

IZA Standpunkte Nr. 17

Googlemetrie und Arbeitsmarkt in der Wirtschaftskrise

Nikos Askitas
Klaus F. Zimmermann

August 2009

Googlemetrie und Arbeitsmarkt in der Wirtschaftskrise

Nikos Askitas

IZA

Klaus F. Zimmermann

Universität Bonn, IZA und DIW Berlin

IZA Standpunkte Nr. 17
August 2009

IZA

Postfach 7240
53072 Bonn

Tel.: (0228) 3894-0
Fax: (0228) 3894-180
E-Mail: iza@iza.org

Die Schriftenreihe "IZA Standpunkte" veröffentlicht politikrelevante Forschungsarbeiten und Diskussionsbeiträge von IZA-Wissenschaftlern, IZA Research Fellows und IZA Research Affiliates in deutscher Sprache. Die Autoren sind für den Inhalt der publizierten Arbeiten verantwortlich. Im Interesse einer einheitlichen Textzirkulation werden Aktualisierungen einmal publizierter Arbeiten nicht an dieser Stelle vorgenommen, sondern sind gegebenenfalls nur über die Autoren selbst erhältlich.

ZUSAMMENFASSUNG

Googlemetrie und Arbeitsmarkt in der Wirtschaftskrise

Die große Wirtschaftskrise hat bisher nur verhaltene Spuren am Arbeitsmarkt hinterlassen. Angesichts der unsicheren weiteren konjunkturellen Entwicklung, der schlechten Auslastung der Arbeitskräfte in den Unternehmen und der hohen Kurzarbeit erwarten viele Beobachter zum Herbst einen dramatischen Anstieg der Arbeitslosigkeit mit einer baldigen Überschreitung der Vier-Millionen-Grenze. Nach Prognosen unter Verwendung von Google-Internetzugriffstatistiken bleibt es aber im Vorfeld der Bundestagswahlen in den Sommermonaten August und September aller Voraussicht nach völlig ruhig. Saisonal bedingt geht die Arbeitslosigkeit sogar zurück. Damit verringert sich die Gefahr, dass eine Arbeitslosenzahl von vier Millionen noch in diesem Jahr erreicht werden wird.

JEL-Codes: C22, C82, E17, E24, E37

Schlagworte: Google, Internet, Suchmaschinen, Schlüsselworte, Arbeitslosigkeit, Prognosen, Evaluation

Kontaktadresse:

Klaus F. Zimmermann
IZA
Postfach 7240
D-53072 Bonn
E-mail: Zimmermann@iza.org

Krisenpropheten beschwören seit langem einen Einbruch der deutschen Beschäftigung und einen drastischen Anstieg der Arbeitslosigkeit. Politik und Bundesagentur für Arbeit bemühen sich, die Entwicklung durch den Einsatz des arbeitsmarktpolitischen Instrumentariums und neuer Arbeitslosigkeitsabgrenzungen im Griff zu behalten. Tatsächlich entwickelte sich die Arbeitslosigkeit bis zum Juli 2009 eher entspannt. Dazu trägt auch das robuste Konsumklima bei, das durch diese niedrige Arbeitslosigkeit und den drastischen Rückgang der Inflation getragen wird. Es besteht allerdings Bedarf, frühzeitig Veränderungen der Lage zu erkennen. Dazu benötigt man geeignete Indikatoren. Sie sind im Internet mit der Suchstatistik von Google rasch und in hoher Frequenz verfügbar. Das IZA hat mit seinem neuen Arbeitslosigkeitsbarometer ökonometrisches Neuland betreten, um die Nützlichkeit dieser Internetdaten im realen Labor der Bundesrepublik Deutschland zu prüfen, nachzuweisen und zu verfeinern. Dies trägt der Tatsache Rechnung, dass immer mehr Menschen das Internet nutzen, insbesondere auch für Arbeitsmarktfragen.

Die Arbeitsmarktlage hat sich in den vergangenen Jahren erstaunlich entwickelt. Angetrieben durch die Arbeitsmarktreformen und eine boomenden Weltwirtschaft wurden in Deutschland insbesondere für Problemgruppen (ältere, ungelernte oder ausländische Arbeitnehmer sowie Langzeitarbeitslose) viele gute Jobs geschaffen. Verbunden damit fiel die saisonunbereinigte Arbeitslosenquote¹ bis zum November 2008 und stieg erst im Dezember leicht an, obwohl die Rezession bereits im zweiten Quartal des vergangenen Jahres einsetzte. Zwar stieg die Arbeitslosenquote bis zum März/April 2009 auf 8,6%, lag aber angesichts des gleichzeitig massiv eingebrochenen Wirtschaftswachstum² nur moderat über den Vorjahreswerten (8,4%/8,1%).

Im Mai fiel die Arbeitslosenquote dann um fast einen halben Prozentpunkt auf 8,2%, im Juni noch einmal auf 8,1%, um im Juli nur sehr geringfügig auf wiederum 8,2% zuzunehmen. Im Wesentlichen ist dies auf eine verzögerte Frühjahrsbelebung zurückzuführen, die im März und April wegen des schlechten Wetters ausgeblieben war. Aufgrund dieser Belebung benötigen saisonabhängige Branchen wie das Gastgewerbe und

¹ Wir verwenden hier grundsätzlich saisonunbereinigte Daten, weil die von uns benutzten Google-Internetaktivitätsdaten saisonunbereinigt sind und Saisonbereinigungen in Zeiten rascher Veränderungen für sich genommen Probleme verursachen können. Wo dies nicht der Fall ist wird dies genau benannt.

² Das reale Bruttoinlandsprodukt war im ersten Quartal 2009 im Vergleich zum Vorjahr um 3,8% zurückgegangen.

der Bau wieder mehr Personal als im Winter. Ohne die vermehrte Inanspruchnahme von Kurzarbeit wäre das Resultat aber wohl auch nicht erreichbar gewesen.

Sommerpause bei der Arbeitslosigkeit?

Die große Wirtschaftskrise hat insgesamt bislang nur verhaltene Spuren am Arbeitsmarkt hinterlassen. Angesichts der unsicheren weiteren konjunkturellen Entwicklung, der schlechten Auslastung der Arbeitskräfte in den Unternehmen und der hohen Kurzarbeit erwarten viele Beobachter jedoch zum Herbst einen dramatischen Anstieg der Arbeitslosigkeit mit einer baldigen Überschreitung der Vier-Millionen-Grenze.

Der Boden der Rezession in Deutschland scheint zwar erreicht zu sein und die wichtigen Absatzregionen in Asien und den Vereinigten Staaten zeigen Zeichen einer Besserung der Lage. Die Struktur der Krise, die vor allem die exportorientierten Investitionsgüterindustrien mit ihren Fachkräften getroffen hat, lässt jedoch erwarten, dass sich erst nach einer konjunkturellen Erholung auch in den west- und osteuropäischen Ökonomien wieder ein deutlich wachsender Bedarf an deutschen Investitionsgütern einstellt.³ Dann bleibt bis tief in das nächste Jahr hinein die zentrale Frage, inwieweit diese Leistungsbranchen ihre massiv unterausgelasteten Fachkräfte halten können. Dafür spricht das Kurzarbeiterprogramm der Bundesregierung und das Eigeninteresse der Firmen, die angesichts des schon vor der Krise bestehenden und auch wegen des langfristig sich abzeichnenden Fachkräftemangels ihre Stammbeschaften so weit und so lange es geht halten wollen.

