

## EMPIRISCHE UNTERSUCHUNG

### 7 Problemstellung und Konzeption der Studie

#### 7.1 Der Untersuchungsgegenstand

Diese Studie macht die Kanzlerpräferenz der (west-)deutschen Wahlbevölkerung zum Gegenstand empirischer Analysen. Die Kanzlerpräferenz wird damit zur *abhängigen Variable* – im Gegensatz zu vielen typischen Wahlstudien, die das Wahlverhalten der Bürger zum Erkenntnisgegenstand erheben und dabei die Kandidatenorientierung als eine *unabhängige* Erklärungsgröße verwenden.

Bei der Kanzlerpräferenz interessiert in dieser Studie nicht der persönliche Entscheidungsablauf beim einzelnen Wähler und seine Abhängigkeiten von sozial-psychologischen oder soziologischen Einflussgrößen, sondern vorrangig die Frage, in wie weit die Massenmedien einen messbaren Effekt auf die Kanzlerpräferenz entwickeln können.

Auf der Seite der Massenmedien musste eine forschungsökonomische Fokussierung stattfinden. *Alle relevanten Medieninhalte* zu untersuchen ist theoretisch wie praktisch unmöglich. Für die Urteilsbildung über Kanzlerkandidaten sind jedoch die Hauptnachrichtensendungen der reichweitenstärksten Fernsehsender besonders ausschlaggebend. Ihre Darstellungsmöglichkeiten, ihre Tagesaktualität und hohe Glaubwürdigkeit aus Sicht der Rezipienten sowie ihre große Reichweite (und damit gesellschaftliche Relevanz), sind nur einige der Argumente, warum eine Untersuchung lohnenswert erscheint (vgl. für Glaubwürdigkeit und Reichweite: STUDIE MASSENKOMMUNIKATION 2000, Z.B. RIDDER/ENGEL 2001).

Untersuchungsgegenstand der vorliegenden Studie ist daher "die Darstellung der Kandidaten in den Fernsehnachrichten und ihre Auswirkungen auf die Kanzlerpräferenz in der Bevölkerung".

#### 7.2 Die Untersuchungsperspektive

Der Untersuchungsgegenstand "die Darstellung der Kandidaten in den Fernsehnachrichten und ihre Auswirkungen auf die Kanzlerpräferenz in der Bevölkerung" wird in der vorliegenden Arbeit mit Hilfe von Medieninhaltsanalysen und Bevölkerungsumfragen untersucht (siehe Kapitel "7.4 Die Untersuchungsmethoden"). Die allgemeine Hypothese lautet dabei:

Die Darstellung der Kanzlerkandidaten in den Fernsehnachrichten hat einen Effekt auf die Bewertung der Kandidaten durch die Bevölkerung (Kanzlerpräferenz). Diese Hypothese muss jedoch differenziert werden. Auf der *Medienseite* (1) werden nur bestimmten Inhalten stärkere (und damit messbare) Effekte zugeschrieben und bei den *Kandidaten* (2) werden für Herausforderer und Amtsinhaber jeweils unterschiedliche Zusammenhänge angenommen, da die beiden Kandidaten mit sehr unterschiedlichen Voraussetzungen "ins Rennen" gehen.

(1) Dass in den Nachrichten vor allem *negative Meldungen* eine größere Aufmerksamkeit und damit ein höheres Effektpotential erlangen, ist in den Kommunikationswissenschaften Bestandteil unterschiedlicher Betrachtungen, sowohl theoretischer Ansätze wie empirischer Bestätigungen. Bei den Praktikern sind Aussagen wie "only bad news are good news" oder "good news are no news" ein Beleg dafür, dass Journalisten ihre Aufgabe vor allem in der Berichterstattung über problematische Zustände und Entwicklungen sehen. Kommunikationswissenschaftliche Ansätze wie die Nachrichtenwerttheorie oder die Selektivitätsforschung haben immer wieder gezeigt, dass "Negativismus" ein wichtiger Nachrichtenfaktor für den Nachrichtenwert einer Nachricht sind bzw. dass negative Meldungen – im Gegensatz zu positiven – Selektionshürden bei der Nachrichtenrezeption besser überwinden können. So hat DONSBACH (1991) zeigen können, dass die Selektion von Nachrichten im Sinne der kognitiven Dissonanz nach FESTINGER (1957) überwiegend für positive (oder neutrale) Inhalte zutrifft. Bei stärker negativen Meldungen werden diese "Schutzbarrieren" der Rezipienten jedoch überwunden und die (Vor-)Einstellungen der Rezipienten wirken sich nicht mehr auf die Rezeption aus. Auf Nachrichten über die Kanzlerkandidaten bezogen heißt das: positive Meldungen über einen Kandidaten werden vornehmlich von seinen Anhängern aufgenommen, negative Aussagen jedoch von allen Rezipienten – nicht nur von seinen Gegnern.<sup>57</sup> Positive Nachrichten können auf das Gesamt-Elektorat also weniger Einfluss entwickeln als negative. Daher lautet die Hypothese für diese Untersuchung nach der ersten Differenzierung: Negative Darstellungen der Kanzlerkandidaten in den Fernsehnachrichten wirken sich negativ auf die Kanzlerpräferenz in der Bevölkerung aus.

---

57 Beim Medium Fernsehen und speziell bei Fernsehnachrichten ist Selektion zudem ein geringeres Problem für Medienwirkungen, da die Ausweichmöglichkeiten des Rezipienten – z.B. im Gegensatz zur Tageszeitung – deutlich beschränkt sind, was vor allem zwei Gründe hat: Die Rezeption kann nur zeit-synchron stattfinden und die Sendeform Nachrichten besteht aus vielen kurzen Beiträgen. Beim "Weg-Zappen" (als typische Selektionsform beim Fernsehen) läuft der Rezipient also Gefahr, andere Inhalte zu verpassen. Außerdem sind Valenzen von Beiträgen für den Rezipienten oft schwerer bzw. erst später im Rezeptionsprozess zu erkennen, so dass Selektion bei Fernsehnachrichten insgesamt unwahrscheinlicher ist.

(2) Die Medieninhalte besitzen jedoch auch ein unterschiedliches Effektpotential je nach Adressat. Der *Amtsinhaber* ist allen Rezipienten seit vielen Jahren sehr gut bekannt. Sie kennen ihn aus zahlreichen Medienauftritten, wo auch seine Arbeit, Amtsführung und Persönlichkeit vielfach Gegenstand der Berichterstattung war. Außerdem hat jeder Rezipient eine Leistungsbilanz des Amtsinhabers im Kopf: gesellschaftliche Zustände, für die er den Kanzler als Chef der Exekutive verantwortlich macht. Hauptkriterium ist hier die wirtschaftliche Lage des Landes (siehe detailliert im Kapitel "4.3 Zum Einfluss der Wirtschaftslage"). Der *Herausforderer* ist dagegen weitaus weniger bekannt. Zwar hat auch er als Oppositionsführer im Bundestag und/oder Chef der größten Oppositionspartei einen höheren Bekanntheitsgrad, jedoch beschränkt sich die Bekanntheit oft auf einige oberflächliche Daten (Name, Herkunft etc.). Als politisch Handelnder und Verantwortlicher ist er oft nur regional bekannt (z.B. Scharping als Ministerpräsident in Rheinland-Pfalz) und daher müssen sich die Rezipienten über diesen Politiker als möglichen zukünftigen Kanzler erst ein Urteil bilden. Bei diesem Urteil sind sie in hohem Maße von den Informationen der Medien abhängig (siehe detailliert im Kapitel "4.1 Der Prozess der Urteilsbildung über Spitzenkandidaten"). Aus dieser ungleichen Rollenverteilung ergeben sich wichtige Konsequenzen für die Aufdeckung möglicher Medieneffekte. Es werden daher zwei differenzierte Hypothesen formuliert:

*Hypothese 1:* Die Kanzlerpräferenz des Herausforderers Rudolf Scharping wird von negativen Nachrichteninhalten negativ beeinflusst. Das gilt insbesondere auch, da der Untersuchungszeitraum in der frühen Phase der Kanzlerkandidatur Scharpings liegt, also in einer Zeit, in der sich die Rezipienten ihr erstes Urteil über den Herausforderer bilden.

*Hypothese 2:* Die Kanzlerpräferenz des Amtsinhabers Helmut Kohl wird *nicht* von den Nachrichteninhalten beeinflusst. Auch diese Hypothese bezieht sich natürlich nur auf den Untersuchungszeitraum. Es soll damit nicht behauptet werden, dass die Meinung der Bürger über Kohl vollkommen unabhängig von der Berichterstattung über ihn in den Nachrichten wäre. Jedoch liegen den Rezipienten bereits so viele Informationen und Meinungen über Kohl vor (medial wie extramedial), dass sie sich eine eigene Meinung über den Amtsinhaber bereits längst gebildet haben und diese Meinung von den Nachrichteninhalten einiger Wochen (Untersuchungszeitraum) kaum verändert werden kann. Dazu müssten in dieser Zeit außergewöhnliche Ereignisse stattfinden (die nicht stattgefunden haben). Die Kanzlerpräferenz für Kohl ist dagegen eher von den Einschätzungen der Rezipienten be-

züglich der wirtschaftlichen Entwicklungen der Bundesrepublik abhängig, da er als Kanzler hierfür verantwortlich gemacht wird.

Zur Überprüfung dieser beiden Hypothesen wird zuerst die Bevölkerungsmeinung sowie die Medienberichterstattung über Kohl und Scharping in den Nachrichten deskriptiv-analytisch dargestellt (Kapitel 8.1 "Kanzlerpräferenzen der Bevölkerung" und 8.2 "Kandidaten in den Fernsehnachrichten"). Anschließend werden mit Hilfe von Zeitreihenanalysen die Auswirkungen der negativen Aussagen über Scharping auf seine Kanzlerpräferenz (Kapitel 9.1 "Herausforderer Scharping"), bzw. die Auswirkungen der Einschätzungen der wirtschaftlichen Perspektiven der BRD auf die Kanzlerpräferenz Kohls untersucht (Kapitel 9.2 "Amtsinhaber Kohl").

