IfW KIEL INSTITUT FÜR WELTWIRTSCHAFT

KIELER

BEITRÄGE ZUR WIRTSCHAFTSPOLITIK

Big Data in der makroökonomischen Analyse

Nr. 32 Februar 2021

Martin Ademmer, Joscha Beckmann, Eckhardt Bode, Jens Boysen-Hogrefe, Manuel Funke, Philipp Hauber, Tobias Heidland, Julian Hinz, Nils Jannsen, Stefan Kooths, Mareike Söder, Vincent Stamer und Ulrich Stolzenburg

Leibniz Gemeinschaft

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über
http://dnb.d-nb.de
abrufbar.

ISBN 978-3-89456-347-9
ISSN 2567–6474

Herausgeber:
Institut für Weltwirtschaft (IfW)
Kiellinie 66, D–24105 Kiel
T +49 431 8814–1
F +49 431 8814–500

Schriftleitung:
Dr. Klaus Schrader

Redaktion:
Kerstin Stark
Marlies Thiessen
Korinna Werner-Schwarz

Das Institut für Weltwirtschaft ist eine rechtlich selbständige Stiftung des öffentlichen Rechts des Landes Schleswig-Holstein

Umsatzsteuer ID:
DE 251899169

Das Institut wird vertreten durch:
Prof. Gabriel Felbermayr, Ph.D. (Präsident)

Zuständige Aufsichtsbehörde:
Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein

Bilder/Fotos:
Cover: © sarawuth - Fotolia

https://www.ifw-kiel.de/de/

Überblick

- Unter dem Schlagwort Big Data werden neue und in Abgrenzung zur üblichen Wirtschaftsstatistik unkonventionelle Datenquellen zusammengefasst. Sie sind sehr umfangreich und sind sehr zeitnah sowie in hoher Frequenz verfügbar. Allerdings weisen diese neuen Daten eine hohe Bandbreite und Komplexität auf, weil sie nicht für die Analyse von ökonomischen Fragestellungen erhoben werden, sondern vielmehr als Nebenprodukt unterschiedlicher Anwendungen anfallen.
- Viele der Datenquellen erfassen tatsächlich getätigte Transaktionen bzw. erhobene Preise (elektronischer Zahlungsverkehr, Scannerdaten, Handelsplattformen) oder bilden die laufende wirtschaftliche Aktivität in anderer Form unmittelbar ab (Satellitenbilder, Verkehrsdaten). Die daraus resultierenden Daten haben ein besonders großes Potenzial für die laufende Konjunkturbeobachtung, insbesondere für die Prognose am aktuellen Rand.
- Andere Datenquellen bilden Stimmungen und Interessen der Nutzer oder die Relevanz bestimmter Themen ab (soziale Medien, Internetsuchanfragen, Presseartikel). Sie weisen einen weniger engen Bezug zur laufenden wirtschaftlichen Entwicklung auf. Aus ihnen können aber Indikatoren abgeleitet werden, die umfassendere Informationen liefern als bereits vorliegende Indikatoren (z.B. für das Konsumentenvertrauen) oder die über andere konventionelle Datenquellen nicht abgebildet werden können (z.B. wirtschaftspolitische Unsicherheit).
- Der erforderliche Aufwand, um die Daten für makroökonomische Analysen nutzbar zu machen, unterscheidet sich je nach Datenquelle. Aus einigen der Datenquellen werden bereits regelmäßig makroökonomisch relevante Indikatoren gebildet (wirtschaftspolitische Unsicherheit, Witterungsbedingungen, Lkw-Fahrleistung), oder die Daten werden so bereitgestellt, dass sie unmittelbar in empirische Analysen eingebunden werden können (Internetsuchanfragen). Sofern die Daten selbst aufbereitet und ausgewertet werden, ist damit oft ein sehr hoher organisatorischer und methodischer Aufwand verbunden. Viele der Daten sind nicht frei oder nur beschränkt verfügbar.
- Die konkreten Potenziale von Big Data für die makroökonomische Analyse sind in solchen Bereichen besonders groß, für die die konventionelle Wirtschaftsstatistik erst mit Verzögerung Indikatoren bereitstellt oder diese nur wenig zuverlässig sind. Für Deutschland, wie für viele andere Volkswirtschaften, gehören beispielsweise die privaten Konsumausgaben oder die Wertschöpfung in den Dienstleistungsbranchen zu diesen Bereichen. Freilich können Big Data aber auch komplementär zu den bereits vorliegenden Indikatoren wertvolle Beiträge liefern, da sie grundsätzlich zeitnäher und in einem höheren Detailgrad verfügbar sind.
- Die Potenziale und Grenzen von Big Data für die makroökonomische Analyse hängen auch von den individuellen Anforderungen der jeweiligen Nutzer ab. So sind für die amtliche Statistik und die Konjunkturanalyse vor allem Datenquellen interessant, die tatsächliche Transaktionen abbilden und dadurch helfen, die laufende wirtschaftliche Aktivität zeitnäher und umfassender abzubilden. Eine mögliche Verstetigung der Datenauswertung, die vielfach mit einem besonders hohen Aufwand verbunden ist, ist hier allerdings ein zentrales Anwendungskriterium. Für vertiefende makroökonomische Analysen ist eine mögliche Verstetigung des Datenzugangs dagegen ein weniger relevantes Anwendungskriterium.
- Angesichts der hohen Komplexität der neuen Datenquellen ist die vorliegende Evidenz vielfach noch nicht ausreichend, um ihren Nutzen konkret beziffern zu können. Die enormen Potenziale zeigen sich jedoch schon alleine daran, dass bereits nahezu alle Facetten des Wirtschaftsgeschehens digital erfasst werden. Diese Daten können bislang aber noch nicht vollständig ausgewertet werden, weil sie noch nicht systematisch gesammelt werden oder Datenschutzrichtlinien dem entgegenstehen. In dem Ausmaß, wie sich der Zugang zu den Daten verbessert und methodische Weiterentwicklungen voranschreiten, werden die Potenziale von Big Data weiter ausgeschöpft werden. Vieles spricht dafür, dass Big Data in vielen Anwendungsfeldern vor allem komplementär zu den Daten der konventionellen Wirtschaftsstatistik zum Einsatz kommen werden.

Schlüsselwörter: Big Data, makroökonomische Analyse, Konjunktur, Konjunktur Deutschland, Machine Learning

Core Results

- Big data describe new and unconventional data sources in contrast to conventional data sources. Data from these new sources usually exhibit a high volume, are often unstructured, and are available in real-time and at high frequency. These new data have a high bandwidth and complexity because they are not collected for the purpose of economic analysis, but are rather by-products of various applications.
- Many of these new data sources contain actual transactions or prices (electronic payment transactions, scanner data, trading platforms) or directly reflect current economic activity in another form (satellite images, traffic data). The resulting data have particularly potential for macroeconomic monitoring and forecasting.
- Other data sources reflect moods and interests of users or the relevance of topics (social media, internet search queries, press articles). They are less closely related to the current economic development. However, this data can be used to build indicators that provide more comprehensive information than existing indicators (e.g. for consumer confidence) or that cannot be constructed with other conventional data sources (e.g. economic policy uncertainty).
- The effort required to make the data usable for macroeconomic analyses differs depending on the data source. Some of the data sources are already regularly used to build macroeconomic indicators (economic policy uncertainty, weather conditions, truck toll mileage), or the data are made available in a ready to use format for empirical analyses (internet search queries). If the data have to be organized and prepared, this often involves a very high organizational and methodological effort. Much of the data is not freely available or is only available to a limited extent.
- The potential of big data for macroeconomic analysis is particularly high in those areas for which conventional economic statistics only provide indicators with a delay, or for which they are not very reliable. For Germany – as for many other national economies – these areas include private consumer spending and value added in the service industries. However, big data can also provide valuable contributions complementary to conventional data, since big data are generally available earlier and in greater detail.
- The potential and limits of big data for macroeconomic analysis also depend on the individual needs of the user. For official statistics and business cycle analysis, for example, data sources that contain actual transactions and thus help to depict current economic activity in a more timely and comprehensive manner are of particular interest. However, a continuous data access, which can be associated with a particularly high effort, is a central application criterion here. For in-depth macroeconomic analyses a continuous data access is a less relevant criterion.
- In view of the high complexity of big data, the available evidence is in many cases not yet sufficient to be able to quantify their benefits concretely. However, the high potential is evident from the fact that almost all aspects of economic activity are already digitally recorded. This data cannot yet be fully exploited because it is not systematically collected or due to data protection guidelines. To the extent that access to the data improves and methodological developments advance, the potential of big data will be further exploited. Big data will be used in many applications above all as a complement to conventional data.

Keywords: Big Data, macroeconomics, business cycle, business cycle Germany, Machine Learning

Inhaltsverzeichnis

Tabellenverzeichnis

Abbildungsverzeichnis und Kasten

BIG DATA IN DER MAKROÖKONOMISCHEN ANALYSE

Martin Ademmer, Joscha Beckmann, Eckhardt Bode, Jens Boysen-Hogrefe, Manuel Funke, Philipp Hauber, Tobias Heidland, Julian Hinz, Nils Jannsen, Stefan Kooths, Mareike Söder, Vincent Stamer und Ulrich Stolzenburg

1 Einleitung[1]

Im Zuge der Digitalisierung fallen zunehmend neue, elektronisch erfasste Daten mit immer größerem Volumen an. Diese Daten werden häufig unter dem Schlagwort Big Data zusammengefasst. Sie stammen meist aus unkonventionellen, unstrukturierten Datenquellen, liegen häufig in Echtzeit vor und sind in der Regel deutlich umfangreicher und in höherer Frequenz verfügbar als solche der konventionellen Wirtschaftsstatistik. Allerdings weisen diese neuen Daten eine hohe Bandbreite und Komplexität auf, weil sie nicht für die Analyse von ökonomischen Fragestellungen erhoben werden, sondern vielmehr als Nebenprodukt unterschiedlicher Anwendungen anfallen (IMF 2017). Zusammengenommen mit ihrem häufig großen Volumen reicht das klassische Instrumentarium der wirtschaftswissenschaftlichen Forschung deshalb nicht immer aus, um diese Daten effizient auszuwerten. Da gleichzeitig auch die Analysemethoden, beispielsweise im Bereich des maschinellen Lernens, fortlaufend weiterentwickelt werden, bieten Big Data zahlreiche neue Möglichkeiten (Varian 2014). Während diese von Unternehmen und in einigen wissenschaftlichen Disziplinen bereits vielfach genutzt werden, spielen sie für makroökonomische Analysen bislang eher eine untergeordnete Rolle.

Motivation

Freilich können Big Data auch für viele makroökonomische Fragestellungen neue Möglichkeiten eröffnen. So ergeben sich für die Konjunkturanalyse und -prognose schon alleine deshalb Potenziale, weil Big Data häufig sehr zeitnah oder sogar in Echtzeit und in hoher Frequenz zur Verfügung stehen. Konventionelle Daten werden dagegen erst mit einiger Verzögerung publiziert. So wird die erste Schnellschätzung zum Bruttoinlandsprodukt für Deutschland erst 30 Tage (bis zum zweiten Quartal 2020: 45 Tage) nach Ablauf eines Quartals vom Statistischen Bundesamt bekanntgegeben. Auch werden die einschlägigen Konjunkturindikatoren in der Regel „nur“ auf Monatsdatenbasis (häufig mit einer ähnlichen zeitlichen Verzögerung) bereitgestellt. Vor diesem Hintergrund können Big Data insbesondere für die Prognose der laufenden Entwicklung (Nowcast) nützlich sein. Verbesserungspotenziale ergeben sich zudem für Bereiche, für die verhältnismäßig wenige zuverlässige Frühindikatoren und Aktivitätsmaße in höherer Frequenz vorliegen. Dazu zählen auch große Bereiche wie die Dienstleistungsbranche, deren Anteil an der gesamten Bruttowertschöpfung in Deutschland rund 70 Prozent ausmacht. Auch für niedrigere Aggregationsebenen, beispielsweise auf der Ebene von Wirtschaftszweigen, sind Daten häufig erst mit großer Verzögerung und in geringerer Frequenz verfügbar. Schließlich spielen Erwartungen und darauf gestützte Handlungspläne der wirtschaftlichen Akteure für die konjunkturelle Entwicklung eine

[1] Dieser Kieler Beitrag zur Wirtschaftspolitik entspricht weitgehend einem Gutachten, das die Autoren im Auftrag des Bundesministeriums für Wirtschaft und Energie im Jahr 2020 abgeschlossen haben.

wichtige Rolle. In dem Maße, wie entscheidungsvorbereitende Aktivitäten – etwa in Form von Suchanfragen – Datenspuren hinterlassen, lassen sich aus den neuen Datenquellen gegebenenfalls neue oder umfassendere Erwartungsindikatoren gewinnen.

Darüber hinaus haben Big Data auch das Potenzial, präzisere Informationen zu liefern, nicht zuletzt da sie im Gegensatz zu amtlichen Statistiken (Döhrn 2019a) in der Regel keinen Revisionen unterliegen und zumindest zum Teil deutlich umfangreichere Datenquellen abbilden, ohne dass dadurch zusätzlicher Erhebungsaufwand entstünde. In diesem Zusammenhang können aus ihnen nicht nur neue Indikatoren gebildet werden, sondern sie können auch komplementär zu bereits vorliegenden Indikatoren einen wertvollen Beitrag für die Konjunkturanalyse und -prognose liefern.

Schließlich können Big Data neue Impulse für die makroökonomische Analyse liefern, beispielsweise indem sich durch zusätzliche Informationen und Indikatoren neue wirtschaftliche Wirkungszusammenhänge aufzeigen lassen oder die Auswirkungen verschiedener Einflussfaktoren auf Konjunktur und Wachstum besser identifiziert werden können. Für Konjunkturprognosen ergeben sich dadurch wiederum große Potenziale für Vorhersagen und Analysen über die kurze Frist hinaus, nicht zuletzt durch Maße zur Einschätzung der Entscheidungsunsicherheit (z.B. Politikunsicherheit). Auch für die Evaluation von Prognosen, die Bewertung wirtschaftspolitischer Maßnahmen oder die Entwicklung makroökonomischer Modelle könnten sich wertvolle Erkenntnisse gewinnen lassen. Angesichts dieser hohen Bandbreite von möglichen Einsatzgebieten werden die Potenziale von Big Data schon seit längerem nicht nur von einzelnen Wissenschaftlern, sondern auch von Zentralbanken (BIS 2015; Cœuré 2017), internationalen Organisationen (IMF 2017; Vereinte Nationen 2020) bzw. Statistikämtern (Eurostat 2020; Statistisches Bundesamt 2020a) diskutiert.

Ziel des Gutachtens

Dieses Gutachten liefert einen systematischen Überblick über die mit Big Data in Zusammenhang stehenden relevanten Daten und deren Potenzial für die makroökonomische Analyse, mit einem besonderen Schwerpunkt auf die Konjunkturbeobachtung und -prognose. Dabei werden nicht nur die Möglichkeiten ausführlich diskutiert, sondern auch die Herausforderungen, die sich aufgrund der Datenverfügbarkeit, Datenbeschaffenheit, methodischer Anforderungen oder der Handhabbarkeit der Daten ergeben können. Das Augenmerk liegt auf Datenquellen, die bereits für Analysen in den Wirtschaftswissenschaften oder verwandten Disziplinen eingesetzt worden sind. Es werden aber auch die Potenziale möglicher zukünftiger Datenquellen besprochen; dazu können Daten zählen, die bereits erfasst, aber noch nicht systematisch aufbereitet werden oder für wissenschaftliche Analysen verfügbar sind. In dem Gutachten wird ein besonderer Fokus auf die deutsche Volkswirtschaft gelegt. Freilich dürften sich die Schlussfolgerungen für Deutschland mit denen für andere fortgeschrittene Volkswirtschaften häufig ähneln, da sowohl die Verfügbarkeit von Big Data als auch die Abdeckung des Wirtschaftsgeschehens durch die konventionelle Wirtschaftsstatistik vielerorts vergleichbar sind.

Woran lassen sich die Potenziale und Grenzen festmachen?

Die Potenziale und Grenzen können anhand zahlreicher Kriterien beurteilt werden. Für die ökonometrische Auswertung spielen die Frequenz und insbesondere der Zeitraum, für den die Daten verfügbar sind, eine wichtige Rolle. So liegen neue Datenquellen häufig nur für kürzere Zeiträume vor als konventionelle Daten. Dadurch kann sich ihr Potenzial für den Einsatz in Zeitreihenmodellen (zumindest vorübergehend) deutlich verringern. Sofern historische Daten nicht verfügbar sind, sondern vom Anwender selbst gesammelt werden müssen, könnte es somit sogar einer Vorlaufzeit von mehreren Jahren bedürfen, bevor neue Datenquellen für quantitative Analysen nutzbar sind. Für die Konjunkturbeobachtung und -prognose ist zudem relevant, wie zeitnah neue Datenpunkte am aktuellen Rand verfügbar sind. In diesem Zusammenhang ist ein weiteres Kriterium, ob ein regelmäßiger Zugang zu den

Daten gewährleistet ist, der es ermöglicht, die darauf aufbauenden Analyseergebnisse und Prognosen zu verstetigen. Eine Verstetigung kann beispielsweise dadurch erschwert werden, dass kein freier Zugang zu den Daten besteht oder die Aufbereitung der Daten sehr aufwendig und zeitintensiv ist. Auch der Umfang und die Datenqualität sind relevante Kriterien. Während die Datenqualität der konventionellen Wirtschaftsstatistik durch die Statistikämter regelmäßig sorgfältig geprüft wird und vereinbarten, häufig international einheitlichen, Mindestanforderungen entspricht, gilt dies für Daten aus unkonventionellen Quellen nicht. Ausschlaggebend für die Datenqualität kann sein, ob die Daten über die Zeit vergleichbar und repräsentativ sind. Speisen sich Daten beispielsweise aus Nutzeranwendungen, so können sich die Nutzungsintensität und das typische Profil der Nutzer im Zeitablauf und zwischen alternativen Datenquellen (z.B. Facebook oder Twitter) unterscheiden. Schließlich ist auch der Aufwand, der betrieben werden muss, um Big Data für makroökonomische Analysen nutzbar zu machen, zum Teil deutlich größer als bei konventionellen Daten. So sind für die Aufbereitung und Auswertung der Daten vielfach Methoden, wie das maschinelle Lernen, nötig, die bislang noch nicht zum Standardinstrumentarium in der wirtschaftswissenschaftlichen Ausbildung zählen. Eine Herausforderung besteht in diesem Zusammenhang auch darin, dass sich die Datenbeschaffenheit je nach Quelle stark unterscheiden kann. Zudem gehen die Anforderungen an die IT-Infrastruktur in Bezug auf die Speicherkapazitäten und die Leistungsfähigkeit zum Teil über die gewöhnlich zur Verfügung stehende Ausstattung hinaus.

Neben diesen Kriterien, die sich zum Teil recht konkret festmachen lassen, gibt es auch noch eine Reihe von weicheren Kriterien. Dazu zählt nicht zuletzt, wie gut das jeweilige Anwendungsfeld bereits durch konventionelle Variablen bzw. Indikatoren abgedeckt ist. So sind die Potenziale für Bereiche, für die derzeit kaum geeignete Indikatoren zeitnah vorliegen, augenscheinlich höher. Dazu zählen neben der Dienstleistungsbranche auch die privaten Konsumausgaben, für die zwar einige Indikatoren vorliegen (z.B. Einzelhandelsumsätze, Verbrauchervertrauen), diese aber nur eine vergleichsweise geringe Prognosegüte aufweisen. Freilich können sich auch für Wirtschaftsbereiche, die durch die konventionelle Indikatorik vergleichsweise gut abgedeckt sind, wie z.B. die Industrie, große Potenziale ergeben, sofern durch Big Data relevante Informationen zeitnäher zur Verfügung gestellt werden können oder ein umfassenderes Bild liefern.

Das Potenzial der neuen Datenquellen soll auch anhand der bereits vorliegenden Literatur abgeschätzt werden. Allerdings lassen sich die Ergebnisse aus der Literatur nicht immer verallgemeinern, sodass das Potenzial selbst dann häufig nicht präzise bewertet werden kann, wenn Big Data bereits auf makroökonomische Fragestellungen angewendet worden sind. So kann die ausgewiesene relative Prognosegüte von neuen Daten davon abhängen, welche Vergleichsmodelle verwendet worden sind. Für den Vergleich werden häufig einfache univariate Ansätze herangezogen, die für mögliche Anwender jedoch in der Regel nicht die relevanten Alternativen darstellen. Auch können die Ergebnisse vom jeweiligen Versuchsaufbau (z.B. Prognosezeitraum, Verwendung von Echtzeitdaten, verwendete Modellklassen und -spezifikationen) abhängen und zwischen Ländern und über die Zeit variieren. Somit sagen Ergebnisse vorliegender Studien eher etwas über die grundsätzliche Eignung der Datenquellen für bestimmte Anwendungsfelder aus und weniger über ihr Potenzial, derzeit verwendete Prognosemethoden oder vorliegende Analyseergebnisse merklich zu verbessern.

Vor diesem Hintergrund hat eine Bewertung der unterschiedlichen Datenquellen hinsichtlich ihrer Potenziale und Grenzen immer subjektive Elemente, zumal sie auch je nach Zielen und Voraussetzungen der Anwender unterschiedlich ausfallen kann.

Abgrenzung von Big Data

Auch wenn das Schlagwort Big Data bereits seit längerem verwendet wird, um neue elektronisch erfasste Datenquellen von konventionellen Daten abzugrenzen, liegt bislang keine allgemein gültige Definition vor. Häufig wird Big Data an fünf recht allgemeinen Eigenschaften festgemacht („5 Vs"; Bello-Orgaz et al. 2016; Blazquez und Domenech 2018). So liegen Big Data in der Regel in einem sehr großen Volumen vor (Volume), das sich aufgrund der hohen Geschwindigkeit und der hohen Frequenz, mit der neue Daten generiert werden, rasch vergrößert (Velocity). Da sie in der Regel als Nebenprodukt von anderen Prozessen anfallen, weisen sie eine hohe Bandbreite und Komplexität bezüglich ihrer Datenbeschaffenheit auf (Variety), die über das von konventionellen Daten bekannte Maß hinausgeht. Auch kann die Datenqualität geringer sein, als man dies aus der Wirtschaftsstatistik gewohnt ist, nicht zuletzt, weil die Daten zum Teil verzerrt oder nicht repräsentativ sind (Veracity). Schließlich kann das technologisch und geschäftlich sehr dynamische Umfeld, aus dem Big Data zum Teil stammen, dazu führen, dass sich Eigenschaften und Abdeckung der Daten im Zeitablauf deutlich verändern oder sogar die längerfristige Verfügbarkeit der Daten in Frage stellen (Volatility).

Datensätze aus dem Umfeld von Big Data lassen sich gegenüber konventionellen Daten auch anhand der Dimensionen spezifischen Eigenschaften abgrenzen. Sie basieren in der Regel auf einem deutlich größeren Querschnitt an Beobachtungen und liegen in deutlich höherer Frequenz vor (zumeist Tagesdaten oder höher) als konventionelle Daten. Dafür ist der Zeitraum, für den sie vorliegen, in der Regel deutlich kürzer als bei konventionellen Daten. Nicht zuletzt, weil sie für makroökonomische Analysen mit konventionellen Daten in Verbindung gesetzt werden, kann dies ihren Nutzen derzeit noch spürbar begrenzen.

Die Vereinten Nationen klassifizieren unterschiedliche Datenquellen aus dem Bereich Big Data anhand des jeweiligen Ursprungs der Daten (Vereinte Nationen 2015). Demzufolge können Big Data in drei Kategorien eingeteilt werden:

Durch Nutzer generierte Daten entstammen im Wesentlichen sozialen Netzwerken.

Prozessbasierte Daten decken neben administrativen Daten vor allem geschäftliche Transaktionen ab.

Maschinengenerierte Daten basieren auf mittels Sensoren erfassten Aktivitäten.

Auch wenn noch keine einheitliche Definition vorliegt, so werden doch eine Reihe von Datenquellen typischerweise dem Schlagwort Big Data zugeordnet. Dazu zählen beispielsweise Daten aus sozialen Medien, Mobilfunkdaten, Satellitendaten oder Zahlungsvorgänge. Andere Daten, die zumindest einige der typischen Eigenschaften aufweisen, werden dagegen nicht immer als Big Data klassifiziert. Dazu zählen beispielsweise Finanzmarktdaten sowie administrative Daten, die sehr umfangreich ausfallen können. Während Finanzmarktdaten bereits seit geraumer Zeit für wissenschaftliche Analysen, auch für makroökonomische Fragestellungen, eingesetzt werden, stehen einer stärkeren Verwendung administrativer Daten neben technischen Hürden auch hohe Anforderungen an den Datenschutz entgegen, die es beispielsweise erschweren, verschiedene Datensätze miteinander zu verknüpfen.

In diesem Gutachten behandelte Datenquellen

In diesem Gutachten werden die Potenziale und Grenzen von Big Data für die makroökonomische Analyse entlang unterschiedlicher Datenquellen analysiert. Aufgrund der hohen Anzahl von möglichen Datenquellen, insbesondere im Bereich der sozialen Medien, kann nur eine Auswahl von Quellen ausführlich behandelt werden. Bei der Auswahl der Datenquellen wurde darauf geachtet, dass die bislang am häufigsten in der wissenschaftlichen Literatur verwendeten Quellen enthalten sind und dass sie die übergeordneten Kategorien, also soziale Medien sowie prozess- und sensorgenerierte Daten, möglichst gut abdecken. So werden Potenziale und Grenzen von Nachrichten und Presseartikeln als

Datenquelle diskutiert. Für Daten aus sozialen Netzwerken werden beispielhaft Facebook und Twitter herangezogen. Suchanfragen werden anhand von Google Trends behandelt. Im Bereich der prozessbasierten Daten werden Girocard- oder Kreditkartentransaktionen, Swift- sowie Scannerdaten diskutiert. Zudem werden Handelsplattformen als mögliche Datenquelle in Betracht gezogen. Für über Sensoren gewonnene Daten werden Fernerkundungsdaten (insbesondere Satellitendaten), Schiffspositionsdaten sowie weitere Verkehrsdaten und Mobilfunkdaten diskutiert. Viele der für diese Datenquellen diskutierten Aspekte dürften sich auf andere Quellen übertragen lassen

Zum Teil werden dem Schlagwort Big Data auch Ansätze zugeordnet, die zwar recht große Datensätze auswerten, diese aber aus konventionellen Daten bestehen. Dazu zählen beispielsweise große Datensätze bestehend aus konventionellen Konjunkturindikatoren, die mit dynamischen Faktormodellen ausgewertet werden, um makroökonomische Größen wie das Bruttoinlandsprodukt zu prognostizieren oder wesentliche Einflussfaktoren zu identifizieren (Bok et al. 2018; Hauber 2018). Diese Ansätze sind bereits umfangreich in der Literatur diskutiert worden und werden in diesem Gutachten nicht behandelt.

Bedeutung von Big Data für die amtliche Statistik

Die amtliche Statistik stellt als Hauptlieferant für objektive und qualitativ hochwertige Daten eine wichtige Grundlage nicht nur für die makroökonomische Analyse, sondern auch für die Wirtschaftspolitik, Unternehmen oder Verwaltung dar. Durch Big Data ergeben sich auch für die amtliche Statistik zahlreiche neue Möglichkeiten und Herausforderungen. Zum Teil ähneln diese denen, die sich für die makroökonomische Analyse ergeben. Zum Teil weichen sie aufgrund des spezifischen institutionellen Hintergrunds nationaler Statistikämter aber auch voneinander ab.

Möglichkeiten ergeben sich beispielsweise, wenn durch Big Data umfassendere Informationen in die bereits regelmäßig veröffentlichten Wirtschaftsstatistiken einfließen und sie so verbessern. Da Big Data oft sehr zeitnah oder sogar in Echtzeit verfügbar sind, können sie zudem dazu beitragen, dass amtliche Statistiken rascher veröffentlicht werden können. Ferner können die amtlichen Statistiken zum Teil detaillierter ausgestaltet werden, also beispielsweise auf niedrigerer Aggregationsstufe Informationen (früher) bereitstellen. Schließlich können durch Big Data auch neue Sachverhalte in der amtlichen Statistik abgebildet werden (Wiengarten und Zwick 2017). Letzteres kann für Statistikämter insbesondere dann relevant sein, wenn die Anforderungen an die von ihnen erstellten Statistiken vom Umfang her steigen. So ist mit der Verständigung über die sogenannten Sustainable Development Goals (SDG) auf der Ebene der Vereinten Nationen vereinbart worden, dass die Statistikämter der Hauptlieferant für objektive und hochwertige Daten zur Beobachtung der Fortschritte sein sollen; diese Ziele umfassen allerdings auch Bereiche, die von der amtlichen Statistik bis dato noch nicht oder nicht vollständig abgedeckt werden. In dem Ausmaß, in dem sich die Erhebung von Statistiken aus Big Data automatisieren lässt und dadurch andere aufwendigere Erhebungsmethoden ersetzt werden können, könnten die Statistikämter leistungsfähiger und die Auskunftsgebenden, insbesondere Unternehmen, entlastet werden.

Big Data gehen in die amtliche Statistik als sekundärstatistische Quelle ein – neben den primärstatistischen Erhebungen, mit denen die statistischen Ämter selbst Daten (beispielsweise von Unternehmen) erheben. Vorteile solcher sekundärstatistischen Quellen sind, dass weder für das Amt noch für Auskunftsgebende zusätzlicher Erhebungsaufwand entsteht, da Big Data als Nebenprodukt von anderen Aktivitäten anfallen (Schnorr-Bäcker 2018). Ein Nachteil solcher Quellen ist, dass die Statistikämter keinen Einfluss auf die Definition der erhobenen Merkmale und die Klassifizierung der Daten haben. Gegenüber den bislang verwendeten Sekundärquellen, die überwiegend der öffentlichen Verwaltung entstammen (z.B. Steuerstatistik oder Arbeitsmarktstatistik), könnte dieser Nachteil bei Big Data zum

Teil noch gravierender ausfallen, da sie in der Regel unsystematischer erfasst werden als Verwaltungsdaten. Dafür sind sie für die amtliche Statistik zeitnäher verfügbar. Ein limitierender Faktor für den Einsatz von Big Data sind die Anforderungen der amtlichen Statistik an die Datenqualität. Die Qualität unterliegt dabei vorgeschriebenen und auf Ebene des Europäischen Statistischen Systems und der Vereinten Nationen vereinbarten Mindestanforderungen, die auch für Big Data gelten. Auch Datenschutzanforderungen können gegen einen Einsatz bestimmter Datenquellen für die amtliche Statistik sprechen, beispielsweise wenn Anwendungen für Big Data mit eigenen Befragungsdaten verknüpft werden sollen. Ferner ist der Aspekt der Verstetigung auch für Statistikämter von zentraler Bedeutung.

In einigen Bereichen setzt das Statistische Bundesamt Big Data oder damit in Zusammenhang stehende Methoden bereits ein bzw. plant diese zeitnah einzusetzen. So bildet das Amt aus Mautdaten einen Konjunkturindikator, der regelmäßig zeitnah veröffentlich wird (Cox et al. 2018). Es verwendet darüber hinaus Web Scraping Methoden für die Messung von Preisen im Online-Handel (Brunner 2014; Blaudow und Seeger 2019). Freilich stellt dies weniger die Berücksichtigung einer neuen Datenquelle dar, sondern vielmehr die automatisierte Erfassung von Preisen, die zuvor manuell erfasst wurden. Auch könnten demnächst Satellitenbilder für die Abschätzung von Ernteerträgen eingesetzt werden. Zudem werden neue mögliche Datenquellen derzeit eingehend daraufhin geprüft, ob sie zukünftig in die amtliche Statistik einfließen können. So untersucht das Statistische Bundesamt, inwieweit Fernerkundungsdaten (insbesondere Satellitendaten) und geografische Informationen im Allgemeinen beispielsweise für die Analyse der Landbedeckung und -nutzung oder zur Erhebung von Daten in Bezug auf die Sustainable Development Goals genutzt werden können (Arnold und Kleine 2017; Gebers und Graze 2019). Auch wird der Einsatz von Scannerdaten zur Preismessung (Bieg 2019), von Web Scraping für die internetgestützte Erfassung offener Stellen (Rengers 2018) und von Mobilfunkdaten für die Bevölkerungs- und Pendlerstatistik (Statistisches Bundesamt 2019; Digitale Agenda) geprüft.[2]

Die Ergebnisse ihrer Analysen bringt das Statistische Bundesamt in ein Projekt des Netzwerks des Europäischen Statistischen Systems ein, das zum Ziel hat, die Eignung von Big Data für die amtliche Statistik zu prüfen (ESSnet Big Data). Innerhalb des Projekts prüfen die beteiligten Ämter weitere Datenquellen; dazu zählen Unternehmensinformationen aus dem Internet, die mittels Web Scraping und Textanalysen ausgewertet werden (Oostrom et al. 2016), Schiffspositionsdaten, intelligente Energiezähler oder Finanztransaktionen (Eurostat 2020).[3]

Insgesamt werden Big Data zukünftig wohl eine zunehmende Bedeutung für die amtliche Statistik erlangen. Dabei dürften sie in vielen Bereichen komplementär zu herkömmlichen Erhebungsmethoden verwendet werden, diese jedoch nicht ersetzen. Da die Anforderungen an die Qualität der Daten und Datenschutzanforderungen für die amtliche Statistik sehr hoch sind, werden Big Data wohl erst nach und nach in die amtliche Statistik integriert werden.

Methoden für Big Data

Für die Analyse von Big Data kommen nur Methoden in Frage, die mit sehr großen Datensätzen umgehen können. Zum Teil liegen dafür bereits in der empirischen Makroökonomik etablierte Werk-

[2] Das Statistische Bundesamt veröffentlicht online ausgewählte Projektergebnisse aus Machbarkeitsstudien mit Bezug zu Big Data unter der Rubrik „EXDAT – Experimentelle Daten“: https://www.destatis.de/DE/Service/EXDAT/_inhalt.html#sprg364080.

[3] Ein Überblick über die Inhalte des ESSnet Big Data Projekts findet sich hier: https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/Main_Page.

Eurostat stellt zudem ausgewählte Ergebnisse aus neuen unkonventionellen Datenquellen zur Verfügung: https://ec.europa.eu/eurostat/web/experimental-statistics.

zeuge vor, wie z.B. Faktormodelle (Stock and Watson 2002; Giannone et al. 2008) oder bayesianische Vektorautoregressionen (Banbura et al. 2010), die nicht zuletzt für Prognosen von makroökonomischen Größen regelmäßig verwendet werden. Sehr viel häufiger kommen bei Big Data aber Methoden des Machine Learning zum Einsatz. Unter dem Begriff Machine Learning wird eine große Bandbreite von Verfahren zusammengefasst, mit deren Hilfe sich Datensätze strukturieren lassen (Unsupervised Learning) oder Variablen prognostiziert oder klassifiziert werden können (Supervised Learning) (Athey und Imbens 2019).

Kennzeichnend für das Machine Learning sind die flexiblen funktionalen Zusammenhänge, die zwischen den Variablen zugelassen werden, um große Datenmengen auszuwerten. Der Übergang zur klassischen Ökonometrie ist dabei fließend. So zählen zum Machine Learning Verfahren, die auf linearen Regressionen aufbauen und so angepasst werden, damit sie aus einer Vielzahl von möglichen erklärenden Variablen, ein geeignetes Modell identifizieren können. Solche Verfahren werden in den Wirtschaftswissenschaften bereits regelmäßig eingesetzt, um beispielsweise Prognosemodelle zu spezifizieren (Bai und Ng 2008). Andere Methoden, die komplexere funktionale Zusammenhänge zwischen den Variablen zulassen, sind in der Makroökonomik bislang weniger stark verbreitet.

Einige Methoden des Machine Learning (Unsupervised Learning) dienen dazu, große Datensätze zu verdichten oder zu klassifizieren, ohne sie mit einer bestimmten Variablen von Interesse in Verbindung zu setzen (Athey und Imbens 2019). Solche Verfahren sind einer weiterführenden empirischen Analyse häufig vorgeschaltet, um komplexe Datensätze besser handhabbar zu machen. Unter den konventionellen empirischen Methoden entspricht die Faktoranalyse einem solchem Ansatz, bei dem der Datensatz zunächst zu Hauptkomponenten zusammengefasst wird, die dann in einem zweiten Schritt beispielsweise als Variablen für Prognosemodelle verwendet werden. Die im Bereich des Machine Learning dafür verwendeten Methoden ähneln häufig Clusteranalysen, bei denen die Variablen anhand von vorgegebenen statistischen Kriterien in Gruppen aufgeteilt werden. Diese Gruppen können dann beispielsweise mit weiterführenden Methoden individuell ausgewertet werden. In Textanalysen werden Texte so zunächst auf bestimmte Themen oder Trends reduziert, um diese darauf aufbauend mit makroökonomischen Variablen in Verbindung zu setzen (Gentzkow et al. 2019).

Viele Methoden des maschinellen Lernens (Supervised Learning) ähneln klassischen Prognoseansätzen, dienen sie doch dazu, basierend auf einer großen Anzahl von möglichen Indikatoren X eine geeignete Funktion f auszuwählen, um den Verlauf einer Variablen y möglichst gut zu erklären bzw. zu prognostizieren (Mullainathan und Spiess 2017): $y = f(X)$. Während für viele empirische makroökonomische Anwendungen die Schätzergebnisse für $f(X)$ im Vordergrund stehen, die Auskunft über die kausalen Zusammenhänge zwischen den Variablen geben können, werden beim Machine Learning Modelle anhand ihrer Prognosegüte für y, ausgewählt. Die resultierenden Schätzergebnisse können somit nicht ohne weiteres ökonomisch interpretiert werden. Die Anwendungsfelder für Machine Learning konzentrieren sich dementsprechend auf Fragestellungen, in denen die Prognosegüte von besonderer Bedeutung ist. Neben einschlägigen Prognoseanwendungen (wie z.B. die Kurzfristprognose für das Bruttoinlandsprodukt) werden sie insbesondere im Bereich Big Data eingesetzt, um komplexe Datensätze, wie Bild-, Audio- oder Video-Dateien, auszuwerten (Mullainathan und Spiess 2017). In diesem Zusammenhang sind sie für makroökonomische Analysen schon alleine deshalb häufig relevant, um umfangreiche bzw. nicht numerische Datensätze, wie beispielsweise Text- oder Satellitendaten, so aufzubereiten, damit sie mit makroökonomischen Zeitreihen in Verbindung gesetzt werden können.

Ein grundsätzliches Problem dabei ist, dass die Zahl der zur Verfügung stehenden erklärenden Variablen häufig sehr groß ist. Dann können zahlreiche Modelle spezifiziert werden, die den Verlauf der abhängigen Variablen schon allein aufgrund der hohen Anzahl an erklärenden Variablen perfekt oder

nahezu perfekt erklären können, ohne dass darauf auf den tatsächlichen Erklärungsgehalt bzw. die Prognosegüte geschlossen werden kann (Overfitting). Um diesem Problem zu begegnen, werden die Methoden dahingehend restringiert bzw. regularisiert, dass der Umfang bzw. die Komplexität der Modelle begrenzt wird (Regularization). Allgemein wird bei Methoden des Machine Learning also eine Verlustfunktion L(·) minimiert, welche die Abweichungen der prognostizierten Werte $\hat{y} = \hat{f}(X)$ von den tatsächlichen Werten enthält, unter Berücksichtigung einer Regularisierung $R(\hat{f})$ (Mullainathan und Spiess 2017):

$$\arg\min_{\hat{f}} L\left(y - \hat{f}(X)\right) \; s.t. \; \mathrm{R}(\hat{\mathrm{f}}) \tag{1}$$

Die Regularisierung kann – mittels eines „Bestrafungsterm" – direkt über die Verlustfunktion erfolgen, indem beispielsweise die zusätzliche Aufnahme von Variablen in das Modell mit einem zusätzlichen Verlust versehen wird. Die Komplexität des Modells kann aber auch unmittelbar regularisiert werden, indem beispielsweise die maximal zulässige Anzahl der geschätzten Parameter festgelegt wird. Insgesamt sind eine Vielzahl verschiedener Formen der Regularisierung möglich, und a priori ist nicht klar, welche zu dem bestmöglichen Ergebnis führt.

Da aus dem (In-Sample) Erklärungsgehalt eines Modells nicht unmittelbar auf seine Prognosegüte geschlossen werden kann, werden für die Modellselektion regelmäßig (Out-of-Sample) Prognosevergleiche herangezogen. Diese können auch dazu dienen, eine geeignete Regularisierung für (1) auszuwählen oder gewichtete Kombinationen aus verschiedenen Modellen zu bilden, die eine höhere Prognosegüte aufweisen als einzelne Modelle; diesen Modellen können dabei auch unterschiedliche Machine Learning Methoden zugrunde liegen (Super Learner). Da Big Data häufig nur für vergleichsweise kurze Zeitreihen vorliegen, wird für die Prognosevergleiche der vorliegende Datensatz zufällig in mehrere Teile (Folds) zerlegt und jeweils eine dieser Teilmengen prognostiziert, während der jeweils übrige Datensatz (Training Sample) für die Schätzung gemäß (1) verwendet wird (Cross Validation). Anders als in klassischen Out-of-Sample-Prognoseevaluationen spielt die zeitliche Ordnung von Schätz- und Prognosezeitraum hier keine Rolle. Cross Validation kann somit auch problemlos für Querschnittsdaten angewendet werden.

Einige der gebräuchlichsten Maschine Learning Methoden basieren auf einer quadratischen Verlustfunktion und ähneln damit linearen Kleinst-Quadrat-Schätzungen. Die Regularisierung erfolgt über einen zusätzlichen Term $\mathrm{R}(\hat{\beta})$ in der Verlustfunktion, der Abweichungen der Schätzparameter β von null mit einem zusätzlichen Verlust versieht:

$$\hat{\beta} \; = \; \arg\min_{\hat{\beta}} \{(y - X\hat{\beta})^2 + \lambda \mathrm{R}(\hat{\beta})\}, \tag{2}$$

wobei der Einfluss dieses zusätzlichen Terms auf die Schätzung über den Parameter λ erfolgt. In der Praxis haben sich unterschiedliche Regularisierungen etabliert. So gehen beim Ridge-Verfahren die quadrierten Schätzparameter in die Verlustfunktion ein. Dies führt im Ergebnis zu einer Verringerung (Shrinkage) der geschätzten Koeffizienten mit dem Ziel die Schätzunsicherheit und somit auch die Prognosefehler zu reduzieren. Beim Lasso-Verfahren gehen dagegen die absoluten Werte der Schätzparameter in die Verlustfunktion ein. Dies führt zu einer Verringerung der Modellgröße, da Koeffizienten auch auf null gesetzt werden (Variablenselektion bzw. Sparsity), wodurch besser nachvollzogen werden kann, welche Variablen als relevant eingestuft werden. In praktischen Anwendungen stößt diese Form der Regularisierung allerdings an Grenzen. So ist die Anzahl der von null verschiedenen Koeffizienten durch die Anzahl der verfügbaren Beobachtungen begrenzt, was in Datensätzen mit besonders vielen Variablen ein Nachteil sein kann. Zudem wird aus einer Gruppe von stark korrelierten

erklärenden Variablen – wie dies beispielsweise in herkömmlichen makroökonomischen Daten häufig der Fall ist – in der Regel nur eine Variable selektiert. Da die ausgewählte Variable häufig aber nicht stabil ist und schon bei kleinen Änderungen des Datensatzes oder der Regularisierung variiert, kann dadurch die Interpretation des Modells erschwert werden. In der Praxis hat sich gezeigt, dass Elastic Net-Ansätze, bei denen die beiden Verfahren kombiniert werden, häufig gute Ergebnisse liefern (Zou und Hastie 2005).

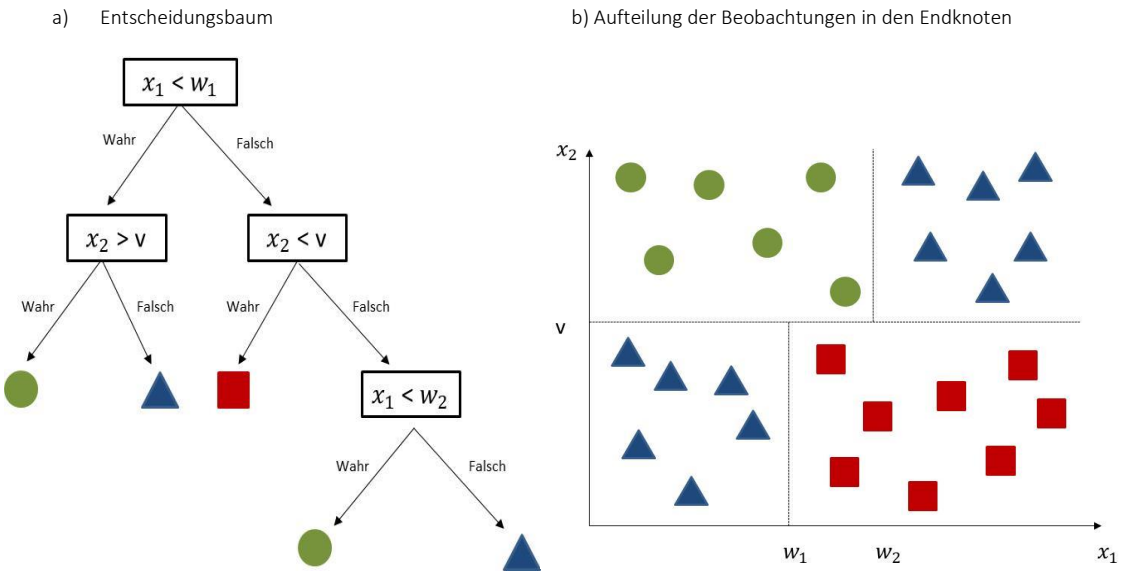
Die Gewichtung zwischen der quadratischen Verlustfunktion und den Abweichungen der Parameter von null muss mittels λ gesondert festgelegt werden. Für Modellkombinationen, wie beim Elastic Net, müssen zusätzliche Gewichte bestimmt werden. Typischerweise geschieht dies über einen Vergleich der Prognosegüte mittels Cross Validation. Solche Methoden werden nicht nur regelmäßig für die Auswertung von Big Data eingesetzt, sondern sind bereits vielfach für die Bearbeitung von konventionellen makroökonomischen Datensätzen angewendet worden, beispielsweise um aus einer großen Zahl von Indikatoren geeignete Prognosemodelle zu spezifizieren (Coulombe et al. 2019; Jung et al. 2018).

Andere im Machine Learning gebräuchliche Verfahren unterstellen keinen linearen Zusammenhang zwischen den Regressoren und der abhängigen Variablen, sondern ermöglichen eine flexiblere Modellierung. Hierzu zählen Regression Trees und die von ihnen abgeleiteten Random Forests sowie Neuronale Netze. Regression Trees unterteilen die Beobachtungen für die abhängige Variable entlang der erklärenden Variablen in Entscheidungsbäumen (Abbildung 1a). Die Modellprognose entspricht dann dem durchschnittlichen Wert der abhängigen Variablen in jedem Endknoten (Abbildung 1b).

Der Erklärungsgehalt der Regression Trees steigt, je mehr Entscheidungsstufen bzw. Knoten zugelassen werden, gleichzeitig erhöht sich dadurch aber auch das Risiko des Overfitting. Um dem entgegenzuwirken, wird die maximale Tiefe des Regression Trees festgelegt oder anhand statistischer Kriterien über die Verlustfunktion bestimmt. Trotz Regularisierung schneiden Prognosen einzelner Regression Trees häufig nicht gut ab, nicht zuletzt, da ihre Ergebnisse sehr sensibel hinsichtlich des jeweiligen Training Samples sind. Vor diesem Hintergrund hat es sich als nützlich erwiesen, verschiedene Regression Trees mittels Random Forrests oder Boosting zu kombinieren.

Random Forests kombinieren eine Vielzahl von Regression Trees, die jeweils auf einer zufällig ausgewählten Teilmenge des verfügbaren Datensatzes beruhen (Breimann et al. 1983; Breimann 1996). Derartige Verfahren umfassen naturgemäß eine Vielzahl möglicher Zusammenhänge und liefern in der Regel robustere Ergebnisse als einzelne Regression Trees (Athey und Imbens 2019). Boosted trees kombinieren hingegen sequentiell aufeinander aufbauende Regression Trees. Die einzelnen Regression Trees werden dabei hinsichtlich ihrer Tiefe vergleichsweise eng begrenzt (Döpke et al. 2017). Regression Trees und insbesondere Random Forrests sind in makroökonomischen Analysen bereits bei der Modellierung von Frühwarnsystemen für Bankenkrisen eingesetzt worden (Alessi und Detken 2018; Blustwein et al. 2020; Holopainen and Sarlin 2017; Tanaka et al. 2016), auch wenn es offenbar vom jeweiligen Analyserahmen abhängt, ob sie spürbare Verbesserungen gegenüber den herkömmlichen Methoden liefern (Beutel et al. 2019).

Abbildung 1:
Regression Tree[a]

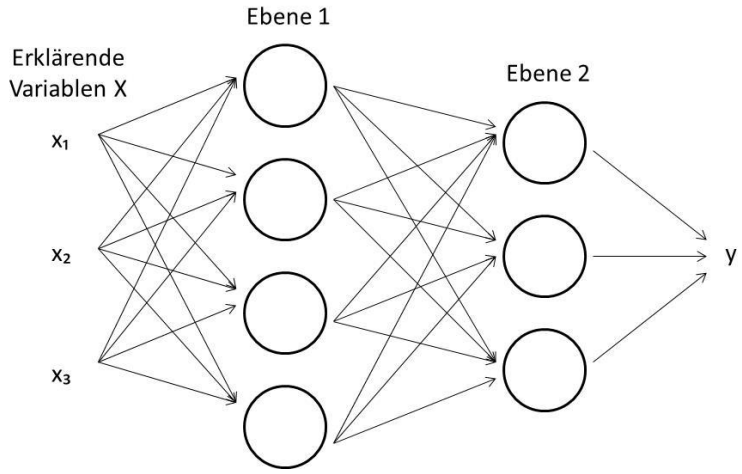


[a]Beispielhafter Regression Tree. Ausgehend von der Unterscheidung ob x_1 und x_2 ober- oder unterhalb von bestimmten Schwellenwerten (*v* und *w*) liegen, werden die Beobachtungen der zu erklärenden Variablen y unterschiedlichen Endknoten zugeteilt.

Quelle: Eigene Darstellung.

Neuronale Netze basieren nicht auf hierarchischen Entscheidungsbäumen, sondern auf miteinander verwobenen Entscheidungsknoten, über die die erklärenden Variablen über mehrere Ebenen hinweg flexibel mit der abhängigen Variablen in Verbindung gesetzt werden (Abbildung 2). Neuronale Netze können sehr komplexe Formen annehmen. Deshalb sind in der Regel umfangreichere Datensätze erforderlich als bei weniger komplexen Methoden, um robuste Ergebnisse zu erzielen. Auch können die Ergebnisse kaum noch nachvollzogen werden, sodass sie stärker als andere Machine Learning Methoden

Abbildung 2:
Grundstruktur eines neuronalen Netzes



Quelle: Eigene Darstellung in Anlehnung an Jung et al. (2018).

als Black Box betrachtet werden. Neuronale Netze werten die Daten in ihrer Gesamtheit ohne Berücksichtigung der zeitlichen Ordnung der Variablen aus, sodass sie für Zeitreihenanalysen ggf. entsprechend angepasst werden müssen. Erste Studien legen nahe, dass Neuronale Netze für die Prognose der gesamtwirtschaftlichen Aktivität auf Basis konventioneller Datensätze geeignet sind (Sachverständigenrat 2019: 54; Lörmann und Maas 2019; Jung et al. 2018).