Die Zahl der Kurzarbeiter hat sich in den letzten Monaten kaum erhöht, wie Abbildung 1 zeigt. Auch im Juni 2009 lag sie bei 1,4 Millionen. Das sind immerhin reichlich 5% aller sozialversicherungspflichtig Beschäftigten; in der Industrie sogar etwa 20%. Der Kurzarbeiterbestand sinkt aktuell – vor allem urlaubsbedingt –, ohne dass dies zu größeren Freisetzungen führen dürfte. Dies dokumentiert bereits die Entwicklung der Arbeitslosigkeit im Juli. Während sich in diesem Monat üblicherweise vermehrte Entlassungen aufgrund des vorhergehenden Quartalsendes in einem deutlichen Anstieg der Arbeitslosenzahlen niederschlagen, ist dieser Zuwachs in diesem Jahr bemerkenswert gering ausgefallen. Für eine Veränderung der konservativen Haltung der Firmen gibt es wenig Anzeichen. Zudem sprechen Kündigungsfristen von durchschnittlich drei Monaten, die in vielen Unternehmen

³ Vgl. Dreger et al. (2009).

existierenden längerfristigen Jobgarantien und in der Krise mancherorts geschlossene Beschäftigungspakte dafür, dass es auch im Herbst allenfalls zu einem verhaltenen Anstieg der Arbeitslosigkeit kommen wird.

Vor dem Hintergrund dieser Entwicklung ist es zur Bestätigung der Einschätzung wichtig, wie die kritischen Monate August und September verlaufen werden. Denn üblicherweise hat die wirtschaftliche Entwicklung einen prägenden Einfluss auf wichtige Wahlen.⁴ Im amerikanischen Präsidentschaftswahlkampf 1992 spielte dieser Gedanke als Slogan „*It's the economy, stupid!*“ eine wichtige Rolle, mit dem die Kampagne von Bill Clinton entscheidende Punkte gewann. Demzufolge sollten in Zeiten einer schweren Wirtschaftskrise wie gegenwärtig auch der Bundestagswahlkampf und die Erfolgchancen der Parteien durch Wirtschaftsthemen geprägt sein. Nichts davon ist zu sehen. Im Gegenteil, alle Parteien halten sich bei diesen Fragen auffallend zurück; es gelingt nicht, die Menschen mit kontroversen Themen zu mobilisieren. Dramatisch anwachsende Arbeitslosenzahlen im Vorfeld der Bundestagswahlen am 27. September 2009 oder eine signifikante Einschätzung einer solchen Entwicklung für die unmittelbare Zeit danach könnten allerdings zu einer größeren Relevanz der Wirtschaftsthemen führen.⁵

Unter Verwendung von Google-Daten über die Internetsuchaktivität legen wir deshalb hierzu eine Einschätzung vor. Dabei werden die gemessenen Suchabfragen zur Bundesagentur für Arbeit, nach Jobs und zu Kurzarbeit benutzt, um die Arbeitslosenquote zu erklären. Das Schätzmodell wird dann für die Prognose benutzt. Der Einsatz der Google-Daten in diesen ökonometrischen Modellen lässt einen weiteren Rückgang der Arbeitslosenquote für August erwarten. Schreibt man die Internetdaten mit ökonometrischen Methoden fort, so kann man auch einen ersten Blick auf den Folgemonat werfen. In der „Glaskugel“ kann auch für September ein weiterer Rückgang der Arbeitslosigkeit identifiziert werden. Das signalisiert insgesamt eine erfreuliche Sommerpause bei der Arbeitsmarktentwicklung.

⁴ Auch in der letzten Präsidentschaftswahl in den USA brachte die Wirtschaftskrise die entscheidende Wende zugunsten des in Wirtschaftsfragen kompetenter erscheinenden Barack Obama gegenüber dessen Kontrahenten John McCain, dessen Kompetenz vor allem in Sicherheitsfragen wahrgenommen wurde. Dieses Phänomen wurde unter Verwendung von Google-Internetzugriffsstatistiken bereits vor der Wahl belegt, vgl. Constant und Zimmermann (2008).

⁵ So gehen etwa die Sommergrundlinien des DIW Berlin von einer deutlichen Verschärfung der Arbeitsmarktlage bereits im Herbst aus, vgl. Dreger et al. (2009).

Stürmisches Wasser für Prognosen

In der gegenwärtigen Wirtschaftskrise haben sich Prognosen als besonders umstritten erwiesen. Bei ihrem Beginn haben Panikvorhersagen die Krise verschärft, da sie insbesondere dazu beitrugen, dass die Investitionspläne der Unternehmen weltweit über Nacht zurückhaltender wurden.⁶ Stimmungen beeinflussen die Realität und Prognosen können deshalb Krisen verstärken. Dies ist ein weltweites Phänomen. In immer kürzerer Folge kam es zu Prognoserevisionen, die letztlich in einem Herdenverhalten der Prognostiker endete – ein typisches Zeichen für mangelnde Informationen im Markt. Dies hatte zunächst mit der Geschwindigkeit zu tun, mit der sich in der globalisierten Welt die negativen Impulse, die aus einem Zusammenspiel einer zyklischen Abschwächung der Weltkonjunktur und dramatischer Krisensignale aus den Finanzmärkten entstanden, über die ganze Welt verbreiteten und insbesondere das Investitionsklima eintrübten. Mit dieser Geschwindigkeit war die traditionelle Konjunkturforschung und die amtliche Statistik überfordert, da sich die Anpassungsprozesse in Tagen oder Wochen vollzogen, und nicht wie sonst üblich in Monaten oder Quartalen.

Aber auch langfristige Prognosen der Wirtschaftsentwicklung wurden problematischer. Schon zu normalen Zeiten sind sie für das Folgejahr wenig treffsicher, wenn sie früh im Jahr gemacht werden.⁷ Deshalb mussten auch alle Vorhersagen für 2010 im Frühjahr 2009 mit besonderer Vorsicht bewertet werden. Das heißt nun gar nicht, dass für 2009 keine Prognosen gemacht werden können. Nachdem der freie Fall der Wirtschaft nach dem ersten Quartal dieses Jahres beendet ist, kann das traditionelle Prognosegeschäft im Gegenteil wieder an Boden gewinnen.

In der Krisenphase und jetzt im Übergang wurden und wird noch mehr als sonst auf „weiche“ Indikatorsysteme wie Stimmungsumfragen und Handelsindizes zurückgegriffen. Auch wenn sie wenig über die weitere Zukunft sagen können, so geben sie doch in normalen Zeiten ein robustes Bild über die Lage, in der sich die Wirtschaft jeweils befindet.⁸ Aber auch weiche Indikatoren lagen in der jetzigen Krise nicht immer rechtzeitig vor. Schließlich ist es die Wirtschaftspolitik selbst, die über Maßnahmenpakete zur Konjunkturstabilisierung und definitorische Änderungen bedeutender Zielgrößen das Prognosegeschäft erschwert. Die weichen Indikatoren haben auch den Nachteil, dass sie nicht sofort mit den interessierenden Zielgrößen in einen quantitativen Zusammenhang

⁶ Vgl. Zimmermann (2008a, 2008b, 2009).

⁷ Vgl. Kholodilin und Silverstovs (2009).

gebracht werden können. Dies gilt insbesondere, wenn sie neue Phänomene aufgreifen, die noch nicht zeitlich lange abgebildet werden können. Bei den hier verwendeten Google-Daten handelt es sich um bisher für wissenschaftliche Zwecke ungenutztes Material, mit dem zunächst Erfahrungen gesammelt werden müssen. Das IZA-Arbeitslosigkeitsbarometer bietet jedoch einen unmittelbaren Vergleich mit der Zielgröße, weil es dem Nutzer die direkte Verknüpfung mit der Arbeitslosenquote ermöglicht. Diese Möglichkeit hat der Betrachter bei anderen Barometern nicht, sondern muss Entwicklungen aus Kurvenverläufen abschätzen.