### **7.3 Der Untersuchungszeitraum**

Um Medieneffekte mit empirischen Untersuchungen aufdecken zu können, müssen die beteiligten Faktoren – im Bereich der politischen Kommunikation z.B. Medienberichterstattung und Bevölkerungsmeinung – über die Zeit hin deutlich unterschiedliche Werte annehmen, da es bei relativ konstanter Berichterstattung und/oder Bevölkerungsmeinung kaum Möglichkeiten gibt, Medieneffekte mit statistischen Analysen "dingfest" zu machen. ZALLER hat dies anschaulich beschrieben (siehe Kapitel 5.3 "Empirische und analytische Randbedingungen

für den Nachweis starker Medieneffekte") und darauf hingewiesen, dass scheinbare Null-Effekte auf der Meinungsseite durchaus das Ergebnis starker Medienwirkungen sein können, nämlich dann, wenn konkurrierende Medieninhalte sich in ihrer Wirkung gegenseitig ausgleichen (ZALLERS "Tau-Zieh-Metapher"). Technisch gesprochen benötigt man also Varianz in den Variablen der Analysemodelle. Diese scheinbar banale Anforderung – ohne Varianz gibt es auch keinen "Anteil erklärter Varianz" ( $R^2$ ) und damit keine empirische Unterstützung der Annahme kausaler Zusammenhänge – lässt jedoch viele empirische Studien ungeeignet erscheinen, Medienwirkungen überhaupt aufspüren zu können. Typische Wahlstudien beziehen sich auf die letzten vier bis maximal zwölf Wochen vor dem Wahltermin, da hier die "heiße Wahlkampfphase" vermutet wird, in der einerseits die Parteien sich besonders stark um Medienpräsenz bemühen und andererseits das Mediensystem selbst eine erhöhte Aufmerksamkeit gegenüber politischen Akteuren aufweist. Und auch

die Wähler – so könnte man vermuten – haben eine umso höhere Aufmerksamkeit gegenüber wahlrelevanten Medieninhalten, je näher der Wahltermin rückt. Aufgrund dieser Annahmen müsste der "Medieninput", also die wahl- und wahlkampfbezogenen Inhalte in den Medien, in dieser Zeit besonders hoch sein. Dies trifft für die absolute Menge der Medienberichterstattung auch zu, für die Varianz der Inhalte jedoch gerade nicht. Verschiedene empirische Studien (siehe Kapitel "4.2 Zum Einfluss des Fernsehens") haben gezeigt, dass mit näher rückendem Wahltermin auch die Berichterstattung über die politischen Akteure ausgeglichener wird. Dafür sind verschiedene Nachrichtenfaktoren verantwortlich. Bei der Berichterstattung über die Kanzlerkandidaten ist zum Beispiel der viel zitierte "Kanzlerbonus" kurz vor der Wahl meist weniger ausgeprägt (oder gar vollkommen verschwunden) als längere Zeit vor dem Wahltermin, z.B. wenn der Herausforderer gewählt wird und erste Wahlkampfaktivitäten stattfinden. Deutschen Journalisten "riecht" die Berichterstattung über Aktivitäten des Herausforderers lange vor dem Wahltermin scheinbar zu sehr nach "Verlautbarungsjournalismus", da der Ereignishintergrund des Herausforderers oft "nur Wahlkampf" ist. Aktivitäten des Amtsinhabers scheinen sich dagegen auf politisches Handeln zu beziehen und werden daher eher berichtet. Journalistische Gebote wie Ausgewogenheit oder Chancengleichheit kommen zu dieser Zeit noch nicht so sehr zum Tragen. Je näher die Wahl rückt desto stärker rücken die Kandidaten jedoch ins Zentrum der Berichterstattung und die Ausgewogenheit der Berichterstattung (bezogen auf die quantitative mediale Aufmerksamkeit für die beiden Kandidaten) nimmt zu. Dazu tragen auch noch andere Nachrichtenfaktoren bei, wie z.B. Spannung, wenn der Ausgang der Wahl noch sehr ungewiss ist. Wenn also keine besonderen Ereignisse in der "heißen Wahlkampfphase" geschehen, so ist die Zeit der starken Medienberichterstattung eine besonders *schlechte* Zeit, um die Wirkung von Medieninhalten aufzudecken, einfach weil – in ZALLERS Terminologie – alle Botschaften in den Medien gleich "laut" sind und somit kein Vorteil für eine Seite entsteht und sich konkurrierende Inhalte ausgleichen. Dieses scheinbare Paradoxon (keine Medienwirkungen bei besonders vielen Medieninhalten) beraubt daher viele Wahlstudien der Möglichkeit, überhaupt Medieneffekte aufdecken zu können (obwohl sie vielleicht vorhanden sind). Nur selten gibt es in den letzten Wochen vor einer Wahl die Situation, dass eine politische Seite besondere Vorteile aus der Nachrichtenlage ziehen kann. Im Wahlkampf 2002 gab es scheinbar eine solche Situation. Zwar liegt aufgrund der Kürze der Zeit noch nicht ausreichende empirische Evidenz vor, jedoch weisen bisherige Beobachtungen darauf hin, dass Gerhard Schröder bei zwei wichtigen Themen, die in der heißen Wahlkampfphase im Zentrum der Berichterstattung standen, komparative Vorteile gegenüber

seinem Herausforderer Edmund Stoiber hatte: Bei den Themen "Oder-Flut" und "Militärische Beteiligung bei einem eventuellen Irak-Konflikt" unterstützte eine Mehrheit der Bevölkerung die Aktivitäten und Aussagen Gerhard Schröders. Aufgrund des Priming-Effekts (nach dem Priming-Konzept wird die Gesamtbewertung eines Kandidaten vor allem anhand der aktuellen Medienthemen vorgenommen), konnte Gerhard Schröder aus der beschriebenen Nachrichtenlage einen positiven Effekt für die Bewertung seiner Person – und damit letztlich auch für die Wahlentscheidung der Wähler – ziehen. Solche Situationen sind in der heißen Wahlkampfphase jedoch äußerst selten, auch deshalb, weil sich Parteistrategen und beauftragte Medien-Experten um die Themen-Agenda der Medien "kümmern" (aus eben diesem Grunde). Im Normalfall ist die Berichterstattung der letzten Wochen vor dem Wahltermin zwar besonders umfangreich, aber auch besonders ausgewogen. Diese Ausgewogenheit bezieht sich dabei sowohl auf die reine Menge der Berichterstattung, die jedem Akteur zugebilligt wird, als auch auf explizite Wertungen, die mit näher rückendem Wahltermin ebenfalls abnehmen (vielleicht weil Journalisten um ihre besondere Verantwortung und/oder Beobachtung durch Politik und Wissenschaft in dieser Zeit wissen).

Aufgrund dieser Überlegungen ist bereits die Auswahl des "richtigen" Untersuchungszeitraums eine analytische Entscheidung – wobei "richtig" nicht manipulativ zu sehen ist, also dass etwa ein einzelner Medieneffekt, den man zu einer bestimmten Zeit besonders offensichtlich "sehen" kann, als Beleg für generelle, allgemeingültigere Medienwirkungen gesucht und missbraucht wird! Ein angemessener Untersuchungszeitraum ist vielmehr ein Zeitabschnitt, in dem deutliche Varianzen in den Variablen auftreten. Diese Varianzen können dann auf ihre Zusammenhänge hin untersucht werden und z.B. Veränderungen in der Bevölkerungsmeinung auf Medienhalte zurückgeführt werden (wobei es sich dabei natürlich immer nur um Partialeffekte handelt). Einen "guten" Untersuchungszeitraum zu finden, ist jedoch im Vorhinein einer Untersuchung kaum machbar. Leider besitzt kein Forscher die hellseherische Gabe zu wissen, wann die Berichterstattung oder die Bevölkerungsmeinung interessanten Veränderungen unterworfen sein wird. Daher kann ein Untersuchungszeitraum – für meine Zwecke – nur retrospektiv festgelegt werden, wenn bestimmte Daten über diese Zeit bereits vorliegen.

Für meine Analysen nutze ich daher die Daten aus dem DFG-Projekt "Politikverdrossenheit und Wählerabwanderungen" in dem *alle Hauptnachrichtensendungen des Jahres 1994* von ARD, ZDF, RTL und SAT.1 aufgezeichnet und inhaltsanalytisch untersucht wurden (RTL und SAT.1 allerdings nur im vierwöchentlichen Wechsel). Dadurch kann ich einerseits in einer Art natürlichem Experiment rückwirkend einen Untersuchungszeitraum festlegen, in

dem wichtige Variablen stärkere Varianzen aufweisen. Andererseits sind inhaltsanalytisch vertiefende Nachkodierungen des Ursprungsmaterials möglich, da auch das Ursprungsmaterial (die täglichen Hauptnachrichtensendungen) für Nacherhebungen zur Verfügung steht.

Als endgültiger Untersuchungszeitraum wurden 10 Wochen im Frühjahr/Sommer des Jahres 1994 gewählt, da hier die stärksten Veränderungen in der Bevölkerungmeinung stattfanden<sup>58</sup>: 18.4. - 26.6.1994.

Neben der Varianz der Kanzlerpräferenzen spricht aber noch ein anderes inhaltliches Argument für diesen "frühen" Untersuchungszeitraum:

Mehrere Studien über Medieninhalte und/oder Wahlkampfeffekte der letzten Bundestagswahlen haben gezeigt, dass Meinungen und Berichterstattung über die Kanzlerkandidaten in den letzten drei Monaten – und teilweise schon früher – relativ konstant und kaum Schwankungen unterlegen waren (z.B. HAGEN/ZEH/BEHRENS 1994, KEPPLINGER/RETTICH 1996). Untersuchungen der Kanzlerpräferenzen weiter vor den eigentlichen Wahlen zeigen jedoch beträchtliche Meinungsveränderungen auf (z.B. KINDELMANN 1994). Den typischen Wahlstudien fehlt für diese frühe Zeit jedoch der Medieninput, da nur die letzte "heiße Wahlkampfphase" inhaltsanalytisch festgehalten wurde. Vor allem für den Herausforderer sind die ersten größeren Auftritte als designierter Kanzlerkandidat jedoch entscheidend, da sich die Rezipienten über den eher unbekanntem Kandidaten in dieser Zeit sehr intensiv ein Urteil bilden, dass oft nur schwer wieder zu verändern ist. Der gewählte Zeitraum ist auch aus dieser Perspektive angemessen, da Scharping bereits der designierte Kanzlerkandidat der SPD war und erste Wahlkampfaktivitäten vollzog. Am Ende des Untersuchungszeitraums wurde er dann offiziell von seiner Partei zum Kanzlerkandidaten gewählt (Parteitag in Halle am 22.6.1994).

---

58 ZALLER geht bei der Auswahl des richtigen Zeitraums von der Varianz in der *unabhängigen* Variable aus, während ich den Zeitraum aufgrund der Varianz in der *abhängigen* Variable festlege, da in meiner Untersuchung die unabhängigen Variablen erst noch durch die Nachkodierung der Fernsehnachrichten erhoben werden.

## 7.4 Die Untersuchungsmethoden

Die empirischen Analysen der vorliegenden Arbeit beruhen auf zwei Datenquellen. Zum einen werden Befragungsdaten aus einem DFG-Wahlforschungsprojekt sekundäranalytisch ausgewertet, zum anderen wird Ursprungsmaterial dieses Projekts – die Hauptnachrichtensendungen der vier reichweitenstärksten Sender ARD, ZDF, RTL und SAT.1 – einer eigenen inhaltsanalytischen "Nachkodierung" unterzogen, da für die spezielle Fragestellung dieser Arbeit Nachrichteninhalte benötigt wurden, die in der Primärerhebung nicht gemessen wurden.

