

DUISBURGER BEITRÄGE zur SOZIOLOGISCHEN FORSCHUNG

No. 3/1999

Sozialwissenschaftliches Praktikum in der amtlichen Statistik

Fallbeispiel: Ergebnisse politischer Wahlen und
Strukturmerkmale der nordrhein-westfälischen Gemeinden

von

Wolfgang Gerß

Die "Duisburger Beiträge zur soziologischen Forschung" werden herausgegeben vom:

Fachbereich 1 - Soziologie
Gerhard-Mercator-Universität Gesamthochschule Duisburg
Lotharstraße 65
D-47048 Duisburg

Ein Verzeichnis aller bisher erschienenen Beiträge befindet sich im Anhang.

ISSN 0949-8516 (Duisburger Beiträge zur soziologischen Forschung)

Inhalt

1. Einleitung.....	3
2. Statistisches Praktikum als Ergänzung des Soziologiestudiums.....	4
3. Landesdatenbank Nordrhein-Westfalen als Informationsquelle	13
4. Auswertung von Daten der nordrhein-westfälischen Gemeinden.....	19
5. Regressionsanalyse der Ergebnisse politischer Wahlen.....	26
Literaturverzeichnis.....	34
Anhang	

1. Einleitung

Für die empirische Sozialforschung ist die Kenntnis und die Nutzung der amtlichen Statistik unerlässlich. Die Notwendigkeit der Einbeziehung der amtlichen Statistik in das Soziologiestudium bedarf daher keiner weiteren Begründung. In diesem Sinne arbeiten das Fach Soziologie der Gerhard-Mercator-Universität Duisburg bereits seit 1974 und das Sozialwissenschaftliche Institut der Heinrich-Heine-Universität Düsseldorf seit 1983 mit dem Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (LDS) bei der Ausbildung von Studenten zusammen. Die Erfahrungen werden sowohl von den Hochschulen als auch vom LDS positiv beurteilt. Als den Bedürfnissen des Lehrgebietes der empirischen Sozialforschung in besonderem Maße angemessen haben sich dabei vier verschiedene Veranstaltungsformen erwiesen, die unter den Bezeichnungen Datenseminare, Lehrforschungsprojekte, ADV-Lehrgänge und Betriebspraktika zur Anwendung kamen (Gerß 1995 b). „Datenseminare“ behandeln Erhebungsmethoden und Ergebnisse sowie organisatorische und rechtliche Aspekte der amtlichen Statistik als Datenquelle für soziale Indikatoren. Diese Veranstaltungen gehören seit mehreren Jahren in beiden Hochschulen als Hauptseminare mit Leistungsnachweisen durch Referate und deren schriftliche Ausarbeitungen zum regelmäßigen Lehrangebot. „Lehrforschungsprojekte“ nutzen die Ergebnisse der amtlichen Statistik für empirische Untersuchungen zu speziellen sozialwissenschaftlichen Fragestellungen. Solche Projekte wurden vorzugsweise dann als Lehrveranstaltungen angeboten, wenn auf vorliegenden methodologischen Arbeiten aufgebaut werden konnte oder wenn die Ergebnisse des Projekts auch für Zwecke des LDS verwendbar waren. „ADV-Lehrgänge“ leiten zur selbständigen Benutzung von Rechenzentren und mit diesen verbundenen Datenbanken unter Verwendung von Dienst- und Auswertungsprogrammen an. Diese Lehrgänge waren vor allem vor dem „PC-Zeitalter“ wichtig, als die meisten Studenten noch wenig Gelegenheit hatten, in der Hochschule am Bildschirmgerät zu arbeiten (Gerß und Jamke 1989). Seitdem das LDS in der maschinellen Ausstattung keinen wesentlichen Vorsprung mehr vor den Hochschulen hat, haben die ADV-Lehrgänge ihre Bedeutung verloren. „Betriebspraktika“ vermitteln Einblicke in die Organisation und Arbeitsweise statistischer Ämter durch Beteiligung der Studenten am normalen Arbeitsprozeß (Gerß 1995 a). Das LDS hat vor über zehn Jahren damit begonnen, aufgrund spezieller Vereinbarungen mit den Universitäten Duisburg und Düsseldorf interessierten Studenten die Möglichkeit eines mehrwöchigen Vollzeitpraktikums in der amtlichen Statistik anzubieten. Diese Praktika sind Gegenstand der folgenden Darstellung. Dabei wird zunächst auf die Organisation, den Ablauf und den Inhalt des Praktikums in seiner Funktion als Ergänzung des Soziologiestudiums eingegangen. Für den produktiven Einsatz der Praktikanten im LDS – wie für die rationelle empirische Forschung allgemein – ist die Nutzung der Landesdatenbank

Nordrhein-Westfalen als Informationsquelle geboten; dazu ist die Kenntnis der typischen „Spielregeln“ dieser Datenbank notwendig. Die Datenbank ist vor allem als Speicher kleinräumiger Daten – bis hinunter zur regionalen Ebene der Gemeinden – konzipiert. Die Praktikanten können dementsprechend an viele LDS-eigene Arbeiten zur Auswertung von Daten der nordrhein-westfälischen Gemeinden anknüpfen. Schließlich werden als Beispiel für die in einem Praktikum gestellte Aufgabe die Ergebnisse politischer Wahlen – Wahlbeteiligung und Stimmenanteile der Parteien – einer großen Anzahl von demographischen, sozialen, wirtschaftlichen und umweltbezogenen Strukturdaten der Gemeinden gegenübergestellt. Dabei werden für Nordrhein-Westfalen die kleinräumigen Wahlergebnisse des „Superwahljahres“ 1994 (Europawahl, Bundestagswahl, Kommunalwahl) und des Jahres 1995 (Landtagswahl) regressionsanalytisch untersucht.

2. Statistisches Praktikum als Ergänzung des Soziologiestudiums

Im Jahre 1987 wurden in der Universität Duisburg Grundsätze für Praktikumsstellen des Faches Soziologie formuliert, die im folgenden auszugsweise sinngemäß zitiert werden (Strasser 1987). Ausgegangen wird dabei von der Feststellung, daß Soziologie sich als Instrument für die Praxis bewähren muß. Die zu diesem Ausbildungszweck durchzuführenden Praktika zukünftiger Hochschulabsolventen müssen in Hochschule und Praktikumsinstitution organisatorisch verankert werden. Zwischen dem Lehrpersonal der Hochschule und dem Management der die Praktika tragenden Institution sollen persönliche Kontakte bestehen. Durch die damit sichergestellte persönliche Betreuung in Hochschule und Trägerinstitution sollen die Praktika zu einem selektiven Instrument der Studienförderung werden. Praktikumsstellen sollen nicht nur zum Kennenlernen der Praxis, sondern auch zu einer Teilnahme am Arbeitsprozeß führen. Es geht vor allem darum, Einblicke in die Arbeits- und Berufswelt im allgemeinen und Kenntnis über Organisation und Arbeitsweise ausgewählter Berufs- und Tätigkeitsfelder zu erhalten, die im bisherigen Studium erworbenen Kenntnisse und Fähigkeiten durch Lösung konkreter Aufgaben anzuwenden und soziale Kompetenzen - z. B. die Fähigkeit zu Zusammenarbeit und Kommunikation – zu erweitern. Die Notwendigkeit einer Praktikantenausbildung wird mit den bei manchen Soziologen festzustellenden Qualifikationsdefiziten begründet, die punktuell unter anderem dort sichtbar werden, wo empirische Untersuchungen zu einem bestimmten Berufsfeld oder statistische Methoden und Forschungstechniken verlangt werden. Dazu können – sicherlich nicht studienfachspezifische – Defizite in der Fähigkeit zur Informationssammlung und -aufbereitung, in der Artikulationsfähigkeit und im Arbeitsverhalten (Pünktlichkeit, Zuverlässigkeit) kommen. Im Praktikum sollen die Studenten die Arbeits- und Berufswelt zwar ausschnitthaft, aber in diesem Rahmen so umfassend wie möglich kennenlernen. Dabei ist anzustreben, die Tätigkeiten der Praktikanten in

Zusammenhang mit Studieninhalten oder in Verbindung mit dem Thema der Diplom- oder Doktorarbeit auszuwählen. Das Praktikum soll vorzugsweise im fortgeschrittenen Stadium des Studiums (Hauptstudium) absolviert werden. Die Dauer soll sechs Wochen nicht unter- und sechs Monate nicht überschreiten, wobei in einer einleitenden Hospitationsphase Aufgaben zu konzipieren sind, die für den Praktikanten persönlich überschaubar und leistbar sind. Der Praktikant hat nach Abschluß seiner Tätigkeit einen Bericht zu erstellen, der den soziakundlichen Erfahrungsprozeß in der Trägerinstitution, das sachliche Ergebnis der eigenen Arbeit, eine sozialwissenschaftliche Würdigung der eigenen Tätigkeit und Perspektiven für das weitere Studium dokumentieren soll.

In ähnlicher Weise wurden in der Universität Düsseldorf Informationen zur Durchführung von Praktika zusammengestellt (Gernand ohne Jahr). Danach soll – ausgehend von der Frage nach der außeruniversitären Verwertbarkeit soziologischen Wissens – die notwendige Herstellung des Praxisbezugs im Sinne einer Berufsfeld- und Anwendungsorientierung durch studienbegleitende Praktika erreicht werden. Der Nutzen des Praktikums liegt für die Studenten unter anderem in der realistischen Wahrnehmung der Arbeitsweltbedingungen und dem Abbau von Ängsten vor der Berufspraxis, in dem Erwerb vertiefender Kenntnisse über Organisation und Arbeitsweise eines Berufsfeldes, in der Vermittlung außerfachlicher Erfahrungen (interdisziplinäre Kooperations- und Kommunikationsfähigkeit, Verhandlungsfähigkeit u. a.) und in der motivationsfördernden Wirkung für den weiteren Studienverlauf hinsichtlich eines zielstrebigen Abschlusses und stärkerer Beachtung praxisnaher Fragestellungen. Auch für die Anbieter von Praktikantenstellen wird ein Nutzen gesehen; sie lernen die Breite und Gestaltungsfähigkeit einer sozialwissenschaftlichen Ausbildung kennen und können den Kontakt zur Universität verwenden, um bisher nicht oder nur unbefriedigend gelöste Probleme als Fragestellungen an die Forschung einzubringen. Im ersten Teil des mindestens sechswöchigen Praktikums (Hospitationsphase) sollen die Studenten in teilnehmender Beobachtung Handlungsroutinen und Entscheidungsprozesse miterleben und nachvollziehen, um dann im weiteren Ablauf zunehmend eigene Arbeitsaufgaben übernehmen zu können. Die gewonnenen Erfahrungen sollen in einem von dem Praktikanten zu erstellenden Bericht nach Abschluß des Praktikums reflektiert werden.

In beiden Universitäten wurden – zunächst speziell für die soziologischen Abteilungen – Praktikumsbüros eingerichtet. Die dort als „Praktikumsbeauftragte“ eingesetzten Mitarbeiter der Hochschule haben die Aufgabe, für beide Seiten – Studenten und Trägerinstitutionen – als Ansprechpartner für alle das Praktikum und dessen Durchführung betreffenden Bereiche zur Verfügung zu stehen und die Praktika mit den Studenten vor- und nachzubereiten. Die Tätigkeit des Praktikumsbüros wurde von der Universität Düsseldorf ausführlicher beschrieben (Heinrich-Heine-Universität Düsseldorf ohne Jahr). Danach sollen durch das Büro

den Studenten alternative Einsatzmöglichkeiten aufgezeigt werden; gleichzeitig sollen die potentiellen Arbeitgeber mit einem vielleicht ungewohnten, aber dennoch für eine Reihe von Arbeitsfeldern interessanten Qualifikationspotential bekanntgemacht werden. So sollen für beide Seiten befriedigende Ergebnisse erzielbar sein: Die Praktikanten liefern einerseits einen für sie sinnvollen Beitrag zu den Aufgaben ihres Einsatzbereiches, was andererseits für die Arbeitgeber verwertbare Arbeitsergebnisse bedeutet. In der Universität Duisburg wurde das anfänglich auf das Fach Soziologie beschränkte Praktikumsbüro in Kooperation mit dem Arbeitsamt Duisburg unter der Bezeichnung AkZent zu einem fachbereichsübergreifenden akademischen Informations- und Dienstleistungszentrum für den gesamten Themenbereich „Studium und Beruf“ erweitert (Gerhard-Mercator-Universität Duisburg 1997). AkZent befaßt sich vor allem mit der Vorbereitung auf den Berufseinstieg während des Studiums – unter anderem Organisation, Vermittlung und Betreuung außeruniversitärer Praktika – und mit der Karriereplanung und Vermittlung von Absolventen.

Auf der Grundlage der geschilderten Überlegungen wandten sich die Sprecher des Faches Soziologie der Universität Duisburg im Jahre 1988 an das LDS mit dem Wunsch nach Bereitstellung von Praktikantenstellen. Das Ergebnis der darauf folgenden Besprechung wurde von den Fachsprechern in einem Protokoll festgehalten, das anschließend vom LDS bestätigt wurde. Auf diese Weise kam durch zwei sich entsprechende einseitige Erklärungen eine Rahmenvereinbarung zwischen Universität und LDS zustande, ohne daß ein formeller Rahmenvertrag geschlossen werden mußte. Mit diesem Vorgehen wurde das Problem umgangen, daß seitens des LDS ein echter Rahmenvertrag nicht möglich gewesen wäre, weil es im Stellenplan keine als solche bezeichneten Praktikantenstellen gibt. Nach dem Protokoll ist das LDS bereit, ständig bis zu zwei Praktikanten der Universität Duisburg gleichzeitig auszubilden. Die Praktikanten lernen das LDS nach einem Ausbildungsplan ausschnitthaft kennen. Sie unterzeichnen zu diesem Zweck eine arbeitsvertragliche Regelung, die alle wesentlichen Rechte und Pflichten des einzelnen Praktikanten für die Zeit des Praktikums festlegt. Die Praktikanten werden im LDS fachlich einem von Fall zu Fall wechselnden Dezernenten und organisatorisch/dienstrechtlich einer betreuenden Bezugsperson zugeordnet. Der Vergabe der Praktikantenstellen geht eine hochschulinterne Ausschreibung durch Aushang voraus. Die Auswahl erfolgt durch Universität und LDS einvernehmlich. Um die vom LDS im Rahmen dieser Vereinbarung eingegangene Selbstverpflichtung zu erfüllen, werden die Praktikanten in den in einem individuellen Ausbildungsplan festzulegenden Tätigkeitsbereichen von den Gruppenleitern und Dezernenten des LDS unterwiesen und nach Möglichkeit auch zu praktischen Arbeiten eingesetzt, wobei der Aufenthalt in einem einzelnen Bereich mindestens eine Woche dauern soll. In dieser Zeit haben die Praktikanten einen schriftlichen Bericht anzufertigen oder eine sonstige Arbeit zu erledigen, nach der sie beurteilt werden können. Die Dezernenten nehmen die Unterweisung persönlich vor oder überwachen sie, stellen

dem Praktikanten ein Thema zur Bearbeitung und beurteilen ihn selbst und das Arbeitsergebnis. Die erwähnte den Praktikanten während des ganzen Praktikums betreuende Bezugsperson stellt den individuellen Ausbildungsplan auf, koordiniert den Einsatz in den Dezernaten, beurteilt den Praktikanten abschließend und stellt nach Abschluß jedes Praktikums eine benotete Bescheinigung des LDS aus.

Im Jahre 1991 wandte sich der Sprecher des Sozialwissenschaftlichen Instituts der Universität Düsseldorf an das LDS mit der Bitte, neben der bekannt gewordenen Kooperation mit der Universität Duisburg auch Studenten der Universität Düsseldorf Praktika zu ermöglichen. Wie im Falle Duisburg – d. h. durch ein vom Hochschulvertreter angefertigtes und anschließend vom LDS bestätigtes Besprechungsprotokoll – wurde daraufhin eine Rahmenvereinbarung gleichen Inhalts zwischen der Universität Düsseldorf und dem LDS getroffen; das LDS ist also bereit, auch im Falle Düsseldorf ständig bis zu zwei Praktikanten gleichzeitig auszubilden.