Saisonale Muster im tatsächlichen Verlauf der Arbeitslosigkeit

Das Ausmaß der Arbeitslosigkeit schwankt im Zeitablauf erheblich, folgt dabei jedoch wiederkehrenden saisonalen Mustern. Abbildung 2 zeigt die amtliche, saisonunbereinigte Arbeitslosenquote für Januar 1992 bis Juli 2009; sie macht deutlich, dass die heute gemessene Arbeitslosigkeit im historischen Vergleich sehr niedrig ist – selbst wenn man die Tatsache ignoriert, dass wir noch in der Mitte einer der schwersten Wirtschaftskrisen seit Jahrzehnten stehen.

Das saisonale Muster ist unverkennbar. Die Arbeitslosigkeit ist in den Wintermonaten besonders hoch, vor allem im ersten Quartal des Jahres. Danach geht sie immer mehr zurück. Zur Jahresmitte, also vor den Ferien und wegen der in vielen Arbeitsverträgen vereinbarten Entlassungstermine zum Quartalsende, zieht sie oft im Juli kurzzeitig etwas an, um in den nächsten Monaten danach zu stagnieren, und im Herbst zu sinken. In den Ferienmonaten im Sommer geht die Arbeitslosigkeit etwas zurück, weil sich Entlassungen in engen Grenzen halten und es in Saisonjobs eher zu Neueinstellungen kommt. Ab September sinkt die Arbeitslosigkeit in der Regel etwas stärker, da vermehrt Einstellungen vorgenommen werden und das neue Ausbildungsjahr beginnt. Der Monat mit der geringsten Arbeitslosigkeit im Jahresverlauf war von 1994 bis 2008 in der Regel der Oktober oder der November, seltener der September. Zum Winter hin steigt die Zahl der Arbeitslosen, weil saisonbedingt die Arbeitskräftenachfrage nachlässt – beispielsweise im vom Tourismus abhängigen Gastgewerbe. In den ersten Wintermonaten eines neuen Jahres muss häufig die Bauproduktion wegen Frosts eingeschränkt werden, was die Arbeitslosenzahlen steigen lässt.

⁸ Vgl. Brenke et al. (2009), S. 16-17.

In Abbildung 3 sind die Jahresabläufe unabhängig vom jeweiligen Jahr dargestellt: Zur besseren Sichtbarkeit wurde jeder Monatswert mit dem Juliwert des jeweiligen Jahres geteilt und dann der Wert in diesem Monat auf Null verschoben. Die Linien in Abbildung 3 zeigen also die prozentuale Abweichung der tatsächlichen Arbeitslosenquote in den einzelnen Monaten des jeweiligen Jahres vom Juliwert. Die Abbildung verdeutlicht, dass es (bis auf eine Ausnahme) in keinem der hier erfassten Jahre im Juli zu einem Rückgang der Arbeitslosigkeit gekommen war.

Aus saisonaler Betrachtungsweise ist dagegen im Monat August ein Rückgang der Arbeitslosigkeit sehr wahrscheinlich. Dies verdeutlicht Abbildung 4, bei der jetzt die Normierung auf diesen Monat erfolgte. Unabhängig von der konjunkturellen Lage fällt im Beobachtungsjahr die Arbeitslosigkeit durchgehend von Juli auf August und danach weiter von August auf September. In beiden Monaten ist also aus saisonalen Gründen mit hoher Wahrscheinlichkeit ein Rückgang der Arbeitslosigkeit zu erwarten. Erst für den Oktober ist das saisonale Muster etwas divers: Es kam in der Vergangenheit, wie Abbildung 5 zeigt, im Übergang vom September sowohl zu Anstiegen (1992 und 1993) wie (in allen anderen Jahren) zu Rückgängen der Arbeitslosigkeit.

Die historischen Erfahrungen mit der Saisonkomponente besagen freilich nicht, dass die Prognose schon entschieden wäre, bevor man sie durchführt. Dass im Beobachtungszeitraum die Arbeitslosigkeit saisonal im August und September immer zurückging, heißt nicht, dass dies ein Naturgesetz wäre, sondern nur, dass diese Entwicklung eher wahrscheinlich ist. Umgekehrt gilt: Sollte eine Prognose von diesem Muster abweichen, würde ein angekündigter Anstieg in diesen Monaten auf ein großes potenzielles Krisenzeichen hinweisen. Das tatsächliche Eintreten eines Anstiegs in diesem Zeitraum wäre ein beklemmendes Signal für den Arbeitsmarkt. Die Anzeichen, wie sie auch der Google-Indikator liefert, sprechen dagegen, dass dieser Fall eintreten wird.

Sollte man angesichts der komplexen Saisonstrukturen bei Prognosen nicht auf saisonbereinigte Arbeitslosenquoten ausweichen? Hierfür spricht wenig, denn zum einen ist es die unbereinigte Zahl, die die breite Öffentlichkeit interessiert, weil sie unmittelbar einleuchtend ist, aber auch für die Politik hoch relevant ist, da sie Wählerstimmungen beeinflusst: Aber es lassen sich zum anderen auch diverse technische Gründe dafür nennen, mit Saisonbereinigungen bei Prognosen zurückhaltend vorzugehen. Sie sind sensibel für die eingesetzten Verfahren, so dass schnell ein stochastischer Prozess künstlich produziert wird, den man dann zu modellieren versucht. Wir verfügen hier zudem nur über kurze Zeitreihen,

die ferner durch ständige Änderungen der Definition der Arbeitslosenquote, durch Arbeitsmarktprogramme und die Modifizierung von Instrumenten wie die Kurzarbeit und aktuell nicht zuletzt durch die besonders schwere Wirtschaftskrise gestört sind. Es kann kaum angenommen werden, dass all dies die Saisonbereinigung nicht beeinflusst. Zur Modellierung der Arbeitslosenquote ist deshalb allemal die unbereinigte Reihe vorzuziehen.

Googlemetrie und Arbeitsmarkt

Inzwischen surfen knapp 70% der Deutschen im Internet und damit mehr als der europäische Durchschnitt.⁹ Internetdaten liefern deshalb eine interessante, bisher praktisch wissenschaftlich ungenutzte Datengrundlage.¹⁰ Sie sind rasch und umfangreich verfügbar und reagieren flexibel auf Änderungen der Rahmenbedingungen. Ihre Nutzung für Arbeitsmarktfragen erscheint aussichtsreich, da nach dem IZA-Evaluationsdatensatz über 86% der Arbeitslosen auch das Internet für ihre Jobsuche nutzen.¹¹ Fast jeder der Arbeitssuchenden hinterlässt also Spuren im Internet.

Mit Googlemetrie kann sich eine wissenschaftliche Subdisziplin herausbilden, die diese Daten zur Formulierung und Beantwortung von Fragen nutzt, die mit den üblichen empirischen Daten so nicht beantwortet werden können. Auf diese Weise kann Googlemetrie einen wichtigen Beitrag dazu liefern, die immer mehr an Relevanz gewinnenden Suchaktivitäten im Internet für Verhaltensanalysen und Prognosen messbar zu machen. Askitas und Zimmermann (2009a) haben hierzu ein ökonometrisches Konzept vorgestellt und auf die Arbeitslosigkeit angewendet. In weiteren Studien haben sie die bisherigen Prognoseerfahrungen dokumentiert und untersucht.¹²

Die Arbeitslosenquote eines Monats wird in der Regel an seinem Ende verkündet. Askitas und Zimmermann (2009a) haben gezeigt, dass sie am besten mit der Google-Internetaktivität der zweiten Hälfte des Vormonats korreliert. Dies lässt sich auch durch die Verwaltungsabläufe in der Bundesagentur für Arbeit belegen, die einen verzögerten Informationstransfer von der Entstehung bis zur Messung von Arbeitslosigkeit implizieren. Die Arbeitslosenquote ist stichtagsbezogen für die Monatsmitte. Die zusammengefassten

⁹ Vgl. (N)Onliner Atlas 2009.