### 7.4.1 Die Befragungsdaten aus dem DFG-Projekt "Wählerwanderung und Politikverdrossenheit"

Die Bevölkerungsdaten für meine Analysen stammen aus dem 1994er DFG-Wahlforschungsprojekt "Wählerwanderung und Politikverdrossenheit".<sup>59</sup> Im Rahmen dieses Projekts im Mehrmethoden-Design befragte das FORSA-Institut das gesamte Jahr 1994 pro Werktag (Montag bis Freitag) jeweils 500 repräsentativ ausgewählte Personen (ADM-Telefon-Mastersample, GÜLLNER 2001: 568). Damit stehen kontinuierlich (werktätlich) eine Reihe von Indikatoren politischer Einstellungen und Verhaltensweisen zur Verfügung, unter anderem die Kanzlerpräferenz und die allgemeinen Wirtschaftserwartungen. Die Vorteile dieser Erhebungsform ("forsa-Wahl-Tracking") sind vor allem "die differenzierte, zeitnahe Erfassung der Entwicklung der öffentlichen Meinung" (GÜLLNER 2001: 568). Damit sind vor allem politische Stimmungsänderungen sehr viel genauer nachvollziehbar. GÜLLNER vergleicht die vorliegende Datenstruktur mit den monatlichen Daten der Forschungsgruppe Wahlen (FGW) für das bekannte ZDF-Politbarometer. In einer Reihe von Beispielen (GÜLLNER 2001: 568-572) kann er zeigen, dass weniger zeitnah erhobene Daten nicht oder nur verzögert in der Lage sind, deutliche Stimmungswechseln nachzuzeichnen. Für eine Kausalanalyse täglicher Medieninhalte wäre dies jedoch ein elementarer Nachteil.

Die Frage nach der Kanzlerpräferenz "Wenn Sie sich für einen der beiden Kanzlerkandidaten entscheiden müssten, wen würden Sie präferieren?" ist in den Kausalanalysen die Basis für die abhängige Variable. Da es als Kandidaten nur Helmut Kohl oder Rudolf

---

59 An dieser Stelle werden nur die Teile des Projekts erläutert, die für die Sekundäranalyse relevant sind. Für eine genauere Beschreibung des Projekts vgl. FRIEDRICHSEN (1997) oder GÜLLNER (2001: 568).



Scharping zu wählen gab, ist der Gewinn des einen Kandidaten meist auch ein Verlust des Gegenkandidaten. Die Variable ist aber weniger dichotom, als es auf den ersten Blick erscheint und Analysen mit dieser Variable daher auch weniger mit den typischen Problemen dichotomer Variablen behaftet (Informationsredundanz der zweiten Alternative): Mehr als ein Drittel der Befragten gibt bei den Befragungen keinen Kandidaten an, sondern verteilt sich auf die zusätzlichen Antwortalternativen "keinen von beiden" und "weiß nicht". Damit kommen auch Phasen vor, in denen die Präferenzwerte *beider* Kandidaten steigen oder fallen. Die Präferenzwerte für Kohl und Scharping verhalten sich nicht vollkommen komplementär zueinander.

#### **7.4.2 Die Inhaltsanalysedaten der Nachkodierung von Fernsehnachrichten**

Innerhalb des beschriebenen DFG-Forschungsprojekts wurden auch Medieninhaltsanalysen vorgenommen. Dabei wurden unter anderem die täglichen Hauptnachrichtensendungen von ARD, ZDF, RTL und SAT.1 (Montag bis Sonntag) für das gesamte Jahr 1994 einbezogen (die privaten Sender RTL und SAT.1 wurden allerdings nur im vierwöchentlichen Wechsel kodiert). Untersuchungseinheit dieser Inhaltsanalyse war der einzelne, thematisch abgegrenzte Nachrichtenbeitrag. Kategorien, die erhoben wurden, waren neben Platzierung und Präsentationsstil des Beitrags vor allem das Hauptthema des Beitrags sowie die beteiligten Akteure (es wurden bis zu fünf aktive und fünf passive Akteure<sup>60</sup> festgehalten). Damit kann für das gesamte Jahr 1994 ausgesagt werden, über welche Akteure und Themen an welchem Tag in den Nachrichtensendungen Tageschau (ARD), heute (ZDF) und RTL-aktuell/SAT.1-Newsmagazin berichtet wurde. Allerdings waren die Ausprägungen bei den Themen – den Anforderungen der Primärerhebung entsprechend – wenig detailliert, und Akteursbeziehungen und -bewertungen wurden nicht analysiert. So bilden die Daten der Primärerhebung z.B. nicht ab, wer genau mit wem bzw. über wen kommuniziert und vor allem auch nicht, mit welcher Bewertung die Akteure ihre Aussagen versehen. Daher wurden (nach Festlegung des Untersuchungszeitraums) mit Hilfe der Primärdaten diejenigen Nachrichtenbeiträge im Untersuchungszeitraum identifiziert, bei denen mindestens einer der Kanzlerkandidaten als Akteur genannt wurde. Durch die Kodierung der ersten

---

60 "Aktiv" ist ein Akteur in der Primärerhebung dann, wenn er selbst eine Aussage macht oder eine Handlung vollzieht. "Passiv" ist ein Akteur, wenn er Gegenstand einer Aussage oder Handlung ist.

fünf aktiven und der ersten fünf passiven Akteure eines Beitrags in der Primärerhebung sind damit alle relevanten Beiträge über die beiden Kanzlerkandidaten erfasst. Es ist kaum vorstellbar, dass ein Nachrichtenbeitrag Inhalte über einen der Kandidaten enthält und der Kandidat nicht als einer der ersten zehn Akteure (fünf aktive und fünf passive Akteure) genannt wird. Solche theoretisch möglichen Extremfälle umfassen sicherlich nicht relevante Anteile der Berichterstattung über die Kandidaten. Kontrollen der Akteurskodierungen der Primärerhebung konnten keine problematischen Fälle ausfindig machen.

Die auf diese Weise identifizierten Beiträge wurden einer inhaltlich vertiefenden Nachkodierung unterzogen, die im Folgenden dargestellt wird.

#### **7.4.2.1 Untersuchungsanlage**

Für die Inhaltsanalyse wurden das gleiche Anfangs- und Enddatum verwendet wie für den Gesamt-Untersuchungszeitraum: 18.4.-26.6.1994 (Kalenderwoche 16-25). Es wurden zehn Wochen lang täglich die Nachrichtenbeiträge der Tagesschau, der heute-Sendung sowie von RTL-aktuell und SAT.1-Newsmagazin (im vierwöchentlichen Wechsel) auf das Vorkommen eines der Kandidaten hin untersucht. Da an jedem Tag drei Nachrichtensendungen (ARD, ZDF, RTL/SAT.1 im Wechsel) kodiert wurden, belief sich das Ursprungsmaterial auf 210 Nachrichtensendungen (70 Tage Untersuchungszeitraum, 3 Sendungen pro Tag). Die *Analyseeinheit* bildet der Nachrichtenbeitrag, der mindestens einen der Kandidaten als Akteur aufweist, identifiziert über die Akteursvariablen der Primärerhebung (*Zugriffskriterium*). Insgesamt kamen Scharping und/oder Kohl im Untersuchungszeitraum in 155 Beiträgen als Akteure vor. *Kodiereinheit* war jedoch nicht der Beitrag, sondern eine inhaltlich abgegrenzte Sinneinheit: die "*mediale Aussage*" über einen der Kandidaten. Als "*mediale Aussage*" (im Folgenden kurz "*Aussage*") gelten einerseits alle verbalen Äußerungen, bei denen einer der Kandidaten *Subjekt* (Aussagenträger) oder *Objekt* (Bezugsperson) der Äußerung ist. Dazu zählen alle *Zitate* und "*O-Töne*" der Kandidaten selbst, sowie alle *Zitate* und *O-Töne*, in denen andere Akteure einen der Kandidaten ansprechen bzw. thematisieren. Andererseits gelten auch *Handlungsbeschreibungen* der Kandidaten als "*mediale Aussagen*" (Beispiele: "Bundeskanzler Kohl traf sich heute mit dem britischen Premierminister John Major", "Scharping wurde auf der DGB-Kundgebung in Hamburg von Demonstranten mit Eiern beworfen"). Mit dieser *Falldefinition* wird die gesamte Kommunikation von und über die Kandidaten sowie die Darstellung ihrer politischen Handlungen in den Fernsehnachrichten festgehalten. Die Kandidaten können als "*aktive Akteure*" auftreten, wenn sie

selbst per Zitat oder O-Ton zu Wort kommen oder als aktiv Handelnde auftreten, oder sie können als "passive Akteure" Gegenstand der Zitate und O-Töne anderer Akteure bzw. Gegenstand der Handlungen anderer Akteure werden. Wann immer einer der Kandidaten genannt wird (mit Namen oder einer ihrer Funktionsbezeichnungen), entsteht eine Aussage, die als neuer Fall festgehalten und inhaltlich kodiert wird.

Aus technischer Sicht war für diese "inhaltliche Tiefenkodierung" ein häufiges und schnelles Vorwärts- und Rückwärts-Bewegen im Nachrichtenmaterial nötig, weshalb die 155 Beiträge von ihrem ursprünglichen analogen Speichermedium (VHS-Cassetten) per Digitalisierung auf den PC übertragen wurden und so als Video-Datei (mpeg-Format) vorlagen. Für die tatsächliche Kodierung wurden die Beitragsdaten der Primärerhebung in eine Datenbank eingespielt, die es erlaubte, per Knopfdruck (Mausklick) zu den einzelnen Beitragsdaten den jeweils zugehörigen Nachrichtenbeitrag in einem Videofenster auf den Bildschirm zu bringen und dort auch direkt die medialen Aussagen zu kodieren.

Auf diese Weise wurden in den 155 Beiträgen insgesamt **716 mediale Aussagen** identifiziert (4,6 Aussagen pro Beitrag im Durchschnitt).

Die wichtigsten Definitionen und Kategorien der Inhaltsanalyse werden im folgenden Kapitel schematisch dargestellt. Längere Ausprägungslisten (z.B. Akteurslisten, Themenlisten) werden aus Lesbarkeitsgründen nicht in dieser Übersicht dargestellt, sondern in den entsprechenden Tabellen im Ergebnisteil bzw. im Codeplan im Anhang.

#### **7.4.2.2 Kategoriensystem**

##### **Definitionen**

- *Analyseeinheit*: Nachrichtenbeitrag, der mindestens einen der Kandidaten als Akteur aufweist, identifiziert über Akteursvariablen der Primärerhebung (=Zugriffskriterium);
- *Kodiereinheit*: nur *Aussagen* mit Kandidat als Subjekt oder Objekt (=Falldefinition);
- *Aussage*: Einheit von Subjekt, Objekt, Aussage-Typ, Thema und Bewertung (siehe Aussagevariablen); wechselt einer dieser Variablen, entsteht ein neuer Fall.

Da Analyse- und Kodiereinheit in dieser Inhaltsanalyse voneinander abweichen und mit *Beitrag* (Analyseeinheit) und *Aussage* (Kodiereinheit) zwei hierarchisch gegliederte Ebenen darstellen, wurde auch die Kodierung auf diesen zwei Ebenen vorgenommen. Kategorien, die sich auf Merkmale des gesamten Beitrags beziehen, wurden auf *Beitragsebene* kodiert und Variablen, die Eigenschaften jeder Aussage bezeichnen auf *Aussagenebene*.

### **Variablen auf Beitragsebene**

Zusätzlich zu den Beitragsangaben aus der Primärerhebung (ID, Datum, Sendung, Platzierung) wurden lediglich die Gesamtlänge des Beitrags, die Länge der bildlichen Darstellung der Kandidaten sowie das Ereignis des Beitrags festgehalten (Tabelle 5).

