

Wahlverhalten und politische Einstellungen in Westeuropa

JOHANNES
GUTENBERG
UNIVERSITÄT
MAINZ

Analysen mit dem Eurobarometer
und den Europäische Wahlstudien

Formales: Scheinerwerb

- Referat und Teilnahme am Seminargespräch.
- Die Lektüre der Pflichttexte ist für alle Teilnehmer *verbindlich*, gelegentlich überprüfe ich Ihren Kenntnisstand.
- Sie dürfen *maximal zwei* Sitzungen versäumen.
- Hausarbeit
 - thematischer Zusammenhang mit dem Seminar
 - explizite Fragestellung
 - bloße Ausformulierung des Referates genügt nicht („genügt nicht“ = „ungenügend“)

Formales: Hausarbeit

- Gliederung, Zitierweise, Literaturverzeichnis etc. bitte entsprechend den üblichen Standards (vgl. Homepage)
- Schriftgröße etc.: Formatvorlage
- Umfang ca. 7.000-9.000 Worte, entspricht etwa 20-25 reinen Textseiten
- Beginn bereits während des Semesters
- **Letzter Abgabetermin: Montag, 8. März 2004. Eine Verschiebung des Termins ist *nicht* möglich**

Formales: Referate

- Es soll
 - genug Zeit für Rückfragen während des Vortrages bleiben
 - pro Sitzung mindestens 20-30 Minuten für eine abschließende Diskussion zur Verfügung stehen
- Grundlage ist in jedem Fall die Pflichtlektüre
- Darüber hinaus können und sollen Sie selbständig Literatur suchen

Weitere Informationen zum Seminar

- <http://www.politik.uni-mainz.de/kai.arzheimer/Lehre-Eurobarometer/>

Wozu überhaupt Eurobarometer / vergleichende Umfrageforschung?

- Normatives (politisches) Interesse
 - Demokratie basiert auf freiwilliger Zustimmung der Regierten
 - Europäische Einigung bis heute primär elitengesteuert
 - Umfrageforschung als Mittel zur Herstellung von Responsivität

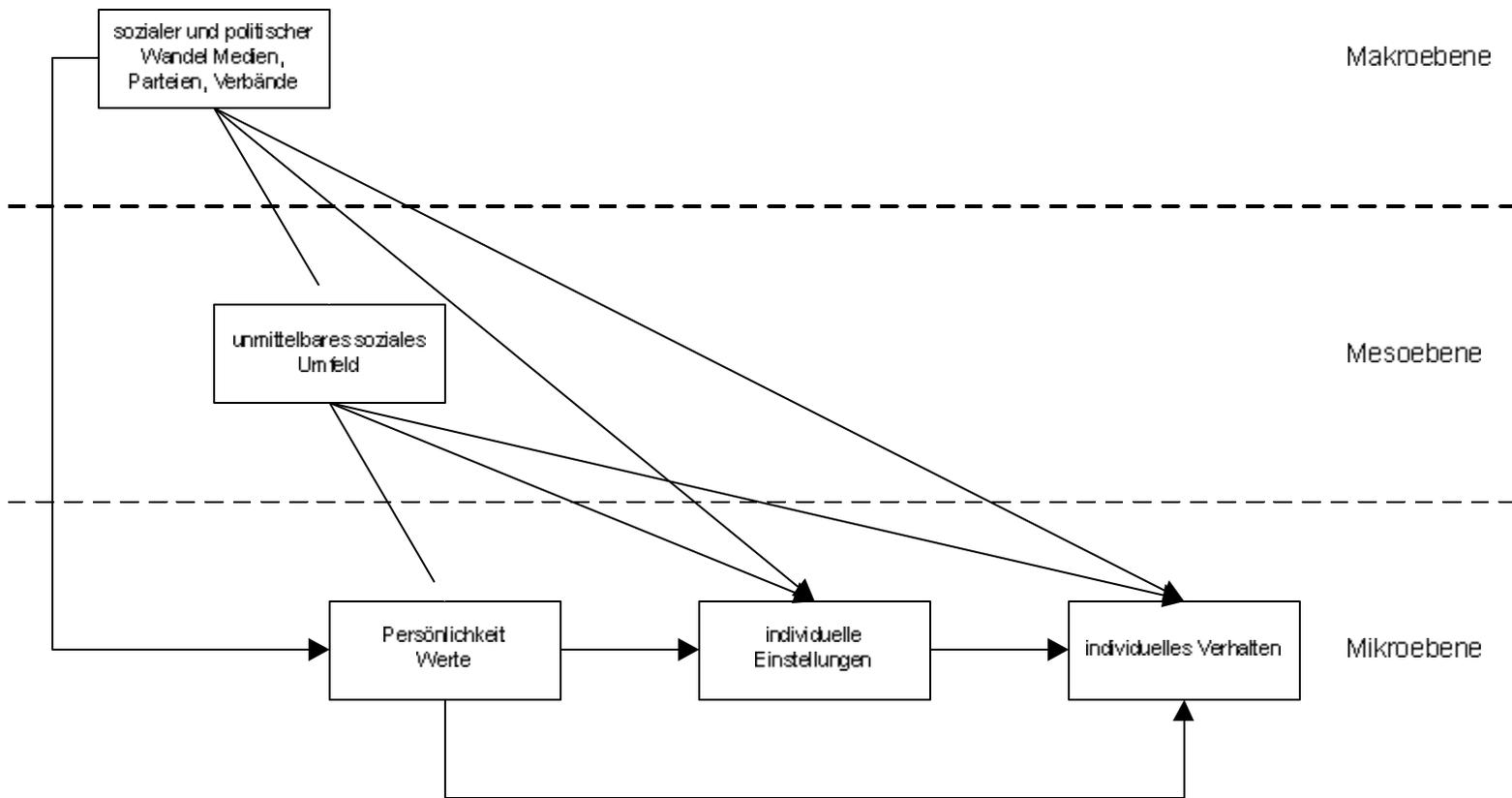
Wozu überhaupt Eurobarometer / vergleichende Umfrageforschung?