¹⁰ Constant und Zimmermann (2008) und Ginsberg et al. (2009) haben diese Datenquelle bereits im Zusammenhang mit den amerikanischen Präsidentschaftswahlen bzw. mit Grippeepidemien erstmals ausgewertet.

¹¹ Vgl. Caliendo et al. (2009). Das ist mehr als bei jeder anderen Alternative. Mögliche (nicht ausschließende) Alternativen sind u.a. (i) Freunde und Familie (84%), (ii) Zeitungsanzeigen (83%) und (iii) Arbeitsagentur (70%). Die lokalen Arbeitsagenturen ermöglichen ihren Kunden vor Ort den Zugang zum Internet.

Wochendaten der zweiten Monatshälfte für die Regressoren werden deshalb als Prädiktoren für die Arbeitslosenquote der Folgeperiode eingesetzt. Für die Indikatorbildung heißt dies aber, dass man zum Zeitpunkt der amtlichen Verkündung einer Arbeitslosenzahl auf Basis dieses Verfahrens bereits eine Prognose für den Folgemonat abgeben kann. Der Informationsvorsprung beträgt also einen Monat.

Die Studie von Askitas und Zimmermann (2009a) hat sich intensiv mit der Auswahl der Schlüsselbegriffe aus der Suchaktivität der Internetbenutzer beschäftigt. Wir verwenden drei alternative Indikatormodelle zur Prognose der Arbeitslosenquote unter Benutzung diverser Schlüsselworte:

1. Arbeitsamt, Arbeitsplatzsuche (Google 1)
2. Arbeitsamt, Arbeitsplatzsuche, Kurzarbeit (Google 2) und
3. Arbeitsplatzsuche, Kurzarbeit (Google 3).¹³

Die Entwicklung der Arbeitslosigkeit lässt sich historisch mit diesen nur wenigen Variablen sparsam und statistisch sehr zufrieden stellend erfassen. Die Internetaktivität bezüglich dieser Schlüsselwörter aus der zweiten Monatshälfte wird mit den Monatsangaben der Bundesagentur für Arbeit am Ende des Folgemonats für den abgelaufenen Monat verknüpft. Das Google-Datenmaterial ist seit Januar 2004 bis zum aktuellen Rand im Prinzip täglich verfügbar. Wir nutzen aber nur die Daten der zweiten Monatshälfte, die wir aggregieren.

Die gegenwärtige komplexe Konjunktur- und Arbeitsmarktlage liefert einen idealen und anspruchsvollen Testrahmen für diesen innovativen Prognose-Ansatz. Zunächst zeigt sich, dass die in Abbildung 6 dargestellten wöchentlich gemessenen Internetzugriffe auf Kurzarbeit mit den Meldungen von Kurzarbeit durch die Firmen aus Abbildung 1 hoch korrelieren. Die Internetzugriffe bieten allerdings wegen der kürzer gemessenen Zeitfrequenz und ihrer unmittelbaren Verfügbarkeit direkt nach Ende des Untersuchungszeitraums entscheidende Vorteile. Offensichtlich ist das öffentliche Interesse im Verlauf des Dezembers 2008 zunächst zurückgegangen und stieg erst zu Beginn des Jahres 2009 auf ein hohes Niveau, das im Februar und März anhielt. Seither geht auch dieser Indikator zurück.

¹² Askitas und Zimmermann (2009b) und (2009c).

¹³ Die genaue Abgrenzung der Schlüsselworte ist: „Arbeitsamt“ = (Arbeitsamt OR Arbeitsagentur); „Arbeitsplatzsuche“ = (Stepstone OR Jobworld OR Jobscout OR Meinestadt OR meine Stadt OR Monster Jobs OR Monster de OR Jobboerse); „Kurzarbeit“ = (Kurzarbeit).

Google-Modelle und Alternativen

Mit den Daten zur Google-Suchaktivität können mit Hilfe ökonomischer Verfahren ökonomisch sinnvolle Verhaltensmodelle geschätzt und für Prognosezwecke eingesetzt werden.¹⁴ Wir führen in dieser Studie zunächst eine Revision und Aktualisierung der bislang verwendeten Modellvarianten durch. Dazu wurden alle Google-Variablen geprüft und einige Selektionsveränderungen vorgenommen. Als Referenzmodell für die Google-Aktivitätsdaten haben wir den DAX verwendet, der sich gegenüber anderen Arbeitsmarktindikatoren als überlegen erwiesen hat.¹⁵ Wir verwenden ihn hier in laufender Periode, ziehen also den ersten verfügbaren Wert eines Monats heran. Er erweist sich grundsätzlich als starke Referenz. Analysezeitraum ist hier Januar 2005 bis Juli 2009.

Tabelle 1 enthält eine Auswertung der geschätzten Modelle. Dabei regressieren wir bei diesem Standardansatz die jährliche Veränderung der Arbeitslosenquote auf ihre 12. Veränderung und bei jeder weiteren Regressorvariablen ist ihre jährliche Veränderung und ihre 12. Verzögerung berücksichtigt. Drei allein auf Google-Daten basierende Auswertungen (Google 1 bis Google 3) werden einem Modell mit dem DAX alleine (DAX-Prognose) und einer Kombination aus Google-Daten und dem DAX (Google 1 + DAX bis Google 3 + DAX) gegenübergestellt.

In Tabelle 1 finden sich ferner das korrigierte Bestimmtheitsmaß R^2 -a und das Bayessche Informationskriterium (BIC) zur Auswahl der besten im Analysezeitraum gefundenen Spezifikationen. Die Gesamtperformance der Modelle ist sehr gut. Die Google-Daten schlagen dabei das DAX-Modell in jeder Spezifikation deutlich, wie diese Maße klar anzeigen.¹⁶ Google 2 mit den Regressoren Arbeitsagentur, Jobsuche, Kurzarbeit erweist sich dabei als das beste Google-Modell im Untersuchungszeitraum. Es wird aber auch klar, dass eine gemeinsame Regression der Google-Daten mit dem DAX noch erhebliche weitere Verbesserungen liefert. Das Modell Google 2 + DAX, in dem der DAX als weiterer Regressor hinzukommt, liefert allerdings die besten Verbesserungsmöglichkeiten.

In Abbildung 7 sind die ex post-Prognosen der besten Modelle in jeder Klasse (Google 2, DAX, Google 2 + DAX) der Entwicklung der Arbeitslosigkeit gegenüber gestellt. Die Modelle bilden die Entwicklung grundsätzlich gut ab. Dabei wird zu Beginn der Krise Ende 2008 der Anstieg der Arbeitslosigkeit überschätzt, danach wird ihr Niveau

¹⁴ Für methodische Details dazu vgl. Askatas und Zimmermann (2009a) und (2009c).

¹⁵ Vgl. Askatas und Zimmermann (2009c). Der um eine Periode verzögerte DAX schlägt z.B. das Ifo-Beschäftigungsbarometer. Der BAX der Bundesagentur für Arbeit ist für diese Zwecke noch nicht einsatzfähig.

unterschätzt. Dabei zeigen sich bei der Richtung keine erheblichen Performance-Unterschiede.