**Tabelle 5: Variablen der Fernsehnachrichtenanalyse auf Beitragsebene**

<b>Variable</b>	<b>Definition/Ausprägungen</b>
ID*	Eindeutige Bezeichnung des Beitrags: Kombination aus Datum/Sendung/Platzierung
Datum*	Sendedatum
Sendung*	ARD Tagesschau, ZDF heute, RTL aktuell, SAT.1-Newsmagazin
Platzierung*	Platzierung (lfd. Nummer) innerhalb der Nachrichtensendung
Länge	Länge des Beitrags (Sekunden)
Kohl/Scharping im Bild	Länge, die Kohl/Scharping im Bild (Standbild oder Filmbeitrag; Sekunden)
Beitragsergebnis	Hintergrund des Beitrags, Informationsanlass

\*: vorkodierte Variablen aus dem DFG-Wahlforschungsprojekt

### **Variablen auf Aussagenebene**

Auf Aussagenebene wurde neben der Identifikationsvariable *Aussage-Nummer* zuerst der *Aussage-Typ* kodiert (Tabelle 6). Wie bereits beschrieben wurden Zitate, O-Töne und Handlungsbeschreibungen von bzw. über die Kandidaten als Aussagen gezählt. Eine besondere Kategorie wurde vergeben, wenn der Journalist selbst Subjekt einer Aussage (Aussagenträger) war ("journalistische Aussage"). Dies hat mehrere Hintergründe: Da der Journalist der Produzent des Beitrags ist und damit auch entscheidet, welche Zitate, O-Töne oder Handlungsbeschreibungen im Beitrag vorkommen, könnte man ihn als den Urheber *jeder* Aussage bezeichnen. Diese "allererste Ur-Ebene" des Aussagenträgers soll jedoch nicht kodiert werden, da sie (a) keine Variable ist, sondern eine Konstante (nämlich "Journalist"), (b) es nicht um mögliche "instrumentelle Aktualisierungen" seitens der Journalisten geht (KEPPLINGER/BROSIUS/STAAB 1992), sondern um die Auswirkungen einer gegebenen Medienberichterstattung auf die Meinung der Rezipienten und Bürger. Der Journalist tritt daher nur als eigenständiger Akteur auf, wenn er selbst eine Aussage über einen Kandidaten macht, die keiner anderen Quelle zugeordnet werden kann. Beispiele: (1) Die Aussage

"Kohl kommt bei den Bürgern gut an" ist eine positive Äußerung des Journalisten über Helmut Kohl, die er keiner anderen Quelle zuweist. Der Rezipient sieht sie also als Meinung des Journalisten, und als solche wird sie kodiert. (2) "Scharping schwächelt im Wahlkampf" ist eine negative Einschätzung des Journalisten, die ebenfalls auf keine andere Quelle erkennbar zurückgeht. Die Aussage hat zwar den Charakter einer Handlungsbeschreibung, aber durch die negative Konnotation wird sie als (negative) journalistische Aussage gewertet. Handlungsbeschreibungen sind neutrale Beschreibungen von Kandidaten-Handlungen wie z.B. "Scharping begann heute seine Wahlkampftour durch die neuen Bundesländer".

Mit den folgenden drei Variablen *Subjekt*, *Objekt*, *Bewertung* wurden die Akteursbeziehungen gemessen: Wer sagt was zu/über wen mit welcher Bewertung. Der thematische Kontext jeder Aussage ist in der Variable *Thema* vermerkt.

**Tabelle 6: Variablen der Fernsehnachrichtenanalyse auf Aussagenebene**

<b>Variable</b>	<b>Definition, Ausprägungen</b>
Aussage-Nummer	Laufende Nummer pro Beitrag
Aussage-Typ	Zitat/O-Ton/Handlungsbeschreibung/ journalistische Aussage (nur hier Journalist als Akteur)
Subjekt	Aussageträger (Journalist nur Akteur, wenn er selbst eine Aussage macht bzw. Wertung abgibt)
Objekt	Person, auf die sich die Aussage bezieht
Bewertung	Bewertung des Objekts durch das Subjekt; positiv/negativ
Thema	Thema der Aussage

Ein typisches Problem von Analysen auf Aussagen- oder Argumentebene sind "verschachtelte" Aussagen und vor allem Bewertungen. Wenn der Journalist z.B. meint, Schröder (damals Ministerpräsident in Niedersachsen) würde Scharping nicht leiden können und der Journalist Schröder dafür positiv bewertet, so liegt eine verschachtelte Aussagen- und Bewertungsstruktur vor: Der Journalist lobt Schröder, der wiederum Scharping kritisiert (nach Meinung des Journalisten). Verschachtelungen von Aussagen und Bewertungen sind noch sehr viel komplexer als im angeführten Beispiel vorstellbar, eine Übersicht gibt WEISS (1989). Für die vorliegende Inhaltsanalyse traf das Problem verschachtelter Aussagen jedoch kaum zu: Aufgrund von Fall-/Aussagendefinitionen, die sehr kleine Sinneinheiten festlegen, kamen verschachtelte Aussagen extrem selten vor. Die wenigen auftretenden Verschachtelungen konnten in einzelne Fälle überführt werden. So gehen zwar einige sehr

wenige komplexe Bewertungsstrukturen verloren, ihr Erkenntnisgewinn ist jedoch generell äußerst zweifelhaft und steht in keinem Verhältnis zur Komplexitätsreduktion, die durch das Aufspalten der verschachtelten Aussagen in Einzelaussagen gewonnen wird (es ist auch fraglich, inwieweit der Rezipient komplex-verschachtelte Bewertungsstrukturen "richtig" auflöst). Für die Aussagenverteilung insgesamt fallen diese wenigen Sonderfälle nicht ins Gewicht.

### **7.4.3 Die empirischen Analysen**

#### **7.4.3.1 Deskriptiv-analytische Betrachtungen von Bevölkerungsmeinung und Nachrichteninhalten**

In einem ersten deskriptiven Schritt wird bei den empirischen Analysen die Kanzlerpräferenz (abhängige Variable späterer Analysen) vor allem grafisch dargestellt (Kapitel 8.1). Die detaillierte Betrachtung bestimmter Phasen sowie die Einbeziehung wichtiger Ereignisse (z.B. Bundespräsidentenwahl, Europawahl) in diese grafischen Betrachtungen tragen jedoch schon analytische Züge.

Auch die Darstellung der Medienberichterstattung über die Kanzlerkandidaten (Kapitel 8.2) weist neben der Deskription, wer was zu wem mit welcher Bewertung sagte, bereits analytische Anteile auf, wenn z.B. Themen zu bestimmten Sinnkomplexen zusammengefasst werden und die Kandidaten auf ihr Abschneiden in diesen Oberthemen hin verglichen werden (Bezug Framing und Priming, siehe Kapitel 8.2.5 und 8.2.6).

Die Darstellung der Kanzlerpräferenzen der Bevölkerung und der Ergebnisse der Inhaltsanalyse (Medienberichterstattung) sind daher keine "rein deskriptiven" Auswertungsschritte, sondern geben bereits wichtige Hinweise auf mögliche Ursachen und Hintergründe und können einige dieser Hinweise zumindest teilweise überprüfen. Dieser Auswertungsteil (Kapitel 8) wird daher als "deskriptiv-analytisch" bezeichnet.

#### **7.4.3.2 Zeitreihenanalysen**

Da kommunikationswissenschaftliche Studien m.E. viel zu selten den Faktor Zeit als wichtige kausaltheoretische Implikation beachten, werde ich bei dieser Untersuchung auch eine zeitreihenanalytische Betrachtung der zeitlichen Abfolge von Ursache (Medieninhalt) und Wirkung (Bevölkerungsmeinung) vornehmen.

Die wenigen vorliegenden zeitreihenanalytischen Studien im Fach kranken zudem daran, dass sie relativ große Zeiteinheiten als Vergleichsbasis wählen (müssen). Dies ist zwar aus forschungsökonomischen und praktischen Gründen nur allzu verständlich, kausaltheoretisch bleibt jedoch das Problem, das Studien die z.B. mit monatlichen Daten arbeiten, Effekte, die wahrscheinlich oft innerhalb von Tage ablaufen, überhaupt nicht in ihrer zeitlichen Reihenfolge abbilden können (siehe dazu auch die Ausführungen zur Datenstruktur im Kapitel 7.4.1 und von GÜLLNER 2001: 568-572).

Selbst die wenigen Studien in der Kommunikationswissenschaft, die mit wöchentlichen Daten arbeiten, können viele der Effekte, die sich Medienforscher typischerweise vorstellen, nicht messen und dementsprechend auch nicht aufdecken. Die möglichen Effekte von Fernsehnachrichten sind dafür ein gutes und relevantes Beispiel: Die Wirkungen dieser tagesaktuellen politischen Sendungen mit großer Reichweite lassen sich bei wöchentlicher Zusammenfassung zeitlich nicht mehr fein genug verfolgen, um Ursache-Wirkungs-Abläufe festzuhalten. Korrelationen – als empirisches Äquivalent für einen vermuteten Kausalzusammenhang – treten so nur noch "zeitgleich" auf: alle Effekte, die innerhalb von wenigen Tagen eventuell in einer klaren zeitlichen Abfolge abgelaufen sind, erscheinen dem Forscher mit wöchentlichen Daten als "gleichzeitig". Zeitgleiche Korrelationen sind jedoch ein weitaus weniger gutes Argument für kausale Zusammenhänge als zeitverschobene Korrelationen. Da Medienwirkungen immer als Prozess zu verstehen sind, der auch sehr schnell (innerhalb von Tagen) ablaufen kann, ist es mein Anliegen, diesen Prozess anhand von täglich erhobenen Daten in entsprechenden Analysemodellen statistisch abzubilden.

Bevor die Ergebnisse der empirischen Analysen dargestellt werden, sollen einige methodische und methodologische Überlegungen den Begriff "Zeitreihenanalyse" näher spezifizieren. Dies ist insofern wichtig, als dass es "die Zeitreihenanalyse" nicht gibt und zeitreihenanalytische Auswertungen in der Kommunikationswissenschaft bislang eher selten sind.

Der Begriff Zeitreihenanalyse ist nicht die Bezeichnung für eine einzelne, konkrete Auswertungsmethode, sondern beschreibt eine Kategorie von Verfahren, mit denen Zeitreihen analysiert werden können. Das statische Pendant dieser Kategorie wäre die "Querschnittsanalyse", unter der ebenfalls eine Fülle unterschiedlicher Datenauswertungsverfahren subsumieren. Charakteristisch für die Zeitreihenanalyse ist die Struktur der Daten, die ihr zugrunde liegen und die sich daraus ergebenden statistischen Schätz-Probleme. Eine Zeitreihe besteht aus Ausprägungen einer Variablen, die in Zeitabständen wiederholt gemessen

werden. Für statistische Analysen sollten Zeitreihen zwei Kriterien genügen: Es sollte sich um "lange" Zeitreihen von mindestens 50 Messpunkten handeln und die Zeitabstände zwischen den Messungen sollte immer gleich sein ("äquidistante Messwiederholungen"; BOX/JENKINS 1976, MCCLEARY/HAY 1980, THOME 1988 UND 1992).<sup>61</sup>

Es ist für die Zeitreihe unerheblich, ob die Messungen von einer einzelnen Person stammen (z.B. im klinischen Bereich), ob es sich um Durchschnittswerte einer Stichprobe (z.B. Panel) oder um Indizes zur Beschreibung von Populationen anhand vieler Stichproben handelt (wie in dieser Analyse, in der die Werte von täglichen, repräsentativen Bevölkerungsstichproben untersucht werden).