- Deskriptives (wissenschaftliches) Interesse
 - Lasswell-Formel: Who Gets What, When and How?
 - Rolle der Bürger bei Beantwortung dieser Frage (politische Kultur, Wahlforschung)
 - Typische Fragen:
 - Wie unterscheiden sich Gesellschaften?
 - Warum unterscheiden sich Gesellschaften?
 - Welche Konsequenzen haben Unterschiede?
 - Welchen Einfluß haben unterschiedliche Kontexte?

Welche Informationen liefern (vergleichende) Umfragen?

- Informationen über...
 - Eigenschaften der Bürger (z. B. Alter, Geschlecht, Bildung)
 - Überzeugungen/Wissen („Wer ist Präsident der Europäischen Kommission?“)
 - Orientierungen/Meinungen (=„Einstellungen“) („Soll Europa eine Verfassung haben?“)
 - Verhalten (retrospektiv/prospektiv, z.B. Wahlbeteiligung in der Vergangenheit, Wahlabsicht in der Zukunft)
- Jeweils in vergleichbarer Form für Stichproben aus den untersuchten Ländern

Welcher implizite Erklärungsansatz steht hinter dieser Form der Umfrageforschung?



Wie sind die Eurobarometer aufgebaut?

- *Eurobarometer ist eine Trend-Studie*
- 1970 und 1971 Meinungsumfragen der EG-Kommission in FR/BE/NL/DE/IT
 - Einstellungen zum Gemeinsamen Markt
 - den Europäischen Institutionen
 - Inglehart-Index
 - weitere Fragen
- 1973 auf LU/DK/IE/GB ausgeweitet, Frageprogramm noch umfangreicher („EB 0“). Ca. 1000 Fälle pro Land
- Seit 1974 halbjährlich (Frühjahr/Herbst) als (Standard-) „Eurobarometer“ durchgeführt, ab 1975 separate Stichprobe für Nordirland (später auch für Ostdeutschland)
- Seit 1980 inkl. GR, ab 1985 inkl. ES/PT, ab 1993 FI, ab 1994 AT/SE; NO sporadisch
- Innerhalb der Standard-EB häufig Themenschwerpunkte, die teilweise repliziert werden (Mediennutzung, Fremdenfeindlichkeit, EP etc.)