Allerdings ist bekannt, dass sich die Performance von Prognosemodellen erheblich anders darstellen kann, wenn man die Güte der Ein-Schritt-Prognosen dieser Modelle betrachtet. Für die Auswertung der Qualität der Ein-Schritt-Prognosen wird in Tabelle 1 MAE definiert als mittlerer absoluter Prognosefehler angegeben. Zunächst ist jetzt die DAX-Prognose allen reinen Google-Modellen (Google 1 bis Google 3) überlegen. Dies ist allerdings nur deshalb der Fall, weil das Google-Modell im Februar 2009 einen außergewöhnlichen Fehler erzeugt hat, der vor allem darauf zurückgeführt werden kann, dass sich trotz der Ausweitung von Kurzarbeit auch die Suchaktivitäten der Internetnutzer in Bezug auf neue Jobs deutlich intensiviert haben.

Verwendet man Google und DAX in einer Regression (Google + DAX), so reduziert sich der Fehler erheblich, beim Google + Dax - Modell im Vergleich zum reinen DAX - Modell um 24%. Weit geringere MAE's erhält man allerdings, wenn man die Prognosen aus den reinen Google und DAX - Modellen einfach mittelt (s. die AVG - Zeilen in Tabelle 1). Auch dies ist ein bekanntes Verfahren. Stützt man sich auf die beste Variante, Google 2 + DAX, mittelt also einfach die Prognosen aus der DAX-Regression und der Google-Regression (mit Arbeitsagentur, Jobsuche, Kurzarbeit), so reduziert sich der Fehler im Vergleich zum reinen DAX-Modell gar um 37%.

Prognosen für August und September

Wir wenden die beschriebenen Verfahren nun auf die Monate August und September 2009 an und beschränken uns dabei auf die beiden alternativen Modelle DAX - Prognose und Google 2 sowie ihre gemittelte Variante. Zwar schlägt Google 2 + DAX, das gemeinsame ökonomische Modell unter Verwendung der Google- und DAX-Daten, die getrennten Modelle für die Ein-Schritt-Prognosen, aber der Mittelwert der Ein-Schritt-Prognosen aus den beiden getrennten Modellen im Untersuchungszeitraum ist besser als die Ein-Schritt-Prognosen von Google 2 + DAX.

Die Prognosen für August ergeben sich dabei aus den vorliegenden Google-Daten für die zweite Monatshälfte im Juli und dem DAX-Wert vom 3. August. Für die September-Prognose liegen dagegen noch keine Angaben für die Regressoren aus dem Monat August vor. Wir haben deshalb die Regressoren mit Zeitreihenverfahren für August fortgeschrieben.

¹⁶ Ein Modell ist umso besser, je größer R^2 und je kleiner BIC ist.

Dies ist für den DAX bekanntlich problematisch, da sich sein Wert vor allem aus dem Fehlerterm (den sogenannten Innovationen) ergibt. Insoweit ist die in der Folge vorgestellte DAX-Prognose eher explorativer Natur. Dies muss bei der Bewertung der Septemberprognose besonders berücksichtigt werden.

Abbildung 8 enthält die Ein-Schritt-Prognosen der betrachteten Modelle und die Realisationen sowie die Ergebnisse am aktuellen Rand. Die Prognosen mit dem reinen Google-Modell liegen seit dem November systematisch unter der tatsächlichen Entwicklung, wohingegen die Prognosen des DAX-Modells seit Mitte letzten Jahres meistens über den tatsächlichen Werten liegen. Im Mittel heben sich diese Fehler weitgehend auf, mit einer gewissen Tendenz zur Unterschätzung am aktuellen Rand.

Prognosefehler gehören zu allen wissenschaftlichen Verfahren. Im Gegensatz zu reinen Barometermodellen, bei denen kein direkter quantitativer Bezug zu einer Zielvariablen gesucht wird, können wir hier Prognosefehler direkt ablesen. Das ist ein Härtestest, dem sich andere Ansätze entziehen. Der ex post-Erklärungsgrad (vgl. Abbildung 7) für die Google-Daten ist ohnehin sehr gut. Es kommt uns allerdings hier besonders auf das Prognosemuster an: Die Google-Daten haben den fallenden Trend von März bis November 2008 gut getroffen; sie haben den Anstieg der Arbeitslosigkeit bis März 2009 danach abgebildet (mit Ausnahme der durch die Kurzarbeit im Februar gekennzeichneten Fehleinschätzung) und auch den Rückgang danach.

Bei der August-Prognose zeigt sich, dass beide Indikatoren von einer Verringerung der Arbeitslosigkeit ausgehen. Dies steht im Einklang mit dem üblichen saisonalen Muster. Es wäre ein großes Krisenzeichen, würde die Prognose hier ausbrechen. Wenn sie es wie hier nicht tut, dann sollte dies auf eine Ruhepause hindeuten.

Etwas anders gelagert sind die Analysen für September. Hier nimmt die Unsicherheit nochmals zu, da die Google- und DAX-Daten selber zunächst prognostiziert werden müssen. Die Google-Indikatoren signalisieren – wieder im Einklang mit dem Saisonmuster – einen weiteren Rückgang der Arbeitslosigkeit. Allerdings zeigt der DAX deutlich aufwärts, wohingegen der Mittelwert aus beiden Verfahren stagniert. Die Prognose des DAX selbst ist aus den oben genannten Gründen unsicher und sollte nicht überschätzt werden. Auch befindet sich die Börse derzeit im Aufwind, was eine positive Wirkung auf Beschäftigung haben sollte. Letztlich bleibt die Einschätzung, dass auch im September von weiterer Ruhe am Arbeitsmarkt ausgegangen werden kann.

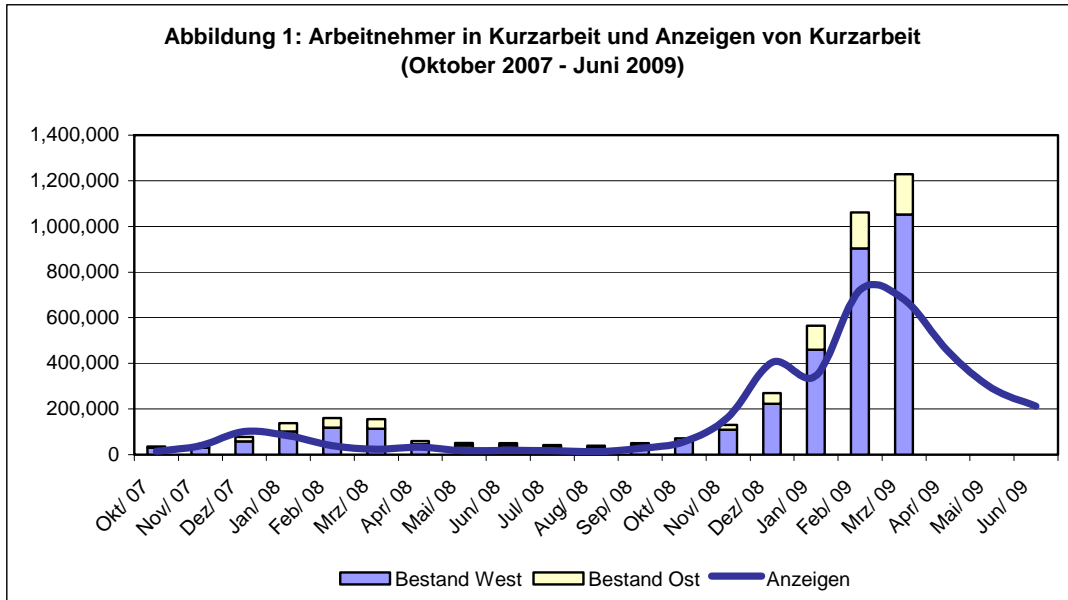
Schlussfolgerungen

Das IZA-Arbeitslosigkeitsbarometer stützt sich auf Suchaktivitätsstatistiken von Google Insight und berücksichtigt somit rasch verfügbare Informationen. Über einen längeren Untersuchungszeitraum von 2005 bis 2009 betrachtet ist die Erklärungskraft der Google-Suchaktivität für die Arbeitslosigkeit beeindruckend – auch im Vergleich zu respektablem anderen Alternativen. Das Modell ist in der Lage, die amtliche Arbeitslosenquote vorherzusagen. Bis einschließlich Januar sind diese Vorhersagen auch eine Periode voraus sehr gut. Erst mit der Einführung der Kurzarbeit und dem Einsatz anderer Arbeitsmarktpolitiken ab Februar 2009 wird das Instrument unsicherer.

