispiele finden sich in der COVID-19-Krise: Während in den USA den staatlichen Stellen die finanzielle Situation US-amerikanischer Firmen und Haushalte sowie die aktuelle Lage am Arbeitsmarkt und des Konsums schnell bekannt war, wurden die deutschen wirtschaftspolitischen Hilfspakete im Frühjahr und Sommer 2020 weitgehend im Blindflug aufgelegt. Eine bessere Dateninfrastruktur hätte es ermöglicht, die Lage der Haushalte und Unternehmen vor der Krise und ihre Betroffenheit durch die Krise zu analysieren sowie die Wirkung der einzelnen Maßnahmen nach ihrer Implementierung zu evaluieren.

Guter Datenzugang ist für Forschung und Politikberatung unverzichtbar

In der Gruppe der Makroökonom:innen wird mithin problematisiert, dass fehlende Forschungsergebnisse für Deutschland die evidenzbasierte Politikevaluation und -beratung beeinträchtigen und zu unnötiger Unsicherheit für die Wirtschaftspolitik führen. Aus Unsicherheit können unzureichend fundierte Entscheidungen resultieren, die mit einer mangelhaften Zielgenauigkeit von Politikmaßnahmen und somit mit überhöhten fiskalischen Kosten einhergehen. Darüber hinaus ermöglichen fehlende und unzureichende Daten es einigen Akteuren, wie z. B. Lobbyorganisationen, die Politik in einer Situation der Unsicherheit leichter zu beeinflussen. Die Verbesserung der Dateninfrastruktur wäre also eine Investition in besser fundierte, zielgenauere und kosteneffizientere Politikentscheidungen.

Unsicherheit aufgrund fehlender Daten und Forschung ist auch für den privaten Sektor ein Problem. Dort kann mangelnde Datenverfügbarkeit und die damit einhergehende Unsicherheit zu verzögerten und ineffizienten Investi-

Prof. Dr. Rüdiger Bachmann ist Professor an der University of Notre Dame in den USA.

Prof. Dr. Almut Balleer ist Professorin an der TU Dortmund und am RWI Essen.

Prof. Dr. Zeno Enders ist Professor an der Ruprecht-Karls-Universität Heidelberg.

Dr. Niklas Garnadt ist Generalsekretär des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung.

Prof. Dr. Heiko Stüber ist Professor an der Hochschule der Bundesagentur für Arbeit in Schwerin.

© Der/die Autor:in 2023. 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.

* Die vorliegende Stellungnahme wurde von der Unterarbeitsgruppe (UAG) „Makroökonomik“ des VfS verfasst und bildet die Position der Mitglieder der UAG sowie der 93 von 500 Teilnehmenden ab, die als Arbeitsgebiet Makroökonomik oder Geldtheorie/-politik angegeben haben (nachfolgend als „Gruppe der Makroökonom:innen“ bezeichnet). Eine längere Fassung dieses Textes findet sich als „Für einen besseren Datenzugang: Stellungnahme der Arbeitsgruppe Makroökonomische Daten“ unter www.socialpolitik.de/de/vfs_stellungnahmen_daten_makro.

tions- und Arbeitsmarktentscheidungen führen (Attentismus). Der gesamtwirtschaftliche Wohlstand wird durch solches Verhalten geschmälert.

Defizite der deutschen Forschungsdateninfrastruktur

Wie die Mitglieder des VfS insgesamt stimmen auch die Makroökonom:innen darin überein, dass fehlende Datenverknüpfungen einer der größten Mängel in der deutschen Dateninfrastruktur sind. Das betrifft sowohl die Verknüpfung von Daten verschiedener öffentlicher Datenproduzenten untereinander (z. B. Statistisches Bundesamt (Destatis), Bundesagentur für Arbeit (BA), Rentenversicherung, Bundesbank, Kreditanstalt für Wiederaufbau (KfW), Steuerbehörden, Gutachterausschüsse für Grundstückswerte) als auch die Verknüpfung dieser Daten mit denen, die durch staatlich geförderte Stellen (z. B. ifo, DIW, ZEW) sowie privat (z. B. GfK, Hoppenstedt, BvD) erhoben werden. Darüber hinaus wird eine fehlende Harmonisierung und Zentralisierung im Datenzugang für die Daten der statistischen Landesämter beklagt sowie eine mangelhafte Koordination der statistischen Landesämter untereinander und mit Destatis.