Wie sind die Eurobarometer aufgebaut?

- Zusätzlich
 - weitere Standard-EB (z.B. 2002); Oversamples
 - „Flash-EB“ zu aktuellen Themen, telefonisch und mit ca. 500 Befragten pro Land, teilweise spezielle Populationen
 - „Monthly Monitor“ 1994-1995
 - „Continuous Tracking Survey“ 1996-1998 (wöchentlich, nicht für Wissenschaft zugänglich)
 - Central and Eastern Eurobarometer 1990-1997, jährliche Befragungen in 20 mittel- und osteuropäischen Ländern, an Standard-EB angelehnt
 - Candidate Countries Eurobarometer 2001-, an Standard-EB angelehnte Befragung in den 13 Kandidaten-Ländern, noch nicht für Wissenschaft zugänglich
 - European Election Studies (`79-`99), größtenteils aus Standard EBs zusammengesetzt
- Einen Überblick über Variablen und Design verschafft man sich am besten auf den Eurobarometer-Seiten des ZA:
http://www.gesis.org/en/data_service/eurobarometer/

In welcher Form liegen die Daten vor?

- Aufbereitung der Daten durch ICPSR (ZA, SSD)
- Evtl. Embargofristen für bestimmte Fragen (bis zu 2 Jahre)
- SPSS-Datensätze und Codebücher (als PDF) können in Deutschland vom ZA bezogen werden
 - Üblicherweise entspricht ein Datensatz einer Befragungswelle (für alle Länder zusammengefaßt), ca. 15.000 Fälle
 - EES, Oversamples, Sonderfälle

In welcher Form liegen die Daten vor?

- partielle Kumulation „Mannheim Eurobarometer Trend File 1970-1999“ (ZA 3521)
 - harmonisierte Codierung, einheitliche Variablennamen und Labels
 - Vereinheitlichung von fehlenden Werten und GewichtungsvARIABLEN
 - basiert auf 70 Standard-EB
 - enthält 132 Variablen, die mindestens 5 mal erhoben wurden
 - basiert auf 70 Standard-EB
 - 877.223 Fälle (Befragte)
- Dem Institut für Politikwissenschaft stehen für Forschung und Lehre zur Verfügung:
 - die Kommissions-Studien
 - die Standard EB 2-57.2 OVR (1970-2002)
 - die Flash-Eurobarometer 2002 (Euro-Attitudes/-Introduction)
 - der Trend File `70-`99
 - die European Election Studies `89-`99 sowie
 - die Central and Eastern EBs 2-8 (`91-`97)

Beispiel Datenmatrix EB 32 (Querschnitt Herbst 1989)

- Fälle gleich Zeilen, Variablen gleich Spalten
- Ca. 1387 Variablen, 23.397 Fälle
- v1 ICPR-Studiennummer, v4 ID, v5 Land(1), v6 GewichtungsvARIABLE, v7 Land(2)
- v10 Eintrag Wählerregister, v11 Entwicklung ind. Wirtschaftslage, v12 Lebenszufriedenheit, v13 Demokratiezufriedenheit, v14/15 Häufigkeit politischer Diskussionen, v16 Inglehart-Index erstes Ziel
- Aufbau des Trendfiles ähnlich, zusätzliche Variable: Befragungszeitpunkt