Freilich gibt es darüber hinaus noch zahlreiche Variationen dieser Methoden bzw. andere Methoden, die unter dem Schlagwort Machine Learning zusammengefasst werden. Viele dieser Methoden stellen die Grundlage dafür dar, um die Potenziale von Big Data nicht nur für makroökonomische Analysen möglichst gut ausschöpfen zu können. Durch die hohe Vielfalt und Flexibilität der Methoden, können sie zwar für eine große Bandbreite von Datenquellen und Analysezielen angewendet werden. Allerdings stellt dies auch den Anwender vor große Herausforderungen. So ist a priori häufig nicht offensichtlich, welche Methode für die jeweilige Datenquelle und Fragestellung optimal geeignet ist. Zudem verfügt der Anwender, selbst nachdem er sich auf eine Methode festgelegt hat und die Modellspezifikation über Cross Validation auswählen möchte, noch über zahlreiche Freiheitsgrade, die das Ergebnis spürbar beeinflussen können. Dazu zählen beispielsweise welche Formen der Regularisierung zugelassen werden, wie die Cross Validation konkret ausgestaltet wird, wie die einfließenden Variablen transformiert werden oder ob theoretische Überlegungen bei der Variablenauswahl schon vorab berücksichtigt werden sollen. Aktuelle Forschungsarbeiten befassen sich u.a. damit, die Methoden hinsichtlich ihrer Transparenz und der expliziten Modellierung kausaler Zusammenhänge zu verbessern (Athey und Imbens 2019). Dies würde ihre Anwendungsmöglichkeiten auch für makroökonomische Analysen noch einmal zusätzlich erhöhen.

Aufbau der Analyse

Im folgenden Kapitel werden die Anwendungsfelder und Potenziale verschiedener Datenquellen aus dem Bereich Big Data in einem vergleichbaren Rahmen vorgestellt und diskutiert. In Kapitel 3 werden kurz mögliche zukünftige Potenziale von Big Data vorgestellt, die derzeit noch nicht umsetzbar wären, beispielsweise, weil die dafür notwendigen Daten noch nicht systematisch gesammelt oder erfasst werden. Schließlich werden in Kapitel 4 die Ergebnisse zusammengefasst und eingeordnet.

2 Anwendungsfelder und Potenziale

Die Potenziale und Grenzen von Big Data für die makroökonomische Analyse werden entlang verschiedener Datenquellen analysiert. Dafür werden vor allem diejenigen Datenquellen betrachtet, die bislang in den Sozialwissenschaften am häufigsten verwendet worden sind. Zudem decken die Quellen die zentralen Kategorien – also soziale Medien sowie sensorgenerierte und prozessbasierte Daten – von Big Data ab, sodass sich viele der Eigenschaften auf andere Datenquellen derselben Kategorie übertragen lassen dürften. Für die Analyse soll die vorliegende Literatur ausgewertet werden und mit Anwendungserfahrungen mit Big Data sowie den Anforderungen und Bedürfnissen an Daten für makroökonomische Analysen in Verbindung gesetzt werden.

Konkret werden die Datenquellen gemäß vorgegebenen Kategorien in einem einheitlichen Rahmen diskutiert. Zu diesen Kategorien zählen die *Datenquellen* und die *Datenbeschaffenheit*, beispielsweise in Hinblick auf Zugangsbeschränkungen, Verfügbarkeit, Frequenz, Abdeckung oder Volumen der Daten. Ferner werden die relevanten *Methoden*, die zur Verarbeitung und zur Auswertung der jeweiligen Daten verwendet werden können, beschrieben und diskutiert. Im Rahmen eines umfangreichen *Literatur-*

überblicks werden die bislang wichtigsten Anwendungsfelder der jeweiligen Datenquellen und die dafür verwendeten Daten und Methoden beschrieben. Dabei werden auch Anwendungsfelder in Betracht gezogen, die nicht unmittelbar mit makroökonomischen Fragestellungen in Verbindung stehen, sofern sie Einsichten bezüglich der Potenziale und Grenzen der Datenquellen für die Konjunkturanalyse und -prognose erwarten lassen. Schließlich werden auf Basis der vorgehenden Aspekte *Potenziale und Grenzen* der Datenquellen für die makroökonomische Analyse mit einem besonderen Fokus auf die Konjunkturforschung für Deutschland beschrieben. Dabei sollen nicht nur die bereits vorliegenden Anwendungsfelder kritisch diskutiert und darauf aufbauend mögliche Weiterentwicklungen aufgezeigt werden, sondern auch mögliche neue Anwendungsfelder beleuchtet werden. Bei alldem werden nicht nur die Möglichkeiten, sondern auch Probleme und Hürden aus Anwendersicht dargestellt.

2.1 Nachrichten und Presseartikel

In den vergangenen Jahren haben Nachrichten und Presseartikel als Datenquelle und Gegenstand quantitativer Forschungsansätze erheblich an Bedeutung gewonnen. Sie beinhalten in vielerlei Hinsicht relevante Informationen, die auch für die makroökonomische Analyse nützlich sind. So bilden sie laufend wirtschaftliche Entwicklungen ab, noch bevor diese in den amtlichen Statistiken oder in den einschlägigen Frühindikatoren sichtbar werden. Ferner können ihre Inhalte das zukünftige Verhalten von wirtschaftlichen Akteuren beeinflussen. Schließlich können sie dazu dienen, relevante Einflussfaktoren zu quantifizieren, für die keine amtlichen Daten vorliegen, wenn es beispielsweise um wirtschaftspolitische Entwicklungen geht. In der makroökonomischen Analyse zählen dazu bislang insbesondere Indikatoren, die die wirtschaftspolitische Unsicherheit abbilden. Nachrichten und Presseartikel werden mit Methoden der computergestützten Textanalyse ausgewertet. In der einfachsten und derzeit am häufigsten angewendeten Form basiert die Analyse auf der Häufigkeit, mit der sich Nachrichten mit bestimmten Themen befassen. Grundsätzlich liegen aber auch Methoden vor, um komplexere inhaltliche Zusammenhänge von Texten auszuwerten. Bislang wurden in den sozialwissenschaftlichen Disziplinen vornehmlich Presseartikel ausgewertet. Allerdings können auch andere Nachrichten, beispielsweise von Unternehmen, relevante Informationsquellen darstellen.

2.1.1 Datenquellen

Viele der führenden Tageszeitungen weltweit haben mittlerweile ihre aktuellen und historischen Printausgaben digitalisiert und machen diese über Online-Archive zugänglich. Im englischsprachigen Raum verfügen beispielweise die Financial Times, der Guardian und die Times (alle Vereinigtes Königreich) oder USA Today, das Wall Street Journal, die New York Times, die Los Angeles Times, die Washington Post, die Chicago Tribune und der Boston Globe (alle Vereinigte Staaten) über digitale Archive. Auflagenstarke Tageszeitungen mit Online-Archiven in Kontinentaleuropa sind zum Beispiel die Süddeutsche Zeitung, die Frankfurter Allgemeine Zeitung und das Handelsblatt in Deutschland, Le Monde und Le Figaro in Frankreich, El Pais und El Mundo in Spanien oder der Corriere Della Sera und La Repubblica in Italien.[4] Auch viele wöchentliche oder zweiwöchentliche Printnachrichtenmagazine (z.B. Der Spiegel in Deutschland oder The Economist im Vereinigtem Königreich) und Sonntagszeitungen (z.B. The Observer im Vereinigten Königsreich) haben ihre Artikel mittlerweile digitalisiert.

[4] Weitere Beispiele sind der Kommersant (Russland), die Times of India (Indien), Folha (Brasilien), Reforma (Mexiko), El Mercurio (Chile), El Tiempo (Kolumbien), Yomiuri Shimbun (Japan), Doga Ilbo (Südkorea), South China Morning Post (Hongkong), Herald Sun (Australien) oder Globe and Mail (Kanada).

Die digitalisierten Artikel dieser Tageszeitungen und Magazine sind zumeist nicht nur über die Webseiten der Zeitungen selbst, sondern über kommerzielle Datenbankanbieter zugänglich, die die Archive von Zeitungen bündeln. Zu den wichtigsten dieser Metaarchive zählen Factiva, Lexis Nexis, Access World News (Newsbank) und ProQuest Historical Newspapers. Factiva umfasst derzeit über 30 000 Quellen aus 200 Ländern in mehr als 25 Sprachen, insbesondere Nachrichtenagenturen (z.B. Dow Jones und Reuters) und Zeitungen (z.B. Wall Street Journal). Der Schwerpunkt von Factiva liegt auf den Themen Wirtschaft und Finanzen. Lexis Nexis umfasst ca. 36 000 Informationsquellen, grob unterteilt in die Bereiche Recht (etwa 1/4 der Quellen) sowie Wirtschaftsnachrichten (etwa 3/4 der Quellen). Letztere umfassen Volltexte aus Tages- und Wochenzeitungen, Zeitschriften, Agenturmeldungen, Brancheninformationen und Marktübersichten. Access World News (Newsbank) bietet publizierte Inhalte von 2 000 Tageszeitungen, Geschäftszeitungen, Magazinen, Fernseh- und Radiosendern und auch Regierungen an. ProQuest Historical Newspapers hat den Fokus auf historischen Mikrofilm- sowie Onlinepublikationen (auch aus dem akademischen Bereich) und verfügt über ca. 9 000 Titel aus den letzten 500 Jahren (etwa 55 Millionen digitalisierte Seiten Text). In den Datenbanken findet sich meist neben Zeitungsartikeln noch eine Vielzahl weiterer Nachrichten- bzw. Agenturmeldungen, Mitschriften von Fernseh- und Radiosendungen, Web- und Bloginhalten sowie Unternehmensnachrichten. Solche Unternehmensinformationen, insbesondere Finanzberichte, werden von spezialisierten Anbietern, wie z.B. Refinitiv EIKON, online zur Verfügung gestellt. Gebündelt sind sie dort kostenpflichtig abrufbar.

Auch die Online-Archive der Tageszeitungen und der kommerziellen Datenbankanbieter sind zum größten Teil kostenpflichtig. Viele Nationalbibliotheken, Universitätsbibliotheken, Bibliotheksverbünde und andere wissenschaftliche Institutionen erwerben aber mittlerweile entsprechende Lizenzen bei den Verlagen und den Drittanbietern. So können individuelle Forscher und Forscherinnen in der Regel nach Anmeldung bei einer dieser Bibliotheken oder Bibliotheksnetzwerke Zugang auf die Archive der individuellen Zeitungen und auch auf die Datenbanken der kommerziellen Metaarchive erhalten.

Mittlerweile werden aus Presseartikeln abgeleitete wirtschaftliche Indikatoren öffentlich bereitgestellt und regelmäßig aktualisiert. Dazu zählen insbesondere Indikatoren zur wirtschaftspolitischen Unsicherheit für mehr als 20 der größten Volkswirtschaften (policyuncertainty.com) sowie Indikatoren zur handelspolitischen Unsicherheit (Trade Policy Uncertainty Index).

2.1.2 Datenbeschaffenheit

Die digitalen Archive der meisten Zeitungen reichen ca. 30 Jahre (in Deutschland z.B. Handelsblatt seit 1986) bis 70 Jahre (in Deutschland z.B. Süddeutsche Zeitung seit 1945 und Frankfurter Allgemeine Zeitung seit 1949 oder in Frankreich Le Monde seit 1944) zurück. Insbesondere im englisch- und französischsprachigen Raum und gerade bei wöchentlichen Formaten bieten einige Online-Archive aber auch Zugang zu Artikeln, die teilweise vor über 100 oder sogar 200 Jahren verfasst wurden (z.B. The Observer seit 1791, The Guardian seit 1821, Le Figaro seit 1826, New York Times seit 1851, Boston Globe seit 1872, Los Angeles Times seit 1881, Wall Street Journal seit 1889).

Die Archive weisen in der Regel aufgrund ihres langen Verfügbarkeitszeitraums und der hohen Frequenz ein immenses Volumen auf (die Frankfurter Allgemeine Zeitung beispielsweise beziffert ihre Gesamtartikelanzahl der vergangenen 25 Jahre auf mehr als 2,7 Millionen). Die Datenqualität ist dabei recht hoch. So enthalten die Texte nur wenige Rechtsschreib- oder Grammatikfehler, die die Textanalyse erschweren, und die Archive verzeichnen in der Regel kaum Lücken. Allerdings sind aktuellere Jahrgänge insbesondere bei einem Zugang über kommerzielle Datenbankanbieter oder über die Kataloge von Universitätsbibliotheken teilweise nicht verfügbar. Die Verzögerungen aufgrund

vertragsrechtlicher Vereinbarungen mit den Verlagen betragen je nach Anbieter typischerweise zwei bis fünf Jahre. Beim Erwerb eines Direktzugangs auf die einzelnen Zeitungsarchive tritt dieses Problem seltener auf. Allerdings müssen hier die Konditionen für die regelmäßige Nutzung von tagesaktuellen Informationen, die beispielsweise für eine Verstetigung von Konjunkturanalysen notwendig ist, ggf. individuell vereinbart werden.

Die meisten Zeitungs-Onlinearchive erlauben neben der Identifikation von relevanten Artikeln mittels der Wörter im Titel, der Autorennamen oder der mit dem Artikel verbundenen Schlagwörter auch die Anzeige aller archivierten Artikel mit dem gesuchten Begriff in ihrem Volltext. So können für die Erstellung einfacher Indizes Zeitreihen über die Anzahl der Artikel mit relevanten Begriffen direkt aus den Archiven erstellt werden. Komplexere Datenreihen, etwa zu Begriffsanzahlen innerhalb von Artikeln oder zu Wortkombinationen, sind aufwendiger zu generieren. Es ist hier oft weitere Software (z.B. Python oder R) nötig, um die Daten in das gewünschte Format zu bringen. Die Datenauswertung ist auch dann zeitintensiv, wenn Daten aus unterschiedlichen Quellen kombiniert werden sollen. Die Suchfilter und die Ergebnisanzeige der Archive sind sehr heterogen und unterscheiden sich teilweise auch deutlich in ihrer Qualität. Der Download der Suchergebnisse bzw. Daten erfolgt – anders als bei modernen quantitativen oder rein webtextorientierten Datenbanken – mit wenigen Ausnahmen (z.B. New York Times) nicht über Programmierschnittstellen (Application Programming Interface, API) oder Massendownload, sondern meist herkömmlich. Oft ist es somit nicht möglich, die Daten direkt in eine externe Software einzulesen, und es müssen separate Downloads für die einzelnen Suchanfragen vorgenommen werden.

Kommerzielle Drittanbieter hingegen ermöglichen nicht nur die Textstellensuche in den Archiven mehrerer Zeitungen simultan, sondern verfügen auch über für Analysezwecke wesentlich anwendungsfreundlichere Oberflächen, etwa Suchfilter und Ergebnisanzeigen sowie vereinfachte Download-Verfahren. Textdateien (z.B. HTML oder PDF) aus verschiedenen Quellen können beispielsweise bei Factiva, LexisNexis oder ProQuest über eigene APIs gezielt, einheitlich formatiert und gebündelt heruntergeladen werden. Diese Metadatenbanken erlauben somit, die verfügbaren Textdaten über diverse Quellen und Herkunftsländer der Quellen hinweg zu erheben, zu selektieren und auszuwerten. Allerdings erfordern internationale Vergleichsstudien aufgrund der unterschiedlichen Sprachen ggf. einen besonders hohen Aufwand und sind daher bislang eher selten anzutreffen.[5]

Aus Presseartikeln abgeleitete Indikatoren zur wirtschaftspolitischen Unsicherheit liegen für viele Volkswirtschaften zumindest seit Anfang der 1990er Jahre vor; ausgewählte Indikatoren, beispielsweise für die Vereinigten Staaten, sind aber auch für einen deutlich längeren Zeitraum verfügbar. Die Indikatoren werden in der Regel auf monatlicher Basis erstellt und zeitnah nach Ablauf eines Monats aktualisiert. Für die Vereinigten Staaten wird ein Indikator zur wirtschaftspolitischen Unsicherheit auf Tagesdatenbasis bereitgestellt.

2.1.3 Methodik

Für die empirische Auswertung von Nachrichten und Presseartikeln müssen die vorhandenen Texte mit quantitativen Variablen in Verbindung gesetzt werden. Aufgrund ihrer inhärent hohen Dimension werden die Texte deshalb häufig vorab präpariert, um ihre Kerninformationen leichter erfassbar zu machen. Für einfachere Anwendungen wie der Messung von Frequenzen, mit denen bestimmte Wörter

[5] Eine Ausnahme stellen die länderspezifischen Indikatoren zur wirtschaftspolitischen Unsicherheit gemäß Baker et al. (2016) dar. Sie basieren allerdings auf wenigen und relativ einfachen Wortkombinationen.

verwendet werden oder Artikel sich mit bestimmten Themen befassen, sind diese vorgelagerten Bearbeitungsschritte freilich nicht immer notwendig.

2.1.3.1 Text als Daten

Ein großer Unterschied von Texten im Vergleich zu anderen Datentypen, die für gewöhnlich in den Wirtschaftswissenschaften genutzt werden, ist ihre hohe inhärente Dimension (Gentzkow et al. 2019; Kelly et al. 2019). So hat die eindeutige Repräsentation eines Textdokuments mit einer Anzahl von je w Wörtern aus einem Wortschatz von p möglichen Wörtern die Dimension p^w. Weil dieselben Phänomene mit unterschiedlichen Wörtern beschrieben werden können, sind Texte überdies wesentlich komplexer als die gängigen quantitativen Daten. Aus diesem Grund sind statistische Methoden zur Analyse von Texten oft eng verwandt mit Methoden zur Analyse von großen und komplexen Datensätzen in anderen Bereichen, wie zum Beispiel dem maschinellen Lernen (Gentzkow et al. 2019).[6]

Für Anwendungen in den Sozialwissenschaften wird die hohe Dimensionalität von Texten zunächst auf eine erfassbare Höchstzahl von Eigenschaften (Features) reduziert (Feature Selection), um sie als numerisches Datenfeld (Matrix) abbilden und statistisch untersuchen zu können. Dafür werden sehr häufig auftretende Wörter ohne inhaltliche Relevanz (sogenannte Stoppwörter wie „ein“ oder „an“) entfernt. Zudem werden auch selten verwendete Wörter sowie Zahlen und Satzzeichen aus dem Text entfernt. Ein rein maschinelles Entfernen dieser Textbausteine kann allerdings problematisch sein, sofern diese für den Untersuchungsgegenstand relevante Informationen enthalten (Athey und Imbens 2019; Kelly et al. 2019). Die Identifikation von besonders informativen Wörtern kann mittels der sogenannten Term Frequency – Inverse Document Frequency (TF-IDF) erfolgen. Diese gewichtet solche Wörter am stärksten, die sehr oft in einem Textdokument, aber nur selten in den anderen verwendeten Textdokumenten vorkommen.

Ferner können die Eigenschaften der Texte auch verändert oder neu erstellt werden (Feature Engineering). Dazu werden typischerweise verschiedene morphologische Wortvarianten mit ihrem gemeinsamen Wortstamm ersetzt (Stemming; siehe z.B. Porter 1980). Die einfachste Repräsentation von Text als Daten ist dann eine Matrix, wobei die Zeilen die einzelnen Textdokumente und die Spalten die einzelnen Wörter des Gesamtvokabulars nach Stemming kennzeichnen. Jede Zelle enthält dann die Häufigkeit eines Wortes in einem Textdokument. Diese Bag-Of-Words (BOW)-Methode basierend auf einzelnen Wörtern ignoriert allerdings die Reihenfolge und die Abhängigkeit der Wörter untereinander und bildet somit den Inhalt der Texte oft nur unzureichend ab. Eine etablierte Methode ist es, das Gesamtvokabular nicht in einzelne Wörter, sondern in Wortfolgen bzw. Wortkombinationen zu unterteilen. Aufgrund praktischer Grenzen werden hier meist lediglich die jeweils direkt benachbarten Wörter (überlappend) miteinander kombiniert (Bigramme), um den Sinn des Textes genauer zu erfassen (z.B. „schwere Rezession“ versus „leichte Rezession“).

Problematisch ist, dass für diese vorbereitenden Schritte zur Analyse von Nachrichten oder Presseartikeln wenig etablierte Regeln vorliegen, sie aber Einfluss auf die Analyseergebnisse haben können. So können die Ergebnisse je nach Auswahl der zugrunde liegenden Quellen oder angewendeten Methoden zur Vereinfachung der Texte variieren. Deshalb werden diese Schritte in der Regel von Sensitivitätsanalysen begleitet, um Informationsverluste zu vermeiden und nicht bereits strukturelle Fehler in dem Datensatz zu generieren, auf dessen Basis anschließend empirische Analysen durchgeführt werden (vgl. Athey und Imbens 2019; Gentzkow et al. 2019). Generell empfiehlt sich bei allen

[6] Zur steigenden Bedeutung von Verfahren aus dem Bereich des maschinellen Lernens für die empirische Wirtschaftsforschung generell siehe Athey und Imbens (2019).

Schritten der Textanalyse die Kombination von computergestützten und manuellen bzw. händischen Verfahren, um wechselseitig methodische Fehler zu identifizieren und Annahmen zu überprüfen (vgl. z.B. das „Human Audit“ in Baker et al. 2016).

2.1.3.2 Analysemethoden

Davon ausgehend, dass die (wie oben beschrieben) als Daten ausgedrückten Texte informativ z.B. über Größen wie wirtschaftspolitische Unsicherheit, das Auftreten von Grippefällen, die Arbeitslosenquote oder die Bewertung von Aktien sind, kann man die generierten Datensätze im Hinblick auf Frequenzen und Verteilungen der Wörter empirisch analysieren, um Aussagen über derartige Größen abzuleiten. Dafür können je nach Fragestellung unterschiedliche Methoden verwendet werden, die sich in ihrer Komplexität unterscheiden.

(1) Wörterbuch-basierte Ansätze sind mit Abstand am populärsten und am einfachsten umsetzbar. Sie finden bereits seit langem bei der Einschätzung von Aktienmarktentwicklungen Anwendung (Cowles 1933; Tetlock 2007; Bollen et al. 2011; Loughran und McDonald 2011; Wisniewski und Lambe 2013), haben aber in den letzten Jahren insbesondere bei der Messung wirtschaftspolitischer (Baker et al. 2016), geldpolitischer (Husted et al. 2017) und jüngst auch handelspolitischer (Arbatli et al. 2019; Caldara et al. 2020; Davis et al. 2019) Unsicherheit Bedeutung erlangt. Wörterbuch-basierte Ansätze abstrahieren in der Regel von jeglicher Inferenzstatistik. Ihr Ziel ist es, basierend auf der Häufigkeit bestimmter Wörter oder Wortkombinationen in ausgewählten Quellen zu verschiedenen Zeitpunkten Indizes für eine ansonsten schwer messbare Variable von Interesse (z.B. politische Unsicherheit) zu bilden. Diese Indizes können dann in weiterführenden Analysen eingesetzt werden. Die Variable von Interesse (z.B. Einschätzung der Börsenentwicklung in den Artikeln des Wall Street Journals) wird hier entlang mehrerer Dimensionen definiert (z.B. positive Einschätzung, negative Einschätzung). Diese Dimensionen werden jeweils durch eine Gruppe bestimmter Wörter oder Wortkategorien aus dem Gesamtvokabular beschrieben, das sogenannte Wörterbuch (z.B. „Aufschwung“, „optimistisch“ etc. für eine positive Einschätzung versus „Abschwung“, „pessimistisch“ etc. für eine negative Einschätzung). Die Anzahl der Artikel mit bestimmten Wörtern (oder die Wortfrequenzen im gesamten Textkorpus) in einem vordefinierten Zeitraum (z.B. Monat) kann dann Aufschluss über die typische Einschätzung der Börsenentwicklung in diesem Zeitraum geben. In der Regel werden die Frequenzen mit der Anzahl der ausgewerteten Wörter oder Artikel normiert, um zu vermeiden, dass Ergebnisse mit dem Umfang der ausgewerteten Texte schwanken. Endogenitätsprobleme können sich ergeben, wenn bestimmte Begriffe in manchen Perioden oder Quellen *per se* überrepräsentiert sind oder Medienberichte sich wiederum auf Indikatoren beziehen, die aus Presseartikeln abgeleitet sind.

(2) Bei Textregressions-Ansätzen wird mittels Regressionsanalyse der Zusammenhang zwischen den Variationen bestimmter Wörter in Texten und der Entwicklung anderer Variablen geschätzt. Die Frage wäre hier dann nicht mehr nur, wie bei den Wörterbuch-basierten Ansätzen, „Wird die Börsenentwicklung im Wall Street Journal (WSJ) positiv oder negativ eingeschätzt?“, sondern vielmehr „Sagt die (negative/positive) Einschätzung der Börsenentwicklung im WSJ die *tatsächliche* (negative/positive) Entwicklung der Börse voraus?“. Solche Zusammenhänge können prinzipiell durch jede Regressionstechnik geschätzt werden, wobei in der Praxis angesichts der hohen Dimensionalität der Daten oft LASSO- oder Ridge-regularisierte lineare oder logistische Regressionsmodelle verwendet werden, deren Berechnungen unter anderem darauf basieren, dass sie (ähnlich der TF-IDF) relevantere Wörter (bzw. Features) bei den Berechnungen systematisch höher gewichten können als andere (siehe auch Gentzkow et al. 2019; Athey und Imbens 2019). Genauer werden in diesen Modellen Koeffizienten von nicht oder nur wenig aussagekräftigen Regressoren (Wörter im Gesamtvokabular) nahe oder gleich null gesetzt. Der Datensatz bzw. das Gesamtvokabular schrumpft sukzessive, bis eine verwertbare

Dimension erreicht ist.[7] Mit Textregressionsansätzen kann umgekehrt auch die Wirkung von Variablen (z.B. Aktienpreise) auf die Attribute von Texten (z.B. Frequenz bestimmter Wörter in Artikeln über den Aktienmarkt) geschätzt werden. Für makroökonomische Analysen spielen diese Ansätze bislang, anders als z.B. in den Politikwissenschaften, eher eine untergeordnete Rolle. Mit ihnen wurde bislang beispielsweise die Autorenschaft (siehe z.B. Stock und Trebbi 2003 zum Urheber des Instrumentenvariablenschätzers) und Subjektivität (siehe z.B. Greenstein et al. 2016 zu Wikipedia-Beiträgen) wissenschaftlicher oder wissenschaftsrelevanter Texte untersucht.

(3) Für Textanalysen werden auch zunehmend komplexere Methoden aus dem Bereich des maschinellen Lernens eingesetzt. Es gibt im Bereich der Computerlinguistik bereits wesentlich komplexere statistische Modelle für eine gehaltvollere Repräsentation von Wörtern als Zahlen, die teilweise auch in der Ökonomie Anwendung finden. Die Verfahren beziehen sich weniger auf die Messung von Effekten selbst, sondern auf die Darstellung und thematische Einordnung (bis hin zu Interpretation) der Textdatensätze. Maschinelles Lernen betrifft also insbesondere die Bereiche der Textselektion und die Auswahl der Wörterbücher. Es können beispielweise mittels der Latent Dirichlet Allocation (LDA) große Textmengen auf bestimmte Trends und Themen reduziert werden. Ein Anwendungsbeispiel für LDA-basiertes Topic Modeling ist der von Müller et al. (2018) entwickelte Uncertainty Perception Indicator (UPI) für Deutschland. Hier werden bestimmte Muster der Berichterstattung über wirtschaftliche Unsicherheit durch die Kategorisierung von Artikeln in thematische Cluster maschinell identifiziert. Dieses automatisierte Verfahren geht deutlich über Ansätze hinaus, bei denen lediglich bestimmte Suchwörter und Themen vorgegeben werden (Müller 2020). Auch ist es mit dem sogenannten Word Embedding bereits möglich, durch Algorithmen auf Basis der Satzsyntax und der Verteilung der Wörter im Text bedeutungsähnliche Wörter zu identifizieren und zu gruppieren (Goldberg und Orwant 2013) und so den Text letztendlich als eine Kombination mathematischer Vektoren abzubilden.[8] Bekannte Verfahren sind sogenannte (Unsupervised) Learning Word Vectors wie z.B. Word2Vec (Mikolov et al. 2013) und GloVe (Pennington et al. 2014). Solche Modelle können die innere Struktur des unterliegenden Textes mitsamt Abhängigkeiten und Bezügen der Sätze bzw. Wörter untereinander (Subtext) fast vollständig erfassen und in Zahlenform ausdrücken, sind aber sehr aufwendig und erfordern viel Programmier- und Rechenleistung. Sie werden daher in den Sozialwissenschaften bisher noch selten genutzt (siehe z.B. Bolukbasi et al. 2016), bieten aber angesichts ihres erfolgreichen Einsatzes in praktischen Anwendungsbereichen der Computerlinguistik (z.B. Spracherkennung und semantische Sprachmodelle, Texterkennung, Übersetzung) Potenziale, um wirtschafts- und sozialwissenschaftliche Fragestellungen mittels Textanalysen besser adressieren zu können.

2.1.4 Bisherige Anwendungen

Es liegen bereits zahlreiche Studien vor, die Textanalysen von Nachrichten und Presseartikel verwenden, um makroökonomische Analysen durchzuführen. Sie werden aber auch regelmäßig in anderen sozialwissenschaftlichen Disziplinen angewendet.

In einer sehr einflussreichen wirtschaftswissenschaftlichen Studie nutzen Baker et al. (2016) Textanalysen von Presseartikeln, um Indikatoren zur wirtschaftspolitischen Unsicherheit (Economic Policy Uncertainty-Index, EPU) zu bilden. Die Idee ist, dass Unsicherheit in Bezug auf die heutige und zukünftige Politik der Regierung Risiken erhöhen und beispielsweise Investitionen verringern kann. Die Autoren

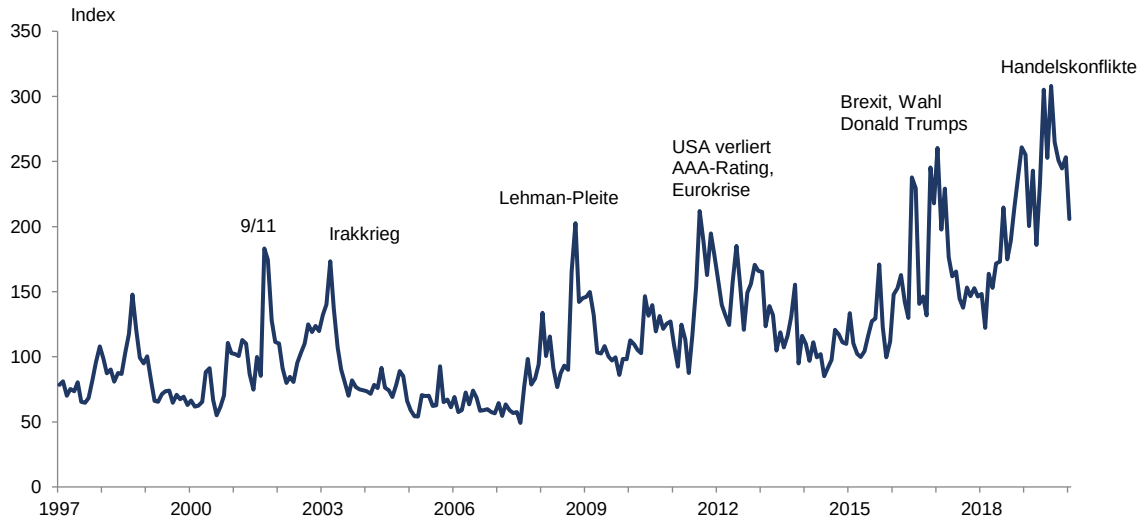
[7] Es ist hierbei allerdings darauf zu achten, dass die Modellparameter nicht so restriktiv sind, dass unter Umständen relevante Faktoren getilgt werden.

[8] Einfache mögliche Kombinationen von Vektoren wären zum Beispiel KÖNIG - MANN + FRAU = KÖNIGIN oder auch BERLIN - DEUTSCHLAND + FRANKREICH = PARIS.

nutzen Texte führender Zeitungen in den Vereinigten Staaten (und anderer Länder, wie Deutschland, Frankreich oder dem Vereinigten Königreich) in meist monatlicher Frequenz, um wirtschaftspolitische Unsicherheit zu messen. Dazu wird die durchschnittliche und normalisierte monatliche Anzahl der Artikel, die jeweils vorgegebene Wörter aus den drei Themenbereichen Volkswirtschaft, Wirtschaftspolitik und der Unsicherheit enthalten, erfasst. Die so gebildeten Indikatoren weisen in der Regel große Ausschläge bei wirtschaftspolitisch bedeutsamen Ereignissen auf. So ist der Indikator für die globale wirtschaftspolitische Unsicherheit beispielsweise nach den Anschlägen vom 11. September 2001 oder dem Brexit-Votum deutlich gestiegen (Abbildung 3). Die Autoren zeigen, dass höhere Unsicherheit mit geringerer Beschäftigung und niedrigeren Investitionen sowohl bei US-Firmen als auch gesamtwirtschaftlich einhergeht. Ferner wurde gezeigt, dass eine höhere globale wirtschaftspolitische Unsicherheit zu einem Rückgang der weltweiten (Ademmer et al. 2019a) und der deutschen (Ademmer und Jannsen 2019) Industrieproduktion führt.

Die Idee von Baker et al. (2016) wurde mittlerweile auf andere Länder, andere Dimensionen von Unsicherheit und auch andere Methoden ausgeweitet. Eine ähnliche Studie mit einem Fokus auf handelspolitische Unsicherheit ist Caldara et al. (2020). Angelehnt an den EPU wird hier ein Trade Policy Uncertainty Index (TPU) gebildet, wobei unter anderem die Berichterstattung in Zeitungsmeldungen als Grundlage dient (neben anderen Variablen wie z.B. Zolldaten). Die Autoren finden, dass ein Anstieg der handelspolitischen Unsicherheit die Investitionen und die wirtschaftliche Aktivität in den Vereinigten Staaten hemmen. Nach ähnlichem Muster wurden bereits Indizes zur Handelsunsicherheit in China (Davis et al. 2019) und Japan (Arbatli et al. 2019) erstellt. Die Idee wurde auch auf geldpolitische (Husted et al. 2017; Arbatli et al. 2019), geopolitische (Caldara und Iacoviello 2019) und migrationspolitische (Baker et al. 2015) Unsicherheit übertragen. In ähnlicher Weise erstellt Azzimonti (2017) einen Indikator zur politischen Polarisierung in den Vereinigten Staaten (Partisan Conflict Index, PCI) aus Artikeln führender US-Zeitschriften für die Jahre von 1891 bis 2017. Die Studie zeigt, dass eine höhere Polarisierung mit einer signifikant geringeren Investitionstätigkeit einhergeht. Mittlerweile werden auch

Abbildung 3:
Globale wirtschaftspolitische Unsicherheit 1997–2020[a]



[a]Monatsdaten; Gewichtung auf Basis von Marktwechselkursen.

Quelle: Economic Policy Uncertainty (2020).

methodisch anspruchsvollere Verfahren verwendet. So entwickeln Müller et al. (2018) basierend auf denselben Suchanfragen wie der deutschen EPU gemäß Baker et al. (2016) mittels des Topic-Modelling-Verfahrens LDA einen multidimensionalen Unsicherheitsindikator, der eine wesentlich detailliertere Analyse der der wirtschaftspolitischen Unsicherheit zugrunde liegenden Einflussfaktoren erlaubt. Azqueta-Gavaldon et al. (2020) verwenden ebenfalls einen solch LDA-basierten Ansatz, um Indikatoren für die wirtschaftspolitische Unsicherheit und ihrer jeweils wichtigsten Komponenten in einem einheitlichen Modellrahmen für die vier größten Volkswirtschaften des Euroraums zu schätzen.

Für die Prognose von makroökonomischen Variablen wurden Presseartikel oder andere Nachrichten bislang kaum verwendet. Eine Ausnahme stellt die Studie von Kelly et al. (2019) dar, in der Zeitungsmeldungen zur Prognose von makroökonomischen Variablen in den Vereinigten Staaten herangezogen werden. Konkret nutzen die Autoren einen Korpus, der aus Nachrichten des WSJ im Zeitraum 1990 bis 2017 besteht. Für eine empirische Analyse handhabbar gemacht werden die Textdokumente, indem die Häufigkeit von Wortpaaren (Bigramme) erfasst wird. Mittels eines Modells werden Zeitreihen gebildet, die zwischen der bloßen Nennung eines Bigramms und der Häufigkeit, mit der es in einem Dokument auftritt, unterscheiden. Diese stellen die Basis für die Prognose gesamtwirtschaftlicher Größen, wie der Beschäftigung, der Industrieproduktion oder der Bauaktivität, dar. Modelle, die die aus den Textdokumenten gebildeten Zeitreihen als Regressoren enthalten, schneiden in einer Prognoseevaluierung besser ab als jene, die ausschließlich auf einem konventionellen makroökonomischen Datensatz beruhen. Dies gilt bereits für einen Prognosehorizont von einem Monat; besonders deutlich verringern sich die Prognosefehler allerdings bei einem Horizont von einem Jahr. Auch für den Nowcast, d.h., die Prognose der obengenannten Variablen in Monat *t* unter Verwendung der in diesem Monat veröffentlichten Zeitungsartikel, schneiden die nachrichtenbasierten Modelle insgesamt recht gut ab. Allerdings fallen die Verbesserungen in der Prognosegüte nicht so deutlich aus wie bei längeren Prognosehorizonten; in einigen Fällen schneiden sie sogar schlechter ab. Thorsrud (2020) nutzt das LDA-Verfahren, um Artikel einer norwegischen Tageszeitung auszuwerten und für die statistische Analyse greifbar zu machen. Durch LDA wird der Korpus, der etwa eine halbe Million Meldungen umfasst, durch 80 Themenkomplexe beschrieben, die jeweils als eine Wahrscheinlichkeitsverteilung über alle in dem Korpus vorkommenden Worte zu verstehen sind. So weist beispielsweise ein Themenkomplex mit hoher Wahrscheinlichkeit Wörter wie Immobilien, Wohnungen oder Quadratmeter auf (Themenkomplex „Immobilienmarkt“); ein anderer enthält eher Wörter wie Zinsen, Inflation oder Zentralbank (Themenkomplex „Geldpolitik“). Diese könnten daher als Immobilienmarkt- oder Geldpolitik-Topic interpretiert werden. Für die Prognosen in dieser Studie ist diese Interpretation aber nicht von Bedeutung. Hierfür werden nur die in den an einem Tag erschienenen Artikeln verwendeten Themenkomplexe gezählt sowie durch eine einfache Saldierung der positiv und negativ konnotierten Wörter (laut Harvard IV-4 Psychological Dictionary) die Tonlage („sentiment“) eines Artikels erfasst. Des Weiteren werden die Zeitreihen vor der Schätzung mittels eines gleitenden Durchschnitts bereinigt. Auf Basis dieser Zeitreihen in Kombination mit der Zuwachsrate des norwegischen Bruttoinlandsprodukts wird dann ein Faktormodell geschätzt. Diese Modellklasse eignet sich für die Prognose besonders gut, da sie mit vielen Zeitreihen umgehen kann und verschiedene Frequenzen (in dieser Anwendung Tages- und Quartalsdaten) berücksichtigen kann. In einer Pseudo-Echtzeitevaluierung schneidet das LDA-Faktormodell deutlich besser ab als naive Verfahren und liefert ähnliche Ergebnisse wie komplexere Prognosekombinationsverfahren.

Traditionell verwendete Klassifikationen von Wirtschaftszweigen können gerade in Zeiten des wirtschaftlichen Wandels ökonomisch relevante Aktivitäten nur unzureichend unterscheiden. Hoberg und Phillips (2015) klassifizieren US-amerikanische Firmen deshalb basierend auf einer Textanalyse der Produktbeschreibungen in den jeweiligen offiziellen Unternehmensangaben in Wirtschaftszweige. Dies bietet mehr Flexibilität bei der Zuordnung und Analyse von Industriezweigen, wenn sich die Struktur von Firmen im Zuge der fortschreitenden Digitalisierung rasch ändert. In einem ähnlichen Ansatz

messen Kelly et al. (2018) mittels Textanalyse die Ähnlichkeit zwischen (neuen und alten) Patentdokumenten, um neue Indikatoren für die Patentqualität zu bilden. Atalay et al. (2017) analysieren Stellenanzeigen mittels Textanalyse, um strukturelle Veränderungen der Arbeitsinhalte am US-Arbeitsmarkt zwischen den Jahren 1960 und 2000 zu identifizieren. Jegadeesh und Wu (2013) nutzen Textregressionen, um die Reaktion von firmenspezifischen Aktienrenditen auf Veränderungen in den Textinformationen im Jahresbericht der jeweiligen Firma zu schätzen. In einer aktuellen Studie werten Hassan et al. (2020) die Mitschriften von Telefonkonferenzen zur Geschäftsentwicklung mittels Verfahren der computergestützten Textanalyse aus, um den Einfluss des Brexit auf Erwartungen und reale Geschäftszahlen (Kosten, Investitionen und Einstellungen) amerikanischer und internationaler Firmen aufzuzeigen. In einem ähnlichen Ansatz (Hassan et al. 2019) werden die protokollierten Telefonkonferenzen hinsichtlich der Intensität, mit der sie sich mit politischen Risiken generell (über einzelne Ereignisse wie den Brexit hinaus) befassen, ausgewertet. Firmen, die sich höheren Risiken ausgesetzt sehen, drosseln ihre Beschäftigung und ihre Investitionen.

Bereits Cowles (1933) hat Nachrichtentexte des WSJ von 1902 bis 1929 genutzt, um Renditen des Dow Jones Industrial Average Index vorherzusagen. In jüngerer Zeit erschienen zahlreiche Studien mit vergleichbaren Ansätzen. Tetlock (2007) nutzt ebenfalls Wörterbuch-basierte Textanalyse einer WSJ-Kolumne für die Erstellung eines „Medien-Pessimismus-Faktors“ bezüglich der Entwicklung des Dow Jones. Er findet transitorische negative Effekte einer Erhöhung dieses Faktors auf die zukünftigen Renditen. Ähnliche Wörterbuch-basierte Studien finden sich in Loughran und McDonald (2011), Bollen et al. (2011) und Wisniewski und Lambe (2013). Manela und Moreira (2015) konstruieren mittels Regressionsanalysen einen nachrichtenbasierten Marktvolatilitätsindex anhand von Texten des WSJ (1890–2009). Sie wenden regularisierte Regressionen an, die es erlauben, Subgruppen von aussagekräftigen Wörtern zu identifizieren, deren Frequenzen dann mit Daten über Turbulenzen an den Finanzmärkten in Verbindung gesetzt werden. Antweiler und Frank (2004) testen, wie informativ Internetpostings von Börsenexperten zu Aktien für die Kursentwicklung sind. Sie klassifizieren automatisiert 1,5 Millionen solcher Internetpostings in Kauf-, Verkauf- und Haltesignale und aggregieren diese Klassifikationen zu einem Index, anhand dessen dann Aktienrenditen vorausgesagt werden können.

Lucca und Trebbi (2011) nutzen Statements des Federal Open Market Committee (Offenmarktausschuss, FOMC) der US-Zentralbank (Fed), um Fluktuationen in Staatsanleihen und Schatzbriefen vorauszusagen. Sie quantifizieren in einem Wörterbuch-basierten Ansatz anhand von Factiva, ob einzelne Wörter in einem Statement und letztendlich das Statement als Ganzes eine positive oder negative Konnotation haben, und wie hoch die Intensität dieser Konnotation ist. Die resultierenden Maßzahlen werden dann als erklärende Variablen in Vektorautoregressive Modelle (VAR-Modelle) eingesetzt. Die Autoren stellen fest, dass Veränderungen im Inhalt der FOMC-Statements eher Veränderungen in der Bewertung einzelner Anleihen erklären, als unerwartete Änderungen in der Federal Funds Rate (Leitzins) dies tun. Ähnliche Ansätze finden sich in Born et al. (2014) mit dem Fokus auf die Auswirkung von aktuellen Reden der Führungskräfte von Zentralbanken und neu erschienenen Finanzstabilitätsberichten auf Renditen und Volatilität der Aktienmärkte sowie in Hansen et al. (2014) mit dem Fokus auf die Effekte von gestiegener Transparenz bei der Protokollierung von FOMC-Meetings auf die Informationseffizienz der Zentralbank.

In den Sozialwissenschaften kommt der Identifikation und Messung von „News Bias“ eine immer wichtigere Bedeutung bei. Dies gilt auch im Hinblick auf Wirtschaftsnachrichten (Larcinese et al. 2011). Führende Medien und insbesondere die großen Tageszeitungen sind seit jeher auch einflussreiche politische Akteure. Ihre Berichterstattung über Politikmaßnahmen und Wirtschaftspolitik prägt maßgeblich die öffentliche Meinung und kann dabei subjektiv bzw. politisch gefärbt sein. Mittels

Textanalysen kann diese Subjektivität identifiziert und quantifiziert werden. Groseclose und Milyo (2005) oder Gentzkow und Shapiro (2010) beispielweise sind einflussreiche Studien, die die ideologische Ausrichtung von bekannten US-Medien, insbesondere Tageszeitungen, mittels Textanalysen analysieren. Die von Politologen konzipierten Computerprogramme Wordfish (für R, siehe Slapin and Proksch 2008) oder Wordscores (für Stata, siehe Laver et al. 2003) erlauben es, politische Positionen anhand von Textdokumenten abzubilden. Kelly et al. (2019) stellen anhand automatisierter Analysen von politischen Reden fest, dass Politiker früher deutlich häufiger bestimmte Phrasen wiederholt haben, um ihre Parteizugehörigkeit unter Beweis zu stellen. Ferner wurde von Greenstein et al. (2016) auch der Grad an Subjektivität in Artikeln des Online-Lexikons Wikipedia mit computergestützten Textanalysen bewertet.

2.1.5 Potenziale und Grenzen

Presseartikel und andere Nachrichten bieten grundsätzlich einen großen Fundus an makroökonomisch relevanten Informationen. Sie werden daher bereits regelmäßig für makroökonomische Analysen eingesetzt und sind insbesondere etabliert, um wirtschaftspolitische Unsicherheit als Alternative zu anderen vorliegenden Unsicherheitsmaßen abzubilden. In diesem Zusammenhang fließen sie bereits regelmäßig in laufende Konjunkturanalysen ein.

Zusätzliche Potenziale bieten sie für die quantitative Prognose von makroökonomischen Variablen. Da sie auch für die zukünftige Entwicklung relevante Informationen abbilden, könnten sie nicht nur für die Prognose der laufenden (Nowcast), sondern auch der zukünftigen Entwicklung nützlich sein. Neben dem Bruttoinlandsprodukt könnten sie dabei auch für die Prognose von disaggregierten Größen, für die zum Teil weniger geeignete Frühindikatoren vorliegen, wie den Unternehmensinvestitionen, den Bauinvestitionen oder dem privaten Konsum relevant sein. Ferner können Presseartikel möglicherweise auch zur frühzeitigen Identifikation von konjunkturellen Wendepunkten beitragen. Hier liegen zwar schon umfangreiche und technisch aufwendige Modelle auf Basis der konventionellen Frühindikatoren vor; die Prognoseunsicherheit ist hier allerdings nach wie vor hoch und die Prognosegüte reicht meist bestenfalls nur einige Monate in die Zukunft. Einen Beitrag zu alledem könnten auch komplexere Methoden für die Auswertung von Texten leisten, die bislang für makroökonomische Fragestellungen noch kaum angewendet worden sind. Bislang liegen allerdings kaum Ergebnisse zur Prognosegüte von Presseartikeln oder Nachrichten vor. Deshalb sind weitere Studien notwendig, um ihren konkreten Nutzen in diesen Bereichen besser abschätzen zu können, auch im Vergleich zu den vorliegenden konventionellen Frühindikatoren.

Neben Presseartikeln stellen auch Unternehmensinformationen (z.B. Geschäftsberichte) eine interessante Informationsquelle für makroökonomische Analysen dar. Sie können helfen, die Auswirkungen von wirtschaftspolitischer Unsicherheit oder andere Einflussfaktoren auf Firmen zu identifizieren und ihr zukünftiges Verhalten vorherzusagen. Dies kann ggf. Prognosen verbessern und zur Evaluation von wirtschaftspolitischen Maßnahmen beitragen. Auch eine disaggregierte Analyse, z.B. nach Wirtschaftszweigen, wäre hier denkbar. Allerdings werden Finanzberichte erst mit einiger Verzögerung veröffentlicht. Zudem werden mit dieser Datenquelle nur ein Teil der Unternehmen, vornehmlich börsennotierte, erfasst. Schließlich hängt es von der jeweiligen Fragestellung ab, ob Textanalysen einen Mehrwert gegenüber den veröffentlichten Unternehmenszahlen zur laufenden und erwarteten Geschäftsentwicklung bieten.

Einer stärkeren Nutzung von Presseartikeln und Nachrichten für makroökonomische Analysen steht der zum Teil recht hohe Aufwand entgegen, der für die Auswertung der Texte betrieben werden muss. Dieser variiert mit dem jeweiligen Analysegegenstand. Während der Aufwand für eine einfache Wörter-

buch-basierte Analyse von Artikeln einer einzelnen oder wenigen Zeitungen noch vergleichsweise gering ist, steigt er mit dem Umfang der Datenquellen und der Komplexität der gewählten Analysemethoden rasch an. Eine Verstetigung der Analysen wird zudem dadurch erschwert, dass der Zugang zu den jeweiligen Datenquellen Beschränkungen unterliegen kann. Schließlich gibt es noch wenig etablierte Vorgehensweisen für Textauswahl und Textauswertungen, sodass die Ergebnisse, wie auch in anderen jüngeren Forschungszweigen, teilweise noch stark von Ad-hoc-Annahmen abhängen. Angesichts der recht großen Potenziale dürfte sich die Bedeutung der Analyse von Presseartikeln und Nachrichten für makroökonomische Fragestellungen aber zukünftig erhöhen.

2.2 Soziale Medien

Soziale Medien werden von einem Großteil der Bevölkerung genutzt, um soziale Kontakte zu pflegen oder Informationen auszutauschen. Dabei hinterlassen die Nutzer Spuren, die für wissenschaftliche Analysen eine wertvolle Datengrundlage bilden können. Aus makroökonomischer Sicht ist besonders interessant, dass sich die Stimmung und die Erwartungen der Nutzer in diesen Daten widerspiegeln dürften. Um diese Daten auszuwerten, muss jedoch ein recht hoher Aufwand betrieben werden, nicht zuletzt, weil sie häufig unstrukturiert und häufig textbasiert sind. Im Folgenden werden die Potenziale und Grenzen von Daten aus sozialen Medien anhand von Facebook (Abschnitt 2.2.1) und Twitter (Abschnitt 2.2.2) diskutiert. Sie gehören zu den beliebtesten sozialen Medien, gewähren zumindest teilweise Zugang zu ihren Daten und sind schon für eine Vielzahl von wissenschaftlichen Studien genutzt worden, auch wenn makroökonomische Fragestellungen dort bislang noch nicht im Fokus standen. Freilich gibt es viele andere Plattformen unter den soziale Medien, die als Datenquelle dienen können. In vielen der für makroökonomische Analysen relevanten Eigenschaften ähneln diese Datenquellen in der Datenstruktur und den damit verbundenen Möglichkeiten und Problemen jedoch Facebook und Twitter. Die Auswahl einer geeigneten Datenquelle wird im Einzelfall von der Fragestellung und vom verfügbaren Datenzugang abhängen.