Zwischen den einzelnen Studenten beider Hochschulen und dem LDS wird jeweils ein bis auf die Terminierung des Praktikums inhaltsgleicher „Vertrag über ein sozialwissenschaftliches Praktikum“ abgeschlossen. Der Vertrag legt Beginn und Ende des Praktikums fest und erklärt die ersten beiden Wochen zur Probezeit, in der beide Seiten jederzeit vom Vertrag zurücktreten können.

Das LDS erklärt sich bereit, . . .

- ... den Studenten für die Dauer des Praktikums zu unterweisen;
- ... in allen den Studenten betreffenden Fragen der Durchführung des Praktikums mit dem Praktikumsbeauftragten der Universität zusammenzuarbeiten;
- ... auf die Eignung des Studenten für das Praktikum zu achten und ggf. mit ihm und dem Praktikumsbeauftragten über die Zweckmäßigkeit der Fortsetzung des Praktikums zu sprechen;
- ... den Praktikumsbeauftragten von einer vorzeitigen Beendigung des Vertrages oder vom Nichtantritt des Praktikums zu unterrichten;
- ... dem Studenten nach Beendigung der praktischen Tätigkeit eine Bescheinigung (Bewertung) zu erteilen, die Angaben über Dauer, Inhalt und Erfolg des Praktikums enthält;
- ... den vom Studenten über den Verlauf des Praktikums zu fertigenden schriftlichen Auswertungsbericht sachlich zu prüfen.

Der Student verpflichtet sich, . . .

- ... die ihm übertragenen Aufgaben sorgfältig und gewissenhaft auszuführen und alle gebotenen Ausbildungsmöglichkeiten wahrzunehmen;
- ... die vom LDS im Rahmen der Ausbildung erteilten Weisungen zu befolgen;

- ... die geltenden Ordnungen des LDS, über die der Student zu Beginn des Praktikums belehrt wurde, zu beachten;
- ... die betriebsübliche Arbeitszeit einzuhalten und bei Fernbleiben das jeweilige Beschäftigungsdezernat unverzüglich zu benachrichtigen;
- ... Materialien und Geräte, die zur Verfügung gestellt werden, sorgfältig zu behandeln;
- ... Vorgänge und Informationen, die ihrer Natur nach oder aufgrund besonderer Anordnung der Geheimhaltungspflicht unterliegen, geheimzuhalten (dies gilt auch nach Ablauf des Praktikums) und die geltenden Datenschutzgesetze zu beachten;
- ... den schriftlichen Praktikumsbericht dem LDS zur Prüfung und Genehmigung vorzulegen.

Der Vertrag bestimmt ferner, daß eine Vergütung vom LDS nicht gezahlt wird. Eine vorzeitige Auflösung des Vertrages ist beiderseitig innerhalb der Probezeit ohne Einhaltung einer Frist und durch den Studenten bei Aufgabe oder Änderung des Ausbildungszieles mit einer Frist von vier Wochen möglich. Die Auflösung wird schriftlich unter Angabe der Auflösungsgründe erklärt. Dem Praktikumsbeauftragten der Universität ist unverzüglich eine Abschrift der Erklärung zu übersenden. Neben den Vertragspartnern erhält auch der Praktikumsbeauftragte eine Ausfertigung des Vertrages; dies gilt auch für das vom LDS auszustellende Zeugnis. Dieses Zeugnis enthält stets einen allgemeinen Teil, in der Regel einen auf die Tätigkeit des jeweiligen Praktikanten ausgerichteten individuellen Teil und eine Gesamtnote nach der Schulnotenskala. Der allgemeine Teil lautet: „Herr (Frau).....hat als Student(in) der Universität ... vom ... bis zum ... (... ganze Arbeitstage) im Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (LDS) in Düsseldorf ein sozial- und wirtschaftsstatistisches wissenschaftliches Praktikum absolviert, das sowohl Information als auch eigene produktive Tätigkeit umfaßte.“ Der individuelle Teil des Zeugnisses enthält die Bezeichnung des Einsatzbereiches innerhalb des LDS, eine Beschreibung der dort vom Studenten ausgeführten speziellen Tätigkeit und eine verbale Beurteilung seiner Leistung und seines Verhaltens.

Bereits im Zusammenhang mit den Rahmenvereinbarungen wurden die möglichen Einsatzstellen für die Praktika von Soziologiestudenten im LDS umschrieben. Diese Einsatzstellen sind im folgenden in Anlehnung an den Geschäftsverteilungsplan des LDS stichwortartig aufgelistet. Dabei handelt es sich um ein flexibles Programm, das Schwerpunktsetzungen nach der jeweiligen Kapazität des LDS unter weitestmöglicher Berücksichtigung der speziellen Interessen der Praktikanten offen läßt.

- Einzelne Statistiken mit besonderer Relevanz für die Soziologie: Volkszählung; Mikrozensus; Bevölkerungsprognosen; Wahlen; Beschäftigungsprognosen.
- Zur Zusammenarbeit des LDS mit dem Statistischen Bundesamt und anderen statistischen Landesämtern: Kostenkalkulation und Terminplanung;

Verbundprogrammierung; Kartei der Unternehmen und Betriebe des Produzierenden Gewerbes; Arbeitskreis Volkswirtschaftliche Gesamtrechnungen der Bundesländer.

- Zur Zusammenarbeit des LDS mit der Kommunalverwaltung (kommunalstatistische Ämter u. a.): Erhebungsmethodik der laufenden Bevölkerungsstatistik; Aufbereitung der laufenden Bevölkerungsstatistik; Bautätigkeitsstatistik; Preisstatistik; Gemeindefinanzstatistik.
- Zur Zusammenarbeit des LDS mit anderen Landesbehörden (v. a. Ministerien): Zahlgeschäfte und andere Verwaltungsaufgaben; Erhebungsmethodik der Statistiken der Sozialleistungen und der Rechtspflege; Aufbereitung der Statistiken der Sozialleistungen und der Rechtspflege; Schulstatistik; Hochschulstatistik; Steuerstatistik; Datei der Zweckzuwendungen aus öffentlichen Mitteln; ADV für die Landtagsverwaltung; ADV für die Schulverwaltung; ADV für die Hochschulverwaltung.
- Anwendung der Stichprobentheorie: Stichprobenpläne; Fehlerrechnungen.
- Hard- und Software-Arbeiten: Technische Planung; Anwendersoftware, Dateienorganisation und Datenabruf; Standardsoftware; spezielle mathematisch-wissenschaftliche Software; Software für Planungstechniken; Hard- und Software zur graphischen Datenverarbeitung.
- Zum Datenverbund: Datenübertragung; Datenfernverarbeitung.
- Zur Archivierung und Veröffentlichung von statistischen Ergebnissen: Datenbanknutzung; Datenbanktechnik; gedruckte Veröffentlichungen; elektronische Information.
- Zum Datenschutz: Objektschutz; Rechtsfragen zu Datenschutz und -sicherung; technische Fragen zu Datenschutz und -sicherung.

Die Unterweisung der Praktikanten über die in dieser Auflistung mit einbezogenen Arbeiten der LDS-Abteilung „Landesdatenverarbeitungszentrale“ wurde bei der Durchführung der einzelnen Praktika modifiziert. Da die Studenten der Sozialwissenschaften in der Regel nicht über genügend fundierte Kenntnisse insbesondere der mehr technischen Seite der Datenverarbeitung verfügen, erschien die Möglichkeit einer praktischen Mitarbeit hier wenig wahrscheinlich; eine Beurteilung der Studenten wäre daher nicht möglich gewesen. Um den Studenten trotzdem einen gewissen Einblick in Fragestellungen der Datenverarbeitung zu geben, wurde vorgesehen, einen zusammenhängenden einwöchigen Informationsrundgang durch ausgewählte Bereiche der Landesdatenverarbeitungszentrale – gelegentlich einschließlich des Bereichs Produktionssteuerung und Maschineneinsatz im eigentlichen Rechenzentrum – durchzuführen. Da auf diese Weise nur eine passive Information ohne sichtbare eigene Leistung der Praktikanten erreichbar ist, bleibt diese „ADV-Woche“ aus der Beurteilung des Praktikumserfolgs ausgeschlossen.

Um den Studenten ihre Entscheidung für die Bewerbung um ein Praktikum zu

erleichtern, wird seit 1991 einmal jährlich ein halbtägiges Programm für den Besuch des LDS angeboten. Das Interesse für diese Veranstaltung hat kontinuierlich auf bis zu ca. 130 Teilnehmer zugenommen. Das Besuchsprogramm enthält einen ständigen und einen variablen Teil. Der ständige Teil umfaßt meist die Vorführung eines Films, der allgemeinverständlich beispielhaft zeigt, „wzu Statistiken gut sind“, sowie Vorträge über „Zugangswege zu statistischen Ergebnissen (Veröffentlichungen und Landesdatenbank)“ und „Zugangsmöglichkeiten, Einstellungsvoraussetzungen und Qualifikationsanforderungen für Bewerber(innen) des höheren Dienstes“. Der variable Teil besteht aus thematisch wechselnden Kurzreferaten und der Beantwortung von Fragen zu den Aufgaben des LDS unter besonderer Berücksichtigung von Einsatzmöglichkeiten für Sozial- und Wirtschaftswissenschaftler. Den bekanntgewordenen Wünschen der die Studenten begleitenden Hochschullehrer entsprechend wurden vor allem die folgenden Themen behandelt: Statistiken der Bevölkerung, der Gesundheit und der Wahlen; Bevölkerungsprognosen; Mikrozensus; Statistiken des Bildungswesens; Quantifizierung des gesamtwirtschaftlichen Arbeitsplatzangebots; Statistiken der Preise und der Lebenshaltung; Statistiken des Handels und des Verkehrs; Steuerstatistiken; Statistiken des Umweltschutzes; Organisationsstruktur des LDS. Wenn genügend Zeit übrigblieb, wurde der Besuch durch eine Besichtigung – vorzugweise der graphischen Datenverarbeitung – abgeschlossen.

Die im LDS beschäftigten Praktikanten entschieden sich weit überwiegend für die vorgegebene Mindestdauer des Praktikums (sechs Wochen bzw. ca. 30 ganze Arbeitstage); in Einzelfällen dauerten die Praktika bis zu zehn Wochen. Mit sehr wenigen Ausnahmen hielten die Praktikanten die ganze geplante Zeit durch. In den Ausnahmefällen wurden die Praktika durch Kündigung seitens der Praktikanten abgebrochen. Die Kündigung erfolgte teilweise aufgrund einleuchtender objektiver Hinderungsgründe, die auf der Seite des Praktikanten unerwartet eintraten. Es kam auch vor, daß ein Praktikant offensichtlich zwar der aus dem normalen Studienbetrieb gewohnten einseitigen Informationsaufnahme in der ADV-Woche, aber nicht der notwendigen zweiseitigen Kommunikation bei der praktischen Arbeit psychisch gewachsen war. Hier wurde das Praktikum kennzeichnenderweise nach der ersten Woche – eben der ADV-Woche – abgebrochen. Das Spektrum der den Praktikanten vom LDS in Aussicht gestellten Einsatzbereiche wurde bisher nicht vollständig genutzt. Vielmehr konzentrierte sich der Einsatz auf bestimmte Bereiche, was nicht auf Kapazitätsengpässe im LDS, sondern auf die speziellen Wünsche der Praktikanten aus der Sicht des Soziologiestudiums zurückzuführen und daher verständlich ist. Die bei weitem bevorzugten Einsatzbereiche sind demnach der Mikrozensus und die Bevölkerungsstatistik, letztere mit besonderer Vorliebe für die Bevölkerungsprognose. Weitere beliebte Bereiche sind vor allem der Umgang mit Datenbank und statistischen Auswertungsprogrammen, die Gesundheitsstatistik, die Erwerbstätigenberechnung und die Umweltschutzstatistik. Speziellere Wünsche der Praktikanten richteten sich insbesondere auf den Einsatz in der Bildungsstatistik, der

Fremdenverkehrsstatistik und der Sozialproduktberechnung. Den Praktikanten wurde grundsätzlich angeboten, die von ihnen zur Beurteilung vorzulegende schriftliche Ausarbeitung entweder in der Form eines Berichts über den Verlauf des Praktikums oder in der Form einer thematischen Abhandlung anzufertigen. Die Praktikanten entschieden sich für die beiden Varianten etwa gleich häufig. Der Praktikumsbericht wurde dabei gelegentlich angereichert, zum Beispiel mit Kommentaren aus persönlicher Sicht, mit inhaltlichen Aussagen zu Methoden und mit einer allgemeinen Beschreibung der Aufgaben des LDS. Soweit die schriftliche Leistung durch eine Abhandlung erbracht wurde, wurden unter anderem die folgenden Themen behandelt: Umweltschutzinvestitionen; Erwerbstätigenzahl in sektoraler und regionaler Gliederung; Prognosen im Bildungsbereich; Entwicklung des Fremdenverkehrs; Methodendiskussion mit eigenen Vorschlägen zum Mikrozensus; Morbiditätsstatistiken (Methodendokumentation, Ergebnisinterpretation, persönliche Erfahrungen); Vergleichsanalyse von Bevölkerungsprognosen (insbesondere Annahmendiskussion verschiedener Ansätze); Arbeitslosigkeit in Nordrhein-Westfalen; Beziehungen zwischen dem Mikrozensus und der Einkommens- und Verbrauchsstichprobe; Erwerbstätigkeit und -einkommen von Frauen und Männern; europäisches Haushaltspanel. Wenn diese Ausarbeitungen inhaltlich und formal bestimmte Anforderungen erfüllten und thematisch in das aktuelle Veröffentlichungsprogramm des LDS paßten, konnte die Veröffentlichung unter den Namen der Verfasser in der Zeitschrift des LDS in Aussicht gestellt werden. So wurde die von einer Praktikantin erstellte deskriptive Analyse des Fremdenverkehrs in Nordrhein-Westfalen veröffentlicht (Bentgens 1992). Auch die faktorenanalytische Bearbeitung von Daten für die Gemeinden des Landes Nordrhein-Westfalen wurde in die Zeitschrift aufgenommen (Düntgen 1997). Die große Menge und die Vielfalt der Daten der amtlichen Statistik bieten eine günstige Möglichkeit, das Praktikum im LDS auch als Einstieg in eine Diplomarbeit zu nutzen. So wurde ein Diplomand durch den Abschluß eines Praktikantenvertrages in die Lage versetzt, für seine Diplomarbeit über kleinräumige Unterschiede der Mortalität sehr detaillierte Daten aus dem Mikrozensus behördintern auswerten zu können (Drogies 1998).

Die den Praktikanten erteilten Noten fielen im allgemeinen bemerkenswert gut aus, was einer Erklärung bedarf. Das Praktikum ist für die Soziologiestudenten kein Pflichtbestandteil des Studiums (abgesehen von gelegentlich vorkommenden Ausnahmefällen, wie bei der Fächerkombination „Sozialwissenschaft“ und „Informationswissenschaft“ in der Universität Düsseldorf). Die im LDS tätigen Praktikanten erhalten dafür weder eine Vergütung noch Fahrkostenersatz. Aus diesen Bedingungen folgt, daß sich fast ausnahmslos Studenten mit ausgeprägtem Interesse an der amtlichen Statistik bewerben, die zudem bereits ziemlich genau wissen, was sie mit dem Praktikum erreichen wollen (Kennzeichnenderweise wurde ein Praktikant unmittelbar nach Abschluß seines im Mikrozensus absolvierten

Praktikums für diese Statistik als Interviewer beschäftigt). Die gute Beurteilung ist vor diesem Hintergrund verständlich.

