

Gerhard Thury

Kalendereffekte in der Industrieproduktion

Kalendereffekte, die durch gesetzliche Feiertage und im Anschluß daran konsumierte Urlaubstage sowie durch die von Jahr zu Jahr variierende Wochentagszusammensetzung der einzelnen Monate ausgelöst werden, verursachen Zufallsschwankungen in der Industrieproduktion. Der Einsatz statistischer Methoden zur Saisonbereinigung und Prognose wird dadurch erschwert oder überhaupt unmöglich gemacht. Eine einfache Korrektur der betroffenen Zeitreihen um die Zahl der Arbeitstage, wie sie das Österreichische Statistische Zentralamt durchführt, löst das Problem nicht.

Das Produktionsvolumen eines Monats hängt von der Zahl der Arbeitstage ab. Die relativ große Zahl der gesetzlichen Feiertage in Österreich muß daher beträchtliche Schwankungen in der Produktionstätigkeit bewirken. Aber es sind nicht nur diese Feiertage selbst, die diese Produktionsschwankungen auslösen. Fast noch entscheidender erscheint die gängige Praxis, im Zusammenhang mit diesen Feiertagen zusätzlich Urlaub zu nehmen, und zwar vor allem dann, wenn solche Tage innerhalb der Woche günstig fallen. Die Tatsache, daß dies nicht in jedem Jahr der Fall ist, führt zu weiteren Verzerrungen. Die Zeit zwischen Weihnachten und Neujahr (oder auch Dreikönig) ist ein klassisches Beispiel. Auch das unterschiedliche Datum der Osterfeiertage — im März oder April — hat deutliche Effekte. Weiters sind die schulfreien und daher „urlaubsanfälligen“ Diens-tage nach Ostern und Pfingsten zu berücksichtigen. Ebenso wären hier die Donnerstag-Feiertage im Frühsommer zu nennen, die viele Erwerbstätige als „Fensterfreitage“ frei nehmen. Auch das Datum dieser Feiertage kann zwischen zwei Monaten — Mai und Juni — wechseln.

In der einschlägigen Literatur werden solche Einflüsse als Kalendereffekte bezeichnet. Sie erschweren die Anwendung von formal statistischen Methoden zur Analyse von Zeitreihen stark oder machen sie überhaupt unmöglich. Die Saisonbereinigung und Prognose mit Hilfe von Zeitreihenmethoden bringt kaum brauchbare Ergebnisse, wenn die zu untersuchenden Reihen nicht von diesen Kalendereffekten bereinigt werden. Das WIFO bereinigt etwa die Industrieproduktion durch Gewichtung der Ori-

nalreihe mit einer ebenfalls vom Statistischen Zentralamt publizierten arbeitstägig bereinigten Reihe im Verhältnis 7 zu 3. Die Gewichte wurden aufgrund langjähriger Erfahrung gewählt. Diese Bereinigung mildert zwar die bei der Saisonbereinigung des Produktionsindex auftretenden Probleme, kann sie aber nicht beseitigen. Eine Prognose des Produktionsindex mit Zeitreihenmethoden wird dadurch jedoch unmöglich gemacht: Man könnte zwar Prognosen der gewichteten Reihe erstellen, aber es gibt keinen Weg, daraus Prognosen der Originalreihe abzuleiten. Und gerade sie interessieren den Praktiker.

Die vorliegende Arbeit sucht daher eine Methode, Kalendereffekte aus der Industrieproduktion so auszuschalten, daß die bereinigte Reihe sowohl für Prognosen als auch für die Saisonbereinigung weiter verwendet werden kann.

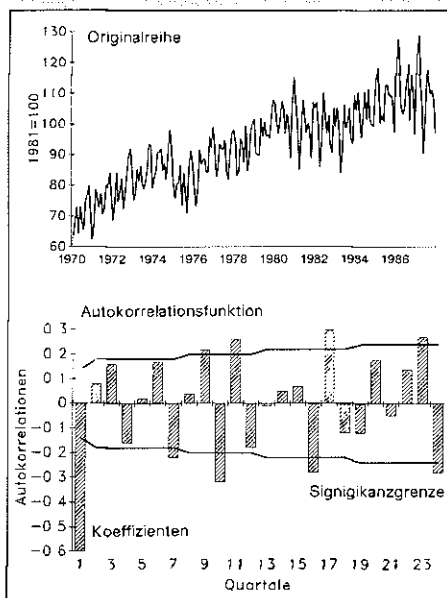
Ein ARIMA-Modell für den Produktionsindex

Die Publikation des Buches von Box — Jenkins (1970) trug entscheidend zur vermehrten Anwendung von