Die Gruppe der Makroökonom:innen beklagt im Speziellen, dass wichtige Daten, wie z. B. gesamtwirtschaftliche Zeitreihen (bei Destatis), nicht benutzerfreundlich vorliegen, es organisatorische Hürden beim Zugang gibt, z. B. keinen Fernzugriff, und relevante Informationen schwer zu finden sind. Öffentlich verfügbare, tief aggregierte Daten werden nicht oder nur unzureichend bereitgestellt. Darüber hinaus fehlt der Zugang zu Daten mancher öffentlicher Institutionen vollständig, etwa zu denen der KfW. Andere Daten werden Forschenden nur mit großer zeitlicher Verzögerung bereitgestellt. Ein Beispiel sind Informationen zur Nutzung von Kurzarbeitergeld in der COVID-19-Krise, die aber zur Beurteilung der Lage der Unternehmen in Krisensituationen oder zur Evaluierung unternehmensgerichteter wirtschaftspolitischer Maßnahmen wichtig wären.

Die Hindernisse beim Zugang und im Umgang mit Daten lassen sich in drei wesentliche Gruppen gliedern:

- *Rechtliche Hindernisse* wie unterschiedliche Datenschutzgesetze und -auslegungen in den Bundesländern, fehlendes Forschungsprivileg (d. h. Forschung darf mit Daten machen, was nicht explizit verboten ist), rechtliche Beschränkungen der Verknüpfbarkeit von Daten, insbesondere über Datensilos hinweg.

- *Technische Hindernisse* wie Datenbereitsteller ermöglichen keinen Fernzugriff (remote), mangelnde technische Möglichkeiten zur Datenverknüpfung wegen fehlender Identifikationen, mangelnde Maschinenlesbarkeit.

- *Organisatorische Hindernisse* wie oft keine einheitlichen Datendefinitionen (Harmonisierung), Daten liegen – vor allem in Krisenzeiten – oft zu spät vor, zum Teil ist der Zugang zu Daten öffentlicher Institutionen zu teuer und liegt, insbesondere bei Standarddatensätzen, deutlich über den Grenzkosten der Bereitstellung, was dem Charakter dieser Daten als öffentliches Gut widerspricht.

Empfehlungen zur Verbesserung der deutschen Forschungsdateninfrastruktur

Bereitstellung von Daten

Für eine Reform hin zu einer benutzerfreundlicheren Bereitstellung von frei verfügbaren Zeitreihen kann man sich an den öffentlich zugänglichen Datenbanken der OECD oder der Federal Reserve Bank of St. Louis (FRED) orientieren. Diese bieten, im Vergleich zu deutschen Datenanbietern, einfache Such- und Downloadfunktionen an, insbesondere was das Zusammenstellen von eigenen Datensätzen aus verschiedenen Zeitreihen und Quellen in einem einheitlichen Format angeht (One-stop-shopping-Prinzip). FRED bietet zudem einfache Analysewerkzeuge für einen ersten Blick auf die Daten an. Viele Forschende nutzen daher selbst für die von Destatis erhobenen Daten der amtlichen Statistik FRED anstatt GENESIS-Online (das Portal von Destatis) als Zugangsweg.