Formal lassen sich für Zeitreihenanalysen drei verschiedene Hypothesenarten unterscheiden (BORTZ 1984: 447):

- (a) Die in einer Zeitreihe entdeckten Regelmäßigkeiten setzen sich auch zukünftig fort (Vorhersagemodelle);
- (b) Ein "Treatment" verändert eine Zeitreihe in einer bestimmten Weise (Interventionsmodelle);
- (c) Änderungen in der Verlaufsform einer Zeitreihe sind auf eine oder mehrere andere Zeitreihen zurückzuführen (Transferfunktionsmodelle).

Zu Punkt (c) Transferfunktionsmodelle sagt BORTZ: "Diese multivariate Variante der Zeitreihenanalyse gestattet die Überprüfung kausaler Hypothesen" (BORTZ 1984: 461). Inwiefern die Zeitreihenanalyse Vorteile bei der Überprüfung kausaler Hypothesen bietet, ist Gegenstand des folgenden, methodologisch orientierten Kapitels.

### **"Mehr Kausalität" durch Zeitreihen-Verfahren?**

"Das Problem der Kausalität ist so alt wie die (abendländische) Wissenschaft, und wie viele wichtige wissenschaftliche (wissenschaftstheoretische) Probleme ist es bis heute nicht "gelöst"" (KIRCHGÄBNER 1981: 9).

Auch an dieser Stelle kann keine Lösung des wissenschaftstheoretischen Problems mit dem Kausalitätsbegriff erfolgen. Es werden nur einige zentrale Gedanken zum Verhältnis von "Kausalität" und "Korrelation" dargestellt.

---

61 Diese Voraussetzungen ermöglichen eine regressionsanalytische Auswertung und die Anwendung von Signifikanztests. Es lassen sich auch Daten auswerten, die diesen Anforderungen nicht genügen. So können bei Kenntnis der Form von zeitlichen Abhängigkeiten in den Variablen auch kürzere Reihen statistisch analysiert werden. Auf andere, eher "qualitativ" orientierte Methoden kann in dieser Arbeit nicht eingegangen werden.



"Kausalität" wurde in der Wissenschaftsgeschichte gelegentlich auf "Prognostizierbarkeit" reduziert (vgl. EBD.). Oder das Kausalitätskonzept (Ursache-Wirkung) wurde aufgrund der Definitionsschwierigkeiten völlig aus dem wissenschaftlichen Diskurs getilgt und man kümmerte sich nicht mehr darum, "ob ein "Zusammenhang" zwischen Variablen kausal interpretiert werden kann oder nicht" (THOME 1988: 93).

"Dennoch kann die Forschung in den empirischen Wissenschaften (und auch die ganze Geschichte der empirischen Wissenschaften) als eine beständige Suche nach kausalen Zusammenhängen betrachtet werden" (KIRCHGÄBNER 1981: 9).

Auch Thome hält die "Vermeidungsstrategie", den Kausalitätsbegriff ganz zu streichen, für die Sozialwissenschaften nicht für sinnvoll:

"Gewiss hat die Soziologie<sup>62</sup> nicht nur die Aufgabe, Kausalwissen zu produzieren; aber gäbe sie den Anspruch auf, auch Theorien über planvoll änderbare soziale Wirklichkeit zu liefern, minderte sie ihre Legitimität. Auf diesen Bedarf an Handlungsorientierung ist das Kausalitätskonzept zugeschnitten, wie immer man es im Einzelnen definieren mag" (THOME 1988: 93).

Es kann an dieser Stelle keine wissenschaftstheoretische Auseinandersetzung mit unterschiedlichen Kausalitätskonzepten erfolgen. Es sollen lediglich einige Aspekte dieser Problematik zum Vergleich von statischen und dynamischen (zeitreihenanalytischen) Methoden diskutiert werden. Dazu ist es ausreichend, auf den gemeinsamen Gedanken der verschiedenen Kausalitätsdefinitionen zu rekurrieren:

"Nicht in der universellen Koinzidenz zweier Phänomene X und Y ist das entscheidende Merkmal ihres kausalen Zusammenhangs zu sehen, sondern darin, das in Y eine Änderung herbeigeführt (oder unterdrückt) werden könnte, wenn man X änderte (oder eine Änderung in X unterdrückte)" (VON WRIGHT 1974: 72).

Statistisch bedeutet dies, dass einfache (zeitgleiche) Korrelationen keine Kausalität beweisen. Diese altbekannte Aussage nennt Thome "ebenso wahr wie irreführend" (THOME 1988: 93). Einerseits sei es richtig, dass Scheinkorrelationen (Korrelationen ohne Kausalbeziehung) nie gänzlich auszuschließen seien und die theoretische Behauptung einer Kausalbeziehung stets über Beobachtbares hinausgehe, andererseits könnten Kausalhypothesen ohne Korrelationen überhaupt nicht überprüft werden.

Wenn wir nun weitergehen und nicht in der "universellen Koinzidenz" (Korrelation) zweier Phänomene X und Y das entscheidende Merkmal ihres kausalen Zusammenhangs sehen, sondern darin, das in Y eine Änderung herbeigeführt werden kann, wenn man X ändert, so zeigt sich, dass diese Veränderung der Phänomene mit Querschnittsdaten überhaupt nicht beobachtet werden kann. Bei Querschnittsanalysen wird die Veränderung simu-

---

62 Das Zitat ist vollständig auf die Kommunikationswissenschaft übertragbar.

liert. Der Wert des Phänomens X ändert sich beim einzelnen Untersuchungsobjekt nicht - es existiert ja nur eine Messung pro Untersuchungsobjekt. Die Simulation erfolgt über die unterschiedlichen Werte für die einzelnen Variablen bei den verschiedenen Untersuchungsobjekten. Diese Interpretation der Daten benötigt aber zwei Voraussetzungen: Zum einen muss bei allen Untersuchungsobjekten der gleiche strukturelle Zusammenhang vorliegen. Dieses Problem kann durch multivariate Modellspezifikation minimiert werden. Zum anderen ist eine Kausalinterpretation aber nur dann sinnvoll, wenn sich die Daten zum Zeitpunkt der Messung in "Ruhelage" ("Äquilibrium") befinden. Jede *vor* der Messung abgelaufene (aber nicht beobachtete) Änderung von X-Werten muss sich bereits voll auf Y ausgewirkt haben. Andernfalls ist mit einer Verzerrung der geschätzten Parameter zu rechnen (THOME 1992b: 81f.). Thome sieht diese Gleichgewichtsvoraussetzung, die in Lehrbüchern und in der Forschungspraxis oft unterschlagen werde, in vielen sozialwissenschaftlichen Untersuchungen als nicht erfüllt an.

Mit der operationalen Definition von Ursache und Wirkung hat der Kausalitätsbegriff eine "Renaissance" erlebt, die vor allem auf die Arbeiten von GRANGER (1969) zurückgeht. Aus dem Konzept der "Granger-Kausalität" sind statistische Tests abgeleitet worden (z.B. das "direkte Granger-Verfahren", das von SMITH 1987 und BROSIUS/KEPPLINGER 1990 benutzt wurde), die als Datenbasis multivariate Zeitreihen voraussetzen, was mit ein Grund für die Bedeutung der multivariaten Zeitreihenmodelle ist (vgl. siehe KIRCHGÄBNER 1983: 229f., SCHMITZ 1989: 164).

Im Gegensatz zu Querschnittanalysen, die, wie hier gezeigt, selbst den grundlegenden Gedanken von Kausalität (Änderung eines Phänomens bewirkt Änderung eines anderen Phänomens) nicht statistisch adäquat abbilden können (Veränderung der Variable X bewirkt Veränderung in Variable Y), können mit zeitreihenanalytischen Verfahren Implikationen kausaltheoretischer Modelle überprüft werden, die bislang ungetestet blieben. Es geht dabei nicht um den "Beweis von Kausalität". Diese Leistung bleibt immer noch dem Experiment mit randomisierten Gruppen vorbehalten. Aber mit Zeitreihenverfahren ist es möglich, Voraussetzungen für Kausalinterpretationen zu überprüfen, die von Querschnittanalysen nicht überprüft werden können. Deshalb können mit dynamischen Analysen kausale Zusammenhänge "besser" überprüft und unterstützt werden, als dies mit statischen Verfahren der Fall ist. BROSIUS/ESSER stellen die Vorteile von Zeitreihenanalysen folgendermaßen dar:

"Auch wenn man also mit Zeitreihendaten Kausalität im experimentellen Sinne nicht nachweisen kann, bieten sich doch statistische Verfahren an, mit denen man zwei wichtige Voraussetzungen für eine kausale Interpretation überprüfen kann.

Die erste Voraussetzung betrifft die zeitliche Dynamik der unabhängigen und abhängigen Zeitreihen. Nach dem allgemeinen Verständnis liegt die Ursache eines Phänomens zeitlich vor einer möglichen Wirkung... Die zweite Voraussetzung betrifft den Einfluss von Drittvariablen, bzw. die Vermeidung von Scheinzusammenhängen" (BROSIOUS/ESSER 1995: 132).

Thome sieht die Vorteile zeitreihenanalytischer Verfahren vor allem in folgenden Aspekten (THOME 1988: 94):

- Die vermutete Exogenität bestimmter Variablen eines Gleichungssystem kann getestet werden;
- die Struktur der Wirkungsverzögerung und des weiteren Wirkungsverlaufs ("Lag-Struktur") kann aufgedeckt werden.

Es soll an dieser Stelle aber auch auf Nachteile von Zeitreihenverfahren eingegangen werden. Diese liegen im Bereich statistischer Schätzmethoden. Während auf theoretischer Ebene die Vorteile eindeutig sind, entstehen bei der Anwendung statistischer Zeitreihenanalysen in der Praxis Probleme bei der Parameterschätzung. Auf Verfahren und spezielle Schätzprobleme wird im nächsten Kapitel detailliert eingegangen.

### **Die Analyse-Strategie nach BOX/JENKINS**

Die Zeitreihen dieser Arbeit werden mit der "BOX/JENKINS-Methode" ausgewertet.

"Die wenigen bisher vorliegenden sozialwissenschaftlichen Anwendungen dieser Methode sind vielversprechend und lassen hoffen, dass der BOX/JENKINS-Ansatz zukünftig häufiger zur Überprüfung langfristiger Veränderungen herangezogen wird" (BORTZ 1984: 447).

Die BOX/JENKINS-Methode ist nicht ein einzelnes Auswertungsverfahren, sondern eine Gruppe statistischer Analyse-Instrumente zur Untersuchung von Zeitreihen. GEORGE E.P. BOX und GWILYM M. JENKINS haben verschiedene zeitreihenanalytische Elemente zusammengetragen und eine "Strategie" zur Untersuchung von Zeitreihen entwickelt (BOX/JENKINS 1976). MCCLEARY/HAY bezeichnen die BOX/JENKINS-Methode als "model-building strategy". Kernstücke dieser Strategie sind ARIMA-Modelle, Kreuzkorrelationen sowie Transferfunktionen, wobei die Transferfunktionen das eigentliche Ziel der BOX/JENKINS-Analyse darstellen, da erst sie die vermuteten Zusammenhänge innerhalb eines Gesamt-Schätzmodells adäquat überprüfen.

