

2 Methodische Grundlagen der ifo Befragungen

ARTEM MARJENKO, STEFAN SAUER, KLAUS WOHLRABE

Aus wirtschaftswissenschaftlicher Sicht bieten die Umfragen des ifo Instituts vor allem aufgrund der Vielzahl an erhobenen Fragestellungen, ihrer hohen Periodizität (monatliche Erhebung der ifo Konjunkturumfragen) sowie des umfassenden Berichtskreises ein enormes Analysepotenzial. So können zum einen zahlreiche Variablen, wie etwa unternehmerische Urteile, Pläne und Erwartungen, für die Konjunkturanalyse genutzt werden. Zum anderen bietet sich ein weites Feld an Möglichkeiten für die empirische Überprüfung von Hypothesen der ökonomischen Theorie sowie für Ex-ante-/Ex-post-Vergleiche auf Makro- und Mikroebene (z.B. Treffsicherheit von Antizipationen oder Verhaltensanalysen).

Anderson und Strigel (1982) charakterisieren die zeitlichen Abläufe von der Konzeption und ersten Durchführung einer Umfrage bis hin zur Entstehung für die Forschung verwertbarer Ergebnisse durch die folgende Unterteilung in verschiedene Phasen, die eine Übersicht über die Nutzung der ifo Erhebungsdaten vermitteln soll. Die Darstellung bezieht sich dabei hauptsächlich auf die über lange Zeiträume regelmäßig erhobenen ifo Konjunktur- sowie Investitionsumfragen:

- **Phase 1:** : Entwicklung und Aufbau der Umfragen. In dieser Phase stehen die Erstellung des Frageprogramms, der Aufbau des Teilnehmerpanels sowie Gewichtungsmethoden im Vordergrund.
- **Phase 2:** Start der Umfrage. Da noch keine längeren Zeitreihen vorliegen, beschränkt sich die Interpretation der Erhebungsergebnisse in der Regel auf einfache Darstellungen der Häufigkeitsverteilungen der eingegangenen Antworten.
- **Phase 3:** Längere Zeitreihen ermöglichen nun Interpretationen und Untersuchungen. Darüber hinaus können erste Treffsicherheitsanalysen der Erwartungen der Befragungsteilnehmer sowie Untersuchungen unternehmerischer Verhaltensweisen durchgeführt werden.
- **Phase 4:** Analysen des saisonalen und zyklischen Verhaltens der Zeitreihen treten stärker in den Vordergrund. Besonderes Interesse gilt der Entwicklung der Zeitreihen an konjunkturellen Wendepunkten.
- **Phase 5:** Konstruktion zusammengesetzter Konjunkturindikatoren und Einbau der Zeitreihen in ökonometrische Prognosemodelle.
- **Phase 6:** Eingehende Studien des Datenmaterials auf Makro- und Mikroebene.

Dieses Kapitel beschäftigt sich zunächst mit der kurzen Erläuterung einiger grundsätzlicher

2 Methodische Grundlagen der ifo Befragungen

konzeptioneller Überlegungen bei der Entwicklung von Befragungen und bezüglich der Datenerhebung sowie mit der Verarbeitung der erhobenen Mikrodaten zu Zeitreihen und aussagekräftigen Indikatoren. Eine ausführliche Beschreibung der verschiedenen Umfragen des ifo Instituts sowie der daraus gewonnenen Informationen und ihrer Relevanz für die Konjunkturanalyse und die Forschung wird anschließend ab Abschnitt 3.1 geliefert.

2.1 Umfragedesigns

Das Design einer Umfrage hängt entscheidend vom Untersuchungsziel ab. Daher müssen vor dem Start einer Umfrage verschiedene grundlegende und konzeptionelle Entscheidungen getroffen werden. Zunächst einmal lässt sich unterscheiden zwischen Fragestellungen, die mit einer einmaligen Befragung untersucht werden sollen, sowie mehrfachem bzw. regelmäßigem Vorlegen der gleichen Fragen bei einem festen Teilnehmerkreis. Solche sogenannten Panelbefragungen sind der Regelfall bei den ifo Befragungen. So wird etwa in den ifo Konjunkturumfragen monatlich ein festes Fragenprogramm verwendet, um bezüglich der verschiedenen abgefragten Variablen Entwicklungen und Veränderungen über die Zeit beobachten und analysieren zu können. Neben den monatlichen Befragungen führt das ifo Institut auch Panelbefragungen mit vierteljährlicher (z.B. Personalleiterbefragung), halbjährlicher (ifo Investitionsumfrage im Verarbeitenden Gewerbe) oder jährlicher (Investitionsumfrage Leasing) Periodizität durch. Einmalige Sonderumfragen zu einem speziellen Thema (sogenannte Querschnitterhebungen) sind dagegen eher der Ausnahmefall, da das monatliche Panel der Unternehmen nicht zu stark belastet werden soll und darf. Stattdessen können beispielsweise im Rahmen der Konjunkturumfragen oder anderer Befragungen einmalige Sonderfragen zu aktuellen wirtschaftlichen Themen gestellt werden (vgl. Abschnitt 3.8).

Während der große Vorteil von Panelbefragungen in der regelmäßigen Abfrage zahlreicher Variablen besteht, so dass sowohl auf der Individualebene der befragten Unternehmen als auch auf aggregierter Ebene Entwicklungen beobachtet werden können, sind der Aufbau und die Aufrechterhaltung der Unternehmenspanels allerdings auch mit einem großen Aufwand verbunden. Da es das Ziel ist, anhand der Erhebungsergebnisse möglichst aussagekräftige und robuste Aussagen über Entwicklungen in der deutschen Wirtschaft zu treffen, muss das Panel der Umfrageteilnehmer ein möglichst gutes Abbild der deutschen Unternehmen darstellen. Vollerhebungen aller deutschen Unternehmen sind selbstverständlich nicht praktikabel, zumal die Teilnahme an den ifo Umfragen auf freiwilliger Basis geschieht und es kein öffentlich zugängliches Unternehmenspanel gibt, welches dafür benutzt werden könnte. Daher handelt es sich bei den Unternehmenspanels um (geschichtete) Stichproben aus allen deutschen Unternehmen; lediglich die jährlich durchgeführte Investitionserhebung Leasing ist als Vollerhebung bei allen deutschen Leasinggesellschaften konzipiert. Die Stichprobengröße muss einen gewissen Umfang haben, um robuste Ergebnisse zu erzeugen. Dabei ist die Abdeckung sowohl aller wichtigen Wirtschaftsbereiche (fachliche Repräsentation) als auch eines ausreichend großen Anteils der Unternehmen (Firmenrepräsentation) wichtig. In die

monatlichen Konjunkturumfragen gehen derzeit etwa 9.000 Unternehmensmeldungen ein, der Aufbau eines weitgehend stabilen Panels dauert in der Regel aber einige Jahre. Darüber hinaus müssen sogenannte Selektionseffekte vermieden werden, bei denen etwa selektive Neuwerbungsmaßnahmen zu systematischen Verzerrungen der Ergebnisse führen. Daher ist bei der Zusammenstellung des Panels darauf zu achten, dass die regionale Zusammensetzung genauso wie die Verteilung der Unternehmen nach Firmengröße und Wirtschaftszweigen in etwa die Realität widerspiegelt. Eine optimale oder Mindestgröße des Panels ist schwer festzulegen, da im klassischen theoretischen Sinne keine Stichprobe gezogen werden kann, aufgrund des fehlenden zugänglichen Unternehmensregisters. Experimente mit unterschiedlichen Teilnehmerzahlen haben ergeben, dass der Gesamtindex mit nahezu identischem Verlauf zum Originalindex auch mit etwas weniger als 1.000 Unternehmen erreicht werden kann. Eine höhere Teilnehmerzahl hat allerdings den Vorteil, dass detaillierte und aussagekräftige Zeitreihen für viele - auch kleinere - Branchen in den verschiedenen Sektoren ermittelt werden können.

Im Zeitverlauf kommt es auch regelmäßig vor, dass Unternehmen aus verschiedensten Gründen über einen längeren Zeitraum oder überhaupt nicht mehr aktiv an der Umfrage teilnehmen. Um diesem sogenannten Nonresponse bzw. der Panelmortalität entgegenzuwirken, ist eine ständige Panelpflege mittels Neuwerbung von Firmen zum Stabilhalten oder Ausbau des Panels auch im Hinblick auf das Vermeiden von systematischen Verzerrungen vonnöten. Verschiedene Studien, wie unter anderem Seiler (2010) und Seiler und Heumann (2013), zeigen, dass es bei den ifo Konjunkturumfragen keine erkennbaren verzerrenden Effekte durch Nonresponse von Unternehmen gibt.