ARIMA-Modellen bei. Eine stichwortartige Einführung in die Theorie der ARIMA-Modelle findet sich im Anhang. Voraussetzung für die Schätzungen von ARIMA-Modellen ist eine genaue Beobachtung der Originalreihe. Abbildung 1 gibt im oberen Teil den monatlichen Index der Industrieproduktion (IP) vom Jänner 1969 bis zum August 1987 wieder. Die Reihe weist beträchtliche Saisonbewegungen auf, deren Muster von Jahr zu Jahr stark zu schwanken scheint. Schwierigkeiten bei der Saisonbereinigung sind daher zu befürchten. Weiters unterliegt die Reihe einem Trend, und ihre Schwankungsbreite nimmt mit steigendem Niveau zu. Mittelwert und Varianz sind über die Zeit also nicht konstant, d. h. die Reihe ist nichtstationär. Die Varianz kann durch eine Transformation der Reihe in natürliche Logarithmen ($\ln$) stabilisiert werden. Trend und Saison sind — wenn sie theoretischen Idealvorstellungen entsprechen — durch Bildung geeigneter Differenzen auszuschalten: der Trend durch Differenzen gegenüber dem Vormonat, die Saisonkomponente über Differenzen gegenüber dem entsprechenden Vorjahresmonat.

Ausgangspunkt der Identifikation eines ARIMA-Modells ist die Analyse der Autokorrelationsfunktion der zu modellierenden Reihe. Abbildung 1 zeigt in ihrem unteren Teil die Autokorrelationsfunktion der transformierten und entsprechend gefilterten Produktionsreihe; die Autokorrelationskoeffizienten für eine Verzögerung von 1 bis 24 Monaten sind als Stäbe und ihre Signifikanzgrenzen als Linien dargestellt. Wenn ein Koeffizient diese Signifikanzgrenze übersteigt, gilt er als signifikant von Null verschieden.

Originalreihe des monatlichen Produktionsindex und Autokorrelationsfunktion für $(1 - B)(1 - B^{12}) \ln IP_t$ Abbildung 1



Die Originalreihe der Industrieproduktion weist beträchtliche Saisonschwankungen auf, und überdies schwankt das Saisonmuster von Jahr zu Jahr stark. Aus der Gestalt der Autokorrelationsfunktion der logarithmierten und differenzierten Reihe kann man schließen, daß es ohne vorherige Korrektur nicht möglich sein wird, ein brauchbares Zeitreihenmodell für die Industrieproduktion zu finden.

Dies trifft für eine größere Zahl der geschätzten Koeffizienten zu. Einige weitere Werte liegen zudem nur knapp unter der Signifikanzgrenze. Das zu schätzende Modell müßte also eine sehr komplexe Struktur aufweisen. Ein entscheidendes Hindernis ist aber die Tatsache, daß in dieser Folge von Autokorrelationskoeffizienten kein Muster zu erkennen ist. Auch die partielle Autokorrelationsfunktion, das zweite wichtige Instrument für die Identifikation eines ARIMA-Modells, liefert keine bessere Information.

Die Gestalt der beiden Autokorrelationsfunktionen weist eindeutig darauf hin, daß der Produktionsindex keinem ARIMA-Modell niedriger Ordnung folgt. Sie impliziert vielmehr, daß jeder Wert der trend- und saisonberei-

nigten Industrieproduktion eine unabhängige und höchst komplexe Linearkombination von Zufallsfehlern ist. In diesem Fall fällt der autoregressive (AR) Teil des ARIMA-Modells weg, und im Moving-average-Teil (MA) wird für jeden signifikanten Autokorrelationskoeffizienten ein eigener Parameter geschätzt. Bei einer größeren Zahl von signifikanten Autokorrelationskoeffizienten, wie dies hier der Fall ist, sind große Schätzprobleme unausweichlich. Das Modell ist überparameterisiert, die Schätzwerte für die einzelnen Parameter sind hochkorreliert und damit sehr unzuverlässig. Die schlechten Prognoseeigenschaften solcher Modelle sind hinreichend bekannt.

Die mit den Schätzergebnissen in Übersicht 1 angeführten Teststatistiken zeigen, daß das Modell für die Erklärung der Industrieproduktion unbrauchbar ist. Der Schätzfehler der Industrieproduktion (SE) ist mit über 4% sehr hoch, die Berechnung brauchbarer Prognosen ist somit kaum möglich. Außerdem weist der hohe Wert der Q -Statistik auf Autokorrelation in den Residuen dieses Modells hin – trotz der bereits relativ großen Zahl von MA-Parametern. Freilich sollte die Schätzung dieses Modells nur den Ausgangspunkt für eventuelle Modellverbesserungen liefern.

Ein so enttäuschendes Ergebnis ist bei der Schätzung von ARIMA-Modellen für makroökonomische Zeitreihen leider keine Ausnahme. Bald nach Einführung dieser Modelle wurde nämlich deutlich, daß sie ökonomische Zeitreihen oft nur unzureichend beschreiben können. In der Folge suchte man die Klasse der ARIMA-Modelle so zu erweitern, daß sie auch diese Zeitreihen adäquat modellieren werden können. Die Form der Autokorrelationsfunktion in Abbildung 1 läßt vermuten, daß die bei der Modellierung des Produktionsindex auftretenden Probleme in Kalendereffekten ihre Ursache haben. Methoden zur

Ausschaltung solcher Effekte schlagen Liu (1980), Bell – Hillmer (1981) und Hillmer (1982) vor.