2.2.1 Facebook

2.2.1.1 Datenquellen

Facebook ist mit ca. 2,4 Milliarden monatlich aktiven Nutzern (statista 2019) das größte soziale Netzwerk weltweit. Zunächst gegründet als soziales Netzwerk für Studierende, ist Facebook zu einem der größten Unternehmen weltweit herangewachsen. Facebook selbst stellt damit nicht nur eine potenzielle Datenquelle dar, sondern ist neben Amazon und Google eine der großen Online-Plattformen. Somit ist Facebook selbst Teil der modernen Wirtschaftsstruktur geworden. Die aktuelle und zukünftige Rolle von Plattformen für die Wirtschaft werden bisher weder in der Forschung noch in der Politik vollends verstanden. Der regulatorische Rahmen ist daher noch in der Entwicklung begriffen und Änderungen unterworfen, die auch die Datenverfügbarkeit beeinflussen. Diese Unsicherheit ist ein relevanter Aspekt, wenn es um die mögliche Verstetigung zukünftiger Analyseergebnisse geht, beispielsweise hinsichtlich der Frage, ob rechtliche Änderungen die Datenverfügbarkeit verringern.

2.2.1.2 Datenbeschaffenheit

Für die Qualität der Daten spielt die Datenherkunft eine große Rolle. Die Teilnahme am Netzwerk ist für die einzelnen Nutzer kostenlos; sie zahlen implizit mit ihren zur Verfügung gestellten Daten. Aufgrund des Netzwerkcharakters weist der Markt Ähnlichkeiten zu einem natürlichen Monopol auf: Je mehr

Personen Mitglied von Facebook sind, desto höher ist der Nutzen für jeden Einzelnen. Angesichts der zentralen Rolle von Facebook für die Kommunikation bestimmter Bevölkerungsgruppen ist es für Einzelpersonen mit sozialen Kosten verbunden, nicht an dem Netzwerk teilzunehmen. Somit verstärkt sich der Anreiz, selbst am Netzwerk teilzunehmen, je größer die Zahl der dort bereits präsenten Freunde und Bekannten ist. Dies hat in der Vergangenheit zu dem hohen Wachstum bei den Mitgliederzahlen von Facebook beigetragen. Heute ist in vielen Ländern mindestens ein Drittel der Bevölkerung Mitglied von Facebook. In den Vereinigten Staaten beträgt der Anteil der Bevölkerung mit einem Facebook-Konto sogar 68 Prozent, stagniert aber seit dem Jahr 2017.[9] Das Geschäftsmodell von Facebook ist, auf Basis dieser Daten Werbeplätze und die für den personalisierten Zuschnitt relevanten Informationen zu verkaufen.

Beim Datenzugang unterscheidet Facebook zunächst nicht zwischen Forschungsanwendungen und kommerziellen Interessen. Es gibt mehrere Möglichkeiten für Werbekunden, automatisiert sowohl Inhalte als auch deren Reichweite zu analysieren oder zu steuern. Über die Plattform ist es auch als nicht zahlender Nutzer möglich, zu untersuchen, welche Bevölkerungsgruppen durch Facebook abgedeckt sind. Der Datenzugang für nicht zahlende Kunden ist allerdings deutlich schlechter als der für zahlende Kunden. So konnten Werbekunden im Jahr 2016 die für sie als Zielgruppe relevanten Nutzer nach etwa 29 000 Kategorien aufschlüsseln, darunter das ungefähre Haushaltseinkommen, politische Affiliationen wie Parteipräferenzen und sogar höchst privates Verhalten wie „breastfeeding in public“ (ProPublica 2016). Der Großteil dieser Indikatoren wird durch einen Facebook-internen Algorithmus generiert, der auf maschinellem Lernen basiert und die Angaben der Nutzer, ihre Kommentare und ihr sonstiges Onlineverhalten auswertet. Für nicht zahlende Kunden sind nur deutlich weniger detaillierte Informationen verfügbar, die es beispielsweise erlauben abzuschätzen, wie viele Personen eines bestimmten Geschlechts in einem geographischen Umkreis (z.B. 10 km um Berlin-Mitte) regelmäßig Facebook nutzen und wie viele davon etwa über 1 000 Euro netto im Monat zur Verfügung haben. Diese Zahlen lassen sich für Forschungszwecke manuell exportieren, das Filtern von Nutzerprofilen ist jedoch nicht möglich.[10]

Der Datenzugang wird über die Facebook Graph API (Application Programming Interface) bereitgestellt. Damit lassen sich Inhalte in das Netzwerk schreiben (z.B. Texte oder Bilder posten) und Informationen auslesen. Daten wie die Kommentare von einzelnen Nutzern und, sofern nicht geschützt, deren persönliche Profile, lassen sich mittels der API ohne Verzögerung herunterladen. Dabei lässt sich prinzipiell der gesamte Zeitraum seit dem Start von Facebook nutzen. Die Datenfülle ist also ausgesprochen groß. Gerade für Daten aus Deutschland ist aber zu bedenken, dass viele Profile (darunter auch kommerzielle und nicht-kommerzielle von Unternehmen oder Lobbygruppen) erst seit dem Jahr 2010 bestehen und Facebook zuvor vor allem im englischsprachigen Raum genutzt wurde.

Durch die große, weltweite Reichweite und die hohen Nutzerzahlen ist das verfügbare Datenvolumen bei Facebook sehr umfangreich. Allerdings hat Facebook einige Grenzen für die automatisierte Nutzung dieser Daten eingezogen. Dies liegt nicht zuletzt an den geltenden Datenschutzstandards. Dadurch ist kein Vollzugriff auf alle Daten möglich; Facebook erlaubt lediglich Zugriff auf die öffentlich freige-

[9] Vgl. PEW Research Center (2018).

[10] Facebook-intern wird auch zukünftiges Verhalten der Nutzer prognostiziert. Dafür wird das Programm „FBLearner Flow“ genutzt, das auf nicht öffentlich bekannten Machine-Learning-Ansätzen beruht. Diese Informationen werden Werbekunden als Produktangebot zur Verfügung gestellt. Daneben bietet Facebook gelegentlich auch andere Anwendungen an. So wurde beispielsweise ein Algorithmus zur Vorhersage von Selbstmordgefahr programmiert, der zum Ziel hat, frühzeitig auf dieses Risiko hinzuweisen. In der Europäischen Union ist er aufgrund der Europäischen Datenschutzverordnung allerdings nicht operabel (Business Insider 2018). Diese prozessierten Daten sind ebenfalls nicht für Forschungszwecke nutzbar.

schalteten Informationen. Informationen von Profilen, die nur für Facebook-Freunde sichtbar oder unsichtbar geschaltet sind, sind dagegen nicht verfügbar. Dies schränkt die Nutzbarkeit der Daten bezüglich von Netzwerken, persönlichen Vorlieben und ähnlichen Informationen massiv ein. Abrufbar sind jedoch Kommentare auf öffentlichen Seiten, wie sie beispielsweise unter Zeitungsartikeln hinterlassen werden (Ademmer et al. 2019b; Ademmer und Stöhr 2019).

Für die Datenabfrage ist grundsätzlich ein eigenes Facebook-Profil notwendig. Ferner wird ein Zugriffschlüssel (Token) benötigt, der von Facebook zur Verfügung gestellt wird. Nicht zuletzt, weil die Nutzung von Detailinformationen davon abhängt, ob Profile öffentlich freigeschaltet sind, dürften die verfügbaren Profildaten weder für die Gesamtbevölkerung noch die Gesamtheit der Facebook-Nutzer repräsentativ abbilden. Somit ist der Nutzen der Daten für Fragestellungen, die auf die Profildaten (z.B. Geschlecht, Vorlieben) abzielen, eher gering. Hinzu kommt, dass die Zahl der Nutzer und deren Nutzungsintensität über die Zeit variieren.

Interagieren die Nutzer mit einem öffentlichen Inhalt einer Facebook-Seite (z.B. „Like“ für das BMWi) oder verfassen Kommentare zu einer dort veröffentlichten Nachricht, so lässt sich diese Interaktion unabhängig vom Datenschutzprofil verwenden. Öffentliche Kommentare (Posts) und „Likes“ sind also vollständig für Analysen nutzbar. In diesem Bereich ist der potenzielle Nutzen der Datenquelle somit hoch, sofern die ggf. fehlende Repräsentativität gegenüber der Gesamtbevölkerung für die jeweilige Fragestellung keine große Relevanz aufweist oder diese adäquat berücksichtigt werden kann.

Neben der Datenabfrage hat Facebook im Februar des laufenden Jahres im Rahmen von Social Science One einen anonymisierten Datensatz zugänglich gemacht, der es beispielsweise erlaubt, zu verfolgen, wie sich geteilte Links auf Facebook verbreiten. So wird beispielsweise erstmals groß angelegte Forschung zur Frage möglich, welchen Einfluss sogenannte „Fake News“ auf Facebook entfalten.

Die Qualität der Posts kann je nach Thema und Anwendergruppe sehr heterogen ausfallen. Der Detailgrad reicht von ausführlichen Texten bis zu einzelnen Emoticons; in der Regel sind sie aber recht kurz und bestehen aus wenigen Sätzen oder einzelnen Zeichen. Ein Problem bei der Auswertung der von privaten Nutzern geschriebenen Textdaten von Facebook, beispielsweise im Vergleich zu Presseartikeln, ist die große Zahl an Rechtschreib- und Grammatikfehlern. Ist ein Begriff falsch geschrieben, wird es einem Auswertungsalgorithmus ohne spezielle Anpassungen nicht möglich sein, diesen mit einem anderen Begriff korrekt in Beziehung zu setzen. So würde ein Textanalyse-Algorithmus die Begriffe „Fehler“ und „Feler“ als separate Begriffe behandeln und Analysen, die dies nicht berücksichtigen, wären mit einem Messfehler behaftet. Wenn Fehler systematisch – also beispielsweise bei bestimmten sozio-demographischen Gruppen – auftreten, kann dies eine systematische Verzerrung der Analyseergebnisse zur Folge haben.

2.2.1.3 Methodik

Die Datenabfrage mittels der Facebook Graph API ist grundsätzlich unproblematisch, zumal eine Vielzahl von Anwendungsbeispielen frei verfügbar ist. Für die Datenabfrage eignet sich beispielsweise die Programmiersprache python. Um die Anzahl nicht kostenpflichtiger Anfragen zu begrenzen, werden kurzfristige und langfristige Autorisierungsschlüssel vergeben. Diese Zugangscodes dienen der Facebook API dazu, zu erkennen, von wem eine Anfrage ausgeht. Für Forschungsanwendungen können kurzfristige Autorisierungsschlüssel unproblematisch generiert werden. Jedoch müssen sie, je nach Volumen der angeforderten Daten, typischerweise alle 30 bis 120 Minuten erneuert werden. Eine automatische Generierung dieser Zugangscodes ist recht umständlich. Für größere Datensammlungen ist deshalb ein recht großer Aufwand notwendig, zumindest sofern sie kostenfrei erfolgen sollen.

Neben Textdaten bietet es sich vor allem an, „Likes“ und ähnliche zählbare Interaktionen zu untersuchen. Ein Beispiel, das für Unternehmen relevant ist, sind „Likes“ für Ankündigungen neuer Produkte. Für einen solchen Ansatz können alle Interaktionen der Nutzer auf einer bestimmten Facebook-Seite heruntergeladen werden. Diese können dann bei Bedarf gefiltert werden, um die Datenmengen zu reduzieren. Über die relativ leicht handhabbare API, stehen viele Möglichkeiten offen, die Daten von vornherein passend zuzuschneiden und somit durch direktes Filtern den Bedarf an Speicherplatz gering zu halten.

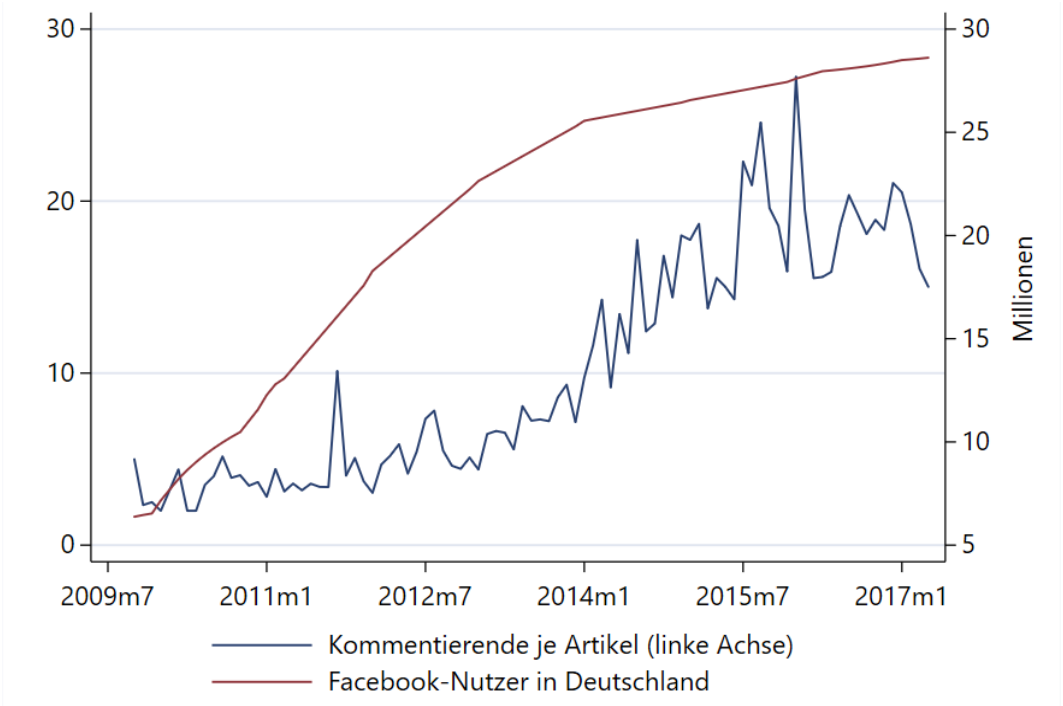
Je nach Form und Fragestellung müssen die Daten ggf. vorab bearbeitet werden (z.B. Textdaten in quantitative Daten umwandeln (Abschnitt 2.1)), um sie für empirische Analysen nutzbar zu machen. Ein Hindernis bei der Prozessierung der Daten entsteht durch die weltweite Verbreitung von Facebook. Die Nutzer verwenden zahlreiche verschiedene Sprachen. Dies stellt für viele länderspezifische Analysen kein Problem dar. In bestimmten Zusammenhängen (z.B. multilinguale Länder oder bei einem Fokus auf Migrantengruppen) kann die Aussagekraft der Daten jedoch eingeschränkt werden. Für eine Verarbeitung der Textinformationen muss dann zunächst die Sprache einer Nachricht identifiziert werden. Allein dieser Schritt ist technisch bereits recht anspruchsvoll. Sofern die Ursprungssprache erfolgreich von einem Algorithmus identifiziert worden ist, kann dann die Sprache mittels automatisierter Werkzeuge in eine einheitliche Arbeitssprache übersetzt werden. Dabei treten jedoch häufig weitere Fehler auf. Vor allem Grammatikfehler und falsche Übersetzungen von Einzelbegriffen sind für die weiterführenden Analysen problematisch und verschlechtern die Datenqualität zum Teil deutlich.

2.2.1.4 Bisherige Anwendungen

Bislang wird Facebook vor allem in den Politikwissenschaften und in benachbarten Feldern wie der politischen Kommunikation als Datengrundlage genutzt, da das Kommunikationsverhalten der Nutzer dort im Zentrum des Interesses steht. So wurde beispielsweise von Stier et al. (2017) untersucht, wie die großen Parteien ihre Online-Kommunikation im Vorlauf der Bundestagswahl 2017 gestaltet haben, wer PEGIDA-Unterstützende anziehen konnte und wie die politische Kommunikation verschiedener Parteien auf Facebook die Themen dieser Bewegung aufgegriffen hat. Wählerverbindungen mit bestimmten Parteien lassen sich durch „Likes“ und andere Interaktionen gut messen. Zugleich ist aber auch die Frage der Repräsentativität wichtig und es ist zu vermuten, dass online aktive Wähler nur einen besonderen Teil der Wählerschaft bestimmter Parteien ausmachen, insbesondere auch weil die Assoziation mit einer politischen Bewegung mit sozialen Kosten verbunden sein kann.

Ein Anwendungsfeld mit vielfältigen Möglichkeiten ist die Analyse von Facebook-Kommentaren zu ökonomischen Themen auf öffentlich zugänglichen Seiten. Sie können als Stimmungsindikator verwendet werden. Ademmer et al. (2019c) sowie Ademmer und Stöhr (2019) untersuchen Kommentare zu Artikeln zum Thema Migration, die auf den Seiten der 100 größten deutschen Lokal- und Regionalzeitungen hinterlassen wurden. Dabei lässt sich beobachten, wie sehr der starke Zustrom von Flüchtlingen nach Deutschland im Jahr 2015 zu einem Anstieg der Berichterstattung und insbesondere zu einem überproportional starken Wachstum des öffentlichen Diskurses der Bürger geführt hat (Abbildung 4). Anhand der Kommentartexte lässt sich nachvollziehen, welche Themen besonders diskutiert werden. Dabei lässt sich beispielsweise feststellen, dass auf dem Höhepunkt des Zustroms zunächst Verteilungsthemen wie fiskalische Kosten und Wohnraumknappheit thematisiert wurden. Kulturelle Aspekte und andere eher migrationsablehnende Themen wurden erst später stärker diskutiert, etwa ab November 2015. Der öffentliche Diskurs verschob sich danach zum Jahreswechsel 2015/2016 nachhaltig zu Sicherheitsthemen. Somit lässt sich sogar tagesgenau untersuchen, wann sich der öffentliche Diskurs der auf Facebook aktiven Bürger verschoben hat (Ademmer und Stöhr 2019). Diese Methodik

Abbildung 4:
Zahl der kommentierenden Facebook-Nutzer je Artikel zum Thema Migration[a]



[a]Monatsdaten. Anzahl der Kommentare verschiedener Nutzer zu Artikeln mit dem Thema Migration.

Quelle: Ademmer et al. (2019c).

lässt sich prinzipiell auf viele andere auch makroökonomische Themen anwenden und gut mittels Eventstudien untersuchen.

In der makroökonomischen Forschung kann Facebook beispielsweise einen Mehrwert bieten, wenn die Netzwerkstrukturen ausgewertet werden. Bailey et al. (2018) nutzen anonymisierte Facebook-Daten für die gesamten Vereinigten Staaten,[11] die zeigen, dass Individuen, die in ihrem Netzwerk einen Anstieg der Hauspreise wahrnehmen, eher von Mietern zu Hausbesitzern werden, größere Häuser kaufen und bereit sind, mehr für eine Immobilie zu bezahlen. Dieser Effekt wird von Veränderungen der erwarteten Entwicklung der Hauspreise getrieben. Neben dieser qualitativ hochwertigen Studie, die aber auf direkte Zusammenarbeit mit Facebook angewiesen war, gibt es bislang allerdings relativ wenige Forschungsarbeiten, die makroökonomische Zusammenhänge untersuchen.

Eine Ausnahme stellen Arbeiten dar, die Facebook-Daten verwenden, um die Stimmung der Nutzer zu erfassen. Dazu können beispielsweise die positiv und negativ konnotierten Wörter in den Statusupdates der Nutzer zu einem Stimmungsindikator aggregiert werden. Siganos et al. (2014) untersuchen den Zusammenhang zwischen der auf Facebook geäußerten durchschnittlichen Stimmung und Aktienmarktentwicklungen in den USA. Sie finden einen positiven Zusammenhang und zeigen, dass mittels der anhand von Facebook-Daten gemessenen Stimmung am Sonntag die Entwicklungen am Aktienmarkt am folgenden Montag besser vorhergesagt werden können.[12] Daas and Puts (2014) analysieren

[11] Dieser umfangreiche Datenzugang ist nur möglich, weil einer der Autoren bei Facebook arbeitet.

[12] Karabulut (2013) und Siganos et al. (2014) argumentieren in diesem Zusammenhang, dass sich Facebook-Daten nutzen lassen, um einen Indikator für das sogenannte „Bruttonationalglück" zu bilden, da in vielen fortgeschrittenen Volkswirtschaften etwa die Hälfte der Bevölkerung Facebook nutzt und man aus ihren Posts und Kommentaren positive und negative Wahrnehmungen („sentiment") extrahieren könne. Zu diesem Zwecke stellte

Facebook-Kommentare aus den Niederlanden und zeigen, dass sie hoch korreliert mit konventionellen umfragebasierten Indikatoren zum Konsumvertrauen sind. Dazu werten sie die Texte und Emoticons von Kommentaren aus und klassifizieren die damit verbundene Empfindung als positiv, negativ oder neutral. Die Korrelation ist im Vergleich zu Daten aus anderen sozialen Medien höher, sie kann allerdings noch gesteigert werden, wenn die Facebook-Daten mit Kommentaren von Twitter kombiniert werden. Sie finden auch, dass die anhand von Facebook-Daten gemessene Konsumentenstimmung systematisch positiver ist als die mittels konventioneller Methoden. Eine Ursache könnte die Selbstselektion der Anwender sein.

2.2.1.5 Potenziale und Grenzen

Facebook ist durch relativ geringen Bezug zu ökonomischen Themen und des recht hohen Aufwands, der für die Auswertung der Daten betrieben werden muss, nicht optimal für makroökonomische Fragestellungen geeignet. Daher gibt es auf diesem Gebiet bisher auch nur wenige Anwendungen.

Grundsätzlich bieten Facebook-Daten jedoch nicht zuletzt aufgrund der hohen Nutzerzahlen ein beträchtliches Potenzial, um das Verhalten und die Stimmung von Konsumenten zu analysieren. Dazu zählt beispielsweise, wie die Nutzer auf Veränderungen in ihrem Umfeld reagieren. Auch kann die Einstellung der Nutzer bezüglich politisch relevanter Aspekte, wie die Akzeptanz der Gemeinschaftswährung Euro, der Klimapolitik oder der Haushaltspolitik, erfasst werden. In diesem Zusammenhang wäre es aus makroökonomischer Sicht von Interesse, wie Konsumenten auf bestimmte Ereignisse oder wirtschaftspolitische Maßnahmen reagieren, beispielsweise auf Unsicherheiten bezüglich der zukünftigen Nutzung von Dieselfahrzeugen, auf eine erhöhte wirtschaftspolitische Unsicherheit oder auf steuerliche Entlastungen. Dies könnte für die Konjunkturanalyse auch dazu beitragen, kurzfristig zwischen angebots- und nachfrageseitigen Schwankungen besser zu unterscheiden. Schließlich sind Facebook-Daten zum Teil bereits dafür verwendet worden, um Indikatoren für das Konsumentenvertrauen zu entwickeln, die Aufschluss über die zukünftigen privaten Konsumausgaben geben können. Diese Indikatoren könnten zum einen dazu beitragen, die bislang gängigen Erhebungen zum Konsumentenvertrauen, die auf Umfragen ausgewählter Haushalte basieren und monatlich veröffentlicht werden, zu prognostizieren. Zum anderen könnten sie angesichts der hohen Nutzerzahlen möglicherweise auch ein umfassenderes und stärker disaggregiertes Bild für das Konsumentenvertrauen liefern.

In der Praxis stellen sich einer stärkeren Nutzung von Facebook-Daten für makroökonomische Analysen allerdings große Hürden entgegen. So ist die Auswertung der Daten mit einem sehr hohen Aufwand verbunden. Eine mögliche Verstetigung von Analysen, wie sie für eine regelmäßige Schätzung des Konsumentenvertrauens erforderlich wäre, ist zudem nicht gewährleistet, da sich jederzeit die Geschäftspraktiken von Facebook oder die Datenschutzrichtlinien ändern und dadurch den Datenzugang erschweren können. Dies spricht dafür, dass Facebook als Datenquelle in absehbarer Zukunft keine große Rolle in der laufenden Konjunkturanalyse spielen dürfte. Zwar können die Facebook-Daten in der Einzelbetrachtung für ausgewählte makroökonomische Fragestellungen eine interessante Grundlage bieten. Jedoch sind auch dafür die Hürden aufgrund der Zugangsbeschränkungen recht hoch. Hinzu kommt, dass die Daten trotz der weiten Verbreitung in der Regel nicht repräsentativ für die Gesamtbevölkerung sein dürften. So ist zu bedenken, dass sich die Nutzer, die regelmäßig Kommentare auf Facebook verfassen, in gewissen Charakteristika von der Gesamtbevölkerung unterscheiden. Zu den Determinanten dieser Selbstselektion ist bisher recht wenig bekannt. Ein vermutlich verallgemeinerbares Ergebnis ist, dass besonders nachrichteninteressierte Personen, die oft weniger Vertrauen in

Facebook zeitweise den sogenannten „Gross National Happiness"-Index bereit (Kramer 2010), der die positiv und negativ konnotierten Wörter in den Statusupdates der Nutzer aggregierte.

klassische Medien haben, häufiger online Nachrichten konsumieren und auch kommentieren (Kalogeropoulous et al. 2017).

2.2.2 Twitter

2.2.2.1 Datenquellen

Twitter ist ein sogenannter Microbloggingdienst, bei dem Nutzer öffentliche oder private Kurznachrichten („Tweets“) im Umfang von bis zu 280 Zeichen (bis November 2017 bis zu 140 Zeichen) verbreiten können. Gegründet wurde Twitter im März 2006 zunächst als „twttr“ und ist seit April 2007 Teil von Twitter Inc. mit Sitz in San Francisco. Twitter kann über einen Browser oder per App via Smartphone und Tablet genutzt werden. Ursprünglich war der Dienst primär gedacht als Medium, um SMS zu publizieren, daher auch die Einschränkung auf wenige Zeichen.

Twitter wird als Kommunikationsplattform von Privatpersonen, Unternehmen und anderen Organisationen genutzt und grenzt sich dabei von anderen sozialen Netzwerken – wie z.B. Facebook – dadurch ab, dass die Netzwerkstrukturen weniger persönliche Kontakte als Interessen widerspiegeln. Insgesamt hat Twitter ca. 340 Mio. aktive Nutzer (Stand Januar 2020, Hootsuite „Digital 2020“). Als Microbloggingdienst hat Twitter damit de facto in großen Teilen der Welt eine marktbeherrschende Stellung.

2.2.2.2 Datenbeschaffenheit

Auf Twitter abgesendete Tweets sind standardmäßig öffentlich. Nur etwa 10 Prozent aller Nutzer haben ihre Profile auf bestimmte Leser eingeschränkt. Insgesamt dürften datenschutzrechtliche Fragen für die Datenverfügbarkeit daher von geringerer Bedeutung sein als bei vielen anderen sozialen Medien.

Der eigentliche Text eines Tweets kann aus Unicode-Zeichen bestehen, insbesondere auch Emojis. Im Text können dazu andere Nutzer mittels @-Zeichen markiert und verknüpft werden. Das Verwenden von sogenannten Hashtags mittels #-Zeichen kategorisiert den Text und ordnet ihn einem Thema zu. Neben dem eigentlichen Text kann dazu ein Foto angefügt werden, und ein Ort angegeben werden, auf den sich der Tweet bezieht. Daneben werden automatisch eine Vielzahl an weiteren Metadaten durch die Benutzeroberfläche, bzw. das Nutzergerät, generiert. So z.B. der Standort, sofern vom Nutzer freigegeben, sowie Sprache des Tweets oder des Betriebssystems. Neben den Tweets selbst wird zudem das sogenannte Friends- und Follower-Netzwerk gespeichert, also die Nutzer, die die Tweets eines anderen Nutzers abonniert haben, bzw. deren Tweets abonniert werden.

Weltweit ist Twitter.com die sechst meistbesuchte Internetseite und liegt damit im Vergleich zu anderen sozialen Netzwerken hinter Facebook (4.) und vor Instagram (7.). Twitter ist grundsätzlich weltweit verfügbar, jedoch ist der Zugriff auf den Dienst in einigen Ländern eingeschränkt bzw. verboten. So ist der Dienst offiziell in China, Nordkorea und Iran blockiert. In Ägypten, Irak, der Türkei und Turkmenistan ist er nur eingeschränkt verfügbar. Daneben macht eine Vielzahl von Ländern Gebrauch von der Möglichkeit, Tweets löschen zu lassen. Dies stellt für die Analyse von Twitter-Daten aus fortgeschrittenen Volkswirtschaften freilich kein Problem dar.

Die Nutzerzahlen des Microbloggingdiensts ist von monatlich ca. 30 Mio. aktiven Nutzern im Jahr 2010 auf knapp 300 Mio. aktive Nutzer in 2015 gestiegen. Seitdem ist die Zahl weitgehend stabil und lag zuletzt bei 340 Mio. Der Anteil von Nutzern pro Land variiert stark. Weltweit nutzen ca. 6 Prozent der über 13-Jährigen den Dienst, wobei dies in Saudi Arabien 53 Prozent, in Japan und Irland 40 Prozent, in Deutschland 7 Prozent und beispielsweise in Indien nur 1 Prozent sind (Stand Januar 2020, Hootsuite „Digital 2020“). Gleichzeitig variiert die Nutzung mit dem Alter und Geschlecht: Während beispielsweise

in Indien 85 Prozent der Nutzer männlich sind, sind in Indonesien 68 Prozent der Nutzer weiblich. Weltweit nutzen etwa 15 Prozent der Altersgruppe zwischen 25 und 34 Twitter, jedoch nur etwa 7 Prozent der über 60-Jährigen. In Deutschland sind laut Statista 33 Prozent der 20 bis 29-Jährigen Internetnutzer bei Twitter aktiv und 20 Prozent der über 60-Jährigen. Insgesamt stellt Twitter somit eine sehr umfangreiche Datenquelle dar. Die Nutzer sind jedoch ähnlich wie bei Facebook nicht repräsentativ für die Gesamtbevölkerung, worauf in Analysen ggf. Rücksicht genommen werden muss. Hinzu kommt, dass sich der Nutzerkreis im Zeitablauf deutlich verändert haben könnte.

Weltweit verwenden ca. 80 Prozent aller Nutzer Twitter von Mobiltelefonen aus. Daneben geht Twitter Inc. davon aus, dass von den 340 Mio. aktiven Nutzern ca. 23 Mio. sogenannte Bots sind. Diese automatisch von Programmen gesteuerten Accounts können beispielsweise Informationen zu Wetter- oder Verkehrsdaten automatisch veröffentlichen. Daneben gibt es aber auch Bots, die menschliche Accounts imitieren und so versuchen, bestimmte Themen, Nachrichten, oder anderen Accounts eine größere Aufmerksamkeit zu verschaffen. Es liegen jedoch Methoden vor, die es erlauben, diese mit hoher Treffsicherheit von Personen zu unterscheiden (Gorodnichenko et al. 2018).

Tweets sind semistrukturierte Daten: teils klar strukturierte Metadaten wie bspw. Sprache des Tweets oder eine Ortsangabe in Form von Koordinaten, teils unstrukturierte Daten wie der Text des Tweets selbst. Daten über Nutzernetzwerke sind bis auf wenige Ausnahmen (wie beispielsweise die nutzereigene Beschreibung) strukturiert. So werden Follower und Freunde durch Nutzer-IDs eindeutig identifiziert.

Twitter ermöglicht einen Zugriff auf die generierten Daten über mehrere APIs. Daneben gibt es einige öffentlich verfügbare historische Datensätze von Drittanbietern, die sich jedoch größtenteils auf bestimmte Hashtags beziehen, zum Teil unvollständige Metadaten aufweisen und nicht aktualisiert werden.

Der API-Zugang zu den Twitter-Daten ist beschränkt in Bezug auf die Anzahl von Nutzern oder Tweets. Es gibt dabei grundsätzlich zwei unterschiedliche Arten des Zugriffs:

Zum einen eine „Realtime Streaming API“ für Echtzeitdaten. Dies bedeutet, dass aktuell abgeschickte Tweets gesammelt werden können, jedoch kein Zugriff auf historische Daten besteht. Der „Stream“ kann dabei gefiltert werden, z.B. nach Hashtags, Orten, Sprache oder Vorhandensein eines Fotos. Die Anzahl der Tweets beschränkt sich normalerweise auf eine Zufallsstichprobe von 1 Prozent. Eine kostenpflichtige Version der API („Decahose“) bietet eine Stichprobe im Umfang von 10 Prozent, jedoch ohne die Möglichkeit, nach den oben genannten Kriterien zu filtern.

Zum anderen eine „Search API“ für historische Daten bis zurück in das Jahr 2006. Dabei kann nach bestimmten Stichwörtern bzw. Hashtags gesucht oder auf sämtliche Tweets von spezifischen Nutzern zugegriffen werden. Auch hier ist wieder zu unterscheiden zwischen der kostenlosen und kostenpflichtiger („Premium“ und „Enterprise“) Varianten. Letztere können Zugriff auf den gesamten Korpus an Tweets geben, erstere hingegen bietet nur eine Stichprobe.

Die erste (Echtzeit-)Variante bietet den Vorteil, eine Zufallsstichprobe zu erhalten, mit der Einschränkung jedoch, dass Zeitreihen mühsam über einen längeren Zeitraum aufgebaut werden müssen. Die zweite Variante ist von Vorteil, wenn sich die zu analysierenden Daten klar auf bestimmte Schlagwörter oder Nutzerkonten eingrenzen lassen.

Neben diesen beiden APIs besteht zudem die Möglichkeit des „Batch Downloads“ von historischen Daten, d.h. dem Herunterladen eines Teils des Archives von Twitter-Daten, also sowohl Tweets als auch historische Netzwerkdaten. Diese Form der Datenbeschaffung ist jedoch kostspielig, gängige Preise belaufen sich auf ca. 10 000 US-Dollar je 1 Mio. Tweets oder je Netzwerk von 10 000 Nutzern.

Die durch die API oder als „Batch Download“ verfügbaren Daten werden im JSON-Format (JavaScript Object Notation) bereitgestellt, die in allen gängigen Programmiersprachen verarbeitet werden können.

Angesichts der hohen Zahl der Nutzer und der täglich gesendeten Tweets ist das verfügbare Datenvolumen, auch wenn man auf die kostenfreie Stichprobe von 1 Prozent der Tweets zurückgreift, sehr umfangreich. Sofern für eine Analyse die Echtzeitdaten selbst gesammelt werden sollen, ist allerdings für viele makroökonomische Analysen eine recht große Vorlaufzeit notwendig, um hinreichend lange Zeitreihen zu generieren.

2.2.2.3 Methodik

Die Form des Managements von Twitter-Daten ist abhängig von der Anwendungsfrage sowie den spezifischen Charakteristika des vorliegenden Datensatzes. In der Regel sind diese Datensätze aber ausgesprochen groß, sodass es einer entsprechenden IT-Infrastruktur bedarf, um sie auswerten zu können.

Eine Besonderheit stellen über eine Streaming API bezogene Daten dar: Sie haben den Vorteil, stetige oder häufig zu aktualisierende Anwendungen zu ermöglichen, da Tweets kontinuierlich eintreffen. Dafür muss jedoch auch die technische Infrastruktur mit Streaming-Daten umgehen und letztlich für den Anwender nutzbar machen können.

Softwareseitig ermöglichen einige moderne Programmiersprachen wie R oder Python einen reibungslosen Einstieg in die Analyse von Twitter-Daten: Mit einer Vielzahl von Softwarenpakten, wie z.B. Tweepy und python-twitter für Python, und twitteR und rtweet für R, lassen sich die APIs direkt aus der jeweiligen Umgebung ansprechen. Häufig verfügen diese zudem über Funktionen und Programme, die die jeweils gelieferten Daten weiterverarbeiten können.

Die großen Datenmengen machen häufig die Anwendung von anspruchsvollen Schätzverfahren erforderlich. Zudem gibt es für die Auswertung der Texte als inhärent qualitative Daten in den Wirtschaftswissenschaften noch vergleichsweise wenig Erfahrungswerte und etablierte Methoden (Abschnitt 2.1). Für die Auswertung des strukturierten Teils der Daten ergeben sich dagegen keine zusätzlichen speziellen methodischen Herausforderungen. Angesichts dieses strukturierten Teils der Daten und der standardisierten Form der Textnachrichten dürften Daten von Twitter im Vergleich zu anderen sozialen Netzwerken, wie Facebook, etwas leichter auszuwerten sein.

2.2.2.4 Bisherige Anwendungen

Twitter-Daten sind in der volkswirtschaftlichen Forschung bisher in einigen Studien verwendet worden. Die Forschungsfelder reichen dabei von Migration, über Finanzmärkte bis hin zu spezifischen makroökonomischen Themen.

Hausmann et al. (2018) verwenden Twitter-Daten, um Flüchtlingsströme von Venezuela in Nachbarländer zu quantifizieren. Dabei nutzen sie georeferenzierte Tweets, anhand derer Flüchtlingsrouten nachvollzogen werden können. Wird ein Nutzer beispielsweise zunächst lediglich in Venezuela geortet, später jedoch ausschließlich in einem anderen Land wie Kolumbien, kann davon ausgegangen werden, dass er das Land verlassen hat. Mittels dieser Heuristik, gekoppelt mit einer statistischen Methode, die für die Häufigkeit der Nutzung und für die Wahrscheinlichkeit, Teil der Stichprobe zu sein, korrigiert, lässt sich ein klares Bild der Flüchtlingsbewegungen zeichnen. Abbildung 5 zeigt die Verteilung und Bewegung venezolanischer Twitter-Nutzer: Jeder Punkt ist ein Nutzer, die Linien zwischen Punkten

Abbildung 5:
Twitter-Nachrichten von Venezolanern in Venezuela und Kolumbien[a]

[a]Daten für die Jahre 2017 und 2018. Jeder Punkt ist ein Nutzer, die Linien zwischen Punkten verbinden unterschiedliche Aufenthaltsorte eines einzelnen Nutzers.

Quelle: Hausmann et al. (2018).

verbinden unterschiedliche Aufenthaltsorte eines einzelnen Nutzers. Dadurch wird sichtbar, dass ein großer Teil der „Mobilität" sich innerhalb Venezuelas abspielt, jedoch ein beträchtlicher Teil ehemaliger Nutzer aus Venezuela später beispielsweise aus Kolumbiens Hauptstadt Bogotá den Dienst nutzt. Mit Hilfe dieser Methode wird die Zahl der insgesamt aus Venezuela Geflüchteten zwischen April des Jahres 2017 und April des Jahres 2018 auf 2,9 Mio. Menschen geschätzt – fast 10 Prozent der venezolanischen Bevölkerung.

Gholampour und van Wincoop (2019) nutzen eine Zeitreihe von Tweets zum Euro/Dollar Wechselkurs über 2 Jahre. Sie zeigen, dass das anhand von Twitter-Nachrichten gemessene „Sentiment" – durch als positiv, negativ oder neutral eingeschätzte Tweets – wertvolle Informationen bietet. Eine Handelsstrategie basierend auf diesen Nachrichten führt in der Studie zu besseren Ergebnissen als ein üblicher „Carry-Trade".

Lehrer et al. (2019) verwenden Twitter-Daten, um den monatlichen Verbrauchervertrauensindex des Conference Boards für die Vereinigten Staaten zu prognostizieren. Konkret berechnen sie auf Basis der in Tweets verwendeten Emojis ein stündliches Stimmungsmaß, dass der Methodik des von IHS Markit entwickelten U.S. Sentiment Index (USSI) ähnelt. Sie evaluieren die Prognosen verschiedener Modellansätze für das Verbrauchervertrauen des jeweils nächsten Veröffentlichungstermins für den Zeitraum von 2015 bis 2017. Sie zeigen zum einen, dass Methoden aus dem Bereich des maschinellen Lernens besser abschneiden als klassische ökonometrische Verfahren. Zum anderen liefern Modelle, die das stündliche aus den Twitter-Daten berechnete Stimmungsmaß verwenden, genauere Prognosen als Modelle, in die lediglich der Monatsdurchschnitt eingeht. Allerdings sind der Stütz- und Evaluierungszeitraum recht knapp bemessen. Des Weiteren zeigen die Autoren, dass eine von ihnen vorgeschlagene Weiterentwicklung für den MIDAS-Ansatz, der es grundsätzlich erlaubt, Variablen mit unterschiedlichen

Frequenzen in einem Modell zu berücksichtigen, besonders geeignet ist, um mit Heterogenitäten von Tages- oder Stundendaten umzugehen. Solche Heterogenitäten können auftreten, wenn Twitter an verschiedenen Tagen oder zu unterschiedlichen Tageszeiten von sich deutlich unterscheidenden Nutzergruppen verwendet wird. Dieser Ansatz könnte auch relevant sein, um andere in sehr hoher Frequenz vorliegende Daten aus dem Umfeld von Big Data auszuwerten.

Auch andere Analysen verwenden Twitter-Nachrichten, um daraus Rückschlüsse für das Verbrauchervertrauen zu ziehen. So nutzen Daas und Puts (2014) Twitter-Daten zur Prognose des Konsumentenvertrauens für die Niederlande. Sie zeigen, dass Facebook-Daten dafür geeigneter sind, auch wenn die zusätzliche Verwendung von Twitter-Nachrichten die Prognose noch etwas verbessert. O'Connor et al. (2010) finden für die Vereinigte Staaten eine recht hohe Korrelation zwischen einem aus Twitter-Nachrichten gebildeten Vertrauensindikator mit dem von Gallup monatlich ausgewiesenen Konsumentenvertrauen. Schließlich veröffentlicht das Wall Street Journal fortlaufend den sogenannten „U.S. Social Sentiment Index". Dieser basiert auf einer automatisierten Auswertung von Tweets hinsichtlich ihrer durch bestimmte Wörter oder Emojis ausgedrückten Stimmung. Die geografische Aufschlüsselung von Twitter-Daten erlaubt es zudem, diesen Stimmungsindikator auch für unterschiedliche Regionen zu erstellen.

Bianchi et al. (2019) nutzen ausschließlich Twitter-Daten des amerikanischen Präsidenten, um Erwartungen über die Geldpolitik in den Vereinigten Staaten zu analysieren. Dabei untersuchen sie die Marktreaktionen auf Tweets, in denen eine Niedrigzinspolitik propagiert wird. Im unmittelbar folgenden Marktgeschehen ist ein Effekt auf die erwartete zukünftige Zinspolitik der Fed festzustellen – was darauf hindeutet, dass die Fed wiederum nicht als komplett unabhängig wahrgenommen wird.

Für Prognosen der realwirtschaftlichen Entwicklung sind Twitter-Daten bisher kaum genutzt worden. Lediglich Indaco (2019) zeigt, dass die Anzahl der in einem Land abgesendeten Twitter-Nachrichten mit dem Bruttoinlandsprodukt korreliert. Allerdings bezieht sich die Analyse nur auf einen sehr kurzen Zeitraum – insgesamt werden 140 Mio. Nachrichten in den Jahren 2012 und 2013 untersucht – und ein Großteil der durch die Nachrichten erklärten Varianz dürfte auf Unterschiede im Bruttoinlandsprodukt über die Länder hinweg zurückzuführen sein. Ob dieser Zusammenhang für die Konjunkturprognose von Bedeutung sein kann, ist somit zweifelhaft.

2.2.2.5 Potenziale und Grenzen

Twitter bietet ähnlich wie Facebook Zugang zu einem umfangreichen Daten-Pool, der insbesondere Stimmungen und Erwartungen der Nutzer abbilden kann. Bisherige Anwendungen in makroökonomischen Analysen sprechen dafür, dass Twitter-Daten dazu geeignet sind, die gängigen Indikatoren zum Konsumentenvertrauen zu prognostizieren. Darüber hinaus könnten mit Twitter-Daten auch Reaktionen auf makroökonomisch relevante Ereignisse oder wirtschaftspolitische Maßnahmen gemessen werden, die es beispielsweise erlauben würden, die jeweiligen Auswirkungen auf die Konjunktur abzuschätzen. Für die makroökonomische Analyse nützliche Informationen könnte die gezielte Auswertung der Nachrichten von bestimmten Nutzergruppen, wie Unternehmen oder Zeitungen, liefern. Hier könnten Twitter-Nachrichten auch komplementär zu anderen Datenquellen, wie Presseartikeln oder anderen sozialen Medien, zum Einsatz kommen. Allerdings liegen hierzu noch keine Erfahrungswerte vor. Auch wären solche Auswertungen mit einem erheblichen Aufwand verbunden, da bisher noch keine automatisierten Methoden für eine solche Zuordnung der Nutzer vorliegen.

Die Auswertung von Twitter-Daten ist mit einem hohen Aufwand verbunden. Im Vergleich zu anderen sozialen Medien bietet Twitter aber einige Vorteile, insbesondere wenn Analysen verstetigt werden sollen, beispielsweise um aus den Daten gebildete Konjunkturindikatoren laufend zu aktualisieren. So

ist die tägliche Aktualisierung vergleichsweise komfortabel möglich. Zudem enthalten die Daten einen recht hohen Anteil an strukturierten Informationen, und die Texte der Nachrichten sind aufgrund des standardisierten Formats wohl zum Teil leichter auszuwerten. Sofern die Daten selbst gesammelt werden sollen, ist für viele makroökonomische Fragestellungen allerdings auch eine recht lange Vorlaufzeit nötig, um eine geeignete Datengrundlage zu schaffen

2.3 Internetsuchanfragen

Internetsuchanfragen bilden das Interesse von Personen an bestimmten Themen ab und können in diesem Zusammenhang nützliche Informationen über ihre Kaufabsichten, Erwartungen, Zuversicht oder Unsicherheit liefern. Da Suchanfragen zudem mittlerweile von einem Großteil der Bevölkerung verwendet werden und Daten dazu zeitnah verfügbar sind, waren sie bereits Grundlage für eine Vielzahl von makroökonomischen Analysen. Insbesondere wurden sie bislang für die Prognose des privaten Konsums und der Arbeitsmarktentwicklung herangezogen. In diesem Zusammenhang ist auch geprüft worden, ob sie das Konsumentenvertrauen abbilden oder prognostizieren können. Die vorliegenden Studien verwenden bislang praktisch ausschließlich Suchanfragen bei Google als Datengrundlage.

2.3.1 Datenquellen

Google stellt eine Vielzahl von Daten öffentlich bereit. Mittels Google Trends kann die Häufigkeit verschiedener Suchanfragen im Internet analysiert werden. Google Correlate erlaubte es Nutzern, Suchbegriffe zu identifizieren, welche mit anderen Suchbegriffen oder vorgegebenen Zeitreihen korrelieren, wurde aber im Dezember 2019 eingestellt. Die dominante Marktstellung von Google spiegelt sich in einem Marktanteil von mehr als 80 Prozent auf Desktopgeräten und 97 Prozent auf Mobil- und Tablet-Geräten wider (Böhme et al. 2020). Der Umstand, dass lediglich Nutzer von Google erfasst werden, schränkt die Repräsentativität von dieser Seite her somit kaum ein. Für Unternehmen bzw. Betreiber einzelner Internetseiten ist zudem Google Analytics interessant, da hier die Besucherstruktur hinsichtlich der Verweildauer sowie in Bezug auf zuvor aufgesuchte Internetseiten oder verwendete Suchmaschinen aufgeschlüsselt wird.

2.3.2 Datenbeschaffenheit

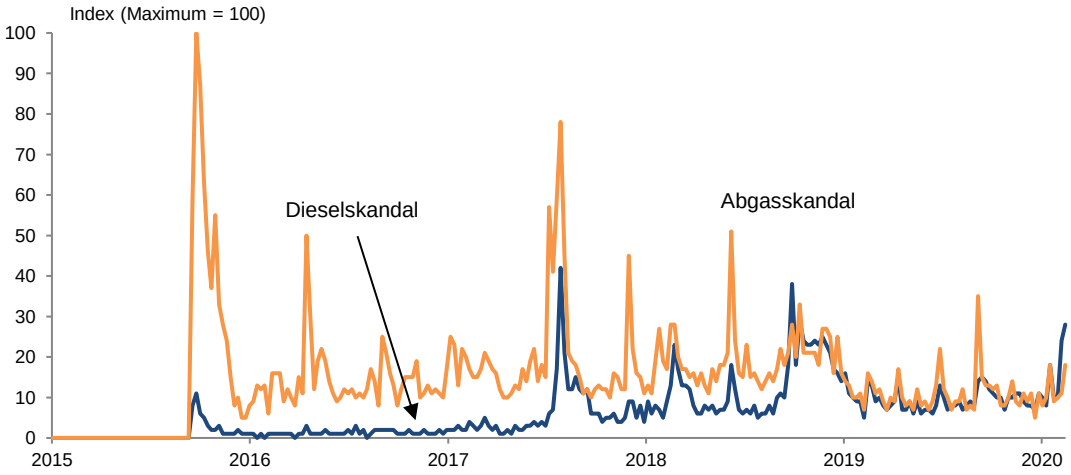
Mit Google Trends kann die Häufigkeit von Begriffen im Rahmen von Suchabfragen ermittelt werden; die Rohdaten werden hierbei aber nicht zur Verfügung gestellt. Zunächst wird die relative Häufigkeit des Suchbegriffs als Verhältnis des Suchvolumens zur Gesamtzahl der regionalen Suchanfragen über den entsprechenden Zeitraum dargestellt. Die Ergebnisse erscheinen dann in einer Skala von 0 bis 100, und bilden die Popularität eines Themas im Vergleich zu allen Suchanfragen der jeweiligen Anfrage ab. Hilfreich ist oftmals, die Suche nach einem Wort auf eine der vorgegebenen Kategorien (z.B. „Autos und Fahrzeuge“) einzugrenzen, um nur Ergebnisse für die tatsächlich beabsichtigte Bedeutung des Worts zu erhalten. Gleichzeitig können bis zu fünf Suchanfragen verglichen werden, wobei jede Anfrage maximal 25 Begriffe beinhalten kann. Bei den Suchergebnissen erhält man separate Treffer für alle Begriffsgruppen der jeweiligen Suchanfrage. Bei “Trending Searches” werden zudem besonders beliebte Suchabfragen der letzten 24 Stunden dargestellt.