Der zweite wichtige Block bei der Konzeption einer Umfrage ist die Erstellung des Frageprogramms. Dabei gilt es zuerst, Themen und Variablen auszuwählen, die sich an den der Umfrage zugrunde liegenden Forschungszielen orientieren, und daraus Fragen zu konstruieren, die interessante und aussagekräftige Ergebnisse versprechen. So werden etwa für die ifo Konjunkturumfragen Variablen abgefragt, die für die Konjunkturbeobachtung im Hinblick auf aktuelle Entwicklungen sowie Pläne und Erwartungen für die Zukunft von Bedeutung sind. Da es auf diesem Feld eine große Anzahl von interessierenden Themen gibt, der Bearbeitungsaufwand bei Unternehmensbefragungen jedoch nicht unnötig hoch sein sollte, um das Teilnehmerpanel nicht zu sehr zu strapazieren, gilt es, die Fragen nach ihrer Wichtigkeit und Dringlichkeit zu gruppieren. Weniger wichtige Fragen sollten daher entweder ganz gestrichen oder zumindest in geringerer Frequenz erhoben werden. Der Fragenkatalog der Konjunkturumfragen beispielsweise wurde deshalb in einen Standardfragenteil mit monatlich wiederkehrenden Fragen sowie lediglich quartalsweise bzw. ein oder zweimal pro Jahr gestellte Fragen eingeteilt. Diese Trennung richtet sich sowohl nach der Bedeutung der Variablen und der Dringlichkeit laufender aktueller Information als auch nach dem Bearbeitungsaufwand für die Teilnehmer. Ein einmal etabliertes Frageprogramm sollte möglichst nur noch selten geändert werden. Bereits kleinere Änderungen im Wortlaut einer Frage können zu einem Strukturbruch in der Zeitreihe führen, weil sich das Antwortverhalten bei den Unternehmen entsprechend verändern könnte.

2 Methodische Grundlagen der ifo Befragungen

Auch das Einstellen, Austauschen und Neuaufnehmen von Fragen auf den Fragebögen sollte gut überlegt sein. Wenn Fragen eingestellt werden, kann die entsprechende Zeitreihe nicht mehr berechnet werden. Daher sollte zuvor stets genau evaluiert werden, ob und von wie vielen Nutzern die aus der Frage gewonnenen Zeitreihen und Indikatoren verwendet werden. Bei neuen Fragen muss immer im Blick bleiben, dass es nicht zu viele werden. Zum anderen dauert es in der Regel einige Jahre, ehe die aggregierten Antworten für eingehendere konjunkturelle Analysen genutzt werden können, da zuerst eine Saisonbereinigung erfolgen muss. Diese ist jedoch erst möglich, wenn genügend Beobachtungen vorliegen.

Des Weiteren sind auch einige wichtige Regeln bezüglich der Gestaltung und Formulierung der Fragen zu beachten. Sie beinhalten zum einen die Verwendung eines einfachen Satzbaus sowie verständlicher und gebräuchlicher Wörter. Zu vermeiden sind dabei unter anderem Doppeldeutigkeiten oder unverständliche Konstrukte, wie etwa doppelte Verneinungen. Zum anderen sollten die Fragen möglichst neutral formuliert werden, um die Teilnehmer nicht durch einseitige oder wertende Fragestellungen ungewollt zu beeinflussen oder gar zu einer Antwortoption zu drängen. Um den Bearbeitungsaufwand der Fragen möglichst gering zu halten, werden den Befragungsteilnehmern wiederum in der Regel Antwortmöglichkeiten vorgegeben. Die Umfragen des ifo Instituts setzen zum allergrößten Teil auf qualitative Fragestellungen mit drei oder vier Antwortmöglichkeiten statt auf die quantitative Abfrage von Zahlen. Die vorgegebenen Antworten müssen dabei alle Möglichkeiten abdecken und sich gegenseitig ausschließen.

Vor der erstmaligen Durchführung einer Umfrage können Pre-Tests mit ausgewählten Firmen hilfreich sein, um eventuell noch bestehende Probleme im Fragebogen, wie etwa unverständliche Fragestellungen, aufzudecken.

Tiefergehende Literatur zum Thema Design von Umfragen und zum Ziehen von repräsentativen Stichproben bieten unter anderem Schumann (2006) sowie Engel u. a. (2012). Zum Design der verschiedenen Umfragen des ifo Instituts sowie zu den jeweiligen Fragenkatalogen liefern die folgenden Kapitel detaillierte Einblicke.

2.2 Befragungskanäle

Grundsätzlich kann die Erhebung von Daten in Befragungen auf verschiedenen Wegen erfolgen: Die gängigsten Methoden sind dabei die Zusendung eines Papierfragebogens an den Umfrageteilnehmer per Post oder Fax, ein E-Mail-Fragebogen zum Ausdrucken, ein Online-Fragebogen (Beantwortung online, Antworten liegen in elektronischer Form vor), sowie Telefon- oder „Face-to-Face“- Interviews. In der Praxis der Unternehmensbefragungen kommen davon in der Regel ausschließlich Papier-, Online- sowie per E-Mail zugesandte und zum Ausdruck gedachte Fragebögen zum Einsatz. Sowohl bei den ifo Konjunkturumfragen als auch bei entsprechenden Unternehmensbefragungen in den anderen EU-Ländern wird überwiegend auf diese drei Befragungskanäle zurückgegriffen (lediglich in Großbritannien, Italien und Spanien werden

zusätzlich auch Telefon-Interviews durchgeführt). Die Charakteristika dieser drei Methoden sind in Tabelle 2.1 zusammengefasst.

Tabelle 2.1: Charakteristika verschiedener Befragungskanäle

	Papier-Fragebogen	Fragebogen per E-Mail zum Ausdrucken	Online-Fragebogen
Kosten	Kosten für Druck und Versand sowie für die Erfassung der Daten	Kosten für Druck und für die Erfassung der Daten	Kosten für Hardware, Programmierung, Softwarelizenzen und Wartung
Bereitstellung von Umfrageergebnissen für die Teilnehmer	Nur eine beschränkte Anzahl an Ergebnisberichten wird per Post/Fax zugesandt	Zusendung mehrerer Ergebnisberichte im E-Mail-Anhang möglich	Alle Ergebnisberichte im Online-Portal zugänglich
Platz für Sonderfragen auf den Fragebögen	Beschränkt	Beschränkt	Unbeschränkt
Plausibilität der Antworten	Prüfung der Angaben auf Plausibilität erforderlich	Prüfung der Angaben auf Plausibilität erforderlich	Nicht plausible Angaben können bereits bei der Eingabe verhindert werden
Sicherheitsaspekte	Telefaxsendungen sind nicht verschlüsselt	Telefaxrücksendungen sind nicht verschlüsselt	Übertragung der Daten über eine verschlüsselte Verbindung, passwortgeschützter Zugang zum Online-Portal

2 Methodische Grundlagen der ifo Befragungen

Wegen der Vielzahl an Vorteilen, die eine Online-Erhebung mit sich bringt, ist das ifo Institut laufend bemüht, möglichst viele Teilnehmer in dieses Befragungsverfahren zu bringen. Dies hatte in den vergangenen Jahren bereits eine spürbare Erhöhung des Anteils der online teilnehmenden Firmen in den Konjunkturumfragen zur Folge. Die in Tabelle 2.2 dargestellte Verteilung der Teilnahmearten (Durchschnitt des Jahres 2017) zeigt jedoch, dass bei einer großen Zahl von Unternehmen nach wie vor eine Präferenz für die papiergestützte Umfrage besteht. Viele empfinden den Umgang mit dem Papierfragebogen als bequemer, mitunter werden auch Bedenken hinsichtlich der Datensicherheit im Online-Verfahren geäußert. Obwohl dieses Risiko in der Realität als sehr gering eingestuft werden kann, bietet das ifo Institut deshalb auch weiterhin die Möglichkeit der Teilnahme per Papierfragebogen an, um der Präferenz vieler Antwortender Rechnung zu tragen. Lediglich vereinzelte Umfragen, wie etwa die Konjunkturumfrage im Versicherungsgewerbe, werden bereits ausschließlich online durchgeführt. Mit der Fortentwicklung der Technik ergeben sich auch immer wieder neue Möglichkeiten für die Durchführung von Umfragen. So könnte etwa in näherer Zukunft auch die Datenerhebung mit Hilfe einer Umfrage-App für mobile Endgeräte eine denkbare Option für das ifo Institut darstellen.