Ausschaltung von Kalendereffekten

Kalendereffekte sind ein mehrschichtiges Problem. Nicht nur die im Zeitablauf variierende Zahl der Arbeitstage pro Monat verursacht die beobachteten erratischen Produktionsschwankungen – eine weitere Ursache liegt in der unterschiedlichen Wochentagszusammensetzung der einzelnen Monate, weil die Arbeitsproduktivität nicht an allen Wochentagen gleich hoch ist. Dieses Problem ist deshalb besonders knifflig, weil die Wochentagszusammensetzung auch für den gleichen Monat von Jahr zu Jahr schwankt. Schaltjahre machen das Problem noch komplexer. Da diese Kalendereffekte weitgehend durch institutionelle Faktoren ausgelöst werden, scheint ihre Ausschaltung wünschenswert. Das kann geschehen, indem man sie durch einen deterministischen Ansatz zu approximieren versucht. Dabei wird unterstellt, daß der Kalendereffekt, KE_t , im Monat t eine Funktion der Anzahl verschiedener Wochentage, T_{it} , im Monat ist:

$$(1) \quad KE_t = \sum_{i=1}^7 \beta_i T_{it},$$

T_{it} , $i = 1, \dots, 6$, steht für die Zahl der Montage (Dienstage usw.) minus Zahl der Sonntage bis zur Zahl der Samstage minus Zahl der Sonntage im Monat t , T_{7t} bezeichnet die Gesamtzahl der Tage im Monat t . Diese Normierung erfolgt, um orthogonale Variable zu erhalten: Die Gesamtzahl der Tage ist nämlich die Summe der einzelnen Wochentage und somit kollinear.

Das Modell für die beobachtete Zeitreihe der Industrieproduktion hat somit folgende Form:

$$(2) \quad \ln IP_t = KE_t + N_t,$$

N_t folgt einem ARIMA-Modell.

Vor der Schätzung der β_i aus den Daten sind die Variablen T_{it} zu konstruieren. Der erste Schritt ist sehr einfach. Die PC-Version des SCA-Systems, die für alle Berechnungen dieser Arbeit verwendet wurde, liefert die Zahl der verschiedenen Wochentage in jedem Monat für einen gewünschten Zeitraum. Diese Aufstel-

Ein einfaches ARIMA-Modell der Industrieproduktion

Übersicht 1

$$(1 - B)(1 - B^{12}) \ln IP_t =$$

$$(1 - 0,775 B + 0,132 B^3 - 0,172 B^7 + 0,257 B^9 - 0,228 B^{11})(1 - 0,748 B^{12}) a_t$$

(0,056) (0,057) (0,056) (0,077) (0,074) (0,054)

$$SE = 0,042$$

$$DW = 2,09$$

$$Q(24) = 64,2$$

Schätzwerte der Wochentagsgewichte

Übersicht 2

$$\begin{aligned}
 & (1 - B)(1 - B^{12}) \ln IP_t = \\
 & 0,005 (1 - B)(1 - B^{12}) T_{1t} + 0,007 (1 - B)(1 - B^{12}) T_{2t} + \\
 & (0,001) \quad (0,002) \\
 & 0,012 (1 - B)(1 - B^{12}) T_{3t} + 0,006 (1 - B)(1 - B^{12}) T_{4t} + \\
 & (0,002) \quad (0,002) \\
 & 0,006 (1 - B)(1 - B^{12}) T_{5t} - 0,018 (1 - B)(1 - B^{12}) T_{6t} + \\
 & (0,002) \quad (0,002) \\
 & 0,029 (1 - B)(1 - B^{12}) T_{7t} + \epsilon_t \\
 & (0,006)
 \end{aligned}$$

$SE = 0,024$ $DW = 2,20$ $Q(24) = 67,4$

lung spiegelt aber die Besonderheiten des österreichischen Kalenders mit seiner relativ großen Zahl von Feiertagen nicht wider. Sie muß daher für österreichische Verhältnisse schritt-

Durch die Einführung von Kalendervariablen in ein ARIMA-Modell gelingt es, erratische Schwankungen der Industrieproduktion auszuschalten. Der Schätzfehler dieses erweiterten Modells sinkt auf die Hälfte. Das Modell kann für Zwecke der Saisonbereinigung und Prognose eingesetzt werden.