„Google Autocomplete“ identifiziert Wort-Kombinationen über eine Vervollständigung von Suchbegriffen und ist somit potenziell nützlich, um sinnvolle Suchkombinationen zu identifizieren. Suchanfragen können auch zeitlich und räumlich eingegrenzt werden: So kann beispielsweise die Häufigkeit

von Suchbegriffen innerhalb bestimmter Zeiträume und/oder innerhalb bestimmter Regionen analysiert werden. Die Daten sind für nahezu alle Länder verfügbar und werden auch für die Welt insgesamt bereitgestellt. Sie sind ab dem Jahr 2004 in unterschiedlichen Frequenzen abrufbar, wobei die aktuellste Information die jeweils letzte Stunde umfasst. Die Ergebnisse für historische Suchanfragen können jedoch durch Anpassungen der zugrundeliegenden Algorithmen durch Google beeinflusst sein sowie dadurch, dass die Ergebnisse jeweils auf einer Stichprobe basieren (Choi und Varian 2012). Abbildung 6 zeigt beispielhaft die relative Häufigkeit von Suchanfragen zu den Begriffen „Abgasskandal" und „Dieselskandal". Dies spricht dafür, dass die Manipulationen verschiedener Hersteller bezüglich des Abgasausstoßes ihrer Fahrzeuge für die Nutzer im Jahr 2015 von besonderem Interesse waren. Insbesondere im Jahr 2017 stieg das Interesse im Zuge drohender Fahrverbote für bestimmte Fahrzeuge nochmals, auch in Verbindung mit dem sogenannten Dieselskandal. Ausschläge bei diesen Suchanfragen können z.B. auf einen geringeren Absatz von Pkws bzw. von Dieselfahrzeugen hindeuten.

Die Daten werden ohne Zeitverzögerung bereitgestellt, sind also in Echtzeit als Zeitreihe verfügbar. Da der Datenzugang zudem vergleichsweise einfach ist, ist eine Verstetigung der Abfragen und der darauf aufbauenden Analysen unproblematisch.

Abbildung 6:
Google-Suchanfragen für die Begriffe „Abgasskandal" und „Dieselskandal"[a]



[a]Wochendaten. Alle Angaben relativ zum Maximum (=100) der beiden Reihen.

Quelle: Google Trends (2020).

2.3.3 Methodik

Insgesamt sind die von Google bereitgestellten Daten leicht und mit jeder gängigen Software zu verarbeiten. Besonders nützlich zur Analyse von Google-Daten ist die frei verfügbare Software R. Diese ermöglicht beispielsweise auch das Extrahieren von täglichen Daten. Zudem existieren zahlreiche Funktionen, wie z.B. *„gtrends"*, mit denen Google Trend-Daten eingelesen und ausgewertet werden können. Zusatzbefehle wie „geo" ermöglichen die Disaggregation der Daten, in diesem Fall nach Regionen. Der Aufwand ist somit vergleichsweise gering. Voraussetzung für eine erfolgreiche Nutzung ist die Identifikation von adäquaten Suchbegriffen, da diese vom Nutzer vorgegeben werden. Dazu können gezielt einzelne Suchbegriffe vom Nutzer vorgegeben werden. Eine Alternative ist im ersten Schritt eine Vielzahl von Suchbegriffen zu verwenden, um die Komplexität in einem zweiten Schritt mittels Modellauswahlkriterien oder gemeinsamen Komponenten zu reduzieren (Götz und Knetsch,

2019). Darüber hinaus bot „Google Correlate“ bis vor kurzem die Möglichkeit, Suchanfragen mit der höchsten Korrelation zu einer vorgegebenen Zeitreihe zu identifizieren.

2.3.4 Bisherige Anwendungen

Suchanfragen bei Google wurden zunächst vor allem in der Medizin verwendet, beispielsweise zur Vorhersage und Überwachung von Erkrankungen (Althouse et al. 2011; Ginsberg et al. 2009), wobei der Informationsgewinn aus der zeitnahen Verfügbarkeit der Daten resultiert.[13] Böhme et al. (2020) verwenden Suchanfragen zur Prognose von Migrationsströmen und testen diesen Ansatz anhand von Migrationsdaten der OECD. Sie zeigen, dass Google-basierte Zeitreihen, die das Interesse an migrationsbezogenen Begriffen wie „Visum“ oder „Arbeit“ messen, einen Prognosegehalt haben. Mit ihrer Hilfe lassen sich die Prognosen von etablierten Migrationsmodellen, die meist allein auf jährlich verfügbaren Makrovariablen aufbauen, verbessern und die teils mehrjährigen Verzögerungen bei der Veröffentlichung neuer internationaler Migrationsdaten somit überbrücken. Donadelli et al. (2018) identifizieren einen Zusammenhang von Unsicherheit, Migrationsängsten und makroökonomischer Entwicklung basierend auf Google-Daten. Bangwayo-Skeete und Skeete (2015) zeigen, dass die Verwendung von Suchanfragen die Prognosen von Tourismusströmen verbessert.

Es existieren auch bereits zahlreiche Anwendungen von Suchanfragen für makroökonomische Fragestellungen. Bislang sind sie hierbei insbesondere für die Prognose von makroökonomischen Variablen und die Konstruktion von Unsicherheits- und Stimmungsindikatoren verwendet worden. So konstruieren Castelnuovo und Tran (2017) auf Basis von Suchanfragen Unsicherheitsmaße für die Vereinigten Staaten und Australien. Die zugrunde liegende Annahme ist, dass Personen im Falle von erhöhter Unsicherheit nach Begriffen suchen, die mit zukünftigen, typischerweise von Unsicherheit geprägten Ereignissen zusammenhängen. Hierzu zählen beispielsweise „Bankenkrise“, „Aktienmarkt“ und „Schuldenpolitik“. Bontempi et al. (2016) verwenden identische Begriffe, um die Effekte von Unsicherheit basierend auf Google-Suchanfragen mit einem auf Zeitungsartikeln basierenden Indikator zu vergleichen. Ihre Ergebnisse zeigen, dass Suchanfragen einen Anstieg von Unsicherheit für einige Kategorien früher als Zeitungen anzeigen. Dies ist beispielweise bei finanzpolitischen Fragen der Fall, während bezüglich geldpolitischer Unsicherheit der anhand von Zeitungsartikeln gebildete Indikator früher ansteigt. Donadelli (2015) verwendet einen ähnlichen Ansatz, um wirtschaftspolitische Unsicherheit basierend auf Google-Daten zu modellieren. Donadelli und Gerotto (2019) analysieren Suchanfragen für Begriffe aus den Bereichen Umwelt, Sicherheit und Politik zur Konstruktion von Unsicherheitsindikatoren und identifizieren einen negativen Effekt der betreffenden Indizes auf die Wirtschaftsaktivität in den USA. Bilgin et al. (2019) konstruieren einen vergleichbaren Index für die Türkei.

Zahlreiche Studien nutzen Suchanfragen, um die Zuversicht und den zukünftigen Konsum von Verbrauchern zu analysieren bzw. zu prognostizieren. Della Penna und Huang (2009) analysieren die Einzelhandelsumsätze auf Ebene von US-Bundesstaaten sowie auf internationaler Ebene und identifizieren vier übergeordnete Suchkategorien bei Google Trends, die mit diesen Variablen korrelieren. Der aggregierte Index, der die identifizierten Kategorien gleichgewichtet, korreliert in hohem Maß (Korrelationskoeffizient von ca. 0,9) mit umfragebasierten Indikatoren für die Konsumentenzuversicht, die von der Universität Michigan oder dem Conference Board veröffentlicht werden. Die weiteren Ergebnisse zeigen, dass der Google-basierte Index als vorlaufender Indikator einen statistisch signifikanten Erklärungsgehalt für die Entwicklung von umfragebasierten Indizes der Konsumentenzuversicht aufweist. Niesert et al. (2019) finden für fünf Volkswirtschaften (Vereinigte Staaten, Vereinigtes Königreich,

[13] Jun et al. (2018) liefern einen umfassenden Überblick über die Verwendung von Google-Daten in verschiedenen wissenschaftlichen Disziplinen.

Kanada, Deutschland und Japan) demgegenüber allerdings nicht, dass Suchanfragen Prognosen für das Konsumentenvertrauen verbessern. Vosen und Schmidt (2011) verwenden Suchdaten für die Prognose des privaten Konsums für die Vereinigten Staaten. Ihren Ergebnissen zufolge sind internetbasierte Suchanfragen besser zur Prognose geeignet als die konventionellen, umfragebasierten Maße der Konsumentenzuversicht der Universität Michigan oder des Conference Boards. Basierend auf diesen Erkenntnissen veröffentlichte das RWI zeitweise einen wöchentlichen Konsumindikator für Deutschland, der auf Suchintensität bei Google zurückgreift, um Prognosen des privaten Konsums erstellen zu können (Schmidt und Vosen 2012). Die Studie von Kholodilin et al. (2010) stellt ebenfalls fest, dass auf Suchanfragen basierende Prognosen des privaten Konsums in den Vereinigten Staaten für den Nowcast genauer sind als die eines einfachen autoregressiven Vergleichsmodells. Eine Studie von Gil et al. (2018) vergleicht verschiedene Indikatoren zum Nowcasting des Konsums in Spanien. Neben traditionellen Indikatoren werden auch Unsicherheitsmaße, Kredikartendaten und Google-Suchabfragen verwendet. Ihre Ergebnisse zeigen, dass Google-basierte Indikatoren und Unsicherheitsindikatoren einen Mehrwert bieten, wenn sie mit traditionellen Indikatoren kombiniert werden. Zudem weisen Google-Suchabfragen auch eine gewisse Prognosegüte über den Nowcast hinaus auf. Choi und Varian (2012) zeigen, dass Google-Abfragen einen Prognosegehalt für die Entwicklung von Automobilverkäufen in den Vereinigten Staaten aufweisen.

Zahlreiche Studien basierend auf Google-Daten widmen sich der Prognose von Arbeitsmarktvariablen. Ausgangspunkt ist eine frühe Studie von Choi und Varian (2009), die zeigt, dass Google-Daten nützlich sind, um die Anmeldungen für Leistungen aus Arbeitslosigkeitsversicherungen zu prognostizieren. D'Amuri und Marcucci (2009) stellen fest, dass um Suchanfragen erweiterte Modelle der US-Arbeitslosenquote bessere Prognosen liefern als lineare und nicht lineare Zeitreihenmodelle oder als Umfragen unter professionellen Prognostikern. D'Amuri und Marcucci (2017) nutzen ebenfalls Google-basierte Modelle zur Prognose der amerikanischen Arbeitslosigkeit. Die korrespondierenden Modelle schneiden besser ab als die vierteljährlichen Prognosen von professionellen Prognostikern oder autoregressive Modelle. Lediglich Modelle basierend auf Erwartungen der Arbeitgeber weisen über einen Prognosehorizont von bis zu einem Jahr eine höhere Prognosegüte auf. Askitas und Zimmermann (2009) analysieren den Zusammenhang zwischen Suchabfragen und Arbeitslosigkeit in Deutschland. Sie identifizieren im Rahmen einer Kointegrationsanalyse einen kontemporären langfristigen Zusammenhang. Die Prognosegüte evaluieren die Autoren allerdings nicht. D'Amuri (2009) und Suhoy (2009) konstatieren ebenfalls, dass internetbasierte Daten die Prognosen der italienischen bzw. israelischen Arbeitslosigkeit verbessern. Baker und Fradkin (2017) konstruieren einen Index, der die Suche nach Arbeitsplätzen (Job Search Activity) in den Vereinigten Staaten basierend auf Google-Suchanfragen abbildet. Sie identifizieren einen signifikanten Effekt von Änderungen bei der Arbeitslosenversicherung auf derartige Anfragen. Insgesamt zeigen ihre Ergebnisse, dass Internetabfragen die Bedingungen am Arbeitsmarkt widerspiegeln. Die international angelegte Studie von Niesert et al. (2019) evaluiert auch monatliche Nowcasts für die Arbeitslosigkeit. Für jede Zeitreihe liegen den Prognosen 60 Google-Kategorien zugrunde. Ihren Ergebnissen zufolge verbessern Suchanfragen die Prognosen für die Arbeitslosigkeit in der Regel ebenfalls. Allerdings ist Deutschland das einzige Land, bei dem keine höhere Prognosegüte für die Arbeitslosigkeit verzeichnet wurde.

Im Zusammenhang mit BIP-Prognosen sind auch Nowcasting-Ansätze, bei denen Daten höherer Frequenz (z.B. täglich oder wöchentlich) für die Prognose von makroökonomischen Indikatoren niedrigerer Frequenz (monatlich oder quartalsweise) herangezogen werden, von hoher Relevanz, um die verfügbaren Informationen zeitnah in die Prognosen einfließen lassen zu können.

Eine aktuelle Studie von Götz und Knetsch (2019) greift auf Google-Daten zur Prognose des Deutschen Bruttoinlandsprodukts für verschiedene Prognosehorizonte zurück. Für die Prognose werden neben

Google-Suchabfragen auch weitere monatliche Indikatoren verwendet. Die vorliegenden Google-Daten werden dafür zu Monatswerten zusammengefasst. Die Prognose für das Bruttoinlandsprodukts setzt sich dabei aus der Prognose der Wertschöpfung in verschiedenen Wirtschaftszweigen zusammen Ausgangspunkt ist eine große Zahl von Zeitreihen aus Suchabfragen für 26 Kategorien und 269 Subkategorien, aus denen mittels unterschiedlicher Methoden ausgewählt wird. Die Suchanfragen werden alternativ oder komplementär zu Umfragedaten verwendet, um zunächst sogenannte „harte" Frühindikatoren (wie die Industrieproduktion) fortzuschreiben, aus denen sich wiederum die Wertschöpfungsprognosen für die Wirtschaftszweige und die BIP-Prognose ergeben. Ihre Ergebnisse zeigen, dass Suchanfragen eine gute Alternative zu Umfragen des Ifo Instituts darstellen. Allerdings verbessern sie Modelle, die bereits die Umfrageindikatoren enthalten, nicht mehr systematisch. Insbesondere für kurze Prognosehorizonte schneiden die um Umfrage- und/oder Google-Daten erweiterten Modelle zwar tendenziell nicht so gut ab. Insgesamt stellen die Google-Daten aber zumindest eine relevante Alternative zu Umfragedaten dar.

Eine aktuelle Studie von Ferrara und Simoni (2019) zeigt, dass Suchabfragen BIP-Nowcasts für den Euroraum vor allem aufgrund ihrer schnelleren Verfügbarkeit verbessern können. Dies ist in den ersten vier Wochen eines Quartals, wenn noch keine konventionellen Frühindikatoren für das Quartal vorliegen, der Fall. Im weiteren Verlauf des Quartals und mit zunehmender Verfügbarkeit anderer Indikatoren ergibt sich allerdings keine zusätzliche Prognoseverbesserung durch die Suchanfragen. Die Ergebnisse verdeutlichen außerdem, dass viele potenziell relevanten Suchbegriffe eine vergleichsweise geringe Korrelation mit dem BIP aufweisen. Ferrara und Simoni (2019) schlagen daher vor, die beim Nowcasting verwendeten Suchbegriffe ex ante auszuwählen bzw. zu begrenzen.

Tabelle 1 fasst die in ausgewählten Studien durchgeführte Art der Identifikation von relevanten Suchbegriffen zusammen. Es wird deutlich, dass in den Studien zumeist im ersten Schritt eine Vielzahl von Suchbegriffen ausgewählt wird, um die relevanten Informationen im zweiten Schritt mittels statistischer Verfahren zu verdichten.

Insgesamt weisen Google-Suchanfragen für verschiedene Variablen einen gewissen Prognosegehalt auf. Im Zusammenhang mit Prognosen für realwirtschaftliche Größen bleibt allerdings letztlich unklar, in welchem Ausmaß Google-Suchanfragen die Prognosen im Vergleich zu anderen Frühindikatoren tatsächlich verbessern. Zahlreiche Studien verwenden nämlich vergleichsweise einfache univariate An-

Tabelle 1:
Klassifizierung von Suchbegriffen

Studie	Verwendeter Suchbegriff
Baker und Fradkin (2017)	Verwendung des Begriffs „Jobs" Analyse von verwandten Begriffen auf Google Trends
Ferrara und Simoni (2019)	26 Kategorien und 269 Subkategorien aus Google Trends Auswahl der relevanten Suchbegriffe basierend auf der SIS-pre-selection-Methode im zweiten Schritt.
Götz und Knetsch (2019)	26 Kategorien und 269 Subkategorien aus Google Trends; Auswahl der relevanten Suchbegriffe: (1) Ad-hoc, (2) Google Correlate, (3) Gemeinsame Komponenten, (4) Shrinkage-Methode (LASSO)
Schmidt und Vossen (2011)	Nicht beobachtbare Faktoren, identifiziert durch Hauptkomponenten der Suchbegriffe bei Google Trends. Verwendung von 605 Kategorien
Niesert et al. (2019)	Verwendung von Google Correlate für jede makroökonomische Zeitreihe zur Identifikation von maximal 50 positiv und negativ korrelierten Suchabfragen.

Quelle: Eigene Zusammenstellung.

sätze als Vergleichsmaßstab, häufig aber nicht Alternativmodelle, die alle relevanten Frühindikatoren enthalten. Um eine Aussage hinsichtlich der Verbesserung der Prognosegüte bereits vorliegender Prognosemodelle treffen zu können, müsste jedoch gerade ein Vergleich mit den bestmöglichen Indikatoren bzw. Modellen vorgenommen werden.

Für die Prognose von Finanzmarktvariablen ist dieser Umstand weniger relevant, da ein Zufallsprozess als Referenz für Prognosen relativ etabliert ist. Hier waren bislang insbesondere Wechselkurse, Aktienpreise und Gold Gegenstand der Analyse. Andrade et al. (2013) verwenden Suchdaten, um Ursachen des starken Anstiegs der chinesischen Aktienkurse im Jahr 2007 zu identifizieren. Ihre Ergebnisse zeigen, dass dem Platzen der Aktienblase ein ungewöhnlich hohes Interesse von Privatanlegern, welches sich in Suchabfragen widerspiegelt, vorausging.

Preis et al. (2013) stellen fest, dass die Suche nach bestimmten Firmennamen mit dem Transaktionsvolumen der Aktien dieser Unternehmen korreliert. Die Ergebnisse von Vlastakis und Markellos (2012) zeigen in diesem Zusammenhang, dass Suchabfragen neben den Transaktionsvolumen auch die Volatilität beeinflussen. Da et al. (2011a) konstruieren einen "Investor Attention"-Index mit Hilfe von Internetabfragen, welcher Aktienkurse gut prognostiziert. Sie kommen in einer weiteren Studie zu dem Schluss (Da et al. 2011b), dass Suchdaten hinsichtlich des Produkts eines Unternehmens besser abschneiden als die Prognosen von Analysten. Insgesamt zeigt sich jedoch, dass Finanzmarktprognosen im Vergleich zu etablierten Vergleichsmethoden, wie der simplen Annahme eines unveränderten Kurses, sich durch Suchanfragen nicht systematisch verbessern lassen. Dieses Ergebnis steht im Einklang mit der bestehenden empirischen Literatur.

Neben der Verwendung für Prognosen greifen einige Studien auch auf Google-Daten zurück, um Schocks zu identifizieren. Beispielsweise verwenden Demir et al. (2020) Suchanfragen für die Türkei, um einen nicht erwarteten Schock in Form von Steuererhöhungen zu identifizieren, indem sie zeigen, dass es zuvor keine erhöhte Zahl von Suchanfragen bei Google zu dieser Steuer gab.

2.3.5 Potenziale und Grenzen

Internetsuchanfragen weisen zahlreiche für makroökonomische Analysen nützliche Eigenschaften auf. Sie können relevante Informationen über die Kaufabsichten, die Zuversicht oder die Unsicherheit von Personen liefern. Sie werden regelmäßig von einem Großteil der Bevölkerung eingesetzt. Zudem sind sie zeitnah verfügbar und können vergleichsweise leicht ausgewertet und stetig aktualisiert werden. Aus diesen Gründen sind sie bereits vielfach für makroökonomische Analysen und insbesondere für Prognosen eingesetzt worden.

Die vorliegenden Studien zeigen, dass Suchanfragen insbesondere geeignet sind, um die privaten Konsumausgaben zu prognostizieren, da sie die Kaufabsichten und Erwartungen von Personen widerspiegeln. Hier ist auch das Potenzial, konventionelle Prognosemodelle zu verbessern, recht hoch, da nur wenige andere zuverlässige Indikatoren für die Kurzfristprognose des privaten Konsums zur Verfügung stehen. Die bereits vorliegenden Ergebnisse deuten darauf hin, dass Suchanfragen auch für die Prognose der Arbeitsmarktentwicklung geeignet sein könnten; hier werden die amtlichen Informationen jedoch bereits rascher veröffentlicht als beispielsweise die Volkswirtschaftlichen Gesamtrechnungen (VGR). Zudem läuft der Arbeitsmarkt in der Tendenz der konjunkturellen Entwicklung eher nach, sodass hier das Verbesserungspotenzial insgesamt geringer ist. Jüngere Studien sprechen dafür, dass Suchanfragen auch die Prognosen des Bruttoinlandsprodukts verbessern können. Hier wäre es interessant zu verstehen, ob Suchanfragen neben den privaten Konsumausgaben und damit zusammenhängend dem Konsumentenvertrauen auch noch Informationen über andere Komponenten des Bruttoinlandsprodukts liefern können. Dazu könnten beispielsweise Indikatoren für die Unsicherheit zählen, die

mittels Suchanfragen zeitnäher erstellt werden können als andere gebräuchliche Indikatoren. Schließlich können Google-Suchanfragen auch dazu dienen, makroökonomisch relevante Ereignisse zu identifizieren, die nicht von den Nutzern antizipiert worden sind, um darauf aufbauend ihre wirtschaftlichen Auswirkungen zu analysieren. Hierfür sind Suchanfragen bislang noch kaum genutzt worden.

Problematisch an Google-Suchanfragen ist, dass sie strukturelle Brüche aufweisen können, wenn sich der zugrunde liegende Suchalgorithmus ändert (Lazer et al. 2014). Ferner können einmal erhaltene Ergebnisse nicht exakt repliziert werden, da die Daten auf einer zufälligen Stichprobe und nicht auf der Gesamtheit aller Suchanfragen beruhen. Die Ergebnisse von Böhme et al. (2020) legen allerdings nahe, dass die resultierenden Revisionen nicht zu substanziellen Änderungen führen. Schließlich kann sich auch der Datenzugang bei Google ändern; so wird der Dienst Google Correlate seit Dezember 2019 nicht mehr bereitgestellt. Interessant wäre es auch, Suchanfragen, die typisch für die unternehmerische Aktivität sind, auszuwerten. Diese dürften aber aufgrund der deutlich größeren Anzahl privater Suchanfragen selbst für sehr spezifische Suchbegriffe bzw. Kombinationen von Begriffen nur schwer zu identifizieren sein.

Alles in allem stellen Suchanfragen eine nützliche Datenquelle für makroökonomische Analysen und insbesondere Prognosen dar. Die vorliegenden Ergebnisse deuten darauf hin, dass sie eher eine komplementäre Funktion zu den gängigen Frühindikatoren haben. Ihr Vorteil liegt wohl vor allem darin, dass sie früher verfügbar sind als andere Indikatoren.

2.4 Online-Handelsplattformen und Scannerdaten

Online-Handelsplattformen und Scannerdaten liefern zeitnah detaillierte Informationen vor allem über Güterpreise. Aus diesem Grund werden sie bereits in der amtlichen Statistik zur Erhebung der Verbraucherpreise eingesetzt. Auch sind diese Daten schon regelmäßig für makroökonomische Analysen eingesetzt worden. Dabei standen bislang die Prognose der Inflation (insbesondere der Nowcast) sowie internationale makroökonomische Fragestellungen im Vordergrund, für die insbesondere Online-Handelspreise in hohem Detailgrad vergleichbare Daten liefern können. Neben Preisanalysen können Daten von Online-Handelsplattformen auch für andere Fragestellungen herangezogen werden. So sind sie bereits dafür verwendet worden, um die Auswirkungen von Steueränderungen auf im Internet verkaufte Waren (Einav et al. 2014) oder Preissetzungs- sowie Verkaufsstrategien zu analysieren (Einav et al. 2013a; Einav 2013b). Ferner wurden bereits Daten von Online-Plattformen für Fahrgemeinschaften ausgewertet, um ethnische Diskriminierung zu studieren (Tjaden et al. 2018). Solche Studien stellen bislang allerdings eher die Ausnahme dar.

2.4.1 Datenquellen

Daten von Online-Handelsplattformen lassen sich mit Web-Scraping-Methoden in Echtzeit erheben. Damit können im Internet öffentlich verfügbare Daten erfasst, aufbereitet und gesammelt werden. (Edelman 2012).

Die Nutzung von Preisdaten auf Online-Handelsplattformen in der makroökonomischen Forschung geht wesentlich auf das „Billion Prices Project“ (BPP) zurück (Cavallo und Rigobon 2016). Im Rahmen dieses Projekts werden seit dem Jahr 2007 in großem Umfang Preise von Online-Händlern in vielen Ländern erfasst. Die aus dem Projekt resultierenden Daten sind in vielen akademischen Beiträgen verwendet worden; diese Daten sind auch für andere Forschungsvorhaben frei verfügbar. Die Firma PriceStats stellt den kommerziellen Arm des BPP dar und bietet kostenpflichtig tägliche Preisindizes sowie Zeitreihen

zur Kaufkraftparität auf Basis von Online-Preisen an. Für konkrete Forschungsvorhaben können diese Daten aber mitunter auch kostenfrei bezogen werden. Ferner kann der Datenzugang über direkte Kooperationen mit Plattformbetreibern erfolgen. Dies stellt bislang jedoch eher eine Ausnahme dar. Beispielsweise hat das RWI in Zusammenarbeit mit der Vermittlungsplattform ImmobilienScout24 regionale Immobilienpreisindizes entwickelt (an de Meulen et al. 2011).

Scannerdaten können von privaten Anbietern, wie z.B. den Marktforschungsinstituten Nielsen oder GfK, kostenpflichtig bezogen werden. Die Datensätze decken in der Regel eine Vielzahl von Produkten, Verkäufern und Regionen ab. Vereinzelt werden auch Scannerdaten einzelner Handelsketten verwendet, z.B. des US-amerikanischen Einzelhändlers Target (Taylor 2009), allerdings dürften solche Daten bezüglich der Kunden oder der Produktpalette weniger repräsentativ sein.

2.4.2 Datenbeschaffenheit

Daten von Online-Handelsplattformen und Scannerdaten sind sehr zeitnah und in hoher Frequenz (z.B. Tagesdaten) verfügbar. Sie umfassen üblicherweise eine große Produktpalette. Die Datenqualität ist hoch, da tatsächliche Preise bzw. andere Merkmale von Produkten erfasst werden. Aus diesem Grund stellen sie auch für die amtliche Statistik eine nützliche Datenquelle dar (Bieg 2019).

Grundsätzlich können die Daten von Online-Handelsplattformen direkt per Web Scraping gesammelt werden. Da auf diesem Wege historische Daten allerdings nachträglich nicht mehr erfasst werden können, sind mitunter lange Vorlaufzeiten nötig, um eine für zeitreihenanalytische Auswertungen geeignete Datengrundlage zu schaffen. Eine Alternative stellen die im Rahmen des BPP erhobenen Daten dar, die prinzipiell frei zugänglich sind und für zahlreiche Länder, darunter auch Deutschland, vorliegen (Tabelle 2). Der zeitliche Umfang ist jedoch ebenfalls begrenzt; die Daten reichen frühestens bis in das Jahr 2008 zurück und sind nicht bis zum aktuellen Rand verfügbar, sodass eine Verstetigung von Analysen auf dieser Basis nicht möglich ist.

Tabelle 2:
Für Forschungsvorhaben genutzte Online-Preis-Datensätze

Studie	Daten/Quellen	Zeitraum
Cavallo (2017) [Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers]	1 200 Produkte der Hersteller wie Galeria Kaufhof, obi, real, REWE, Saturn untersucht.	2015 bis 2016
Cavallo und Rigobon (2016) [The Billion Prices Project: Using Online Prices for Measurement and Research]	Vorjahresrate des täglichen Verbraucherpreisindizes für Deutschland auf Basis von Onlinepreisen.	2010 bis 2015
Cavallo et al. (2014) [Currency Unions, Product Introductions, and the Real Exchange Rate]	Preise aller Produkte von Herstellern IKEA, ZARA, H&M und Apple	2008 bis 2015
Cavallo et al. (2018) [Using Online Prices for Measuring Real Consumption across Countries]	Tägliche Online-Preise für Nahrungsmittel und Getränke, Treibstoffe sowie Elektrogeräte großer deutscher Einzelhändler.	2010 bis 2017

Quelle: Eigene Zusammenstellung.

Über den Anbieter PriceStats haben Forscher Zugriff auf Verbraucherpreisindizes und Kaufkraftparitäten auf Tagesdatenbasis. Diese liegen aber offenbar auch nur mit einer gewissen Verzögerung vor (Cavallo und Rigobon 2016). Der Zugang zu den zugrunde liegenden Rohdaten war für Forschungszwecke in der Vergangenheit ebenfalls möglich.

Scannerdaten sind ausschließlich über kommerzielle Anbieter erhältlich. Auch hinsichtlich der Datenbeschaffenheit gibt es Unterschiede zu Online-Handelsplattformen. So werden die Daten in der Regel bereits vom Anbieter aggregiert und auf Wochen- oder Monatsdatenbasis zur Verfügung gestellt. Ein weiterer Unterschied besteht in dem Umfang des Datensatzes. So enthalten Scannerdatensätze üblicherweise neben Produktkennziffern (wie der Global Trade Item Number) auch Informationen zu Herstellern, Marken, nominalen Umsätzen und Absatzmengen (Blieg 2019). Dies ist ein wesentlicher Vorteil gegenüber Daten von Online-Handelsplattformen, da insbesondere die gehandelten Mengen für die Konjunkturanalyse von Bedeutung sein können, und es erlauben, absatzgewichtete Durchschnittspreise zu berechnen. Zudem lassen die Mengenangaben Rückschlüsse auf die Zusammensetzung des typischen Warenkorbs zu und ermöglichen es, Substitutionseffekte bei der Preismessung zu berücksichtigen.

2.4.3 Methodik

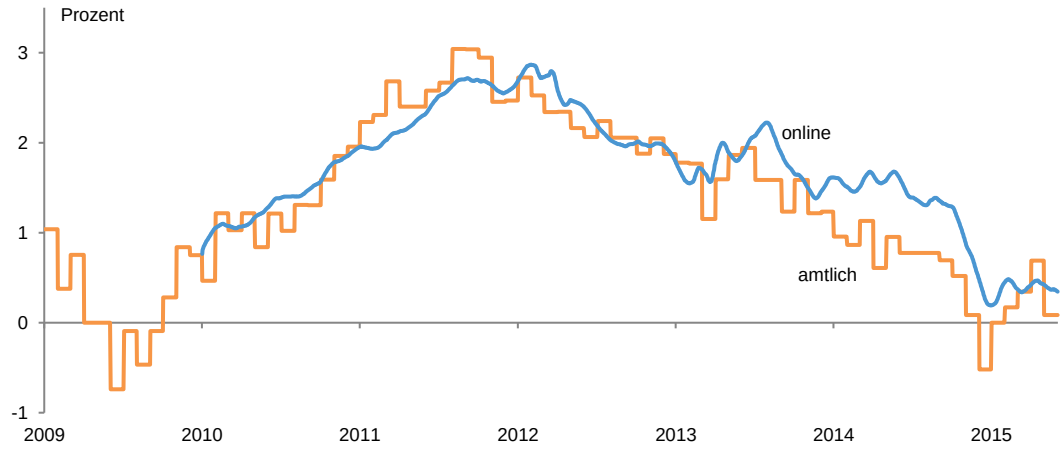
Die Organisation und Analyse der Daten von Online-Handelsplattformen oder Scannerdaten weisen grundsätzlich keine speziellen Anforderungen im Vergleich zu konventionellen makroökonomischen Zeitreihen, die in der Konjunkturanalyse und -forschung genutzt werden, auf. Lediglich die große Menge an Daten – angesichts der hohen Anzahl erfasster Produkte und der hohen Frequenz, mit der Preise insbesondere im Online-Handel erhoben werden können – stellt hinsichtlich Speicherkapazitäten, ökonometrischer Methodik oder Bearbeitung mit gängiger Software eine Herausforderung dar.

Für das Web Scraping sind prinzipiell keine besonderen IT-Kenntnisse erforderlich. Einfache Datenabfragen können mit browserbasierten Tools oder Programmiersprachen wie R oder Python bewältigt werden. Allerdings ist eine Verstetigung der Datenerhebung recht aufwendig. Hierfür sind unter Umständen spezifische IT-Kenntnisse und Softwarelösungen erforderlich, insbesondere um die Konsistenz der erhobenen Daten zu gewährleisten. Problematisch ist hier beispielsweise, wenn sich die zugrunde liegenden Internetseiten ändern und deshalb die Datenabfrage angepasst werden muss oder die Produkte ausgetauscht werden (Blaudow 2019). Wie umfangreich dieser Folgeaufwand ist, dürfte je nach Anwendung deutlich variieren. Seitens des Urheberrechts ist Web Scraping wohl unproblematisch, sofern die Erhebung auf eine Stichprobe begrenzt ist und kein wesentlicher Teil der jeweiligen Datenbank erfasst wird (Brunner 2014). In der Praxis ist es möglich, dass Online-Händler versuchen, Web Scraping zu verhindern, beispielsweise um die Preistransparenz einzuschränken oder Ladezeiten für Kunden zu minimieren. Die Tatsache, dass Web Scraping bereits von statistischen Ämtern oder Anbietern wie PriceStats intensiv genutzt wird, spricht jedoch dafür, dass dies für das Sammeln von Daten kein wesentliches Hindernis darstellt.

2.4.4 Bisherige Anwendungen

Daten von Online-Handelsplattformen und Scannerdaten sind bislang vor allen für Preisanalysen und -prognosen verwendet worden. In einer der ersten Analysen, die Daten von Online-Handelsplattformen nutzt, vergleicht Cavallo (2013) für fünf lateinamerikanische Länder Verbraucherpreisindizes auf Basis von online verfügbaren Supermarktpreisen mit amtlichen Zahlen. Nur in Argentinien weichen die so berechneten Inflationsraten wesentlich von der amtlichen Statistik ab und bestätigen damit den Verdacht, dass die argentinischen Behörden die Inflation seit dem Jahr 2007 systematisch zu niedrig ausgewiesen hatten. Darauf aufbauend zeigt Cavallo (2017), dass in einer Vielzahl von Ländern und einer breiten Produktpalette die Laden- und Online-Preise sehr ähnlich sind: Im Durchschnitt stimmen knapp 70 Prozent der „offline“ und online gemessenen Preisniveaus überein. Auch für Deutschland weisen die

Abbildung 7:
Verbraucherpreise und Onlinepreise in Deutschland[a]



[a]Tagesdaten; Veränderung gegenüber dem Vorjahr.

Quelle: Statistisches Bundesamt (2020b); Cavallo und Rigobon (2016).

online erhobenen Preise einen recht hohen Gleichlauf mit dem amtlichen Verbraucherpreisindex auf (Abbildung 7).

Der enge Zusammenhang zwischen täglich verfügbaren Online-Preisdaten und dem offiziellen Verbraucherpreisindex kann auch für Prognosen genutzt werden, insbesondere da Online-Preise schneller verfügbar sind. Aparicio und Bertolotto (2020) zeigen für eine Vielzahl von fortgeschrittenen Volkswirtschaften, dass Modelle, die Online-Preise enthalten, treffsicherere Inflationsprognosen liefern als Expertenbefragungen. Dies gilt für den Nowcast, also die Prognose der Inflationsrate des laufenden Monats. Dabei schneiden die Modelle mit Online-Daten auch für Deutschland recht gut ab und weisen Prognosefehler auf, die etwa zehn Prozent niedriger sind als von Modellen ohne Online-Preise. Auch für einen Prognosehorizont von bis zu drei Monaten, bei denen die zeitliche Verfügbarkeit eine geringere Rolle spielen dürfte, helfen Online-Preise, die Prognosen in einigen Ländern zu verbessern (unter anderem wiederum für Deutschland). Als Ursache für die Verbesserungen der Prognosen auch bei längeren Prognosehorizonten stellen die Autoren die These auf, dass neben methodischen Aspekten der Preisstatistik Preisanpassungen im Online-Handel mit weniger Kosten verbunden sind und sich eine Veränderung der grundlegenden Preisdynamik dort früher als bei den Ladenpreisen bemerkbar macht.

Weitere Arbeiten auf Basis der BPP-Daten beschäftigen sich mit Fragestellungen der (internationalen) Makroökonomie wie der Häufigkeit und der Verteilung von Preisänderungen, den Auswirkungen von internationalen Handelskonflikten und dem Gesetz des einheitlichen Preises. Hierzu nutzen Cavallo et al. (2014) die Verfügbarkeit von Online-Preisen für identische Produkte großer, multinationaler Konzerne wie IKEA, H&M und Zara, um zu untersuchen, unter welchen Bedingungen diese in verschiedenen Ländern zum selben Preis gehandelt werden. Ihren Ergebnissen zufolge kommt es in den meisten Ländern zu deutlichen Abweichungen vom Gesetz des einheitlichen Preises. Nur zwischen Ländern, die dieselbe Währung nutzen – wie im Euroraum oder in Ländern, deren Volkswirtschaften „dollarisiert" sind wie Ecuador – stimmen die Preise für identische Güter überein. Die Ursache dafür ist vermutlich, dass Kunden Preise dann leichter vergleichen können. Cavallo et al. (2019) untersuchen die Auswirkungen des im Jahr 2018 begonnenen Handelskonflikts zwischen den Vereinigten Staaten und der Volksrepublik China auf Warenpreise. Die Autoren verwenden dafür die im Vergleich zur amtlichen Statistik wesentlich detaillierteren Daten von Online-Handelsplattformen. Dadurch können sie exakt die Preise

der von Strafzöllen betroffenen Produkte analysieren. Sie stellen fest, dass es zwar im Zuge der Zollerhöhungen zu höheren Importpreisen in den Vereinigten Staaten gekommen ist, diese jedoch nur vereinzelt oder in geringerem Maße auch zu Preisanstiegen im Einzelhandel, sondern vorrangig zu geringeren Margen geführt haben.

Wie häufig und wie stark Unternehmen ihre Preise anpassen, spielt eine wichtige Rolle in makroökonomischen Modellen, nicht zuletzt, da das Ausmaß von Preisrigiditäten die Wirksamkeit der Geldpolitik beeinflussen kann. Cavallo (2018) zeigt, dass die Verteilung von täglichen Preisänderungen im Online-Handel bimodal ist, mit sehr vielen großen Preiserhöhungen und -senkungen und nur wenigen kleinen Preisanpassungen. Dies steht im Kontrast zu Verteilungen, die auf Basis amtlicher Verbraucherpreisindizes oder Scannerdaten geschätzt werden. Als Erklärung wird hierfür die Zeitaggregation bzw. die Messfrequenz angeführt: Während die Online-Preise tagesgenaue Preisänderungen abbilden, werden die Scannerdaten zum Teil bereits vom Anbieter auf Wochenfrequenz aggregiert. Dadurch können in den Daten fälschlicherweise häufiger kleine Preisänderungen ausgewiesen werden. Als Beleg für diese Hypothese wird gezeigt, dass bei Aggregation der Online-Daten auf die Wochenfrequenz die Verteilung deutlich symmetrischer wird und sich jener auf Basis von wöchentlichen Scannerdaten annähert.

Eichenbaum et al. (2014) zeigen anhand von Scannerdaten eines großen US-amerikanischen Einzelhändlers für den Zeitraum von 2004 bis 2006, dass kleine Preisänderungen in der Tat sehr selten vorkommen. Auch sie erklären die Abweichungen zu früheren Ergebnissen durch Probleme bei der Aggregation der Daten. Anhand eines Scannerdatensatzes, bei dem auch die tatsächlichen Ladenpreise erfasst wurden und die Preisreihen somit keine Messfehler aufweisen, zeigen die Autoren, dass die Anzahl der kleinen Preisanpassungen (definiert als Veränderungen geringer als fünf Prozent) um 80 Prozent niedriger als bei aggregierten Daten ausfällt.

Weitere Studien auf Basis von Scannerdaten befassen sich mit methodischen Aspekten der Preismessung, wie der Erfassung von Substitutionseffekten oder dem Einfluss von Qualitätsänderungen auf die Verbraucherpreisstatistik (Silver and Heravi 2005). In einer anderen Studie werden Preisunterschiede im Binnenmarkt des Euroraumes analysiert (EZB 2015). Sie stellt fest, dass es bei Lebensmitteln zu deutlichen Abweichungen von dem Gesetz des einheitlichen Preises und zu deutlicher Marktsegmentierung entlang nationaler Grenzen kommt. Einen Fokus auf gehandelte Mengen setzt Taylor (2009), der die kurzfristigen Auswirkungen der Finanzmarktturbulenzen im Zuge der Großen Rezession auf den privaten Verbrauch untersucht. Anhand von täglichen Scannerdaten eines US-amerikanischen Einzelhändlers analysiert er die gehandelten Mengen rund um den Kollaps der Investmentbank Lehman Brothers und kommt zu dem Schluss, dass es zwar zu einer Verlangsamung des Absatzes im Einzelhandel kam, dieser jedoch nicht mit Ereignissen an bestimmten Tagen in Verbindung gebracht werden kann, sondern sich vielmehr graduell vollzog. Einen ähnlichen Ansatz wählen Pandya und Venkatesan (2016), um die Auswirkungen des Irak-Kriegs im Jahr 2003 auf das Konsumverhalten zu untersuchen und insbesondere, wie erfolgreich Boykottaufrufe gegenüber Produkten oder Firmen aus Frankreich waren. Dazu verwenden sie einen Scannerdatensatz des Anbieters Information Resources Inc., der Transaktionen von über 1 000 Supermärkten erfasst und zeigen, dass Marken, deren Name Französisch klingt, signifikant Marktanteile verloren haben.

Daten von Online-Handelsplattformen und Scannerdaten werden bereits von einer Vielzahl von nationalen Statistikbehörden für die Messung der Verbraucherpreise eingesetzt. In Deutschland plant das Statistische Bundesamt die manuelle Erfassung von Online-Preisen für 10 000 Produkte, die bereits in die Verbraucherpreisstatistik einfließen, durch Web Scraping-Ansätze zu automatisieren und die Stichprobe der so erfassten Preise auszuweiten (Blaudow und Seeger 2019). Eine Herausforderung für die Methodik stellt dabei die größere Anzahl an unterschiedlichen Preisen pro Produkt und Monat dar

und wie diese aggregiert werden. Dafür müssten die Gewichte geschätzt werden, da online nur Preise, nicht aber Transaktionsvolumen erfasst werden.

Aus diesen Gründen hat das Statistische Bundesamt die Nützlichkeit von Scannerdaten anhand eines wöchentlichen Datensatzes der Firma Nielsen untersucht. Konkret wurden auf Basis der Scannerdaten Preisindizes für Lebensmittel und Getränke in den Jahren 2015 und 2016 erstellt und mit jenen der amtlichen Verbraucherpreisstatistik verglichen (Bieg 2019). Eine Herausforderung ist hierbei offenbar, die Vielzahl an Produkten den entsprechenden Unterkategorien des Verbraucherpreisindexes zuzuordnen. Traditionelle ökonometrische Ansätze wie eine logistische Regression oder Methoden aus dem Bereich des maschinellen Lernens scheinen für diese Aufgabe allerdings gut geeignet. Insgesamt weist die Untersuchung teils erhebliche Unterschiede in den Preisniveaus einzelner Güter aus. Als möglicher Grund hierfür wird angeführt, dass beispielsweise Sonderangebote mit Scannerdaten besser abgebildet werden können.

Ebenfalls mit der Preismessung, wenngleich nicht zu amtlichen Zwecken, beschäftigen sich an de Meulen et al. (2011), die auf Basis von Kauf- und Mietangeboten des Portals ImmobilienScout24 Immobilienpreisindizes berechnen. Die große Zahl an Beobachtungen sowie die detaillierten Informationen zu den angebotenen Häusern und Wohnungen erlauben es, Preisentwicklungen auf regionaler Ebene zu erfassen und dabei für Qualitätsunterschiede zu kontrollieren.

2.4.5 Potenziale und Grenzen

Daten von Online-Handelsplattformen und Scannerdaten werden bereits regelmäßig für makroökonomische Analysen genutzt. Für die amtliche Statistik können mit Web Scraping-Methoden Online-Preise automatisiert erfasst werden; sie erlauben es dabei auch, die Stichproben zu vergleichsweise geringen Kosten zu vergrößern. Ähnliches gilt für die Nutzung von Scannerdaten, wobei ein wesentlicher Vorteil darin besteht, dass anhand der gehandelten Mengen zum einen die tatsächlich gezahlten Preise besser gemessen und zum anderen Substitutionseffekte quantifiziert werden können. Für die Konjunkturanalyse eignen sich Online-Preise insbesondere für die zeitnahe Diagnose der Preisentwicklung. Bisherige Studien legen nahe, dass die Berücksichtigung von Online-Preisen zu exakteren Prognosen der Inflationsrate führen kann, wenngleich die Verbesserungen quantitativ überschaubar sind. Ein Hindernis für die Schätzung von Prognosemodellen stellt der kurze Zeitraum dar, für den Online-Preise vorliegen. Selbst wenn auf die Daten kommerzieller Anbieter wie PriceStats zurückgegriffen wird, reichen die Zeitreihen für Deutschland nur etwa zehn Jahre zurück. Neben dem technischen und fachlichen Aufwand sind kurze Beobachtungszeiträume auch ein wesentliches Hindernis für die eigene Erhebung von Online-Preisen, da sich die resultierenden Zeitreihen für viele ökonometrische Analysen erst nach erheblicher Vorlaufzeit nutzen lassen. Nichtsdestotrotz lassen sich Online-Preise für Querschnittsanalysen oder Event-Studien nutzen, da das Fehlen längerer Zeitreihen bei solchen Anwendungen weniger ins Gewicht fällt. Zusätzliche Potenziale könnten in der Nutzung von Daten von Online-Handelsplattformen, die bislang noch nicht ausgewertet wurden, liegen. So könnten branchenspezifische Portale wie MyHammer.de potenziell nützliche Informationen über bestimmte Wirtschaftsbereiche oder Verwendungskomponenten des Bruttoinlandsprodukts liefern oder Daten auf Online-Jobbörsen zur Analyse und Prognose des Arbeitsmarkts verwendet werden.

2.5 Elektronischer Zahlungsverkehr

Der elektronische Zahlungsverkehr umfasst monetäre Transaktionen, die elektronisch abgewickelt werden. Dazu zählen insbesondere Zahlungen, die per Giro- bzw. Kreditkarte vorgenommen werden und

Überweisungen. Daten zum elektronischen Zahlungsverkehr bilden tatsächlich getätigte Transaktionen – vielfach in konsumnahen Bereichen – umgehend ab und bieten somit ein hohes Potenzial, um die laufende Entwicklung der privaten Konsumausgaben und ggf. des Bruttoinlandsprodukts zu beobachten. In diesem Zusammenhang sind die Daten auch für die amtliche Statistik von Interesse, um die Einzelhandelsumsätze bzw. die Konsumaktivität rascher und umfassender zu messen. Da sie grundsätzlich tagesaktuell und regional disaggregiert erfasst werden, könnten mit Daten zum elektronischen Zahlungsverkehr darüber hinaus auch die Auswirkungen verschiedener makroökonomisch relevanter Ereignisse analysiert werden, die auf gesamtwirtschaftlicher Ebene oder in niedriger Frequenz nur schwer zu erfassen sind. Allerdings sind die Daten nicht frei verfügbar und der Zugang kann nur durch individuelle Vereinbarungen mit den jeweiligen Zahlungsdienstleistern erfolgen.

2.5.1 Datenquellen

Als Datenquelle für Bargeldabhebungen oder „point of sale“-Transaktionen kommen private Kartenanbieter in Betracht. Auf internationaler Ebene zählen dazu insbesondere Kreditkartenanbieter wie American Express, Mastercard und VISA. Ferner stellen Betreiber von Zahlungssystemen eine mögliche Datenquelle dar. Hier werden von Land zu Land jedoch unterschiedliche Zahlungssysteme genutzt. Bislang sind beispielsweise Daten des kanadischen Zahlungssystembetreibers Interac für makroökonomische Analysen verwendet worden (Galbraith und Tkazc 2007; 2013; 2015). Für Deutschland ist Girocard ein großer Anbieter von Bankkarten sowie Euro1/EBA ein gebräuchliches Zahlungssystem. Grundsätzlich könnten solche Daten auch von Geschäftsbanken zur Verfügung gestellt werden. Eine weitere Quelle ist der Telekommunikationsanbieter SWIFT, der von Banken und Brokern genutzt wird und dessen standardisierte Nachrichten Transaktionen weltweit erfassen.

2.5.2 Datenbeschaffenheit

Daten zum elektronischen Zahlungsverkehr werden umgehend erfasst und sind somit grundsätzlich in hoher Frequenz verfügbar (z.B. Tagesdaten). Ferner weisen sie keine Messfehler auf, sodass die Datenqualität von dieser Seite her im Vergleich zu anderen Datenquellen ausgesprochen hoch ist. Sie basieren auf tatsächlich getätigten Transaktionen oder in Bezug auf Bargeldabhebungen in der Regel zeitnah geplanten Transaktionen. Aufgrund der hohen Bedeutung von Kartenzahlungen oder Überweisungen für die Zahlungsvorgänge insgesamt bilden sie einen wesentlichen Teil aller getätigten monetären Transaktionen ab. In Deutschland lag der Anteil der Giro- oder EC-Karten-Umsätze am gesamten Zahlungsverkehrsumsatz im Jahr 2017 bei 35 Prozent; Kreditkarten spielen mit sechs Prozent eine geringere Rolle (Deutsche Bundesbank 2019). Sie bilden somit auch einen beträchtlichen Teil der von privaten Haushalten getätigten Konsumausgaben ab. Insbesondere in diesem Zusammenhang können sie auch wertvolle Informationen zu den Umsätzen in den Dienstleistungsbranchen oder für die gesamtwirtschaftliche Aktivität liefern. Grundsätzlich beinhaltet der elektronische Zahlungsverkehr auch Unternehmenstransaktionen; die Daten könnten somit auch diesbezüglich eine wertvolle Informationsquelle darstellen. Hierfür liegen allerdings noch keine Anwendungsbeispiele vor.

Daten zum elektronischen Zahlungsverkehr bilden getätigte Transaktionen umfassend ab und sind von dieser Seite somit auch repräsentativ. Allerdings werden die unterschiedlichen Zahlungsformen unterschiedlich stark genutzt. So werden Kreditkartenzahlungen in Deutschland beispielsweise auch im internationalen Vergleich recht wenig verwendet. Zudem hat die Nutzung der unterschiedlichen Zahlungsformen im Zeitablauf stark variiert. Beispielsweise sind Kartenzahlungen in den vergangenen Jahrzehnten in vielen Ländern zunehmend beliebter geworden. Hinzu kommt, dass sofern nur auf die Daten eines einzelnen Anbieters zurückgegriffen werden kann, der Umfang der erfassten Transaktionen auch

mit der jeweiligen Marktposition variiert. In diesem Zusammenhang können auch Veränderungen des Kundennetzwerks, beispielsweise infolge von Firmenübernahme, Analysen erschweren (Aladangady et al. 2019). Schließlich können auch Gebührenanpassungen, Abgabenänderungen oder regulatorische Änderungen die Datenqualität beeinflussen. Insgesamt hängt die Repräsentativität der zur Verfügung stehenden Daten somit auch jeweils von der konkreten Datenquelle ab.