Für den benutzerfreundlichen Zugang zu Mikrodaten bietet sich zum einen ein Blick nach Frankreich an. Hier ermöglichen die französischen Behörden, insbesondere das Statistikamt INSEE, die CASD Technologie zum Fernzugriff auf vertrauliche Mikrodaten in einer sicheren Umgebung (siehe www.casd.eu/en). Zum anderen können die skandinavischen Staaten als technisches Vorbild dienen, die ebenfalls seit längerem erprobte Remote-Zugriffswege bereitstellen.

Darüber hinaus wäre die Bereitstellung von öffentlich verfügbaren, tief aggregierten Daten hilfreich, z. B. Informationen zu Haushalten und Firmen für Industrien und Regionen, wie in den Business Dynamics Statistics oder den Longitudinal Employer-Household Dynamics des U.S. Census Bureau. Hinzu kommen regionale Daten zu Preisindizes sowie zu Immobilienpreisen und -transaktionen. Dabei ist auf einheitliche Datendefinitionen und Datenformate über die Gebietskörperschaften hinweg zu achten. Eine umfassendere Maschinenlesbarkeit amtlicher Statistiken wäre wünschenswert.

Um die Übersicht über die deutsche Forschungsdatenlandschaft zu verbessern, wäre eine öffentlich finanzierte und zugängliche Plattform als Forschungsdatenregister zielführend. Auf dieser Plattform sollten alle in Deutsch-

land verfügbaren Forschungsdaten kartografiert werden. Dazu sollten insbesondere Metadatenreports, Datenstrukturfiles sowie Informationen über den Datenzugang bereitgestellt werden. Der Träger dieser Plattform könnte zudem Wissenschaftler:innen hinsichtlich ihrer Befähigung zum Umgang mit schwach anonymisierten Forschungsdaten zertifizieren, um den Datenzugang zu beschleunigen. Der RatSWD könnte als Träger einer solchen Plattform agieren.

Der Datenzugang sollte dabei grundsätzlich für zertifizierte Wissenschaftler:innen ermöglicht werden. Dabei sollte die Datenfernverarbeitung als Standardzugangsweg zu formal anonymisierten Daten etabliert werden. Dafür sind Anpassungen der Statistikgesetzgebung vorzunehmen. So sollte ein Fernzugriff beispielsweise in § 16 Abs. 6 BStatG für den Zugang zu amtlichen Daten und in § 75 SGB X für den Zugang zu Daten der BA verankert werden. Die Preisgestaltung des Datenzugangs sollte sich an den Grenzkosten der Datenbereitstellung orientieren und nicht, wie derzeit oft üblich, an den Durchschnittskosten. Finanzielle Zugangshürden können so abgebaut werden.

Alle gesetzlichen Regelungen und organisatorischen Arrangements, die zur Löschung von Daten führen, sind abzuschaffen. Lange Zeitreihen und Verknüpfbarkeit über lange Zeiträume hinweg sind nicht nur für Makrodaten, sondern gerade auch für Mikrodaten essenziell für die Forschung.

Verknüpfung von Daten

Die Verknüpfung von Daten über Datensilos hinweg ist ein zentrales Anliegen der Forschung und Politikberatung. Dazu müssen die rechtlichen Voraussetzungen für die Verknüpfung von Registern und Datensätzen verschiedener Datenproduzenten untereinander geschaffen werden. Dafür müsste insbesondere § 13a BStatG um weitere öffentliche Datenproduzenten neben der Bundesbank erweitert werden, um eine Verknüpfung der Daten mit der amtlichen Statistik zu ermöglichen. Für die Umsetzung muss es einen eindeutigen Identifikator auf Personen-, Haushalts-, Betriebs- und Unternehmensebene geben, wie dies in vielen europäischen, vor allem den skandinavischen Ländern der Fall ist.

