

WAHR

SCHNEIN

LICHT

WAHR SCHEIN LICH

PROGNOSEN AUF DEM PRÜFSTAND

FABIAN KRÜGER

Prognosen sollen Aussagen über zukünftige Ereignisse und deren Wahrscheinlichkeit ermöglichen – und das natürlich so treffsicher wie möglich. Doch nicht nur im Kontext ökonomischer Kennzahlen wie Wirtschaftsleistung und Inflation kommt es immer wieder vor, dass Vorhersagen mehr oder weniger deutlich vom tatsächlichen Ergebnis abweichen. Das führt zu einer grundlegenden Frage: Wie kann Prognoseunsicherheit gemessen und kommuniziert werden?

A

Am 7. November 2016, dem Vortag der 58. Präsidentschaftswahl in den USA, waren sich die Beobachter weitgehend einig: Hillary Clinton würde die Wahl gewinnen und ins Weiße Haus einziehen. Die renommierte „New York Times“ bezifferte die Wahrscheinlichkeit für einen Sieg Clintons auf 85 Prozent; die „Huffington Post“ und das „Princeton Election Consortium“ waren sich mit 98 Prozent beziehungsweise mehr als 99 Prozent noch sicherer. Einzig der Journalist und Prognostiker Nate Silver („Die Berechnung der Zukunft“) äußerte sich vorsichtiger: Er schätzte die Wahrscheinlichkeit für einen Clinton-Sieg auf „nur“ 71 Prozent und schrieb noch am 4. November, dass Clintons vermeintlicher Vorsprung auf ihren Rivalen Donald J. Trump möglicherweise das Produkt fehlerhafter Wahlumfragen sein könnte.

Das Wahlergebnis vom 8. November 2016 ist hinlänglich bekannt und wirft – mit der Ausnahme Silvers – kein

sonderlich positives Licht auf die Prognosen. Wie kann es sein, dass verschiedene Teams renommierter Experten zu so unterschiedlichen Wahrscheinlichkeitseinschätzungen kommen? Im konkreten Fall fußte Silvers konservativere Prognose darauf, dass er sich der Fehlbarkeit von Wahlumfragen sehr bewusst war: Er hielt es für möglich, dass die Umfragen deutlich vom tatsächlichen Ergebnis abweichen – und das in mehreren US-Bundesstaaten gleichzeitig. Darüber hinaus führt das Beispiel der US-Wahl 2016, die die generelle Unsicherheit von Prognosen der breiten Allgemeinheit ins Bewusstsein gebracht hat, zu einer grundlegenden Frage: Wie kann Prognoseunsicherheit überhaupt gemessen und kommuniziert werden? Mit dieser Frage beschäftige ich mich seit einigen Jahren – als Wirtschaftsstatistiker allerdings nicht im Bereich der Wahlprognosen, sondern vor allem im Kontext ökonomischer Kennzahlen wie Wirtschaftsleistung und Inflation.

Prognoseunsicherheit stößt in den Wirtschaftswissenschaften – wie auch in vielen anderen Disziplinen – erst seit einigen Jahren auf breiteres Interesse. Zuvor waren sogenannte Punktprognosen die Regel, die in der öffentlichen Berichterstattung auch heute noch dominieren. Unter einer Punktprognose versteht man grob einen „gewöhnlich zu erwartenden“ Wert. So schrieb tagesschau.de am 27. Juni 2017, der Internationale Währungsfonds erwarte für „2017 (...) ein US-Wirtschaftswachstum von 2,1 Prozent“. Doch was bedeutet diese Zahl? Anders als

„Die Herausforderung von Prognosen besteht darin, zukünftige Trends und Tendenzen zu erkennen.“

das oben betrachtete Beispiel, in dem es nur zwei mögliche Ergebnisse gibt (Hillary Clinton gewinnt die Wahl oder sie gewinnt die Wahl nicht) ist die Wachstumsrate der US-Wirtschaft eine Zahl, die im Prinzip sehr viele verschiedene Werte annehmen kann, wie etwa -1,7 Prozent, 0,1 Prozent oder 2,4 Prozent. Gehen die Experten des Internationalen Währungsfonds davon aus, dass aus allen möglichen Werten auf jeden Fall eine Wachstumsrate von 2,1 Prozent eintreten wird? Sicherlich nicht – erfahrene Statistiker, ob aus Politik, Privatwirtschaft oder Wissenschaft, wissen, dass fast alle Prognosen mit einer gehörigen Portion Unsicherheit behaftet sind. Doch diese Unsicherheit systematisch zu messen und zu kommunizieren, ist eine große Herausforderung – sowohl aus praktischer als auch aus wissenschaftlicher Sicht.