Tabelle 2.2: Anteil der Befragungskanäle in den ifo Konjunkturumfragen

	Papier-Fragebogen	Per E-Mail zugesandter, zum Ausdruck gedachter Fragebogen	Online-Fragebogen
Verarbeitendes Gewerbe	31%	-	69%
Handel	41%	-	59%
Bauhauptgewerbe	56%	-	44%
Dienstleistungen	17%	14%	69%

2.3 Aggregation der Mikrodaten zu Zeitreihen

Im Rahmen der regelmäßigen Konjunkturumfragen werden die Mikrodaten, d.h. die Antworten der Unternehmen, zu Zeitreihen aggregiert, welche dann veröffentlicht werden. Erst dann ist eine Interpretation und Nutzung im Rahmen der Konjunkturanalyse sinnvoll. Die Mikrodaten werden vor allem im Rahmen der Forschung genutzt (vgl. hierzu Kapitel 8). Eines der Grundprinzipien der Aggregation in den ifo Umfragen erfordert, dass die Antworten zunächst ein individuelles Gewicht bekommen. Dieses repräsentiert die Größe des Unternehmens. In den ifo Umfragen bekommen größere Firmen ein höheres Gewicht als kleinere. In einem zweiten Schritt werden Branchen- und Sektorgewichte vergeben. Branchen, die einen größeren Beitrag zur Wertschöpfung in der deutschen Wirtschaft leisten, erhalten entsprechend ein höheres Gewicht.

Wie bereits erwähnt, basiert der Großteil der regelmäßig gestellten Fragen auf einer Dreierskala. Es gibt unterschiedliche Verfahren, diese Art von Daten zu aggregieren (vgl. auch Pesaran und Weale (2005)). Das ifo Institut nutzt seit der Einführung der Umfragen im Jahr 1949 die Saldenmethodik, welche auf Anderson (1952b) zurückgeht. Dabei wird der relative Anteil der negativen Antworten (oder die dritte Kategorie) von dem relativen Anteil der positiven (oder der ersten Kategorie) abgezogen. Die neutrale oder zentrale Kategorie ist somit nur indirekt berücksichtigt. Haben zum Beispiel 30% der Unternehmen positive Geschäftserwartungen, 50% unverändert (eine neutrale) und 20% der Unternehmen negative angegeben, so beträgt der Saldo in diesem Beispiel +10. Hierfür wurde der relative Anteil von 20% (negativ) von dem positiven Anteil von 30% abgezogen. Der Anteil der neutralen Antworten von 50% bleibt unberücksichtigt. Der Saldo kann dementsprechend zwischen -100 und +100 liegen. Im ersten Fall sind alle Antworten negativ und im zweiten alle positiv. Fällt der Saldo positiv aus, so hat eine Mehrheit der Unternehmen die entsprechende Kategorie gewählt. Genau genommen gilt diese Mehrheit nur für die Unternehmen, welche nicht die neutrale Kategorie gewählt haben. In dem obigen Beispiel wird die neutrale Kategorie am häufigsten gewählt. Bei der Interpretation des Saldos ist zu beachten, dass ein bestimmter Saldo durch unterschiedliches Antwortverhalten entstehen kann. So kann ein Saldo von +10 z.B. durch eine Antwortverteilung 10% (+), 90% (=) und 0% (-) entstehen. Eine andere mögliche Verteilung ist 50%, 10% und 40%. Diese Liste ließe sich beliebig fortsetzen. Das Ausmaß der Streuung der Antworten ist auch ein Maß der Unsicherheit der Unternehmen mit Blick auf eine bestimmte Konjunkturvariable (vgl. Abschnitt 9.6). Die Saldenmethodik kann auch für gewichtete Antworten verwendet werden. Dies ist exemplarisch in Tabelle 2.3 dargestellt. Dabei werden für die Berechnung der relativen Anteile entsprechend die individuellen Gewichte der Unternehmen verwendet. In dem Beispiel errechnet sich der ungewichtete Saldo als $50 - 16,67 = 33,33$, während der gewichtete Saldo 30 beträgt.

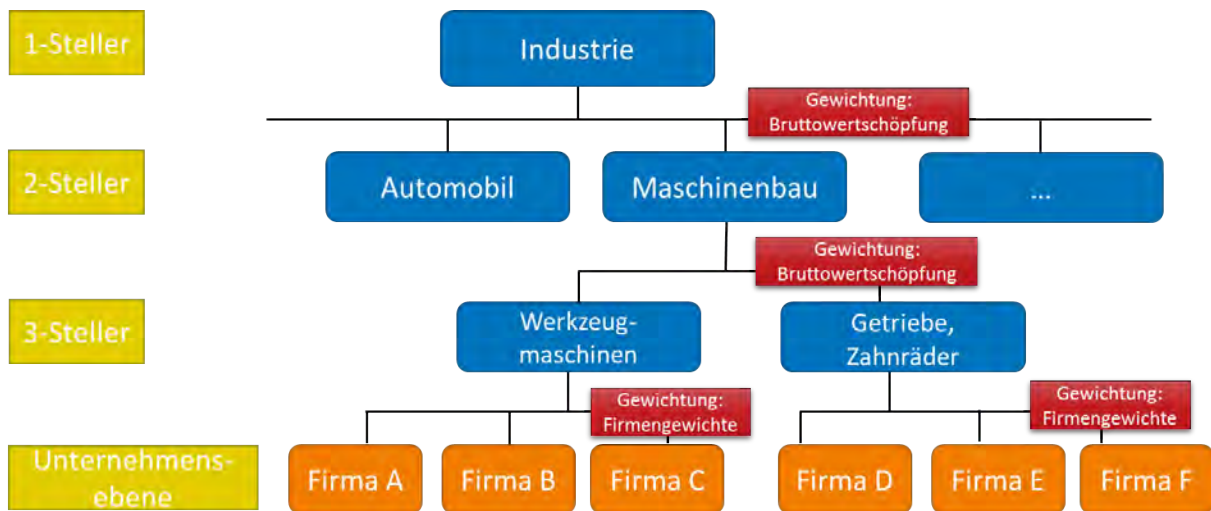
Tabelle 2.3: Beispiellaggregation

Betrieb	Antwort	Gewicht	Antwortverteilung		
			+	=	-
A	befriedigend	5		5	
B	gut	2	2		
C	gut	4	4		
D	schlecht	2			2
E	befriedigend	5		5	
F	gut	2	2		
Verteilung ungewichtet			3	2	1
Summe der Gewichte		20	8	10	2
Prozentanteile (ungewichtet)		100%	50%	33,33%	16,67%
Prozentanteile (gewichtet)		100%	40%	50%	10%