In der Praxis kann die Verknüpfung von Daten aus verschiedenen Quellen durch Datenmittlerstrukturen ermöglicht werden. Idealerweise erfolgt dies durch eine Erweiterung des gesetzlichen Auftrags des Statistischen Bundesamtes beispielsweise durch die Ansiedlung eines mit entsprechenden personellen und finanziellen Ressourcen unterlegten Dateninstituts am Destatis als Datenmittler und -verknüpfer. Als Vorbild könnte das jüngst errichtete Mikrodatenzentrum in Österreich dienen. Dabei sollten mindestens die Daten der öffentlichen Verwaltung, Daten der amtlichen Statistik und von Forschungseinrichtungen für die Wissenschaft aufbereitet, zugänglich und verknüpfbar sein. Idealerweise integriert man aber auch Daten aus der Wirtschaft und von Unternehmen.

Verbesserung der rechtlichen Rahmenbedingungen

Die Entwicklung der Statistikgesetzgebung hin zu einem outputorientierten Gesetzgebungsrahmen sollte fortgesetzt werden. Dies bedeutet eine Fokussierung auf Informationsbedarfe und Dienstleistungen statt auf Datenquellen. Bislang müssen gemäß § 9 BStatG für amtliche Statistiken Rechtsvorschriften – oft fachstatistische Einzelgesetze – beschlossen werden, die detailliert definieren, welche Merkmale von welchen Subjekten erhoben werden dürfen. Das bedeutet allerdings, dass zur Anpassung von Statistiken und Erhebungen sowie für neue Nutzungswege oftmals gesetzliche Anpassungen notwendig sind. Dies reduziert die Agilität der amtlichen Statistik erheblich. Eine outputorientierte Anpassung der Statistikgesetzgebung, insbesondere eine Anpassung von § 9 BStatG, könnte es stattdessen der amtlichen Statistik ermöglichen, stärker im eigenen Ermessen die allgemein definierten Aufgaben gemäß BStatG zu erfüllen, ohne auf fachstatistische Einzelgesetze zurückgreifen zu müssen. Insbesondere sollte der amtlichen Statistik die Freiheit gegeben werden, die notwendigen Erhebungen inklusive Merkmalen und Stichproben selbst zu definieren, die zur Erfüllung eines per Gesetz definierten Informationsbedarfs notwendig sind.

Aus Sicht der Forschung wäre die gesetzliche Verankerung eines Forschungsauftrags für Destatis im BStatG zielführend, um das Dienstleistungsangebot für die Wissenschaft zu verbessern. Die Übertragung eines Forschungsauftrags an Destatis könnte sich dabei am U.S. Census Bureau, der Bundesagentur für Arbeit und der Bundesbank orientieren. Diese Institutionen verfügen mit dem Center for Economic Studies, dem Institut für Arbeitsmarkt- und Berufsforschung (IAB) sowie dem Forschungszentrum der Bundesbank über eigene Forschungszentren. Ein solcher Forschungsauftrag, der die Gründung eines Forschungszentrums einschlösse, hätte mindestens drei Vorteile. *Erstens* würden Forscher:innen direkt bei Destatis mit den Forschungsdaten arbeiten und könnten so unmittelbar an Lösungen für die unweigerlich in der Praxis auftretenden Datenprobleme arbeiten. *Zweitens* würde bei einer Evaluation des Forschungsauftrags das Eigeninteresse von Destatis an der Aufbereitung, Bereitstellung und Nutzung von Forschungsdaten steigen. *Drittens* würde qualitativ hochwertige wissenschaftliche Forschung mit deutschen Daten direkt bei Destatis stattfinden. Die oben genannten Forschungszentren waren in den vergangenen Jahren allesamt an Publikationen in den

besten internationalen Fachzeitschriften beteiligt. Zudem haben das IAB und die von ihm administrierte Dateninfrastruktur maßgeblich zum besseren Verständnis des deutschen Arbeitsmarkts beigetragen.