3. Landesdatenbank Nordrhein-Westfalen als Informationsquelle

Ein wichtiger Bereich der Praktikantenausbildung bezieht sich auf den Umgang mit der im LDS geführten Landesdatenbank Nordrhein-Westfalen. Dementsprechend wurden mehrere Berichte zu den im Praktikum erworbenen Kenntnissen über die Datenbank und die bei der Benutzung dieses Instrumentes angefallenen eigenen Erfahrungen verfaßt. Je nach Studienschwerpunkt und persönlichem Interesse wurden dabei verschiedene Aspekte in den Vordergrund gestellt. Die vorwiegend informationswissenschaftliche Ausrichtung führte zu einer Betonung der formalen Seite der Informationsverarbeitung und Datenbankrecherche, ohne daß dabei auf den Inhalt der gespeicherten Informationen eingegangen wurde (Freimann 1993). In einem anderen Fall wurden die Verfahren des Datenabrufs an einem von der Praktikantin selbst gewählten inhaltlichen Schwerpunkt – der Arbeitslosigkeit – demonstriert (Werner 1997). Eine für den Praktikanten und das betreuende Dezernat gleichermaßen besonders ergiebige Situation trat ein, wenn die Datenbankschulung in ohnehin vorgesehene sozialwissenschaftliche Auswertungsarbeiten des LDS eingebunden werden konnte. Dieser Fall war bei der regionalanalytischen Untersuchung von Ergebnissen politischer Wahlen gegeben (Weinem 1998), die in den folgenden Abschnitten dargestellt wird. Alle Praktikanten, zu deren Ausbildungsplan die Benutzung der Datenbank gehörte, nahmen nach Möglichkeit an einem meist einwöchigen Datenbank-Intensivkurs teil, der für die Bediensteten des LDS regelmäßig stattfindet.

Die Anfänge der Landesdatenbank Nordrhein-Westfalen liegen über dreißig Jahre zurück (Plewa 1985). Im Jahr 1967 wurde von der Landesregierung der Auftrag zum Aufbau einer „Strukturdatenbank“ erteilt. Der Anlaß war die Aufstellung eines Generalverkehrsplans, wozu umfangreiche Datenbestände zusammengetragen werden mußten. Diese Bestände umfaßten nicht nur verkehrsspezifische, sondern auch allgemeine soziodemographische und ökonomische Daten. Alle Daten sollten auf maschinenlesbaren Datenträgern zur beliebigen Weiterverarbeitung verfügbar sein. Über den befristeten Bedarf des Generalverkehrsplans hinaus sollten die Daten der dauerhaften Unterstützung von Planungs- und Entscheidungsprozessen dienen. Die Realisierung des Generalverkehrsplans mußte aus Mangel an Haushaltsmitteln bald darauf eingestellt werden, die Strukturdatenbank hat jedoch überlebt. Diese Datenbank wurde im damaligen Statistischen Landesamt Nordrhein-Westfalen eingerichtet. Diese Behörde lieferte den weitaus größten Teil der benötigten Daten und besaß ein leistungsfähiges Rechenzentrum mit erfahrenem Personal. Die günstigen Bedingungen im Landesamt waren notwendig, weil Nordrhein-Westfalen als erstes Bundesland – und auch früher als der Bund – ein solches Vorhaben in Angriff nahm und weil die von den Software-Lieferanten angebotenen Datenbanksysteme auf die internen Informationsbedürfnisse von Unternehmen zugeschnitten, aber für die hier zu verfolgenden Zwecke nicht verwendbar waren. Somit mußte ein hauseigenes System entwickelt werden. Dabei waren Regelungen

über alle im Zusammenhang mit der Datenbank stehenden technischen Aufgaben sowie über die fachliche Betreuung – Analyse des statistischen Informationsbedarfs, Datenbeschaffung, -übernahme und -dokumentation, Erledigung der Benutzeranfragen – zu treffen. Die gesetzliche Grundlage der Datenbank wurde einige Jahre später mit dem „Gesetz über die Organisation der automatisierten Datenverarbeitung in Nordrhein-Westfalen“ (ADVG) vom 12.2.1974 geschaffen. Dieses Gesetz führte die offizielle Bezeichnung „Landesdatenbank“ ein und änderte den Namen der Behörde von „Statistisches Landesamt“ in „Landesamt für Datenverarbeitung und Statistik“ (LDS). Das ADVG schreibt vor, daß „ausgewählte Daten“ in die Datenbank einzuspeichern sind, wobei aber das LDS in Detailfragen der Datenauswahl nicht gebunden ist. Zur Datenauswahl mußte ursprünglich ein aus Vertretern des Landtags, der Landesverwaltung und der kommunalen Verwaltungen zusammengesetzter Beirat gehört werden. In der geltenden Neufassung des ADVG vom 9.1.1985 (Gesetz- und Verordnungsblatt für das Land Nordrhein-Westfalen S. 41) ist der Beirat nicht mehr enthalten. Nach wie vor gesetzlich vorgeschrieben ist, daß die Landesdatenbank grundsätzlich jedermann zur Verfügung steht. Dabei sind die üblichen statistischen Geheimhaltungsbestimmungen – Verbot der Weitergabe von Daten über einzelne Personen oder Institutionen sowie der Weitergabe von Daten, aus denen solche Einzelangaben rekonstruiert werden können – und die Datenschutzbestimmungen – Verbot der Speicherung personenbezogener Einzeldaten – einzuhalten. Das Recht zum selbständigen Direktzugriff auf die Daten der Landesdatenbank wird vom ADVG ausdrücklich neben der Landesregierung auch dem Landtag zugesprochen.

Die Auswahl der in die Landesdatenbank aufzunehmenden Daten erfolgt benutzerorientiert aufgrund von Bedarfsermittlungen bei den potentiellen Hauptnutzern. Im übrigen werden die folgenden Rahmenbedingungen beachtet:

- Die einzuspeichernden Daten sollen auf regional niedriger Aggregationsstufe – nach Möglichkeit auf Gemeinde-, zumindest aber auf Kreisebene – verfügbar sein, um der detaillierten regionalstatistischen Analyse und bei Bedarf der Umrechnung auf nichtadministrative Gebietseinheiten dienen zu können.
- Neben den Daten aus den in mehrjährigem Turnus durchgeführten Großzählungen (Volkszählung, Arbeitsstättenzählung, Wohnungszählung u. a.) sollen zur Ermöglichung von Zeitreihenanalysen lange Reihen mit jährlicher Periodizität aufgebaut werden.
- Soweit verfügbar, sollen aus allen wichtigen Lebensbereichen Daten gespeichert werden, die nicht unbedingt aus der amtlichen Statistik stammen müssen, sondern auch aus anderen Quellen übernommen werden können.

Die drei genannten Bedingungen sollen soweit wie möglich gleichzeitig erfüllt werden; d. h. angestrebt wird die Speicherung von jährlichen Daten zu verschiedensten Merkmalen für die einzelnen Gemeinden. Für besondere Bedürfnisse – insbesondere der kurzfristigen Wirtschaftsbeobachtung – werden daneben Daten mit monatlicher oder vierteljährlicher Periodizität gespeichert, wobei hier auf die

regionale Aufgliederung der Landeswerte verzichtet wird. Die in der Datenbank vorrangig berücksichtigten Merkmalsbereiche sind: Natürliche Umwelt; Bevölkerung; Gesundheitswesen; Unterricht, Bildung und Kultur; politische Wahlen; Erwerbstätigkeit; Land- und Forstwirtschaft; Unternehmen und Arbeitsstätten; produzierendes Gewerbe; Umweltschutz; Bautätigkeit, Gebäude und Wohnungen; Handel, Gastgewerbe und Fremdenverkehr; Verkehr; Geld- und Kreditwesen; öffentliche Sozialleistungen; öffentliche Finanzen und Steuern; Preise; Löhne und Gehälter; volkswirtschaftliche Gesamtrechnungen. Die zur Einspeicherung in die Landesdatenbank ausgewählten Daten stammen zu etwa 90% aus dem LDS selbst. Die verbleibenden 10% extern bezogener Daten stammen ebenfalls von Trägern der amtlichen Statistik wie dem Kraftfahrtbundesamt und dem Landesarbeitsamt. Bereits in der Phase der Übernahme der Daten in die Datenbank – also nicht erst später beim Abruf der gespeicherten Daten – wird die Einhaltung der statistischen Geheimhaltungsbestimmungen geprüft. Dafür stehen spezielle EDV-Programme zur Verfügung, die sicherstellen, daß außer den Angaben, die sich auf ein oder zwei Berichtsstellen beziehen, auch aggregierte Angaben von mehreren Berichtsstellen dann geheimgehalten werden, wenn der Anteil einer Berichtsstelle deutlich überwiegt. Die Geheimhaltungsprogramme verhindern auch den Rückschluß auf geheimzuhaltende Angaben durch Differenzbildung, indem sie weitere Datenwerte sperren. In der Landesdatenbank sind nach dem gegenwärtigen Stand zu ca. 5 Mill. Merkmalskombinationen 3 Mrd. Daten sowie 5 Mill. Zeitreihenglieder gespeichert. Datenspeicherungen erfolgen im allgemeinen ad hoc, im Bedarfsfall täglich. Die Aktualität des Datenbestands entspricht meist dem Aufbereitungsstand der zugrundeliegenden Fachstatistik.

Wenn ein Interessent sich mit einem von der Datenbank zu erfüllenden Datenwunsch an das LDS wendet, so muß er nicht selbst die Datenbanktechnik benutzen. Sein Auftrag wird vielmehr bis zum vereinbarten Endprodukt von den Mitarbeitern der im LDS bestehenden „Leitstelle für Information und Landesdatenbank“ ausgeführt. Dieser als indirekt zu bezeichnende Zugang zur Datenbank erfordert vom Empfänger der Daten keine Kenntnisse über den Umgang mit der EDV im allgemeinen und der Landesdatenbank im besonderen. Die Leitstelle übernimmt die gesamte technische Umsetzung der keiner Formvorschrift unterliegenden Bestellung. Der indirekte Zugang war in den Jahren der Aufbauphase der einzige mögliche und wird nach wie vor von den weitaus meisten Kunden der Landesdatenbank beschritten. Daneben gibt es eine verhältnismäßig geringe, aber allmählich zunehmende Anzahl von Anwendern im direkten Zugang. Dazu wird eine benutzerseitig installierte Datenstation – bestehend aus Bildschirmgerät und Drucker – benötigt, die über eine festgeschaltete oder anwählbare Postleitung zur Landesdatenverarbeitungszentrale im LDS in Verbindung steht und den dortigen Großrechner im Teilnehmerbetrieb benutzt. Die ersten auf diese Weise direkt an die Datenbank angeschlossenen Stellen waren der Landtag, verschiedene Ministerien, die Bezirksregierungen in ihrer Funktion als Bezirksplanungsbehörden, der Kommunalverband Ruhrgebiet und einige

wissenschaftliche Institute. Zu den Hilfsmitteln für die Datenbanknutzer gehört ein Definitionskatalog, der die in der Datenbank vorkommenden Begriffe so weitgehend erläutert, daß dem Benutzer eine sachgerechte Interpretation der Daten möglich wird. Unverzichtbar ist der sog. Datenbestandskatalog. Dieser dokumentiert den gesamten Datenbestand der Landesdatenbank in sachlicher und zeitlicher Gliederung und ist somit ihr systematisches Inhaltsverzeichnis. Der Direktnutzer muß die im Datenbestandskatalog enthaltenen Abrufnummern bei der Formulierung seiner Abrufspezifikation verwenden. Anstelle der früheren Buch- oder Mikrofiche-Fassungen des Datenbestandskatalogs ermöglichen komfortable Programme dem Bearbeiter am Bildschirm im Dialog mit dem System, in flexibler Weise Katalogauszüge zu erstellen, die unter Berücksichtigung eingegebener Randbedingungen die zur jeweiligen Problemlösung benötigten Datenbestände herausfiltern und beschreiben.

Die Komponenten eines Datenbanksystems sind die Datenbasis, der Computer und die Software. Die letztgenannte Komponente umfaßt die Computerprogramme, die zur Verwaltung der Datenbestände sowie zu ihrer Bearbeitung und Verrechnung und zur Darstellung der aus der Datenbank gewonnenen Informationen zur Verfügung stehen. Die anspruchsvollen und leistungsfähigen Auswertungsprogramme machen den eigentlichen Wert eines Datenbanksystems aus. Neben den zahlreichen Dienstprogrammen, durch die die Arbeitsmöglichkeiten der Direktnutzer ständig erweitert, verbessert und komfortabler gestaltet werden, gehören dazu insbesondere die Verfahren zur rechnerisch-analytischen Auswertung und zur anschaulichen Präsentation. Die in der Landesdatenbank verfügbaren Analyseprogramme wurden zum Teil selbst entwickelt und zum Teil auf dem Software-Markt erworben. Die datenbankeigenen Programme enthalten – modular aufgebaut und beliebig miteinander verknüpfbar – alle Operationen der Matrizenrechnung, wobei die Matrixelemente einzeln, als Vektor oder als Intervall bearbeitbar sind. Aufgrund dieses Vorrats an Funktionen zur beliebigen Ausgestaltung maßgeschneideter eigener Methoden durch den Benutzer konnte auf die feste Installation fachspezifischer Funktionsprogramme verzichtet werden. Im übrigen kann der Benutzer über den Daten- und Funktionsbestand der Landesdatenbank hinaus auch seine eigenen Daten und Funktionen einbringen und von der Datenbank verarbeiten lassen. Für analytisch- statistische Auswertungsaufgaben bietet die Datenbank die Programmpakete SPSS und SAS an. Zur Analyse von Zeitreihen ist das sog. Census-X11-Verfahren aufrufbar. Nach Abschluß der Datenverrechnung können die Ergebnisse entsprechend den Wünschen des Benutzers in frei wählbarer Form ausgegeben werden. Am häufigsten wird die Form der Tabelle gewünscht, die nach individueller Konzeption – auch entsprechend den Standards der DIN 55301 „Richtlinien für die Gestaltung statistischer Tabellen“ sowie verwendbar als Druckvorlage für jede Art von Publikation – ein- oder mehrfarbig ausgedruckt werden kann. Zum Leistungskatalog der Datenbank gehören auch graphische Darstellungsmöglichkeiten. Als sog. „Business Graphics“ können Verlaufskurven

und Kreis- oder Säulendiagramme erzeugt und ausgedruckt werden. Für höhere Ansprüche an die Gestaltung und zeichnerische Qualität können die aus der Datenbank bereitgestellten Daten offline auf einem Plotter graphisch weiterverarbeitet werden. Die Herstellung thematischer Karten für Verwaltungsbezirke (Gemeinden, Kreise, Arbeitsamtsbezirke) oder für nichtadministrative Raumeinheiten (z.B. Arbeitsmarktregionen) ist online realisierbar, wobei Klassengrenzen, Schraffuren, Farben, Maßstabsveränderungen und Ausschnittvergrößerungen frei gestaltet werden können. Außer in ausgedruckter Form werden alle Produkte der Landesdatenbank auf Wunsch auch auf maschinenlesbaren Datenträgern zur weiteren Verarbeitung durch den Benutzer zur Verfügung gestellt. Die Datenträger können Magnetbänder oder Disketten sein. Unabhängig von der Form der Datenpräsentation und vom gewählten Datenträger werden den Benutzern der Datenbank auf der Basis der abgerufenen Datenmenge Kosten berechnet. Außerhalb der Datenbank anfallende Verfahrensanwendungen (z.B. Plotter-Einsatz) werden unabhängig von der Datenmenge individuell kalkuliert. Nur für die Behörden und Einrichtungen des Landes Nordrhein-Westfalen ist die Benutzung der Landesdatenbank nach einem Erlaß des für die Dienstaufsicht über das LDS zuständigen Innenministeriums unentgeltlich.