2 Methodische Grundlagen der ifo Befragungen

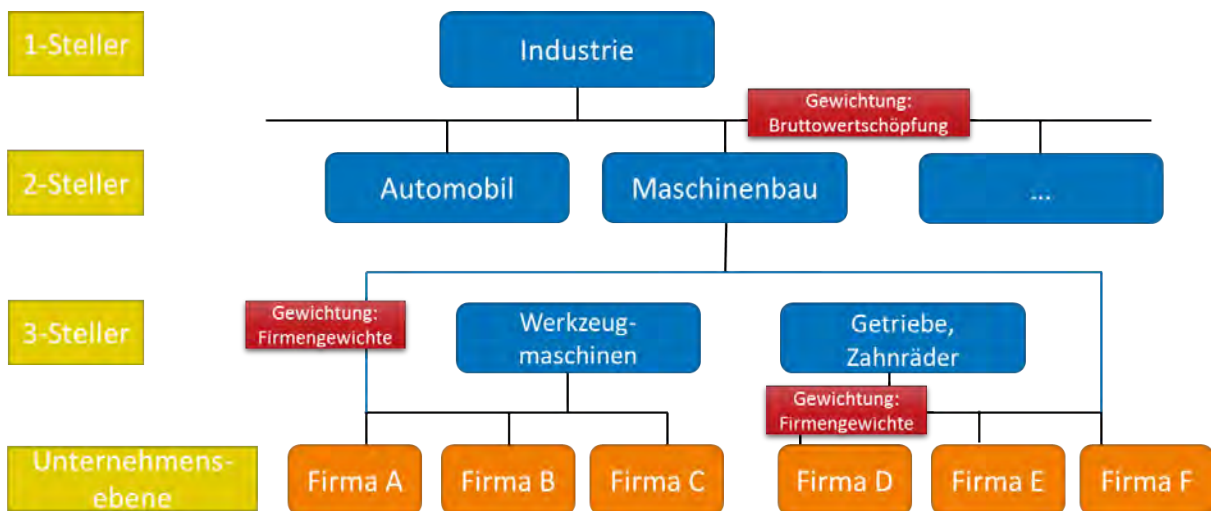
Die Aggregation im Rahmen der Saldenmethodik erfolgte bis März 2018 in einer Art Baumstruktur, die sich nach der offiziellen Wirtschaftszweigklassifikation WZ 2008 des Statistischen Bundesamtes richtet. Ein Ausschnitt dieser Baumstruktur ist exemplarisch in Abbildung 2.1 für das Verarbeitende Gewerbe bis zur dritten Ebene dargestellt. Die oberste Stufe, hier das gesamte Verarbeitende Gewerbe, ist der sogenannte 1-Steller. Auf der 2-Steller-Ebene sind dann die bekannten Industriezweige verzeichnet, wie die Automobilbranche oder der Maschinenbau. Unterhalb dieser Ebene weist die Gliederung unterschiedliche Aggregationstiefen auf, die sich zum Teil bis auf die 6-Steller-Ebene erstrecken. Die Aggregation der Antworten innerhalb dieser Baumstruktur erfolgt dabei von unten nach oben. Jede Firmenmeldung wird auf der untersten Ebene der Hierarchie einem Wirtschaftszweig zugeordnet und bekommt dabei ein individuelles Gewicht zugewiesen. Im Verarbeitenden Gewerbe wird dieses anhand von Mitarbeiterzahlen bestimmt. Anschließend werden die Salden gemäß ihrem Bruttowertschöpfungsanteil zur nächsthöheren Ebene aggregiert. Im Beispiel in Abbildung 2.1 werden etwa die Ergebnisse des Werkzeugmaschinenbaus sowie der Hersteller von Getrieben und Zahnrädern (plus die anderen Unterbereiche des Maschinenbaus) zum Saldo für den Maschinenbau zusammengefasst. Zum Schluss werden alle Salden auf 2-Steller-Ebene ebenfalls mit einer sich an den Bruttowertschöpfungsanteilen orientierenden Gewichtung zum gesamten Verarbeitenden Gewerbe aggregiert. So intuitiv die Aggregation gemäß der Baumstruktur in Abbildung 2.1 auch ist, hat sie doch einige praktische Nachteile. Das Unternehmenspanel in den ifo Konjunkturumfragen bleibt über die Zeit nicht konstant. Unternehmen scheiden aus oder werden neu aufgenommen. Dies hat auch Auswirkungen auf die Besetzung der einzelnen Wirtschaftszweige. Zum einen können in einigen Bereichen im Zeitablauf zu wenige Unternehmen enthalten sein. Der entsprechende Saldo würde dann die wirtschaftliche Entwicklung womöglich nicht mehr akkurat darstellen. So sind Salden auf Basis weniger Antworten volatil als diejenigen mit vielen Antworten. Zum anderen ist es möglich, dass die Saldenwerte eher hohe (extreme) Saldenwerte wie z.B. + 100 annehmen. Es bestünde somit die Möglichkeit, dass dadurch auch die oberen Aggregate verzerrt werden. Ein weiterer wichtiger Aspekt ist der Umstand, dass manche Antworten gar nicht erst berücksichtigt werden können, weil nicht genügend Unternehmen vorhanden sind, um ein eigenes (Unter-)Aggregat zu bilden. Aufgrund der genannten Punkte müssen die Aggregationsvorschriften permanent gepflegt und angepasst werden. Um die Problematik etwas abzumildern, wurde im Jahr 2018 eine Änderung in der Zuordnung der Mikrodaten zur Saldenberechnung vorgenommen. Abbildung 2.2 stellt diese beispielhaft vor. Der wichtigste Unterschied ist, dass zur Berechnung der Salden auf der 2-Steller-Ebene keine Salden von der 3-Steller-Ebene mehr verwendet werden. Das heißt, alle Mikrodaten, die dem Maschinenbau zugeordnet sind, gehen direkt mit ihrem Firmengewicht in die Saldenberechnung des Maschinenbaus ein. Diese Idee pflanzt sich bis in die unteren Ebenen fort. Alle Mikrodaten etwa aus dem Bereich „Werkzeugmaschinen“, unabhängig von weiteren darunter liegenden 4- oder 5-Steller-Aggregaten, werden entsprechend verwendet. Diese Vorgehensweise hat den Vorteil, dass die Pflege des Aggregationsschemas einfacher wird und mögliche Verzerrungen aus zu gering besetzten Unteraggregaten minimiert werden. Zusätzlich können alle Antworten jederzeit verwendet werden, da sie durchweg einem 2-Steller-Aggregat zugeordnet werden können.

Abbildung 2.1: Aggregation bis März 2018



Quelle: Darstellung des ifo Instituts

Abbildung 2.2: Aggregation ab April 2018



Quelle: Darstellung des ifo Instituts

Die Zeitreihen, die mit Hilfe der Saldenmethodik auf Basis der Mikrodaten berechnet wurden, schwanken zwischen den theoretischen Werten -100 und +100 und liegen auch zur Analyse für

2 Methodische Grundlagen der ifo Befragungen

die Öffentlichkeit in dieser Form vor. Jede Zeitreihe kann jedoch auch in einen Index umgerechnet werden, wie es beim ifo Geschäftsklimaindex Deutschland geschieht (vgl. Abschnitt 4.1). Dabei wird ein Basisjahr gewählt und der entsprechende Durchschnitt der Zeitreihe für dieses Jahr. Dieser wird dann gleich 100 gesetzt. Alle Beobachtungen werden dann mit Hilfe eines Dreisatzes relativ zu diesem Wert umgerechnet. Eine indexierte Zeitreihe schwankt somit um den Wert 100. Die Indexberechnung folgt einer Tradition in der Wirtschaftspresse, wo die Indexierung zur besseren Vergleichbarkeit von Zeitreihen mit unterschiedlichen Skalen angewendet wird.

Die Indexierung ist jedoch nicht ohne Nachteil. Es fehlt ein natürlicher Referenzwert. Bei den Salden ist dies die Nulllinie. Ein positiver Saldo entspricht einer positiven Konjunktur-entwicklung und vice versa. Bei einer indexierten Reihe wird manchmal fälschlicherweise angenommen, dass der Referenzwert für eine positive bzw. negative Entwicklung 100 sei. Dies gilt jedoch nur in einem bestimmten Ausnahmefall, nämlich wenn im Basisjahr der Mittelwert 0 war. Der Referenzwert ist mit dem Mittelwert des Basisjahres gegeben, welcher aber nur schlecht aus einer grafischen Aufbereitung abzulesen ist. Deshalb wird empfohlen, vor allem die Veränderung des ifo Geschäftsklimaindex zu betrachten und diesen mit seinem historischen Mittelwert zu vergleichen.

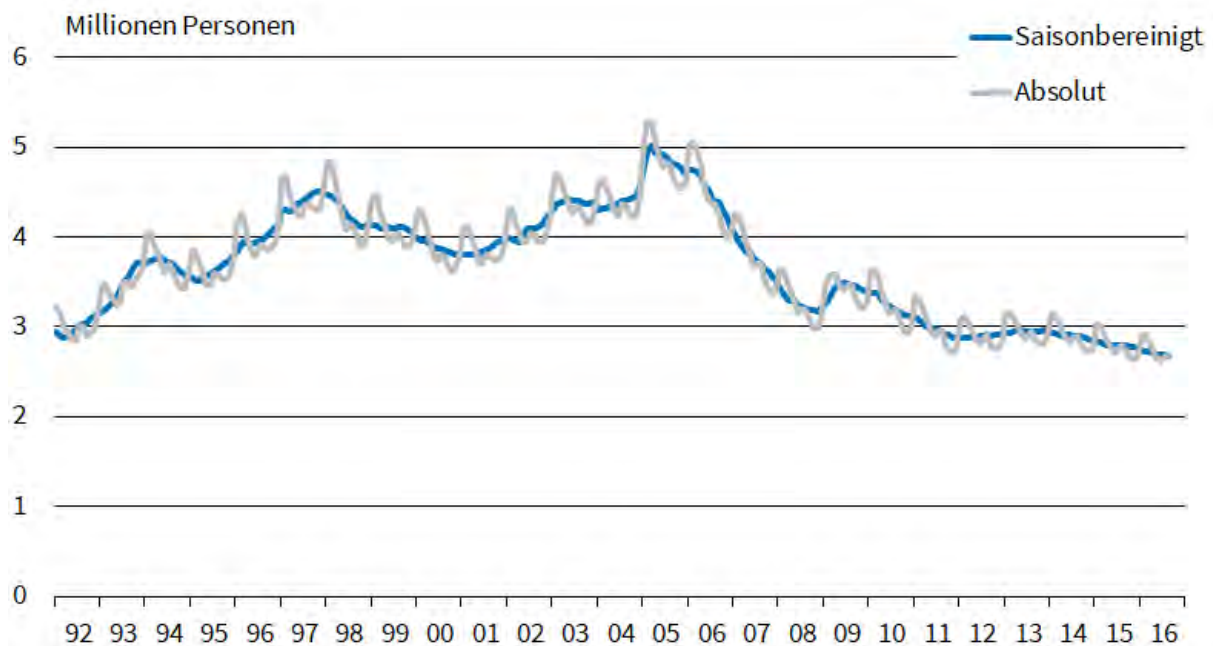
Ein weiterer Nachteil der Indexierung ist, dass die Niveaus von verschiedenen Indizes nicht miteinander verglichen werden können, weil sich in der Regel ihre Mittelwerte im Basisjahr unterscheiden. Ist z.B. der Index für den Großhandel größer als der entsprechende Index für den Einzelhandel, könnte gemeinhin angenommen werden, dass es dem Großhandel wirtschaftlich besser geht. Diese Interpretation gilt aber nur, wenn der Abstand der Mittelwerte aus dem Basisjahr kleiner ist als der Niveauunterschied bei den Indizes. Es wird deshalb beim Vergleich verschiedener Zeitreihen aus den ifo Konjunkturumfragen die Betrachtung der entsprechenden Salden empfohlen.

