

Verwaltungsdaten – der Beipackzettel zu einem Wundermittel für die amtliche Statistik

Peter Weigl, Vizepräsident des Statistischen Bundesamtes

Es scheint, als brauche die amtliche Statistik dringend ein Wundermittel, um eine Schwäche zu kurieren, nämlich die von den Befragten als überbordend empfundene Belastung durch Erhebungen. Das Wundermittel ist schnell identifiziert: **Verwaltungsdatennutzung** soll den Patienten auf die Beine bringen oder besser: ihm Beine machen, seine Daten aktueller und für Bürger und Unternehmen quasi geräuschlos zu produzieren.

Ein so stark wirkendes Medikament ist natürlich rezeptpflichtig und es muss einen Beipackzettel enthalten. An diesem Beipackzettel schreiben – unter anderem – jene Kolleginnen und Kollegen im Statistischen Bundesamt und in den Statistischen Ämtern der Länder, die auf der Basis des Verwaltungsdatenverwendungsgesetzes vom Oktober 2003 die Eignung der Verwaltungsdaten für Wirtschaftsstatistiken, besonders für Konjunkturstatistiken, prüfen. Ihren Berichten verdanken wir die Kenntnis der bisher erforschten Wirkungen und Nebenwirkungen der Verwaltungsdatennutzung.

Es ist bemerkenswert, wie die Diskussion um die Nutzung von Verwaltungsdaten die amtliche (und nur die amtliche) Statistik wieder zu ihren Anfängen führt. In diesen Anfängen emanzipierte sich die statistische Erfassung erst ansatzweise von der durch Verwaltungshandeln getriebenen Erfassung. War es nicht Stolz der Statistiker, selbst Daten zu erheben? Ernst Wagemann, ehemals Präsident des statistischen Reichsamtes sowie Gründer und Leiter des Instituts für Konjunkturforschung, sah es 1952 jedenfalls noch so:

„Es bleibt gewiß die polizeiliche Anmeldung, es bleibt die Paßkontrolle, es bleibt ein raffiniertes Formularwesen. Das sind jedoch behördliche Einrichtungen, und der Amtsschimmel behält seine überkommene Gangart bei. Notgedrungen muss die Statistik auf diesem Schimmel reiten.“

(Ernst Wagemann, Die Zahl als Detektiv (1952), S. 15)

Wagemann verweist auf die „überkommene Gangart“ des Amtsschimmels und wir Nachgeborenen müssen zugeben, dass dieser Schimmel, wenn er auch seinen Trab beschleunigt hat, kaum jemals einen Vorreiter tragen kann.

Teil 1: Nebenwirkungen Eignung der Verwaltungsdaten

Aktualität. Hier tritt die erste schwerere Nebenwirkung des Wundermittels „Verwaltungsdatenverwendung“ zutage: Gefahr mangelnder Aktualität. Verwaltungsdaten mögen alle Informationen enthalten, die wir brauchen, aber häufig nicht zu dem Zeitpunkt, zu dem wir sie brauchen. Aus der Sicht des Statistikers ist dieses Manko den Verwaltungsdaten geradezu immanent. Die Untersuchungen im „Labor“ für Verwaltungsdatennutzung haben indessen ergeben, dass dieses Manko nicht so gravierend sein muss, als dass es zum K. O.-Kriterium der Verwaltungsdatennutzung avancieren könnte. Es kommt, wie so oft, auf den Zweck an: Für einige Branchen lassen sich Konjunkturstatistiken nach heutigen Aktualitätsanforderungen auch aus Verwaltungsdaten herstellen, für andere nicht.

Datenschutz. Als nächste Nebenwirkung steht auf unserem Beipackzettel die Gefährdung des Datenschutzes. Verwaltungsdaten dienen in aller Regel konkreten, vorab definierten Verwaltungszwecken, die ein bestimmtes Handeln der Verwaltung auslösen. Kurz: Aus Verwaltungsdaten folgt meist ein konkretes Verwaltungshandeln. Daten und Handeln sind in vorab definierter Weise verknüpft.

Aus Daten für statistische Zwecke folgt demgegenüber kein konkret vorab definiertes Verwaltungshandeln. Es ist gerade das Wesen von Daten für statistische Zwecke, dass sie in der Regel existieren, ohne dass daraus ein bestimmtes Handeln folgen muss, vor allem keines, das sich auf das Individuum bezieht.