Die Landesdatenbank hat sich bereits seit vielen Jahren zu einem unverzichtbaren Instrument der automatisierten Informationsbereitstellung für öffentliche Verwaltung, Wirtschaft, Wissenschaft und Einzelpersonen entwickelt. Technisch ist sie auf einem Informationsverarbeitungssystem der Herstellerfirma IBM implementiert. Hierbei handelt es sich um ein Großsystem des Typs 9021-832, das 180 Mill. Befehle in einer Sekunde ausführen kann. Die Frage, ob der mit großen Anstrengungen über Jahre vorgenommene Aufbau der Landesdatenbank erfolgreich war, ist eindeutig zu bejahen. Die Datenbank hat sich am Markt durchgesetzt. Die Frequentierung hat seit Einrichtung der Datenbank fast immer kontinuierlich zugenommen. Die stärkste Nutzergruppe ist nach wie vor die Landesverwaltung. Zunehmendes Interesse haben aber auch die Kommunalverwaltung, Unternehmen und wissenschaftliche Institutionen.

Die Ergiebigkeit der Landesdatenbank für die Praktikantenausbildung im LDS resultiert daher, daß dieses Instrument einerseits für die rationelle Auswertung der amtlichen statistischen Daten zur empirischen Forschung unentbehrlich ist und andererseits hohe Anforderungen an den Direktnutzer stellt. Der selbständige Umgang mit der Landesdatenbank ist so schwierig, daß im allgemeinen eine Einarbeitungszeit von etwa zehn Tagen erforderlich ist. Im folgenden wird der Bericht des in der Phase der Datenbeschaffung und -bereitstellung an der Analyse der Wahlergebnisse beteiligten Praktikanten über seine Arbeit mit der Datenbank in gekürzter Form sinngemäß wiedergegeben (Weinem 1998). Aufgabe war es, für jede der 396 Gemeinden des Landes Nordrhein-Westfalen zu jeweils einem bestimmten Berichtsjahr 120 Merkmalswerte zusammenzustellen. Die Merkmale mußten zum

größten Teil für den Untersuchungszweck passend umgerechnet werden. Im Datenbestandskatalog der Landesdatenbank ist jedem Merkmal eine Datenbeschreibung zugeordnet, die Angaben enthält zu:

- Zeit (kennzeichnet den zeitlichen Bezug der Daten)
- Wertart (gibt an, ob es sich um Stichtagswerte, Monatswerte usw. handelt)
- Datenquelle (gibt an, wer die Daten liefert hat)
- Ermittlungsverfahren (verweist auf Details des Arbeitsprogramms des LDS)
- Datentiefe (gibt die kleinste Gebietseinheit an, zu der Daten vorliegen)
- Merkmalsträgergruppe (erläutert die Art der regionalen Aufgliederung des Landes)
- Maßeinheit (Anzahl, m^2 , DM usw.)
- Dezimalcharakter (gibt die Stellengenauigkeit an)
- Datum des Gebietsstandes (Zeitpunkt der letzten Veränderung)

Zur Lösung der gestellten Aufgabe mußten im wesentlichen vier Funktionen der Datenbank (DATENBER, RECH, SFAUS, UMFORM) angewendet werden. Mit der Funktion DATENBER werden die benötigten Daten recherchiert und in einer temporären Arbeitsdatei zur Weiterverarbeitung bereitgestellt. Dabei kann man sich einen detaillierten Überblick über die gespeicherten Daten verschaffen. Die Arbeitsdatei geht zwar als solche nach Beendigung der Datenbanksitzung verloren, kann aber als Prozedurdatei gesichert werden. Zu den abgespeicherten Prozeduren können auch Kommentare eingegeben werden. Die Einsatzmöglichkeiten der Funktion RECH sind sehr vielfältig, die Handhabung ist aber für den Benutzer zum Teil gewöhnungsbedürftig. Im hier beschriebenen Fall wurde diese Funktion hauptsächlich zur spaltenweisen Berechnung verwendet. Wenn zum Beispiel durch DATENBER zwei Spalten mit Merkmalswerten bereitgestellt werden, muß zur Funktion RECH zunächst angegeben werden, daß genau mit diesen Spalten und diesen Werten weitergearbeitet werden soll. Mit RECH können dann z. B. die Werte zweier Spalten addiert oder durcheinander dividiert werden; die Summen und Quotienten stehen jeweils in einer neugebildeten dritten Spalte. Eigenwilliger wird es, wenn mit Nachkommastellen und Rundungen gearbeitet wird. So könnten z. B. zwei Nachkommastellen gesetzt werden, wobei die letzte Stelle zu runden ist. Der Anwender muß dann selbst die Nachkommastellen setzen und durch Multiplikation mit Tausend (auf der rechten Seite des Kommas) dafür sorgen, daß drei zusätzliche Stellen bereitstehen, und zwar eine Stelle zum Wegrunden und zwei, die als Nachkommastellen angezeigt werden. RECH arbeitet demnach nicht wie ein Taschenrechner. Mit der Funktion SFAUS können die Ergebnisse einer Verarbeitungsprozedur am Bildschirm angezeigt oder angedruckt werden, wobei verschiedene Optionen (Zeilen- oder Spaltenanordnung, Sortierung, Spaltenbreite usw.) wählbar sind. Es ist zweckmäßig, nach jedem DATENBER und jedem RECH ein SFAUS zu einer Bildschirmanzeige anzuwenden, um die Ausgangsdaten sofort mit den Rechenergebnissen zu vergleichen. Da beim Aufbau längerer Prozeduren fast immer Fehler anfallen, ist dieses Verfahren sinnvoll, wenn es auch den Prozedurablauf wegen der eingeschobenen Bildschirmanzeigen verlangsamt. Mit der

Funktion UMFORM werden schließlich die Ergebnisse einer Prozedur in eine Ausgabedatei geschrieben. Über die genannten vier wichtigsten Funktionen hinaus bietet die Landesdatenbank dem Direktnutzer weitere Möglichkeiten der Datenverarbeitung. So können eigene standardisierte Berichte erstellt und eigene Daten in das Landesdatenbankformat umgeformt und mit den Beständen der Landesdatenbank verknüpft werden. Die Datenbank kann mit anderer (PC- oder Großrechner-) Software in Verbindung gebracht werden. Weitere Funktionen beziehen sich speziell auf die Darstellung von inoperablen Werten und die Aufzeichnung von Programmen.

4. Auswertung von Daten der nordrhein-westfälischen Gemeinden

Die vorrangige Aufgabe der Landesdatenbank besteht darin, zu sehr vielen unterschiedlichen Merkmalen laufend fortgeschriebene Daten für kleine Gebietseinheiten bereitzustellen. Die Berücksichtigung kleiner Gebietseinheiten ist im übrigen generell die ureigenste Aufgabe aller statistischen Landesämter, während dem Statistischen Bundesamt im allgemeinen nur die Präsentation von Daten für das Bundesgebiet und die Bundesländer obliegt. Die kleinsten für die amtliche Statistik relevanten Gebietseinheiten sind in der Regel die Gemeinden. Im LDS werden seit langem multivariate Regionalanalysen auf der Basis von Gemeindedaten durchgeführt. Dies ist wegen des großen Datenbeschaffungs- und -verarbeitungsaufwands nur mit Hilfe der Datenbank möglich. Es liegt nahe, die Praktikantenausbildung über die Datenbankschulung hinaus mit den Regionalanalysen zu verbinden und die Praktikanten so an laufenden Forschungsarbeiten des LDS teilnehmen zu lassen.

Die im LDS durchgeführten regionalanalytischen Arbeiten beziehen sich unter anderem auf die für Zwecke der regionalisierten Strukturpolitik im Rahmen der Landesplanung gebildeten Regionen des Landes Nordrhein-Westfalen (Gerß 1994). Jede Region soll für ihr Gebiet ein Entwicklungskonzept formulieren, wobei wirtschaftliche, soziale, demographische und umweltbezogene Interessen zu berücksichtigen sind. Die Regionen sind in einem politischen Entscheidungsprozeß abgegrenzt worden, der sich grundsätzlich nach den Grenzen der öffentlichen Verwaltung (Regierungsbezirke, Kreise) orientierte. Da nicht einzelne Gemeinden nach ihrer Ähnlichkeit zu Regionen zusammengefaßt wurden, sondern die Abgrenzung der Regionen grundsätzlich losgelöst von speziellen Merkmalsausprägungen der einzelnen Gemeinden durch die Grenzen der überkommunalen Verwaltungsebenen bestimmt ist, stellt sich die Frage, welche Kennzeichen den Regionen – über die triviale Tatsache hinaus, daß die zu ihnen gehörenden Gebietsteile aneinander grenzen – jeweils eigentümlich sind. Clusteranalytisch betrachtet sollte die Zusammenfassung von Gebieten zu einer

Region dazu führen, daß innerhalb dieser Region verhältnismäßig geringe Unterschiede bestehen, die Region mit ihren Teilen sich dagegen von anderen Regionen deutlich unterscheidet. Die Untersuchungen des LDS behandeln erstens die Frage, welche Merkmale allgemein bei der gegebenen Aufteilung des Landes in Planungsregionen zur Beschreibung der regionalen Unterschiede geeignet sind. Zweitens wird die Frage untersucht, ob es aus der Sicht jeweils einer einzelnen Region diese charakterisierende Merkmale gibt. Die Ausgangsdaten der Untersuchungen des LDS beziehen sich stets auf die einzelnen Gemeinden und damit auf die kleinsten Gebietseinheiten, für die laufend amtliche statistische Daten anfallen. Mit dieser kleinräumigen Betrachtung wird sichergestellt, daß die für die Beurteilung der Eignung eines Merkmals zur regionalen Abgrenzung wichtige intraregionale Streuung sichtbar bleibt. Verwendet werden nur stets aktualisierbare Daten, d. h. solche, die entweder mit höchstens jährlicher Periodizität anfallen oder von denen bekannt ist, daß sie sich mindestens einige Jahre lang – seit dem letzten Berichtsjahr – nicht oder nur unwesentlich ändern. Demnach werden Daten aus Volkszählungen und anderen weit zurückliegenden oder nur in großen Zeitabständen wiederholten Erhebungen nicht einbezogen. Alle Merkmalswerte werden als Prozentzahlen oder Pro-Kopf-Zahlen o.ä. ausgedrückt, so daß die Werte von Gemeinde zu Gemeinde unabhängig von deren unterschiedlicher Größe vergleichbar werden. Insgesamt werden in diesen Arbeiten des LDS 100 Merkmale mit Werten für jede der 396 Gemeinden des Landes Nordrhein-Westfalen verwendet. Diese Ausgangsdaten

stammen aus dem normalen Aufbereitungsprogramm der amtlichen Statistik und werden in der Regel entweder gemeindeweise laufend veröffentlicht oder sind in dieser kleinräumigen Gliederung in der Landesdatenbank gespeichert. Sie sind damit – abgesehen von geheimzu haltenden Individualangaben – grundsätzlich auch für externe Nutzer verfügbar. Nur wenige Daten – Aggregate der Sozialproduktberechnung – sind nach Gemeinden nicht veröffentlicht bzw. nicht für externe Nutzer abrufbar und stehen damit in dieser kleinsten regionalen Gliederung nur zur behördlichen Weiterverrechnung zur Verfügung. Die für die – im folgenden Abschnitt beschriebene – Untersuchung der Wahlergebnisse verwendeten Daten beziehen sich bei den meisten Merkmalen auf das jeweilige Wahljahr (1994 bzw. 1995); nur auf davon abweichende Berichtsjahre wird in den folgenden Auflistungen besonders hingewiesen. Die jeweils durch einen Buchstaben und eine Nummer gekennzeichneten 100 Merkmale verteilen sich zu je 25 auf vier Merkmalsgruppen: Demographische Merkmale (D1 bis D25), soziale Merkmale (S1 bis S25), wirtschaftliche Merkmale (W1 bis W25), umweltbezogene Merkmale (U1 bis U25). Die 100 Merkmale wurden bei der Analyse der Wahlergebnisse vollständig als potentielle erklärende Variablen berücksichtigt. Die Merkmalswerte wurden von dem Praktikanten selbständig zusammengestellt und zeitlich den Wahljahren angepaßt (Weinem 1998).

Die demographischen Merkmale beschreiben die Bevölkerungsstruktur und -bewegung. Sie beziehen sich neben der Gesamtheit der Einwohner auf verschiedene Teilgesamtheiten, die sich aus der Gliederung nach der Ausländereigenschaft, dem Geschlecht, bestimmten Altersgruppen und der Erwerbspersoneneigenschaft ergeben. Die Datenquellen sind die Fortschreibung des Bevölkerungsstandes, die Statistik der natürlichen Bevölkerungsbewegung und die Wanderungsstatistik. Bei den demographischen Daten handelt es sich überwiegend um – in der Terminologie der Landesdatenbank – „endgültige Jahreswerte“ bzw. „offene Realdaten“. Die Daten zur Ausländereigenschaft (namentlich D1, D4, D5 und D6) sind wegen ihrer möglichen Fehlerhaftigkeit als eingeschränkt aussagefähig gekennzeichnet. Zu allen demographischen Merkmalen können die Grunddaten für die hier relevanten Berichtsjahre 1994 und 1995 vollständig aus der Landesdatenbank übernommen werden. Aus den Grunddaten wurden meist Anteilswerte (D1 bis D19), in einigen Fällen Werte je 1000 Einwohner (D20, D21, D24 und D25) und in zwei Fällen prozentuale Veränderungen (D22 und D23) berechnet.

- D1 Anteil der Ausländer an der Gesamtbevölkerung (%)
- D2 Anteil der Ausländer an den Zuzügen (%)
- D3 Anteil der Ausländer an den Fortzügen (%)
- D4 Anteil der Ausländer an den Einwohnern im Alter von 18 bis unter 65 Jahren (%)
- D5 Anteil der Ausländer an den Lebendgeborenen (%)
- D6 Anteil der Männer an den ausländischen Einwohnern (%)
- D7 Anteil der Männer an der Gesamtbevölkerung (%)
- D8 Anteil der Männer an den Einwohnern im Alter ab 65 Jahren (%)
- D9 Anteil der Männer an den Zuzügen (%)
- D10 Anteil der Männer an den Fortzügen (%)
- D11 Anteil der Einwohner im Alter von unter 18 Jahren (%)
- D12 Anteil der Personen im Alter von unter 18 Jahren an den Zuzügen (%)
- D13 Anteil der Personen im Alter von unter 18 Jahren an den Fortzügen (%)
- D14 Anteil der Einwohner im Alter ab 65 Jahren (%)
- D15 Anteil der Personen im Alter ab 65 Jahren an den Zuzügen (%)
- D16 Anteil der Personen im Alter ab 65 Jahren an den Fortzügen (%)
- D17 Anteil der Erwerbspersonen an den Zuzügen (%)
- D18 Anteil der Erwerbspersonen an den Fortzügen (%)
- D19 Anteil der männlichen Personen an den Lebendgeborenen (%)
- D20 Anzahl der Lebendgeborenen je 1000 Einwohner
- D21 Anzahl der Gestorbenen je 1000 Einwohner
- D22 Prozentuale Veränderung der Einwohnerzahl gegenüber dem Vorjahr
- D23 Prozentuale Veränderung der Einwohnerzahl im Zehnjahreszeitraum
- D24 Anzahl der Zuzüge je 1000 Einwohner
- D25 Anzahl der Fortzüge je 1000 Einwohner

Zu den sozialen Merkmalen werden hier vor allem statistische Ergebnisse zu Sozialleistungen, Kennzahlen des Bildungswesens, Angaben über Einkommen und Zahlungsschwierigkeiten als Wohlstandsindikatoren sowie Daten über die Wohnungssituation und das Gesundheitswesen gerechnet. Die Datenquellen sind in erster Linie die Statistik der Jugendhilfe, die Statistik des allgemeinen Schulwesens, die Statistik der Konkurs- und Vergleichsverfahren, die Statistik der Bau fertigstellungen, die Fortschreibung des Wohngebäude- und Wohnungsbestandes, die Wohngeldstatistik und die Krankenhausstatistik. Neben diesen Bundesstatistiken werden die Sozialproduktberechnungen der statistischen Landesämter verwendet. Soweit die Daten aus der Landesdatenbank stammen, handelt es sich um „endgültige Stichtagswerte“ als „offene Realdaten“. Einige Merkmale (S1 bis S4) sind bisher nur für das Jahr 1994 in der Datenbank mit Werten belegt; diese Werte werden hier als auch für 1995 gültig angenommen. Die Daten zur Sportplatzfläche (S5) wurden seit dem Berichtsjahr 1988 nicht mehr aktualisiert. Mehrere Merkmale gehören nicht zum Bestandskatalog der Datenbank; diese Daten wurden aus Unterlagen der Fachdezernate des LDS entnommen (S14 bis S18, S22 bis S25). Dabei waren zum Teil verhältnismäßig aufwendige Umformungen notwendig, damit die automatisierte Weiterverarbeitung möglich wurde.