Daten zum elektronischen Zahlungsverkehr sind nicht frei verfügbar. Ihre Nutzung muss mit dem jeweiligen Anbieter vereinbart werden und kann mit erheblichen Kosten verbunden sein. Bislang sind sie für makroökonomische Analysen vornehmlich als bereits aggregierte Zeitreihen verwendet worden. Sofern disaggregierte Daten verwendet werden, kann dies mit einem erheblichen Zusatzaufwand verbunden sein. So sind SWIFT-Nachrichten häufig sehr umfangreich und erfassen über 100 verschiedene Arten von Transaktionen. Dies stellt nicht nur besondere Anforderungen an die IT-Kapazitäten und -Kenntnisse der Anwender, sondern ggf. auch an die Einhaltung von Datenschutzrichtlinien (IMF 2017).

2.5.3 Methodik

Sofern die Daten aggregiert zur Verfügung gestellt werden, können sie grundsätzlich mit den gängigen empirischen Methoden ausgewertet werden. Allerdings sind einige der Probleme, die sich regelmäßig auch bei den konventionellen makroökonomischen Variablen ergeben, bei Daten zum elektronischen Zahlungsverkehr deutlich ausgeprägter und müssen ggf. explizit methodisch berücksichtigt werden.

Methodische Schwierigkeiten können entstehen, wenn Tagesdaten für die Analyse verwendet werden sollen, da Umsätze des elektronischen Zahlungsverkehrs ein ausgeprägtes Wochenmuster haben. So werden üblicherweise an Freitagen und Samstagen deutlich mehr Transaktionen abgewickelt als unter der Woche. Saison- und Kalenderbereinigungsverfahren sind vorwiegend für Daten auf Monats- oder Quartalsdatenbasis entwickelt und erprobt worden. Da Tageszeitreihen besonders volatil sind, und mehr Ausreißer und Strukturbrüche aufweisen, ist für eine entsprechende Bereinigung eine Vielzahl von Parametern zu schätzen (Ladiray et al. 2018). Ollech (2018) hat ein Verfahren entwickelt, das unter anderem solche Wocheneffekte erfasst, und wendet dies zur Saisonbereinigung des täglichen Bargeldumlaufs an.

In bisherigen makroökonomischen Analysen wurde daher oft mit Zeitreihen gearbeitet, die auf Monatsdatenebene aggregiert worden sind. Zwar liegen Daten zum elektronischen Zahlungsverkehr auch dann noch zeitnäher vor als konventionelle Indikatoren wie die Einzelhandelsumsätze. Jedoch geht dadurch auch ein bedeutender Vorteil – nämlich die zeitnahe tägliche Verfügbarkeit – verloren. Eine Lösung für dieses Problem stellen möglicherweise zeitvariierende oder regimeabhängige MIDAS-Ansätze dar, mit denen Variablen mit unterschiedlichen Frequenzen (also zum Beispiel Tagesdaten zum elektronischen Zahlungsverkehr und monatliche Einzelhandelsumsätze) in einem einheitlichen Modellrahmen ausgewertet werden können. Dazu können diese Ansätze dahingehend weiterentwickelt werden, dass die auf höherer Frequenz vorliegenden erklärenden Variablen je nach Zeitpunkt unterschiedliche Koeffizienten aufweisen (Lehrer et al. 2019). Somit würden beispielsweise bestimmte Transaktionsvolumina an Wochenenden oder unter der Woche mit unterschiedlichen monatlichen Umsatzentwicklungen assoziiert werden.

Eine weitere methodische Hürde rührt daher, dass ein etwaiger Zusammenhang von Daten zum elektronischen Zahlungsverkehr mit makroökonomischen Variablen je nach der konkreten Datenquelle und des Analysezeitraums im Zeitablauf großen Schwankungen unterworfen sein könnte. Bisherige Arbeiten haben das Problem umgangen, indem die Zeitreihe der Transaktionen um einen Trend bereinigt wurde oder der Stützzeitraum entsprechend verkürzt wurde. Allerdings werden durch Letzteres empirische Auswertungen erschwert. Zudem stellt sich dann insbesondere die Frage, inwie-

weit der geschätzte Zusammenhang auch zukünftig variieren wird, was Aussagen über die Prognosegüte erschwert.

2.5.4 Bisherige Anwendungen

Daten zum elektronischen Zahlungsverkehr sind bislang vor allem zur Prognose, insbesondere dem Nowcast, der privaten Konsumtätigkeit sowie zur Analyse der Auswirkungen makroökonomisch bedeutsamer Ereignisse verwendet worden.

Mehrere Studien nutzen Daten des kanadischen Zahlungssystembetreibers Interac zu Debitkartentransaktionen. So zeigen Galbraith und Tkacz (2007), dass die auf Quartalsebene aggregierten, vom Betreiber erfassten Umsätze (Transaktionswerte) und die aggregierte Anzahl der Transaktionen (Transaktionsvolumen) Consensus-Prognosen für den privaten Konsum und das Bruttoinlandsprodukt hätten verbessern können. Allerdings umfasst der Evaluierungszeitraum nur fünf Jahre, sodass sich die Ergebnisse nur schwer verallgemeinern lassen. In einer weiteren Studie analysieren die Autoren auf Basis desselben Datensatzes die ökonomischen Auswirkungen von Terroranschlägen und Epidemien (Galbraith and Tkacz 2013). Hierzu nutzen sie die tägliche Verfügbarkeit der Zahlungsverkehrsdaten, um im Stil einer Event-Studie Veränderungen des Transaktionsvolumens nach bedeutsamen Ereignissen (wie den Terroranschlägen vom 11. September 2001 oder den SARS-Ausbruch im Jahr 2003) abzuschätzen. Eine Schwierigkeit stellen in diesem Zusammenhang die starken innerwöchentlichen Schwankungen der Transaktionen dar. Die Autoren verwenden deshalb zum Vergleich Perioden mit einer vergleichbaren Abfolge von Wochentagen. Galbraith und Tkacz (2015) verwenden zusätzlich die Daten für Kreditkartentransaktionen und überprüfen die Prognosegüte dieser Variablen für das Bruttoinlandsprodukt in Kanada. Generell schneiden Modelle, die Daten zum elektronischen Zahlungsverkehr als zusätzliche erklärende Variable aufnehmen, nicht wesentlich besser ab als solche, die nur die Arbeitslosenquote und den Composite-Leading-Indikator der OECD verwenden. Innerhalb der Zahlungsverkehrsdaten weisen Transaktionen mit Debitkarten zwar die höchste Prognosegüte auf. Aber auch hier sind die Verbesserungen gegenüber dem Vergleichsmodell gering und treten nur für Nowcasts auf, die zu Beginn des zu prognostizierenden Quartals erstellt werden; zu späteren Zeitpunkten verbessert die Berücksichtigung von Zahlungsverkehrsdaten die Prognosen nicht mehr.

Weitere Untersuchungen zu den Prognoseeigenschaften des elektronischen Zahlungsverkehrs sind hauptsächlich von Zentralbanken durchgeführt worden. So nutzen Aprigliano et al. (2019) elektronische Transaktionen zur Prognose des italienischen Bruttoinlandsprodukts, des privaten Verbrauchs, der Bruttoanlageinvestitionen und der Bruttowertschöpfung im Dienstleistungsgewerbe. Als Datenquelle dienen das von der italienischen Zentralbank betriebene Zahlungsverkehrssystem BI-COMP sowie das vom Eurosystem betriebene TARGET-2. Die auf Monatsebene aggregierten Transaktionswerte von Giro- und Kreditkarten beider Systeme werden mit einem makroökonomischen Datensatz kombiniert und daraus Faktoren geschätzt. Darauf aufbauend werden im Rahmen eines MIDAS-Ansatzes Prognosen erstellt und evaluiert. Insgesamt zeigt sich hinsichtlich der Prognosegüte kein klares Bild: So schneidet das Modell, das die Transaktionswerte enthält, für das Bruttoinlandsprodukt im Zeitraum vom dritten Quartal 2011 bis zum zweiten Quartal 2015 für den Nowcast deutlich besser ab als ein Modell ohne Transaktionswerte. Schließt man die große Rezession allerdings in den Evaluierungszeitraum mit ein, so kehrt sich das Ergebnis um und die Modelle ohne Transaktionswerte liefern treffsicherere Prognosen. Auch für die übrigen Variablen und für längere Prognosehorizonte lassen sich zwar meist durch die Verwendung von Transaktionsdaten geringe Verbesserungen der Prognosegüte erzielen. Signifikante Unterschiede sind allerdings – wohl auch wegen des kurzen Evaluierungszeitraums – kaum festzustellen.

Ökonomen der niederländischen Zentralbank (DNB) untersuchen die Prognoseeigenschaft von Girokartentransaktionen hinsichtlich des privaten Verbrauchs (Verbaan et al. 2017). Neben MIDAS-Regressionen nutzen die Autoren auch eine mit Transaktionswerten erweiterte Version des DNB-eigenen Modells DELFI zur Prognose des privaten Verbrauchs. Eine kombinierte Prognose der beiden Ansätze liefert Verbesserungen von etwa 20 Prozent für die Prognose des privaten Verbrauchs im jeweils kommenden Quartal.

Duarte et al. (2016) vergleichen die Prognosegüte einer Vielzahl von verschiedenen MIDAS-Spezifikationen – sowohl mit täglichen als auch monatlichen Transaktionsdaten – für den privaten Konsum in Portugal. Die Modelle mit Tagesdaten schneiden generell schlechter ab als Modelle, die auf Einzelhandelsumsätzen oder dem Verbrauchervertrauen basieren. Mit monatlichen Transaktionsdaten können hingegen für Prognosen des laufenden Quartals deutliche Verbesserungen erzielt werden.

Gil et al. (2018) gehen der Frage nach, welchen zusätzlichen Wert alternative Datenquellen gegenüber herkömmlichen „harten" Indikatoren (wie Einzelhandelsumsätze oder sozialversicherungspflichtig Beschäftigte) und „weichen" Indikatoren (wie Einkaufmanagerindizes) für die Prognose des privaten Verbrauchs in Spanien liefern. Als alternative Datenquellen verwenden sie den elektronischen Zahlungsverkehr, verschiedene Unsicherheitsmaße und Google-Suchanfragen. Für den elektronischen Zahlungsverkehr werden sowohl der Wert als auch das Volumen von elektronischen Transaktionen und Bargeldabhebungen berücksichtigt. Die Prognosen werden mit einem multivariaten Zustandsraummodell erstellt, dass die Berücksichtigung von unterschiedlichen Frequenzen der verwendeten Variablen ermöglicht. Der Evaluierungszeitraum beginnt im ersten Quartal des Jahres 2008 und endet im vierten Quartal des Jahres 2017; die betrachteten Prognosehorizonte reichen dabei vom Nowcast bis zu vier Quartalen, wobei jeweils drei Datenstände innerhalb eines Quartals verglichen werden. Dadurch wird in der Evaluierung berücksichtigt, dass Zahlungsverkehrsdaten zeitnäher verfügbar sind als beispielsweise die Einzelhandelsumsätze oder die Beschäftigung. Um die Prognoseeigenschaften der verschiedenen Indikatorengruppen zu analysieren, vergleichen die Autoren zunächst nur Modelle, die jeweils eine der Indikatorengruppen enthalten. Hier schneidet der Zahlungsverkehr insgesamt am besten ab; im Vergleich zu Modellen, die auf „harten" Indikatoren basieren, sind die Verbesserungen aber gering und insgesamt liefern die Modelle keine signifikant höhere Prognosegüte als ein einfacher Random-Walk. In einem weiteren Schritt werden Indikatorengruppen miteinander kombiniert, z.B. „harte" Indikatoren und Zahlungsverkehrsdaten. Bei diesen kombinierten Modellen schneiden die Zahlungsverkehrsdaten sehr schlecht ab. Die besten Prognosen liefert ein Modell, das herkömmliche „harte" und „weiche" Indikatoren enthält. Nur dieses Modell liefert signifikant bessere Prognosen als ein Random-Walk. Schließlich kommen Prognosekombinationsverfahren zum Einsatz, die Modelle einzelner Indikatoren aus den verschiedenen Gruppen zusammenführen. Hier zeigen sich signifikante Verbesserungen gegenüber einem Random-Walk für fast alle Horizonte (Nowcast sowie für Prognosen der kommenden beiden Quartale). Allerdings sind auch hier die Verbesserungen von Modellen, die alternative Datenquellen verwenden, nur gering gegenüber jenen, die ausschließlich konventionelle Indikatoren verwenden. Insgesamt sind die Ergebnisse uneinheitlich und klare Verbesserungen, die auf die Verwendung von Zahlungsverkehrsdaten zurückzuführen sind, nicht auszumachen.

Aladangany et al. (2019) konstruieren auf Basis von Kreditkartentransaktionen eines privaten Zahlungssystembetreibers einen monatlichen Index der Einzelhandelsumsätze, der möglichst genau den amtlichen Zahlen in den Vereinigten Staaten entsprechen soll. Dieser wird genutzt, um die Auswirkungen des sogenannten „shutdowns" der öffentlichen Verwaltung im Winter 2018/2019 zu erfassen, als die Erhebung und Veröffentlichung von amtlichen Daten verzögert wurde. Des Weiteren liegen die Transaktionsdaten auch auf regionaler Ebene vor. Somit lassen sich auch die Auswirkungen von regional begrenzten Naturkatastrophen auf das Konsumverhalten abschätzen. So beziffern die Autoren den

kurzfristigen Rückgang des Bruttoinlandsprodukts aufgrund der Hurrikane Harvey und Irma auf ungefähr 0,2 Prozent im dritten Quartal des Jahres 2017.

Daten zum elektronischen Zahlungsverkehr können auch Einsichten in die Wirksamkeit wirtschaftspolitischer Maßnahmen liefern. Agarwal und Qian (2014) untersuchen anhand von Giro- und Kreditkartentransaktionen, wie Haushalte auf einen nicht antizipierten Einkommensanstieg reagieren. Bezogen werden die Daten von einer großen singapurischen Bank und sie umfassen Kontostände sowie Giro- und Kreditkartenumsätze von 200 000 repräsentativ ausgewählten Kunden. Ausgangspunkt für die Analyse ist das „Growth Dividend Program" der singapurischen Regierung im Jahr 2011, im Zuge dessen alle Staatsbürger, die älter als 21 Jahre waren, eine Einmalzahlung in Höhe von 428 bis 624 US-Dollar oder etwa 18 Prozent des monatlichen Medianeinkommens erhielten. Für die Analyse von besonderer Bedeutung ist dabei, dass in den Medien über die Pläne der Regierung nicht vorab berichtet wurde und sie daher wohl nicht erwartet wurden. Zudem wurden die Zahlungen zwar bereits im Februar angekündigt, Haushalte konnten jedoch erst ab Ende April auf die Mittel zugreifen. Dadurch lässt sich neben den Gesamtauswirkungen einer nicht antizipierten temporären Einkommenserhöhung auch die Bedeutung von Ankündigungseffekten untersuchen. Die Autoren zeigen, dass mit 80 Prozent ein Großteil des zusätzlichen Einkommens ausgegeben wurde. Ein bedeutender Teil der Ausgaben – knapp 20 Prozent – erfolgte dabei bereits in den ersten Monaten, als die Haushalte über die Zahlungen informiert, diese jedoch noch nicht auf ihren Konten eingegangen waren. Des Weiteren zeigen die Autoren, dass insbesondere jene Haushalte ihren Konsum deutlich ausweiteten, die über wenige liquide Vermögenswerte oder niedrige Kreditkartenlimits verfügten.

2.5.5 Potenziale und Grenzen

Daten zum elektronischen Zahlungsverkehr bilden getätigte Transaktionen unmittelbar und umfassend ab. Zudem liegen sie grundsätzlich zeitnäher vor als andere Frühindikatoren. Somit sind sie insbesondere für den Nowcast von makrökonomischen Größen geeignet. Potenziale für längere Prognosehorizonte könnten sich dadurch ergeben, dass diese Prognosen auf einem besseren Nowcast aufsetzen oder dass die tägliche Frequenz, mit der die Daten ausgewertet werden können, dazu verwendet wird, um statistische Überhänge für die Folgeperioden zu berechnen. Außerdem könnten die Transaktionen im laufenden Quartal Informationen über die zukünftige Aktivität beinhalten, wenn sie implizit Informationen über noch nicht vorliegende Variablen, wie z.B. die Verbraucherstimmung, enthalten. Sie dürften insbesondere dazu geeignet sein, die Einzelhandelsumsätze, die privaten Konsumausgaben oder das Bruttoinlandsprodukt zu prognostizieren. Dafür sind sie auch in der Mehrzahl der vorliegenden makroökonomischen Studien verwendet worden. Sie können aber auch helfen, um kurzfristige Auswirkungen von makroökonomisch relevanten Ereignissen zu analysieren; insbesondere, wenn dafür tagesaktuelle bzw. disaggregierte Daten erforderlich sind. Darüber hinaus könnten sie auch nützlich sein, um die Aktivität in Dienstleistungsbranchen abzuschätzen oder spezifisch die Transaktionen zwischen Unternehmen zu analysieren. Auch könnte mit ihnen versucht werden, makroökonomische Phänomene, wie Schwankungen der Umlaufgeschwindigkeit des Geldes, abzubilden. Für all dies liegen allerdings noch keine Analysen vor. Auch für Deutschland gibt es bislang keine Anwendungen, die auf Daten des elektronischen Zahlungsverkehrs zurückgreifen.

Daten zum elektronischen Zahlungsverkehr sind jedoch nicht frei verfügbar. Der Zugang muss mit dem jeweiligen Anbieter vereinbart werden und kann sehr kostspielig sein. Vor allem ein dauerhafter Zugang zu den Daten, der insbesondere für die Konjunkturanalyse notwendig wäre, dürfte für Statistikämter oder Zentralbanken, die dazu zum Teil auf von ihnen selbst erfasste Daten zurückgreifen könnten, leichter zu erzielen sein als für einzelne Forscher oder Forschergruppen. Zudem sind die vorliegenden Ergebnisse in der Literatur bezüglich der Prognosegüte der Daten zum elektronischen Zahlungsverkehr

gemischt. Dies könnte damit zusammenhängen, dass es in den bisherigen Analysen noch nicht gelungen ist, die grundsätzlichen Vorteile der Daten, insbesondere ihre zeitnahe Verfügbarkeit, vollständig für die Prognose auszuschöpfen.

2.6 Fernerkundungsdaten

Fernerkundungsdaten liefern eine große Bandbreite von Informationen, die für wissenschaftliche Analysen von Interesse sind. Dazu zählen beispielsweise Daten zu Witterungs- und Klimabedingungen, zur Landnutzung oder zum Nachtlicht. Fernerkundungsdaten weisen drei große Vorteile auf:

1. Sie sind anderweitig nicht oder nur in geringerer Qualität verfügbar.
2. Sie sind in einer sehr hohen regionalen Auflösung verfügbar.
3. Sie sind konsistent für eine sehr große räumliche Abdeckung verfügbar.

Aus diesen Gründen sind sie nicht nur in verschiedenen wissenschaftlichen Disziplinen bereits regelmäßig eingesetzt worden; sie werden auch zunehmend als Datenquelle für die amtliche Statistik in Betracht gezogen. Für makroökonomische Fragestellungen sind bislang vor allem Nachtlichtdaten als Maß für die wirtschaftliche Aktivität genutzt worden sowie Daten zu den Witterungsbedingungen, um den Einfluss der Witterung auf die Konjunktur abzuschätzen.

2.6.1 Datenquellen

Fernerkundungsdaten werden durch die Messung und Interpretation der von der Erdoberfläche ausgehenden Energiefelder gewonnen. Die meisten dieser Daten werden mithilfe von Satelliten gemessen. Man spricht daher oft von Satellitendaten als Synonym für Fernerkundungsdaten. Grundsätzlich ist aber auch eine Messung von anderen Flugobjekten wie Flugzeugen oder Raumstationen aus möglich. Für Wetter- und Klimadaten kommen neben Fernerkundungsdaten alternativ auch Wetterstationen als Datenquelle in Frage.

Nachdem die Nutzung von Satellitendaten vor zwanzig Jahren noch wenigen Experten vorbehalten war, hat sich die Menge und Verfügbarkeit der Daten zuletzt rasant erhöht. Die Weltorganisation für Meteorologie (WMO), eine Sonderorganisation der Vereinten Nationen, listet im Rahmen des Observing Systems Capability Analysis and Review Tool (OSCAR) derzeit 170 aktive oder geplante Satellitenprogramme von 85 Raumfahrtbehörden.

Satellitendaten werden von unterschiedlichen Anbietern zum Teil kostenpflichtig zur Verfügung gestellt (Kasten 1). Zu den kostenfreien Datenportalen zählen u.a.:

- ESA (Europäische Weltraumorganisation): Copernicus oder Observing the Earth,
- CEOS (Committee on Earth Observations Satellites): IDN Portal,
- NASA (National Aeronautics and Space Administration): MODIS-Daten oder Visible Earth,
- USGS (U.S. Geological Survey): Earth Explorer oder GloVis,
- GLCF (Global Land Cover Facility): Maryland-Server.

Satellitendaten zur nächtlichen Emission künstlichen Lichts (Nachtlichtdaten) werden aus zwei Quellen gewonnen: Dem „Operational Linescan System“ Sensor an Bord von Satelliten des Defense Meteorological Satellite Program (DMSP) der US-Luftwaffe sowie dem „Visible Infrared Imaging Radiometer Suite“ (VIIRS) Sensor an Bord des „Suomi National Polar-orbiting Partnership“ Wettersatelliten. Die daraus resultierenden Daten sind beispielsweise bei der Earth Observations Group (Lichtdaten bei Nacht) kostenfrei zugänglich.

Kasten 1:
Eine detaillierte Beschreibung ausgewählter Satellitenprogramme als Datenquellen

Das Copernicus-Programm

Das Copernicus Programm besteht vornehmlich aus der so genannten Sentinel-Familie, einer Serie von Erdbeobachtungssatelliten, die von der europäischen Weltraumagentur (European Space Agency ESA) betrieben werden (Tabelle K1). Sie werden ergänzt durch Messsysteme am Boden, in der Luft und in Gewässern. Die Copernicus-Dienste analysieren und verarbeiten diese Datenfülle zu Informationsprodukten, wie tagesaktuelle Karten für die Überwachung der Meere und der Schifffahrt, für das Katastrophen- und Krisenmanagement oder für das Monitoring von Veränderungen der Atmosphäre und des Klimas, von Vegetation und der Landnutzung. Zu den angebotenen Diensten gehört es ebenfalls, vergleichbare historische Datenreihen zur Verfügung zu stellen, die es erlauben, Zusammenhänge und Trends zu identifizieren. Die Copernicus-Dienste enthalten derzeit Daten für die folgenden Bereiche: Landüberwachung (z.B. Vegetationsindizes, Oberflächentemperatur), Überwachung der Meeresumwelt, Katastrophen- und Krisenmanagement, Beobachtung der Atmosphäre und des Klimawandels. Die einzelnen Datenprodukte der Copernicus Dienste können über Datenportale nach einer Registrierung teilweise kostenfrei bezogen werden.

Auf dem deutschen Zugangsportal zu den Copernicus-Daten (CODE-DE) ist es außerdem möglich, weiterführende Informationen aus den Sentinel-Satellitendaten zu erstellen. So können zum Beispiel Oberflächentemperaturen oder Pflanzenwachstumsindizes über eine Karte unter der Auswahl des Datums und des jeweiligen räumlichen Ausschnitts des Orbits bezogen werden. Die Frequenz der Daten richtet sich nach der Überflughäufigkeit des jeweiligen Satelliten (im Fall von Sentinel 2 z.B. alle fünf Tage). Um die Daten für ökonomische Analysen zu verwenden, ist aber häufig noch ein weiterer Schritt notwendig, denn in der Regel bedarf es dafür einer größeren räumlichen (z.B. Deutschland) und einer bestimmten zeitlichen Abdeckung (z.B. monatliche Durchschnitte). Diese Level-3-Analysen können mithilfe der auf CODE-DE angebotenen sogenannten Prozessoren grundsätzlich durchgeführt werden. Dieses Angebot und die verfügbaren Daten werden stetig erweitert.

Um Datentransfer und Datenspeicherung der Copernicus-Daten auf Nutzerseite zu erleichtern, werden auf CODE-DE zudem zusätzlich Cloud-basierte Dienste angeboten. Diese Dienste oder Plattformen (DIAS) ermöglichen den Zugriff auf Copernicus-Daten und bieten Computerressourcen und Algorithmen an, um die Daten verarbeiten zu können, ohne sie vorher speichern zu müssen. Zurzeit werden fünf alternative DIAS-Plattformen angeboten, die sich nach ihrem Daten- und Softwareangebot, in der Portal-, Nutzungs- und Prozessierungsumgebung sowie in den Preismodellen unterscheiden.

Nachtlichtdaten

Satellitendaten über die nächtliche Emission künstlichen Lichts auf der Erdoberfläche werden aus zwei Quellen gewonnen: dem „Operational Linescan System“ (OLS) Sensor an Bord von Satelliten des Defense Meteorological Satellite Program (DMSP) der US Luftwaffe, und dem „Visible Infrared Imaging Radiometer Suite“ (VIIRS) Sensor an Bord des „Suomi National Polar-orbiting Partnership“ (Suomi NPP) Wettersatelliten. Satelliten des DMSP umkreisen seit Mitte der 1960er Jahre mehrmals täglich die Erde und erfassen die Intensität der Lichtemissionen der Erde (vgl. u.a. Elvidge et al. 1997; Elvidge et al. 2001; Chen und Nordhaus 2011; Henderson et al. 2012; Michalopoulos und Papaioannou 2018). Diese Satelliten dienten der Luftwaffe ursprünglich dazu, Wetterdaten zu sammeln. Als Nebenprodukt erfassen sie in wolkenlosen Nächten aber auch Lichtemissionen und -Reflektionen von der Erdoberfläche. Diese Satellitenbilder wurden durch das geophysische Datenzentrum der US-Wetter- und Ozeanografiebehörde (NOAA) aufbereitet. Ziel war es dabei, die Abstrahlung von künstlichem, menschlich verursachtem Licht auf der Erdoberfläche bei Nacht in dem für das menschliche Auge sichtbaren und infrarotnahem (visible and near-infrared; VNR) Bereich zu erfassen. Während der Aufbereitung wurden die Daten unter anderem um Bewölkung, natürliche Lichtquellen (z.B. Polarlichter), Blitze, Reflektionen von Mondlicht und Sonneneinstrahlung sowie Verzerrungen durch Atmosphärenstaub und Waldbrände bereinigt.[a] Die bereinigten Daten wurden über alle Nächte, aus denen verwertbare Daten vorlagen, zu Jahresdurchschnitten gemittelt. Die DMSP-Lichtdaten sind als Jahresdaten von 1992 bis 2013 auf den Interseiten der NOAA,[b] oder der Colorado School of Mines[c] frei verfügbar. Die Standardversion der Daten („Average Visible, Stable Lights, & Cloud Free Coverages“) misst die Intensität stabiler Lichtquellen menschlichen Ursprungs mit einer Pixelgröße von 30 Bogensekunden (in Deutschland ca. 600x600m). Die Strahldichte wird mit einer Punkteskala von 0 DN (average visible band digital number) bis zu einem maximalen Wert (obere Abschneidegrenze) von 63 DN wiedergegeben.

Tabelle K1:
Auswahl der derzeit bei Copernicus gelisteten aktiven Satelliten

Die derzeit bei Copernicus gelisteten aktiven Satelliten	Frequenz	Auflösung	Beschreibung
Sentinel-1	6 Tage	9 m–40 m	Liefert unabhängig von Helligkeit und Wolkenbedeckung, lückenlos Daten über die Land- und Wasseroberflächen der Erde. Die hohe Auflösung und Wiederkehrrate ermöglicht zeitnahe Aufnahmen von Überschwemmungen an Land oder Ölverschmutzungen auf dem Meer, ebenso wie von Bodenbewegungen oder der Vegetationsdichte.
Sentinel-2	5 Tage	10 m–60 m	Aufnahmen im sichtbaren und infraroten Spektrum, für die Beobachtung der Landoberflächen optimiert. Die hohe Auflösung von bis zu 10m und die Abtastbreite von 290 km sind geeignet, um Veränderungen der Vegetation zu beobachten und Erntevorhersagen zu erstellen, Waldbestände zu kartieren oder das Wachstum von Wild- und Nutzpflanzen zu bestimmen. Die Aufnahmen werden auch für Küsten- und Binnengewässer eingesetzt, um etwa das Algenwachstum zu beobachten oder den Sedimenteintrag in Flussdeltas nachzuverfolgen.
Sentinel-3	< 2 Tage	300 m–1020 m	Sentinel-3 wurde zur Beobachtung der Weltmeere entwickelt. Die Satelliten nutzen ein Paket aus fünf Instrumenten, die präzise die Temperatur, Farbe und den Pegel der Meeresoberfläche bestimmen können. Daraus lassen sich Erkenntnisse über Unterwasserströmungen, Wellenhöhen oder Nährstoffverteilung in den Weltmeeren ableiten. Die Messdaten dienen auch zur Bestimmung des Energiehaushaltes der Erde und der Wasserqualität oder der Umweltverschmutzung an den Küsten. Die Daten eignen sich aber auch zur Bestimmung der Oberflächentemperatur an Land.
Sentinel-5P	1 Tag	7–68 km	Messung von Atmosphärengasen weltweit. Spurengase wie Methan, Ozon und Stickstoff, ebenso wie Aerosole, spielen für das Wetter, den Klimawandel und die Luftreinheit eine wichtige Rolle. Die Daten sind für die Wissenschaft, für die Wettervorhersage und zur Bestimmung der Luftqualität von Bedeutung.
Aeolus	7 Tage	50 km	Weltweite Bestimmung von Windvektoren bis zu einer Höhe von 30 km.
CryoSat-2		369 Tage	Mit CryoSat können die Veränderungen der Eisdicke der polaren Eisschilde und des Meereises genau und flächendeckend bis in hohe Breiten beobachtet werden.
DEIMOS-1	3 Tage	22 m	Liefert multispektrale Bilder einer räumlichen Auflösung von etwa 20 Metern und einer Aufnahmebreite von 650 km. Der Satellit ist auf Land- und Forstwirtschaft sowie Überwachungsanwendungen zugeschnitten.
GRACE-FO		regionale Skala	Vermessung des Erdgravitationsfelds und seiner zeitlichen Veränderungen
Landsat 7 und 8	16 Tage	Panchromatisch: 15 m, VIS-SWIR: 30 m	Aufnahmen im sichtbaren und infraroten Spektrum, die für die Beobachtung der Landoberflächen und ihrer Veränderungen geeignet sind.
Meteosat	alle 15 min	1–3 km	Geostationäre Wettersatelliten für Europa.
PlanetScope-Konstellation / Dove-Satelliten	1 Tag	3 m	Tägliche Aufnahmen der Erde (außer bei Wolken) mit einer hohen Auflösung von drei Metern pro Pixel.
Proba-V	2 Tage	300 m–1 km	Aufnahmen zur Beobachtung der Landbedeckung und des Vegetationswachstums auf der Erde.
CSA/MDA	24 Tage	1–100 m	Daten werden u.a. zur Landnutzungskartierung, zur Hochwasserkartierung, Schiffsüberwachung und zum Gewässermonitoring eingesetzt.
Rapideye	1–5 Tage	5 m	Erdbeobachtungssystem mit hoher Auflösung.

Quelle: Eigene Darstellung auf Basis von Copernicus (2020).

Neben der mangelnden Aktualität schränken verschiedene Faktoren die Datenqualität und die Aussagekraft der DMSP-Daten zum Teil erheblich ein (vgl. u.a. Elvidge et al. 2013; Gibson et al. 2019). Erstens bestehen Unschärfen im Bereich sehr niedriger Lichtintensitäten (DN nahe null). Dort fallen vor allem schwache Lichtquellen in dünnbesiedelten Regionen Filtern zum Opfer, mit denen natürliches Restlicht von Mond und Sonne herausgefiltert wird (Gibson et al. 2019: 6–7). Zudem verzerrt die obere Abschneidegrenze bei einem DN-Wert von 63 die ausgewiesene Lichtintensität vor allem in größeren Städten nach unten (Hsu et al. 2015). Dort können DN-Werte von über 5000 DN erreicht werden (Bluhm und Krause 2018). Diese Zensierung der Daten betrifft zwar weltweit nur einen kleinen Prozentsatz der Pixel (Henderson et al. 2012: 999), dürfte aber in relativ dicht besiedelten Ländern wie Deutschland überdurchschnittlich sein. In den Niederlanden beispielsweise wurden im Jahr 2010 laut Bluhm und Krause (2018: 7) die tatsächlichen Lichtintensitäten auf 17 Prozent der Gesamtfläche um durchschnittlich mehr als 50 Prozent unterschätzt.[d] Zweitens gibt es Überlagerungseffekte zwischen benachbarten Pixeln (Small et al. 2011; Abrahams et al. 2018). Diese resultieren vor allem daher, dass (i) auch Licht aus benachbarten Pixeln einfangen wird, (ii) die Geokodierung der Pixel im Satelliten unpräzise ist, und (iii) starke Lichtquellen indirekt auch Objekte in benachbarten Pixeln beleuchten können.[e] Drittens sind die Daten der sukzessive eingesetzten sechs Satelliten nur eingeschränkt vergleichbar, weil die Datenqualität mit der technischen Beschaffenheit und dem Alter der eingesetzten Sensoren erheblich variieren (Henderson et al. 2012: 999, Hsu et al. 2015). Stichprobenartige Vergleiche haben Unterschiede von bis zu 29 Prozent ergeben (Gibson et al. 2019: 6).

Der VIIRS-Sensor an Bord des Suomi NPP-Wettersatelliten liefert aktuellere und qualitativ höherwertige Nachtlichtdaten in höherer Auflösung als das DMSP. Er erfasst seit dem Jahr 2011 Daten über den Energiehaushalt, die Vegetation, die Landnutzung und den Wasserkreislauf der Erde. Die VIIRS-Daten werden durch die Erdbeobachtungsgruppe (Earth Observations Group, EOG) des Payne Institute for Public Policy an der Colorado School of Mines (Golden, Colorado) bearbeitet, um Streulicht, Blitze, Mondlicht und Wolken herauszufiltern und zu Monats- und Jahreswerten gemittelt. Die VIIRS-Lichtdaten sind als Monats- oder Jahresdurchschnitte für den Zeitraum von April 2012 bis zum aktuellen Rand verfügbar. Sie werden mit einer zeitlichen Verzögerung von rund zwei Monaten veröffentlicht. Sie sind kostenlos auf der Internetseite der Earth Observations Group verfügbar.[f] Zusätzlich zur Strahldichte mit einer geografischen Auflösung von 15 Bogensekunden (in Deutschland 300x300 Meter) enthalten die Daten auch eine Angabe über die Zahl der Beobachtungen, die in den jeweiligen monatlichen Mittelwert eingeflossen sind. Die monatlichen Daten werden in zwei Varianten angeboten. In einer Variante („vcm“) werden nur die Beobachtungen in die Monatsdurchschnitte einbezogen, die nicht durch Streulicht beeinflusst sind. In der anderen Variante („vcmsl“) werden diese Beobachtungen einbezogen, wenn sie durch einen Streulichtfilter korrigiert wurden. Die vcmsl-Mittelwerte basieren also tendenziell auf mehr Beobachtungen, können aber dadurch verzerrt sein, dass der Streulichtfilter auch menschlich verursachte Lichtquellen abschwächt. Grundsätzlich weisen die VIIRS-Daten ähnliche Probleme auf wie die DMSP Daten; allerdings sind sie hier deutlich weniger ausgeprägt.

[a]Vgl. Via Internet (11.2.2020) <https://www.ngdc.noaa.gov/eog/gcv4_readme.txt>. — [b]Vgl. Via Internet (11.2.2020) <https://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>. — [c]Vgl. Via Internet (11.2.2020) <https://eogdata.mines.edu/dmsp/downloadV4composites.html>. — [d]Um die tatsächlichen Lichtintensitäten zu schätzen, wurden verschiedene Ansätze entwickelt, die unter anderem die sporadischen, für nur sieben Jahre verfügbaren Satellitenaufnahmen („Global Radiance Calibrated Nighttime Lights”) nutzen, bei denen die Zensierung deaktiviert wurde (z.B. Bluhm und Krause 2018). Allerdings sind diese Daten durch andere Messfehler verunreinigt. Zudem erfordern die Schätzungen restriktive Annahmen. Bluhm und Krause (2018: 7) nehmen beispielsweise an, dass die Stadtgrößenstruktur dem Zipfschen Gesetz folgt. Dieses Gesetz fordert, dass die von der Einwohnerzahl her *n*'t größte Stadt ein *n*'tel mal so groß ist wie die größte Stadt in diesem Land, dass also die zweitgrößte Stadt halb so groß, die drittgrößte ein Drittel so groß, die viertgrößte ein Viertel so groß u.s.w. ist. Für Deutschland trifft dieses Gesetz aber nur bedingt zu, insbesondere, da hierzulande viele Städte eine ähnliche Zahl von Einwohnern haben. — [e]Erst seit kurzer Zeit sind erste Verfahren verfügbar, die versuchen, diese Überlagerungen zu korrigieren (Abrahams et al. 2018; Cao et al. 2019; Liu et al. 2019; Hu et al. 2019). Diese Verfahren wurden jedoch bisher nur auf die zensierten Daten angewandt. Verfahren zur simultanen Korrektur von Zentrierung und Überlagerungseffekten sind bisher offenbar nicht verfügbar. — [f]Vgl. EOG data via Internet (11.2.2020): <https://eogdata.mines.edu/download_dnb_composites.html>. Bis Oktober 2019 wurden die Daten durch die NOAA bereitgestellt, via Internet (11.2. 2020): <https://www.ngdc.noaa.gov/eog/viirs/download_dnb_composites.html>. Dort werden sie jedoch nicht mehr aktualisiert.

Eine zentrale Quelle für Klima- und Wetterdaten für Deutschland und Europa ist die Datenbank des Deutschen Wetterdienstes (DWD). Die Daten stammen weitestgehend von den Messstationen des DWD und werden nur zum Teil durch Satellitendaten ergänzt.

2.6.2 Datenbeschaffenheit

Satelliten-basierte Fernerkundungsdaten werden von den Datenanbietern in unterschiedlichen Bearbeitungsgraden (Level) zur Verfügung gestellt. Die unmittelbar von den Satelitensensoren gemessenen Rohdaten (Level 0) sind in der Regel nicht unmittelbar für die Analyse ökonomischer Fragestellungen anwendbar, sondern müssen dafür zunächst aufbereitet und ausgewertet werden. Für die sogenannte Level 1-Daten werden die Rohdaten in der Regel nur leicht angepasst; etwa werden die unterschiedlichen Blickwinkel der Satellitendaten berücksichtigt, eine Projektion zu Koordinatensystemen hergestellt oder unterschiedliche Messungen kombiniert, um wolkenfreie Aufnahmen zu erhalten (Donaldson and Storeygard 2016). Level 2-Daten sind bereits zu den jeweils gewünschten Variablen weiterverarbeitet. Hierfür werden die gemessenen Werte mit Hilfe von Machine Learning Algorithmen interpretiert und ausgewertet. Teilweise können (physikalische) Parameter direkt aus der von den Satelliten gemessenen elektromagnetischen Strahlung gewonnen werden, zum Teil werden zur Berechnung der Daten aber auch zusätzliche Informationen (z.B. Umgebungsbedingungen wie die Temperatur oder die Helligkeit des Nachtlichts) hinzugezogen. Für andere Daten bedarf es einer weiteren Klassifizierung der gemessenen Strahlungswerte in diskrete Kategorien (Donaldson and Storeygard 2016). So werden Landnutzungskarten mittels Algorithmen (welche zuvor anhand von Trainingsdaten entwickelt wurden) erstellt, die anhand der Strahlungsinformationen unterschiedliche Landnutzungskategorien identifizieren.[14] Level 3-Daten umfassen in der Regel bereits finale Kartenprodukte wie etwa die monatliche Durchschnittstemperatur in Europa oder Nachtlichtdaten für verschiedene Regionen (Abbildung 8).

Geostationäre Satelliten beobachten durchgehend eine bestimmte Region der Erde. Heliosynchrone Satelliten liefern dagegen Informationen für die gesamte Erde. Die Höhe, mit der ein Satellit die Erde umkreist, entscheidet dabei, in welcher Auflösung und mit welcher Frequenz die Daten verfügbar sind; je höher die Auflösung, desto geringer die Frequenz. Viele heliosynchrone Satelliten liefern zumindest einmal innerhalb weniger Tage Rohdaten, allerdings können ungünstige Witterungsbedingungen dazu führen, dass die Daten nicht durchgängig erfasst werden. Der jeweilige Bearbeitungsgrad der Daten beeinflusst, mit welcher Verzögerung sie dem Nutzer zur Verfügung gestellt werden können. Manche Daten sind bereits nach wenigen Tagen verfügbar (zum Beispiel Temperaturdaten), aufwendigere Analysen können mehrere Monate oder sogar Jahre in Anspruch nehmen (zum Beipiel für Landnutzungskarten). Nachtlichtdaten vom DMSP liegen auf Jahredatenbasis für 1992 bis 2013 vor. Aktuellere und qualitativ hochwertigere Daten liefert Suomi NPP. Sie liegen auf monatlicher Basis seit dem Jahr 2012 vor und werden mit einer Verzögerung von etwa 2 Monaten aktualisiert. Auch für andere Satellitendaten unterscheidet sich der verfügbare Zeitraum zum Teil deutlich. Einige Daten liegen bereits ab den 1970er Jahren vor. Auch bei der Auflösung der Satellitenbilder (Größe eines Pixels) gibt es Unterschiede. Einige Satelliten liefern Bilder mit einer Auflösung von bis zu 30 cm je Pixel.

Der DWD stellt umfassende Wetter- und Klimadaten in unterschiedlichen Frequenzen (täglich bis jährlich) und räumlichen Abgrenzungen zur Verfügung. Die verfügbaren Daten umfassen zum Beispiel Durchschnittstemperaturen und Temperaturspitzen, Niederschlags- und Schneeparameter, Windgeschwindigkeiten, Sonnenscheindauer oder Bodentemperatur.

[14] Siehe zum Beispiel die jährlich herausgegebenen Corine Land Cover Maps: https://land.copernicus.eu/pan-european/corine-land-cover.

Abbildung 8:
Nachtlicht im Mittelmeerraum im Jahr 2016

Quelle: NASA Earth Observatory (2020).

2.6.3 Methodik

Werden Fernerkundungsdaten in einem gebräuchlichen Raster- oder Polygondatenformat zur Verfügung gestellt, können sie mit der gängigen Geoinformationssystem (GIS) Software (z.B. das lizenzpflichtige ArcGIS von Esri oder die Open Source Software Qgis) verarbeitet werden. Dort können beispielsweise Änderungen der räumlichen Projektion der Daten oder eine weitere räumliche Aggregation vorgenommen werden. Innerhalb der GIS-Software können die Daten mit anderen ökonomischen Daten, die einen räumlichen Bezug haben, kombiniert werden. Möglich wäre zum Beispiel eine Kombination zu den sogenannten NUTS-Regionen (*Nomenclature des unités territoriales statistiques*), den räumlichen Bezugseinheiten der amtlichen Statistik in den Mitgliedstaaten der Europäischen Union.

Liegen ökonomische Daten auf einer NUTS-Ebene vor, können diese Daten den NUTS-Grenzen zugeordnet und als Karte dargestellt werden. Darüber hinaus können (zum Beispiel mithilfe der Software ArcGis) auf Basis der Satellitendaten Statistiken für die NUTS-Regionen berechnet werden. Die aufbereiteten Daten können zudem zur weiteren Analyse in gängige Statistik-Software (z.B. R oder Stata) übertragen werden. Dafür stehen bereits Programme zur Verfügung, die kostenfrei in die jeweilige Software importiert werden können.

Das Verarbeiten der Geodaten bedarf grundsätzlich Kenntnisse der Geografie, z.B. in Bezug auf Koordinatensysteme und räumlichen Projektionen als auch der Geoinformatik zum Umgang mit der relevanten Software (z.B. ArcGis, QGis, Python). Bei der Kombination unterschiedlicher Datensätze ist zusätzlich eine Vereinheitlichung von räumlichen und zeitlichen Bezügen zur Vergleichbarkeit der Daten notwendig.

So werden viele Fernerkundungsdaten derzeit größtenteils maximal im Prozessionslevel 2 zur Verfügung gestellt. Bei Copernicus liegen beispielsweise die aktuellen Oberflächentemperaturen des jeweils

vergangenen Monats nicht als komplette Karte für Deutschland vor, sondern in Form von Kacheln (einzelne Beobachtungsausschnitte vorgegeben durch die Umlaufbahnen der Satelliten). Um eine durchgehende Karte der monatlichen Durchschnittstemperaturen in Deutschland zu erstellen, müssen somit zunächst die für Deutschland relevanten Kacheln ausgewählt werden. Die Zusammensetzung der einzelnen Kacheln und das Berechnen der Durchschnittstemperatur aller Beobachtungen in einem Monat müssen selbst vorgenommen werden. Die Verarbeitung der Kacheln kann z.B. mit ArcGis oder Python vorgenommen werden, ist aber aufgrund der Vielzahl einzelner Kacheln und unterschiedlicher Beobachtungszeitpunkte aufwendig. Zukünftig sollen solche Prozessionsschritte auch innerhalb der Copernicus-Plattform vorgenommen werden können. Diese und andere Dienste sowie das Angebot an zusätzlichen bereits fertigen Datenprodukten befinden sich noch im Aufbau. Je nachdem in welcher Frequenz, Auflösung und Prozessionstiefe die Daten zur Verfügung stehen, kann die Verarbeitung und Aktualisierung somit noch sehr aufwendig sein.

Angesichts dieser Herausforderungen ermöglicht der DWD derzeit einen weitaus einfacheren Zugang auf Wetterdaten. Dort werden die Daten der einzelnen Messstationen bereitgestellt. Um ganz Deutschland abdecken zu können, werden die Werte der einzelnen Messstationen zwar interpoliert; für Analysen bezogen auf größere räumliche Ebenen (wie etwa das gesamte Bundesgebiet) sollte das allerdings kaum eine Rolle spielen. Die Daten des DWD lassen sich mit der gängigen GIS-Software darstellen und weiterverarbeiten.

2.6.4 Bisherige Anwendungen

Fernerkundungsdaten nehmen unterschiedliche Funktionen in wirtschaftswissenschaftlichen Analysen ein. Für makroökonomische Fragestellungen wurden sie bislang vor allem herangezogen, um die wirtschaftliche Aktivität insbesondere mittels Lichtdaten zu messen sowie um den Einfluss der Witterung auf die Konjunktur abzuschätzen.

Mittlerweise gibt es mehr als 150 wissenschaftliche Publikationen, die Lichtdaten verwenden (Gibson et al. 2019: 3). In der ökonomischen Literatur werden bisher fast ausschließlich die DMSP-Daten genutzt (Gibson et al. 2019).[15] Ihre spezifischen Vorteile entfalten Lichtdaten vor allem in Studien für Entwicklungs- und Schwellenländer, wo sie in Ermangelung hochwertiger wirtschaftsstatistischer Daten zur Approximation des Pro-Kopf-Einkommens verwandt werden. Ferner erlauben sie makroökonomische Analysen subnationaler Einheiten auf globaler Ebene (z.B. Small et al. 2011; Gennaioli et al. 2013). Sie sind für alle Orte weltweit (mit Ausnahme der Polregionen) in vergleichbarer Qualität und über einen verhältnismäßig langen Zeitraum (22 Jahre) verfügbar. Schließlich werden Lichtdaten auch zur Analyse und Prognose der kleinräumigen Nachfrage nach Energie oder Elektrizität verwandt, wiederum vor allem in Entwicklungs- und Schwellenländern (z.B. He et al. 2012; Tripathy et al. 2018).

Ihr Nutzen für die Analyse makroökonomischer Zusammenhänge in fortgeschrittenen Volkswirtschaften ist allerdings umstritten. Im Gegensatz zu Henderson et al. (2012), die die Verwendung der DMSP-Daten für ökonomische Fragestellungen empfohlen und popularisiert haben, kommen Chen und Nordhaus (2011) Chen und Nordhaus (2019), Nordhaus und Chen (2015), Addison und Stewart (2015) und Leßmann et al. (2015) zu pessimistischeren Einschätzungen. Sie stellen anhand der älteren DMSP-Daten fest, dass die Strahldichte zwar im Querschnitt mit den Pro-Kopf-Einkommen korreliert ist, dies aber vor allem für urbane Regionen und weniger für periphere Regionen mit niedriger Lichtintensität gilt. Als Proxies für Einkommensentwicklungen über die Zeit eignen sich die DMSP-Lichtdaten diesen Studien

[15] Die Überblicksartikel von Donaldson und Storeygard (2016), Michalopoulos und Papaioannou (2018) und Gibson et al. (2019) fassen die verschiedenen Facetten dieser Literatur zusammen.

zufolge jedoch kaum. Der Zusammenhang zwischen den jeweiligen Zuwachsraten ist eher schwach ausgeprägt und in einigen Schätzungen auch nicht statistisch signifikant. Auch Bickenbach et al. (2016b) zeigen, dass Lichtdaten weder in Schwellenländern noch in fortgeschrittenen Volkswirtschaften verlässliche Aussagen über die Zuwachsrate der Pro-Kopf-Einkommen ermöglichen.

Die VIIRS-Daten sind qualitativ hochwertiger als die DMSP-Daten (Elvidge et al. 2013),[16] was sich auch in einer höheren Korrelation mit dem BIP niederschlägt (Chen und Nordhaus 2019; Gibson et al. 2019). Allerdings wecken empirische Studien auch für die VIIRS-Daten Zweifel daran, ob dieser Zusammenhang tatsächlich stabil ist, insbesondere, ob er weitgehend unabhängig von der Besiedlungsdichte ist.[17] Für Analysen mit Lichtdaten auf nationaler Ebene, etwa im Rahmen von Konjunkturprognosen, wäre eine Instabilität im räumlichen Querschnitt vermutlich vernachlässigbar, wenn es einen engen stabilen Zusammenhang zwischen den Veränderungsraten des Bruttoinlandsprodukts und der Lichtintensität gäbe. Chen und Nordhaus (2019) finden für US-Bundesstaaten und für Metropolregionen zwischen den Jahren 2014 und 2016 allerdings nur einen schwachen Zusammenhang. Die geschätzten Parameter für die Lichtintensität sind zwar positiv, deuten aber nur auf einen quantitativ schwachen Zusammenhang hin und sind teilweise nicht signifikant von null verschieden.

Auch andere Satelliten-basierte Daten sind zur Messung ökonomischer Aktivität herangezogen worden. Katona et al. (2018) zeigen, dass mit hochauflösenden Satellitenbildern die Parkplatzauslastung von Einzelhändlern zeitnah bestimmt werden kann, die wiederum Auskunft über die jeweiligen Umsätze liefert. Die Rohdaten für die Analyse stammen von Digital Globe und enthalten Bilder mit einer hohen Auflösung von 50 cm, die täglich für die jeweils etwa gleiche Tageszeit vorliegen.[18] Der Datensatz umfasst 4,7 Millionen tägliche Beobachtungen der Parkplätze von etwa 67 000 Einzelhandelsgeschäften von 44 großen US-Einzelhändlern im Zeitraum von 2011 bis 2017. Die Satellitenbilder werden genutzt, um die Auslastung der Parkplatzflächen in den Geschäften abzuschätzen. Es zeigt sich, dass die Parkplatzauslastung auf Quartalsebene einen signifikanten Erklärungsgehalt für die Umsatzentwicklung der US-Einzelhändler besitzt.