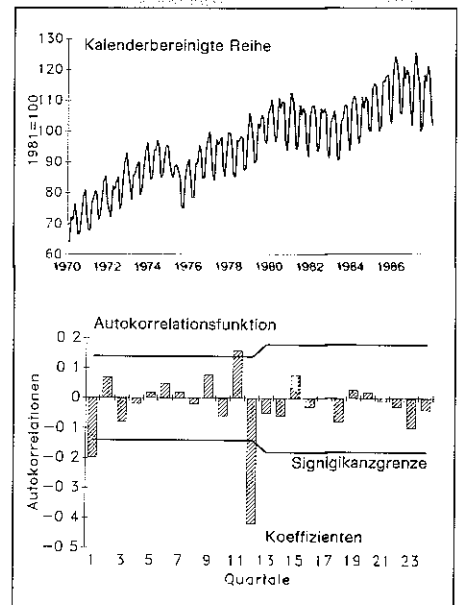
weise adaptiert werden (es sollte aber nicht verhehlt werden, daß dadurch ein Element von Willkür eingeführt wird). Zunächst werden die gesetzlichen Feiertage als Sonntage behandelt und Gleichung (2) versuchsweise geschätzt. Für N_t wird vorerst ein sehr einfaches ARIMA-Modell unterstellt: $(1 - B)(1 - B^{12}) N_t = a_t$. Das Ergebnis dieses Schätzversuchs ist vielversprechend. Eine genauere Analyse zeigt aber, daß im Mai, Juni und Dezember noch immer größere Schätzfehler auftreten — offenbar wurden die Kalendereffekte noch nicht vollständig ausgeschaltet. Der verbliebene Effekt resultiert daraus, daß Arbeitnehmer in Zusammenhang mit Feiertagen Urlaubstage konsumieren. Anhand der Information über die jeweilige Konjunkturlage, die mitentscheidend sein dürfte, ob zusätzlicher Urlaub gewährt wird und ob Feiertage eingearbeitet werden, konnte rasch eine Wochentagszusammensetzung der einzelnen Monate gefunden werden, die brauchbare Schätzwerte der β_t lieferte (Übersicht 2). Alle geschätzten Parameter sind signifikant von Null verschieden und haben das erwartete Vorzeichen. Die allgemeine

Gestalt der Gewichtungsfunktion erscheint plausibel. Die Gewichte steigen in der ersten Wochenhälfte kontinuierlich an und fallen vom Höchstwert am Mittwoch ebenso kontinuierlich wieder ab. Das Gewicht für den Samstag ist erwartungsgemäß negativ.

Mit Hilfe dieser Gewichte kann man nun die verschiedenen Kalendereffekte aus der Originalreihe des Produktionsindex ausschalten. Der so korrigierte Index der Industrieproduktion ist im oberen Teil von Abbildung 2 wiedergegeben. Im Vergleich mit Abbildung 1 wird die Verbesserung deutlich. Die korrigierte Produktionsreihe weist einen sehr regelmäßigen Verlauf auf, vor allem ihr Saisonmuster scheint viel stabiler zu sein. Durch die Bereinigung gelingt es also tatsächlich, Saisoneffekte und Einflüsse, die ihre Ursache in Kalendervariationen haben, zu trennen. Dieses Ergebnis wird auch durch die Gestalt der Autokorrelationsfunktion der entsprechend differenzierten neuen Produktionsreihe (im unteren Teil von Abbildung 2) eindrucksvoll untermauert. Vernachlässigt man den Koeffizienten für Lag 11 als Zufallsprodukt, so entspricht die Autokorrelations-

Kalenderbereinigte Industrieproduktion und Autokorrelationsfunktion für $(1 - B)(1 - B^{12}) \ln IP_t$

Abbildung 2



Der Erfolg der Kalenderbereinigung ist nicht zu übersehen: Die Ursprungsreihe besitzt ein stabiles Saisonmuster, die Gestalt der Autokorrelationsfunktion für die entsprechend transformierte Reihe belegt, daß die kalenderbereinigte Industrieproduktion einem relativ einfachen Zeitreihenmodell folgt.

funktion genau einem multiplikativen saisonalen ARIMA-Modell. Das Modell für N_t ist somit doch etwas komplizierter als zunächst vermutet (Übersicht 3).

Alle Parameter — sowohl die Wochentagsgewichte als auch die beiden Moving-average-Parameter — werden nun simultan geschätzt. Die Schätzwerte sind durchwegs signifikant von Null verschieden und haben das erwartete Vorzeichen. Die Erweiterung des Modells hat den Schätzfehler gegenüber dem ursprünglichen ARIMA-Modell aus Übersicht 1 um mehr als

ARIMA-Modell für den Index der Industrieproduktion unter Ausschaltung von Wochentags- und sonstigen Kalendereffekten

Übersicht 3

$$\begin{aligned}
 & (1 - B)(1 - B^{12}) \ln IP_t = -0,001 + \\
 & (0,000) \\
 & 0,005 (1 - B)(1 - B^{12}) T_{1t} + 0,006 (1 - B)(1 - B^{12}) T_{2t} + \\
 & (0,001) \quad (0,002) \\
 & 0,013 (1 - B)(1 - B^{12}) T_{3t} + 0,007 (1 - B)(1 - B^{12}) T_{4t} + \\
 & (0,002) \quad (0,002) \\
 & 0,005 (1 - B)(1 - B^{12}) T_{5t} - 0,018 (1 - B)(1 - B^{12}) T_{6t} + \\
 & (0,002) \quad (0,001) \\
 & 0,029 (1 - B)(1 - B^{12}) T_{7t} + (1 - 0,243 B)(1 - 0,674 B^{12}) a_t \\
 & (0,007) \quad (0,068) \quad (0,053)
 \end{aligned}$$