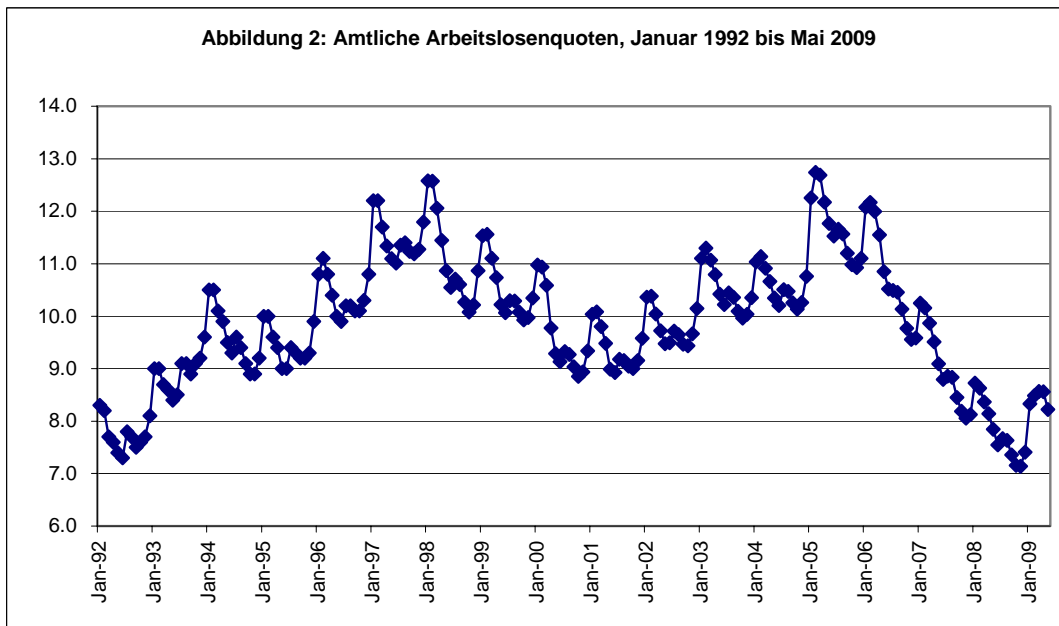
Die zu Beginn dieses Beitrages formulierten Erwartungen hinsichtlich der Entwicklung der Arbeitslosigkeit bis zur Bundestagswahl im September dieses Jahres werden durch Prognosen, die auf Google-Aktivitätsdaten sowie auf der Entwicklung des DAX beruhen, bestätigt. Es finden sich keine Hinweise für den Beginn eines deutlichen oder gar dramatischen Anstieg der Arbeitslosigkeit mit einer baldigen Überschreitung der Vier-Millionen-Grenze im Herbst diesen Jahres, wie er von manchen Beobachtern erwartet wird. Den Schätzungen zufolge kommt es im August zu einem leichten Rückgang der Arbeitslosigkeit. Für den September lassen erste Abschätzungen vermuten, dass sich die Arbeitslosenquote nicht über das August-Niveau bewegt, eher darunter bleibt. Die Frühindikatoren signalisieren nicht, dass die Wirtschaftskrise so stark agiert, dass traditionelle saisonale Muster außer Kraft gesetzt werden. Es bleibt also bei einer leichten saisonbedingten Absenkung der Arbeitslosigkeit in den Monaten August und September. Also: Sommerpause bei der Arbeitslosigkeit.

Literatur

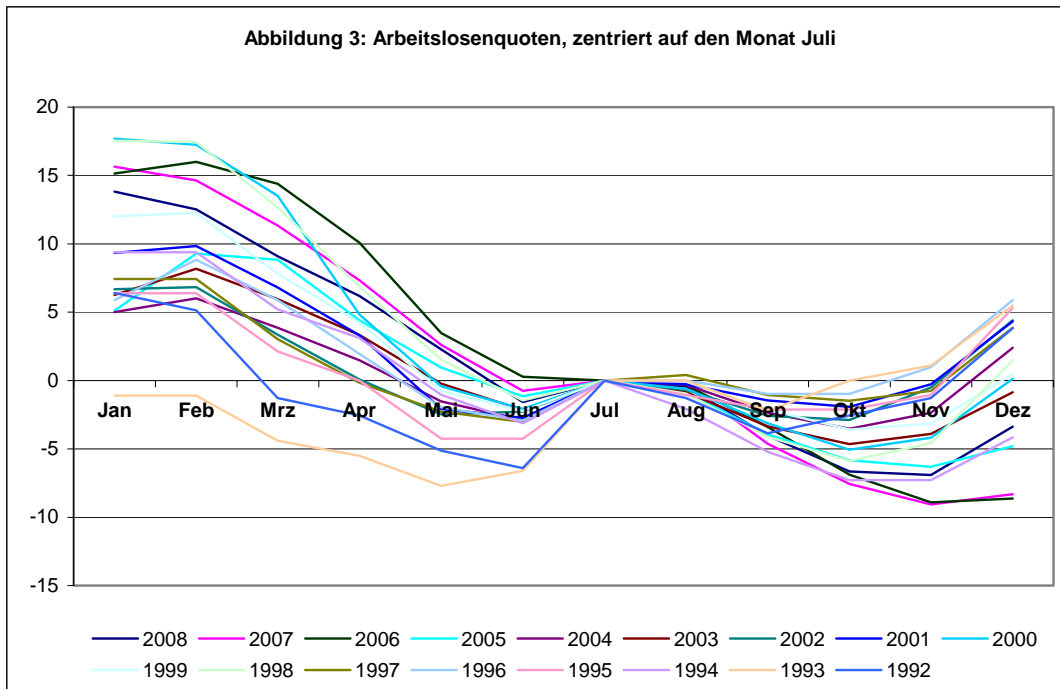
- Askatas, N. und Zimmermann, K. F. (2009a): Google Econometrics and Unemployment Forecasting, *Applied Economics Quarterly*, 55, 107-120.
- Askatas, N. und Zimmermann, K. F. (2009b): Prognosen aus dem Internet: Weitere Erholung am Arbeitsmarkt erwartet, *DIW Wochenbericht* 76, 402-408.
- Askatas, N. und Zimmermann, K. F. (2009c): Googlemetrie und Arbeitsmarkt, *Wirtschaftsdienst* 89, 489-496.
- Brenke, K., Dreger, C., Kooths, S., Kuzin, V., Weber, S. und Zinsmeister, F. (2009): Grundlinien der Wirtschaftsentwicklung 2009/2010, *DIW Wochenbericht* 76, 2-35.
- Caliendo, M., Falk, A., Kaiser, L. C., Schneider, H., Uhlendorff, A., Van den Berg, G. und Zimmermann, K. F. (2009): The IZA Evaluation Dataset. IZA Bonn, mimeo.
- Constant, A. und Zimmermann, K. F. (2008): Im Angesicht der Krise: US-Präsidentenwahlen in transnationaler Sicht, *DIW Wochenbericht* 75, 688-701.
- Dreger, C., Alvarez-Plata, P., Bernoth, K., Brenke, K., Kooths, S., Kuzin, V., Weber, J., Weber, S. und Zinsmeister, F. (2009): Tendenzen der Wirtschaftsentwicklung 2009/2010, *DIW Wochenbericht* 76, 490-529.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. und Brilliant, L. (2009): Detecting Influenza Epidemics using Search Engine Query Data, *Nature* 457, 1012–1014.
- Initiative D21, (N)Onliner Atlas 2009.
- Kholodilin, K. A. und Siliverstovs, B. (2009): Geben Konjunkturprognosen eine gute Orientierung?, *DIW Wochenbericht* 76, 207-213.
- Zimmermann, K. F. (2008a): Schadensbegrenzung oder Kapriolen wie im Finanzsektor?, *Wirtschaftsdienst* 12, 18-20.
- Zimmermann, K. F. (2008b): Weiterbildung in der Krise, *IZA Compact*, Dezember 2008, 12.
- Zimmermann, K. F. (2009): Prognosekrise: Warum weniger manchmal mehr ist, *Wirtschaftsdienst* 2, 86-90.