Ähnlich wie die Studien basierend auf den Nachtlichtdaten analysieren Marx et al. (2019) die Qualität der Bebauung in Kiberia, einem Slum in Kenias Hauptstadt Nairobi mithilfe des reflektierten Tageslichts der Metalldächer in unterschiedlichen Metallqualitäten. Die Analyse der Qualität der Bebauung soll dabei zeigen, dass ethnische Klientelsysteme im lokalen Wohnungsmarkt eine Rolle spielen. Die hohe Auflösung der Satellitendaten von 50 cm erlaubt eine Analyse einzelner Dächer in Kombination mit lokalen Umfragedaten. Die Daten decken den Zeitraum zwischen Juli 2009 und August 2012 ab.

---

[16] Die DNB-Sensorik wurde explizit für die Erfassung menschlich erzeugten künstlichen Lichts konzipiert. Die Strahldichte wird nicht durch eine Obergrenze zensiert und die Fähigkeit des Sensors, schwache Lichtquellen zu erfassen, ist deutlich größer, was freilich auch zur Folge hat, dass die unerwünschte Reflektion natürlichen Lichts schwieriger herauszufiltern ist (Gibson et al. 2019). Unschärfen durch Überlagerungseffekte werden deutlich verringert, weil das „Sichtfeld" der Sensorik (ground instantaneous field of view) mit 742x742 Metern 45-mal höher als die des OLS ist und beim Schwenken zu den Seiten konstant bleibt.

[17] Gibson et al. (2019: 7) verwenden für diese Modellregressionen Daten für Indonesien, dessen VGR eine vergleichsweise hohe Qualität aufweist, sodass Schätzfehler beim BIP als gering angenommen werden können. Sie regressieren das jährliche preisbereinigte BIP von 497 indonesischen Regionen auf die für diese Regionen vom VIIRS gemessene jahresdurchschnittliche Strahldichte des Lichts in den Jahren 2015 und 2016 und finden eklatante Unterschiede zwischen urbanen und peripheren Regionen. Während die Elastizität des BIPs in Bezug auf die Strahldichte für die Städte auf rund 0,9 geschätzt wird und statistisch hoch signifikant ist, wird sie für die ländlichen Regionen nur auf weniger als 0,1 geschätzt und ist nicht signifikant von null verschieden. Chen und Nordhaus (2019) stellen ähnliche Schätzungen für die Vereinigten Staaten an und finden, dass die Elastizität auf der Ebene von Bundesstaaten (0,84) signifikant niedriger ist als auf der Ebenen von Metropolregionen (1,1).

[18] Die Bezugskosten für solch hochauflösende Satellitenbilder sind bislang in der Regel recht hoch.

Eine weitere Anwendung von Satelliten-Daten ist die Abschätzung der Ressourcennutzung. In ökonomischen Analysen wird zum Beispiel mithilfe satelliten-gestützter Landnutzungsdaten die Frage nach den Treibern von Landnutzungsänderungen, insbesondere Abholzungsaktivitäten in tropischen Ländern, analysiert. Hierfür werden unterschiedliche, die Abholzung begünstigende Faktoren (z.B. Entfernung zu Straßen, physikalische Indikatoren der Bodenqualität) mit dem jeweiligen Landnutzungsraster auf Pixelebene in Zusammenhang gebracht (vgl. Busch und Ferretti-Gallon 2017 für eine Metaanalyse der Treiber von Abholzung in tropischen Wäldern). Solche Analysen bilden auch die Grundlage für die Berechnung zukünftiger Abholzungstrends im Rahmen des REDD-Programms (Reducing Emissions from Deforestation and Forest Degradation) der Vereinten Nationen.

Schließlich gibt es eine umfangreiche Literatur zum Einfluss der Witterungsbedingungen oder des Klimas auf die wirtschaftliche Aktivität. Die Daten zu Witterungsbedingungen entstammen dabei meist Satelliten oder Wetterstationen. Die Ergebnisse in der Literatur deuten darauf hin, dass Schwankungen der Temperatur oder des Niederschlags einen sichtbaren Einfluss auf die wirtschaftliche Aktivität eines Landes haben können. So sprechen internationale Vergleichsstudien dafür, dass höhere Temperaturen in der Tendenz mit niedrigeren Zuwachsraten des Bruttoinlandsprodukts einhergehen (Dell et al. 2012; Hsiang 2010). Insgesamt sind die wirtschaftlichen Auswirkungen solcher Schwankungen der Witterungsbedingungen in Entwicklungs- und Schwellenländern tendenziell größer als in fortgeschrittenen Volkswirtschaften (Dell et al. 2012). Das Wetter kann dabei neben seinem natürlichen Einfluss auf die landwirtschaftliche Erzeugung auch über andere Kanäle, wie den Energiekonsum oder gesundheitliche Auswirkungen, die wirtschaftliche Aktivität eines Landes beeinflussen.

Besonders große wirtschaftliche Auswirkungen können Extremwetterereignisse, wie Dürren oder Überflutungen und Stürme sowie Extremtemperaturen, haben (Felbermayr und Gröschl 2014). Die konkreten Auswirkungen variieren dabei freilich stark mit der jeweiligen Intensität des Ereignisses. Solche Extremwetterereignisse können auch in fortgeschrittenen Volkswirtschaften zu einem signifikanten Rückgang des Bruttoinlandsprodukts führen.

Für Deutschland wurde bislang vor allem gezeigt, dass sich ändernde Witterungsbedingungen zu kurzfristigen Schwankungen der wirtschaftlichen Aktivität führen können. Neben Umfragedaten von Bauunternehmen sind die Witterungsbedingungen dafür beispielsweise über die Durchschnittstemperaturen, die Anzahl der Eistage oder die Schneehöhe erfasst worden. Ein recht enger Zusammenhang besteht zwischen Witterung und der Bauaktivität (BMWi 2014). So führen ungünstige Witterungsbedingungen kurzfristig zu signifikant niedrigeren Bauinvestitionen. Zwar wird der Einfluss der Witterung auf die Bauaktivität in der Regel bereits bei Konjunkturprognosen berücksichtigt. Allerdings argumentiert Döhrn (2014), dass Prognosen in der Vergangenheit geringer Fehler hätten aufweisen können, wenn Temperaturschwankungen besser berücksichtigt worden wären. Ungewöhnlich milde Winter gehen demnach tendenziell mit zu optimistischen Prognosen einher. Ferner können ungünstige Witterungsbedingungen kurzfristig auch zu einem Anstieg der Arbeitslosigkeit führen (Hummel et al. 2015; Döhrn und an de Meulen 2015). Schließlich können auch andere außergewöhnliche Wetterereignisse die Konjunktur in Deutschland beeinflussen. So zeigen Ademmer et al. (2019c), dass Niedrigwasserperioden des Rheins temporär die Industrieproduktion in Deutschland signifikant drücken. Die Niedrigwasserperiode im Rhein im Herbst 2018 hat die Industrieproduktion demnach in der Spitze um knapp 2 Prozent gemindert.

2.6.5 Potenziale und Grenzen

Grundsätzlich weisen Fernerkundungsdaten zahlreiche Eigenschaften auf, die Potenziale für makroökonomische Analysen bieten. Sie liefern eine große Bandbreite potenziell relevanter Informationen, die durch andere Quellen nicht oder kaum abgedeckt werden können. Auch sind sie häufig zeitnah

verfügbar. Schließlich liefern sie weltweit objektive Informationen in vergleichbarer Qualität und in einem regionalen Detailgrad, für den die herkömmliche Wirtschaftsstatistik keine Daten bereitstellt. Aus diesen Gründen prüfen auch Statistikämter mehr und mehr, ob Fernerkundungsdaten einen wertvollen Beitrag für die amtlichen Statistiken liefern können.

Für makroökonomische Fragestellungen sind Fernerkundungsdaten bislang vor allem eingesetzt worden, um mittels Nachtlichtdaten die wirtschaftliche Aktivität in verschiedenen Regionen abzuschätzen. Allerdings ist insbesondere für fortgeschrittene Volkswirtschaften fraglich, ob Nachtlichtdaten geeignet sind, um kurzfristige Schwankungen der wirtschaftlichen Aktivität abzubilden. In die Konjunkturanalyse fließen ferner regelmäßig Informationen über Witterungsbedingungen ein, um deren kurzfristige Einflüsse auf die Konjunktur, insbesondere die Bauaktivität, zu berücksichtigen. Darüber hinaus steckt die Nutzung von Fernerkundungsdaten für makroökonomische Analysen und insbesondere für die Konjunkturbeobachtung und -prognose bisher noch in den Kinderschuhen. Machbarkeitsstudien wie das „Smart Business Cycle Statistics"-Projekt von Eurostat deuten zwar darauf hin, dass hochauflösende Satellitenbilder in Kombination mit Machine Learning Methoden grundsätzlich geeignet sein könnten, um die laufende wirtschaftliche Aktivität zeitnah zu beobachten. So können damit Handelsintensitäten (anhand von Aufnahmen von Häfen oder anderen Infrastruktureinrichtungen), Einzelhandelsumsätze (Parkplatzauslastung vor Einzelhandelsgeschäften), Bauaktivitäten (Baustellen) oder Ernteerträge abgeschätzt werden (Rosenski und Schartner 2018). Auch das Statistische Bundesamt hat sich an diesen Machbarkeitsstudien beteiligt.[19] Allerdings bedarf es noch weiterer umfangreicher Studien, um besser abschätzen zu können, wie groß das Potenzial von Fernerkundungsdaten tatsächlich für die Konjunkturanalyse ist, auch im Vergleich zu den bereits verfügbaren konventionellen Frühindikatoren oder anderen Datenquellen aus dem Bereich Big Data.

Einer stärkeren Nutzung für makroökonomische Analysen steht der sehr hohe Aufwand entgegen, der betrieben werden muss, um Fernerkundungsdaten auszuwerten. Ohne fundierte geowissenschaftliche Vorkenntnisse ist bereits die Identifikation des für die jeweilige Analyse geeigneten Datensatzes und der notwendigen Prozessschritte zum Aufbereiten der Daten eine Herausforderung. Zudem bedarf es zur Nutzung der Daten oft des Wissens aus anderen Fachrichtungen. Schließlich erhöht sich der Aufwand darüber hinaus noch einmal deutlich, wenn die ökonomische Aktivität anhand von Satellitenbildern in sehr hoher Auflösung in Verbindung mit Methoden des maschinellen Lernens erfasst werden soll.

2.7 Verkehrsdaten

Das Verkehrsaufkommen ist eng mit der wirtschaftlichen Aktivität verknüpft. Da Verkehrsdaten umgehend erfasst werden, können sie frühzeitig Informationen über die laufende wirtschaftliche Entwicklung liefern. Der Güterverkehr dürfte dabei einen weitaus engeren Bezug zur wirtschaftlichen Aktivität aufweisen als der Personenverkehr. In Deutschland wurden im Jahr 2018 etwa 80 Prozent der Güter – gemessen anhand der Beförderungsmengen – über den Straßenverkehr transportiert (Tabelle 3). Deutlich geringer waren die Anteile der per Eisenbahn (8,8 Prozent), per Seeverkehr (7,2 Prozent) oder per Binnenschifffahrt (4,9 Prozent) transportierten Güter. Der Anteil der Luftfracht war mit 0,1 Prozent sehr gering. Auch in anderen fortgeschrittenen Volkswirtschaften sind Lastkraftwagen (Lkw) das klar dominierende Verkehrsmittel im Güterverkehr. Für den grenzüberschreitenden Transport aus und nach Deutschland spielt dagegen der Seeverkehr mit etwa 45 Prozent die größte Rolle. Während der Luftverkehr hier wiederum nur einen sehr geringen Anteil aufweist, entfallen auf die übrigen Verkehrsträger Anteile zwischen 10 Prozent und 25 Prozent.

[19] Vgl. z.B. Arnold und Kleine (2017) und Statistisches Bundesamt, Smart Business Cycle Statistics mit Satellitendaten: https://www.destatis.de/DE/Service/EXDAT/Datensaetze/satellitendaten.html.

Tabelle 3:
Anteile an Beförderungsmengen nach Verkehrsträgern in Deutschland 2018

In Prozent	Gesamt	Innerhalb Deutschlands	Grenzüberschreitend			Durchgangs-verkehr
			Gesamt	Versand	Empfang	
Eisenbahn	8,8	7,0	16,2	17,7	15,0	56,0
Binnenschifffahrt	4,9	1,5	21,1	16,3	24,7	38,6
Seeverkehr	7,2	0,1	45,4	41,9	48,0	0,0
Luftverkehr	0,1	0,0	0,7	0,9	0,6	0,3
Straßenverkehr inländischer Lastkraftwagen	79,0	91,4	16,6	23,1	11,7	5,0

Quelle: Statistisches Bundesamt (2020c); eigene Berechnungen.

Allerdings ist für die wirtschaftliche Aktivität der Wert der Fracht relevanter und diese unterscheidet sich je nach Verkehrsträger deutlich. So ist der Wert im Schienenverkehr am Geringsten (1 400 €/t), für Schiff (2 000€/t) und Straße (2 900/t) etwas höher und im Luftverkehr (86 000€/t) um den Faktor 30–50 höher als bei den übrigen Verkehrsträgern (BVL 2019: 11). Somit spielt die Luftfracht für den Außenhandel mit einem Anteil von rund 11,2 Prozent des wertmäßigen Handelsvolumens eine nicht unbedeutende Rolle; im Handel mit Übersee (Amerika, Asien, Afrika, Australien und Ozeanien) liegt der Anteil der Luftfracht sogar bei rund einem Drittel (ebd.: 12). Bei weltweiten grenzüberschreitenden Warenströmen macht der Seeverkehr mit rund 80 Prozent den wesentlichen Anteil der Transportmenge aus (UNCTAD 2019). In Hinblick auf den Warenwert ist der Luftverkehr mit rund 40 Prozent Anteil neben dem Seeverkehr allerdings ein fast ebenso bedeutender Verkehrsträger (BPB 2017a; BPB 2017b).

2.7.1 Straßenverkehr: Lkw-Maut und Verkehrszählung

2.7.1.1 Datenquellen

Seit Beginn des Jahres 2005 sind Lkw auf deutschen Autobahnen mautpflichtig. Mautpflichtige Fahrten werden per GPS erfasst und automatisiert an den Betreiber TollCollect übermittelt, der die Daten zeitversetzt an das Bundesamt für Güterverkehr weiterleitet. Die Rohdaten zur Mauterhebung sind nicht frei verfügbar. Allerdings werden vom Bundesamt für Güterverkehr rund 15 Tage nach Ende jedes Berichtsmonats die aggregierten monatlichen Mautstatistiken veröffentlicht (BAG 2020). Die dortigen Veröffentlichungen reichen bis ins Jahr 2007 zurück und der Berichtsumfang je Kalendermonat wurde im Zeitverlauf etwas ausgeweitet. So wird z.B. die Fahrleistung getrennt nach Emissionsklasse, Achsenanzahl, zulässigem Gesamtgewicht, nationaler Herkunft bzw. Zulassung des Lkw ausgewiesen, zudem wird die Anzahl der Grenzüberfahrten von mautpflichtigen Fahrzeugen berichtet. Das Statistische Bundesamt veröffentlicht darüber hinaus im Rahmen der amtlichen Konjunkturstatistik inzwischen einen monatlichen Lkw-Fahrleistungsindex, der bereits etwa eine Woche nach dem Berichtsmonat veröffentlicht wird.[20]

Darüber hinaus werden Daten über automatische Verkehrszählungen erfasst: Auf Autobahnen und Bundesstraßen in Deutschland werden an über 1 900 automatischen Zählstellen alle passierenden Fahrzeuge gezählt. Es werden dabei verschiedene Fahrzeugarten – etwa Pkw, Busse und Lkw – getrennt erfasst. Die Daten werden von den jeweiligen Bundesländern erhoben und quartalsweise an das

[20] Seit kurzem veröffentlicht das Statistische Bundesamt mit einer Verzögerung von etwa einer Woche auch tägliche Daten des Lkw-Fahrleistungsindex.

Bundesamt für Straßenwesen übermittelt, das für jede der Messstationen zur Verkehrszählung stündliche Daten ab dem Jahr 2003 veröffentlicht. Allerdings stehen die Daten bislang erst mit einer Verzögerung von mehr als einem Jahr zur Verfügung. Vergleichbare Zählsysteme gibt es auch vielen anderen Ländern.

2.7.1.2 Datenbeschaffenheit und Methodik

Die Mauterhebung erfasst passierende Lkw auf allen Bundesstraßen und Autobahnen. Die Maut, die zunächst nur Fahrzeuge mit zulässigem Gesamtgewicht ab 12t und deutsche Autobahnen betraf, wurde schrittweise auf Lkw ab 7,5t und inzwischen auch auf alle Bundesstraßen ausgeweitet. Damit existiert praktisch eine Vollerhebung in diesem Bereich (Cox et al. 2018). Die frei verfügbaren aggregierten Daten unterscheiden sich in ihrer Beschaffenheit nicht von gängigen makroökonomischen Zeitreihen und können mit üblichen empirischen Methoden ausgewertet werden.

Bezüglich der Verkehrszählung werden grundsätzlich für jede der rund 1 900 Zählstellen stündliche Daten in strukturierter Form vom Bundesamt für Straßenwesen zur Verfügung gestellt. In den Daten wird die Zahl der passierenden Fahrzeuge nach Kategorie (z.B. Pkw oder Lkw) und nach Fahrtrichtung unterschieden. Aufgrund der hohen zeitlichen Verzögerung, mit der die Daten bereitgestellt werden, sind sie für die laufende Konjunkturanalyse allerdings ungeeignet. Zudem stellt die Verarbeitung und Analyse der Daten eine zusätzliche Herausforderung dar.[21]

2.7.1.3 Bisherige Anwendungen

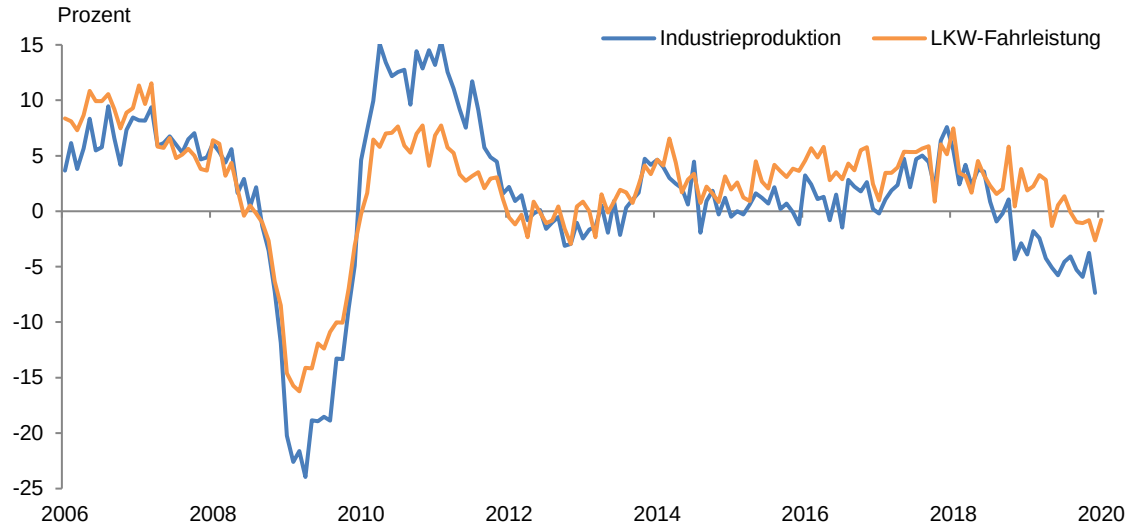
Die Mauterhebungen sind bereits in mehreren Studien für die Konjunkturanalyse und insbesondere für die Prognose der Industrieproduktion verwendet worden. Askitas und Zimmermann (2013) attestieren den Maut-Daten ein hohes Potenzial für die Konjunkturanalyse. Sie schätzen den Zusammenhang zwischen der Vorjahresveränderung verschiedener Varianten der Lkw-Fahrleistung und der Produktion und finden für den Zeitraum von 2008 bis 2010 – in dem die Produktionszahlen stark schwankten – einen sehr hohen Erklärungsgehalt für die Produktion. Allerdings vergleichen sie die Prognosegüte der Maut-Daten nicht mit anderen Frühindikatoren und führen auch keine Saisonbereinigung durch. Döhrn (2011) bereinigt die Maut-Daten zunächst um Saisoneffekte und ermittelt die Prognoseeigenschaften für die monatliche Industrieproduktion in einer Reihe von Modell-Spezifikationen. Dabei konzentriert sich die Analyse auf die mautpflichtige Fahrleistung und nicht auf die Anzahl der erfassten Mautfahrten, da die Fahrleistung offenbar höher mit der Industrieproduktion korreliert ist (Döhrn 2011: 12). Insgesamt schneidet die Maut-Fahrleistung in keiner der gewählten Spezifikationen besser ab als das ifo-Geschäftsklima für das Verarbeitende Gewerbe, das als umfragebasierter Indikator sogar noch früher (vor Ablauf des Berichtsmonats) zur Verfügung steht. Positiv wird bewertet, dass die Lkw-Fahrleistung aufgrund ihrer geringen Revisionsanfälligkeit möglicherweise künftige Revisionen der amtlichen Zahlen zur Industrieproduktion antizipieren könne.

Das Bundesamt für Güterverkehr berechnet einen monatlichen Lkw-Fahrleistungsindex (Cox et al. 2018). Der Index ist sehr zeitnah verfügbar, mittlerweile bereits rund eine Woche nach dem Berichtsmonat. Er ist zudem kaum revisionsanfällig, lediglich im Folgemonat treten geringfügige Revisionen auf, da für einen kleinen Teil der erfassten Lkw-Flotte Daten mit Verzögerung nachgeliefert werden. Der

[21] In einem vom Bundesministerium für Verkehr und Infrastruktur geförderten Projekt wurden die Daten vereinheitlicht, eine Programmierschnittstelle (API) für Entwickler zur Verfügung gestellt und eine prototypische Visualisierung der Daten entwickelt: https://www.bmvi.de/SharedDocs/DE/Artikel/DG/mfund-projekte/visualisierung-von-daten-der-automatischen-verkehrszaehlung-verkehrsvis.html.

Index bildet die Fahrleistung in mautpflichtigen Kilometern ab und hat ein Saisonmuster, das dem Muster der Industrieproduktion stark ähnelt (ebd.: 19). Insgesamt weisen Fahrleistungsindex und Industrieproduktion in Vorjahresveränderungen einen recht hohen Gleichlauf auf (Abbildung 9). Auch die monatlichen Zuwachsraten der Industrieproduktion und des Fahrleistungsindex – jeweils saison- und kalenderbereinigt – weisen für den Zeitraum von 2006 bis 2020 eine recht hohe Korrelation von 0.56 auf. Da der Fahrleistungsindex etwa einen Monat vor der Industrieproduktion vorliegt, spricht dies dafür, dass er grundsätzlich für die Prognose der Produktion geeignet ist. Eine empirische Auswertung der Prognosegüte für die Industrieproduktion, auch im Vergleich zu anderen Frühindikatoren, liegt jedoch noch nicht vor.

Abbildung 9:
Lkw-Fahrleistungsindex und Industrieproduktion 2006–2020[a]



[a]Monatsdaten, saison- und kalenderbereinigt; Veränderung gegenüber dem Vorjahr.

Quelle: Deutsche Bundesbank (2020); Statistisches Bundesamt (2020d); eigene Berechnungen.

Eine Autobahnmaut existiert auch in Österreich, wo Lkw seit dem Jahr 2004 mautpflichtig sind. Die Österreichische Nationalbank berechnet aus der Lkw-Fahrleistung einen monatlichen Indikator für die Exportentwicklung, der rund zwei Wochen nach jedem Berichtsmonat veröffentlicht wird. Für den Zeitraum bis zum März des Jahres 2009 wies dieser arbeitstägig bereinigte Indikator für die Vorjahresraten insbesondere für den Export einen hohen Erklärungsgehalt auf (Fenz und Schneider 2009). Die Fahrleistung auf Strecken durch industriell geprägte Regionen wie der Steiermark (Korridor Süd), die weniger von Durchgangsverkehr betroffen sind, lieferte zudem den höchsten Erklärungsgehalt für die Industrieproduktion. Eine Analyse der Prognosegüte des monatlich veröffentlichten Exportindikators für die amtlichen Exportzahlen liegt allerdings nicht vor.

Daten der Straßenverkehrszählung sind in Deutschland bisher nicht systematisch für makroökonomische Analysen ausgewertet worden. Doch auch in anderen Ländern gibt es vergleichbare Systeme zur digitalen Erfassung des Verkehrsaufkommens und die daraus resultierenden Daten wurden zum Teil bereits für Konjunkturanalysen genutzt. Rowland et al. (2019) werten die Verkehrszählungen im Vereinigten Königreich aus und aggregieren die Daten zu Indikatoren zur Verkehrsaktivität im Umkreis der zwölf wichtigsten britischen Seehäfen sowie landesweit. Besonderes Augenmerk wird dabei auf größere Fahrzeuge gelegt, die im Rahmen der Verkehrszählung gesondert erfasst werden. Insgesamt ist die Streuung der so gebildeten Indikatoren hoch und die Korrelation mit der wirtschaftlichen Aktivität

offenbar gering. Zudem sind die Daten mit einer Verzögerung von zwei Monaten derzeit nicht vor anderen Frühindikatoren verfügbar.

Der „ANZ-NZ Truckometer“ (ANZ Bank 2020) misst die Verkehrsdichte auf wichtigen Verbindungsstraßen Neuseelands und aggregiert diese Daten in einen „Heavy Traffic Index“ (Lkw) und einen „Light Traffic Index“ (Pkw und Vans). Die zugrunde liegenden Daten werden bereits sieben Tage nach dem Berichtsmonat von der New Zealand Transport Agency veröffentlicht. Der Index für den Lkw-Verkehr, der seit dem Jahr 2002 verfügbar ist, weist einen hohen Gleichlauf mit dem Bruttoinlandsprodukt auf. Der saisonbereinigte Index wird monatlich ca. zehn Tage nach dem Berichtsmonat veröffentlicht und lediglich im Folgemonat nur leicht revidiert. Mit den Indikatoren werden auch Prognosen zur Entwicklung der Wirtschaftsleistung im jeweiligen Quartal erstellt und veröffentlicht. Die Prognosegüte ist allerdings noch nicht systematisch ausgewertet worden. Daten aus denselben Verkehrszählungen gehen zudem in einen Nowcast für das Bruttoinlandsprodukt ein (Susnjak und Schumacher 2018). Neben den Verkehrsdaten gehen auch Daten aus anderen Quellen in ein Modell ein, das diese mittels Methoden des maschinellen Lernens auswertet und Prognosen für die gesamtwirtschaftliche und regionale wirtschaftliche Aktivität liefert. Die Auswertung bisheriger Prognosen liegt jedoch nur für einige Quartale vor und es fehlt auch hier eine systematische Analyse der Prognosegüte im Vergleich zu anderen gängigen Frühindikatoren.

Zur Kurzfristprognose der finnischen Wirtschaftsleistung verwenden Fornaro und Luomaranta (2020) Daten zum Lkw-Verkehrsaufkommen aus der automatischen Verkehrszählung an rund 500 Messpunkten. Die Daten liegen jeweils gegen Mitte des Folgemonats vor und erlauben somit bereits frühzeitig eine BIP-Schnellschätzung auf Monatsbasis. Alle Variablen – neben dem Bruttoinlandsprodukt auch umfragebasierte Umsatzzahlen von Unternehmen, die als zusätzlicher Indikator verwendet werden – gehen als Vorjahresraten in die Prognosemodelle ein. Dadurch kann zwar eine Saisonbereinigung der Daten umgangen werden, es verstellt aber den Blick auf die Prognosegüte hinsichtlich der Zuwachsrate der saisonbereinigten wirtschaftlichen Aktivität gegenüber dem Vorquartal, die für Konjunkturprognosen von größerem Interesse ist. Im Ergebnis haben die BIP-Prognosen auf Basis der Umsatzzahlen offenbar eine höhere Prognosegüte als die Daten der Verkehrszählung.

2.7.1.4 Potenziale und Grenzen

Grundsätzlich dürfte ein Zusammenhang zwischen Verkehrsaufkommen und wirtschaftlicher Aktivität – insbesondere zur Industrieproduktion – bestehen, der aufgrund der frühen Verfügbarkeit von Verkehrsdaten für die Prognose nutzbar gemacht werden kann. Zwar geben die verfügbaren Daten keine Auskunft über die Lkw-Ladungen, aber die vorliegenden Studien deuten darauf hin, dass das Lkw-Verkehrsaufkommen grundsätzlich für die Prognose der wirtschaftlichen Aktivität geeignet ist – auch wenn noch unklar ist, wie die Prognosegüte im Vergleich zu anderen Frühindikatoren zu bewerten ist. Für Deutschland liegen dazu noch keine jüngeren Studien vor. Darüber hinaus könnten Verkehrsdaten insbesondere bei größeren makroökonomischen Schwankungen oder bezüglich der Auswirkungen bedeutender makroökonomischer Ereignisse frühzeitig wertvolle Signale liefern.

Zusätzliche Potenziale könnte eine Auswertung der detaillierten Rohdaten liefern; bislang sind in makroökonomischen Studien nur aggregierte Indikatoren für das Verkehrsaufkommen verwendet worden. Sie könnten helfen, um genauer zwischen binnenwirtschaftlichem und grenzüberschreitendem Lkw-Verkehr sowie Durchgangsverkehr, der für die inländische wirtschaftliche Aktivität keine größere Bedeutung haben dürfte, zu unterscheiden. Dadurch könnte insbesondere der Erklärungsgehalt für die inländische Produktion erhöht werden. Zudem könnte der Informationsgehalt der Daten vergrößert werden, wenn sie mit regionalen Daten (beispielsweise zur jeweiligen Industriestruktur oder dem

Standort von Flug- und Seehäfen) in Verbindung gesetzt würden. Das Verkehrsaufkommen könnte dann mit der wirtschaftlichen Bedeutung der jeweiligen Region gewichtet werden, um so treffsichere Signale für die laufende wirtschaftliche Entwicklung zu erhalten. Dies wäre freilich mit einem erheblichen Aufwand verbunden. Analysen, in welchem Ausmaß die dadurch gewonnenen Informationen über den Informationsgehalt der bereits zeitnah verfügbaren Indikatoren hinausgehen, stehen noch aus.

2.7.2 Schiffspositionsdaten

Mit Schiffspositionsdaten können die Bewegungen aller größeren Frachtschiffe weltweit verfolgt werden. Damit stellen sie eine relevante Informationsquelle für makroökonomische Analysen dar, befördern doch Containerschiffe, Tanker und Massengutfrachter einen Großteil der weltweit gehandelten Güter. Allerdings sind diese Daten bislang noch kaum in der wirtschaftswissenschaftlichen Forschung verwendet worden.

2.7.2.1 Datenquellen

Schiffspositionsdaten werden in großer Menge durch das Automatische Identifikationssystem (AIS) erhoben, das im Jahr 2006 weltweit eingeführt wurde. Alle Frachtschiffe mit einer Größe über 500 Bruttoregistertonnen (BRT), unabhängig davon ob sie Container, Öl oder andere Güter befördern, tragen AIS-Sender an Bord, die mehrmals pro Minute Informationen versenden. Auf internationalen Routen gilt diese Regel bereits für Schiffe ab 300 BRT und damit effektiv für alle Frachtschiffe.[22] Satelliten und terrestrische Stationen sammeln diese Daten. Während die Daten einzelner terrestrischer Stationen zum Teil frei verfügbar sind, werden die Daten mit weltweiter Abdeckung in der Regel von kommerziellen Anbietern, wie MarineTraffic und VesselTracker, gesammelt und aufbereitet.

2.7.2.2 Datenbeschaffenheit

AIS-Daten sind grundsätzlich sehr zeitnah verfügbar. Häufig bieten Agenturen historische Daten zu Schiffspositionen rückwirkend bis zum Jahr 2014 an. Zwar liegen vereinzelt auch Daten für die Zeit vor dem Jahr 2014 vor; dann ist jedoch keine weltweite Abdeckung mehr gewährleistet. Grundsätzlich werden AIS-Daten tagesaktuell durch terrestrische Stationen und Satelliten erhoben. Eine Verarbeitung durch Agenturen erfolgt je nach Unternehmen ebenfalls tagesaktuell oder mit einer Verzögerung von etwa zwei Wochen.

Neben den einfachen Positionsdaten, die neben den Identifikationsnummern der Schiffe (IMO und MMSI Nummer) auch Geschwindigkeit, Fahrtrichtung und Tiefgang enthalten, bieten Unternehmen, wie MarineTraffic und VesselTracker, auch sogenannte „Port Calls“-Informationen kostenpflichtig an. Diese geben darüber Auskunft, wann Schiffe an Häfen haltgemacht haben, enthalten jedoch keine Information über die umgeschlagenen Güter. Gemeinsam mit Positionsdaten lassen die „Port Call“-Informationen die Rekonstruktion der Bewegungsprofile der Schiffe zu.

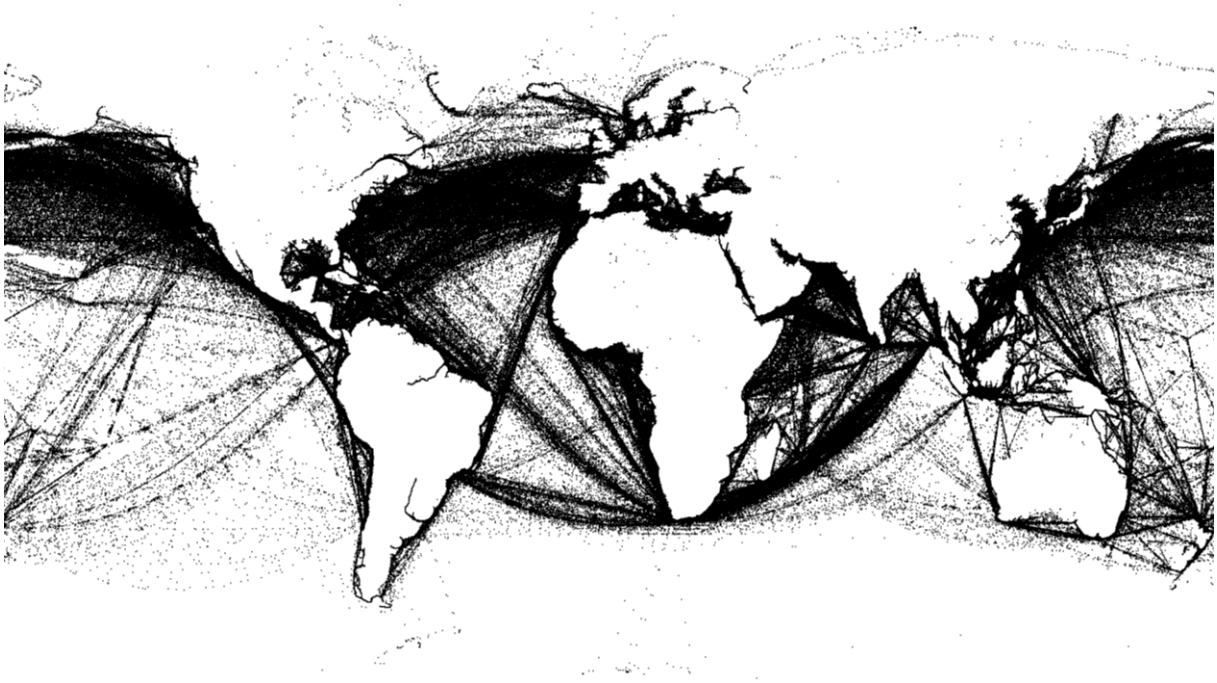
Ein typischer Datensatz könnte die folgenden Maße annehmen: Um das Bewegungsprofil der gesamten Containerschifffflotte in einem Jahr zu rekonstruieren, bezieht man tägliche Positionsdaten der Containerschiffe und deren „Port Calls“. Bei etwa 5 500 Containerschiffen weltweit entspricht dies etwa einem Datensatz von 2 Millionen Schiffspositionen und 500 000 „Port Calls“. Dieser Datensatz ist ausreichend, um die Bewegung der Schiffe gut nachzuvollziehen. Bei Bedarf können auch Positionen

[22] Bereits kleinere Containerschiffe mit einer Kapazität für etwa 100 bis 200 Container weisen eine Größe von über 2 500 BRT auf.

mit einer höheren Frequenz als einer Position pro Tag angefordert werden. Zwar sind bei perfekter Abdeckung über 1 000 Positionen am Tag pro Schiff denkbar, allerdings werden auf hoher See in der Regel nicht alle Radiosignale des AIS-Systems durch Satelliten erfasst.

Die weltweite Abdeckung mit terrestrischen Stationen und Satelliten ist mittlerweile so gut, dass ein komplettes tagesaktuelles Bild der gesamten Schiffsflotte auf allen Weltmeeren gewährleistet ist (Abbildung 10). Die Datenqualität der Rohdaten ist somit sehr gut.

Abbildung 10:
Tägliche Positionen der Containerschiffflotte 2014–2016



Quelle: Fleetmon (2020); eigene Berechnungen.

2.7.2.3 Methodik

Die Rohdaten zu den Schiffspositionen können auf unterschiedliche Weisen ausgewertet werden. So können lange Zeitreihen von Positionsdaten analysiert werden, indem man retrospektiv beispielsweise in Europa einlaufende Schiffe mit ihren alten Aufenthaltsorten verknüpft. Dadurch entstehen Verbindungen zwischen den vergangenen Halten und dem Ankunftsland. Hierfür sind vergleichsweise einfache Algorithmen auf Programmiersprachen wie Python und R ausreichend. Bei umfassenderen Untersuchungen bieten sich Sprachen wie C++ an, und der Einsatz von Servern mit sehr großem Arbeitsspeicher ist empfehlenswert.

Für umfangreichere Analysen können die Schiffspositionsdaten mit anderen Datensätzen verknüpft werden. Beispielsweise ist es sinnvoll, den Schiffen ihre Frachtkapazität und ihren technisch möglichen Tiefgang über Schiffsdatenbanken zuzuordnen. Damit lässt sich abschätzen, wie stark ein Schiff an jedem beliebigen Zeitpunkt beladen ist bzw. wie schwer die geladene Menge ist. Eine Ergänzung um geografische Daten, wie den Umrissen von Kontinenten, ermöglicht es zudem, die gefahrene Distanz zwischen Positionspunkten des Schiffes zu approximieren.

Mittels weiterführender Methoden, wie dem maschinellen Lernen, können die AIS-Daten mit anderen Daten in Verbindung gesetzt werden, um so die Anwendungsfelder zu erweitern: So können Programme

„lernen“, eine Momentaufnahme der Rohdaten mit gesamtwirtschaftlichen Größen wie der Export- oder Importmenge zu assoziieren. Durch den Vergleich mit historischen Daten könnte damit beispielsweise ein Zusammenhang zwischen der Anzahl der Containerschiffe im Suezkanal und des zukünftigen Importvolumens in Deutschland hergestellt werden. Methoden des maschinellen Lernens sind vor allem aufgrund des großen Datenvolumens für diese Aufgabenbereiche besonders nützlich.

2.7.2.4 Bisherige Anwendungen

AIS-Daten werden historisch bedingt vor allem in der Logistikbranche selbst verarbeitet. Beispielsweise nutzen Häfen die Positionen von sich nähernden Schiffen, um Abläufe an den Terminals zu koordinieren und Staus zu vermeiden. Ebenfalls finden die Daten Anwendung in Softwarelösungen, die Schiffe auf See bei der Navigation unterstützen.

Ökonomische Analysen haben bisher nur in Einzelfällen AIS-Daten verwendet. So schätzen Adland et al. (2017) die Ölexporte eines Landes mit Hilfe von AIS-Daten für Öltanker ab. Sie finden eine hohe Korrelation zwischen der Anzahl von Öltankern, die ein Land verlassen, und den Ölexporten dieses Landes. Der Zusammenhang ist besonders ausgeprägt, wenn keine anderen Transportmittel wie Pipelines zur Verfügung stehen. Brancaccio et al. (2017) entwickeln ein Modell für die Massengutfrachterflotte, mit dessen Hilfe die Forscher Handelskosten für Massengüter schätzen. Die AIS-Daten liefern dafür Informationen, wie viele leere Frachter sich in der Nähe von Häfen befinden. Schließlich verwenden Heiland et al. (2019) Schiffspositionsdaten, um abzuschätzen, wie sich die Erweiterung des Panama-Kanals im Jahr 2016 auf den Welthandel ausgewirkt hat. Sie zeigen, dass sich der Handel zwischen Ländern, für die der Kanal eine relevante Handelsroute darstellt, sichtbar erhöht hat.

2.7.2.5 Potenziale und Grenzen

Schiffspositionsdaten bieten grundsätzlich ein recht großes Potenzial für makroökonomische Analysen und Konjunkturprognosen, nicht zuletzt aufgrund ihrer großen Abdeckung und zeitnahen Verfügbarkeit. Allerdings sind ihrem Informationsgehalt auch Grenzen gesetzt. So geben die AIS-Daten zwar Auskunft darüber, welchen Weg ein Schiff nimmt und wie stark es beladen ist, sie beinhalten jedoch keine weiteren Informationen über die Ladung, und es ist nicht eindeutig nachvollziehbar, wann und wo Güter geladen bzw. gelöscht werden. Da Schiffe in einem Hafen sowohl be- als auch entladen werden können, lässt sich die Gesamtmenge der in einem Hafen gelöschten Waren nicht ermitteln. Ein Containerschiff, das z.B. in Hamburg Ladung löscht kann diese Ladung in allen vormals angelaufenen Häfen aufgenommen haben. Gleichzeitig kann ein Container aus Shanghai nach Deutschland gelangen, indem Logistikunternehmen den Container von verschiedenen Häfen wie Hamburg, Rotterdam oder Genua per Land weitertransportieren.

Gerade moderne Analysemethoden könnten aber Muster aus den Bewegungsprofilen und dem Tiefgang der Schiffe ableiten, um daraus auf die Exporte und Importe zu schließen, noch bevor andere Frühindikatoren, wie der Containerumschlagsindex, dazu vorliegen. Gegenüber dem Containerumschlag an einzelnen Häfen könnten mittels Schiffspositionsdaten zudem die geplanten Routen der Schiffe prognostiziert werden, um somit Informationen für die zukünftige Handelsaktivität zu erhalten. Aufgrund ihrer großen Abdeckung können Schiffspositionsdaten sowohl Informationen für einzelne Länder als auch für den Welthandel liefern.

2.7.3 Übrige Logistikaktivitäten

Da ein Großteil des internationalen Warenverkehrs auf dem Seeweg abgewickelt wird, kann auch der Warenumschlag in wichtigen Seehäfen – z.B. der Containerumschlag – zur Abschätzung der Entwicklung im Welthandel herangezogen werden. Im Hinblick auf den Wert der Waren ist auch der Luftverkehr ein wichtiger Verkehrsträger. Bislang liegen jedoch kaum makroökonomische Studien vor, die auf diese Daten zurückgreifen.

2.7.3.1 Datenquellen

Viele bedeutende Häfen der Welt weisen Kennzahlen und Statistiken online aus, allerdings häufig mit großen Verzögerungen und nicht als Zeitreihen. Die Hafendatenbank des Instituts für Seeverkehrswirtschaft und Logistik Bremen (ISL) enthält dagegen strukturierte, vergleichbare Informationen für rund 400 Seehäfen weltweit. Darunter sind Zeitreihen zum Containerumschlag und für ausgewählte Güterarten, die teilweise bis in das Jahr 1980 zurückreichen. Allerdings ist die Datenbank nicht kostenfrei verfügbar, und die online zugänglichen Informationen beschränken sich auf eine Klassifizierung von Seehäfen nach Größenkategorien und die jährlichen Anteile verschiedener Güterarten am Containerumschlag. Individualisierte Auswertungen der Hafendatenbank können beim ISL angefordert werden.

Ähnlich wie die Seehäfen bieten auch einzelne Flughäfen ausgewählte Statistiken an, die allerdings für Konjunkturanalysen nicht umfassend und aktuell genug sind. Die International Air Transport Association (IATA) sammelt Frachtdaten von dem Großteil (82 Prozent) der weltweit aktiven Luftfahrgesellschaften und bietet Statistiken – etwa zum monatlichen Frachtaufkommen nach Länderpaaren – kostenpflichtig zum Download an. Wie zeitnah diese Daten bereitgestellt werden ist nicht direkt ersichtlich, allerdings veröffentlicht die IATA aber rund 35 Tage nach jedem Berichtsmonat eine Pressemeldung mit einigen Statistiken zu den Entwicklungen im weltweiten Luftverkehr. Zum selben Zeitpunkt veröffentlicht auch der Flughafenverband ADV eine statistische Auswertung des Passagier- und Transportaufkommens an deutschen Flughäfen.

Amtliche Daten zum Güterverkehrsaufkommen über die Verkehrsträger Binnenschifffahrt und Eisenbahn, aber auch zur Seeschifffahrt und zum Luftverkehr, werden für Deutschland mit zwei bis drei Monaten Verzögerung nach jedem Berichtsmonat vom Statistischen Bundesamt bereitgestellt.

2.7.3.2 Datenbeschaffenheit und Methodik

Die Rohdaten zu Seeverkehrs- und Luftfrachtvolumen sind nicht frei zugänglich. Die Angaben aus verschiedenen weltweiten Häfen zum Containerumschlag, die über die Hafendatenbank des ISL zugänglich sind, beziehen sich allerdings auf international standardisierte Containergrößen und sind daher vergleichbar. Mit einer geeigneten Gewichtung für die Seehäfen, die die wirtschaftliche Bedeutung der Weltregionen im Welthandel abbilden, können die Informationen aus verschiedenen Häfen zusammengeführt werden (Döhrn und Maatsch 2012).

2.7.3.3 Bisherige Anwendungen

Das RWI veröffentlicht in Zusammenbarbei mit dem ISL einen monatlichen Containerumschlagindex. Der Index liegt ab Januar 2007 vor und ist so konstruiert, dass er eine möglichst enge Korrelation mit dem Welthandelsvolumen aufweist. Er basiert auf Daten zum Containerumschlag in über 80 Seehäfen, die rund 60 Prozent des weltweiten Containerumschlags abdecken (Döhrn 2019b). Jede Weltregion wird dabei durch einige Seehäfen abgebildet, die je nach wirtschaftlicher Bedeutung der Region für den

Welthandel gewichtet sind. Etwa 20 Tage nach jedem Berichtsmonat sind Daten aus rund 40 Seehäfen bekannt, die gewichtet etwa zwei Drittel der im Index abgebildeten Häfen ausmachen und daher eine Schnellschätzung der Entwicklung des monatlichen Frachtvolumens ermöglichen. Bei der darauffolgenden Veröffentlichung einen Monat später (ca. 55 Tage nach dem Berichtsmonat) liegen Daten zu fast allen Seehäfen vor. Der Indexwert wird dann zum Teil deutlich revidiert; zu weiteren nennenswerten Revisionen kommt es danach in der Regel nicht mehr. Döhrn (2019b) berichtet über notwendige Anpassungen der Indikatorberechnung im Zeitverlauf, etwa aufgrund von Änderungen im Berichtskreis oder durch die Berücksichtigung des chinesischen Neujahrsfestes im Rahmen der Saisonbereinigung. Die Indikatorhistorie wurde im Verlauf dieser Neudefinitionen neu berechnet und ist insofern immer konsistent mit den jeweils aktuellen Veröffentlichungen. Eine Auswertung der Prognosegüte des Indikators für den Welthandel liegt bislang nicht vor.

2.7.3.4 Potenziale und Grenzen

Der weltweite Containerumschlag wird durch die Berechnung des RWI/ISL-Containerumschlagindex bereits als Datenquelle genutzt. Die dahinter liegenden detaillierteren Informationen könnten dazu dienen, den Containerumschlag für einzelne Länder oder Regionen zeitnah zu beobachten. Freilich wird Seefracht in der Regel mit anderen Verkehrsträgern zum Teil auch grenzüberschreitend weiterbefördert, sodass ein enger Zusammenhang mit der regionalen wirtschaftlichen Aktivität nicht gegeben sein muss.

Daten für einen Großteil des weltweiten Luftfrachtaufkommens sind bei der IATA kostenpflichtig verfügbar. Potenziale für die makroökonomische Analyse könnten sich dadurch ergeben, dass per Luftfracht recht spezielle Waren (insbesondere hochpreisige Elektronik, Maschinen und Pharmazeutika) transportiert werden. Sofern diese Waren spezifische zyklische Eigenschaften aufweisen, könnte dies für Konjunkturprognosen genutzt werden. In diesem Zusammenhang könnten Luftfrachtdaten auch wertvolle Informationen zur Aktivität in einzelnen Wirtschaftszweigen liefern. Schließlich wäre denkbar, dass auf Luftfracht (unabhängig von den transportierten Waren) als Transportmittel über den Konjunkturzyklus hinweg unterschiedlich stark zurückgegriffen wird. Allerdings liegen noch keine makroökonomischen Analysen auf Basis dieser Daten vor.

2.8 Mobilfunkdaten

Die hohe Verbreitung von Mobilfunkanschlüssen hat zu einem zunehmenden Interesse von Wissenschaftlern an den daraus resultierenden Daten geführt. Die weltweite Zahl der Smartphone-Nutzer belief sich im Jahr 2018 auf rund 2,9 Milliarden. Mobilfunkdaten werden insbesondere dazu verwendet, um die Bewegungsprofile von Nutzern auszuwerten. Einige Studien versuchen aber auch, von der Nutzung von Mobilfunkanschlüssen auf die wirtschaftliche Aktivität zu schließen.

2.8.1 Datenquellen

Mobilfunkdaten werden von den Mobilfunkanbietern erfasst und können von diesen kostenpflichtig bezogen werden. Die Internationale Fernmeldeunion (ITU), die sich als Sonderorganisation der Vereinten Nationen technischen Aspekten der Telekommunikation widmet, stellt zudem kostenlos umfangreiche internationale Daten zur Telekommunikation zur Verfügung. Diese umfassen z.B. die jährliche weltweite Nutzung von Mobiltelefonen bezüglich Haushalten, Ländern und Regionen. Die

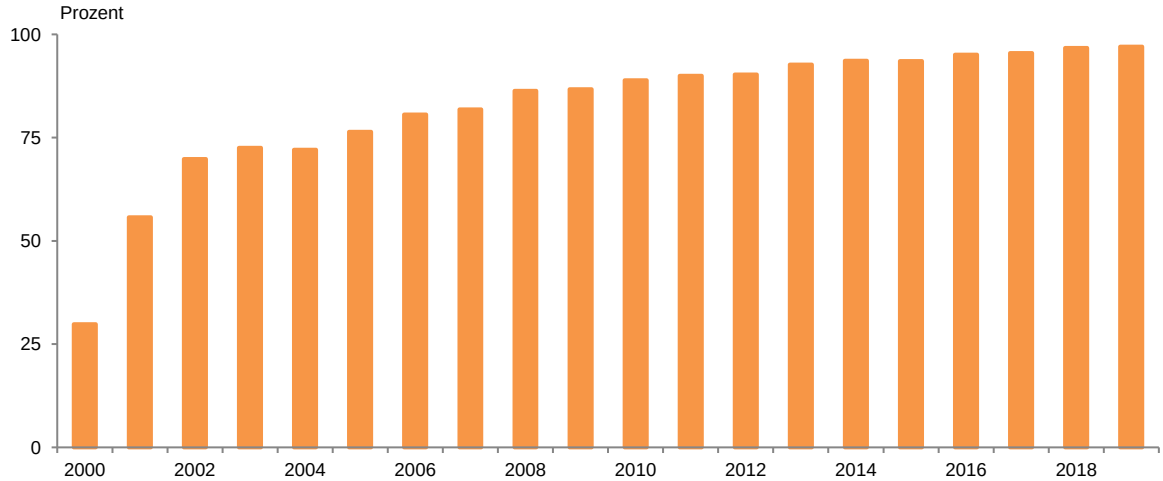
Daten werden relativ zeitnah bereitgestellt. So waren vorläufige Daten für das Jahr 2019 bereits im November 2019 verfügbar.