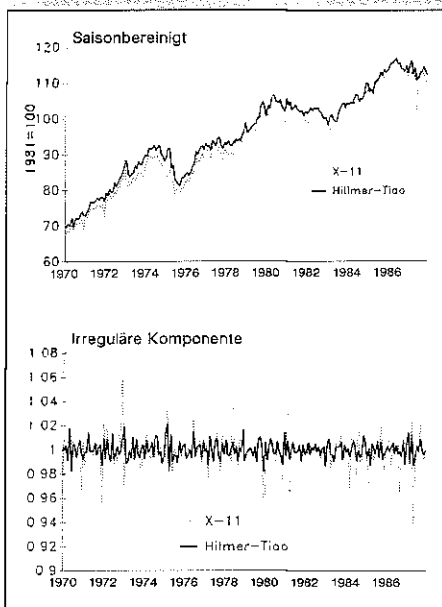
$SE = 0,0189$ $DW = 2,02$ $Q(24) = 18,2$

50% verringert, mit 1,9% weist er nun eine durchaus akzeptable Größe auf. Weiters deutet die Q -Statistik des neuen Modells nicht mehr auf Autokorrelation in den Residuen hin. Das Modell scheint somit den zeitlichen Verlauf der Produktion zufriedenstellend zu beschreiben und kann daher gefahrlos für die Saisonbereinigung und Prognose herangezogen werden.

Saisonbereinigung

Ein kurzer Überblick über eine auf einem ARIMA-Modell basierende Methode der Saisonbereinigung findet sich im Anhang. Der saison- und kalenderbereinigte Index der Industrieproduktion ist in Abbildung 3 wiedergegeben. Als Vergleichsbasis dient die traditionell saisonbereinigte Produktionsreihe des WIFO¹⁾. Die beiden Reihen scheinen nicht sonderlich zu divergieren. Eine genauere Analyse und vor allem ein Vergleich der beiden irregulären Komponenten (Abbildung 3) macht doch signifikante Unterschiede deutlich. Die Reihen bewe-

Saisonbereinigter Index - Abbildung 3 der Industrieproduktion



Die traditionell saisonbereinigte Reihe des WIFO und die anhand der kalenderbereinigten Reihe errechneten saisonbereinigten Werte divergieren kaum. Eine genauere Analyse und vor allem ein Vergleich der beiden irregulären Komponenten fördert jedoch signifikante Unterschiede zutage.

gen sich zwar über die ganze Untersuchungsperiode parallel, doch besteht durchwegs ein Niveauunterschied von 2 bis 3 Indexpunkten. Die traditionelle Saisonbereinigung des

Die vom WIFO erstellte und die auf einem Modellansatz basierende saisonbereinigte Reihe der Industrieproduktion scheinen auf den ersten Blick nicht sonderlich zu divergieren. Eine genauere Analyse und vor allem ein Vergleich der irregulären Komponenten bringt jedoch signifikante Unterschiede zutage.

WIFO scheint das Niveau der Industrieproduktion systematisch zu unterschätzen. Zudem ist die aus einem Modellansatz berechnete Reihe etwas glatter als die WIFO-Reihe, wie man aus einem Vergleich der irregulären Komponenten deutlich erkennt.

Prognose

Übersicht 4 liefert Anhaltspunkte für die Prognosegenauigkeit der beiden ARIMA-Modelle für die Industrieproduktion. Da monatliche Prognosen der Industrieproduktion aus anderen

Die Qualität der mit einem um Kalendereffekte erweiterten ARIMA-Modell erstellten Prognosen ist ausgezeichnet. Sie sind den Vorhersagen anhand eines einfachen Modells deutlich überlegen.

Quellen nicht vorliegen, dient das ursprüngliche ARIMA-Modell als Vergleichsbasis für das erweiterte Modell. Die verglichenen Prognosen sind echte Ex-ante-Vorhersagen. Beginnend mit Jänner 1980 werden Prognosen der Industrieproduktion für einen bzw. sechs Monate voraus errechnet. Für die einstufigen Prognosen wird das Modell z. B. bis Dezember 1979 geschätzt und der Wert für Jänner 1980 vorausgesagt, für sechsstufige Vorhersagen nur bis Juli 1979, wenn der Wert für Jänner 1980 prognostiziert werden soll. So werden bis August 1987 jeweils 92 prognostizierte Werte des Produktionsindex mit

Prognoseverhalten

Übersicht 4

	Prognosegenauigkeit	
	Einfaches ARIMA-Modell	ARIMA-Modell mit Kalendereffekten
1 Monat voraus		
r^2	0,395	0,909
Mittlerer quadratischer Fehler	0,076	0,029
Ungleichheitskoeffizient	0,637	0,064
Regressionskoeffizient	0,777	1,050
U^M	0,000	0,000
U^R	0,051	0,022
U^D	0,949	0,978
6 Monate voraus		
r^2	0,516	0,923
Mittlerer quadratischer Fehler	0,067	0,026
Ungleichheitskoeffizient	0,492	0,078
Regressionskoeffizient	0,891	1,007
U^M	0,000	0,000
U^R	0,016	0,001
U^D	0,984	0,999

einem Prognosehorizont von einen bzw. sechs Monaten errechnet. Diese beiden Stichproben bilden die Basis für die Berechnung der Theil'schen Maße der Prognosegüte²⁾.