In der Praxis finden sich jedoch bis heute insgesamt sehr wenige Zeitreihenuntersuchungen. Die seltenen Längsschnitt-Studien verwenden wiederum selten das BOX/JENKINS-Verfahren und von den wenigen Studien, die es tun, führen viele das Verfahren nicht vollständig durch, sondern bleiben bei den Kreuzkorrelationen stehen, obwohl

BOX/JENKINS diese nur als Hilfsmittel und Zwischenschritt ansehen, der mit einigen statistischen Problemen behaftet ist und deshalb nicht als Endergebnis einer Analyse fungieren sollte. ARIMA-Modelle, Kreuzkorrelationen und Transferfunktionen haben in der BOX/JENKINS-Strategie also unterschiedliche Funktionen.

*ARIMA-Modelle* beschreiben einzelne Reihen (univariate Zeitreihenanalyse) bzw. den Fehlerterm eines multivariaten Modells, *Kreuzkorrelationen* liefern Hinweise auf bivariate, zeitversetzte Verknüpfungen und mit *Transferfunktionen* wird der Zusammenhang zwischen zwei oder mehr Messwiederholungsreihen in einem flexiblen, multivariaten Gesamt-Modell untersucht (synonym: multivariate Zeitreihenanalyse, ARIMAX). Dabei ist die multivariate Vorgehensweise untrennbar mit der univariaten verknüpft: Bevor der Zusammenhang zwischen zwei Zeitreihen ermittelt werden kann (Kreuzkorrelationen/Transferfunktionen), müssen für die einzelnen Reihen univariate Analysen vorgenommen werden (ARIMA-Modelle) (ROTTLEUTHNER-LUTTER 1986: 82f.). Erst dann können unverzerrte Kreuzkorrelationen errechnet werden, die Hinweise auf die Beschaffenheit der Transferfunktion geben und damit das Verbindungsstück zwischen uni- und multivariaten Modellen bilden.

Die Analyse-Strategie von BOX/JENKINS soll die typischen Schätzprobleme bei Zeitreihen lösen. Die Probleme ergeben sich aus den speziellen Eigenschaften von Zeitreihen, unter anderem "Trend" (Stationarität) sowie "Autokorrelationen".

Unterliegen zwei Zeitreihen einem gemeinsamen oder ähnlichen *Trend*, der durch andere Größen verursacht wird, die nicht im Modell spezifiziert sind, so kann sich eine Scheinkorrelation ergeben. Diesem Problem kann relativ einfach begegnet werden, indem entweder die "Trendursache" in das Modell aufgenommen wird (was theoretisch wie statistisch das Beste wäre) oder indem jede Reihe differenziert wird.<sup>63</sup>

Die zweite systematische Komponente, die Scheinzusammenhänge beim Vergleich zweier (oder mehrerer) Zeitreihen produzieren können, sind die so genannten *Autokorrelationen*. Autokorrelationen sind "serienimmanente" oder "serieninterne" Korrelationen zwi-

---

63 Andere Verfahren der Trendbereinigung existieren zwar auch, werden aber eher kritisch bewertet. So wird die Trendbereinigung mittels OLS-Regression über die Zeit von MCCLEARY/HAY als sehr problematisch beschrieben ("almost always inappropriate", 1980: 32). Allerdings sollte auch die Differenzierung nicht automatisch erfolgen - ein scheinbarer systematischer *Trend* kann auch nur Ausdruck eines stochastischen (und damit unsystematischen) *Drifts* sein. "Unless there is a strong theoretical basis (or empirical evidence) for assuming that a time series process trends deterministically, there are great advantages to be gained by modeling it stochastically" (MCCLEARY/HAY 1980: 36). BOX/JENKINS bemerken dazu, dass das Problem von Trend versus Drift auf den Unterschied zwischen "fitting" und "modeling" verweist. Weitere Ausführungen zum Problem des Differenzierens würden den Rahmen dieser Arbeit sprengen. Es sei auf den Originaltext (BOX/JENKINS 1976) sowie auf das vor allem didaktisch hervorragende Lehrbuch von MCCLEARY/HAY 1980 verwiesen.

schen den Originalwerten einer Zeitreihe (zum Zeitpunkt  $t_0$ ) mit ihren zeitversetzten Werten ( $t-1$ ,  $t-2$ ,  $t-3$  usw.). Hohe Autokorrelationen bedeuten, dass die benachbarten Werte einer Zeitreihe stark miteinander zusammenhängen. In einem statistischen Schätzmodell muss den Autokorrelationen besondere Aufmerksamkeit geschenkt werden, da sie die Voraussetzungen vieler "herkömmlicher" Schätzverfahren verletzen. Bei OLS-Regressionen wird z.B. vorausgesetzt, dass die verschiedenen Werte einer Variablen bei den Untersuchungsobjekten *nicht* miteinander korrelieren. Statistisch ausgedrückt muss die Kovarianz der Residuen eines Regressionsmodells gleich Null sein. Die Verletzung dieser Voraussetzung führt zu ineffizienten und/oder verzerrten Parameterschätzungen.

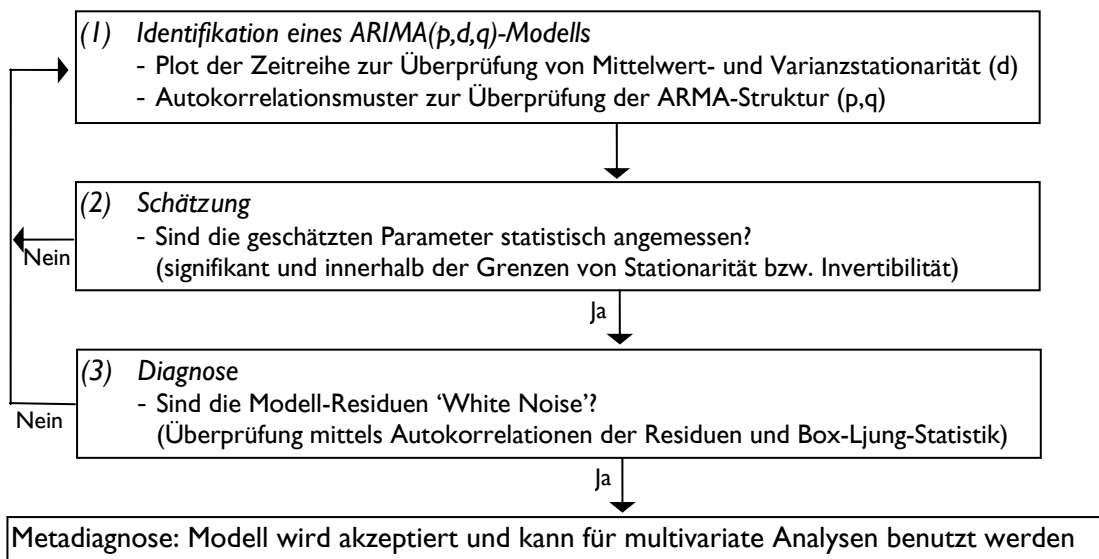
Um Trend und Autokorrelationen kontrollieren zu können, haben BOX/JENKINS die ARIMA-Modelle konzipiert. ARIMA steht für *AutoRegressive Integrated Moving Average*, was drei verschiedene Modellteile bezeichnet: Der Teil "Integrated" steht für die Differenzierung, der eine Zeitreihe möglicherweise unterzogen werden muss (Trendbereinigung, s.o.). Ihr Grad wird im ARIMA-Modell mit dem Buchstaben  $d$  beschrieben. Die beiden Teile "Autoregressive" und "Moving Average" stehen für zwei Möglichkeiten, Autokorrelationen zu modellieren. Im Modell werden diese Teile mit  $p$  bzw.  $q$  bezeichnet, so dass ein ARIMA( $p,d,q$ )-Modell jenes Modell beschreibt, mit dem man "sämtliche Bestandteile aus Zeitreihen entfernen [kann], die eine kausale Interpretation verhindern" (BROSIUS/ESSER 1995: 135).<sup>64</sup> Der mathematische Hintergrund der ARIMA-Modelle ist komplex. Er wird hier nicht weiter dargestellt, was außer an Platzkapazitäten auch an inhaltlichen Gründen liegt: ARIMA-Modelle bilden in der BOX/JENKINS-Strategie nur die notwendige Vorstufe für weitergehende Zusammenhangsanalysen. Ihre geschätzten Parameter selbst sind nicht Grundlage kausaler Interpretationen.

Die univariate "ARIMA-Modellbildung" erfolgt in einem iterativen Prozess mit drei Schritten, die sich graphisch folgendermaßen darstellen lassen (Abbildung 19; diese Beschreibung praktischer Analyseschritte beinhaltet einige Termini, die bislang nicht besprochen wurden. Sie werden im folgenden Absatz erläutert).

---

64 Bei Zeitreihen muss auch auf saisonale Schwankungen geachtet werden. Liegt eine saisonale Komponente in einer Serie vor, so wird auch für die Saisonkomponente ein ARIMA-Modell angepasst. Das "Gesamt-ARIMA-Modell" wird in der Literatur öfter SARIMA genannt oder es wird von einem "general ARIMA( $p,d,q$ )( $P,D,Q$ )<sub>s</sub>-model" gesprochen, wobei die Großbuchstaben ( $P,D,Q$ ) das ARIMA-Modell für den saisonalen Anteil beschreiben. Bei meiner Darstellung verzichte ich auf die gesonderte Beschreibung der Probleme mit Saisonkomponenten, da sie den nicht-saisonalen entsprechen und in meinen Analysen keine saisonalen Strukturen vorliegen.

**Abbildung 12: Die "ARIMA-Modellbildungsstrategie"**



Quelle: MCCLEARY/HAY 1980: 92; eigene Bearbeitung

Um Zeitreihen-Parameter schätzen zu können, müssen die Reihen wenigstens *schwachstationär* sein, d.h. *Mittelwert* und *Varianz* dürfen nicht systematisch variieren. Während bereits diskutiert wurde, wie eine Reihe mittelwert-stationär gemacht werden kann (Differenzierung), wird auf das Problem der Varianz-Stationarität nur hingewiesen (siehe dazu z.B. THOME 1994). In meiner konkreten Analyse spielt dieses Problem keine Rolle.

Die Parameter einer Zeitreihen-Modellschätzung müssen zudem nicht nur signifikant sein, sondern auch innerhalb der *Grenzen der Stationarität* bzw. *Invertibilität* liegen. Diese Grenzen ergeben sich aus der Dynamik der geschätzten Modelle. Bestimmte Parameter-Kombinationen führen zu einem im Zeitverlauf instabilen Modell – der Effekt einer Variablen auf die andere würde sich z.B. über die Zeit exponentiell steigern, was zu instabilen und unplausiblen Vorhersagen führt.