2.8.2 Datenbeschaffenheit

Mobilfunkdaten bestehen üblicherweise aus sogenannten Call Detail Records (CDR). Diese liefern eine Kennung des Mobilgeräts, den Aufenthaltsort, Art der Verbindung (Daten, Anruf usw.) sowie Uhrzeit und Dauer des Anrufs (Ricciato et al., 2015). Die Daten stehen grundsätzlich also in einem sehr hohen Detailgrad und in sehr hoher Frequenz zur Verfügung. Einige Datensätze erlauben durch die Verbindung von Mobilfunkdaten und Zielfunkmasten auch, die Mobilität von Nutzern zu verfolgen. Ein weiterer Vorteil der Daten ergibt sich aus der zumeist sehr breiten Abdeckung sowie der unmittelbaren Verfügbarkeit. Jährliche Daten liegen frühestens ab dem Jahr 1990 vor. Daten höherer Frequenz sowie detailliertere Daten sind zumeist erst ab dem Jahr 2005 verfügbar.

Vor allem in den Entwicklungsländern sind Mobiltelefone weit mehr verbreitet als Internetanschlüsse. In afrikanischen Volkswirtschaften liegt der Anteil der Haushalte mit Internetanschluss bei unter 10 Prozent, wohingegen mehr als 80 Prozent der Haushalte Zugang zu Mobiltelefonen haben. Bereits im Jahr 2014 lag laut der Internationalen Fernmeldeunion (ITU) die durchschnittliche Quote von Mobilfunkabonnenten weltweit bei 96,4 pro 100 Einwohner, mit niedrigeren Durchschnittswerten in Asien (89,2) und Afrika (69,3). In Deutschland hat sich der Anteil der Haushalte mit Mobiltelefonen seit Anfang der 2000er Jahre deutlich erhöht (Abbildung 11). Im Jahr 2019 lag der Anteil bei 97 Prozent, wobei die absolute Anzahl der Anschlüsse die Einwohnerzahl übersteigt.

Abbildung 11:
Anteil der privaten Haushalte in Deutschland mit einem Mobiltelefon 2000–2019[a]



[a]Jahresdaten.

Quelle: Statistisches Bundesamt (2020e).

2.8.3 Bisherige Anwendungen

Bislang sind Mobilfunkdaten für makroökonomische Analysen vornehmlich eingesetzt worden, um aus der Nutzung von Mobilfunkanschlüssen Rückschlüsse auf den Wohlstand oder die Wirtschaftsleistung zu ziehen. Freilich dürfte dies vor allem für Entwicklungs- und Schwellenländer relevant sein. Aus Sicht der amtlichen Statistik können Mobilfunkdaten zur Erhebung von Bevölkerungs- und Pendlerstatistiken

beitragen. Für makroökonomische Prognosen, insbesondere in fortgeschrittenen Volkswirtschaften spielen Mobilfunkdaten bislang kaum eine Rolle.

Einige Studien versuchen, einen Zusammenhang zwischen Mobilfunknutzung und wirtschaftlicher Aktivität herzustellen. So korrelieren regionale Muster der Telefon- und Internetnutzung offenbar mit dem regionalen Wohlstand und der Wirtschaftsleistung (Eagle et al. 2010; Llorente et al. 2015). Eine Studie von Deloitte (2012) nutzt Daten des Visual Networking Index von Cisco für 14 Länder und kommt zu dem Ergebnis, dass eine Verdoppelung der mobilen Datennutzung zu einem Anstieg der Zuwachsrate des Bruttoinlandsprodukts pro Person um 0,5 Prozent führt. Da die Daten lediglich auf Jahresbasis und für einen sehr kurzen Zeitraum verfügbar sind, basiert die Studie allerdings auf sehr wenigen Beobachtungen. Zudem wird nicht berücksichtigt, dass die Kausalität auch umgekehrt verlaufen könnte, also dass eine höhere wirtschaftliche Aktivität zu einer stärkeren mobilen Datennutzung führen könnte (Baldacci et al. 2016).

Minges (2016) liefert in diesem Zusammenhang eine Übersicht über Studien hinsichtlich ökonomischer Effekte einer verbesserten Internet- und Telefonanbindung. Thompson und Garbacz (2011) analysieren den Zeitraum von 2005 bis 2009 für 40 Volkswirtschaften mit niedrigem Einkommen und schätzen, dass eine Zunahme bei der Festnetz-Breitbandversorgung bei Haushalten um zehn Prozentpunkte das Bruttoinlandsprodukt pro Haushalt um 0,8 Prozent erhöht, wohingegen eine Zunahme der mobilen Breitbandverbindung für Haushalte um zehn Prozentpunkte das BIP pro Haushalt um 0,5 Prozent verringert. Diesen Unterschied erklären die Autoren durch potenziell unproduktive Anwendungen von mobilen Breitbandverbindungen.

Gruber und Koutroumpis (2011) analysieren die Auswirkungen der mobilen Telekommunikation auf das Wirtschaftswachstum basierend auf jährlichen Daten aus 192 Ländern über den Zeitraum von 1990 bis 2007. Während in Ländern mit niedrigem Einkommen der Beitrag des Mobilfunks zum jährlichen BIP-Wachstum 0,1 Prozent beträgt, liegt dieser bei Ländern mit hohem Einkommen bei 0,2 Prozent. Sie identifizieren zudem positive Auswirkungen auf das Produktivitätswachstum. Freilich lassen solche Studien keine Rückschlüsse auf die Konjunktur oder Wachstum in fortgeschrittenen Volkswirtschaften wie Deutschland zu, in denen bereits seit vielen Jahren ein Großteil der Bevölkerung Mobilfunk-anschlüsse nutzt.

Es findet bisweilen auch ein Vergleich der Effekte eines verbesserten Zugangs zu Festnetz, Mobilfunk und Internet auf das Bruttoinlandsprodukt pro Kopf statt. Die Auswirkungen eines Internetanschlusses werden hierbei zumeist am stärksten eingeschätzt, wobei in Entwicklungs- und Schwellenländern die Effekte einer verbesserten Mobilfunkanbindung überwiegen können (Minges 2016). Allerdings basieren all diese Studien auf einer sehr geringen Anzahl an Beobachtungen und ihnen liegen vergleichsweise einfache Methoden zugrunde, sodass unklar ist, ob sie sich verallgemeinern lassen.

Mobilfunkdaten sind auch dazu verwendet worden, um die Entwicklung am Arbeitsmarkt zu prognos-tizieren. Toole et al. (2015) und Sundsøy et al. (2016) identifizieren von Arbeitslosigkeit betroffene Personen und prognostizieren Veränderungen der aggregierten Arbeitslosenquoten unter Verwendung von Call Detail Records (CDRs) von Mobiltelefonen. Grundlage ist z.B. die Beobachtung, dass von Entlassungen betroffene Personen signifikante Rückgänge im Sozialverhalten und in der Mobilität nach dem Verlust des Arbeitsplatzes aufzeigen.

Dong et al. (2017) nutzen Bewegungsprofile und Geo-Positionierungsdaten von Mobiltelefonen, um daraus auf die Beschäftigung und die Konsumaktivität in China zu schließen. Dazu identifizieren sie zunächst Beschäftigte anhand ihres Bewegungsprofils zu typischen Arbeitszeiten und bringen diese Information in Verbindung mit den jeweiligen Standorten der Nutzer. Sie evaluieren ihren Ansatz anhand von zwei angekündigten Massenentlassungen sowie zwei regional begrenzten Phasen eines

starken Beschäftigungsaufbaus und zeigen, dass diese Entwicklungen durch Mobilfunkdaten abgebildet werden können. Auf vergleichbare Art und Weise kombinieren sie die Bewegungsprofile mit Standorten, an denen Einzelhandelsgeschäfte stark vertreten sind. Sie zeigen, dass ihr aus diesen Informationen gebildeter Konsumentenindex eine recht hohe Korrelation mit den Umsätzen ausgewählter Unternehmen aufweist.

Mobilfunkdaten wurden zuletzt auch verstärkt verwendet, um die Auswirkungen der Corona-Pandemie auf Mobilität und soziale Interaktionen zu analysieren (Alexander and Karger 2020; Alfaro et al. 2020). Eine aktuelle Studie von Goolsbee und Syverson (2020) greift auf Mobilfunkdaten zurück, um die Auswirkungen der Lockdown-Maßnahmen auf die wirtschaftliche Entwicklung in den Vereinigten Staaten zu analysieren. Die Daten umfassen Mobilitätsdaten von rund 45 Millionen Mobilfunknutzern, mit denen die Besucherfrequenz in etwa 2,25 Millionen Geschäften gemessen wird. Die Autoren kommen zu dem Ergebnis, dass von dem im März und April durchschnittlich zu verzeichnenden Besucherrückgangs um bis zu 60 Prozent lediglich zu sieben Prozentpunkten unmittelbar auf die staatlichen Maßnahmen zurückgeführt werden kann. Individuelle Verhaltensanpassungen, um eine Infektion zu vermeiden, haben demgegenüber eine deutlich größere Rolle gespielt. Eine Verringerung von Lockdown-Maßnahmen führt demnach für sich genommen auch nur wieder zu einem vergleichsweise geringen Anstieg der wirtschaftlichen Aktivität.

Ferner sind Mobilfunk- und Internetdaten dazu verwendet worden, um sozioökonomische Eigenschaften von Personen abzuleiten. Demzufolge spiegeln Mobilitätsmuster von Personen ihre sozialen Interaktionen wider. Empirische Ergebnisse für Afrika zeigen, dass es möglich ist, daraus auf den Wohlstand von Personen sowie auf die Einkommensverteilung über Regionen mittels Mobilfunkdaten zu schließen (Blumenstock et al. 2015).

Mobilfunkdaten erlauben auch Aussagen über die Bevölkerungsverteilung (Deville et al. 2014; Laurila et al. 2012). Mobilfunkdaten werden in diesem Zusammenhang insbesondere in Gebieten eingesetzt, in denen traditionelle demografische Erhebungen besonders teuer oder gefährlich sind; beispielsweise in afrikanischen Staaten (Blondel et al. 2015; Ricciato et al. 2015). Sie können dort in einigen Fällen zur Verbesserung nationaler Statistiken beitragen (Blumenstock et al. 2015). Smith-Clarke et al. (2014) setzen beispielsweise Anrufdaten ein, um den Lebensstandard in Entwicklungsländern abzuschätzen. Sie verwenden aggregierte Aufzeichnungen über Anrufdetails von Mobilfunkteilnehmern und extrahieren Merkmale, die stark mit aus Volkszählungsdaten abgeleiteten Armutsindizes korrelieren.

Aber auch in fortgeschrittenen Volkswirtschaften werden Mobilfunkdaten von den Statistikämtern zur Beobachtung der Bevölkerungsentwicklung eingesetzt. Das Statistische Bundesamt kooperiert beispielsweise mit Tochterunternehmen der Deutschen Telekom AG, um die Tages- und Wohnbevölkerung mit Hilfe der Mobilfunkdaten bundesweit valide abzubilden und zu schätzen. Konkret wird zum Beispiel die Korrelation von Mobilfunkaktivitäten des Jahres 2017 und den Bevölkerungszahlen des Zensus 2011 analysiert. Aktuell erfolgen erste Analysen basierend auf anonymisierten Daten. Die Ergebnisse legen den Schluss nahe, dass die Bevölkerung mit den vorliegenden Mobilfunkdaten teilweise gut abgebildet wird. Um eine bundesweite Repräsentativität der Daten sicherzustellen, müssten nach Einschätzung des Statistischen Bundesamts möglichst Daten aller Mobilfunkanbieter in Deutschland verfügbar gemacht werden (Statistisches Bundesamt 2019). In einem ähnlichen Zusammenhang können Daten von Mobiltelefonen helfen, um Tourismus- oder Migrationsströme zu beobachten.

2.8.4 Potenziale und Grenzen

Mobilfunkanschlüsse werden von einem Großteil der Bevölkerung genutzt. Die daraus resultierenden Daten sind grundsätzlich regional sehr disaggregiert, in hoher Frequenz und zeitnah verfügbar. Aus

diesem Grund können sie insbesondere für die Erhebung von Bevölkerungsstatistiken einen wertvollen Beitrag leisten. Zwar gibt es Ansätze, von der Mobilfunknutzung auf die Arbeitsmarktentwicklung zu schließen. Allerdings liegen Daten für den Arbeitsmarkt zeitnah vor und die vorliegenden Indikatoren für die Prognose scheinen derzeit zuverlässiger zu sein. Ein Zusammenhang zwischen der Verbreitung oder Nutzungsintensität und der wirtschaftlichen Aktivität in fortgeschrittenen Volkswirtschaften dürfte aufgrund der bereits weiten Verbreitung von Mobilfunkanschlüssen und des unklaren Zusammenhangs nur schwer herzustellen sein.

Zusätzliches Potenzial könnten Mobilfunkdaten haben, wenn bei einschneidenden, wirtschaftlich relevanten Ereignissen anhand der Bewegungsprofile der Mobilfunknutzer zeitnah auf die Auswirkungen und den zeitlichen Verlauf geschlossen werden kann; beispielsweise, wenn aufgrund bestimmter Ereignisse die Mobilität erkennbar eingeschränkt wird. Ferner könnten die Bewegungsprofile mit Informationen zu wirtschaftlich relevanten Standorten, wie Einzelhandelszentren, verknüpft werden, um daraus auf die Konsumaktivität zu schließen. Bislang liegt allerdings noch keine belastbare wissenschaftliche Evidenz dazu vor, inwieweit solche recht aufwendigen Verfahren der Konjunkturanalyse zu Gute kommen würden. Vor diesem Hintergrund scheinen die Potenziale von Mobilfunkdaten für makroökonomische Analysen derzeit eher begrenzt zu sein.

3 Zukünftige Potenziale von Big Data für die makroökonomische Analyse

Die Potenziale von Big Data für die makroökonomische Analyse werden sich zukünftig weiter vergrößern. Dies betrifft sowohl Daten, die zwar bereits erfasst werden, aber derzeit aufgrund verschiedener Hürden (wie z.B. Datenschutz, technischer oder methodischer Hemmnisse) noch nicht systematisch verwendet werden können, als auch neue Daten, die im Zuge der fortschreitenden Digitalisierung oder durch neue Geschäftsmodelle anfallen werden. Diese Potenziale beziehen sich nicht nur auf die laufende Konjunkturbeobachtung, sondern auch auf die Beantwortung struktureller makroökonomischer Fragestellungen, die für Konjunkturprognosen und die Ableitung adäquater wirtschaftspolitischer Maßnahmen relevant sind.

Bereits jetzt wird ein Großteil der erbrachten Wertschöpfung und der Preisentwicklungen digital erfasst. So dokumentieren und steuern viele Unternehmen ihre Aktivitäten auf der Basis einschlägiger Unternehmenssoftware, in der alle relevanten Aspekte ihrer Tätigkeit elektronisch einfließen. Statistikämter und andere Datenanbieter arbeiten daran, diese Daten sukzessive in ihre Erhebungen zu integrieren. Für eine stärkere Nutzung dieser Datenquellen gibt es derzeit noch verschiedene Hürden. Dazu zählen Fragen des Datenschutzes nicht nur bezüglich des systematischen Auswertens dieser Daten, sondern auch im Hinblick auf die Verknüpfung mit bereits vorliegenden Datensätzen. Zudem wäre es zunächst für die erhebenden Statistikämter und auch für die Auskunftsgebenden mit einem erheblichen zusätzlichen Aufwand verbunden, all diese Daten systematisch zu sammeln. Sind die notwendigen Umstellungen jedoch einmal erfolgt, würde sich der Aufwand deutlich verringern, da der Datenaustausch dann vollständig automatisiert erfolgen könnte. Schlussendlich könnte die Auswertung dieser Daten eine zeitnahe und nahezu vollständige Erfassung der Wertschöpfung, auch nach Wirtschaftsbereichen getrennt, und der Preise ermöglichen. In diesem Zuge würden dann auch Datenrevisionen, die bislang die Analyse der laufenden Konjunkturentwicklung erschweren, weitgehend entfallen. In der Praxis werden solche Ansätze jedoch nur nach und nach in die amtliche Statistik integriert werden können. Hier wären vor allem solche Bereiche von Interesse, die eine relativ große Bedeutung für die

wirtschaftliche Entwicklung haben und gleichzeitig von bislang verfügbaren Daten und Indikatoren verhältnismäßig schlecht abgedeckt sind. Entstehungsseitig zählen dazu die Dienstleistungsbranchen, verwendungsseitig die privaten Konsumausgaben und die Vorratsveränderungen. Insgesamt könnten sich dadurch auch Konjunkturprognosen, die auf der Einschätzung der laufenden Entwicklung aufbauen, spürbar verbessern.

Neben der besseren Beobachtung der konjunkturellen Entwicklung bieten Big Data auch Potenziale, um strukturelle makroökonomische Zusammenhänge besser abzuschätzen und verstehen zu können. Darauf aufbauend könnten bessere Konjunkturprognosen erstellt und zielgerichteter wirtschaftspolitische Maßnahmen abgeleitet werden.

Sowohl für die Konjunkturprognose als auch für die Wirtschaftspolitik ist relevant, ob die laufende wirtschaftliche Entwicklung vornehmlich durch angebotsseitige oder nachfrageseitige Schwankungen geprägt ist. So hat die Diagnose, ob Ölpreis-Schwankungen durch das Angebot oder die Nachfrage verursacht werden, maßgeblichen Einfluss auf die zukünftigen wirtschaftlichen Auswirkungen (Kilian 2009). In vielen anderen Bereichen ist es derzeit jedoch gerade am aktuellen Rand schwer, zwischen angebots- oder nachfrageseitig verursachten Schwankungen zu unterscheiden.

In Deutschland besteht keine originäre Lagerstatistik für die Konjunkturanalyse am aktuellen Rand. Daher stellen die Vorratsveränderungen in den Erstveröffentlichungen der VGR im Wesentlichen eine auf Plausibilität geprüfte Residualgröße dar, die entsprechend revisionsanfällig ist (Wollmershäuser 2016). Big Data-Konzepte, die an den Warenwirtschaftssystemen der Unternehmen ansetzen, haben das Potenzial, diese Lücke zu schließen. Zudem würde neben der Lagerveränderung auch die Lagerhöhe für die statistische Messung zugänglich. Bestandsdaten können dann herangezogen werden, um Anhaltspunkte darüber zu gewinnen, ob ein Lageraufbau geplant (Bewegung hin zum längerfristigen Mittel) oder ungeplant (Bewegung weg vom längerfristigen Mittel) erfolgt, um so auf angebots- oder nachfrageseitig bedingte Schwankungen der Konjunktur zu schließen.

Im Zentrum der Konjunkturanalyse stehen Auslastungsschwankungen der gesamtwirtschaftlichen Produktionskapazitäten. Während der Arbeitseinsatz über die geleisteten Stunden grundsätzlich konzeptionell adäquat als Stromgröße erfasst wird, liegt für den Faktor Sachkapital nur eine Bestandsrechnung vor, wobei auch hier eine Messung der tatsächlich in den Produktionsprozess eingehenden Kapitaldienste wünschenswert wäre (Bickenbach et al. 2016a). Hier könnte zukünftig anhand der sich ausbreitenden Sensorik im Kontext des „Internet-of-Things“ angesetzt werden. Sensoren verfolgen die Leistungsabgabe einzelner Maschinen primär für Wartungs- und Ersatzentscheidungen der Hersteller bzw. Nutzer der Ausrüstungsgüter. Damit liefern sie aber zugleich auch für die konjunkturelle Einschätzung wichtige Signale über die tatsächliche Inanspruchnahme der installierten sachlichen Produktionskapazitäten. Zur Aggregation der Einzelsignale könnte in erster Annäherung eine Gewichtung über die Anschaffungspreise vorgenommen werden. Ferner ließe sich mit individuellen Leistungsdaten von Kapitalgütern die Abschreibungsrechnung verfeinern, die bislang auf recht groben Annahmen über die Lebensdauer und das Abgangsmuster von Ausrüstungsgüterklassen beruht.

B2B-Handelsplattformen verfolgen den Zweck, Anbieter und Nachfrager zusammenzubringen. Auf diese Weise entstehen sowohl Daten, die auf die Nachfrageintensität schließen lassen (Anfragen auf Angebote), als auch Daten, die die Anbieterreaktion (Abgabe von Angeboten) abbilden. Diese Aktivität ist der Auftragsvergabe vorgelagert. Gegenüber der Auftragseingangsstatistik ergibt sich damit ein zeitlicher Vorlauf, der durch die Echtzeit-Erfassung noch vergrößert wird. Wichtiger noch kann die Information über die Nachfrage- und Angebotsintensität sein, die im Konjunkturverlauf deutlich schwanken dürfte. Während sich im Auftragseingang nur realisierte Vertragsabschlüsse niederschlagen, können über die vorgelagerten Suchprozesse auch nachfrage- und angebotsseitige Triebkräfte identi-

fiziert werden. So könnte in der Hochkonjunktur ein nachlassender Auftragseingang auch dadurch bedingt sein, dass die angefragten Unternehmen aufgrund stark ausgelasteter Kapazitäten kaum noch Neugeschäft akquirieren. Über die Absolutwerte von Anfragen und Angeboten könnten somit auch entsprechende Verhältniszahlen (Angebot je Anfrage) einen im Konjunkturzyklus musterhaften Verlauf aufweisen. Die nach eigenen Angaben führende deutsche B2B-Plattform „wer-liefert-was.de“ verzeichnet täglich 75 000 Suchanfragen. Daraus ließe sich ein monatlich verfügbarer, gleichlaufender Seismograph der Geschäftsanbahnung konstruieren, der einen engen Zusammenhang zur ökonomischen Aktivität innerhalb des Unternehmenssektors aufweisen dürfte.

Die im Zuge der Input-Output-Rechnung gemessene Vorleistungsverflechtung liegt bislang nur auf jährlicher Basis und mit erheblicher (mehrjähriger) Publikationsverzögerung vor. Damit sind diese Daten für die originäre Konjunkturanalyse unbrauchbar (sie können freilich unter der Annahme der Strukturkonstanz für abgeleitete Untersuchungen herangezogen werden). Im Bereich der intraindustriellen Lieferverflechtungen können B2B-Plattformen wichtige Anhaltspunkte zur laufenden Aktivität geben. Perspektivisch hat die Auswertung der in den Unternehmenssoftwaresystemen abgelegten Transaktionsdaten das Potenzial für eine Input-Output-Rechnung in Echtzeit.

Zukünftige Potenziale von Big Data für makroökonomische Analysen können sich auch aus der Kombination unterschiedlicher Datenquellen ergeben. Bislang sind diese Potenziale nur schwer abschätzbar, da dazu noch kaum belastbare Studien vorliegen. Eine Hürde für die Kombination verschiedener Datenquellen ist dabei der hohe Aufwand, der betrieben werden muss, um die ohnehin schon komplexen Datensätze konsistent miteinander in Verbindung zu setzen.

Eine Möglichkeit, den Nutzen von Mobilfunkendaten für makroökonomische Analysen zu erhöhen, ist die daraus resultierenden Bewegungsprofile der privaten Nutzer mit ökonomisch relevanten Standortdaten, wie sie beispielsweise von digitalen Kartendiensten zur Verfügung gestellt werden, in Verbindung zu setzen. So könnte das Besucheraufkommen in Einkaufszentren als hochfrequenter Indikator für die Konsumaktivität der privaten Haushalte dienen. Bei sehr detailgetreuen Standortdaten kann so ggf. auch zwischen verschiedenen Verwendungen der Konsumausgaben (z.B. Dienstleistungen und Waren) unterschieden werden. Verstärkt werden könnte die daraus gewonnen Konjunktursignale, wenn die Bewegungsprofile – von datenschutzrechtlichen und technischen Hürden abgesehen – mit Suchanfragen der Nutzer kombiniert würden, wobei eine konkrete Suchanfrage nach einem Konsumgut verbunden mit einem Aufenthalt in einem entsprechenden Geschäft ein stärkeres Signal für eine tatsächlich getätigte Transaktion darstellen könnte. In einem vergleichbaren Ansatz könnten die Bewegungsprofile auch mit Standortdaten für Arbeitsplätze in Verbindung gesetzt werden, wobei aus den Zeitangaben der Profile ggf. Rückschlüsse darauf gezogen werden können, ob es sich in einem Geschäft um einen Kunden oder Angestellten handelt. Die daraus resultierenden Daten könnten sowohl Informationen für den Arbeitsmarkt als auch für die Wertschöpfung der Unternehmen liefern. Ansätze für solche Auswertungen bieten zur Verfügung stehende Daten zur Frequenz von Passanten in deutschen Innenstädten. Erste quantitative Auswertungen für China deuten das grundsätzliche Potenzial der Verknüpfung dieser Datenquellen für makroökonomische Analysen an (Dong et al. 2017). Zudem haben Goolsbee und Syverson (2020) für die Vereinigten Staaten regionale Mobilfunkdaten in Verbindung mit Standortdaten von Geschäften ausgewertet, um zu zeigen, dass der Rückgang der Besucherzahlen nach dem Beginn der Corona-Pandemie im Wesentlichen auf freiwillige Verhaltensanpassungen und nur zu einem kleinen Teil auf die staatlichen Lockdown-Maßnahmen zurückzuführen ist. Allerdings sind noch weitere Studien notwendig, um das tatsächliche Potenzial besser abschätzen zu können.

Auch die Verknüpfung von Bewegungsprofilen von Lkw mit ökonomisch relevanten Standortdaten könnten zusätzliche Informationen liefern, die über den bereits durch das Statistische Bundesamt

regelmäßig veröffentlichen Lkw-Fahrleistungsindex hinausgehen. Dazu könnten zunächst Bewegungsprofile für jede einzelne erfasste Lkw-Fahrt erstellt werden. Wahrscheinliche Start- und Endpunkte ließen sich dabei für jeden Transport automatisiert – ggf. mittels GPS-Daten und geeigneter Heuristiken – abschätzen. Diese Informationen könnten anhand von Standortdaten verschiedenen wirtschaftlichen Aktivitäten zugeordnet und entsprechend ihrer Bedeutung gewichtet werden, wobei die ökonomisch relevanten Standorte beispielsweise anhand von digitalen Kartendiensten oder ggf. Satellitendaten identifiziert werden. Die Zahl der Transporte im Umfeld von Industrieunternehmen könnte dann ein Indikator für den Grad der industriellen Aktivität darstellen, je nach Datenverfügbarkeit auch disaggregiert nach Wirtschaftszweigen. Die Transportaktivität im Umfeld von Groß- und Einzelhandelszentren könnte frühzeitig Signale für die Einzelhandelsumsätze oder den privaten Konsum liefern. Aus grenzüberschreitenden Transporten bzw. Transportaktivitäten an Logistikstandorten für See- oder Flughäfen könnten Frühindikatoren für den internationalen Handel gebildet werden.

Auch Satellitendaten könnten durch eine systematische Verknüpfung mit Standortdaten an Informationsgehalt für makroökonomische Analysen gewinnen. So könnte die über Satellitenbilder erfasste Aktivität auf Parkplätzen für Industriestandorte oder Einkaufszentren bzw. in Häfen als Indikator für die Wertschöpfung, den Konsum oder den internationalen Handel dienen. Erste Analysen für die Vereinigten Staaten deuten darauf hin, dass durch solche Kombinationen Indikatoren für die Umsatzentwicklung von Einzelhändlern erstellt werden können. Um das tatsächliche Potenzial für die Konjunkturanalyse abschätzen zu können, sind jedoch weitere Analysen notwendig, so wie sie jetzt beispielsweise vom Statistischen Bundesamt für Deutschland in Form einer Machbarkeitsstudie vorgenommen werden. Das Potenzial hängt hier freilich auch davon ab, wie rasch andere Kombinationen von Datenquellen systematisch ausgewertet werden können. Sollten beispielsweise umfangreiche Bewegungsprofile von Privatpersonen und Lkw mit Standortdaten verknüpft werden können, so müsste sich erst herausstellen, in welchen Bereichen Satellitendaten einen zusätzlichen Nutzen bringen könnten.

4 Zusammenfassung und Fazit

Unter dem Schlagwort Big Data werden neue und in Abgrenzung zur üblichen Wirtschaftsstatistik unkonventionelle Datenquellen zusammengefasst. Die zugrunde liegenden Daten weisen gewisse Gemeinsamkeiten auf. So sind sie sehr umfangreich, liegen zum Teil nur in unstrukturierter Form vor und sind sehr zeitnah und in hoher Frequenz verfügbar. In vielen Eigenschaften unterschieden sie sich jedoch auch. So decken sie unterschiedliche Bereiche des Wirtschaftsgeschehens ab, erfordern unterschiedlich großen Aufwand, um sie für makroökonomische Zwecke auszuwerten und sind unterschiedlich leicht zugänglich.

Viele der Datenquellen erfassen tatsächlich getätigte Transaktionen bzw. erhobene Preise (elektronischer Zahlungsverkehr, Scannerdaten, Handelsplattformen) oder bilden die laufende wirtschaftliche Aktivität in anderer Form unmittelbar ab (Satellitenbilder, Verkehrsdaten). Die daraus resultierenden Daten haben ein besonders großes Potenzial für die laufende Konjunkturbeobachtung, insbesondere für die Prognose am aktuellen Rand (Nowcast). Sie sind auch für die amtliche Statistik von besonderem Interesse, da sie zum Teil unmittelbar in die Erhebungen zur wirtschaftlichen Aktivität einfließen oder sie auf anderem Wege verbessern oder ergänzen können. Da sie auch in höherer Frequenz und detaillierter vorliegen als die Daten der konventionellen Wirtschaftsstatistik, können sie zudem auch für tiefergehende makroökonomische Analysen eine wertvolle Grundlage darstellen.

Andere Datenquellen bilden Stimmungen und Interessen der Nutzer oder die Relevanz bestimmter Themen ab (soziale Medien, Internetsuchanfragen, Presseartikel). Sie weisen einen weniger engen Bezug zur laufenden wirtschaftlichen Entwicklung auf. Aus ihnen können aber Indikatoren abgeleitet werden, die umfassendere Informationen liefern als bereits vorliegende Indikatoren (z.B. für das Konsumentenvertrauen) oder die über andere konventionelle Datenquellen nicht abgebildet werden können (z.B. wirtschaftspolitische Unsicherheit). Diese Indikatoren können auch für die zukünftige wirtschaftliche Entwicklung nützliche Informationen liefern oder neue Möglichkeiten für vertiefende makroökonomische Analysen eröffnen.

Auch der erforderliche Aufwand, um die Daten für makroökonomische Analysen nutzbar zu machen, unterscheidet sich je nach Datenquelle. Einige der Datenquellen werden bereits regelmäßig ausgewertet, um aus ihnen makroökonomisch relevante Indikatoren zu bilden (wirtschaftspolitische Unsicherheit, Witterungsbedingungen, Lkw-Fahrleistung) oder sie werden bereits so bereitgestellt, dass sie unmittelbar in empirische Analysen eingebunden werden können (Internetsuchanfragen). Diese Datenquellen sind bislang für makroökonomische Analysen besonders häufig verwendet worden und zum Teil auch in laufende Konjunkturanalysen integriert. Sofern die Daten selbst ausgewertet werden, ist damit oft ein deutlich größerer organisatorischer und methodischer Aufwand verbunden. Hier bieten sich Kooperationen mit Experten aus anderen wissenschaftlichen Disziplinen an, um diesen Aufwand zu verringern (z.B. im Bereich der Satellitendaten). Schließlich sind viele der Daten nicht frei oder nur beschränkt verfügbar. Die Nutzung der Daten muss dann mit dem jeweiligen Anbieter vereinbart werden und kann mit hohen Kosten verbunden sein. Der Aufwand und die Verfügbarkeit der Daten kann – insbesondere für die Konjunkturanalyse, für die häufig ein stetiger Datenzugang erforderlich ist – ein Hindernis darstellen.

Die konkreten Potenziale von Big Data für die makroökonomische Analyse hängen auch davon ab, in welchem Umfang und wie zeitnah geeignete Alternativen durch die konventionelle Wirtschaftsstatistik bereitgestellt werden. Folglich sind die Potenziale in solchen Bereichen besonders groß, für die die konventionelle Wirtschaftsstatistik erst mit Verzögerung Indikatoren bereitstellt oder diese nur wenig zuverlässig sind. Für Deutschland, wie für viele andere Volkswirtschaften, gehören beispielsweise die privaten Konsumausgaben oder die Wertschöpfung in den Dienstleistungsbranchen zu diesen Bereichen. Freilich können Big Data aber auch komplementär zu den bereits vorliegenden Indikatoren wertvolle Beiträge liefern, da sie grundsätzlich zeitnäher und in einem höheren Detailgrad verfügbar sind oder das Wirtschaftsgeschehen aufgrund der umfangreichen Datengrundlage unter Umständen umfassender abbilden können.

Die Potenziale und Grenzen von Big Data für die makroökonomische Analyse hängen schließlich auch von den individuellen Voraussetzungen und Anforderungen der jeweiligen Nutzer ab. So sind für die amtliche Statistik vor allem Datenquellen interessant, die tatsächliche Transaktionen abbilden und dadurch helfen, die laufende wirtschaftliche Aktivität zeitnäher und umfassender abzubilden. Solche Datenquellen sind auch für die Konjunkturanalyse von besonderem Interesse. Hier sind zudem Datenquellen, die Aufschluss über die Stimmung von Konsumenten oder Unternehmen liefern, relevant, da sie auch für die zukünftige wirtschaftliche Entwicklung wertvolle Informationen liefern können. Eine mögliche Verstetigung des Datenzugangs, der vielfach mit einem besonders hohen Aufwand verbunden ist, ist für die Konjunkturanalyse allerdings ein zentrales Anwendungskriterium. Dabei dürften beispielsweise Zentralbanken einen großen Wert darauf legen, die Diagnose der laufenden Preisentwicklung zu verbessern, auch wenn Informationen dazu von der amtlichen Statistik schon recht zeitnah zur Verfügung gestellt werden und der Nowcast der Inflation für die allgemeine Konjunkturforschung eine geringere Priorität hat. Für vertiefende makroökonomische Analysen ist eine mögliche Verstetigung des Datenzugangs ein weniger relevantes Anwendungskriterium.

Vor diesem Hintergrund unterscheiden sich auch die in diesem Gutachten betrachteten Datenquellen hinsichtlich ihrer möglichen Anwendungsfelder und ihrer Vor- und Nachteile (siehe auch Tabelle 4):

- Nachrichten und Presseartikel werden bereits regelmäßig verwendet, um Indikatoren zur wirtschaftspolitischen Unsicherheit zu bilden. Darüber hinaus könnten sie auch quantitative Prognosen für das Bruttoinlandsprodukt oder einzelner Komponenten (z.B. Unternehmensinvestitionen) verbessern. Dafür könnten sich auch bislang für makroökonomische Analysen wenig verwendete komplexere Methoden zur Textauswertung als nützlich erweisen. Auch eine intensivere Auswertung von Unternehmensnachrichten könnte makroökonomische Analysen bereichern. Insgesamt liegen dazu aber bislang kaum wissenschaftliche Studien vor.
- Daten aus sozialen Medien sind bislang kaum für makroökonomische Analysen verwendet worden. Sie bieten sich insbesondere an, um Stimmungen der privaten Haushalte zu messen, um daraus Indikatoren für das Konsumklima (Verbrauchervertrauen, Erwartungen oder Unsicherheit) zu bilden. Dafür liefern sie eine umfassende Datengrundlage. Aus heutiger Sicht scheint Twitter als Datenquelle für makroökonomische Analysen etwas besser geeignet als Facebook, da die Daten leichter ausgewertet werden können. Allerdings ist auch hier der erforderliche Aufwand recht hoch, um eine für empirische Analysen geeignete Datengrundlage zu schaffen.
- Internetsuchanfragen können anhand von Google Trends zeitnah und mit eher geringem Aufwand ausgewertet werden. Angesichts des hohen Marktanteils von Google dürften die dort bereitgestellten Informationen getätigte Suchanfragen umfassend abbilden. Sie sind bereits häufig für makroökonomische Studien eingesetzt worden und bieten sich vor allem dazu an, um die privaten Konsumausgaben zu prognostizieren und entsprechende Indikatoren für das Konsumklima zu erstellen. Einzelne Studien deuten darauf hin, dass sie auch für Prognosen des Bruttoinlandsprodukts nützlich sein könnten. Insgesamt sprechen die vorliegenden Analysen dafür, dass der wichtigste Vorteil von Suchanfragen als Datenquelle gegenüber den Indikatoren der konventionellen Wirtschaftsstatistik in ihrer zeitnahen Verfügbarkeit liegt.
- Daten von Online-Handelsplattformen und Scannerdaten sind bislang vor allem für Preisanalysen verwendet worden. Ihr Vorteil liegt darin, dass sie tatsächliche Preise zeitnah und in hoher Detailtreue abbilden. Aus diesem Grund werden sie bereits für die amtliche Statistik genutzt. Scannerdaten enthalten neben Preisen auch andere Informationen, insbesondere die gehandelten Mengen. Sie könnten somit auch nützlich sein, um die Einzelhandelsumsätze zeitnah zu erfassen. Allerdings sind sie nur kostenpflichtig erhältlich. Preisdaten von Online-Handelsplattformen können per Web Scraping auch direkt gesammelt werden. Hier ist jedoch häufig eine gewisse Vorlaufzeit notwendig, um eine ausreichende Datengrundlage für empirische Analysen zu schaffen.
- Daten zum elektronischen Zahlungsverkehr bilden die aktuelle wirtschaftliche Aktivität unmittelbar ab. Deswegen haben sie ein großes Potenzial, um die laufende Entwicklung beim Bruttoinlandsprodukt und insbesondere bei den privaten Konsumausgaben zeitnah zu erfassen. Der Zugang zu diesen Daten muss mit dem jeweiligen Anbieter vereinbart werden. Die vorliegenden Studien zur Prognosegüte im Vergleich zu konventionellen Indikatoren liefern allerdings noch kein klares Bild. Dies könnte zum einen daran liegen, dass ein grundsätzlicher Vorteil dieser Daten, nahezu tagesaktuell verfügbar zu sein, noch nicht vollständig ausgeschöpft wurde und zum anderen, dass die Daten auf Tagesbasis eine sehr hohe Fluktuation aufweisen. Ferner könnten Strukturveränderungen bei der Nutzung unterschiedlicher Zahlungsmittel oder Schwankungen von Marktanteilen auftreten, insbesondere wenn die Daten auf einem einzelnen Anbieter beruhen, wodurch es erschwert würde, auf gesamtwirtschaftliche Entwicklungen zu schließen.

Tabelle 4:
Übersicht Big-Data-Quellen[a]

Daten	Quellen	Mögliche Anwendungsfelder in der makroökonomischen Analyse	Vor- und Nachteile
Nachrichten und Presseartikel	Online-Archive einzelner Zeitungen Kommerzielle Datenbanken (Factiva, Lexis Nexis, Access World News, ProQuest) Einzelne abgeleitete Indikatoren kostenfrei verfügbar (z.B. wirtschaftspolitische Unsicherheit)	• Unsicherheitsindikatoren • Prognose BIP und ggf. Komponenten • Prognose konjunkturelle Wendepunkte • Finanzmarktprognosen	+hohe Datenqualität +direkter Bezug zur aktuellen wirtschaftlichen Entwicklung -aufwendig, insbesondere bei komplexeren Textanalysemethoden -wenig standardisierte Verfahren
Soziale Medien	Facebook (freigeschaltete Profile und Kommentare auf öffentlichen Seiten) Twitter (Stichprobe täglich frei verfügbar; größere Datensätze kostenpflichtig)	• Konsumentenvertrauen • Unsicherheitsindikatoren • Unmittelbare Nutzerreaktion auf Ereignisse oder Wirtschaftspolitik • Finanzmarktprognosen	+umfassende Datenquelle +kann unmittelbar Stimmungen von Nutzern abbilden -große Datensätze umständlich zu generieren und aufwendig auszuwerten -kaum direkter Bezug zu makroökonomischen Themen
Internetsuchanfragen	Google Trends	• Prognose privater Konsum, BIP • Konsumentenvertrauen • Unsicherheitsindikatoren • Identifikation von Schocks	+Zeitreihen können ohne weitere Verarbeitung genutzt werden +flexibel und leicht zu verarbeiten -kaum zusätzlicher Prognose-gehalt ggü. anderen Indikatoren -kein Zugang zu den Rohdaten
Onlinehandels-Plattformen	Web Scraping Billion Prices Project (BPP) PriceStats (kostenpflichtig)	• Preiserhebungen für die amtliche Statistik • Nowcast für Verbraucherpreise • Makro-Analysen: z.B. Gesetz des einheitlichen Preises, Auswirkungen von Handelskonflikten, Preissetzungsverhalten von Unternehmen	+es werden tatsächliche Preise erfasst +sehr disaggregiert zum Teil international vergleichbar verfügbar - keine lange Zeitreihen frei verfügbar - nur Preise, keine gehandelten Mengen
Scannerdaten	Private Anbieter, z.B. GfK oder Nielsen (kostenpflichtig)	• Nowcast Verbraucherpreise, Einzelhandelsumsatz, privater Verbrauch • Preismessung • Eventstudien zu den realwirtschaftlichen Auswirkungen besonderer Ereignisse	+es werden tatsächliche Preise erfasst +sehr disaggregiert verfügbar +Neben Preisen auch gehandelte Mengen verfügbar -Zugang kostenpflichtig -keine langen Zeitreihen
Zahlungsvorgänge	Private Anbieter (VISA, Mastercard, SWIFT) Zahlungssystembetreiber (Girocard, Euro1, TARGET2)	• Nowcast privater Konsum, BIP, ggf. BWS in Dienstleistungsbranchen • Auswirkungen von makroökonomisch bedeutsamen Ereignissen	+umfassend, keine Messfehler +direkter Bezug zu wichtigen makroökonomischen Größen -stark eingeschränkter Zugang -anfällig für Strukturbrüche -bislang wenig Evidenz für hohe Prognosegüte
Fernerkundungsdaten	Mehr als 170 aktive oder geplante Satellitenprogramme Teilweise kostenfreier Zugang (z.B. ESA, CEOS, NASA, USGS,GLCF)	• Einfluss von Witterungsbedingungen auf Konjunktur • Satellitenbilder zum Nowcast der wirtschaftlichen Aktivität (z.B. BIP, Bauinvestitionen, Einzelhandel, Güterhandel, Ernteerträge) • Nachtlichtdaten zur Erfassung der wirtschaftlichen Aktivität	+breite Palette von Informationen +häufig weltweit in regional gleichbleibender Qualität verfügbar -Auswertung der Daten sehr aufwendig -Satellitenbilder in sehr hoher Auflösung in der Regel kostenpflichtig -kaum Erfahrungswerte für Prognosegüte der Daten

Fortsetzung *Tabelle 4*

Daten	Quellen	Mögliche Anwendungsfelder in der makroökonomischen Analyse	Vor- und Nachteile
Schiffpositionsdaten	Kommerzielle Anbieter (z.B. MarineTraffic, VesselTracker)	• Abschätzung von Handelskosten • Nowcast und ggf. zukünftige Entwicklung der weltweiten Handelsaktivität • Nachtlichtdaten zur Erfassung der wirtschaftlichen Aktivität	+umfassende Abdeckung der weltweiten Frachter -Auswertung recht aufwendig -Be- und Entladen von Schiffen an Häfen nicht nachvollziehbar -wenig Erfahrungswerte für makroökonomische Anwendungen
Verkehrsdaten	Rohdaten zum Straßenverkehr (Verkehrszählung, Maut) nicht frei verfügbar Daten einzelner Häfen oder Flughäfen frei verfügbar. Umfassende Datenbanken kostenpflichtig Einzelne aggregierte Indikatoren zeitnah frei verfügbar	• Nowcast Industrieproduktion und internationaler Handel	+auch bereits aggregierte Indikatoren zeitnah verfügbar -kaum Erfahrungswerte für Prognosegüte der verfügbaren Indikatoren
Mobilfunkdaten	Mobilfunkanbieter (kostenpflichtig)	• Amtliche Statistik (Bevölkerungsstatistik, Pendlerstatistik)	+regional sehr disaggregiert verfügbar -kaum Anwendungsfelder für makroökonomische Analysen

[a]Für alle Datenquellen gilt: Die Daten zeitnah erfasst und sind grundsätzlich auch zeitnah verfügbar. Nicht frei verfügbare Daten werden zum Teil von Anbietern für Forschungszwecke oder im Zuge von Kooperationen zugänglich gemacht.

Quelle: Eigene Darstellung.

- Fernerkundungsdaten bzw. Satellitendaten liefern eine Vielfalt von Informationen in vergleichbarer Qualität und in hoher Detailtreue. Für makroökonomische Analysen sind bislang vornehmlich Daten zu Witterungsbedingungen eingesetzt worden, um ihren Einfluss auf die Konjunktur abzuschätzen. Darüber hinaus sind Nachtlichtdaten verwendet worden, um Rückschlüsse auf die wirtschaftliche Aktivität zu ziehen. Bei Nachtlichtdaten bestehen allerdings erhebliche Zweifel, ob sie (insbesondere für fortgeschrittene Volkswirtschaften) geeignet sind, kurzfristige Schwankungen der wirtschaftlichen Aktivität gut abzubilden. Grundsätzlich könnte auch die tägliche Auswertung von Satellitenbildern dazu geeignet sein, um die laufende wirtschaftliche Aktivität zeitnah zu erfassen, zumindest in dem Ausmaß, wie diese optische Spuren hinterlässt. In diesem Zusammenhang könnten sie auch dazu dienen, die Entwicklung in spezifischen Bereichen abzubilden, beispielsweise, wenn sich anhand der optischen Erfassung von Baustellen Rückschlüsse auf die Bauaktivität ziehen lassen. Hierzu liegen bislang aber kaum Erkenntnisse vor. Viele Anbieter stellen Fernerkundungsdaten kostenfrei zur Verfügung, allerdings ist mit ihrer Auswertung ein erheblicher Aufwand verbunden.
- Schiffpositionsdaten bilden die Bewegungen von Transportschiffen weltweit umfassend ab. Sie können deshalb dazu dienen, den Welthandel oder die Handelsaktivität einzelner Länder zu erfassen. Sie sind bei verschiedenen kommerziellen Anbietern zeitnah erhältlich. Hinderlich ist, dass Schiffpositionsdaten keine Informationen über die jeweilige Ladung beinhalten und insbesondere parallel stattfindende Be- und Entladungen an einem Hafen nicht nachvollzogen werden können. Allerdings könnten zumindest auf Basis des Tiefgangs der Schiffe Rückschlüsse auf das Gewicht der Ladung gezogen werden. Auch könnte auf Basis der Positionsdaten prognostiziert werden, welche Häfen die Schiffe zukünftig anlaufen werden. Bislang sind Schiffspositionsdaten noch kaum für

makroökonomische Analysen verwendet worden, sodass belastbare Ergebnisse hinsichtlich ihres Nutzens noch ausstehen.

- Weitere Verkehrsdaten, insbesondere zum Straßenverkehr, werden bereits in recht hohem Detailgrad unmittelbar erfasst. Gerade Daten zum Lkw-Verkehr bieten dabei das Potenzial, Informationen über die laufende wirtschaftliche Entwicklung zu liefern. Rohdaten zum Lkw-Verkehr sind zumindest zeitnah nicht frei zugänglich. Jedoch werden aggregierte Daten auf Grundlage der Mauterhebungen (beispielsweise vom Statistischen Bundesamt) frühzeitig auf monatlicher Basis veröffentlicht. Bislang liegen recht wenige Studien zum Nutzen dieser Daten für die Konjunkturprognose vor. Insgesamt spricht aber einiges dafür, dass sie insbesondere für die Prognose der Industrieproduktion tauglich sind. Detailgetreuere Auswertungen zum Straßenverkehr könnten zudem Aufschlüsse über die grenzüberschreitende Handelsaktivität geben.
- Mobilfunkdaten bilden die Bewegungsprofile von Personen detailliert und zeitnah ab. Aus diesem Grund können sie wertvolle Informationen für die amtliche Statistik liefern, beispielweise bezüglich der regionalen Bevölkerungsstatistik oder Pendlerbewegungen. Sie sind nicht frei verfügbar; der Zugang muss mit den jeweiligen Anbietern vereinbart werden. Mobilfunkdaten sind für makroökonomische Analysen, insbesondere für fortgeschrittene Volkswirtschaften, bislang kaum eingesetzt worden. Ein Grund dafür könnte sein, dass ihr Potenzial dafür aus heutiger Sicht recht begrenzt erscheint. Zukünftig könnten sie an Relevanz gewinnen, wenn die mittels Mobilfunkdaten erfassten Bewegungsprofile der Nutzer mit wirtschaftlich relevanten Standortdaten in Verbindung gesetzt würden.

Alles in allem bieten Big Data große Potenziale, die Konjunkturanalyse – hier insbesondere den Nowcast – und makroökonomische Analysen zu bereichern. Vielfach ist es allerdings noch mit einem erheblichen Aufwand verbunden, die Informationen aus diesen neuen Datenquellen auszuwerten. In diesem Zusammenhang ist die vorliegende Evidenz vielfach noch nicht ausreichend, um ihren konkreten Nutzen beziffern zu können. In dem Ausmaß, wie sich der Zugang zu den Daten verbessert und methodische Weiterentwicklungen voranschreiten, werden diese Potenziale weiter ausgeschöpft werden. Aus heutiger Sicht spricht einiges dafür, dass die Anwendung von Big Data in vielen Anwendungsfeldern vor allem komplementär zu den Daten der konventionellen Wirtschaftsstatistik erfolgen wird.