v1	v4	v5	v6	v7	v10	v11	v12	v13	v14	v15	v16	v17
1752	1	belgium	1.0908	belgium	present a	same	very sati	not very	from time	occasiona	fight ris	maintaini
1752	2	belgium	.8069	belgium	present a	same	fairly sa	not at al	often	never	fight ris	maintaini
1752	3	belgium	.834	belgium	present a	same	fairly sa	fairly sa	from time	frequentl	fight ris	freedom o
1752	4	belgium	1.2368	belgium	present a	same	fairly sa	fairly sa	from time	occasiona	fight ris	freedom o
1752	5	belgium	.7549	belgium	present a	better	very sati	fairly sa	rarely	occasiona	freedom o	fight ris
1752	6	belgium	1.0432	belgium	present a	better	dk	fairly sa	often	never	fight ris	giving pe
1752	7	belgium	.7885	belgium	present a	same	fairly sa	fairly sa	never	occasiona	fight ris	freedom o
1752	23390	united ki	2.3624	great bri	present a	inap., fo	very sati	very sati	never	never	freedom o	maintaini
1752	23391	united ki	.9488	great bri	present a	inap., fo	very sati	fairly sa	from time	never	maintaini	freedom o
1752	23392	united ki	1.769	great bri	present a	inap., fo	fairly sa	very sati	from time	occasiona	fight ris	maintaini
1752	23393	united ki	1.1469	great bri	present a	inap., fo	not at al	not at al	rarely	occasiona	giving pe	fight ris
1752	23394	united ki	1.9318	great bri	present a	inap., fo	fairly sa	not very	often	occasiona	giving pe	freedom o
1752	23395	united ki	1.8181	great bri	present a	inap., fo	very sati	not very	from time	occasiona	giving pe	freedom o
1752	23396	united ki	1.4571	great bri	present a	inap., fo	very sati	fairly sa	from time	occasiona	giving pe	fight ris
1752	23397	united ki	2.0249	great bri	present a	inap., fo	very sati	fairly sa	often	frequentl	giving pe	freedom o

Beispiel Codebuch EB 32

V11 [Q2] PERSONAL SITUATION - EVOLUTION NEXT YEAR

Location: 38 MD1: 4
Width: 1 MD2: 8

Q.2

As far as you are concerned, do you think that next year -
1990 - will be better or worse than 1989?

1. Better
2. Worse
3. Same
4. DK
8. Inap., not asked in Form B
9. Not ascertained

- v11 ind. Wirtschaftslage

	F	B	NL	D	I	L	DK	IRL	GB	NIRL
1	339	370	377	246	528	94	343	517	364	106
%	36.8	39.6	39.8	23.3	54.8	32.8	35.6	57.0	41.1	38.5
2	113	100	95	130	201	22	127	137	287	69
%	12.3	10.7	10.0	12.3	20.9	7.7	13.2	15.1	32.4	25.1
3	469	465	475	682	235	171	493	253	234	100
%	50.9	49.7	50.2	64.5	24.4	59.6	51.2	27.9	26.4	36.4
4	76M	63M	95M	73M	68M	15M	37M	71M	62M	16M
8	1000M	997M	994M	1081M	1056M	301M	1000M	967M	968M	291M
9	3M	4M		5M						1M
Sum	2000	1999	2036	2217	2088	603	2000	1945	1915	583

Beispiel Codebuch EB 32

SAMPLING AND FIELDWORK

From October 11 through November 22, 1989, International Research ASSOCIATES OF EUROPE (INRA (EUROPE)) CARRIED OUT THE FIELDWORK for this Euro-Barometer, at the request of the Survey, Research, Analysis Unit of the Directorate General X, Information, Communication and Culture, of the Commission of the European Communities. INRA (Europe) is a European network of market and public opinion research agencies, coordinated by the European Co-ordination Office (ECO), Avenue R. Vandendriessche 18, B-1150 Brussels. In all member states, fieldwork was conducted on the basis of detailed and uniform instructions prepared by the ECO.

INRA (Europe) developed equivalent French and English questionnaires for this Euro-Barometer (shown in Appendix B). These questionnaires were translated into other languages by the firms responsible for interviewing in each country.

In all, Euro-Barometer 32 interviewed 23,397 citizens of the 12 countries in the European Community. Respondents were ages 15 and over.

Two questionnaires, Form A and Form B, were used in the survey. Each was administered to one half of the respondents. The Form A questionnaire comprised almost all of the questions asked by Euro-Barometer 32. Form B contained a subset of the questions in Form A, plus some additional questions. The "Questionnaire Comparison Table", below, lists the questions contained in each questionnaire.

Welche Analysemöglichkeiten gibt es?