Die Kunst, das eigene Nicht-Wissen in eine Zahl zu fassen

Das mathematische Vokabular für die Messung von Prognoseunsicherheit stammt aus der Wahrscheinlichkeitsrechnung. Eine Wahrscheinlichkeit drückt aus, was zum Zeitpunkt der Prognose als wahr erscheint. Dieser Schein ist nicht als trügerischer oder falscher Schein zu verstehen. Vielmehr bezeichnet er eine nach bestem Wissen und Gewissen erstellte und in Zahlen ausgedrückte Einschätzung. Eine Wahrscheinlichkeit kann sehr diffus sein: Die Wahrscheinlichkeit, dass eine faire Münze „Kopf“ anzeigt, liegt bei 50 Prozent. Sie kann auch sehr konkret sein: Die Wahrscheinlichkeit, dass bei zehnmalem Münzwurf mindestens einmal „Kopf“ erscheint, liegt bei mehr als 99,9 Prozent. Bei der Messung von Prognoseunsicherheit besteht die Kunst darin, das eigene Nicht-Wissen in eine Zahl (oder je nach mathematischer Spielart in mehrere Zahlen) zu fassen. Dies ist bei ökonomischen oder auch politikwissenschaftlichen Prognosen deutlich schwieriger als im Beispiel des Münzwurfs: Die Wahrscheinlichkeiten für verschiedene Ausgänge sind nicht per Konstruktion vorgegeben, sondern werden geschätzt, wofür eine überwältigende Vielzahl von statistischen Methoden und Datenquellen zur Verfügung steht. Diesen spannenden Vorgang und seine Besonderheiten möchte ich im Folgenden darstellen.

Zunächst ein Beispiel: Abbildung 1 zeigt die Wachstumsrate der europäischen Wirtschaftsleistung von 1996 bis Anfang 2017. Gemessen wird die Wirtschaftsleistung mit dem Bruttoinlandsprodukt, kurz BIP; mit der „europäischen Wirtschaft“ sind hier die 19 Länder gemeint, die den Euro als Währung übernommen haben. Seit der großen Wirtschafts- und Finanzkrise, deren Talsohle in Europa im ersten Quartal 2009 erreicht war, ist die europäische Wirtschaft um durchschnittliche 0,6 Prozent pro Quartal gewachsen; zuletzt erreichte sie im ersten Quartal 2017 eine Wachstumsrate von 1,9 Prozent. Doch wie wird sie sich in Zukunft weiterentwickeln? Anders als in der Physik gibt es keine Naturgesetze, die zur Klärung dieser Frage



PROF. DR. FABIAN KRÜGER ist seit Oktober 2016 als Juniorprofessor für Empirische Volkswirtschaftslehre am Alfred-Weber-Institut für Wirtschaftswissenschaften der Universität Heidelberg tätig. Von 2004 bis 2009 studierte er Volkswirtschaftslehre an der Universität Konstanz, an der er 2013 auch promoviert wurde. Im Anschluss war er bis 2016 als Postdoktorand in Statistik am Heidelberger Institut für Theoretische Studien (HITS) tätig. Seine Forschung befasst sich mit theoretischen und praktischen Aspekten ökonomischer Prognosen und wurde in führenden internationalen Fachzeitschriften veröffentlicht.

beitragen können. Vielmehr ist die Wirtschaftsleistung eine wilde Mischung aus Konsumentenverhalten, politischen Entscheidungen, internationalen Handelsverflechtungen und vielem mehr – kurz: Sie ist ein Ergebnis menschlichen Handelns. Wie und mit welcher Genauigkeit lässt sich eine solche Zeitreihe vorhersagen?