Manchmal scheinen statistische Daten aber doch bestimmte, vorab definierte Handlungen auszulösen. Denken Sie etwa an die so genannten „Maastricht-Kriterien“ oder die Formel zur Rentenanpassung. Beide beziehen sich auf die statistischen Daten aus den amtlichen Volkswirtschaftlichen Gesamtrechnungen. Bei näherer Betrachtung zeigt sich jedoch, dass in beiden Fällen immer ein politischer Spielraum besteht, was das konkrete Handeln angeht. Die statistischen Daten haben letztlich nur informativen Charakter und erzwingen kein Handeln. Zumindest lassen sie einen weiten Ermessensspielraum.

Diese Unterscheidung von „Verwaltungszweck“ und „statistischem Zweck“ ist fundamental für das Selbstverständnis der amtlichen Statistik, weil sie den sensiblen Bereich des Datenschutzes berührt und weil sie einer der Garantien für die Neutralität und damit Glaubwürdigkeit der Statistik ist. Daten eines Individuums, die einmal in die Sphäre der amtlichen Statistik gelangt sind, sollen keine Rückwirkungen auf das betreffende Individuum haben. Darauf müssen sich Auskunft gebende verlassen können. Dieser Grundsatz der Zweckbindung von Daten gilt aber auch umgekehrt: Für personenbezogene Daten bestimmt das Bundesdatenschutzgesetz, dass diese Daten ohne Einwilligung der Betroffenen grundsätzlich nicht zu anderen Zwecken verwandt werden dürfen, auch nicht zu statistischen Zwecken. Das ist Ausdruck des Grundrechts auf informationelle Selbstbestimmung: Jeder Mensch soll Herr über die Daten bleiben, die ihn betreffen.

Nun gibt es natürlich auch Daten über juristische Personen und für diese gilt das Bundesdatenschutzgesetz nicht. Dennoch sind auch Daten juristischer Personen – etwa Unternehmen – in vielen Fällen vor weiterer Verwendung geschützt, etwa durch das Steuergeheimnis. Die Statistik darf ohne ausdrückliche gesetzliche Regelung diese Daten nicht verwenden.

Außerdem führt das Sammeln von Daten zu großen Datenmengen und je größer die Menge der Daten, desto gefährdeter sind sie. Das Datenschutzrecht hat deshalb das Prinzip der Datensparsamkeit entwickelt.

Die Verwaltungsdatenverwendung führt die Statistik also auf ein juristisch schwieriges Feld.

Begriffliche Unterschiede zwischen Verwaltung und Statistik

Eine nicht zu unterschätzende Nebenwirkung sind begriffliche Unterschiede. So stimmen die Begriffe „Umsatz“ und „Beschäftigte“ in Verwaltungsdaten nicht mit den international harmonisierten Begriffsabgrenzungen der Konjunkturstatistik überein.

Die Verwaltungsdaten beschränken sich derzeit bei Beschäftigten auf sozialversicherungspflichtig und geringfügig entlohnte Beschäftigte. Zu Selbständigen und mithelfenden Familienangehörigen liegen beispielsweise keine Verwaltungsdaten vor. Bei der Umsatzsteuer werden alle Umsätze erfasst, die auch besteuert werden. Ein Teil dieser Umsätze besteht aus außerordentlichen oder betriebsfremden Erträgen, die bei Konjunkturbetrachtungen nicht einbezogen werden. Nicht immer gelingt es, aus den Steuerdaten den statistischen Umsatz adäquat abzuleiten. Andererseits enthalten die Steuerdaten sicher sorgfältigere Korrekturen als die Primärerhebungen.

Verzerrungen durch Antwortausfälle maßgeblicher Berichtseinheiten

Die Untersuchungen haben gezeigt, dass im Dienstleistungssektor (konkret: Verkehr, Nachrichtenübermittlung und sonstige unternehmensbezogene Dienstleistungen) insbesondere Ausfälle von Meldungen großer Einheiten Probleme bereiten. Dies führt in der Praxis bei der Verwaltungsdatenverwendung in einigen Bereichen zu hohen Revisionen der aktuellen Angaben zwei Monate nach Abschluss des Berichtszeitraums, besonders bei den Umsatzdaten. Im kleinteilig strukturierten Handwerk bereiten die Antwortausfälle dagegen weniger Probleme. Durch ihre große Zahl schlagen Probleme bei einzelnen Einheiten nicht in gleichem Maße auf die konjunkturellen Veränderungsdaten durch und Schätzprobleme mitteln sich bei der Aggregation heraus.