2.4 Saisonbereinigungsverfahren

2.4.1 Einführung und grundlegende Modellansätze

In zahlreichen ökonomischen Statistiken und Zeitreihen sind innerhalb eines Jahres regelmäßig wiederkehrende Muster sichtbar, die die interessierenden trendmäßigen und konjunkturbedingten Strukturen überlagern. So schwankt etwa die Arbeitslosenquote beträchtlich im Jahresverlauf, mit ihrem Höchststand in den Wintermonaten (Abbildung 2.3), oder steigen die Einzelhandelsumsätze in jedem Jahr beträchtlich in der Vorweihnachtszeit. Ursachen für derartige saisonale Veränderungen sind zumeist in jahreszeitlichen oder institutionellen Einflüssen, wie etwa Lufttemperaturen, Feiertagen oder Ferien, zu finden. Um falsche Rückschlüsse zu vermeiden, können in Zeitreihen Bewegungen, die sich im Jahresrhythmus wiederholen, durch eine Saisonbereinigung ausgeschaltet werden.

Abbildung 2.3: Arbeitslosenzahlen in Deutschland



Quelle: Bundesagentur für Arbeit.

© ifo Institut

Zusammenfassende Überblicke über die historische Entwicklung von Saisonbereinigungsverfahren und deren Methodik finden sich unter anderem bei Hylleberg (1992) sowie bei Edel u. a. (1997). Die grundlegende Idee ist es, die Ursprungszeitreihe (U) in systematische Komponenten und einen irregulären Rest zu zerlegen und anschließend den Saisoneffekt herauszurechnen. Jede dieser Komponenten beinhaltet unterschiedliche Einflüsse auf die Zeitreihe (vgl. Tabelle 2.4). Bei den systematischen Komponenten in den Ursprungswerten der Zeitreihen handelt es sich üblicherweise um eine Trendkomponente (T), die auf den langfristigen Entwicklungspfad der Reihe zurückzuführen ist, und eine Konjunkturkomponente (K), in der zyklische Bewegungen mit einer Periodenlänge von mehreren Jahren zum Ausdruck kommen. Diese beiden Komponenten werden zumeist zur sogenannten glatten Trend-Konjunktur-Komponente (TK) zusammengefasst. Daneben zählt auch die Saisonkomponente (S), die sich im Jahresrhythmus wiederholt, zu den systematischen Einflüssen. Optional kann das Komponentenmodell zusätzlich um eine Witterungskomponente (W), die außergewöhnliche Wetterverhältnisse oder die unterschiedliche Intensität von Schnee und Frost im Winter erklärt, und/oder um eine Arbeitstagekomponente (A), die die Anzahl der Werktage eines Monats beziehungsweise Quartals berücksichtigt, erweitert werden. Im irregulären Rest (I) sind alle zuvor nicht erfassten, vor allem kurzfristigen und zufälligen Einflüsse auf die Zeitreihe enthalten, wie zum Beispiel Auswirkungen von Streiks sowie Kurzfristreaktionen auf wirtschaftspolitische Maßnahmen oder sonstige wichtige Ereignisse.

Die einzelnen Komponenten können auf unterschiedliche Arten zueinander in Beziehung

2 Methodische Grundlagen der ifo Befragungen

Tabelle 2.4: Komponentenzerlegung einer Zeitreihe

Komponente	Ursache
Glatte Trend-Zyklus Komponente	Langfristiger Trend und konjunkturelle Bewegungen
Saisonkomponente	Jährlich wiederkehrende Schwankungen
Witterungskomponente (optional)	Außergewöhnliche Wettereinflüsse
Arbeitstagekomponente (optional)	Kalendereffekte
Irregulärer Rest	Kurzfristige und zufällige Einflüsse

gesetzt werden. Als gängigste Modellansätze haben sich in der Praxis die additive Zeitreihenzerlegung, bei der sich die Komponenten zu jedem Zeitpunkt t zum Ursprungsreihenwert U_t aufsummieren, sowie die multiplikative Zerlegung herauskristallisiert:

$$\begin{aligned} \text{Additive Zerlegung} : U_t &= TK_t + S_t + (W_t) + (A_t) + I_t \\ \text{Multiplikative Zerlegung} : U_t &= TK_t \cdot S_t \cdot (W_t) \cdot (A_t) \cdot I_t \end{aligned}$$

Eine saisonbereinigte Reihe enthält dann nach dem Herausrechnen der Saison sowie gegebenenfalls von Witterungs- und/oder Arbeitstageeffekten nur noch die glatte Trend-Konjunkturkomponente und die irreguläre Komponente. Dies geschieht im additiven Modell durch Subtraktion der Saisonkomponente und im multiplikativen Ansatz entsprechend durch Division.

Da es nicht möglich ist, die genauen Ausprägungen der einzelnen Komponenten einer Zeitreihe zu beobachten, gibt es auch keine „objektiv richtige“ Komponentenzerlegung (vgl. Abberger und Nierhaus (2009)). In der Praxis kommt infolgedessen eine Vielzahl von Saisonbereinigungsverfahren zur Anwendung, die zum Teil große methodische Unterschiede aufweisen. International am häufigsten angewendet werden die vom U.S. Bureau of the Census entwickelten sogenannten Census-Modelle, denen gleitende Durchschnittsverfahren zugrunde liegen. In Deutschland finden diese Verfahren unter anderem bei der Deutschen Bundesbank, dem Statistischen Bundesamt, in den Statistiken der Bundesagentur für Arbeit sowie bei der Bereinigung der Zeitreihen der ifo Konjunkturumfragen Anwendung. Das Statistische Bundesamt verwendet zusätzlich das auf Regressionsansätzen basierende Berliner Verfahren BV4.1 (Speth 2004) zur Bereinigung seiner Zeitreihen, das Statistische Amt der Europäischen Union (Eurostat) greift dagegen überwiegend auf das ursprünglich in der Banco de España entwickelte Verfahren TRAMO-SEATS zurück (Gomez und Maravall 1996).

2.4.2 Das X-13ARIMA-SEATS Verfahren

Die sogenannten Census-Verfahren wurden im U.S. Bureau of the Census entwickelt und sind die weltweit von Statistikämtern, Universitäten und sonstigen Institutionen am meisten genutzten Saisonbereinigungsverfahren. Der grundlegende Ansatz dazu wurde mit der X-11-

Methode (Shiskin u. a. 1967) gelegt, die eine Weiterentwicklung in der X-11-ARIMA-Methode (Dagum 1980) fand. Darin brachte die Option für das Anreichern der Reihen mit Prognosen und Rückrechnungen aus ARIMA-Modellen (AutoRegressive Integrated Moving Average) vor der eigentlichen Saisonbereinigung eine Verbesserung der Schätzeigenschaften am Reihenrand mit sich. Eine Fortentwicklung des ARIMA-Ansatzes durch das US Census Bureau erfolgte in der Version X-12-ARIMA, die zusätzlich weitere Diagnosewerkzeuge zur Beurteilung der Leistungsfähigkeit der Saisonbereinigung bot. Die neueste Version der Census-Verfahren ist das X-13ARIMA-SEATS-Verfahren (U.S. Bureau of the Census 2011; citebureau2013time), welches unter anderem auch bereits bei den ifo Konjunkturumfragen zum Einsatz kommt.