- S1 Anzahl der verfügbaren Plätze in Kindertageseinrichtungen je 100 Kinder aller Altersgruppen bis unter 14 Jahren (nur 1994)
- S2 Anzahl der tätigen Personen in Kindertageseinrichtungen je 100 verfügbare Plätze (nur 1994)
- S3 Anzahl der verfügbaren Plätze in Einrichtungen der Jugendhilfe je 100 Personen im Alter von 14 bis unter 18 Jahren (nur 1994)
- S4 Anzahl der tätigen Personen in Einrichtungen der Jugendhilfe je 100 verfügbare Plätze (nur 1994)
- S5 Nutzbare Sportfläche (km²) der Sportplätze je 1 000 Einwohner am 1.7.1988
- S6 Anteil der Ausländer an den Grundschülern (%)
- S7 Anteil der Hauptschüler an den Schülern des siebten Schuljahrgangs (%)
- S8 Anteil der Realschüler an den Schülern des siebten Schuljahrgangs (%)
- S9 Anteil der Gymnasialschüler an den Schülern des siebten Schuljahrgangs (%)
- S10 Anteil der Personen ohne Hauptschulabschluß an den Schulabgängen aller weiterführenden allgemeinbildenden Schulen (%)
- S11 Anteil der Abiturienten (allgemeine Hochschulreife) an den Schulabgängen (%)
- S12 Anteil der Ein- und Zweifamilienhäuser an den fertiggestellten Wohngebäuden (%)
- S13 Wohnfläche je fertiggestellte Wohnung (m²)
- S14 Verfügbares Einkommen je Einwohner (DM)

- S15 Bruttoerwerbs- und -vermögenseinkommen je Einwohner (DM)
- S16 Anzahl der aufgestellten Betten in allgemeinen Krankenhäusern je 1 000 Einwohner
- S17 Anzahl der hauptamtlichen Ärzte in allgemeinen Krankenhäusern je 1 000 aufgestellte Betten
- S18 Anzahl der Krankenschwestern und Krankenpfleger in allgemeinen Krankenhäusern je 1 000 aufgestellte Betten
- S19 Anzahl der Grundschüler je Grundschullehrer
- S20 Anzahl der Wohnräume je vorhandene Wohnung
- S21 Anzahl der Wohnungen je vorhandenes Wohngebäude
- S22 Prozentuale Abweichung des Wohngelds vom Bundesdurchschnitt
- S23 Anzahl der aufgestellten Betten je allgemeines Krankenhaus
- S24 Anzahl der Totgeborenen je 1 000 entbundene Frauen in allgemeinen Krankenhäusern
- S25 Angemeldete Forderungen in Konkursverfahren von natürlichen Personen (prozentuale Veränderung gegenüber dem Vorjahr)

Die wirtschaftlichen Merkmale haben neben mengenmäßigen zum Teil auch in Geld bewertete Einheiten. Sie erfassen die Inanspruchnahme von Flächen für ausschließlich wirtschaftliche Tätigkeiten, die Branchenstruktur, die Bedeutung und Struktur der Arbeitnehmerschaft, die Betriebsgrößen, die öffentlichen Finanzen und – in gesamtwirtschaftlicher Sicht – die Wirtschaftskraft. Spezielle Indikatoren werden für das Beherbergungsgewerbe verwendet. Die Datenquellen sind die Erhebung der Katasterflächen bei den Vermessungsverwaltungen, die Bodennutzungserhebung, die Arbeitnehmerstatistik der Bundesanstalt für Arbeit, die Erhebungen bei Betrieben im Bergbau und Verarbeitenden Gewerbe, die Gemeindefinanzstatistik und die Statistik der Beherbergung im Reiseverkehr. Neben diesen Bundesstatistiken werden die Sozialproduktberechnungen der statistischen Landesämter und die landeseigenen Berechnungen des LDS über die Beschäftigtenzahlen verwendet. Fast alle wirtschaftlichen Daten stammen aus der Landesdatenbank, aus der sie als „endgültige Stichtagswerte“ bzw. als „geschätzte Jahresmittelwerte“ – jeweils „offene Realdaten“ – zu entnehmen sind. Nur die Daten aus der Sozialproduktberechnung (W15, W16 und W23) wurden außerhalb der Datenbank vom Fachdezernat des LDS bereitgestellt. Mit Ausnahme der Bruttowertschöpfung zu Faktorkosten (W15), die nur für die geraden Jahre ermittelt wird – und dann auch für das jeweils folgende ungerade Jahr als gültig angenommen werden muß –, liegen alle wirtschaftlichen Daten für die beiden hier relevanten Berichtsjahre 1994 und 1995 vor. Allerdings hat sich bei einem Merkmal (W14) zwischen 1994 und 1995 die Definition – in vernachlässigbar geringem Umfang – geändert.

- W1 Anteil der Betriebsfläche an der Katasterfläche (%)
- W2 Anteil der Landwirtschaftsfläche an der Katasterfläche (%)

- W3 Anteil der Landwirtschaft, Forstwirtschaft und Fischerei an der Gesamtzahl der Beschäftigten (%)
- W4 Anteil von Energieversorgung, Wasserversorgung und Bergbau an der Gesamtzahl der Beschäftigten (%)
- W5 Anteil von Verarbeitendem Gewerbe und Baugewerbe an der Gesamtzahl der Beschäftigten (%)
- W6 Anteil von Handel, Verkehr und Nachrichtenübermittlung an der Gesamtzahl der Beschäftigten (%)
- W7 Anteil von Kreditinstituten, Versicherungsgewerbe, Dienstleistungsunternehmen und freien Berufen an der Gesamtzahl der Beschäftigten (%)
- W8 Anteil von Gebietskörperschaften, Sozialversicherung, Organisationen ohne Erwerbszweck und privaten Haushalten an der Gesamtzahl der Beschäftigten (%)
- W9 Anteil der sozialversicherungspflichtig beschäftigten Arbeitnehmer an der Gesamtzahl der Beschäftigten (%)
- W10 Anteil der Männer an den sozialversicherungspflichtig beschäftigten Arbeitnehmern (%)
- W11 Anteil der Personen mit Dienstleistungsberufen an den sozialversicherungspflichtig beschäftigten Arbeitnehmern (%)
- W12 Anteil der Facharbeiter an den sozialversicherungspflichtig beschäftigten Arbeitnehmern (%)
- W13 Anteil der Teilzeitbeschäftigte an den sozialversicherungspflichtig beschäftigten Arbeitnehmern (%)
- W14 Anzahl der Beschäftigten je Betrieb im Bergbau und Verarbeitenden Gewerbe (1994; für 1995 Bergbau, Gewinnung von Steinen und Erden und Verarbeitendes Gewerbe)
- W15 Bruttowertschöpfung zu Faktorkosten je Erwerbstätigen (DM; nur 1994)
- W16 Bruttoanlageinvestitionen je Erwerbstätigen (DM)
- W17 Realsteuerkraft je Einwohner (DM)
- W18 Steuereinnahmekraft je Einwohner (DM)
- W19 Landwirtschaftlich genutzte Fläche je landwirtschaftlichen Betrieb (Ar)
- W20 Anzahl der Betten in Beherbergungsbetrieben je 1 000 Einwohner
- W21 Durchschnittliche Aufenthaltsdauer in Hotels garnis, Gasthöfen und Pensionen (Tage je Person)
- W22 Anteil der ausländischen Gäste an den Übernachtungen in Beherbergungsbetrieben (%)
- W23 Bruttowertschöpfung zu Marktpreisen je Erwerbstätigen (DM)
- W24 Schulden der Gemeinde je Einwohner (DM)
- W25 Hebesatz der Gewerbesteuer nach Ertrag und Kapital (%)

Die umweltbezogenen Merkmale beziehen sich auf die aus der Sicht des Umweltschutzes primär als ungünstig (U1 und U2) oder als günstig (U3, U4 und U5) zu beurteilende Inanspruchnahme von Flächen, die Belastung der Fläche durch Besiedlung (U19 und U20) und durch gewerblichen Verbrauch an fossilen Brennstoffen (U21), die Belastung der Bevölkerung durch den Straßenverkehr, Information über den Anfall und die Behandlung von festen Abfällen und Abwasser sowie auf Umweltschutzinvestitionen. Die Datenquellen sind die Erhebung der Katasterflächen bei den Vermessungsverwaltungen, die Bodennutzungserhebung, die Statistik des Kraftfahrzeugbestandes, die Statistik der Straßenverkehrsunfälle, die Statistik der Abfallbeseitigung, die Statistik der Wasserversorgung und der Abwasserbeseitigung, die Statistik der Investitionen für Umweltschutz im Produzierenden Gewerbe, die Fortschreibung des Bevölkerungsstandes, die Statistik der Baufertigstellungen und die Erhebungen bei Betrieben im Bergbau und Verarbeitenden Gewerbe. Die aus der Landesdatenbank übernommenen umweltbezogenen Daten sind „endgültige Stichtagswerte“ oder „endgültige Jahreswerte“, jeweils als „offene Realdaten“. Alle Abwassermerkmale (U11 bis U17 und U25) stammen – außerhalb der Datenbank – aus Unterlagen des Fachdezernats des LDS, die nur für das Berichtsjahr 1995 vorliegen. Dagegen sind die Daten zur Abfallbeseitigung (U8, U9, U10, U23 und U24) in der Datenbank gespeichert; hier wurden auf Anraten des Fachdezernats zur Gewährleistung der Zuverlässigkeit nur die Ergebnisse für das Jahr 1993 verwendet. Auch alle anderen Daten wurden problemlos und vollständig der Datenbank entnommen. Die sich auf 1995 bzw. 1993 beschränkenden Daten müssen als für die hier relevanten Untersuchungsjahre gültig angenommen werden.

- U1 Anteil der Gebäudefläche und der Gebäuden untergeordneten Freifläche an der Katasterfläche (%)
- U2 Anteil der Verkehrsfläche an der Katasterfläche (%)
- U3 Anteil der Erholungsfläche (überwiegend Grünanlagen) an der Katasterfläche (%)
- U4 Anteil der Waldfläche an der Katasterfläche (%)
- U5 Anteil des Dauergrünlands an der landwirtschaftlichen Nutzfläche (%)
- U6 Anzahl (Bestand) der Kraftfahrzeuge je 1 000 Einwohner
- U7 Anzahl der Straßenverkehrsunfälle je 1 000 Einwohner
- U8 Abfallaufkommen in allen Abfallgruppen ohne besonders überwachungsbedürftige Abfälle je an die öffentliche Abfallbeseitigung angeschlossenen Einwohner (kg; nur 1993)
- U9 An öffentliche Abfallbeseitigungs- und -behandlungsanlagen gelieferte Abfallmengen je Einwohner (kg; nur 1993)
- U10 Aufkommen an besonders überwachungsbedürftigen Abfällen gemäß Abfallbestimmungsverordnung je Einwohner (kg; nur 1993)
- U11 Anteil der an öffentlichen Kläranlagen angeschlossenen Einwohner (%; nur 1995)

- U12 Anteil der im Wege der öffentlichen Abwasserbeseitigung direkt in ein Gewässer oder den Untergrund abgeleiteten Schmutzwassermenge an der Schmutzwassermenge insgesamt (%; nur 1995)
- U13 Anteil des im Wege der öffentlichen Abwasserbeseitigung in biologischen Kläranlagen behandelten Abwassers am behandelten Abwasser insgesamt (%; nur 1995)
- U14 Anteil des im Wege der öffentlichen Abwasserbeseitigung in biologischen Kläranlagen mit weitergehender chemisch-physikalischer Behandlung behandelten Abwassers am behandelten Abwasser insgesamt (%; nur 1995)
- U15 Anteil der im Bergbau und Verarbeitenden Gewerbe direkt in ein Gewässer oder den Untergrund abgeleiteten Abwassermenge an der Abwassermenge insgesamt (%; nur 1995)
- U16 Anteil der bei betriebseigenen Abwasseranlagen (Bergbau und Verarbeitendes Gewerbe) in biologischen Kläranlagen behandelten Abwassermenge an der diesen Kläranlagen zugeführten Abwassermenge insgesamt (%; nur 1995)
- U17 Anteil der Abwassermenge mit biologischer und weitergehender chemisch-physikalischer Behandlung bei betriebseigenen Abwasseranlagen (Bergbau und Verarbeitendes Gewerbe) an der diesen Kläranlagen zugeführten Abwassermenge insgesamt (%; nur 1995)
- U18 Anteil der Investitionen für Umweltschutz an den Gesamtinvestitionen des Produzierenden Gewerbes (ohne Baugewerbe) (%)
- U19 Anzahl der Einwohner je km^2
- U20 Rauminhalt bei Baufertigstellungen von Wohn- und Nichtwohngebäuden (Neubau und Wiederaufbau) je km^2 (m^3)
- U21 Energieverbrauch insgesamt (außer der zur Stromerzeugung verbrauchten Energie aus Kohle, Heizöl und Gas) im Bergbau und Verarbeitenden Gewerbe je km^2 (Tonnen Steinkohleeinheiten)
- U22 Anteil der Unfälle mit Personenschaden an der Gesamtzahl der Straßenverkehrsunfälle (%)
- U23 Kapazität von öffentlichen Abfalldeponien (Restvolumen) je Einwohner (m^3 ; nur 1993)
- U24 Abfallaufkommen (ohne besonders überwachungsbedürftige Abfälle) aus Produzierendem Gewerbe und Krankenhäusern je Einwohner (kg; nur 1993)
- U25 Konzentration der Abwasserschädlichkeit bei öffentlichen Kläranlagen vor der Behandlung, berechnet anhand des chemischen Sauerstoffbedarfs in Milligramm je Liter des gereinigten Abwassers (nur 1995)

5. Regressionsanalyse der Ergebnisse politischer Wahlen

Die Untersuchung der kleinräumigen Wahlergebnisse ist aus drei Gründen in besonderem Maße geeignet, mit einem sozialwissenschaftlichen Praktikum im LDS

verbunden zu werden. Erstens sind die Wahlergebnisse nach Gemeinden in der Landesdatenbank gespeichert und können von dort durch die Praktikanten – nach der im vorvorigen Abschnitt erwähnten Schulung – zur Weiterverrechnung mit statistischen Auswertungsprogrammen bequem übernommen werden. Zweitens ist das Thema, die Wahlergebnisse im Zusammenhang mit den sie möglicherweise bestimmenden Merkmalen der Gemeinden – entsprechend der Auflistung im vorigen Abschnitt – zu betrachten, von hervorragender soziologischer Bedeutung. Drittens bietet die hier gegebene Zusammenführung einer großen Zahl von Variablen die Möglichkeit, die im Studium erworbenen Kenntnisse über multivariate statistische Verfahren auf für die empirische Forschung realistische Datenmengen anzuwenden.