Unzureichende Qualität der Verwaltungsdaten für statistische Zwecke

Für uns Statistiker sind manchmal Punkte zentral, die beim Zustandekommen der Verwaltungsdaten eher nachrangig sind. Ein Beispiel ist die Wirtschaftszweiguordnung der Unternehmen und Betriebe. Zwar verwenden Verwaltung und Statistik einheitliche Klassifikationen, dennoch weichen im Einzelfall die konkreten Zuordnungen von Unternehmen zu Wirtschaftszweigen zum Teil erheblich voneinander ab. Trotz dieser Nebenwirkung können wir das Mittel „Verwaltungsdaten“ den Konjunkturstatistiken verabreichen, weil sie sich auf den für Konjunkturbetrachtungen vorrangig relevanten höher aggregierte Ebenen deutlich schwächer auswirkt als auf den tief gegliederten Ebenen: Die Grundzuordnung zu einer Branche stimmt oft überein, aber je mehr es ins Detail geht, umso stärker weichen die Wirtschaftszweiguordnungen von Verwaltungen und Statistik voneinander ab. Der Statistische Beirat fordert an dieser Stelle schon seit längerem eine bessere Berücksichtigung statistischer Belange bei der Ausgestaltung von Verwaltungsregistern.

Länderübergreifende Antworteinheiten, sogenannte Mehrländerunternehmen, verwischen das länderspezifisch gewünschte Ergebnis

Eine weitere Nebenwirkung, die wir gefunden haben: So genannte Mehrländerunternehmen, verwischen das länderspezifisch gewünschte Ergebnis. Große Einheiten sind oft Mitglieder von Organschaften, oft auch Mehrländerunternehmen, bei denen die Umsatzdaten der Finanzverwaltung nicht in einer Form vorliegen, die den Anforderungen der Unternehmensstatistiken entspricht. Bei Organschaften bilden mehrere Unternehmen zusammen einen Steuerzahler, der eine Meldung abgibt. Für die Unternehmensstatistik müssen daher Angaben für die einzelnen Unternehmen geschätzt werden. Bei den Mehrländerunternehmen enthalten die Steuerdaten keine Aufteilung der Umsätze nach Bundesländern. Für eine länderscharfe Zuordnung von Unternehmen bzw. Unternehmensteilen müssen daher Umsätze mit Hilfe von Schätzverfahren aus den Verwaltungsdaten

ermittelt werden. Dies ist methodisch schwierig und gelingt nur mit begrenzter Aussagefähigkeit. Da große Einheiten gleichzeitig einen hohen Einfluss auf die Konjunkturentwicklung in ihrem Wirtschaftszweig haben, führen diese Probleme bei großen Einheiten zu einem zusätzlichen, hohen Risiko einer nicht adäquaten Darstellung der konjunkturellen Entwicklung. Dies gilt in besonderem Maße für die Konjunkturdarstellung auf Länderebene.

Teil 2: Besser neu konzipieren als ersetzen: „Ersatzfähigkeit“ bemisst sich immer nach dem Einsatzzweck

Ob und wie Verwaltungsdaten als Wundermittel wirken, hängt auch davon ab, ob wir dieses Mittel bloß als Surrogat verstehen für Ingredienzien, die wir bisher durch unsere Primärerhebungen gewonnen haben. Oder ob wir in Verwaltungsdaten eine genuin neue Zutat erblicken, die nach einem neuen, adäquaten Rezept verlangt. So wie im alten China der Tofu nicht als Bratwürstchenersatz erfunden wurde, sind auch Verwaltungsdaten Daten sui generis, die nicht geschaffen wurden, um sie anstelle der Daten aus Primärerhebungen auf den Grill der Statistik zu legen.

Unser Labor ist also auch ein Kochstudio. Wir lernen darin, mit Verwaltungsdaten zu kochen. Die erwähnten Schwächen der Verwaltungsdaten für konjunkturstatistische Zwecke bedeuten für uns nicht, dass wir diese Zutat vollkommen ablehnen. Gerades das Potential der Verwaltungsdaten zur Entlastung der Befragten nehmen wir sehr ernst. Deshalb untersuchen wir nicht nur die Verwaltungsdaten per se mit Blick auf ihre Qualität. Wir haben in den letzten Jahren in erheblichem Umfang Verknüpfungsmodele und Schätzmodelle entwickelt, mit denen wir die Schwächen der Verwaltungsdaten für statistische Zwecke zumindest zum Teil kompensieren. Eine wichtige Rolle spielt dabei das Unternehmensregister, das wir bei der Auswertung der Verwaltungsdaten als zentrale ergänzende Informationsquelle heranziehen.