Die Saisonbereinigung mit dem X-13ARIMA-SEATS-Verfahren basiert in ihren Grundzügen auf folgender Methodik (vgl. Tabelle 2.5): Neben der Wahl der Verknüpfung der Komponenten (zumeist additiv oder multiplikativ) müssen zunächst keine expliziten Modellspezifikationen getroffen werden. In einem vorbereitenden Schritt werden die Zeitreihen an den Rändern mit einem ARIMA-Ansatz durch Prognosen verlängert. Dadurch kann die Saisonbereinigung komplett mit symmetrischen Saison- und Trendfiltern durchgeführt werden. Dies hilft bei der Vermeidung von Verzerrungen der Ergebnisse durch asymmetrische Filter. Vor Beginn der eigentlichen Saisonbereinigung können darüber hinaus bereits Witterungs- oder Arbeitstageeffekte sowie Extremwerte, die bei der Bestimmung der Saisonfaktoren zu einem zu starken Einfluss des irregulären Rests und somit zu Verzerrungen führen könnten, per Regressionsansatz aus den Zeitreihen eliminiert werden.

Der Berechnung der Zeitreihenkomponenten, die den Kern der Saisonbereinigung darstellt, liegen verschiedene Filterverfahren auf Basis von gleitenden Durchschnitten zugrunde. Dabei wird zuerst mit Hilfe eines Trendfilters die glatte Komponente geschätzt und aus der Zeitreihe herausgerechnet, so dass nur noch die Saisonkomponente und der irreguläre Rest übrig bleiben. Diese beiden Komponenten zusammen werden auch als rohe Saisonkomponente bezeichnet. Aus dieser rohen Saisonkomponente wird mit einem Saisonfilter, der wiederum auf einem gleitenden Durchschnittsverfahren basiert, die irreguläre Komponente ausgeschaltet, und man erhält schließlich eine Schätzung der Saisonkomponente. Durch Eliminierung dieser Saisonkomponente aus der Ursprungszeitreihe ergibt sich die saisonbereinigte Zeitreihe. Dieses Verfahren wiederholt sich über mehrere Iterationsstufen, bis die endgültige Komponentenzerlegung und somit die saisonbereinigte Zeitreihe feststeht.

Abschließend bietet das X-13ARIMA-SEATS Verfahren verschiedene Diagnoseinstrumente und Kennzahlen, die die Qualität der Saisonbereinigung beschreiben und Hinweise auf eventuelles Optimierungspotenzial bei der Wahl der Filter oder anderer Einstellungen geben können. So wird unter anderem getestet, ob gewählte Regressoren zur Eliminierung von Arbeitstage- oder sonstigen Effekten geeignet sind, oder ob noch systematische Muster in den bereinigten Reihen zu erkennen sind.

Die Census-Verfahren haben zahlreiche Vorteile. Neben der großen Auswahl an zusammenfassenden Kennzahlen und Kontrollstatistiken sowie den ARIMA-Ansätzen zur Reduzierung der

2 Methodische Grundlagen der ifo Befragungen

Randwertproblematik bieten vor allem die verschiedenen Optionen bei der Durchführung der einzelnen Schritte ein hohes Maß an Flexibilität. Gleichzeitig ist dies aber auch ein Ansatzpunkt für Kritiker, welche die Übersichtlichkeit des Verfahrens bemängeln oder gar eine Möglichkeit für Manipulationen sehen (vgl. Sauer und Wohlrabe 2015). Eine ausführliche Dokumentation der verwendeten Einstellungen ist daher unerlässlich.

Tabelle 2.5: Grundzüge des X-13ARIMA-SEATS-Verfahrens

	Schritte	Methode
Optionale vorbereitende Bearbeitung der Zeitreihen	Verlängerung der Zeitreihen zur Vermeidung von Randwertproblemen	ARIMA-Ansätze
	Eliminierung von Witterungs- und Kalendereffekten bzw. von Extremwerten	Regressionsansätze
Iterativer Kernprozess der Saisonbereinigung	Schätzung der glatten Trend-Konjunktur-Komponente	Trendfilter mit gleitenden Durchschnitten
	Berechnung der rohen Saisonkomponente	Herausrechnen der geschätzten glatten Komponente aus der Ursprungsreihe
	Schätzung der Saisonkomponente	Saisonfilter mit gleitenden Durchschnitten
	Ermittlung der saisonbereinigten Zeitreihe	Herausrechnen der geschätzten Saisonkomponente aus der Ursprungsreihe
Diagnose der Qualität der Bereinigung	Überprüfung auf noch vorhandene Saison- oder Kalendereffekte	statistische Tests und Spektralanalyse
	Überprüfung der verwendeten Einstellungen	Berechnung verschiedener Kennzahlen

Modellwahl der Komponentenzerlegung

Die Entscheidung für die Art der Komponentenzerlegung fällt in der Regel auf einen der beiden gängigsten Ansätze, das additive oder das multiplikative Modell. Beim additiven Modell

summieren sich die einzelnen Komponenten zur Ursprungszeitreihe auf. Die Saisonkomponente und der irreguläre Rest sind demnach absolute Abweichungen vom Niveau der Trend-Konjunktur-Komponente. Im multiplikativen Modell wird die Ursprungszeitreihe dagegen als Produkt dargestellt, so dass die einzelnen Komponenten relative Einflüsse widerspiegeln. Ein Saisonfaktor von 1,05 in einem Monat würde also das Zeitreihenniveau um fünf Prozent anheben. Der absolute Anstieg ist damit abhängig vom aktuellen Niveau der Zeitreihe. Das multiplikative Modell impliziert demzufolge einen Zusammenhang der saisonalen Schwankungsmuster mit dem Reihenniveau. Kann kein solcher Zusammenhang festgestellt werden, ist das additive Modell vorzuziehen. Dieses bietet zudem den Vorteil einer einfacheren Interpretation der einzelnen Zeitreihenkomponenten. Darüber hinaus ist das multiplikative Modell lediglich auf Zeitreihen mit ausschließlich positiven Werten anwendbar, was beispielsweise auf viele Zeitreihen der ifo Konjunkturumfragen nicht zutrifft.

Wahl der Trend- und Saisonfilter

Zur Annäherung der glatten Trend-Konjunktur-Komponente werden verschiedene Filter verwendet, die auf gleitenden Durchschnittsverfahren (Moving Averages) basieren (vgl. Ladiray und Quenneville 2001 sowie Ladiray und Quenneville 2001). Gleitende Durchschnitte einer Zeitreihe X_t lassen sich in folgender Form darstellen:

$$M[X_t] = \sum_{k=-p}^{+f} \theta_k X_{t+k}$$

Ausschlaggebend für das Ergebnis des Filterverfahrens ist zum einen die Wahl der Gewichte θ_k und zum anderen die Festlegung der Parameter p und f , die den Stützbereich des gleitenden Durchschnitts bestimmen. Zumeist wird ein symmetrischer Stützbereich gewählt, das heißt $p = f$, so dass der gesamte Stützbereich eine ungerade Anzahl an Werten enthält. Als Standardeinstellung für die gleitenden Durchschnitte werden für Monatsdaten in der Regel 13 Werte gewählt ($p = f = 6$), um Schwingungen mit Perioden von weniger als einem Jahr aus der Zeitreihe zu filtern. In den ersten Iterationsstufen kommt der 2x12-gleitende Durchschnitt zum Einsatz. Dieser minimiert die Verzerrungen durch evtl. in der Reihe vorhandene Extremwerte, da er stärker glättet. Auf der letzten Iterationsstufe wird dagegen ein Henderson-Filter verwendet (vgl. Henderson 1916). Dessen Variabilität hängt entscheidend vom gewählten Stützbereich ab. Je kürzer dieser gewählt wurde, desto weniger glatt verläuft die Trend-Konjunktur-Komponente. In Tabelle 2.6 finden sich die Gewichte der Trendfilter für verschiedene Stützbereiche. Das Gewichtungsschema resultierte aus den beiden von Henderson gestellten Bedingungen, nämlich dass der Filter Polynome vom Grade kleiner als 4 unverändert lässt und dass er die Varianz der dritten Differenzen der Zeitreihe minimiert (vgl. Abberger und Nierhaus 2009).