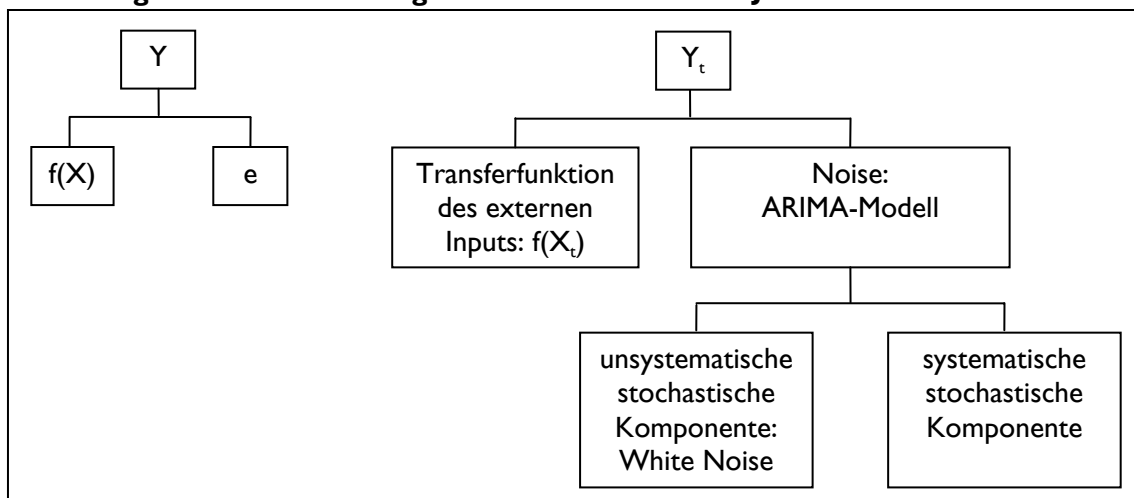
Sind die Parameter jedoch "statistisch akzeptabel" (vgl. MCCLEARY/HAY 1980: 92), so wird im Diagnoseschritt anhand der *Modell-Residuen* die Angemessenheit des theoretischen ARIMA-Modells überprüft. Dazu wird die theoretische, im ARIMA-Modell spezifizierte Zeitreihe von der tatsächlichen, beobachteten Zeitreihe abgezogen. Die Differenz sind die so genannten Residuen, die ihrerseits weder Trend noch Autokorrelationen aufweisen dürfen. Weisen die Residuen aber immer noch diese systematischen Bestandteile auf, so war das ARIMA-Modell falsch identifiziert worden und der iterative Modellbildungsprozess muss erneut durchgeführt werden. Beinhalten die Residuen aber keinerlei systematischen Bestandteile, so werden sie "White Noise" ("Weißes Rauschen") genannt. Die Varianz des

so erzeugten "White Noise" hängt nicht mehr von der internen Dynamik der Reihe ab, sondern kann auf andere Faktoren zurückgeführt und mit ihnen erklärt werden.

Nach dieser kurzen Einführung in den univariaten Teil der BOX/JENKINS-Methode, wird nun der multivariate Teil der Strategie dargestellt, der den Zusammenhang zwischen Zeitreihen über die Transferfunktionen abbildet. MCCLEARY/HAY haben die Analyse-Verfahren von BOX/JENKINS in sechs Analyseschritte aufgeteilt. Wie schon beim univariaten Vorgehen handelt es sich um eine iterative Prozedur, die bei negativer Bewertung einzelner Komponenten einen Rückschritt und Neuversuch vorsieht.

Die statistischen Hintergründe der einzelnen Schritte werden hier nicht theoretisch besprochen, sondern im Ergebnisteil am konkreten Beispiel vorgeführt. Des Weiteren wird die Analyse-Strategie der Übersichtlichkeit wegen nur bivariat dargestellt. Die multivariate Erweiterung der Strategie bedeutet keine grundsätzliche Veränderung des Modells. Bevor die Analyse-Strategie diskutiert wird, ist ein Vergleich zwischen der BOX/JENKINS-Methode und einem klassischen Regressionsmodell hilfreich:

**Abbildung 13: Klassisches Regressionsmodell vs. BOX/JENKINS-Methode**



Quelle: ROTTLEUTHNER-LUTTER 1986: 89; eigene Bearbeitung

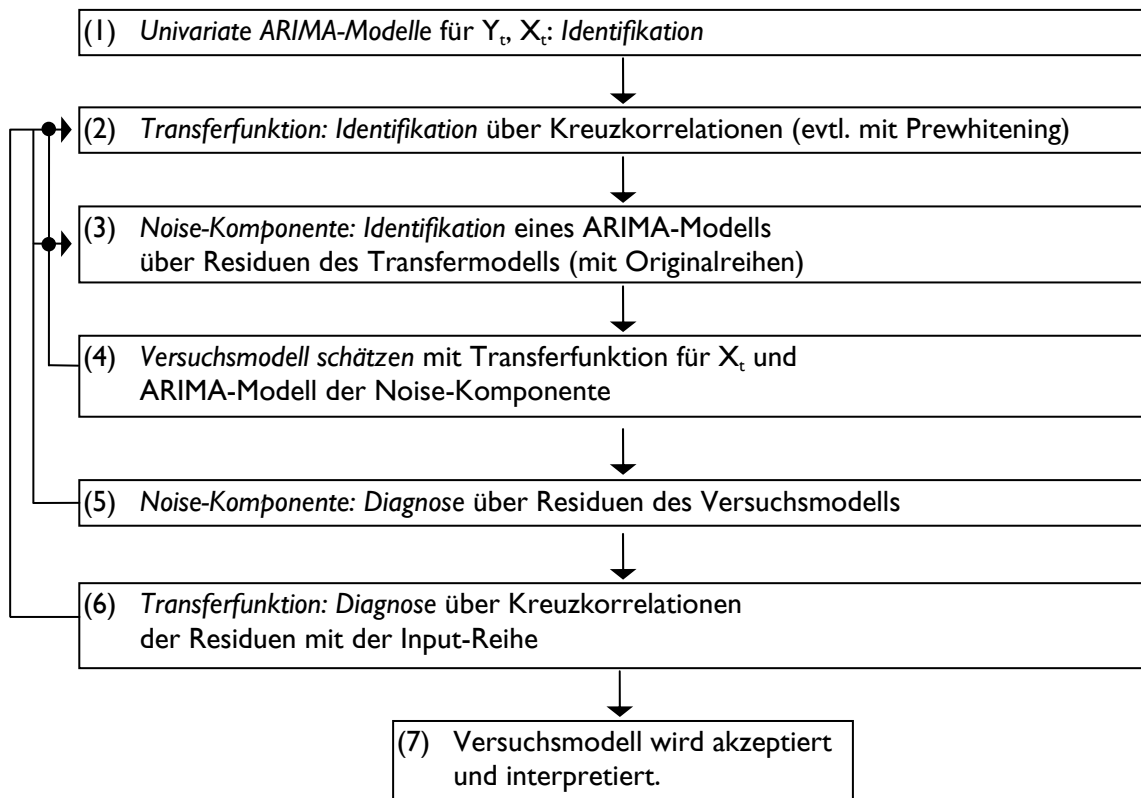
Bei der klassischen Regression wird die abhängige Variable  $Y$  als Funktion der unabhängigen Variable  $X$  und eines Fehlerterms  $e$  aufgefasst. Der Fehlerterm muss für eine effiziente, unverzerrte Schätzung unabhängig bzw. nicht autokorreliert sein (ROTTLEUTHNER-LUTTER 1986: 86).

Bei der BOX/JENKINS-Methode wird die Output-Reihe  $Y_t$  als (Transfer-)Funktion des Input  $X_t$  und des "Noise" aufgefasst, wobei die interne Abhängigkeitsstruktur der Noise-

Komponente in einem ARIMA-Modell spezifiziert wird (EBD.: 88f.). Aus dem Schaubild ergibt sich, dass bei der Zeitreihenanalyse im Gegensatz zur Regression zwei Modellteile geschätzt werden müssen: die Transferfunktion sowie die Noise-Komponente.

Den Anforderungen von Zeitreihendaten entsprechend ist die Analyse-Methode meiner Arbeit kein einzelnes, feststehendes Modell, sondern ein iteratives Verfahren, dass sich in die Prozeduren *Identifikation*, *Schätzung* und *Diagnose* einteilen lässt. Dabei werden *Identifikation* und *Diagnose* für beide Modellanteile (Transferfunktion und ARIMA-Noise-Modell) in einzelnen Schritten durchgeführt, während die *Schätzung* der beiden Anteile simultan in einem "Gesamtmodell" erfolgt. Zusammen mit der vorausgehenden, notwendigen univariaten Analyse der beteiligten Zeitreihen ergeben sich folgende Analyse-Schritte:

**Abbildung 14: Die multivariate "Modellbildungsstrategie" für Zeitreihenanalysen**



Quelle: McCLEARY/HAY 1980: 251; eigene Bearbeitung

In einer ersten "Voruntersuchung" werden die univariaten Modelle für die einzelnen Zeitreihen ermittelt (1). Die univariate ARIMA-Modellierung erfolgt ebenfalls in den iterativen Phasen *Identifikation*, *Schätzung* und *Diagnose* (siehe Abbildung 12). Die Ergebnisse dieser Analyse zeigen u.a., ob die Reihen stationär sind oder transformiert werden müssen.



In Schritt (2) und (3) werden dann die beiden Komponenten des Zeitreihenmodells *identifiziert*. Über den Zusammenhang von Input und Output, der in der Transferfunktion abgebildet werden soll, geben die Kreuzkorrelationen zwischen den Zeitreihen erste Anhaltspunkte. Die Kreuzkorrelationen können jedoch nicht mit den Original-Reihen durchgeführt werden, wenn die unabhängige Variable (Input-Reihe) mit Autokorrelationen behaftet ist, weil diese die Kreuzkorrelationen verzerren würden.

*"...the CCF<sup>65</sup> can be interpreted only when the causer variable is a white noise process... When the causer series is not white noise, the CCF will reflect both between-series dependencies and within-series dependencies" (MCCLEARY/HAY 1980: 246).*

Die Transformation, die durchgeführt werden muss, um die Kreuzkorrelationen interpretierbar zu machen, wird "Prewhitening" ("Vorweißen") genannt: Eine Zeitreihe mit systematischen Bestandteilen wird mittels ARIMA-Modell in einen White-Noise-Prozeß transformiert. Auf welche Art und Weise das "Prewhitening" durchgeführt werden sollte, ist jedoch umstritten, bzw. verschiedene Autoren benutzen verschiedene Varianten oder "weißen" gar nicht vor – was jedoch unwillkürlich zu Verzerrungen des Kreuzkorrelationsmusters führt. Bei den "Methodikern" in der Literatur, also bei den Autoren, die sich speziell mit der Methode der Zeitreihenanalyse befassen, ist die angemessene Form des Prewhitening jedoch unumstritten (z.B. BOX/JENKINS 1976, MCCLEARY/HAY 1980, KMENTA 1986, THOME 1988, 1992): Um zu einem unverzerrtem Kreuzkorrelationsmuster zu gelangen, muss zuerst die Input-Reihe mittels eines ARIMA-Modells vorgeweißt werden. Dieses Modell wird dann (invertiert) auf die Output-Reihe angewendet. In der Praxis sind dann auch andere, weniger angemessenere Verfahren verwendet worden, was vielleicht auch an der technischen Umsetzbarkeit liegt: Das statistische Auswertungsprogramm SPSS bietet keine automatische Prozedur zum Prewhitening an. Das Vorweißen muss "von Hand" vorgenommen werden (Erstellung und Umformung der Formel aus dem Ergebnis des ARIMA-Modells der Input-Reihe zur Berechnung der vorgeweißten Output-Reihe). Je nach ARIMA-Modell ist die Lösung nicht trivial bzw. kann nicht mit diesem Vorgehen ermittelt werden.