Eine wichtige Quelle zur Erstellung wirtschaftlicher Prognosen sind Expertenumfragen, wie sie zum Beispiel von der Europäischen Zentralbank (EZB) durchgeführt werden. Hier geben jedes Quartal rund 50 Wirtschaftsstatistiker aus der Privatwirtschaft sowie aus Forschungseinrichtungen ihre Einschätzung der zukünftigen wirtschaftlichen Entwicklung ab. Während die EZB-Umfrage erst seit Einführung des Euro Anfang 1999 durchgeführt wird, existiert eine ähnliche Umfrage in den USA bereits seit 1968. Ein Blick in die Vergangenheit zeigt, dass die Umfrageprognosen im Allgemeinen mindestens gleich präzise waren wie rein statistische Prognosen; auf kurzen Prognosehorizonten (zum Beispiel für die nächste veröffentlichte Wachstumsrate) ergeben sich aber zum Teil deutliche Vorteile zugunsten der Umfragen. Dieser Befund lässt sich dadurch erklären, dass die Umfrageteilnehmer einerseits selbst statistische Modelle verwenden und zudem auf aktuelle Hintergrundinformationen zurückgreifen können, die zum Teil nur schwer in ein formales statistisches Modell einbezogen werden können. Ein weiterer Vorteil von Umfrageprognosen liegt darin, dass die einzelnen Teilnehmer sich zum Teil grundsätzlich voneinander unterscheiden – etwa hinsichtlich der statistischen Modelle und Datenquellen, die sie verwenden. Statistische Auswertungen haben ergeben, dass diese Heterogenität höchst wünschenswert ist: So schneidet eine Kombination aller Prognosen auf lange Sicht meist besser ab als die einzelnen Prognosen. Dieser Effekt – oft als „Weisheit der Vielen“ (englisch „Wisdom of the Crowd“) bezeichnet – wurde in vielen verschiedenen Anwendungsbeispielen quer über die Fachgebiete beobachtet.

Spiegel menschlichen Verhaltens

Trotz der erwähnten Vorteile von Umfrageprognosen sollte deren Informationsgehalt stets kritisch hinterfragt, statistisch ausgewertet und mit anderen Verfahren abgeglichen werden. Während diese Aussage wohl für fast alle Datenquellen gilt, erscheint sie mir in Bezug auf wirtschaftliche Zeitreihen besonders relevant, da deren statistische Eigenschaften sich im Zeitverlauf ändern können. Kennzahlen wie das BIP und die Inflationsrate spiegeln letztlich menschliches Verhalten wider; folgerichtig machen gesellschaftliche und institutionelle Entwicklungen auch vor diesen Kennzahlen nicht halt. So erreichte beispielsweise die deutsche Inflationsrate in den 1970er-Jahren Werte um acht Prozent, während heute Werte von weniger als zwei Prozent die neue Normalität darstellen. Im Rückblick ist man geneigt, solche Entwicklungen als logisch und unvermeidlich

„Kennzahlen wie BIP und Inflationsrate spiegeln letztlich menschliches Verhalten wider; folgerichtig machen gesellschaftliche und institutionelle Entwicklungen auch vor ihnen nicht halt.“

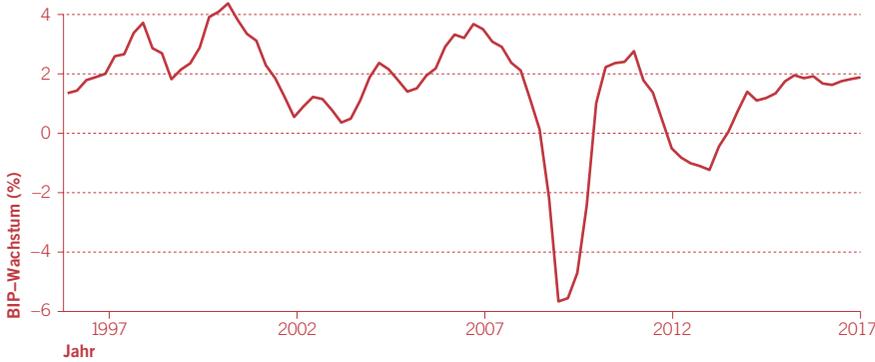


Abbildung 1
Wirtschaftswachstum im Euroraum (reales BIP, in %, 1996–2017)