Dem Saisonfilter zur Berechnung der Saisonkomponente liegt ebenfalls ein gleitendes Durch-

2 Methodische Grundlagen der ifo Befragungen

Tabelle 2.6: Gewichte der Trendfilter

Filter	θ_{t-6}	θ_{t-5}	θ_{t-4}	θ_{t-3}	θ_{t-2}	θ_{t-1}	θ_t	θ_{t+1}	θ_{t+2}	θ_{t+3}	θ_{t+4}	θ_{t+5}	θ_{t+6}
2x12-gleitender Durchschnitt	1/24	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/24
13-Term-Henderson-Filter	-0.019	-0.028	0.000	0.066	0.147	0.214	0.240	0.214	0.147	0.066	0.000	-0.028	-0.019
9-Term-Henderson-Filter			-0.041	-0.010	0.118	0.267	0.331	0.267	0.118	-0.010	-0.041		
7-Term-Henderson-Filter				-0.059	0.059	0.294	0.413	0.294	0.059	-0.059			
5-Term-Henderson-Filter					-0.073	0.294	0.559	0.294	-0.073				

schnittsverfahren zugrunde. Dazu werden zunächst aus der rohen Saisonkomponente 12 Unterreihen gebildet, in denen jeweils die Werte gleicher Kalendermonate zusammengefasst sind (bei Quartalswerten werden dementsprechend 4 Unterreihen gebildet). Für die Schätzung der monatspezifischen Saisoneffekte stehen standardmäßig 3x3-, 3x5-, 3x9- sowie 3x15-gleitende Durchschnitte zur Verfügung. Außerdem steht ein stabiler Saisonfilter zur Auswahl, der nur den Mittelwert aller Werte des gleichen Monats bildet. Die Anwendung dieses Filters unterstellt, dass das Saisonmuster konstant ist und sich über die Jahre nicht verändert. Die Gewichte verschiedener Filter sind in Tabelle 2.7 ersichtlich. Kürzere Stützbereiche haben dabei eine höhere Sensibilität für Schwankungen im Saisonmuster. Der Vorteil von größeren Stützbereichen liegt wiederum in der besseren Stabilität der Ergebnisse; allerdings wächst auch das Risiko, nicht alle Saisonausschläge aus der Reihe zu eliminieren.

Tabelle 2.7: Gewichte der Saisonfilter

Filter	θ_{t-5}	θ_{t-4}	θ_{t-3}	θ_{t-2}	θ_{t-1}	θ_t	θ_{t+1}	θ_{t+2}	θ_{t+3}	θ_{t+4}	θ_{t+5}
3x3-gleitender Durchschn.				1/9	2/9	3/9	2/9	1/9			
3x5-gleitender Durchschn.			1/15	2/15	3/15	3/15	3/15	2/15	1/15		
3x9-gleitender Durchschn.	1/27	2/27	3/27	3/27	3/27	3/27	3/27	3/27	3/27	2/27	1/27
stabiler Filter	Mittelwert über alle Werte mit gleichem Kalendermonat										

Behandlung des Witterungseinflusses

Viele Zeitreihen weisen einen engen Zusammenhang mit der Witterung auf. Vor allem im Baugewerbe, aber auch in anderen Bereichen - wie etwa der Verkehrswirtschaft oder manchen Handelsbereichen - kann schlechtes Wetter stark bremsend wirken. Der Einfluss der Witterung kann mit Hilfe einfacher Regressionsrechnungen abgeschätzt werden. Als abhängige Variable werden dabei beispielsweise die Anzahl der Eistage in einem Monat oder andere vom Deutschen Wetterdienst veröffentlichte Klimaindikatoren herangezogen (vgl. Hielscher und Enkelmann 2014).

Kritiker wenden jedoch gegen eine Witterungsberreinigung ein, dass der Einfluss der Witterung in der saisonbereinigten Reihe erkennbar bleiben sollte, da er als wesentlicher Bestimmungsfaktor in die Konjunkturbeurteilung eingeht. Auch das Statistische Amt der Europäischen Union gibt in seinen Richtlinien zur Saisonberreinigung vor, dass Witterungseinflüsse in der bereinigten Reihe nicht unterdrückt werden sollten (vgl. ESS Eurostat 2009). Die meisten wirtschaftlichen Zeitreihen, wie etwa die vom Statistischen Bundesamt veröffentlichten amtlichen Statistiken,

werden daher ohne Witterungsberreinigung ausgewiesen.

Behandlung des arbeitstäglichen Einflusses

Die unterschiedliche Anzahl der Arbeitstage eines Monats kann einen weiteren Sondereinfluss auf den Verlauf von ökonomischen Zeitreihen darstellen (vgl. Bell und Hillmer 1983). Insbesondere auf Produktions-, Auftragseingangs- oder Umsatzreihen üben arbeitstägliche Effekte eine signifikante Wirkung aus. Bei Reihen, in denen ein signifikanter Kalendereffekt sichtbar wird, ist daher eine arbeitstägliche Bereinigung sinnvoll. Die resultierenden Zeitreihen sind dann unabhängig von der Länge beziehungsweise der Zusammensetzung der Monate und Quartale (Anzahl der Montage, Dienstage etc., Anzahl der Arbeitstage beziehungsweise der Wochenend- und Feiertage).

Die Berechnung der Arbeitstagekomponente erfolgt analog zur Witterungskomponente anhand eines Regressionsansatzes. Dazu verwendet man als unabhängige Variable in der Regel eine Reihe mit der jeweiligen Anzahl der Arbeitstage pro Monat beziehungsweise Quartal oder daraus abgeleitete Größen. Für Deutschland ist eine solche Arbeitstagereihe in der Volkswirtschaftlichen Gesamtrechnung des Statistischen Bundesamtes verfügbar (vgl. Speth 2004).

2.4.3 Beurteilung der Leistungsfähigkeit von Saisonbereinigungsverfahren

Vor dem Hintergrund abweichender Zeitreihenverläufe bei der Verwendung unterschiedlicher Bereinigungsmethoden stellt sich die Frage, welches Verfahren beziehungsweise welche Einstellungen des Verfahrens am besten geeignet sind. Da der „wahre“ saisonbereinigte Verlauf in der Realität nicht beobachtbar ist, kann die Güte einer Zeitreihenzerlegung lediglich anhand vergleichender statistischer Kriterien, wie etwa der Stabilität der Ergebnisse am aktuellen Rand, oder durch die Analyse des Zusammenhangs mit Zielzeitreihen beurteilt werden (vgl. Goldrian und Lehne 1999). Dabei können die unterschiedlichen Kriterien für eine Zeitreihe von unterschiedlich großer Bedeutung sein.

Die wichtigste Eigenschaft einer saisonbereinigten Zeitreihe ist per Definition, dass sie keine erkennbaren Saisoneffekte mehr enthält. Neben der rein grafischen Betrachtung der Reihen, die schon erste Aufschlüsse über mögliche sich jährlich wiederholende Strukturen geben kann, können die Ergebnisreihen mit einem F-Test (vgl. Higginson 1976) beziehungsweise mit einer Spektralanalyse auf noch vorhandene Saisonalität überprüft werden. Das X-13ARIMA-SEATS-Programm etwa führt diese Tests standardmäßig aus und gibt gegebenenfalls eine Warnung aus. Die Parametereinstellungen von Saisonbereinigungsverfahren sollten daher für die jeweiligen Reihen stets so gewählt werden, dass in den Resultaten keine Saisonstrukturen mehr erkennbar sind.

Darüber hinaus gibt es gerade für vielbeachtete Indikatoren einige wünschenswerte Qualitätsmerkmale und Prognoseeigenschaften. So soll der Indikator unter anderem nach der Publikation keinen größeren Revisionen ausgesetzt sein sowie frühzeitige und möglichst

2 Methodische Grundlagen der ifo Befragungen

eindeutige Hinweise auf den Konjunkturverlauf enthalten (vgl. Abberger und Wohlrabe 2006). Bei allen Saisonbereinigungsverfahren kommt es zu mehr oder weniger starken Änderungen der bereinigten Reihen, wenn die Ergebnisse neuer Erhebungsmonate miteinbezogen werden (vgl. Sauer und Wohlrabe 2015). Da sich Werte in der Folgezeit nach ihrer Erstveröffentlichung in der Regel aber so wenig wie möglich ändern sollen, erfolgt die Beurteilung der Eignung eines Saisonbereinigungsverfahrens unter anderem anhand von Kriterien, die auf das Ausmaß der durch die Einbeziehung aktueller Reihenwerte verursachten Revisionen abzielen. Das Bereinigungsverfahren muss demnach ein hohes Maß an Stabilität der Ergebnisse gewährleisten, insbesondere am aktuellen Rand. Darüber hinaus sollten sich anfangs angezeigte konjunkturelle Tendenzen oder Wendepunkte im Nachhinein nicht mehr ändern. Das heißt, wenn etwa ein Wert von einem Monat zum nächsten ansteigt, sollte diese Entwicklung auch nach eventuellen Revisionen bestehen bleiben. Zur Quantifizierung des Ausmaßes der Revisionen können verschiedene Vergleichsmaße herangezogen werden. Diese werden unter anderem von Schips und Stier (1975) diskutiert.