Das um Kalendereffekte erweiterte Modell ist dem ursprünglichen ARIMA-Modell sowohl bei einstufigen als auch bei sechsstufigen Vorhersagen eindeutig überlegen (Übersicht 4): Der Korrelationskoeffizient zwischen vorhergesagten und tatsächlichen Änderungen der Industrieproduktion ist für das erweiterte Modell mit 0,9 fast doppelt so groß wie für das einfache Modell. Umgekehrt ist die Wurzel aus dem mittleren quadratischen Prognosefehler, der über die Größenordnung der Vorhersagefehler Aufschluß gibt, für das einfache ARIMA-Modell fast dreimal so groß wie für das erweiterte. Die Unterschiede in den Ungleichheitskoeffizienten, die das Prognoseverfahren an einer naiven „No-change“-Extrapolation messen, sind noch ausgeprägter, und auch hier ist das erweiterte Modell eindeutig überlegen. Regressiert man die tatsächlichen Änderungen der Industrieproduktion auf die prognostizierten, so erhält man einen Regressionskoeffizienten von 1 — ein Beleg für perfekte Vorhersagen. Eine Zerlegung des mittleren quadratischen Fehlers in Komponenten zeigt, daß keine Gruppe der getesteten Prognosen durch eine lineare Transformation verbessert werden könnte. Für die guten Vorhersagen des erweiterten Modells

¹⁾ Das WIFO gewichtet für die Saisonbereinigung die Originalreihe und eine um die Zahl der Arbeitstage bereinigte Reihe der Industrieproduktion im Verhältnis 7 zu 3 zusammen und bereinigt die so entstehende Reihe mit CENSUS X-11.

²⁾ Eine detaillierte Beschreibung dieser Maßzahlen findet der interessierte Leser in Theil (1966).

hat dieses Ergebnis keine Bedeutung. Für die schlechten Prognosen des ursprünglichen Modells jedoch ist es sehr enttäuschend, weil eine Verbesserung angebracht wäre. Das um Kalendereffekte erweiterte ARIMA-Modell prognostiziert also auf kurze Sicht überaus zufriedenstellend, und mehr sollte man von einem Zeitreihenmodell nicht erwarten. Weil die Wochentagszusammensetzung des Kalenders für die Zukunft bekannt ist, kann das Modell auch in der Praxis für die Prognose eingesetzt werden. Die erforderlichen Anpassungen sollten geringen Aufwand verursachen, weil sich die Zusammensetzung wiederholt und man daher Erfahrungen aus der Vergangenheit gezielt einsetzen kann.

Zusammenfassung

Die Industrieproduktion weist wegen der von Monat zu Monat variierenden Zahl der Arbeitstage starke monatliche Schwankungen auf. Die in Österreich relativ hohe Zahl von gesetzlichen Feiertagen verursacht zusätzliche Probleme. Saisonbereinigung und vor allem Prognose des Produktionsindex werden durch solche Kalendereffekte sehr erschwert. Eine einfache Bereinigung um die Zahl der Arbeitstage, wie sie das Österreichische Statistische Zentralamt vornimmt, kann die Reihe nicht ausreichend glätten.

Die vorliegende Arbeit zeigt, daß es möglich ist, diese Kalendereffekte wirksam auszuschalten. Ausgangspunkt der Bereinigung ist ein Kalender, der die für Österreich charakteristische Wochentagszusammensetzung der einzelnen Monate eines Jahres wiedergibt. Diese Wochentagszusammensetzung hat nämlich entscheidenden Einfluß auf das Produktionsvolumen eines Monats: An Samstagen, Sonn- und Feiertagen ist die Produktionstätigkeit bekanntlich viel geringer als an Werktagen, und sie variiert auch zwischen den einzelnen Werktagen. Diese Tatsache kann man für die Kalendereinigung der Industrieproduktion nutzen, indem man die Produktion eines Monats zur Zahl der Montage, Diensttage usw. in diesem Monat in Beziehung setzt. Wenn auf einen bestimmten Werktag ein Feiertag fällt, ist er vernünftiger-

weise als „Sonntag“ zu zählen. So kommt man zur tatsächlichen Zahl der Werktage in jedem Monat. Die Tatsache, daß gern an Feiertage Urlaubstage angeschlossen werden, ist etwas schwerer in den Griff zu bekommen. Auch dieses Problem kann aber durch gezieltes Probieren relativ rasch gelöst werden. Allerdings führt man dadurch ein Element von Willkür in die Analyse ein. Mit etwas zusätzlichem Arbeitsaufwand könnte diese subjektive Komponente jedoch deutlich reduziert werden.