Bei der in den folgenden Ausführungen dargestellten Regressionsanalyse wird die Eignung aller im vorigen Abschnitt aufgelisteten 100 Merkmale als unabhängige Variablen multipler Regressionsgleichungen geprüft. Für einzelne Gemeinden fehlende Werte einiger dieser Merkmale werden dabei durch das jeweilige arithmetische Mittel aller Gemeinden ersetzt. Die abhängigen Variablen (Y) sind die Wahlbeteiligung und die (Zweit-) Stimmenanteile der vier größten Parteien bei der Europawahl 1994, der Bundestagswahl 1994, der Kommunalwahl 1994 und der Landtagswahl 1995. Die Einheiten, für die die Variablenwerte gemessen werden, sind die 396 Gemeinden des Landes Nordrhein-Westfalen. Insgesamt ergeben sich die wie folgt bezeichneten 20 abhängigen Variablen:

	Wahlbeteiligung (%)	Stimmenanteil (%)			
		CDU	SPD	F.D.P.	GRÜNE
Kommunalwahl 1994	Y1	Y2	Y3	Y4	Y5
Bundestagswahl 1994	Y6	Y7	Y8	Y9	Y10
Europawahl 1994	Y11	Y12	Y13	Y14	Y15
Landtagswahl 1995	Y16	Y17	Y18	Y19	Y20

Dementsprechend gibt es 20 multiple Regressionsgleichungen. Die unabhängigen Variablen jeweils einer dieser Gleichungen werden in einer schrittweisen Prozedur ermittelt (Schuemer, Ströhlein und Gogolok 1990, S. 130). Dabei wird aus den 100 Variablen in einem sukzessiven Prozeß eine nach bestimmten Kriterien als optimal angesehene Teilmenge herausgegriffen. Dieser Auswahlprozeß beginnt mit null unabhängigen Variablen. Für jede potentielle unabhängige Variable wird ein F-Wert berechnet, der den Anteil dieser Variablen an der Varianz der abhängigen Variablen – und damit ihren Beitrag zur Erklärung der abhängigen Variablen – darstellt, falls die untersuchte unabhängige Variable als solche in die Regressionsgleichung aufgenommen würde. Der zum F-Wert korrespondierende Wahrscheinlichkeitswert P wird mit dem zu spezifizierenden Signifikanzniveau verglichen. Wenn keine Variable

gefunden wird, die einen signifikanten F-Wert aufweist, wird der Auswahlprozeß abgebrochen. Andernfalls wird die Variable mit dem größten F-Wert als unabhängige Variable in die Regressionsgleichung aufgenommen. Anschließend beginnt der Auswahlprozeß erneut für die verbleibenden Variablen. Auf diese Weise gehen sukzessive eine, zwei, drei usw. unabhängige Variablen in der Reihenfolge ihres Erklärungsbeitrags in die Regressionsgleichung ein. Bei diesem Verfahren bleibt jede einmal aufgenommene Variable bei den folgenden Auswahlschritten in der Gleichung enthalten. Alle Berechnungen werden in direkter Kombination mit der Datenbereitstellung aus der Landesdatenbank mit dem Programmpaket SAS durchgeführt (SAS Institute Inc. 1989, S. 1351-1456), und zwar mit der Prozedur REG. Als Variablenauswahltechnik wird FORWARD verwendet. Die umfassenden Regressionsrechnungen mit den – aus dem Auswahlprozeß hervorgegangenen – als optimal angesehenen unabhängigen Variablen werden mit der Methode NONE durchgeführt. Als Option, die die Parameterschätzungen betrifft, wird STB verwendet. Dadurch werden standardisierte Regressionskoeffizienten berechnet. Diese beziehen sich auf die auf arithmetisches Mittel null und Standardabweichung 1 standardisierten Variablen, die dimensionslos sind und somit nicht mehr die unterschiedlichen Maßeinheiten der ursprünglichen Variablen widerspiegeln. Die standardisierten Regressionskoeffizienten sind im Hinblick auf den Einfluß der unabhängigen Variablen, denen sie zugeordnet sind, auf die abhängige Variable direkt miteinander vergleichbar (Tiede 1987, S. 176). Je größer der Betrag des standardisierten Regressionskoeffizienten innerhalb einer multiplen Regressionsgleichung ist, desto stärker beeinflußt die jeweilige unabhängige die abhängige Variable.

Der Abbildung 1 liegen die aus der schrittweisen Regression resultierenden F- und P-Werte zugrunde, die sich bei zunehmender Anzahl der Regressoren ergeben. Zur besseren Übersichtlichkeit beschränkt sich die Darstellung jeweils auf denjenigen Bereich der Regressorenanzahl, in dem sich die verschiedenen F- bzw. P-Werte am deutlichsten voneinander unterscheiden. Zu demselben Zweck sind zusammengehörige Werte trotz der diskreten Skalierung der horizontalen Achse stetig (geradlinig) miteinander verbunden. Mit wachsender Anzahl der Regressoren nehmen die F-Werte ab und die P-Werte zu. Das im ersten Teil der Abbildung 1 dargestellte Beispiel der Bundestagswahl 1994 zeigt, daß von den untersuchten fünf Merkmalen dieser Wahl der Stimmenanteil der F.D.P. bei jeweils bestimmter Regressorenanzahl im allgemeinen zu dem höchsten F-Wert (bzw. dem geringsten P-Wert) führt und damit in diesem Sinne durch das Regressionsmodell am besten erklärt wird. Ebenfalls relativ gut wird der Stimmenanteil der SPD erklärt, nicht ganz so gut die Wahlbeteiligung. Die Erklärung der Stimmenanteile der GRÜNEN und der CDU gelingt dagegen am schlechtesten. In diesem Zusammenhang mag eine Rolle spielen, daß die traditionelle Bindung des Wählerpotentials der GRÜNEN an diese Partei eher gering ist, so daß aus taktischen Erwägungen oft auch andere Parteien gewählt werden (Neutag 1998). Die für die Bundestagswahl 1994 festgestellte

unterschiedliche Erklärungsqualität des Regressionsmodells in Bezug auf die fünf Merkmale herrscht auch bei den anderen – hier nicht graphisch dargestellten – Wahlen vor. Bei der Kommunalwahl 1994 werden ebenfalls die Stimmenanteile von SPD und F.D.P. besser erklärt als die von CDU und GRÜNEN; die Wahlbeteiligung erreicht hier auffallend hohe F-Werte. Auch bei der Landtagswahl 1995 werden die Anteile von SPD und F.D.P. besser erklärt als die von CDU und GRÜNEN. Die Europawahl 1994 weicht von dem üblichen Bild teilweise ab, indem zwar der F.D.P.-Anteil besonders gut und der CDU-Anteil besonders schlecht, dagegen der GRÜNEN-Anteil eher besser als der SPD-Anteil erklärt wird. Die Wahlbeteiligung erreicht bei der Landtagswahl und vor allem bei der Europawahl verhältnismäßig geringe F-Werte. Nach dem im zweiten Teil der Abbildung 1 dargestellten Beispiel des Stimmenanteils der SPD ergibt sich für dieses Merkmal unter den vier untersuchten Wahlen bei jeweils bestimmter Regressorenanzahl im allgemeinen in der Bundestagswahl 1994 der höchste, in der Europawahl 1994 der zweithöchste, in der Landtagswahl 1995 der dritthöchste und in der Kommunalwahl 1994 der geringste F-Wert. Die Erklärungsqualität des Regressionsmodells ist für die anderen – hier nicht graphisch dargestellten – Merkmale in ähnlicher Weise von Wahl zu Wahl unterschiedlich. Auch die Anteile von F.D.P. und GRÜNEN werden in der Europawahl und der Bundestagswahl besser erklärt als in der Landtagswahl und der Kommunalwahl. Etwas abweichend davon erklärt das Regressionsmodell den CDU-Anteil in der Landtags- und der Europawahl im allgemeinen besser als in der Bundestags- und der Kommunalwahl. Auffällig ist, daß die Stimmenanteile aller vier Parteien in der Kommunalwahl am schlechtesten erklärt werden. Dies mag einleuchtend sein, da in der Kommunalwahl häufiger als in den anderen Wahlen die individuelle Entscheidung für oder gegen einen bestimmten Kandidaten vor der Parteienpräferenz – die zu den statistisch erfaßten Strukturmerkmalen der Gemeinden in Beziehung gebracht wird – den Ausschlag gibt. Dagegen wird die Wahlbeteiligung in der Bundestags- und der Kommunalwahl – die im Jahr 1994 in Nordrhein-Westfalen am selben Tag stattfanden – eher besser erklärt als in der Landtags- und der Europawahl.

Die Rangnummern der Tabelle 1 ergeben sich, indem gesondert für jede abhängige Variable alle 100 potentiellen unabhängigen Variablen nach ihren F-Beträgen ohne Rücksicht auf deren Signifikanz vom größten (Rang 1) bis zum kleinsten Wert (Rang 100) mit Kardinalzahlen belegt werden. Wenn jede der vier Variablengruppen D, S, U und W insgesamt eine gleich große Bedeutung zur Erklärung jeweils einer Wahlvariablen Y hätte, müßten ihre Rangnummern gleichmäßig auf die Zahlenbereiche 1 bis 25, 26 bis 50, 51 bis 75 und 76 bis 100 verteilt sein. In Wirklichkeit bestehen in dieser Verteilung große Unterschiede, die durch einen Korrelationskoeffizienten ausgedrückt werden können. Dazu werden für jede einzelne Wahlvariable die zu diesem Zweck mit den Nummern 4 (für den angegebenen Zahlenbereich 1 bis 25), 3 (Bereich 26 bis 50), 2 (Bereich 51 bis 75) und 1 (Bereich 76 bis 100) belegten Rangnummernklassen mit der Anzahl der in die

jeweilige Klasse fallenden D- bzw. S- bzw. U- bzw. W-Variablen korreliert. Für die abhängige Variable Y1 gehören z. B. acht D-Variablen in die Rangnummernklasse 4, sieben D-Variablen in die Klasse 3, sechs D-Variablen in die Klasse 2 und vier D-Variablen in die Klasse 1. Das ergibt einen Korrelationskoeffizienten von + 0,983. Die Variablengruppe D hat damit eine extrem große Bedeutung zur Erklärung von Y1. Dagegen gehören für Y1 von den U-Variablen vier in die Rangnummernklasse 4, sechs in die Klasse 3, fünf in die Klasse 2 und zehn in die Klasse 1, was einen Korrelationskoeffizienten von -0,835 ergibt. Die Variablengruppe U ist dementsprechend für die Erklärung von Y1 sehr unbedeutend. Generell zeigen positive Koeffizienten relativ große, negative Koeffizienten relativ geringe Erklärungsanteile an. In diesem Sinne ist die Gruppe der demographischen Merkmale insgesamt zur Erklärung aller Variablen der Wahlbeteiligung wichtiger als die drei anderen Merkmalsgruppen. Diese Dominanz ist bei den verglichenen Wahlen verschieden stark ausgeprägt. Der Korrelationskoeffizient r der D-Gruppe hat für Y1 den Wert + 0,983, für Y6 + 0,929, für Y16 + 0,529 und für Y11 + 0,513. Die umweltbezogenen Merkmale sind für die drei Variablen der Wahlbeteiligung Y11($r = -0,939$), Y1(-0,835) und Y6(-0,775), die wirtschaftlichen Merkmale für die Variable Y16(-0,477) relativ am unwichtigsten. Noch einheitlicher sind die Ergebnisse für die Variablen der SPD-Anteile, die alle von der Gruppe der wirtschaftlichen Merkmale am besten (Y8 $r = + 0,939$; Y3 + 0,887; Y13 + 0,775; Y18 + 0,718) und von der Gruppe der umweltbezogenen Merkmale am wenigsten (Y3 -0,923; Y8 -0,832; Y18 -0,832; Y13 -0,801) erklärt werden. Auch für die Anteile der GRÜNEN sind die Ergebnisse ziemlich eindeutig. Hier sind für drei Wahlvariablen die sozialen Merkmale (Y10 $r = + 0,835$; Y20 + 0,775; Y5 + 0,529) und für die vierte Variable die demographischen Merkmale (Y15 + 0,674) zur Erklärung am wichtigsten; für alle GRÜNEN-Anteile sind die wirtschaftlichen Merkmale (Y15 -0,983; Y10 -0,775; Y20 -0,757; Y5 -0,258) am unwichtigsten. Zur Erklärung der CDU-Anteile liefern in je zwei Fällen die demographischen (Y7 $r = +0,927$; Y2 + 0,328) und die wirtschaftlichen Merkmale (Y17 + 0,944; Y12 + 0,564) den größten Beitrag, während in drei Fällen die umweltbezogenen (Y12 -0,969; Y7 -0,816; Y2 -0,640) und in einem Fall die sozialen Merkmale (Y17 -0,614) die geringste Bedeutung haben. Am uneinheitlichsten sind die Ergebnisse für die Variablen der F.D.P.-Anteile. Hier wird Y4 durch die D-Gruppe ($r = +0,947$), Y9 durch die S-Gruppe (+0,775), Y14 durch die W-Gruppe (+0,757) und Y19 ebenfalls durch die W-Gruppe (+0,832) am besten erklärt; die geringste Bedeutung hat für Y4 (-0,602) die W-Gruppe, für Y9 (-0,680) die U-Gruppe, für Y14(-0,378) die D-Gruppe und für Y19 (-0,569) die U-Gruppe.

Die Abbildung 2 beruht auf den zu jeder einzelnen Wahlvariablen für 1 bis 100 Regressoren schrittweise berechneten F-Werten. Für jedes Wahlvariablenpaar wird der Pearson-Korrelationskoeffizient über alle 100 zugehörigen F-Werte-Paare berechnet. Aus den 20 Wahlvariablen ergeben sich 190 verschiedene Korrelationskoeffizienten. Diese werden nach ihren Absolutbeträgen absteigend

geordnet. Es kommen Beträge von nahezu 1 bis nahezu null vor. Der stärkste Zusammenhang der F-Werte wird für das Variablenpaar Y6/Y1 (Wahlbeteiligung bei der Bundestagswahl 1994 und der Kommunalwahl 1994) ausgewiesen. Da die F-Werte die Wichtigkeit der einzelnen Regressoren zur Erklärung der abhängigen Variablen des Regressionsmodells ausdrücken und damit die Brauchbarkeit dieses Modells bewerten, bedeutet die extrem starke Korrelation, daß die Variablen Y6 und Y1 in gleicher Weise und gleich gut erklärt werden. Dies ist im Hinblick auf die Gleichzeitigkeit der Bundestags- und der Kommunalwahl plausibel. Ein fast gleich starker Zusammenhang der F-Werte besteht für das Variablenpaar Y12/Y7 (Stimmenanteil der CDU bei der Europawahl 1994 und der Bundestagswahl 1994). Nahezu null beträgt der Korrelationskoeffizient dagegen z. B. für die Variablenpaare Y18/Y10 (Stimmenanteile der SPD bei der Landtagswahl 1995 und der GRÜNEN bei der Bundestagswahl 1994) sowie Y17/Y1 (Stimmenanteil der CDU bei der Landtagswahl 1995 und Wahlbeteiligung bei der Kommunalwahl 1994). Dementsprechend werden diese Variablen jeweils durch das Modell in unterschiedlicher Weise bzw. verschieden gut erklärt. In der graphischen Darstellung werden die 20 Wahlvariablen nach den Beträgen der Korrelationskoeffizienten ihrer F-Werte hierarchisch-agglomerativ klassifiziert. Die senkrechte Achse gibt die Korrelationsbeträge an, nach denen die Wahlvariablen in dem von unten nach oben zu betrachtenden Dendrogramm schrittweise zusammengefaßt werden. Um das Bild leichter überschaubar zu machen, sind die Beträge in dieser Darstellung auf die Werte null, 0,05, 0,10 usw. bis 0,95, 1,00 gerundet. Die Zusammenfassung aller 20 Wahlvariablen wird in sechs Schritten erreicht. Zu den Variablen des ersten Schrittes (Y1, Y6, Y7, Y8, Y12, Y13, Y17), die beim Korrelationsbetrag 1 zusammengefaßt werden, gehören auch zwei (Y1, Y17), die miteinander nicht korrelieren. Dieses scheinbar paradoxe Ergebnis kommt dadurch zustande, daß die Korrelation sowohl von Y1 als auch von Y17 zu jeweils anderen Variablen des ersten Zusammenfassungsschrittes extrem stark ist. Auf die plausible Korrelation der Wahlbeteiligungen bei der Bundestags- und der Kommunalwahl (Y6 und Y1) wurde bereits hingewiesen. Ebenso einleuchtend ist die nahezu 1 betragende Korrelation des CDU-Anteils bei der Landtagswahl (Y17) zu den CDU-Anteilen bei der Bundestagswahl (Y7) und der Europawahl (Y12). Zu den im zweiten Zusammenfassungsschritt – beim Korrelationsbetrag 0,95 – über den ersten Schritt hinaus hinzukommenden Variablen (Y2, Y3, Y10, Y11, Y15, Y18) gehören zwei (Y10, Y18) miteinander nicht korrelierende. Auch hier liegt die Erklärung in der extrem starken Korrelation dieser beiden zu anderen – bereits zusammengefaßten – Variablen. So korreliert der GRÜNEN-Anteil bei der Bundestagswahl (Y10) plausibel (Betrag 0,95) mit dem GRÜNEN-Anteil bei der Europawahl (Y15). Der SPD-Anteil bei der Landtagswahl (Y18) korreliert ebenfalls plausibel (Betrag 0,95) mit den CDU-Anteilen bei der Landtagswahl (Y17) und der Bundestagswahl (Y7). Hier sei nochmals betont, daß die Korrelation sich nicht auf die ursprünglichen Werte der Variablen, sondern auf deren F-Werte bezieht, so daß ein hoher Korrelationsbetrag gleiche Erklärungsqualität des Regressionsmodells für die beiden