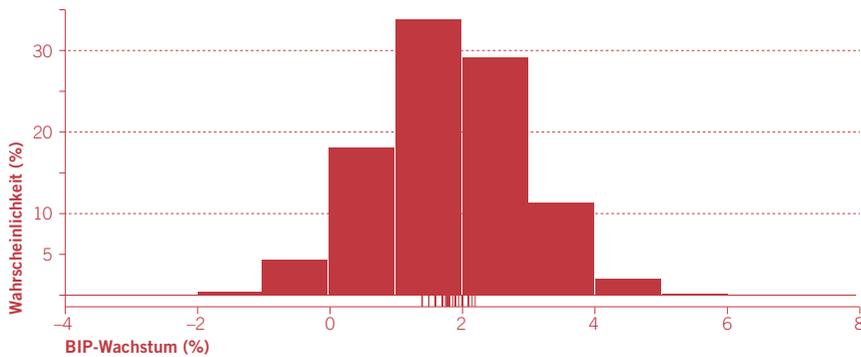


Abbildung 2
Umfragebasierte Prognose für das BIP-Wachstum im ersten Quartal 2018

anzusehen. Die Herausforderung von Prognosen besteht jedoch gerade darin, zukünftige Trends und Tendenzen zu erkennen. Der dänische Philosoph Sören Kierkegaard trifft den Nagel auf den Kopf, wenn er schreibt: „Es ist ganz wahr, was die Philosophie sagt, dass das Leben rückwärts verstanden werden muss. Aber darüber vergisst man den andern Satz, dass vorwärts gelebt werden muss.“ Neben den beschriebenen wirtschaftlichen Transformationsprozessen gibt es auch praktische Gründe, warum sich die Güte von Umfrageprognosen mit der Zeit ändern kann, etwa durch den Ein- und Austritt von Umfrageteilnehmern oder durch technische Änderungen im Format der Umfrage. Zusammengefasst ist die Analyse ökonomischer Expertenbefragungen also ein hochrelevantes Forschungsthema, zu dem am Heidelberger Alfred-Weber-Institut für Wirtschaftswissenschaften intensiv geforscht wird.

In einem aktuellen Forschungsprojekt mit Wirtschaftswissenschaftlern der Lancaster University in Großbritannien gehen wir der Frage nach, wie man die Prognoseunsicherheit in ökonomischen Expertenbefragungen messen kann. Ein Beispiel ist die EZB-Umfrage vom 21. Juli 2017, in der die Teilnehmer das Wirtschaftswachstum im ersten Quartal 2018 prognostizieren sollen. Der letzte Kenntnisstand der Umfrageteilnehmer ist die Wachstumsrate des ersten Quartals 2017 – also genau der letzte Datenpunkt aus Abbildung 1. Die Punktprognosen der 50 Umfrageteilnehmer bewegen sich zwischen 1,4 Prozent und 2,2 Prozent. Doch wie lässt sich aus diesen Zahlen eine Einschätzung der zugrunde liegenden Prognoseunsicherheit gewinnen? Um diese Frage zu beantworten, greifen wir auf mathematisch-statistische Verfahren zurück, die definieren, wodurch sich eine „vernünftige“ Einschätzung der Prognoseunsicherheit auszeichnet. Grob gesagt sollte die angegebene Prognoseunsicherheit so klein wie möglich sein – gleichzeitig aber auch groß genug, um den Schwankungen in den Daten Rechnung zu tragen, die man in der Vergangenheit beobachten konnte. Basierend auf diesem Prinzip haben wir ein einfaches statistisches Modell entwickelt, mit dem sich die Unsicherheit der vorliegenden Punktprognosen messen lässt. Das Modell schätzt aufgrund vergangener Daten, wie stark sich die Punktprognosen von den später tatsächlich beobachteten Werten unterscheiden. Die Einfachheit des Modells ist beabsichtigt: Zu komplizierte Modelle lassen sich oft von rein zufälligen Mustern in den Daten auf eine falsche Fährte locken. Dies wird in der Statistik als „Überanpassung“, englisch „overfitting“, bezeichnet. Abbildung 2 veranschaulicht das Modell anhand der aktuellen EZB-Umfrage. Die Höhe der Balken steht hier für die Wahrscheinlichkeit, dass die BIP-Wachstumsrate in einem bestimmten Bereich landet. So wird die Wahrscheinlichkeit für eine Rate zwischen null Prozent und einem Prozent auf knapp 20 Prozent geschätzt und die Wahrscheinlichkeit für eine Rate zwischen einem Prozent und zwei Prozent auf knapp 35 Prozent.