Datenquelle: Bundesagentur für Arbeit, eigene Berechnungen.
Hinweis: Anzeigen von Kurzarbeit sind Meldungen der Firmen.

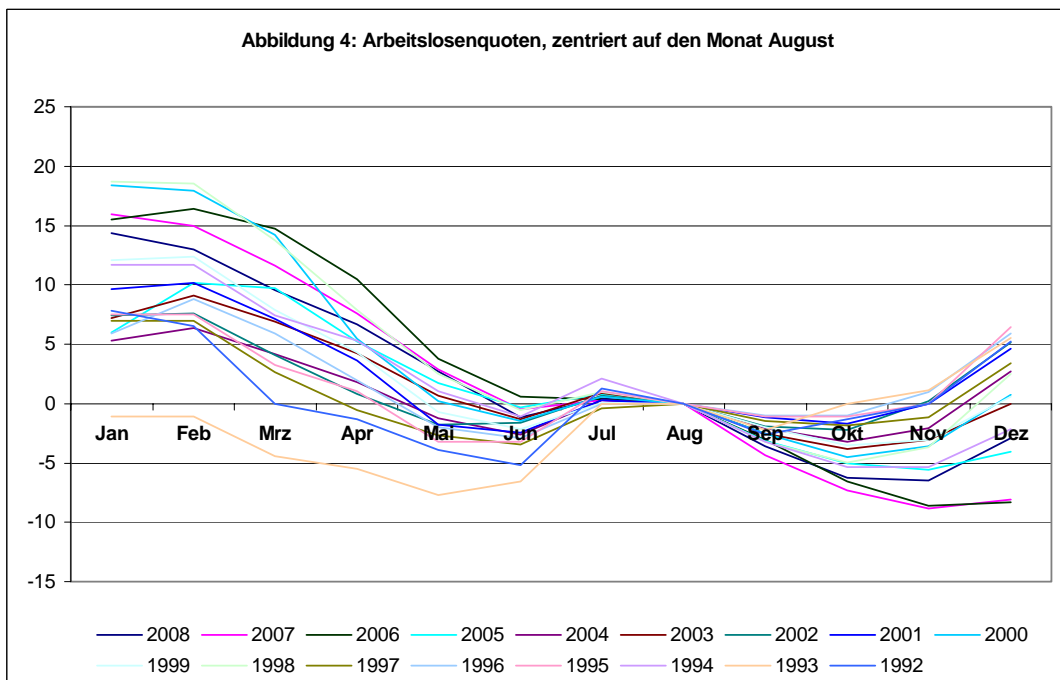


Datenquelle: Bundesagentur für Arbeit.
Hinweis: Amtliche Arbeitslosenquoten in Prozent.



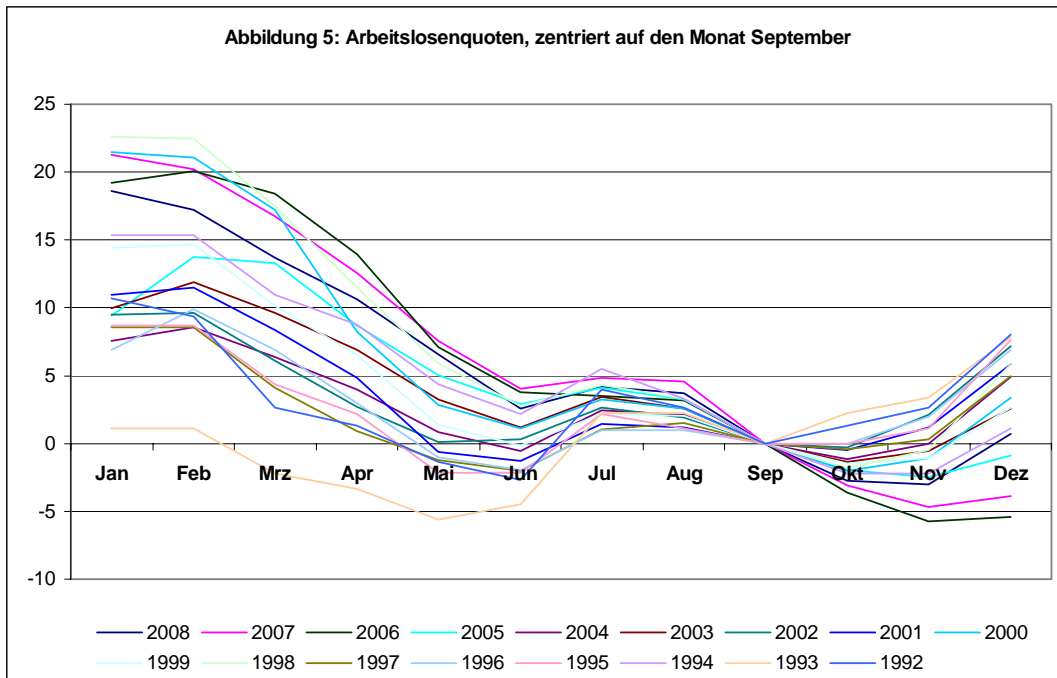
Datenquelle: Bundesagentur für Arbeit, eigene Berechnungen.

Hinweis: Auf den Monat Juli zentrierte Arbeitslosenquoten der Jahre 1992 bis 2008. Abweichungen in Prozent.



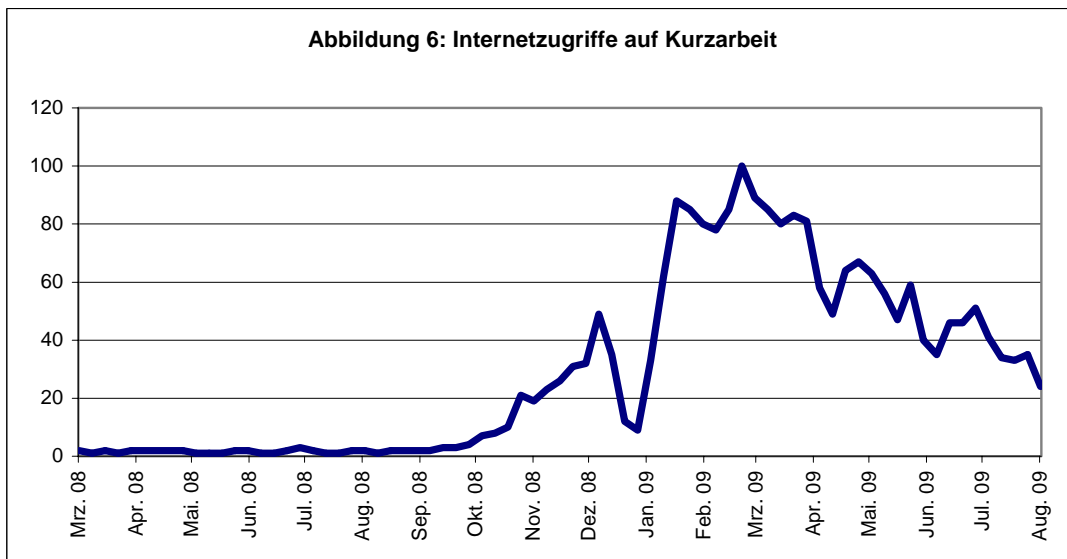
Datenquelle: Bundesagentur für Arbeit, eigene Berechnungen.