Darüber hinaus wäre eine gesetzliche Verankerung der existierenden Forschungsdatenzentren (FDZ) der Statistischen Ämter der Länder und des Bundes im BStatG als Kernaufgabe des Statistischen Verbunds wichtig. Bisher werden die FDZ im BStatG nicht explizit erwähnt und die Bereitstellung der Daten an externe Forschende wird lediglich indirekt über das BStatG geregelt. Mit einer gesetzlichen Verankerung würde die Bedeutung insbesondere des FDZ bei Destatis hervorgehoben und explizit darauf hingewiesen, dass für die Datenbereitstellung eine entsprechende Einrichtung (FDZ) und entsprechende personelle und finanzielle Ressourcen notwendig sind. Damit würde *erstens* der langfristige Betrieb der FDZ gesichert. *Zweitens* würde damit das aktuell bestehende Problem der Selbstfinanzierung der FDZ und die damit vergleichsweise hohen Kosten für die Datennutzer:innen behoben. Eine öffentliche Finanzierung wäre insofern gerechtfertigt, als dass die Forschungsergebnisse der gesamten Gesellschaft zugutekommen.

Im Bereich der Wirtschaftsstatistiken wird die Verfügbarkeit in der Vergangenheit erhobener Daten durch die Löschvorschrift bei Wirtschaftsstatistiken (§ 13a BStatG) erschwert. Diese sieht vor, dass Identifikatoren (und möglicherweise mithilfe dieser Identifikatoren generierte Datensätze) nach 30 Jahren gelöscht werden müssen. Dies ist für die empirische Wirtschaftsforschung und die Politikberatung problematisch, da oft langlaufende Zeitreihen für die Identifikation von Wirkungsmechanismen erforderlich sind. Hier sollte zumindest ein Wissenschaftsprivileg verankert werden, sodass die Löschvorschrift nicht für Daten gilt, die der Wissenschaft für Forschungszwecke zur Verfügung gestellt werden.[1]

Ohnehin sollte ein Wissenschafts- oder Forschungsprivileg als zentraler Bestandteil des geplanten Forschungsdatengesetzes verankert werden. Dieses sollte bestehende Einzelregelungen zum Datenzugang, -verwendung und -schutz überschreiben und bundeseinheitlich regeln. Insbesondere sollte das Forschungsprivileg der wissenschaftlichen Forschung gegenüber dem Datenschutz ein hohes Gewicht bei der Datennutzung sicherstellen, da wissenschaftliche Forschung maßgeblich zum Gemeinwohl beiträgt. Zudem sollten großzügige Verknüpfungsregeln für Daten erreicht werden. Ein Forschungsdatengesetz würde auch zur Rechtssicherheit beitragen, da zurzeit oft unterschiedliche Auffassungen der beteiligten Stellen bezüglich der Auslegung des bestehenden Rechts Datenverknüpfungen verhindern.

Erhebung neuer Daten und Verfügbarkeit bereits erhobener Daten

Der kostengünstigste Weg, um neue Daten zur Verfügung zu stellen, besteht in der Aufbereitung bereits erhobener Daten. Dies betrifft vor allem Daten aus Verwaltungsprozessen, sogenannte administrative Daten. Aus datenmethodologischer Sicht sind administrative Daten Umfragedaten zudem oft wegen geringerer Messfehler und höherer Rücklaufquoten qualitativ überlegen. Nichtsdestotrotz sind neben administrativen Daten hochqualitative Umfragedaten essenziell, um ökonomische Entscheidungen bestmöglich zu verstehen. Hierzu gehören z. B. Daten zu Erwartungen und Einschätzungen in Bezug auf ökonomische Größen und Zusammenhänge. Dafür braucht es einen Ausbau von Umfragedaten von offizieller Stelle, etwa nach dem Vorbild des U.S. Census Pulse Survey oder des Survey of Consumer Expectations der Federal Reserve Bank of New York, und eine intelligente Verknüpfung dieser mit den administrativen Daten.