betroffenen Variablen ausdrückt, wobei die Korrelation der Variablenwerte positiv oder negativ sein kann. Unter den im dritten bis sechsten Zusammenfassungsschritt hinzukommenden Variablen (Y4, Y5, Y9, Y14, Y16, Y19, Y20) befinden sich die F.D.P.-Anteile bei der Kommunalwahl (Y4), der Landtagswahl (Y19), der Bundestagswahl (Y9) und der Europawahl (Y14) sowie die GRÜNEN-Anteile bei der Kommunalwahl (Y5) und der Landtagswahl (Y20). Es sei daran erinnert, daß die F.D.P.-Anteile durch das Modell im allgemeinen besonders gut, die GRÜNEN-Anteile dagegen relativ schlecht erklärt werden (Abbildung 1). Die Klassifizierung der Abbildung 2 stellt dagegen nicht für die einzelnen Variablen den Erklärungswert an sich dar, sondern zeigt, ob Gruppen von Variablen gleich oder unterschiedlich gut bzw. schlecht erklärt werden.

Die Tabelle 2 enthält für die einzelnen Wahlvariablen diejenigen Regressoren, die bei schrittweiser Regression bis zum Abbruchkriterium $P=0,001$ in das Modell aufgenommen werden, also mit ihren standardisierten Regressionskoeffizienten höchstsignifikant sind. Dabei zeigen sich innerhalb der hier unterschiedenen Gruppen von Wahlvariablen – Wahlbeteiligung, CDU-Anteile, SPD-Anteile, F.D.P.-Anteile, GRÜNEN-Anteile – meist typische Einflüsse, die in den verglichenen vier Wahlen wiederkehren. Die Wahlbeteiligung der Kommunalwahl 1994 (Y1), der Bundestagswahl 1994 (Y6) und der Europawahl 1994 (Y11) wird am stärksten – und zwar negativ – von der Variablen S21 (Anzahl der Wohnungen je vorhandenes Wohngebäude) beeinflußt. Auf die Wahlbeteiligung der Landtagswahl 1995 (Y16) ist der Einfluß – ebenfalls negativ – der Variablen U1 (Anteil der Gebäudefläche und der Gebäuden untergeordneten Freifläche an der Katasterfläche) am stärksten. Demnach ist die Wahlbeteiligung – durchaus erwartungsgemäß – in den großstädtischen Verdichtungsgebieten verhältnismäßig niedrig und in Gebieten mit geringer Siedlungsdichte hoch. Die Variable S20 (Anzahl der Wohnräume je vorhandene Wohnung) übt bei allen Wahlen den stärksten Einfluß auf die Stimmenanteile sowohl der CDU (in positiver Richtung) als auch der SPD (in negativer Richtung) aus. Je größer die Wohnungen sind, desto mehr Wähler hat die CDU und desto weniger Wähler die SPD. Die dominierende – und zwar jeweils positiv wirkende – Einflußgröße auf die Stimmenanteile der GRÜNEN ist bei der Kommunalwahl (Y5), der Bundestagswahl (Y10) und der Landtagswahl (Y20) die Variable S22 (Prozentuale Abweichung des Wohngelds vom Bundesdurchschnitt) und bei der Europawahl (Y15) die Variable S15 (Bruttoerwerbs- und -vermögenseinkommen je Einwohner). Je höher das Einkommens- und das Mietenniveau ist, desto besser sind die Wahlchancen der GRÜNEN. Die Stimmenanteile der F.D.P. sind viel weniger eindeutig vorhersagbar als die der anderen Parteien. Bei der Kommunalwahl (Y4) und der Europawahl (Y14) werden die F.D.P.-Anteile am stärksten (positiv) von der Variablen U6 (Anzahl der Kraftfahrzeuge je 1 000 Einwohner), bei der Bundestagswahl (Y9) am stärksten (ebenfalls positiv) von der Variablen S15 (Bruttoerwerbs- und -vermögenseinkommen je Einwohner) und bei der Landtagswahl (Y19) am stärksten

(negativ) von der Variablen W17 (Realsteuerkraft je Einwohner) beeinflußt. Dies deutet darauf hin, daß die F.D.P. außerhalb der Ballungszone (bevorzugte Wohnlage für „Besserverdienende“, stärkere Abhängigkeit vom privaten Kraftfahrzeug) größere Wahlchancen hat als in den stärker gewerblich geprägten und daher realsteuerstärkeren Ballungskernen.

Literaturverzeichnis

- Bentgens, J.(1992): Der Fremdenverkehr 1990. Statistische Rundschau Nordrhein-Westfalen 44 (1): 15-18.
- Drogies, G. (1998): Regionale Unterschiede der Mortalität und deren Entwicklung in Nordrhein-Westfalen 1988 bis 1996. Diplomarbeit Universität Duisburg, nicht veröffentlicht.
- Düntgen, A.(1997): Faktorenanalytische Zusammenfassung verschiedener Merkmale der Gemeinden NRWs. Statistische Rundschau Nordrhein-Westfalen 49 (3): 155-166.
- Freimann, K.(1993): Informationsverarbeitung des Kommunikationssystems der Landesdatenbank – Informationsvermittlung und Datenbankrecherche. Praktikumsbericht, nicht veröffentlicht.
- Gerhard-Mercator-Universität Duisburg (1997): AkZent – Akademisches Zentrum für Studium und Beruf. Rundbrief.
- Gernand, D.(ohne Jahr): Informationen zur Durchführung von Praktika. Universität Düsseldorf, nicht veröffentlicht.
- Gerß, W (1994): Kennzeichen der Planungsregionen des Landes Nordrhein-Westfalen. Statistische Rundschau Nordrhein-Westfalen 46 (11): 565-574.
- Gerß, W. (1995a): Soziologiestudenten als Praktikanten in der amtlichen Statistik. Statistische Rundschau Nordrhein-Westfalen 47(1): 16-21.
- Gerß, W. (1995b): Ausbildung in empirischer Sozialforschung durch Hochschule und statistisches Amt. Allgemeines Statistisches Archiv 79(4): 425-438.
- Gerß, W., und Jamke, V. (1989): Praxisnahe ADV-Ausbildung von Soziologiestudenten. Statistische Rundschau Nordrhein-Westfalen 41(5): 219-223.
- Heinrich-Heine-Universität Düsseldorf (Sozialwissenschaftliches Institut) (ohne Jahr): Informationen zum Praktikumsbüro. Rundbrief.
- Neutag, J.(1998): Das Wählerpotential von BÜNDIS 90/DIE GRÜNEN seit der bundesweiten Parteigründung im Jahr 1980. Magisterarbeit Universität Düsseldorf, nicht veröffentlicht.
- Plewa, H.-E.(1985): Die Landesdatenbank Nordrhein-Westfalen – Ein modernes Auskunftsinstrument der amtlichen Statistik. Statistische Rundschau für das Land Nordrhein-Westfalen 37 (6): 383-390.

- SAS Institute Inc. (1989): SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2. Cary (North Carolina).
- Schuemer, R., Ströhlein, G., und Gogolok, J. (1990): Datenverarbeitung und statistische Auswertung mit SAS, Band II (Komplexe statistische Analyseverfahren). Gustav Fischer Verlag, Stuttgart und New York.
- Strasser, H. (1987): Grundsätze für Praktikums- und Stipendiatenstellen des Faches Soziologie. Universität Duisburg, nicht veröffentlicht.
- Tiede, M. (1987): Statistik – Regressions- und Korrelationsanalyse. R. Oldenbourg Verlag, München.
- Weinem, M. (1998): Datenrecherche und Datenbereitstellung mit der Landesdatenbank NRW. Praktikumsbericht, nicht veröffentlicht.
- Werner, A. (1997): Zur Lage der Arbeitslosigkeit in Nordrhein-Westfalen – Die Verfahren der Datenbeschaffung im Informationssystem Landesdatenbank. Praktikumsbericht, nicht veröffentlicht.

Anhang

Tabelle 1: Rangnummern der unabhängigen Variablen (D, S, U, W) nach ihrer Bedeutung zur Erklärung jeweils einer abhängigen Variablen (Y).

1 = Variable mit dem größten, 100 = Variable mit dem geringsten Erklärungsanteil.

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Y14	Y15	Y16	Y17	Y18	Y19	Y20
D1	53	15	47	84	20	57	19	13	50	3	50	29	25	89	3	59	36	11	61	20
D2	61	99	41	4	41	70	46	31	27	16	79	56	32	32	8	40	70	66	72	32
D3	77	83	5	40	23	75	6	8	55	97	81	3	7	56	88	50	38	46	88	4
D4	6	69	37	16	13	8	9	20	20	54	49	33	28	60	28	1	9	7	31	44
D5	12	43	73	28	26	9	21	35	83	95	28	9	9	13	95	38	6	6	39	5
D6	80	44	25	6	69	94	26	6	78	14	82	23	31	47	16	60	16	10	78	11
D7	48	52	18	3	73	38	41	69	5	69	19	70	53	39	41	58	77	96	41	63
D8	49	54	82	72	75	10	52	76	59	46	59	65	50	44	64	49	76	59	80	97
D9	22	66	46	89	5	17	97	82	90	65	90	86	91	54	45	51	25	27	67	46
D10	89	9	94	78	39	88	27	39	84	36	21	49	41	59	62	44	20	79	53	66
D11	33	73	72	21	44	40	30	10	46	44	45	20	16	88	13	89	50	34	18	68
D12	63	26	21	38	99	58	38	11	33	20	77	34	23	58	32	56	11	14	6	6
D13	28	36	55	17	74	27	23	32	88	45	35	30	36	92	50	26	91	61	7	79
D14	20	5	28	12	90	20	2	5	13	62	75	6	5	30	26	28	3	2	59	16
D15	60	22	92	42	4	69	28	55	37	29	12	27	57	25	24	80	26	41	84	10
D16	92	42	30	54	88	84	69	57	40	34	36	69	80	23	23	41	48	65	19	51
D17	2	97	13	71	66	2	59	38	67	30	2	41	19	62	39	2	85	87	71	81
D18	41	48	49	81	93	37	89	99	99	56	73	93	87	87	49	14	86	60	51	23
D19	51	72	84	60	35	50	86	91	58	79	57	96	85	46	58	94	59	52	99	94
D20	21	18	20	24	81	29	14	84	18	53	62	4	72	36	30	83	10	20	52	37
D21	5	38	66	49	79	3	18	34	38	80	4	28	26	11	91	5	28	42	16	27
D22	47	75	23	37	68	48	49	27	24	22	69	39	34	34	29	70	96	49	75	9
D23	57	30	78	30	61	61	51	18	76	86	11	97	76	90	71	4	73	93	33	74
D24	29	33	83	98	37	33	42	49	63	31	29	47	48	83	86	54	32	88	95	55
D25	15	56	74	69	10	22	24	43	61	84	39	37	60	69	25	15	31	55	54	64
S1	3	60	48	57	54	5	25	29	48	15	5	17	30	97	59	33	23	24	85	53
S2	62	16	8	99	2	63	8	25	31	2	88	2	10	33	2	35	7	17	29	3
S3	35	19	3	46	63	35	15	14	51	76	24	11	21	75	54	17	33	21	77	43
S4	72	68	57	7	87	53	63	86	21	33	72	79	92	16	81	37	84	71	87	59
S5	8	29	24	88	77	14	16	15	56	35	53	16	12	17	33	90	56	76	45	41
S6	44	40	26	43	65	45	79	45	36	74	67	80	46	10	56	82	88	77	79	45
S7	32	90	98	97	12	36	91	89	81	72	7	91	96	71	85	29	30	22	55	85
S8	23	10	93	53	84	21	37	65	65	87	20	25	65	57	60	46	22	56	35	19
S9	50	61	52	15	71	51	84	90	68	96	16	75	82	51	96	92	79	75	74	50
S10	65	91	36	90	55	60	88	77	25	55	41	90	86	68	42	9	63	98	63	18
S11	94	35	99	73	31	87	87	95	91	71	76	99	95	98	57	27	18	30	58	65
S12	19	24	62	31	29	13	32	50	71	43	17	46	49	79	47	20	89	67	9	33
S13	79	27	56	47	17	68	44	33	93	38	25	38	33	95	46	81	90	86	49	42
S14	13	76	7	10	38	19	34	23	7	19	33	43	20	12	4	47	80	28	20	87
S15	14	50	10	48	24	12	58	16	1	6	14	45	15	1	7	11	57	9	1	21
S16	64	87	80	66	46	67	95	96	85	89	78	85	88	26	78	65	8	26	34	72
S17	67	94	68	29	15	74	82	54	42	7	54	92	64	67	22	99	98	51	21	25
S18	55	79	67	5	42	52	75	73	28	21	60	67	66	52	52	97	71	95	26	84
S19	87	82	53	41	43	81	62	88	22	5	66	12	94	22	9	31	27	35	14	8
S20	16	1	1	74	85	16	1	1	15	41	27	1	1	78	20	63	1	1	32	80
S21	1	65	42	26	50	1	31	46	41	11	1	36	54	19	5	85	42	48	12	34
S22	69	20	4	2	1	79	17	4	3	1	94	13	4	2	1	13	93	31	15	1
S23	25	6	17	51	70	26	7	12	86	24	89	8	13	99	97	67	74	54	56	98
S24	4	92	54	13	30	4	65	64	49	42	6	61	63	6	65	74	34	74	36	52
S25	86	67	31	93	51	76	78	60	35	82	95	68	58	77	94	24	83	84	38	31