Die entscheidende Frage ist dabei, in welchem Maße es uns gelingt, die Nebenwirkungen der Verwaltungsdaten so zu beherrschen, dass unser Endprodukt – die Statistik – nach wie vor unseren hohen Qualitätsmaßstäben gewachsen ist. Es hat sich gezeigt, dass das manchmal allein mit Schätzverfahren möglich ist, aber nicht immer. Eine pauschale Bewertung, ob sich Verwaltungsdaten für Konjunkturstatistiken eignen oder nicht, gibt es deshalb nicht. Vieles hängt von den Anforderungen der Datennutzer ab, vom Ausmaß der Nebenwirkungen der Verwaltungsdaten in einem bestimmten Bereich und von den spezifischen Möglichkeiten der Linderung oder Beseitigung dieser Nebenwirkungen.

Für die Konjunkturstatistiken über den Einzelhandel wirkt unser Wundermittel Verwaltungsdaten zum Beispiel ganz anders, als es sich viele vielleicht gedacht hatten. Als Ersatzstoff für Primärdaten völlig unwirksam, können die Verwaltungsdaten statt dessen dazu dienen, das statistische Unternehmensregister zu pflegen, um die in diesen Branchen notorischen Strukturveränderungen durch entsprechende Anpassung der Stichprobe abzufedern.

Sinnvoller ist es deshalb, an einigen Stellen neue Statistiken zu konzipieren, die zumindest in wesentlichen Teilen auf Verwaltungsdaten basieren. Das kann auch parallel zum Ersatz von Verwaltungsdaten innerhalb bestehender Erhebungen geschehen. Für die Konjunkturstatistiken im Dienstleistungsbereich, im Ausbaugewerbe und Teile des Handels und Gastgewerbes haben wir ergänzende Primärerhebungen entwickelt (bzw. sind gerade dabei), um angesichts der Schwächen der Verwaltungsdaten mit Mix-Modellen aus Verwaltungsdaten und Primärerhebung dennoch unseren Qualitätsmaßstäben Rechnung zu tragen.

Ein anderes Paradebeispiel für dieses Vorgehen ist – im Bereich der Bevölkerungsstatistik – die Konzeption des neuen Zensus 2011. Ich sage offen, dass viele der hier Anwesenden als Forschende schmerzlich die eine oder andere Information vermissen werden, die sie sich aus dieser ersten Volkszählung seit 1987 erhofften. Gerade diese freiwillige oder erzwungene

Selbstbeschränkung auf das, was mit Verwaltungsdaten geleistet werden kann, macht es möglich, das Potenzial der Verwaltungsdaten zu erschließen.

Ganz kurz gesagt: Direkt fragen kann jeder. Fordern, dass man „einfach Verwaltungsdaten auswerten“ möge, kann auch jeder. Aber die in der Regel fest begrenzten Möglichkeiten der Verwaltungsdaten in Einklang zu bringen mit den in der Regel variablen Informationsbedürfnissen, das ist die Kunst des Ausgleichs, den Statistiker und Nutzer nur dann finden, wenn sie Kompromisse schließen. Um diese Kompromissbereitschaft werde ich hier ausdrücklich, denn aus meiner Sicht und aus Sicht der statistischen Ämter überhaupt lohnt es sich. Das zeigt der Blick auf die Wirkungen der Verwaltungsdatennutzung in unserem Beipackzettel.

Ist Verwaltungsdatenverwendung mit den Worten von Ernst Wagemann wirklich „notgedrungenes Reiten auf dem Amtsschimmel“? Es ist Zeit, endlich die positiven Wirkungen dieses Mittels zu erwähnen. Da ist zuvorderst die **Entlastung der Befragten** zu nennen. Der Paradigmenwechsel von der Primärerhebung zur Nutzung von Verwaltungsdaten, praktisch am ehesten zur verwirklichen in einer Kombination mehrerer Erhebungsmodi und mehrerer Erhebungskanäle, verringert stets signifikant die Menge der Respondenten, die direkt Daten bereitstellen müssen.

Die statistischen Ämter setzen das Medikament „Verwaltungsdatenverwendung“ seit Anfang des Jahres in der vierteljährlichen Handwerksstatistik ein. Sie haben damit einen großen Behandlungserfolg erzielt, nämlich nichts weniger als den vollständigen Ersatz der bisherigen Primärerhebung, zu der rund 41 000 Handwerksunternehmen berichten mussten. In Zahlen ausgedrückt, entlastet das die Auskunftgebenden jährlich um rund 3,3 Millionen Euro.