- Untersuchung einer Gesellschaft zu einem Zeitpunkt
- Vergleich...
 - einer Gesellschaft über einen Zeitraum hinweg
 - mehrerer Gesellschaften zu einem Zeitpunkt
 - über die Zeit und über Gesellschaften hinweg
- Aggregation oder Analyse von Individualdaten; Berücksichtigung von Kontextvariablen
- Eine Analyse *individueller* Veränderungen ist *nicht* möglich (keine Paneldaten)

Welche Probleme bereiten die EB-Studien?

- Indikatoren-Äquivalenz
 - innerhalb der EU geringeres Problem als sonst
 - aber z.B. nationale Wahlzyklen
- Bedeutungsveränderung über die Zeit
- Nationale Unterschiede beim Sampling und der sonstigen Durchführung
- Beschränkungen des Trend-Designs
- Kaum Informationen über soziales Umfeld, Nutzung von Medieninhalten etc.
- Kein Einfluß auf Gestaltung und Auswahl der Indikatoren

Welche Studien werden im Kurs vorgestellt?

- Analyse hauptsächlich oder ausschließlich auf EB-Daten gestützt
- Schwerpunkt auf Mikro-Ebene
 - Werte/Wertorientierungen
 - politische Einstellungen, insbesondere gegenüber politischen Institutionen auf der nationalen und europäischen Ebene
 - Wählerverhalten

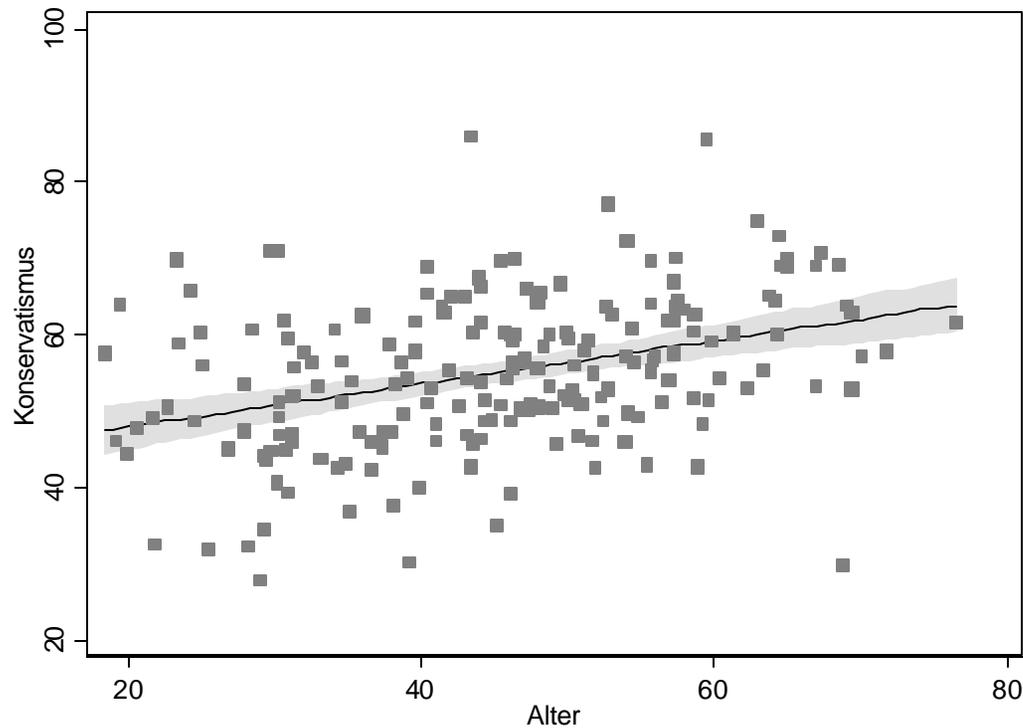
Welche methodischen Kenntnisse brauchen Sie?

- Die meisten der im Kurs vorgestellten Studien sind modellorientiert
- Zusammenhänge zwischen einer abhängigen und mehreren unabhängigen Variablen werden beschrieben oder überprüft
- Beim zugrundeliegenden statistischen Modell handelt es sich meist um Varianten des Verfahrens der multiplen Regression („the workhorse of most political science research“)
- Frischen Sie Ihre Kenntnisse zur Einfachregression auf und lesen Sie zumindest das Regressionskapitel aus Backhaus et al.

Wie sieht das Grundmodell der multiplen Regression aus?