(UN)CERTAIN

PUTTING FORECASTS TO THE TEST

FABIAN KRÜGER

Economic forecasts are surrounded by considerable uncertainty. While this basic fact has long been recognised, the systematic measurement and communication of forecast uncertainty has only recently entered the centre stage of economic research. Measuring forecast uncertainty requires the forecaster to quantify her or his own ignorance. In terms of methodology, this task requires specialised statistical tools that evaluate how well measures of forecast uncertainty correspond to observed data. Economists have recently applied and refined such tools.

Consider, for example, the euro area's gross domestic product (GDP), which is an important measure of economic activity. In the first quarter of 2009, at the height of the recent financial crisis, the euro area's GDP shrank by 5.7%. Since then, it has grown by an average of 0.6%, with the most recent growth rate being 1.9%. Unlike in other sciences, there are no laws of nature in economics that could provide guidance on the future course of GDP growth. Instead, GDP measures human behaviour as reflected in the economic decisions of consumers, firms, and policy-makers. How – and how precisely – can we forecast the future development of this economic indicator?

In my recent research, I have constructed measures of uncertainty for expert forecasts recorded in specialised surveys, such as are conducted by the European Central Bank (ECB) and other policy institutions. They have been found to perform similar to, or better than, purely statistical forecasting methods. ●

PROF. DR FABIAN KRÜGER has held the Junior Professorship of Empirical Economics at Heidelberg University's Alfred Weber Institute for Economics since October 2016. From 2004 to 2009, he studied macroeconomics at the University of Constance, where he also obtained his PhD. Following his transfer to Heidelberg, he worked as a post-doctoral researcher for statistics at the Heidelberg Institute for Theoretical Studies (HITS) until 2016. Dr Krüger's research focuses on theoretical and practical aspects of economic forecasts; his findings have been published in leading international economic journals.

“The challenge of forecasts lies in recognising future trends and tendencies.”

**„Es ist ganz wahr, was die
Philosophie sagt,
dass das Leben rückwärts
verstanden werden muss.
Aber darüber vergisst man den
andern Satz, dass
vorwärts gelebt werden muss.“**

Sören Kierkegaard

Abbildung 2 zeigt auch die Punktprognosen der einzelnen Umfrageteilnehmer; diese sind als senkrechte rote Striche („Teppich“) am unteren Bildrand zu sehen. Die Spannweite dieser Punktprognosen (1,4 Prozent bis 2,2 Prozent) ist relativ gering, und sollte nicht als Einschätzung von Prognoseunsicherheit missverstanden werden.

Das Grundprinzip, vergangene Prognosefehler zur Abschätzung zukünftiger Prognosefehler zu verwenden, stammt übrigens aus der Meteorologie. Dort gilt es, die Prognosen computerbasierter Wettermodelle mit den später tatsächlich gemessenen Beobachtungen abzugleichen. Das Prinzip kann überall dort angewandt werden, wo es systematische Daten zu vergangenen Prognosen und den dazugehörigen Beobachtungen gibt. Die Schwierigkeit liegt, wie so oft, im Detail: Inwieweit sind vergangene Muster noch aktuell? Wie interagieren verschiedene Formen von Prognoseunsicherheit? Welche theoretischen und praktischen Eigenschaften haben statistische Verfahren zur Messung von Unsicherheit? Diese Fragen definieren – auch im Hinblick auf Digitalisierung und „Big Data“ – eine hochaktuelle und herausfordernde Forschungsagenda. ●