Das Prewhitening an sich stellt keinen inhaltlich-interpretierbaren Teil der Zeitreihenanalyse dar. Es ist aber zur Modellidentifikation von Bedeutung. Deshalb möchte ich nur kurz die Auswirkungen des Prewhitening auf die Gesamtanalyse problematisieren: Die Modellidentifikation im Rahmen der BOX/JENKINS-Methode gelingt nur, wenn die Varianz der Input-Reihe im Vergleich zu der Varianz der Output-Reihe, die durch das "Noise"

---

65 CCF: Cross-Correlation Function (Kreuzkorrelationen)

erzeugt wird, relativ groß ist oder/und eine große Menge von Messzeitpunkten vorliegt (BOX/JENKINS 1976: 387, THOME 1992b: 109). "Prewhitening" kann die Modellidentifikation daher unter Umständen erschweren. Das "Herausfiltern von Scheinzusammenhängen" dreht das Problem eventuell um: An die Stelle von Scheinkorrelationen treten nun "Scheinunabhängigkeiten" (KIRCHGÄBNER 1981: 145 und 1983: 221f., SCHMITZ 1989: 172f.). So kann z.B. durch die Datentransformation des Vorweißens eventuell der Anteil der Messfehler an der Gesamtvarianz erhöht werden. Diese und andere Auswirkungen des Prewhitening können dazu führen, daß die strukturelle Beziehung zwischen zwei Variablen mittels Kreuzkorrelationen nur schwer (oder gar nicht) identifiziert werden kann, wenn nur auf signifikante Korrelationen geschaut wird. Daher sollten Kreuzkorrelationen nicht als "Ergebnis" betrachtet werden, sondern nur als Zwischenschritt auf dem Weg zu einem Gesamt-Modell, das die inhaltlichen Beziehungen zwischen den Variablen in den Transferfunktionen ausdrückt. Zwar sollen die Kreuzkorrelationen Hinweise auf die Beschaffenheit dieser Transferfunktion liefern, jedoch sind dazu nicht unbedingt signifikante Korrelationen nötig. Interessanter als die absolute Höhe der Kreuzkorrelationen ist deren Struktur. Es sollten die Korrelationen der positiven und geringen lags höher sein als die negativen lags, da dies auf einen zeitlichen Vorsprung der Input-Variable hindeutet (inhaltlich heißt das: die Ursache liegt zeitlich vor der Wirkung). Wie gesagt müssen diese Korrelationen aber nicht unbedingt signifikant sein. Auch nicht-signifikante Lags können sich in der Transferfunktion als signifikante Parameter herausstellen. Das liegt daran, dass die Kreuzkorrelationen zur Vermeidung von Scheinabhängigkeiten das beschriebene "Prewhitening" praktizieren, was zu "Scheinunabhängigkeiten" führen kann – bei der Schätzung der Transferfunktionen hingegen werden wieder die Originaldaten verwendet und nicht die vorgeweißten Reihen. Daher geben die Kreuzkorrelationen zwar Hinweise auf die Beschaffenheit der Transferfunktion, sie tun dies aber nicht ausschließlich durch signifikante Werte (aus diesen Ausführungen wird auch deutlich, warum Kreuzkorrelationen nicht als Ergebnis einer Kausalanalyse taugen). MCCLEAR/HAY (1980) geben daher den Ratschlag, die theoretisch plausibelsten Transferfunktionen zu schätzen, auch wenn keine signifikanten Kreuzkorrelationen vorliegen. Theoretisch plausibel sind vor allem die zeitgleichen sowie die um ein oder zwei Zeitpunkte versetzten Beziehungen: die Wirkung einer Ursache wirkt sich meist direkt ("zeitgleich")<sup>66</sup> aus oder in den direkt folgenden Zeiteinheiten. Ursache-

---

66 Wobei "zeitgleich" lediglich heißt, dass Ursache und Wirkung im gleichen Erhebungs-Zeit-Intervall abgelaufen sind. Bei monatlichen Daten könnte also durchaus die Ursache eine Woche vor der Wirkung

Wirkungsstrukturen die komplexer sind (stärker verzögerte Wirkungen, Wirkungen in mehreren, aber nicht aufeinander folgenden Zeitintervallen), bedürften einer speziellen Wirkungshypothese, die meist nicht vorliegt. Daher empfehlen MCCLEARC/HAY, systematisch alle Transferfunktionen mit den lags 0, 1 und 2 zu testen – unabhängig davon, was die Kreuzkorrelationen ergeben haben. So wird das Problem der "Scheinunabhängigkeiten" umgangen. In der Regel geben Kreuzkorrelationen aber trotzdem hilfreiche Hinweise auf die Beschaffenheit der Transferfunktion – sei es durch signifikante Werte oder durch ihre Struktur.

Nachdem die Transferfunktion mit Hilfe der Kreuzkorrelationen (oder über systematische Tests) identifiziert wurde (2), muss das Fehlermodell für die "Noise"-Komponente spezifiziert werden (3). Dazu wird das Transfermodell geschätzt und die Residuen dieser Schätzung auf das ihnen zugrundeliegenden ARIMA-Modell überprüft. Das ARIMA-Modell für den Fehlerterm wird wieder nach der univariaten ARIMA-Modellierungsstrategie spezifiziert.

Wenn beide Modellteile (Transferfunktionen und ARIMA-Modell des Fehlerterms) identifiziert wurden, kann ein "Versuchsmodell" geschätzt werden (4). Sind die Parameter der einen oder anderen Komponente nicht statistisch akzeptabel (signifikant und innerhalb der Grenzen von Stationarität bzw. Invertibilität), so muss die entsprechende Komponente neu identifiziert werden (Rückschritt zu (2) und/oder (3)).

Ist das Versuchsmodell statistisch akzeptabel, so muss es anschließend zwei "Diagnose-Checks" passieren: Die Residuen des Versuchsmodells (Noise-Komponente) müssen einen White-Noise-Prozeß darstellen (5). Liegen in den Residuen jedoch noch systematische Anteile vor (ersichtlich z.B. aus der BOX-LJUNG-Statistik der Autokorrelationen), war das ARIMA-Modell unzureichend spezifiziert und muß neu identifiziert werden (Rückschritt zu (3)). Der zweite Diagnose-Check bezieht sich auf die Angemessenheit der Transferfunktion. Wenn sie den Zusammenhang zwischen Input und Output richtig modelliert, dürfte zwischen Input und Modellresiduen keinerlei Zusammenhang mehr bestehen (6). Existieren jedoch noch systematische Beziehungen zwischen Input und Residuen, so war die Transferfunktion nicht adäquat formuliert und muss neu identifiziert

---

stattgefunden haben – in den Kreuzkorrelationen könnte trotzdem nur die "zeitgleiche" Korrelation signifikant sein, da sich Ursache und Wirkung im selben Monat (=Erhebungsintervall) abspielt haben.

werden (Rückschritt zu (2)). Ob noch ein Zusammenhang zwischen Input und Modell-Residuen besteht, wird über das Kreuzkorrelogramm der beiden Reihen ermittelt.<sup>67</sup>

Das Versuchsmodell muss also nicht nur über statistisch akzeptable Parameter verfügen, sondern zusätzlich Diagnose-Test passieren, die die Angemessenheit der einzelnen Modell-Komponenten überprüfen. Ist auf diese Weise ein Modell spezifiziert worden, das sowohl den statistischen wie auch den diagnostischen Anforderungen entspricht, wird es als endgültiges (Gesamt-)Modell akzeptiert und kann zur inhaltlichen Interpretation der Daten herangezogen werden (7).

Diese Beschreibung des BOX/JENKINS-Verfahrens macht deutlich, dass Zeitreihenanalysen höhere Anforderungen sowohl an die zugrunde liegenden Daten als auch an die statistischen Auswertungsverfahren stellen als viele Querschnittsverfahren. Diese hohen Anforderungen sind vielleicht der Grund für die eher seltenen Zeitreihenanalysen in der wissenschaftlichen Forschung. Die Erhebung von Zeitreihendaten ist aufwendiger (und daher teurer) und die Auswertung ist in den Sozialwissenschaften noch nicht so standardisiert und in bekannte Statistik-Programme implementiert wie bei vielen Querschnittsverfahren.<sup>68</sup> Trotzdem wäre eine verstärkte Anwendung von Zeitreihenanalysen wünschenswert. Nicht als Konkurrenz oder gar Ersatz der Querschnittbetrachtungen, sondern als deren Ergänzung, um kausalthoretische Zusammenhänge besser empirisch absichern zu können.

Abschließend sei noch auf die verwirrende Begriffsvielfalt bei der Bezeichnung von Zeitreihenverfahren hingewiesen. BOX/JENKINS (1976) nannten ihre univariaten Beschreibungen von Zeitreihen "ARIMA"-Modelle. Mit ARIMA-Modellen werden also nur einzelne Zeitreihen untersucht und beschrieben. Geht es um den Zusammenhang zwischen zwei (oder mehreren) Zeitreihen, existieren in der Literatur verschiedene Bezeichnungen, die sich zwar teilweise auf Spezialfälle beziehen, aber trotzdem alle mit der dargestellten BOX/JENKINS-Strategie untersuchen lassen. BOX/JENKINS sprechen selbst bei ihren multivariaten Zeitreihenanalysen von "transfer function models" (1976), womit sie besonders die Transferfunktionen betonen, da in diesen Funktionen der Wirkungsmechanismus – der

---

67 Wegen statistischer Anforderungen muss der Input für diesen Diagnose-Schritt wieder vorgeweißt werden.

68 Das BOX/JENKINS-Verfahren stellt auch nur *eine* Möglichkeit der statistischen Längsschnittanalyse dar und ist nicht unumstritten. Allerdings entzündet sich die Kritik (z.B. KIRCHGÄBNER 1983) m.E. vor allem an den Kreuzkorrelationen, wenn sie als Endergebnis präsentiert werden. Das war von BOX/JENKINS jedoch nicht vorgesehen.

Zusammenhang zwischen unabhängiger und abhängiger Variable – ausgedrückt wird. Daneben existieren jedoch noch eine Reihe weiterer Bezeichnungen: BOX/TIAO 1975 sprechen von "ARIMAX", wobei "X" für die Einbeziehung unabhängiger Variablen (üblicherweise bezeichnet mit  $X_1$ ,  $X_2$  etc.) steht. PANKRATZ (1991) nennt diese Analyseform "dynamic regression", womit der Vergleich zur bekannteren Regression aus der Querschnittanalyse hervorgehoben wird. Weitere Bezeichnungen sind "intervention model", "interrupted time series models", "regression models with ARMA errors" oder "BOX-TIAO models". Wie gesagt bezeichnen manche dieser Begriffe Sonderfälle, sie sind jedoch alle mit der BOX/JENKINS-Strategie im Rahmen ihrer Transferfunktionsmodelle untersuchbar.

Auch die Begriff abhängige und unabhängige Variable bekommen bei Zeitreihenanalysen noch einige Synonyme an die Seite: Variablen werden generell auch oft als "Reihe" (engl.: series) bezeichnet, da sie eine zeitlich geordnete Reihe darstellen. Die abhängige Variable wird auch als "output" oder "respondent variable" (oder series) bezeichnet, die unabhängige als "input" oder "predictor variable" (oder series).