Hinweis: Auf den Monat August zentrierte Arbeitslosenquoten der Jahre 1992 bis 2008. Abweichungen in Prozent.



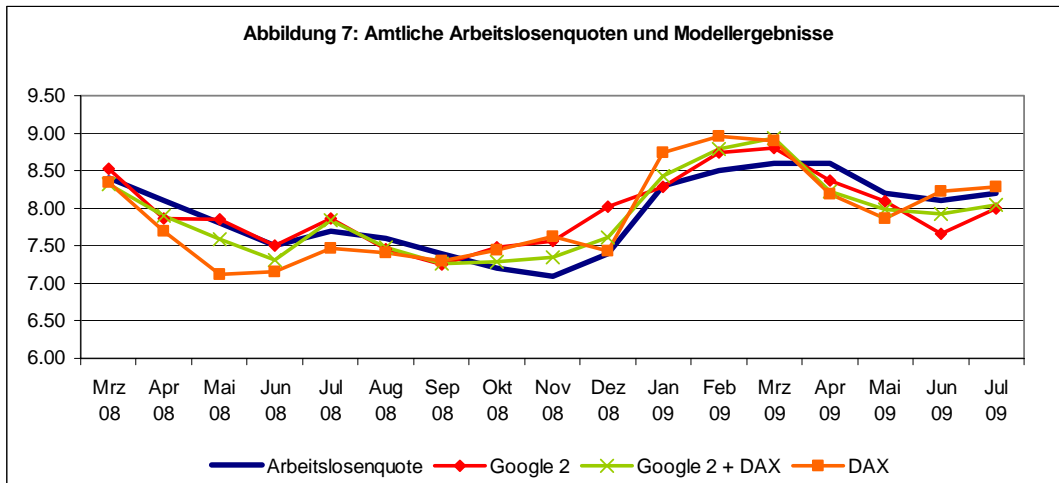
Datenquelle: Bundesagentur für Arbeit, eigene Berechnungen.

Hinweis: Auf den Monat September zentrierte Arbeitslosenquoten der Jahre 1992 bis 2008. Abweichungen in Prozent.

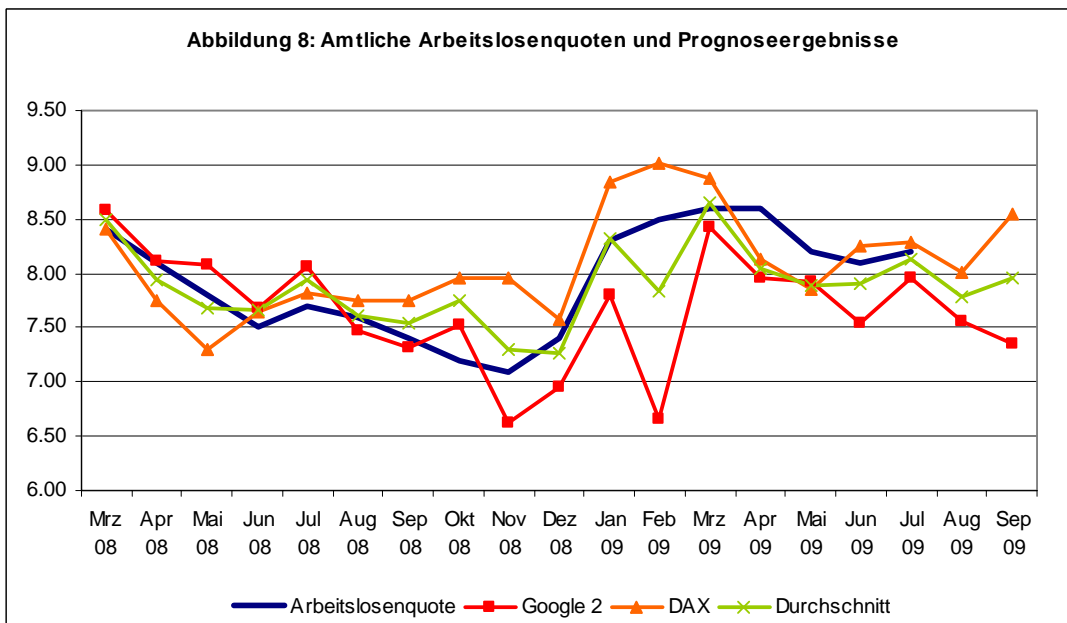


Datenquelle: Google Insights, eigene Berechnungen.

Hinweis: Wöchentliche (Sonntag bis Samstag) Angaben der Suchaktivitäten im Internet.



Datenquelle: Bundesagentur für Arbeit, eigene Berechnungen.



Datenquelle: Bundesagentur für Arbeit, eigene Berechnungen.

Tabelle 1: Auswertung der Regressionsmodelle und Ein-Schritt-Prognosen

<i>Modell</i>	<i>Arbeitsamt</i>	<i>Kurzarbeit</i>	<i>Jobsuche</i>	<i>DAX</i>	<i>R²-a</i>	<i>BIC</i>	<i>MAE</i>
Google 1	L*** + K +		L*** - K* -		0.885	62.138	0.424
Google 2	L*** + K +	L** + K*** +	L*** - K** -		0.939	32.911	0.395
Google 3		L*** + K*** +	L*** - K -		0.925	38.823	0.446
Google 1 + DAX	L*** + K +		L*** - K*** -	L*** - K*** -	0.955	17.175	0.261
Google 2 + DAX	L** + K +	L + K*** +	L*** - K*** -	L*** - K -	0.968	4.292	0.309
Google 3 + DAX		L + K*** +	L*** - K*** -	L*** - K** -	0.957	16.177	0.396
DAX - Prognose				L*** - K*** -	0.906	47.275	0.314
AVG(Google 1,DAX)							0.280
AVG(Google 2,DAX)							0.217
AVG(Google 2,DAX)							0.278

Datenquelle: Arbeitsamt.de, Ifo-Institut und Google Insights. Die benutzte amtliche monatliche Arbeitslosenquote ist saisonunbereinigt. Die genaue Abgrenzung der Schlüsselworte ist: „Arbeitsamt“ = (Arbeitsamt OR Arbeitsagentur); „Arbeitsplatzsuche“ = (Stepstone OR Jobworld OR Jobsout OR Meinestadt OR meine Stadt OR Monster Jobs OR Monster de OR Jobboerse); „Kurzarbeit“ = (Kurzarbeit).

Hinweis: Die Regressionsmodelle nach Gleichung (1) wurden mit Monatsdaten für die Periode Januar 2005 bis Mai 2009 nach der Methode der kleinsten Quadrate geschätzt. K repräsentiert die Veränderung, L die 12. Verzögerung der jeweiligen Variablen. +,- ist das Vorzeichen des geschätzten Koeffizienten und * nennt sein statistisches Signifikanzniveau (* P < 0,05, ** P < 0,01, **** P < 0,001). Die Ein-Schritt-Prognosen wurden für den Zeitraum März 2008 bis Juni 2009 durchgeführt. R2-a ist das korrigierte Bestimmtheitsmaß, BIC das Bayessche Informationskriterium, MAE der mittlere absolute Fehler und AVG die Mittelwerte der entsprechenden Prognosen.