Ein so adaptierter Kalender kann erfolgreich zur Ausschaltung störender Sondereinflüsse aus der Industrieproduktion verwendet werden. Im Gegensatz zur Originalreihe verläuft die korrigierte Reihe sehr regelmäßig und weist vor allem ein stabiles Saisonmuster auf. Die Saisonbereinigung bereitet dann keine Schwierigkeiten mehr. Überdies läßt sich die korrigierte Reihe durch ein relativ einfaches ARIMA-Modell hinreichend genau beschreiben. Dieses Modell kann zur Prognose des Produktionsindex herangezogen werden, wobei eine beachtliche Prognosegenauigkeit zu erwarten ist.

Anhang

ARMA-Modelle

Stationarität ist eine Grundvoraussetzung dafür, daß ein stochastischer Prozeß durch ein autoregressives Moving-Average-Modell (ARMA-Modell) dargestellt werden kann. Ein Prozeß ist stationär, wenn Mittelwert und Varianz konstant sind und die Autokovarianzen

$$\begin{aligned}\gamma_k &= \text{Cov}(y_t, y_{t-k}) = \\ &= E(y_t - \mu)(y_{t-k} - \mu)\end{aligned}$$

und Autokorrelationen

$$\rho_k = \frac{\text{Cov}(y_t, y_{t-k})}{[\text{Var}(y_t) \text{Var}(y_{t-k})]^{1/2}} = \frac{\gamma_k}{\gamma_0}$$

nur von der Zeitdifferenz k abhängen. Die Autokorrelationen (als Funktion dieser Zeitdifferenz k betrachtet) werden in der Literatur als Autokorrelationsfunktion bezeichnet.

Ein stationärer Prozeß y_t kann als ARMA-Modell formuliert werden:

$$\begin{aligned}y_t - \varphi_1 y_{t-1} - \dots - \varphi_p y_{t-p} = \\ = a_t + \vartheta_1 a_{t-1} + \dots + \vartheta_q a_{t-q}\end{aligned}$$

In diesem Ausdruck bezeichnet $\{a_t\}$ eine Folge von unkorrelierten Zu-

fallsvariablen mit Mittelwert 0 und Varianz σ^2 . Eine solche Folge wird als „weißes Rauschen“ bezeichnet. Unter Verwendung des Verschiebungsoperators $B^k y_t = y_{t-k}$ lautet der Ausdruck:

$$\begin{aligned}(1 - \varphi_1 B - \dots - \varphi_p B^p) y_t = \\ = (1 + \vartheta_1 B + \dots + \vartheta_q B^q) a_t\end{aligned}$$

oder, noch kompakter,

$$\varphi(B) y_t = \vartheta(B) a_t$$

Die Koeffizienten φ_i und ϑ_i können aus den beobachteten Werten geschätzt werden. Dazu muß man aber zuerst die Ordnung der Polynome $\varphi(B)$ und $\vartheta(B)$ bestimmen. Dies geschieht durch Inspektion der Autokorrelationsfunktion (ACF) und der partiellen ACF, die je nach Grad dieser Polynome verschiedene Gestalt haben. Die ACF wurde bereits zu Beginn eingeführt. Die PACF mißt die zusätzliche Korrelation zwischen y_t und y_{t-k} , nachdem der Einfluß der dazwischenliegenden Variablen $y_{t-1}, \dots, y_{t-k+1}$ ausgeschaltet wurde. Reine AR- bzw. MA-Prozesse sind relativ leicht zu identifizieren, weil ACF und PACF für sie eine genau definierte Gestalt haben. Die ACF eines AR(p)-Prozesses ist eine Kombination gedämpfter Exponential- und Sinusschwingungen. Die PACF wird Null für Verzögerungen größer als die Ordnung p des Prozesses. Für MA(q)-Prozesse gilt umgekehrt: Die ACF bricht für Verzögerungen größer als q ab; die PACF ist eine Kombination von Exponential- und Sinusschwingungen. Die Identifikation von gemischten ARMA(p, q)-Prozessen gestaltet sich etwas schwieriger, aber auch dafür gibt es inzwischen geeignete Methoden.

Viele ökonomische Zeitreihen sind leider nicht stationär und unterliegen überdies starken Saisonschwankungen. Für diese Fälle sind die multiplikativen saisonalen autoregressiven integrierten Moving-average-Modelle (ARIMA-Modelle) vorgesehen:

$$\begin{aligned}\varphi(B) \Phi(L) (1 - B)^d (1 - L)^D y_t = \\ = \vartheta(B) \Theta(L) a_t\end{aligned}$$