Darüber hinaus dürften die Potenziale von Big Data für makroökonomische Analysen zukünftig noch weiter zunehmen. So werden bereits nahezu alle Facetten des Wirtschaftsgeschehens digital erfasst. Diese Daten können aber noch nicht vollständig ausgewertet werden, weil sie noch nicht systematisch gesammelt werden oder Datenschutzrichtlinien dem entgegenstehen. Abgesehen davon würden sie es vermutlich bereits jetzt erlauben, die laufende wirtschaftliche Entwicklung sehr genau zu erfassen. Die derzeitigen Initiativen der Statistikämter, anderer wirtschaftspolitischer Institutionen und in der wissenschaftlichen Forschung lassen darauf schließen, dass dieses Potenzial soweit möglich nach und nach gehoben werden wird. Diese detailgetreuen Informationen könnten es zukünftig auch ermöglichen, das Verständnis für makroökonomische Zusammenhänge zu vertiefen und darauf aufbauend auch Konjunkturprognosen zu verbessern und zielgerichteter wirtschaftspolitische Maßnahmen abzuleiten. Zu diesen Zusammenhängen zählen beispielsweise die Lieferverflechtungen zwischen Unternehmen, die bessere Unterscheidung zwischen nachfrage- und angebotsseitigen Einflüssen auf die Konjunktur sowie eine genauere Einschätzung über die gesamtwirtschaftliche Kapazitätsauslastung.

Literatur

Abrahams, A., C. Oram und N. Lozano-Gracia (2018). Deblurring DMSP Nighttime Lights: A New Method Using Gaussian Filters and Frequencies of Illumination. *Remote Sensing of Environment* 210: 242–258.

Addison, D., und B. Stewart (2015). Night Time Lights Revisited: The Use of Night Time Lights Data as a Proxy for Economic Variables. World Bank Policy Research Working Paper No. 7496. World Bank, Washington, D.C.

Ademmer, M., und N. Jannsen (2019). Globale Unsicherheit und deutsche Konjunktur. *Wirtschaftsdienst* 99 (7): 519–520.

Ademmer, E., und T. Stöhr (2019). The Making of a New Cleavage? Evidence from Social Media Debates about Migration. Kiel Working Papers 2140. Kiel Institute for the World Economy, Kiel

Ademmer, M., J. Beckmann und N. Jannsen (2019a). Zu den Auswirkungen des jüngsten Anstiegs der globalen wirtschaftspolitischen Unsicherheit. IfW-Box 2019.7. Institut für Weltwirtschaft, Kiel.

Ademmer, E., A. Leupold und T. Stöhr (2019b). Much Ado about Nothing? The (Non-)politicisation of the European Union in Social Media Debates on Migration. *European Union Politics* 20 (2): 305–327.

Ademmer, M., N. Jannsen, S. Kooths und S. Mösle (2019c). Niedrigwasser bremst Produktion. *Wirtschaftsdienst* 99 (1): 79–80.

Adland, R., H. Jia und S.P. Strandenes (2017). Are AIS-based Trade Volume Estimates Reliable? The Case of Crude Oil Exports. *Maritime Policy and Management* 44 (5): 657–665.

Agarwal, S., und W. Qian (2014). Consumption and Debt Response to Unanticipated Income Shocks: Evidence from a Natural Experiment in Singapore. *American Economic Review* 104 (12): 4205–4230.

Aladangady, A., S. Aron-Dine, W. Dunn, L. Feiveson, P. Lengermann und C. Sahm (2019). From Transactions Data to Economic Statistics: Constructing Real-time, High-frequency, Geographic Measures of Consumer Spending. NBER Working Paper 26253. National Bureau of Economic Research, Cambridge, Mass.

Alessi, L., und C. Detken (2018). Identifying Excessive Credit Growth and Leverage. *Journal of Financial Stability* 35: 215–225.

Alexander, D., und E. Karger (2020). Do Stay-at-Home Orders Cause People to Stay at Home? Effects of Stay-at-Home Orders on Consumer Behavior. Federal Reserve Bank of Chicago, Working Paper WP 2020-12. Chicago, Ill.

Alfaro, L., E. Faia, N. Lamersdorf und F. Saidi (2020). Social Interactions in Pandemics: Fear, Altruism, and Reciprocity. NBER Working Paper 27134. National Bureau of Economic Research, Cambridge, M.A.

Althouse B.M., Y.Y. Ng und D.A.T. Cummings (2011). Prediction of Dengue Incidence Using Search Query Surveillance. *Plos Neglected Tropical Diseases* 5 (5): 1–6.

an de Meulen, P., T.K. Bauer, M. Micheli und T. Schmidt (2011). Ein hedonischer Immobilienpreisindex auf Basis von Internetdaten 2007–2011. RWI Projektberichte. Rheinisch-Westfälisches Institut für Wirtschafts-forschung (RWI), Essen.

Andrade, S.C., J. Bian und T.R. Burch (2013). Analyst Coverage, Information, and Bubbles. *Journal of Financial and Quantitative Analysis* 48 (5): 1573–1605.

Antweiler, W., und M.Z. Frank (2004). Is All that Talk Just Noise? The Information Content of Internet Stock Message Boards. *Journal of Finance* 59 (3): 1259–1294.

ANZ Bank (Australia and New Zealand Banking Group) (2020). ANZ Truckometer. Via Internet (11.2.2020): <https://www.anz.co.nz/about-us/economic-markets-research/truckometer/>.

Aparicio D., und M.I. Bertolotto (2020). Forecasting Inflation with Online Prices. *International Journal of Forecasting*. Via Internet (20.2.2020): <https://www.sciencedirect.com/science/article/abs/pii/S0169207019301530>.

Aprigliano, V., G. Ardizzi und L. Monteforte (2019). Using Payment System Data to Forecast Economic Activity. *International Journal of Central Banking* 15 (4): 55–80.

Arbatli, E., S.J. Davis, A. Ito und N. Miake (2019). Policy Uncertainty in Japan. NBER Working Paper 23411. National Bureau of Economic Research, Cambridge, Mass.

Arnold, S., und S. Kleine (2017). Neue Wege der Geodatennutzung: Perspektiven der Fernerkundung für die Statistik. Geplante Erprobung der Nutzung von Satellitendaten für Flächenstatistik und Ernteerhebungen. WISTA – *Wirtschaft und Statistik* 5: 31–36.

Askitas, N., und K.F. Zimmermann (2009). Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly* 55 (2): 107–120.

Askitas, N., und K.F. Zimmermann (2013). Nowcasting Business Cycles Using Toll Data. *Journal of Forecasting* 32 (4): 299–306.

Atalay, E., P. Phongthiengtham, S. Sotelo und D. Tannenbaum (2017). The Evolving U.S. Occupational Structure. Working Paper 12052017. Washington Center for Equitable Growth. Washington, D.C.

Athey, S., und G.W. Imbens (2019). Machine Learning Methods that Economists Should Know About. *Annual Review of Economics* 11: 685–725.

Azqueta-Gavaldon, A., D. Hirschbühl, L. Onorante und L. Sainz (2020). Economic Policy Uncertainty in the Euro Area: An Unsupervised Machine Learning Approach. ECB Working Paper 2359. European Central Bank, Frankfurt am Main.

Azzimonti, M. (2017). Partisan Conflict and Private Investment. *Journal of Monetary Economics* 93: 114–131.

BAG (Bundesamt für Güterverkehr) (2020). Mautstatistik. Via Internet (11.2.2020): <https://www.bag.bund.de/DE/Navigation/Verkehrsaufgaben/Statistik/Mautstatistik/mautstatistik_node.htm>.

Bai, J., und S. Ng (2008). Forecasting Economic Time Series Using Targeted Predictors. *Journal of Econometrics* 146 (2): 304–317.

Bailey, M., R. Cao, T. Kuchler und J. Stroebel (2018). The Economic Effects of Social Networks: Evidence from the Housing Market. *Journal of Political Economy* 126 (6): 2224–2276.

Baker, S.R., N. Bloom und S.J. Davis (2015). Immigration Fears and Policy Uncertainty. *Voxeu.org* vom 15.12.2015. Via Internet (20.2.2020): <https://voxeu.org/article/immigration-fears-and-policy-uncertainty>.

Baker, S.R., N. Bloom und S.J. Davis (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics* 131 (4): 1593–1636.

Baker, S.R., und A. Fradkin (2017). The Impact of Unemployment Insurance on Job Search: Evidence from Google Search Data. *Review of Economics and Statistics* 99 (5): 756–768.

Baldacci E., D. Buono, G. Kapetanios, S. Krische, M. Marcellino, G.L. Mazzi und F. Papailias (2016). Big Data and Macroeconomic Nowcasting: From Data Access to Modelling. Via Internet (22.11.2019): <https://ec.europa.eu/eurostat/documents/3888793/7753027/KS-TC-16-024-EN-N.pdf/50a4e9ea-d88e-4d59-bf37-2c6200f757b0>.

Banbura, M., D. Giannone und L. Reichlin (2010). Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics* 25 (1): 71–92.

Bangwayo-Skeete, P.F., und R.W. Skeete (2015). Can Google Data Improve the Forecasting Performance of Tourist Arrivals? Mixed-data Sampling Approach. *Tourism Management* 46: 454–464.

Bello-Orgaz, G., J.J. Jung und D. Camacho (2016). Social Big Data: Recent Achievements and New Challenges. *Information Fusion* 28: 45–59.

Beutel, J., S. List und G. von Schweinitz (2019). Does Machine Learning Help Us Predict Banking Crises? *Journal of Financial Stability* 45 (C). DOI: 10.1016/j.jfs.2019.100693.

Bianchi, F., H. Kung und T. Kind (2019). Threats to Central Bank Independence - High-Frequency Identification with Twitter. NBER Working Paper 26308.

Bickenbach, F., E. Bode, M. Söder und P. Nunnenkamp (2016a). Night Lights and Regional GDP. *Review of World Economics* 152 (2): 425–447.

Bickenbach, F., E. Bode, J. Boysen-Hogrefe, S. Fiedler, K.-J. Gern, H. Görg, D. Groll, C. Hornok, N. Jannsen, S. Kooths, C. Krieger-Boden und M. Plödt (2016b). Produktivität in Deutschland – Messbarkeit und Analyse der Entwicklung. Kieler Beiträge zur Wirtschaftspolitik 12. Institut für Weltwirtschaft, Kiel.

Bieg, M. (2019). Nutzung von Scannerdaten in der Preisstatistik – eine Untersuchung anhand von Marktforschungsdaten. *WISTA – Wirtschaft und Statistik* 2: 25–38.

Bilgin, M.H., E. Demir, G. Giray, G. Karabulu und H. Kaya (2019). A Novel Index of Macroeconomic Uncertainty for Turkey Based on Google-Trends. *Economics Letters* 184(C).

BIS (Bank for International Settlements) (2015). Central Bank Use of and Interest in Big Data. Irving Fisher Committee report. Via Internet (22.11.2019): <https://www.bis.org/ifc/publ/ifc-report-bigdata.pdf>.

Blaudow, C., und D. Seeger (2019). Fortschritte beim Einsatz von Web Scraping in der amtlichen Verbraucherpreisstatistik – Ein Werkstattbericht. *WISTA – Wirtschaft und Statistik* 2019 (4): 19–30.

Blazquez, D., und J. Domenech (2018). Big Data Sources and Methods for Social and Economic Analyses. *Technological Forecasting and Social Change* 33: 99–113.

Blondel V., A. Decuyper und G. Krings (2015). A Survey of Results on Mobile Phone Datasets Analysis. *EPJ Data Science* 4 (1): Artikel-Nr. 10.

Bluhm, R., und M. Krause (2018). Top Lights – Bright Cities and their Contribution to Economic Development. CESifo Working Paper 7411. München.

Blumenstock, J.E., G. Cadamuro und R. On (2015). Predicting Poverty and Wealth from Mobile Phone Metadata. *Science* 350 (6264): 1073–1076.

Blustwein, K., M. Buckmann, A. Joseph, M. Kang, S. Kapadia und Ö. Simsek (2020). Credit Growth, the Yield Curve and Financial Crisis Prediction: Evidence from a Machine Learning Approach. Bank of England Working Paper 848. London.

BMWi (Bundesministeriums für Wirtschaft und Energie) (2014). Witterungseffekte im Bausektor. Monatsbericht 12-2014, Berlin.

Böhme, M.H., A. Gröger und T. Stöhr (2020). Searching for a Better Life: Predicting International Migration with Online Search Keywords. *Journal of Development Economics* 142(C).

Bok, B., D. Caratelli, D. Giannone, A.M. Sbordone und A. Tambalotti (2018). Macroeconomic Now-casting and Forecasting with Big Data. *Annual Review of Economics* 10 (1): 615–643.

Bollen, J., H. Mao und X. Zeng (2011). Twitter Mood Predicts the Stock Market. *Journal of Computational Science* 2 (1): 1–8.

Bolukbasi, T., K.-W. Chang, J.Y. Zou, V. Saligrama und A.T. Kala (2016). Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings. *CoRR*, abs/1607.06520.

Bontempi, M.E., R. Golinelli und M. Squadrani (2016). A New Index of Uncertainty Based on Internet Searches: A Friend or Foe of other Indicators? Working Paper 1062. Dipartimento Scienze Economiche, Università di Bologna.

Born, B., M. Ehrmann und M. Fratzscher (2014). Central Bank Communication on Financial Stability. *Economic Journal* 124 (577): 701–734.

BPB (Bundeszentrale für politische Bildung) (2017a). Seefracht – in absoluten Zahlen. Via Internet (11.2.2020): <https://www.bpb.de/nachschlagen/zahlen-und-fakten/globalisierung/52531/seefracht>.

BPB (Bundeszentrale für politische Bildung) (2017b). Luftfracht. Via Internet (11.2.2020) <www.bpb.de/nachschlagen/zahlen-und-fakten/globalisierung/52528/luftfracht>.

Brancaccio, G., M. Kalouptsidi und T. Papageorgiou (2017). Geography, Search Frictions and Endogenous Trade Costs. NBER Working Paper 23581. National Bureau of Economic Research, Cambridge, Mass.

Breiman, L., J. Friedman, R. Oshen und C. Stone (1983). *Classification and Regression Trees*. New York.

Breiman, L. (1996). Random Forests. *Machine Learning* 45: 5–32.

Brunner, K. (2014). Automatisierte Preiserhebung im Internet. *WISTA – Wirtschaft und Statistik* 4: 258–261.

Busch, J., und K. Ferretti-Gallon (2017). What Drives Deforestation and What Stops it? A Meta-analysis. *Review of Environmental Economics and Policy* 11 (1): 3–23.

Business Insider (2018). Inside Facebook's Suicide Algorithm: Here's How the Company Uses Artificial Intelligence to Predict your Mental State from your Posts. Via Internet (11.2.2020): <https://www.businessinsider.de/facebook-is-using-ai-to-try-to-predict-if-youre-suicidal-2018-12?r=US&IR=T>.

BVL (Bundesverband der Deutschen Luftverkehrswirtschaft) (2019). Report Luftfahrt und Wirtschaft 2019. Via Internet (11.2.2020): <https://www.bdl.aero/de/publikation/report-luftfahrt-und-wirtschaft/>.

Caldara, D., und M. Iacoviello (2019). Measuring Political Risk. Working Paper. Board of Governors of the Federal Reserve Board, Washington, D.C.

Caldara, D., M. Iacoviello, P. Molligo, A. Prestipino und A. Raffo (2020). The Economic Effects of Trade Policy Uncertainty. *Journal of Monetary Economics* 109: 38–59.

Cao, X., Y. Hu, X. Zhu, F. Shi, L. Zhuo und J. Chen (2019). A Simple Self-adjusting Model for Correcting the Blooming Effects in DMSP-OLS Nighttime Light Images. *Remote Sensing of Environment* 224: 401–411.

Castelnuovo, E., und T. Duc Tran (2017). Google it Up! A Google Trends-based Uncertainty Index for the United States and Australia. *Economics Letters* 161(C): 149–153.

Cavallo, A. (2013). Online and Official Price Indexes: Measuring Argentina's Inflation. *Journal of Monetary Economics* 60 (2):152–165.

Cavallo, A., B. Neiman und R. Rigobon (2014). Currency Unions, Product Introductions, and the Real Exchange Rate. *Quarterly Journal of Economics* 129 (20): 529–595.

Cavallo, A., und R. Rigobon (2016). The Billion Prices Project: Using Online Prices for Measurement and Research *Journal of Economic Perspectives* 30 (2): 151–78.

Cavallo, A. (2017). Are Online and Offline Prices Similar? Evidence from Large Multi-channel Retailers. *American Economic Review* 107 (1): 283–303.

Cavallo, A. (2018). Scraped Data and Sticky Prices. *The Review of Economics and Statistics* 100(1):105–119.

Cavallo, A., W.E. Diewert, R.C. Feenstra,R. Inklaar und M.P. Timmer (2018). Using Online Prices for Measuring Real Consumption across Countries. American Economic Review Papers and Proceedings 108: 483–87.

Cavallo, A., G. Gopinath, B. Neiman und J. Tang (2019). Tariff Passthrough at the Border and at the Store: Evidence from US Trade Policy. NBER Working Paper 26396. National Bureau of Economic Research, Cambridge, Mass.

Chen, X., und W.D. Nordhaus (2011). Using Luminosity Data as a Proxy for Economic Statistics. *Proceedings of the National Academy of Sciences* 108 (21): 8589–8594.

Chen, X., und W. Nordhaus (2019). VIIRS Night Time Lights in the Estimation of Cross-Sectional and Time-series GDP. *Remote Sensing* 11 (9): 1057–1068.

Choi, H., und H. Varian (2009). Predicting Initial Claims for Unemployment Benefits. Via Internet (22.12.2019): <https://ai.googleblog.com/2009/07/predicting-initial-claims-for.html>.

Choi, H., und H. Varian (2012). Predicting the Present with Google Trends. *The Economic Record* 88 (Special Issue June): 2–9.

Cœuré, B. (2017). Policy analysis with big data. Speech at the conference on "Economic and Financial Regulation in the Era of Big Data" organized by the Bank of Franc. Via Internet (14.11.2019): <https://www.ecb.europa.eu/press/key/date/2017/html/ecb.sp171124.en.html>.

Copernicus (2020). Die Sentinel-Satellitenfamilie. Via Internet (11.2.2020). <https://www.d-copernicus.de/daten/satelliten/>.

Coulombe, P.H., D. Stevanovic und S. Surprenant (2019). How is Machine Learning Useful for Macroeconomic Forecasting?,CIRANO Working Papers 2019s-22, CIRANO.

Cowles, A. (1933). Can Stock Market Forecasters Forecast? *Econometrica* 1 (3): 309–324.

Cox, M., M. Berghausen, S. Linz, C. Fries und J. Völker (2018). Digitale Prozessdaten aus der Lkw-Mauterhebung – Neuer Baustein der amtlichen Konjunkturstatistik. *WISTA – Wirtschaft und Statistik* 6: 11–32.

Daas, P.J.H., und M.J.H. Puts (2014). Social Media Sentiment and Consumer Confidence. ECB Statistics Paper Series 5. European Central Bank, Frankfurt am Main.

D'Amuri, F. (2009). Predicting Unemployment in Short Samples with Internet Job Search Query Data. MPRA Paper 18403, University Library of Munich. München.

D'Amuri, F., und J. Marcucci, (2009). Google it! Forecasting the US unemployment rate with a Google job search index. ISER Working Paper Series 2009-32. Institute for Social and Economic Research, Essex.

D'Amuri, F., und J. Marcucci (2017). The Predictive Power of Google Searches in Forecasting US Unemployment. *International Journal of Forecasting* 33: 801–816.

Da, Z., J. Engelberg und P. Gao (2011a). In Search of Attention. *The Journal of Finance* 66 (5): 1461–1499.

Da, Z., J. Engelberg und P. Gao (2011b). In Search of Fundamentals. Working Paper. University of Notre Dame, St. Joseph County, Indiana, and University of North Carolina at Chapel Hill.

Davis, J.S., D. Liu und X.S. Sheng (2019). Economic Policy Uncertainty in China Since 1949: The View from Mainland Newspapers. Working Paper. O.O.

Dell, M., B.F. Jones und B.A. Olken (2012). Temperature Shocks and Economic Growth: Evidence from the Last Half century. *American Economic Journal: Macroeconomics* 4 (3): 66–95.

Della Penna, N., und H. Huang (2009). Constructing Consumer Sentiment Index for U.S. Using Google Searches. Working Papers 2009-26. University of Alberta, Department of Economics. Alberta, Kanada.

Deloitte (2012). What is the Impact of Mobile Telephony on Economic Growth? GSM Association. Via Internet (14.2.2020): <https://www.gsma.com/publicpolicy/wp-content/uploads/2012/11/gsma-deloitte-impact-mobile-telephony-economic-growth.pdf>.

Demir, B., B. Javorciky, T.-K- Michalskiz und E. Örs (2020). Financial Constraints and Propagation of Shocks in Production Networks. CESifo Working Paper 8607. München.

Deutsche Bundesbank (2019). *Payment Behavior in Germany in 2017: Fourth Study of the Utilisation of Cash and Cashless Payments*. Frankfurt am Main.

Deutsche Bundesbank (2020). Produktion in der Industrie nach Hauptgruppen. Via Internet (4.2.2020): <https://www.bundesbank.de/de/statistiken/konjunktur-und-preise/-/saisonbereinigte-wirtschaftszahlen-804168?contentId=804136#anchor-804136>.

Deville P., C. Linard, S. Martin, M. Gilbert, F.R. Stevens, A.E. Gaughan, V.D. Blondel und A.J. Tatem (2014). Dynamic Population Mapping Using Mobile Phone Data. *Proceedings of the National Academy Sciences* 111 (45): 15888–15893.

Döhrn, R. (2011). Analysen und Berichte – Konjunkturindikatoren. Die Mautstatistik: Keine „Wunderwaffe" für die Konjunkturanalyse. *Wirtschaftsdienst* 91 (12): 863–868.

Döhrn, R., und S. Maatsch (2012). Der RWI/ISL-Containerumschlag-Index – Ein neuer Frühindikator für den Welthandel. *Wirtschaftsdienst* 92 (5): 352–354.

Döhrn, R. (2014). Weshalb Konjunkturprognostiker regelmäßig den Wetterbericht studieren sollten. *Wirtschaftsdienst* 94 (7): 487–491.

Döhrn, R., und P. an de Meulen (2015). Weather, the Forgotten Factor in Business Cycle Analysis. Ruhr Economic Papers 539. Essen.

Döhrn, R. (2019a). Revisionen der Volkswirtschaftlichen Gesamtrechnungen und ihre Auswirkungen auf Prognosen. *AStA Wirtschafts- und Sozialstatistisches Archiv* 13 (2): 99–123.

Döhrn, R. (2019b). Sieben Jahre RWI/ISL-Containerumschlag-Index – ein Erfahrungsbericht. *Wirtschaftsdienst* 99 (3): 224–226.

Donadelli, M. (2015). Google Search-based Metrics, Policy-related Uncertainty and Macroeconomic Conditions. *Applied Economics Letters* 22 (10): 801–807.

Donadelli, M., L. Gerotto, M. Lucchetta und Daniela Arzu (2018). Migration Fear, Uncertainty, and Macroeconomic Dynamics. Working Papers 2018: 29. Department of Economics, University of Venice „Ca' Foscari", Venedig.

Donadelli, M., und L. Gerotto (2019). Non-macro-based Google Searches, Uncertainty, and Real Economic Activity. *Research in International Business and Finance* 48 (C): 111–142.

Donaldson, D., und A. Storeygard (2016). The View from Above: Applications of Satellite Data in Economics. *Journal of Economic Perspectives* 30 (4): 171–198.

Dong, L., S. Chen, Y. Cheng, Z. Wu, C. Li und H. Wu (2017). Measuring Economic Activity in China with Mobile Big Data. EPJ Data Science, December 2017: 6–29. Via Internet (11.2.2020): <https://link.springer.com/article/10.1140/epjds/s13688-017-0125-5>.

Döpke, J., U. Fritsche und C. Pierdzioch (2017). Predicting Recessions with Boosted Regression Trees. *International Journal of Forecasting* 33(4): 745–759.

Duarte, C., P.M.M. Rodrigues und A. Rua (2016). A Mixed Frequency Approach to Forecast Private Consumption with ATM/POS Data. Working Papers 2016-01. Banco de Portugal, Lissabon.

Eagle, N., M. Macy und R. Claxton (2010). Network Diversity and Economic Development. *Science* 328: 1029–1031.

Economic Policy Uncertainty (2020). Global Economic Policy Uncertainty Index. Via Internet (11.2.2020): <https://www.policyuncertainty.com/global_monthly.html>.

Edelman, B. (2012). Using Internet Data for Economic Research. *Journal of Economic Perspectives* 26 (2):189–206.

Eichenbaum, M., N. Jaimovich, S. Rebelo und J. Smith (2014). How Frequent Are Small Price Changes? *American Economic Journal: Macroeconomics* 6 (2): 137–155.

Einav, L., C. Farronato, J. Levin und N. Sundaresan (2013a). Sales Mechanisms in Online Markets: What Happened to Online Auctions? Stanford University. Mimeo.

Einav, L., T. Kuchler, J. Levin und N. Sundaresan (2013b). Learning from Seller Experiments in Online Markets. NBER Working Paper No. 17385. National Bureau of Economic Research, Cambridge, Mass.

Einav, L., D. Knoepfle, J. Levin und N. Sundaresan (2014). Sales Taxes and Internet Commerce. *American Economic Review* 104 (1): 1–26.

Elvidge, C.D., K.E. Baugh, E.A. Kihn, H.W. Kroehl und E.R. Davis (1997). Mapping City Lights with Nighttime Data from the DMSP Operational Linescan System. *Photogrammetric Engineering & Remote Sensing* 63 (6): 727–734.

Elvidge, C.D., M.L. Imhoff, K.E. Baugh, V.R. Hobson, I. Nelson, J. Safran, J.B. Dietz und B.T. Tuttle (2001). Night-time Lights of the World: 1994–1995. *ISPRS Journal of Photogrammetry and Remote Sensing* 56: 81–99.

Elvidge, C.D, K.E. Baugh, M. Zhizhin und F.-C. Hsu (2013). Why VIIRS Data are Superior to DMSP for Mapping Night Time Lights. *Proceedings of the Asia-Pacific Advanced Network* 35 (1): 62–69.

Eurostat (2020). ESSnet Big Data. Via Internet (15.1.2020): <https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/Main_Page>.

EZB (Europäische Zentralbank) (2015). Grocery Prices in the Euro Area: Findings from the Analysis of a disaggregated Price Dataset. ECB Economic Bulletin, Issue 1. Frankfurt am Main.

Felbermayr, G., und J. Gröschl (2014). Naturally Negative: The Growth Effects of Natural Disasters. *Journal of Development Economics* 111: 92–106.

Fenz, G., und M. Schneider (2009). A Leading Indicator of Austrian Exports Based on Truck Mileage. *Monetary Policy & Economy* Q1/09: 44–52.

Ferrara, L., und A. Simoni (2019). When are Google Data Useful to Nowcast GDP? An Approach via Pre-selection and Shrinkage. Working Paper 717. Banque de France, Paris.

Fleetmon (2020). Kostenpflichtige Zusammenstellung von Schiffsverkehrsdaten basierend auf Vessel Position Live Tracking Daten und der Vessel Database. Via Internet (11.2.2020): <https://www.fleetmon.com/services/data-request/>.

Fornaro, P., und H. Luomaranta (2020). Nowcasting Finnish Real Economic Activity: a Machine Learning Approach. *Empirical Economics* 58 (1): 1–17.

Galbraith, J.W., und G. Tkacz (2007). Electronic Transactions as High-Frequency Indicators of Economic Activity. Arbeitspapier 2007-58. Bank of Canada, Ottawa.

Galbraith, J.W., und G. Tkacz (2013). Analyzing Economic Effects of September 11 and Other Extreme Events Using Debit and Payments System Data. *Canadian Public Policy* 39 (1): 119–134.

Galbraith, J.W., und G. Tkacz (2015). Nowcasting GDP with Electronic Payments Data. Arbeitspapier 2015-10. Europäische Zentralbank, Frankfurt am Main.

Gebers, K., und P. Graze (2019). Statistische Datengewinnung durch die Nutzung geografischer Informationen. *WISTA – Wirtschaft und Statistik* 4: 11–18.

Gennaioli, N., R. La Porta, F. Lopez-de-Silanes und A. Shleifer (2013). Human Capital and Regional Development. *Quarterly Journal of Economics* 128 (1): 105–164.

Gentzkow, M., und J.M. Shapiro (2010). What Drives Media Slant? Evidence from U.S. daily newspapers. *Econometrica* 78 (1): 35–72.

Gentzkow, M., B.T. Kelly und M. Taddy (2019). Text as Data. *Journal of Economic Literature* 57 (3): 535–574.

Gholampour, V. und E. van Wincoop (2019). Exchange Rate Disconnect and Private Information: What Can We Learn from Euro-Dollar Tweets? *Journal of International Economics* 119: 111–132.

Giannone, D., L. Reichlin und D. Small (2008). Nowcasting: The Real-time Informational Content of Macroeconomic DEdsata. *Journal of Monetary Economics* 55 (4): 665–676.

Gibson, J., S. Olivia und G. Boe-Gibson (2019). A Test of DMPS and VIIRS Night Lights Data for Estimating GDP and Spatial Inequality for Rural and Urban Areas. Working Paper in Economics 11/19. University of Waikato, Hamilton, Neuseeland.

Gil, M., J.J. Pérez, A.J. Sánchez und A. Urtasun (2018). Nowcasting Private Consumption: Traditional Indicators, Uncertainty Measures, Credit Cards and Some Internet Data. Working Papers 1842. Banco de España, Madrid.

Ginsberg, J., M.H. Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski und L. Brilliant (2009). Detecting Influenza Epidemics Using Search Engine Query Data. *Nature* 457: 1012–1014.

Götz, T.B., und T.A. Knetsch (2019). Google Data in Bridge Equation Models for German GDP. *International Journal of Forecasting* 35 (1): 45–66.

Goldberg, Y., und J. Orwant (2013). A Dataset of Syntactic-Ngrams over Time from a Very Large Corpus of English Books. In M. Diab, T. Baldwin und M. Baroni (Hrsg.), *Second Joint Conference on Lexical and Computational Semantics (*SEM)*. Volume 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity. Association for Computational Linguistics, Stroudsburg, Pa.

Google Trends (2020). Google-Suchanfragen für die Begriffe „Abgasskandal“ und „Dieselskandal“. Via Internet (17.2.2020): <https://trends.google.com/trends>.

Goolsbee, A., und C. Syverson (2020). Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020. Becker Friedmann Institute Working Paper 2020-80, Chicago, Ill.

Gorodnichenko, Y., T. Pham und O. Talavera (2018). Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #USElection. NBER Working Paper 24631. National Bureau of Economic Research, Cambridge, Mass.

Greenstein, S., Y. Gu und F. Zhu (2016). Ideological Segregation among Online Collaborators: Evidence from Wikipedians. NBER Working Paper 22744. National Bureau of Economic Research, Cambridge, Mass.

Groseclose, T., und J. Milyo (2005). A Measure of Media Bias. *Quarterly Journal of Economics* (4): 1191–1237.

Gruber, H., und P. Koutroumpis (2011). Mobile Telecommunications and the Impact on Economic Development. *Economic Policy* 26: 387–426.

Hansen, S., M. McMahon und A. Prat (2014). Transparency and Deliberation within the FOMC: A Computational Linguistics Approach. Discussion Papers CEPDP 1276. Centre for Economic Performance, London.

Hassan, T.A., S. Hollander, L. van Lent und A. Tahoun (2019). Firm-Level Political Risk: Measurement and Effects. *The Quarterly Journal of Economics* 134 (4): 2135–2202.

Hassan, T.A., S. Hollander, L. van Lent und A. Tahoun (2020). The Global Impact of Brexit Uncertainty. NBER Working Paper 26609. National Bureau of Economic Research, Cambridge, Mass.

Hauber, P. (2018). Zur Kurzfristprognose mit Faktormodellen und Prognoseanpassungen. IfW-Box 2018.5. Institut für Weltwirtschaft, Kiel.

Hausmann, R., J. Hinz und M.A. Yıldırım (2018). Measuring Venezuelan Emigration with Twitter. Kiel Working Paper 2106.

He, C., Q. Ma, T. Li, Y. Yang und Z. Liu (2012). Spatiotemporal Dynamics of Electric Power Consumption in Chinese Mainland from 1995 to 2008 Modeled Using DMSP/OLS Stable Nighttime Lights Data. *Journal of Geographical Sciences* 22 (1): 125–136.

Heiland, I., A. Moxnes, K.H. Ullveit-Moe und Y. Zi (2019). Trade from Space: Shipping Networks and The Global Implications of Local Shocks. Working Paper. Via Internet (15.1.2020): <http://inga-heiland.de/pdfs/HMUZ2019.pdf> .

Henderson, J.V., A. Storeygard und D.N. Weil (2012). Measuring Economic Growth from Outer Space. *American Economic Review* 102 (2): 994–1028.

Hoberg, G., und G.M. Phillips (2015). Text-based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy* 124 (5): 1423–1465.

Holopainen, M., und P. Sarlin (2017). Toward Robust Early-Warning Models: A Horse Race, Ensembles and Model Uncertainty. *Quantitative Finance* 17 (12): 1–31.

Hsiang, S.M. (2010). Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America. Proceedings of the National Academy of Sciences 107 (35): 15367–15372.

Hsu, F.-C., K.E. Baugh, T. Ghosh, M. Zhizhin und C.D. Elvidge (2015). DMSP-OLS Radiance Calibrated Nighttime Lights Time Series with Intercalibration. *Remote Sensing* 7 (2): 1855–1876.

Hu, Y., X. Cao und J. Chen (2019). Quantitative Evaluation for the Blooming Effect of Nighttime Light Data in China. 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, 28.7.–2.8.2019. Via Internet (11.2.2020): <https://doi.org/10.1109/IGARSS.2019.8899107>.

Hummel, M., A. Vosseler, E. Weber und R. Weigand (2015). Frost und Schnee: Wie das Wetter den Arbeitsmarkt beeinflusst. IAB-Kurzbericht 2/2015. Nürnberg.

Husted, L., J.H. Rogers und B. Sun (2017). Monetary Policy Uncertainty. Working Paper. Board of Governors of the Federal Reserve Board, Washington, D.C.

IMF (International Monetary Fund) (2017). Big Data: Potential, Challenges and Statistical Implications. IMF Staff Discussion Note, September, SDN/17/06. Via Internet (22.10.2019): <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2017/09/13/Big-Data-Potential-Challenges-and-Statistical-Implications-45106>.

Indaco (2019). From Twitter to GDP: Estimating Economic Activity from Social Media. Manuscript. City University of New York.

Jegadeesh, N., und D. Wu (2013). Word Power: A New Approach for Content Analysis. *Journal of Financial Economics* 110 (3): 712–729.

Jun, S.-P., H.S. Yoo und S. Choi (2018). Ten Years of Research Change Using google Trends. *Technological Forecasting and Social Change* 130 (May): 69–87.

Jung, J.-K., M. Patnam und A. Ter-Martirosyan (2018). An Algorithmic Crystal Ball: Forecasts-based on Machine Learning. IMF Working Paper 18/230, Washington D.C.

Kalogeropoulos, A., S. Negredo und I. Picone (2017). Who Shares and Comments on News? A Cross-national Comparative Analysis of Online and Social MediaPparticipation. *Social Media+ Society* 3 (4): 1–12.

Karabulut, Y. (2013). Can Facebook Predict Stock Market Activity? AFA 2013 San Diego Meetings Paper.

Katona, Z., M. Painter, P.N. Patatoukas und J. Zeng (2018). On the Capital Market Consequences of Alternative Data: Evidence from Outer Space. In 9th Miami Behavioral Finance Conference.

Kelly, B.T., A. Manela und A. Moreira (2019). Text Selection. NBER Working Paper 26517. National Bureau of Economic Research, Cambridge, Mass.

Kelly, B.T., D. Papanikolaou, A. Seru und M. Taddy (2018). Measuring Technological Innovation over the Long Run. NBER Working Paper 25266. National Bureau of Economic Research, Cambridge, Mass.

Kholodilin, K.A., M. Podstawski und B. Siliverstovs (2010). Do Google Searches Help in Nowcasting Private Consumption? A Real-Time Evidence for the US. Discussion Papers of DIW Berlin 997. Deutsches Institut für Wirtschaftsforschung, Berlin.

Kilian, L. (2009). Not All Oil Price Shocks are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review* 99 (3): 1053–1069.

Kramer, A.D.I. (2010). An Unobtrusive Behavioral Model of "Gross National Happiness." Via Internet (12.2.2020): <https://research.fb.com/wp-content/uploads/2016/11/an-unobtrusive-behavioral-model-of-gross-national-happiness.pdf>.

Ladiray, D., G.-L. Mazzi, J. Palate und T. Proietti (2018). Seasonal Adjustment of Daily and Weekly Data. In D. Ladiray und G.-L. Mazzi (Hrsg.), *Handbook of Seasonal Adjustment.* Eurostat.

Larcinese, V., R. Puglisi und J. M. Snyder (2011). Partisan Bias in Economic News: Evidence on the Agenda-setting Behavior of U.S. Newspapers. *Journal of Public Economics* 95 (9): 1178–1189.

Laurila, J.K., D. Gatica-Perez, I. Aad, O. Bornet, T.M.T. Do, O. Dousse, J. Eberle und M. Miettinen (2012). The Mobile Data Challenge: Big Data for Mobile Computing Research. In *Proceedings of the Pervasive and Mobile Computing*: 26–30. Trento, Italy.

Laver, M., K. Benoit und J. Garry (2003). Extracting Policy Positions from Political Texts Using Words as Data. *American Political Science Review* 97: 311–331.

Lazer, D., R. Kennedy, G. King und A. Vespignani (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science* 343 (6176): 1203–1205.

Lehrer, S.F., T. Xie und T. Zeng (2019). Does High Frequency Social Media Data Improve Forecasts of Low Frequency Consumer Confidence Measures? NBER Working Papers 26505. National Bureau of Economic Research.

Leßmann, C., A. Seidel und A. Steinkraus (2015). Satellitendaten zur Schätzung von Regionaleinkommen – Das Beispiel Deutschland. *ifo Dresden berichtet* 22 (6): 35–42.

Liu, X., A. de Sherbinin und Y. Zhan (2019). Mapping Urban Extent at Large Spatial Scales Using Machine Learning Methods with VIIRS Nighttime Light and MODIS Daytime NDVI Data. *Remote Sensing* 11 (10): 1247.

Llorente A., M. Garcia-Herranz, M. Cebrian, E. Moro (2015). Social Media Fingerprints of Unemployment. *PLoS ONE* 10 (5): e0128692.

Lörmann, J., und B. Maas (2019). Nowcasting US GDP with Artificial Neural Networks. MPRA Paper 95459.

Loughran, T., und B. McDonald (2011). When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance* 66 (1): 35–65.

Lucca, D.O., und F. Trebbi (2011). Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements. University of British Columbia, Mimeo.

Manela, A., und A. Moreira (2015). News Implied Volatility and Disaster Concerns. *Journal of Financial Economics* 123 (1): 137–162.

Marx, B., T.M. Stoker und T. Suri (2019). There is no Free House: Ethnic Patronage in a Kenyan Slum. *American Economic Journal: Applied Economics* 11 (4): 36–70.

Michalopoulos, S., und E. Papaioannou (2018). Spatial Patterns of Development: A Meso Approach. *Annual Review of Economics* 10 (1): 383–410.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado und J. Dean (2013). Distributed Representations of Words and Phrases and Their Compositionality. In C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani und K.Q. Weinberger (Hrsg.), *Proceedings of the 26th International Conference on Neural Information Processing Systems*. Red Hook: Curran Associates: 3111–3119.

Minges, M. (2016). Exploring the Relationship Between Broadband and Economic Growth. World Development Report background papers Washington, D.C.: World Bank Group.

Müller, H. (2020). Donald Trump als wirtschaftspolitischer Unsicherheitsfaktor. *ifo Schnelldienst* 73 (1): 26–29.

Müller, H., G. von Nordheim, K. Boczek, L. Koppers und J. Rahnenführer (2018). Der Wert der Worte – Wie digitale Methoden helfen, Kommunikations- und Wirtschaftswissenschaft zu verknüpfen. *Publizistik* 63 (4): 557–582.

Mullainathan, S. und J. Spiess (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives* 31(2): 87–106.

NASA Earth Observatory (2020). Night Light Maps Open Up New Applications. Via Internet (11.2.2020). <https://earthobservatory.nasa.gov/images/90008/night-light-maps-open-up-new-applications>.

Niesert, R., J. Oorschot, C. Veldhuisen, K. Brons und R.-J. Lange (2019). Can Google Search Data Help Predict Macroeconomic Series? Tinbergen Institute Discussion Paper 2019-021/III.

Nordhaus, W., und X. Chen (2015). A Sharper Image? Estimates of the Precision of Nighttime Lights as a Proxy for Economic Statistics. *Journal of Economic Geography* 15 (1): 217–246.

O'Connor, B., R. Balasubramanyan, B.R. Routledge und N.A. Smith (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. Carnegie Mellon University, Research Showcase. Via Internet (11.2.2020): <https://www.researchgate.net/publication/221297841_From_Tweets_to_Polls_Linking_Text_Sentiment_to_Public_Opinion_Time_Series>.

Ollech, D. (2018). Seasonal Adjustment of Daily Time Series. Diskussionspaper 41/2018. Deutsche Bundesbank, Frankfurt am Main.

Oostrom, L., A.N. Walker, B. Staats, M. Slootbeek-Van Laar, S. Ortega Azurdu und B. Rooijakkers (2016). Measuring the Internet Economy in The Netherlands: A Big Data Analysis. CBS Discussion Paper 2016|14.

Pandya, S.S., und R. Venkatesan (2016). French Roast: Consumer Response to International Conflict – Evidence from Supermarket Scanner Data. *The Review of Economics and Statistics* 98 (1):42–56.

Pennington, J., R. Socher und C. Manning (2014). GloVe: Global Vectors for Word Representation. In A. Moschitti, B. Pang und W. Daelemans (Hrsg.), *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Stroudsburg: Association for Computational Linguistics: 1532–1543.

PEW Research Center (2018). Social Media Use in 2018. Via Internet (11.2.2020): <https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018>.

Porter, M.F. (1980). An Algorithm for Suffix Stripping. *Program* 14 (3): 130–137.

Preis, T., H.S. Moat und H.E. Stanley (2013). Quantifying Trading Behavior in Financial Markets Using Google Trends. *Nature*, Scientific Reports 3, Article Number 1684.

ProPublica (2016). Facebook Doesn't Tell Users Everything it Really Knows about them. Via Internet (11.2.2020): <https://www.propublica.org/article/facebook-doesnt-tell-users-everything-it-really-knows-about-them>.

Rengers, M. (2018). Internetgestützte Erfassung offener Stellen. *WISTA – Wirtschaft und Statistik* 5: 11–33.

Ricciato, F., P. Widhalm, M. Craglia und F. Pantisano (2015). Estimating Population Density Distribution from Network-based Mobile Phone Data. Technical Report, Joint Research Centre.

Rosenski, N., und C. Schartner (2018). Remote Sensing Data for Better Statistics. Paper prepared for the 16th Conference of the International Association of Official Statisticians (IAOS), OECD Headquarters, Paris, 19-21 September 2018. Via Internet (12.2.2020): <https://www.makswell.eu/event_attachments/19_21-9-2018/iaos-oecd2018_rosenski_schartner.pdf>.

Rowland, E., A. Eidukas, S. Campbell, L. Nolan, D. Elliot, S. Del-Chowdhury (2019). Faster Indicators of UK Economic Activity: Road Traffic Data for England. Data Science Campus, Office for National Statistics. Via Internet (11.2.2020): <https://datasciencecampus.ons.gov.uk/projects/faster-indicators-of-uk-economic-activity-road-traffic-data-for-england/>.

Sachverständigenrat (2019). Den Strukturwandel meistern. Jahresgutachten 2019-20, Wiesbaden.

Schmidt, T., und S. Vosen (2012). A Monthly Consumption Indicator for Germany Based on Internet Search Query Data. *Applied Economics Letters* 19 (7): 683–687.

Schnorr-Bäcker, S. (2018). Möglichkeiten und Grenzen der Nutzung von Big Data in der amtlichen Statistik. In A. Blätte, J. Behnke, K.-U. Schnapp und C. Wagemann (Hrsg.), *Computational Social Science - Die Analyse von Big Data*. Nomos: Baden-Baden: 315–355.

Siganos, A., E. Vagenas-Nanos und P. Verwijmeren (2014). Facebook's Daily Sentiment and International Stock Markets. *Journal of Economic Behavior & Organization* 107 (B): 730–743.

Silver, M., und S. Heravi. (2005). A Failure in the Measurement of Inflation: Results from a Hedonic and Matched Experiment Using Scanner Data. *Journal of Business & Economic Statistics* 23 (3): 269–281.

Slapin, J.B., und S.-O. Proksch (2008). A Scaling Model for Estimating Time-Series Party Positions from Texts. *American Journal of Political Science* 52 (3): 705–722.

Small, C., C.D. Elvidge, D. Balk und M. Montgomery (2011). Spatial Scaling of Stable Night Lights. *Remote Sensing of Environment* 115 (2): 269–280.

Smith-Clarke, C., A. Mashhadi und L. Capra (2014). Poverty on the Cheap: Estimating Poverty Maps Using Aggregated Mobile Communication Networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 511–520. ACM.

statista (2019). Number of Monthly Active Facebook Users Worldwide as of 4th Quarter 2019. Via Internet (24.10.2019): <www.statista.com/statistics/264810/number-ofmonthly-active-facebook-users-worldwide>.

Statistisches Bundesamt (2019). Digitale Agenda des Statistischen Bundesamts. Via Internet (8.1.2019): <https://www.destatis.de/DE/Service/OpenData/Publikationen/digitale-agenda.pdf?__blob=publicationFile>.

Statistisches Bundesamt (2020a). EXDAT – Experimentelle Daten. Via Internet (13.2.2020): <https://www.destatis.de/DE/Service/EXDAT/_inhalt.html>.

Statistisches Bundesamt (2020b). Verbraucherpreisindizes für Deutschland. Fachserie17, Reihe7 (lfd. Jgg.).

Statistisches Bundesamt (2020c). Verkehr. Fachserie 8, Reihe 1.1 (lfd. Jgg.).

Statistisches Bundesamt (2020d). Lkw-Maut-Fahrleistungsindex. Via Internet (4.2.2020). <https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=42191-0001#abreadcrumb>.

Statistisches Bundesamt (2020e). Ausstattung privater Haushalte mit Informations- und Kommunikationstechnik - Deutschland. Via Internet (11.2.2020): <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Ausstattung-Gebrauchsgueter/Tabellen/liste-infotechnik-d.html>.

Stier, S., L. Posch, A. Bleier und M. Strohmaier (2017). When Populists Become Popular: Comparing Facebook Use by the Right-wing Movement Pegida and German Political Parties. *Information, Communication & Society* 20 (9): 1365–1388.

Stock, J.H., und F. Trebbi (2003). Retrospectives Who Invented Instrumental Variable Regression? *Journal of Economic Perspectives* 17 (3): 177–194.

Stock, J.H., und M. Watson (2002). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics* 20(2): 147–162,

Suhoy, T. (2009). Query Indices and a 2008 Downturn: Israeli Data. Discussion Paper 2009.06. Bank of Israel, Jerusalem.

Sundsøy, P., J. Bjelland, B.A. Reme, E. Jahani, E. Wetter und L. Bengtsson (2016). Estimating Individual Employment Status Using Mobile Phone Network Data. arXiv preprint: 1612.03870.

Susnjak, T., und C. Schumacher (2018). Towards Real-Time GDP Prediction. Technical Report Nr. 1. Via Internet (11.2.2020): <http://www.gdplive.net/TechnicalDetails>.

Tanaka, K., T. Kinkyo und S. Hamori (2016). Random Forests-Based Early Warning System for Bank Failures. *Economics Letters* 148: 118–121.

Taylor, J.D. (2009). Analysis of Daily Sales Data during the Financial Panic of 2008. Arbeitspapier. Via Internet (20.2.2020): <https://web.stanford.edu/~johntayl/Onlinepaperscombinedbyyear/2009/Analysis_of_Daily_Sales_Data_During_the_Financial_Panic_of_2008.pdf>.

Tetlock, P. (2007). Giving Content to Investor Sentiment:The Role of Media in the Stock Market. *Journal of Finance 62* (3): 1139–1168.

Thompson Jr., H.G., und C. Garbacz (2011). Economic Impacts of Mobile Versus Fixed Broadband. *Telecommunications Policy* 35 (11): 999–1009.

Thorsrud, L.A. (2020). Words are the New Numbers: A Newsy Coincident Index of the Business Cycle. *Journal of Business & Economic Statistics* 38: 393–409.

Tjaden, J.D., C. Schwemmer und M. Khadjavi (2018). Ride with Me - Ethnic Discrimination in Social Markets. *European Sociological Review* 34 (4): 418–432.

Toole, J., Y.-R. Lin, E. Muehlegger, D. Shoag, M. Gonzalez und D. Lazer (2015). Tracking Employment Shocks Using Mobile Phone Data. *Journal of the Royal Society Interface* 107 (12): 150–185.

Tripathy, B.R., H. Sajjad, C.D. Elvidge, Y. Ting, P.C. Pandey, M. Rani und P. Kumar (2018). Modeling of Electric Demand for Sustainable Energy and Management in India Using Spatio-Temporal DMSP-OLS Night-Time Data. Environmental Management 61 (4): 615–623.

UNCTAD (2019). Review of Maritime Transport 2019. United Nations publication. Sales no. E.19.II.D.20.

Varian, H.R. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives* 28 (2): 3–28.

Verbaan, R., W. Bolt und C. van der Cruijsen (2017). Using Debit Card Payments Data for Nowcasting Dutch Household Consumption. DNB Working Paper 571. De Nederlandsche Bank, Amsterdam.

Vereinte Nationen (2015). Revision and Further Development of the Classification of Big Data. Discussion Paper for the 2015 Global Conference on Big Data for Official Statistics. Global Working Group on Big Data for Official Statistics Task Team on Cross-Cutting Issues.

Vereinte Nationen (2020). Big Data. Via Internet (5.1.2020): <https://unstats.un.org/bigdata/>.

Visa (2019). Visa's UK Consumer Spending Index – April 2019. Via Internet (20.2.2020): <http://www.mynewsdesk.com/uk/visa/documents/visas-uk-consumer-spending-index-april-2019-88072>.

Vlastakis, N. und R.N. Markellos (2012). Information Demand and Stock Market Volatility. *Journal of Banking & Finance* 36 (6): 1808–1821.

Vosen, S., und T. Schmidt (2011). Forecasting Private Consumption: Survey-Based vs. Google Trends. *Journal of Forecasting* 30: 565–578.

Wall Street Journal (2019). U.S. Social Sentiment Index. Via Internet (15.2.2019): <http://graphics.wsj.com/twitter-sentiment/#methodology-anchor>.

Wiengarten, L., und M. Zwick (2017). Neue Digitale Daten in der amtlichen Statistik. *WISTA – Wirtschaft und Statistik* 5: 19–30.

Wisniewski, T.P., und B.J. Lambe (2013). The Role of Media in the Credit Crunch: The Case of the Banking Sector. *Journal of Economic Behavior and Organization* 85 (1): 163–175.

Wolmershäuser, T. (2016). Vorhersage der Revisionen der Vorratsveränderungen mit Hilfe der ifo Lagerbeurteilung. *ifo Schnelldienst* 69 (7): 26–32.

Zou, H., und T. Hastie (2005). Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67(2): 301–320.

Institut für Weltwirtschaft
Kiel Institute for the World Economy
ISBN 978-3-89456-347-9
ISSN 2567-6474
www.ifw-kiel.de