Und noch einer zweiten Statistik haben wir die Verwaltungsdaten als Therapeutikum gegen eine hohe Belastung der Auskunftgebenden verordnet: Die Konjunkturstatistik in bestimmten Dienstleistungsbereichen wird seit dem

zweiten Quartal 2007 nicht mehr aus den Daten einer dezentralen Stichprobenerhebung erstellt, sondern aus Verwaltungsdaten und Daten aus einer Primärerhebung, die wir im Rahmen eines Mix-Modells zusammenführen. Die Zahl der berichtspflichtigen Unternehmen hat sich dadurch von etwa 40 000 auf 4 000 reduziert.

Der Einsatz von Verwaltungsdaten trägt also wesentlich bei zur wirtschaftlichen Datenproduktion, denn Wirtschaftlichkeit bemisst sich natürlich auch nach dem Aufwand für die Befragten. Gerade in Unternehmen ist gesparte Zeit bares Geld wert. Die Untersuchungen im Rahmen der Bürokratiekostenmessung zeigen das sehr deutlich – jene eben erwähnten 3,3 Millionen Euro Ersparnis sind Ergebnis einer fundierten Berechnung, die vom tatsächlichen Arbeitsaufwand für das Ausfüllen der Fragebogen, dem entsprechenden Lohnsatz und eventuell zusätzlich anfallenden Sachkosten ausgeht.

Verbesserung der Qualität von Statistiken durch Abgleich mit Verwaltungsdaten

Daneben eröffnet die Verwendung der unterjährigen Verwaltungsdaten in einigen Bereichen neue, ergänzende Perspektiven. Ich möchte an dieser Stelle nur auf die Möglichkeiten einer Verbesserung des Unternehmensregisters durch aktuellere Daten oder auf das Potenzial hinweisen, für die vierteljährlichen Volkswirtschaftlichen Gesamtrechnungen für einige Wirtschaftszweige, die bisher von Konjunkturstatistik nicht abgedeckt sind, zusätzliche Basisdaten zur Verfügung zu stellen.

So gesehen mutiert der Amtsschimmel für die Statistik zwar nicht zum Pegasus, er könnte uns aber in zuverlässigem Trab weit tragen. Wer wollte bestreiten, dass schlecht geritten immer noch besser ist, als gut gelaufen?

Synthese, Brücke zur Vergangenheit (Prof. Strecker), Abschluss

Neue Erhebungsformen, die Mix-Modelle und die verstärkte Verwaltungsdatennutzung können uns gedanklich noch einmal zurückführen in die Zeit der Erneuerungen und wissenschaftlichen Methoden nach dem 2. Weltkrieg, jener Epoche des Aufschwungs der Stichprobenerhebungen, deren Apologet der schon zitierte Ernst Wagemann war. Damals galt die Vollerhebung als Maß aller Dinge und Stichprobenerhebungen stellten die Produzenten und Nutzer vor neue Herausforderungen. Prof. Heinrich Strecker, langjähriges Mitglied unserer Gesellschaft (er wurde vorgestern – am 13. September – 86 Jahre alt), hat ihre Anwendung in den fünfziger Jahren in der Agrarstatistik exemplarisch erprobt und beschrieben.

Das ist so lange her, dass nur die wenigen Menschen im gesegneten Alter – wie der verehrte Herr Prof. Strecker – es bezeugen können. Die Dokumente des Diskurses von damals zeigen aber, dass seinerzeit etwas Neues gewagt wurde, und zwar gegen erhebliche Widerstände. Das Neue hat sich dann – glücklicherweise – durchgesetzt und scheint uns heute selbstverständlich.

Vielleicht werden jene von uns, die das hohe Alter eines Prof. Strecker erreichen, selbst als Zeugen von der Zeit berichten können, in der man sich über das Für und Wider der Verwaltungsdatenverwendung gestritten hat. Wenn sich der heute erkennbare Trend durchsetzt, wird das Ausschöpfen vorhandener Datenquellen dann gängige Praxis sein.

Meine sehr geehrten Damen und Herren, allen Teilnehmerinnen und Teilnehmern dieser Veranstaltung wünsche ich interessante Erkenntnisse und der gesamten Statistischen Woche einen guten Verlauf! Vielen Dank!