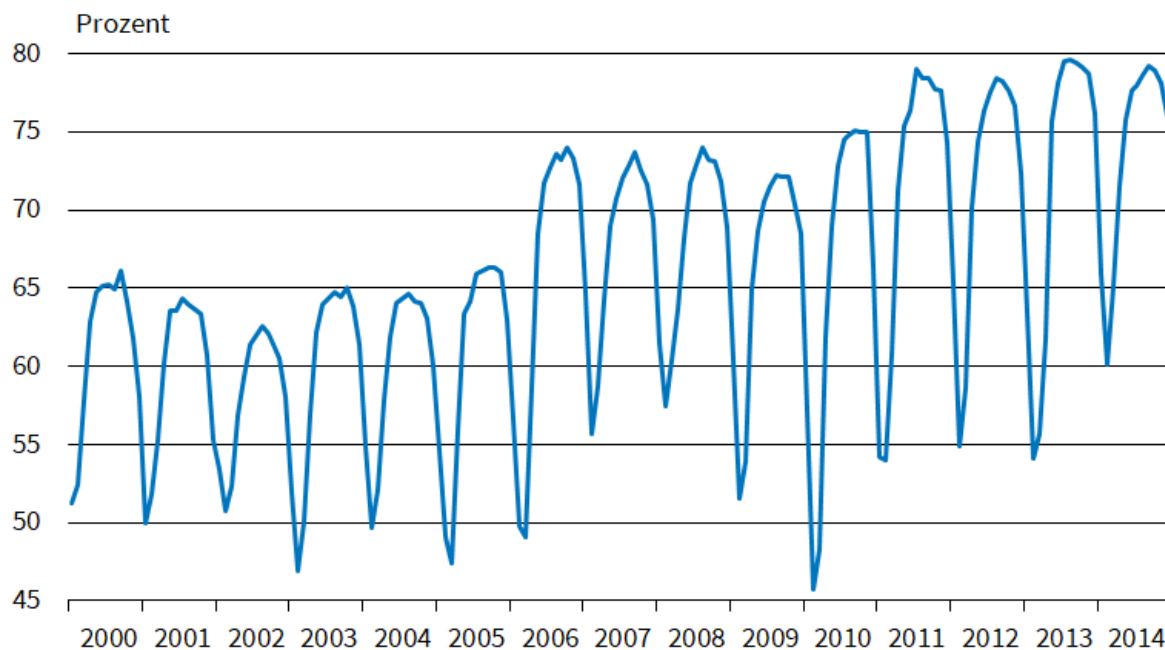
Die Stabilität eines Verfahrens stellt zwar ein Kriterium mit sehr hoher Priorität dar, dabei sollten allerdings nicht die Auswirkungen auf die Aussagekraft sowie die Prognoseeigenschaften der Zeitreihen aus den Augen verloren werden. Besonders der Zusammenhang mit interessierenden Zielgrößen, beispielsweise im Rahmen von Konjunkturprognosen, ist in dieser Hinsicht von großer Bedeutung. Eine Möglichkeit zur Berechnung der Prognosegüte ist bei Henzel (2015) beschrieben.

Prognosegüte und Stabilität sind jedoch zumeist entgegengesetzte Ziele und daher nur schwer miteinander vereinbar, weshalb bei der Wahl des Saisonbereinigungsverfahrens und dessen Einstellungen oftmals ein Kompromiss gefunden werden muss. Dabei muss die Wichtigkeit der beiden Kriterien jeweils in Abhängigkeit von der Fragestellung individuell definiert werden.

2.4.4 Zur Saisonbereinigung in den ifo Konjunkturumfragen

Viele Zeitreihen aus den ifo Konjunkturumfragen weisen eine deutlich erkennbare Saisonfigur auf. Abbildung 2.4 veranschaulicht dies am Beispiel des von den Umfrageteilnehmern gemeldeten Auslastungsgrades der Gerätekapazitäten im Bauhauptgewerbe. Dieser verringert sich in jedem Jahr in den Wintermonaten spürbar und beginnt nach dem Einsetzen einer für das Baugewerbe günstigeren Witterung im Frühjahr wieder auf ein erheblich höheres Niveau zu steigen. Aufgrund derartiger Saisonmuster ist zur besseren Interpretation der Zeitreihen der ifo Konjunkturumfragen eine Saisonbereinigung notwendig. Dafür wurde über Jahrzehnte das ASA-II-Verfahren verwendet, das in seinen Grundzügen aus einer Zusammenarbeit des ifo Instituts mit dem HWWA-Institut für Wirtschaftsforschung in Hamburg sowie dem Rheinisch-Westfälischen Institut in Essen entstanden ist (vgl. Danckwerts u. a. 1970) und in verschiedenen Weiterentwicklungen speziell auf die Gegebenheiten der Zeitreihen der ifo Konjunkturumfragen abgestimmt wurde (vgl. Goldrian 1973a,b). Zu Beginn des Jahres 2015 wurde das Saisonbereinigungsverfahren im Rahmen der Konjunkturumfragen auf das zuvor

Abbildung 2.4: Kapazitätsauslastung im Bauhauptgewerbe in %. Ursprungswerte (nicht saisonbereinigt).



Quelle: ifo Konjunkturumfragen.

© ifo Institut

beschriebene und international am weitesten verbreitete X-13ARIMA-SEATS-Verfahren umgestellt (vgl. Sauer und Wohlrabe 2015). Zum einen ist damit eine bessere Vergleichbarkeit zur offiziellen Statistik gewährleistet, und zum anderen bietet dieses Verfahren den Vorteil, dass auf die individuellen Gegebenheiten der verschiedenen Zeitreihen, wie zum Beispiel sich verändernde Saisonmuster, flexibler eingegangen werden kann.

So konnten beispielsweise die Stützbereiche der Saisonfilter für alle Reihen individuell eingestellt werden. Sie wurden jeweils so groß wie möglich gewählt, um weitestgehende Stabilität der Reihen gegenüber Revisionen zu gewährleisten, jedoch unter der Bedingung, dass im Anschluss kein signifikantes Saisonmuster mehr in der bereinigten Reihe zu erkennen ist. Dabei erwies sich auch der Saisonfilter, der von einem stabilen Saisonmuster ausgeht, für einige Zeitreihen als angemessen. Die Filtereinstellungen werden regelmäßig überprüft und können bei Änderungen im Saisonmuster gegebenenfalls ohne großen Aufwand angepasst werden.

Internationalen Standards folgend wird keine Witterungsbereinigung mehr durchgeführt, aufgrund des Nachweises von Kalendereffekten in einigen Zeitreihen mittels Spektralanalysen beinhaltet die Saisonbereinigung jedoch zusätzlich eine arbeitstägliche Bereinigung. Diese Kalendereffekte sind allerdings nur in einigen Variablen nachweisbar, die sich auf den aktuellen Monat beziehungsweise auf den Vormonat beziehen. Die arbeitstägliche Bereinigung beschränkt sich daher auf die aktuelle Geschäftslage sowie im Verarbeitenden Gewerbe

2 Methodische Grundlagen der ifo Befragungen

auf die Vormonatsentwicklung von Produktion, Nachfragesituation und Auftragsbeständen. Andere Variable - wie Geschäftserwartungen, Produktions- oder Personalpläne -, die sich auf zukünftige Entwicklungen beziehen, weisen dagegen keine arbeitstäglichen Effekte auf und werden deshalb lediglich saisonbereinigt.

Analysen des neuen Verfahrens haben ergeben, dass in Bezug auf die Eigenschaften der bereinigten Zeitreihen sehr gute Ergebnisse erzielt werden. Dies spiegelt sich unter anderem in dem nur geringfügigen Ausmaß an Revisionen sowie einem größeren Zusammenhang mit interessierenden wirtschaftlichen Größen wider (vgl. Sauer und Wohlrabe 2015 sowie Henzel 2015).