Hier ist $L = B^s$, s bezeichnet die Periode der Saisonalität, d. h. $s=4$ für Quartalsdaten und $s=12$ für Monatsdaten. ARIMA-Modelle unterscheiden sich von ARMA-Modellen also durch die Präsenz von zusätzlichen Differenzenoperatoren, die Stationarität der zu modellierenden Zeitreihen si-

chern sollen Ihre Ordnung ist in der Regel $d=D=1$

Wenn die Ordnung der einzelnen Polynome festgelegt ist, kann das Modell geschätzt werden. Anschließend wird getestet, ob die Residuen des geschätzten Modells tatsächlich unkorrelierte Zufallsfehler sind. Dafür verwendet man die Q -Statistik von *Ljung — Box* (1978). Übersteigt sie einen kritischen Wert, so ist die Null-Hypothese zu verwerfen. Man muß in einer neuen Identifikationsphase versuchen, die Modellspezifikation zu verbessern³⁾

Saisonbereinigung auf der Basis eines Modellansatzes

Obwohl saisonbereinigte Daten in der laufenden Konjunkturanalyse häufig verwendet werden, weiß man über die Eigenschaften der zu ihrer Berechnung verwendeten Methoden oft nur unzureichend Bescheid. CENSUS X-11, die am häufigsten eingesetzte Methode, ermittelt saisonbereinigte Werte mit Hilfe gleitender Durchschnitte. Die einzige Rechtfertigung für dieses Vorgehen scheint zu sein, daß es in der Mehrzahl der Fälle zu brauchbaren Resultaten führt. Eine fundierte theoretisch-statistische Begründung für die Verwendung von gleitenden Durchschnitten gibt es hingegen nicht. Die auf einem Modell-

ansatz basierenden Saisonbereinigungsverfahren sollen diesen Mangel beheben.

Ziel jeder Saisonbereinigung ist es, eine beobachtete Zeitreihe y_t in Trend (T_t), Saison (S_t) und irreguläre Komponente (I_t) zu zerlegen:

$$(3) \quad y_t = T_t + S_t + I_t.$$

Durch die Anwendung von Transformationen auf die Originalreihe läßt sich dieser restriktiv erscheinende Ansatz beliebig erweitern. Im Gegensatz zu empirischen Methoden der Saisonbereinigung, wie CENSUS X-11, werden aber in dem auf einem ARIMA-Modell basierenden Ansatz der Trend, die Saison und die irreguläre Komponente exakt definiert. Man unterstellt, daß sie ARIMA-Modellen folgen:

$$(4) \quad \begin{aligned} \Phi_T(B) T_t &= \Theta_T(B) b_t, \\ \Phi_S(B) S_t &= \Theta_S(B) c_t, \\ \Phi_I(B) I_t &= \Theta_I(B) d_t, \end{aligned}$$

b_t , c_t und d_t bezeichnen voneinander unabhängig verteilte Zufallsfehler jeweils mit Mittelwert 0 und Varianz σ_b^2 , σ_c^2 und σ_d^2 . Wenn die Polynome (4) bekannt wären, könnte man Methoden der Signalextraktion verwenden, um Schätzwerte für T_t , S_t und I_t zu erhalten. Diese Polynome sind in der Regel nicht bekannt, und daher ist eine direkte Schätzung von Trend, Saison und irregulärer Komponente

nicht möglich. Man kann aber das ARIMA-Modell der beobachteten Reihe y_t schätzen. Unter gewissen A-priori-Restriktionen kann dieses Modell dazu dienen, die oben erwähnten Polynome zu bestimmen. Danach helfen Methoden der Signalextraktion, Schätzwerte für Trend, Saison und irreguläre Komponente zu berechnen⁴⁾

Literaturhinweise

- Abraham B.**, **Ledolter J.** Statistical Methods for Forecasting. Wiley, New York, 1983.
- Bell W. R.**, **Hillmer S. C.** „Modelling Time Series with Calendar Variation“. University of Kansas School of Business Working Paper, 1981.
- Box G. E. P.**, **Jenkins G. M.**, Time Series Analysis: Forecasting and Control. Holden Day, San Francisco, 1970.
- Burman, J. P.** „Seasonal Adjustment by Signal Extraction“. The Journal of the Royal Statistical Society, Series A, 1980, 143(3), S. 321-337.
- Hillmer S. C.**, „Forecasting Time Series with Trading Day Variation“. Journal of Forecasting, 1982, 1, S. 385-395.
- Hillmer S. C.**, **Tiao G. C.** „An ARIMA Model-Based Approach to Seasonal Adjustment“. JASA, 1982, 77, S. 63-70.
- Liu L. M.** „Analysis of Time Series with Calendar Effects“. Management Science, 1980, 26, S. 106-112.
- Ljung G. M.**, **Box G. E. P.** „On a Measure of Lack of Fit in Time Series Models“. Biometrika, 1978, 65, S. 297-303.
- Theil H.** Applied Economic Forecasting. North-Holland, Amsterdam, 1966.

³⁾ Eine anschauliche Einführung in den Bau von Zeitreihenmodellen geben u. a. *Box — Jenkins* (1970) und *Abraham — Ledolter* (1983).

⁴⁾ Eine ausführliche Beschreibung von auf einem Modellansatz basierenden Verfahren der Saisonbereinigung findet sich in *Burman* (1980) und *Hillmer — Tiao* (1982).