Aus makroökonomischer und wirtschaftspolitischer Sicht besteht mithin der Bedarf einer hochfrequenten Haushaltsumfrage, die zeitnah die ökonomische Lage (Arbeitsmarkt, Konsum, Erwartungen, Einkommen, Vermögen, Verschuldung etc.) der deutschen Haushalte abbildet. Eine solche Umfrage könnte in Anlehnung an den U.S. Current Population Survey (CPS) geschaffen werden.[2] Der Rechtsrahmen sollte mit der Möglichkeit gestaltet werden, kurzfristig und anlassbezogen neue Umfragemodule einzuführen.

Darüber hinaus fehlen amtliche Daten zur Vermögensverteilung in Deutschland. Diese Datenlücke könnte vergleichsweise kostengünstig durch eine Verknüpfung des Mikrozensus mit administrativen Daten und die Entwicklung geeigneter Schätzverfahren zur Abschätzung des Vermögens anhand von administrativen Daten adressiert werden. Dazu sollte schnellstmöglich darauf hingewirkt werden, einen datensatzübergreifenden Identifikator zu finden, etwa die Steuer-ID. Eine Auskunftspflicht über Unternehmensbeteiligungen und eine Nennung der Handelsregisternummer dürfte für eine Verknüpfung mit Unternehmensbilanzdaten geeignet sein.

1 Der diesjährige Nobelpreis an Wirtschaftshistorikerin Claudia Goldin für die Erforschung des Arbeitsmarktes von Frauen zeigt, wie wichtig lang zurückreichende Daten sind.

2 Die Einkommens- und Verbrauchsstichprobe (EVS) wird nur alle fünf Jahre durchgeführt und auch ein auf Quartalsfrequenz umgestelltes Sozio-oekonomische Panel (SOEP) kann in Krisenzeiten zu langsam sein, zumal es auch hier oft deutlichen Meldeverzug gibt.

Bei neuen Gesetzgebungsvorhaben sollte künftig für Vorhaben ab einem zu definierenden Volumen eine Pflicht zur begleitenden Datenerhebung, Datenbereitstellung und zeitnahen wissenschaftlichen Evaluation verankert werden. Zudem sollten Forschungsklauseln, also forschungsfreundliche, barrierefreie Zugangsregeln für die unabhängige wissenschaftliche Forschung geschaffen werden. Idealerweise sollte bei Maßnahmen, die Unternehmen oder Haushalte betreffen, der Identifikator der Betroffenen ermittelt werden. So ließen sich die Daten mit Informationen aus der amtlichen Statistik verknüpfen.

Nicht zuletzt wäre eine Aufbereitung historischer makroökonomischer Daten wertvoll. So sind manche Daten zu den Volkswirtschaftlichen Gesamtrechnungen erst ab dem Jahr 1991 in zugänglicher Form verfügbar, obwohl ihre Erhebung durch Destatis viel früher angefangen hat. Eine Bereitstellung der Daten ab den 1950er Jahren, wie es in den USA der Fall ist, wäre für ein verbessertes Verständnis langfristiger makroökonomischer Zusammenhänge wichtig. Zudem wären beispielsweise für die makroökonomische Investitions- und Kapitalstockforschung die langlaufenden Investitions- und Kapitalstockreihen hilfreich, die teils bis in das 19. oder sogar 18. Jahrhundert zurückgehen. Diese Daten könnten von Destatis zur Verfügung gestellt werden, beispielsweise im Rahmen des Programms „Experimentelle Statistiken“.

Title: *Macro Data: Strengthening Official Statistics*

Abstract: *Despite some progress, German macroeconomists are generally dissatisfied with access to important data in Germany. Legal, technical and organisational obstacles abound, in particular as far as mergers of data sets from heterogeneous sources are concerned. This article suggests a number of concrete measures, including legal changes, to improve the situation. Easy access to a variety of data sources is a necessary condition for the design, evaluation, and thus, for the improvement of macroeconomic policies. Better informed decisions are also less prone to be influenced by lobbyists.*