Fortsetzung Tabelle 1

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Y14	Y15	Y16	Y17	Y18	Y19	Y20
U1	91	11	33	86	60	93	35	19	19	70	86	24	35	49	68	3	29	40	57	76
U2	76	86	44	35	21	80	60	94	32	18	97	58	98	40	21	95	69	97	70	13
U3	40	71	76	32	6	32	66	71	94	17	9	77	55	81	34	19	92	33	90	7
U4	7	45	91	92	83	7	81	85	74	66	40	83	40	42	75	21	13	19	98	88
U5	11	41	61	83	86	15	48	59	87	10	44	48	52	86	11	30	53	92	82	22
U6	17	64	70	1	97	11	50	92	2	51	23	40	39	3	31	62	82	73	96	61
U7	95	81	75	22	19	95	99	67	11	23	63	52	47	20	92	98	15	16	89	14
U8	27	53	100	100	14	28	100	100	100	100	61	54	100	66	80	22	100	100	69	95
U9	34	95	88	34	25	41	93	98	57	40	70	84	93	70	35	61	51	78	42	35
U10	100	80	60	68	100	100	74	40	64	98	100	95	74	38	90	100	52	36	100	70
U11	26	8	85	19	34	25	5	37	10	8	8	10	42	7	10	6	4	25	2	26
U12	73	39	45	63	27	72	47	78	34	60	84	66	97	28	74	88	45	72	13	39
U13	38	17	96	61	48	42	90	81	60	73	26	82	81	27	83	7	64	70	5	86
U14	18	34	86	18	94	24	94	74	53	27	18	88	70	29	14	8	66	80	11	47
U15	78	32	38	75	36	73	92	53	12	13	48	87	59	43	19	78	68	99	25	15
U16	74	74	59	55	57	62	73	51	80	81	85	63	71	76	48	96	67	50	81	36
U17	85	59	77	58	40	89	61	62	54	32	31	60	61	85	66	39	61	69	91	93
U18	93	58	97	56	8	92	53	61	79	48	96	44	84	74	12	32	44	64	40	12
U19	37	47	9	85	89	39	45	47	23	92	71	50	37	41	76	42	37	29	46	82
U20	66	77	40	9	58	64	80	52	77	91	91	98	89	84	84	25	55	37	48	92
U21	99	46	58	25	49	97	72	66	39	68	92	73	75	50	67	53	60	62	43	90
U22	82	63	39	52	52	86	55	83	96	61	47	51	69	61	38	79	14	8	44	75
U23	71	49	95	95	32	77	64	75	43	88	74	71	67	93	69	76	46	39	92	73
U24	81	100	32	82	91	90	56	93	70	93	58	100	44	100	100	71	95	85	93	100
U25	52	96	65	65	95	56	71	70	97	52	87	81	77	91	51	68	99	83	94	67
W1	39	89	79	23	72	43	76	17	6	12	65	55	8	15	18	36	94	12	10	17
W2	31	88	16	27	80	34	39	41	75	78	38	42	45	64	82	43	35	32	22	56
W3	75	37	22	8	18	78	22	36	82	25	10	19	38	73	27	10	24	58	27	24
W4	83	7	2	70	64	85	4	2	16	47	22	5	2	45	40	75	5	3	30	69
W5	9	85	35	94	28	23	96	22	17	9	68	76	6	9	93	55	41	5	37	78
W6	96	93	69	50	82	83	83	97	89	28	98	94	99	65	55	45	78	68	86	30
W7	98	70	87	59	53	82	54	87	52	99	30	64	43	96	17	87	58	38	65	38
W8	97	25	71	20	92	99	77	68	92	50	80	57	83	21	63	93	97	81	97	99
W9	70	12	15	44	33	71	11	63	45	75	15	14	68	14	61	16	19	53	28	71
W10	68	23	43	79	7	55	70	58	14	67	46	59	56	24	73	57	39	44	66	60
W11	90	31	90	87	56	98	68	56	66	64	3	72	62	55	79	12	75	91	68	89
W12	30	98	81	11	9	30	85	80	8	90	13	89	90	5	98	18	40	45	64	49
W13	45	84	64	77	3	49	36	79	95	4	42	74	79	35	6	73	49	90	47	2
W14	10	78	34	91	22	6	98	72	98	37	93	78	73	82	43	84	62	47	24	62
W15	46	21	50	80	45	47	29	28	72	83	56	35	22	94	77	34	54	43	83	58
W16	56	4	89	64	76	54	33	42	47	59	43	31	27	48	37	23	72	89	76	54
W17	42	55	11	14	11	44	67	24	62	49	83	62	29	4	44	48	65	82	3	29
W18	59	62	29	33	59	65	40	44	4	26	99	32	78	8	15	69	43	94	4	91
W19	54	28	14	67	67	66	20	30	26	58	52	22	17	53	36	72	12	18	73	57
W20	58	3	19	76	47	59	3	3	9	77	55	7	3	37	53	86	2	4	8	77
W21	84	14	63	45	98	91	43	26	69	39	64	26	24	63	72	64	47	13	17	40
W22	24	2	12	39	62	18	12	7	30	94	34	18	18	31	99	91	87	57	50	48
W23	43	57	51	96	16	46	57	48	73	63	37	53	51	80	89	66	81	63	60	28
W24	36	51	27	62	96	31	10	21	29	57	32	21	14	18	70	52	21	15	23	83
W25	88	13	6	36	78	96	13	9	44	85	51	15	11	72	87	77	17	23	62	96

Tabelle 2: Standardisierte Regressionskoeffizienten bei schrittweiser "Vorwärts"-Regression bis zum Abbruchkriterium $P = 0,001$

Abhängige Variablen	Koeffizienten der unabhängigen Variablen nach ihren Beträgen geordnet					
Y1	-0,253 S21	+0,211 S24	-0,197 D4	-0,195 D21	+0,191 S1	+0,168 D17
noch: Y1	+0,127 U4					
Y2	+0,500 S20	+0,273 W20	+0,212 W22	+0,202 S23	-0,183 D14	+0,148 W16
Y3	-0,648 S20	-0,360 S22	+0,227 W4	-0,208 S3	+0,186 W25	-0,167 D3
Y4	+0,399 U6	+0,234 S22	-0,186 D2	-0,182 D7		
Y5	+0,264 S22	+0,202 S2	+0,183 W13			
Y6	-0,321 S21	-0,223 D21	+0,205 S24	+0,198 S1	+0,170 D17	-0,143 W14
Y7	+0,669 S20	+0,259 W20	-0,234 D14	-0,192 W4	-0,163 S2	+0,163 S23
noch: Y7	+0,160 D3	+0,141 U11				
Y8	-0,607 S20	-0,442 S22	-0,260 W20	+0,253 W4	+0,230 W25	+0,178 D14
noch: Y8	-0,157 S5	-0,153 S23	-0,138 W22	-0,134 D3	-0,127 S3	
Y9	+0,142 S15	-0,982 S14	+0,239 U6	-0,186 W18	+0,176 S22	-0,173 D7
noch: Y9	-0,166 W12	-0,140 U11	-0,127 W1	+0,122 W20		
Y10	+0,471 S22	-0,274 D1	+0,244 S2	+0,173 W13	-0,162 S19	
Y11	-0,278 S21	-0,250 D21	+0,211 W11	+0,177 S24	+0,164 S1	+0,129 D17
Y12	+0,622 S20	+0,239 W20	-0,196 D14	+0,185 S23	-0,177 S2	-0,172 W4
noch: Y12	+0,156 D3	+0,143 D20	+0,132 U11			

Fortsetzung Tabelle 2

Abhängige Variablen	Koeffizienten der unabhängigen Variablen nach ihren Beträgen geordnet					
Y13	-0,499 S20	-0,393 S22	-0,280 W20	+0,237 D14	+0,225 W4	+0,199 W25
noch: Y13	-0,192 D3	-0,166 S5	+0,163 W1	-0,139 S23	+0,128 W5	
Y14	+0,462 U6	+0,352 S15	+0,326 S22	-0,245 W17	-0,241 W12	+0,132 S24
Y15	+0,819 S15	-0,640 S14	+0,343 S21	+0,322 S22	+0,230 S2	-0,217 D1
noch: Y15	+0,206 W13	-0,163 D2				
Y16	-0,276 U1	-0,271 D4	+0,254 D17	-0,230 D23	-0,159 D21	
Y17	+0,641 S20	+0,350 D4	+0,289 W20	-0,264 D5	-0,201 D14	-0,155 W4
noch: Y17	+0,142 U11	+0,123 S16	-0,122 S2			
Y18	-0,574 S20	-0,464 D4	+0,377 D5	-0,289 W20	+0,231 W4	+0,194 D14
noch: Y18	+0,183 W5	-0,139 U22	-0,132 S15	+0,126 D6		
Y19	-0,963 W17	+0,894 W18	-0,328 U11	+0,325 S15		
Y20	+0,369 S22	+0,242 U3	-0,212 D5	+0,177 S2	-0,170 D3	+0,159 W13

Abb. 1

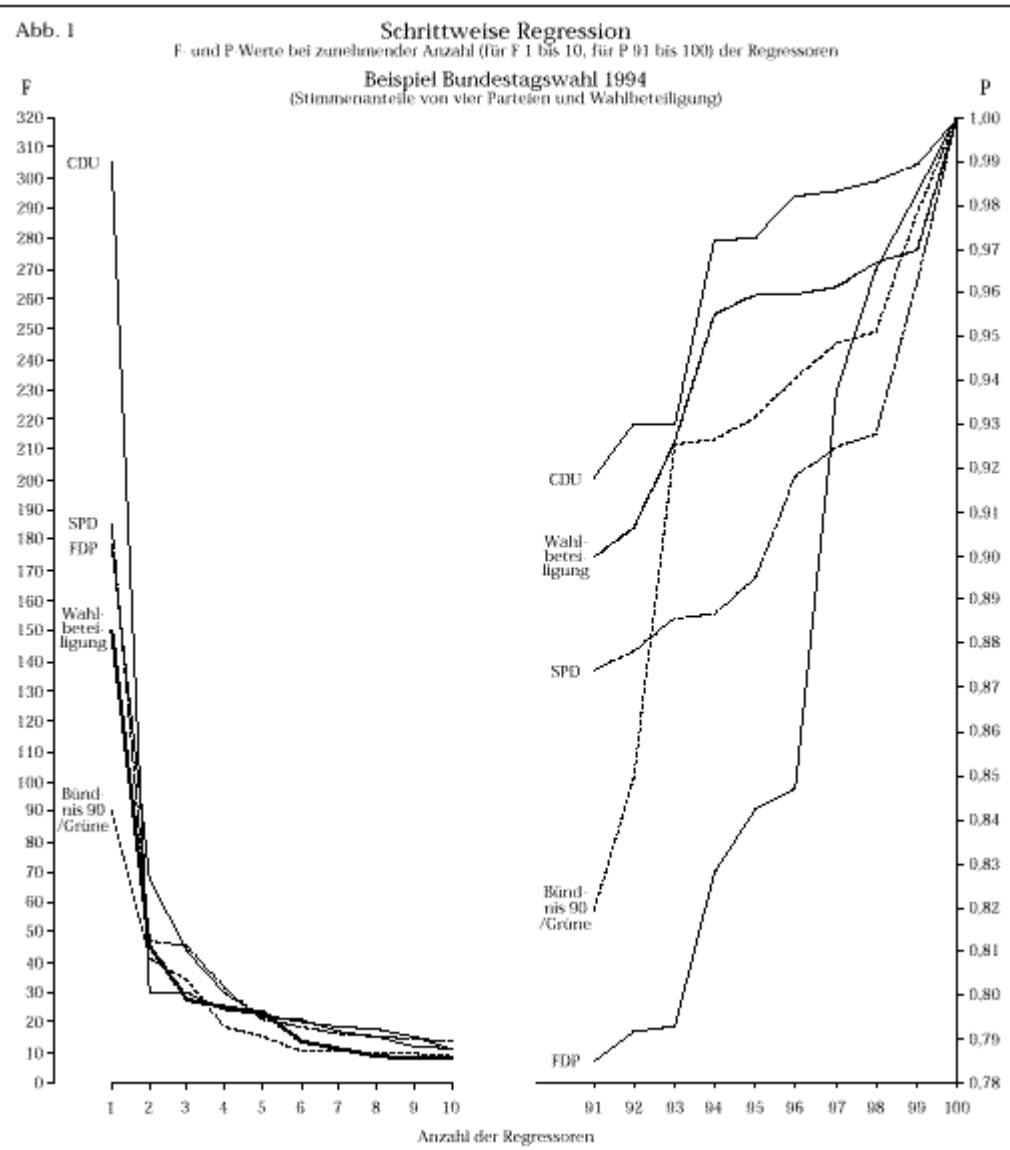


Abb. 1

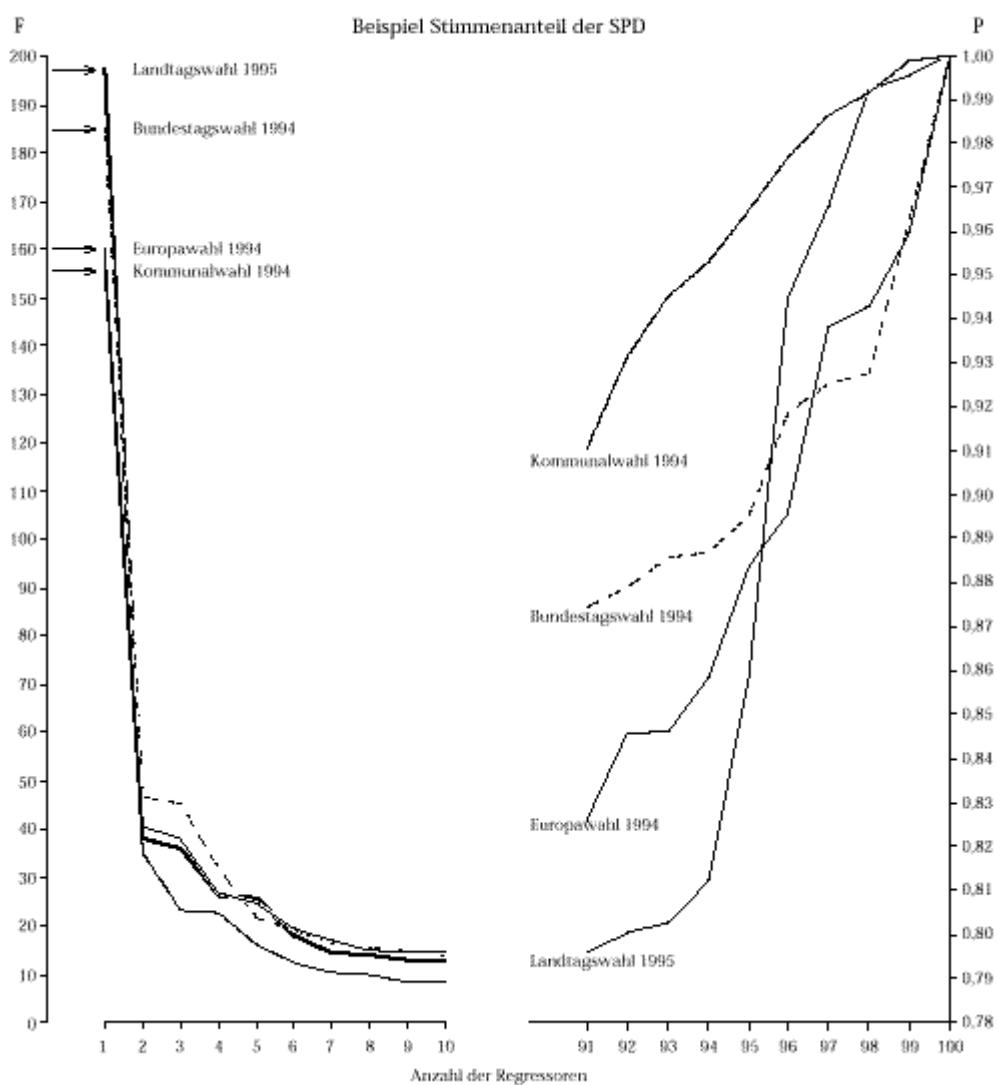
Fortsetzung...
...der Regressionsanalyse

Abb. 2

Hierarchische Klassifikation der Wahlvariablen Y1 bis Y20
nach den Beträgen der Pearson-Korrelationskoeffizienten ($|r|$)
ihrer - für 1 bis 100 Regressoren berechneten - F-Werte