- $y = a + \beta_1 \times x_1 + \beta_2 \times x_2 \dots + \varepsilon$ bzw.
- $y = a + b_1 \times x_1 + b_2 \times x_2 \dots + e$
- D.h.: Die Ausprägung einer abhängigen Variablen wird auf eine Konstante und die Ausprägung mehrerer unabhängiger Variablen zurückgeführt
- Koeffizienten (b_1, b_2 etc.) beschreiben, welche Veränderungen der abhängigen Variablen zu erwarten ist, wenn die entsprechende unabhängige Variable variiert und *alle anderen unabhängigen Variablen konstant gehalten werden*
- Konstante gibt den erwarteten Wert der abhängigen Variablen an, wenn alle unabhängigen Variablen gleich null sind (nicht immer sehr anschaulich; evtl. unabhängige Variablen zentrieren)
- Der Fehlerterm erfaßt alle *zufälligen* Einflüsse (u.a. alle sonstigen Variablen)

Warum multiple Regression?



- $\text{Konservatismus} = 42,3 + 0,28 * \text{Lebensalter}$: bei einem Altersunterschied von 10 Jahren ist ein um 2,8 Punkten höherer Skalenwert zu erwarten → Je älter, desto konservativer

Warum multiple Regression?

- Soziale Phänomene lassen sich selten auf eine einzige Ursache zurückführen
- Bei bivariater Betrachtung können „Scheinkorrelationen“ auftreten
 - In unserem Beispiel: Möglicher Zusammenhang zwischen Alter, Bildung (0-12) und Konservatismus
 - Zusammenhang zwischen Alter und Konservatismus kommt eventuell durch niedrigere Bildung der älteren Befragten zustande
- Lösung 1: Statistische Kontrolle durch Gruppenvergleich: Aufwendig und unübersichtlich
- Lösung 2: Multiple (multivariate) Regression:
 - $y = a + b_1x_1 + b_2x_2 \dots \Leftrightarrow \text{Konservatismus} = a + b_1\text{Alter} + b_2\text{Bildung}$
 - Simultane Schätzung für zwei oder mehr Einflußfaktoren!
 - Wechselseitige Kontrolle der Einflußfaktoren

Warum multiple Regression?

- Beispiel
 - Bivariat-1: Konservatismus= $42,3+0,28*\text{Lebensalter}$
 - Bivariat-2: Konservatismus= $73,1-3*\text{Bildung}$
 - Multivariat: Konservatismus= $84,4-0,17*\text{Lebensalter}-3,6*\text{Bildung}$!
- Das bedeutet
 - Bildung hat „in Wirklichkeit“ (d.h. bei Kontrolle des Alters) einen stärker negativen Einfluß auf Konservatismus als zunächst erkennbar
 - Für jede beliebige Altersgruppe nimmt der erwartete Konservatismus-Wert pro Punkt auf der Bildungsskala um 3,6 Punkte ab
 - Der Einfluß des Alters ist in „Wirklichkeit“ (bei Kontrolle der Bildung) schwach negativ!
 - Für jede beliebige Bildungsgruppe nimmt der erwartete Konservatismus-Wert pro Lebensjahr um 0,17 Punkte ab
 - Bei niedriggebildeten jungen sind die höchsten, bei hochgebildeten älteren Bürgern die niedrigsten Werte zu erwarten
 - Koeffizienten beziehen sich auf „natürliche“ Einheiten. Durch Standardisierung können Koeffizienten eventuell leichter vergleichbar gemacht werden
- Bivariate Analysen führen oft in die Irre. Generell müssen alle relevanten unabhängigen Variablen berücksichtigt werden, damit man zu validen Ergebnissen kommt

Wie wird die Regressionsgleichung bestimmt?

- „Methode der kleinsten Quadrate“, d.h. gesucht werden Werte für $a+b_1+b_2\dots$, die die Summe der quadrierten Abweichungen zwischen beobachteten und errechneten x -Werten minimieren
- Dafür existiert eine analytische Lösung
- Mit Bezug auf dieses Verfahren wird die „normale“ lineare Regression in den Studien oft als OLS („Ordinary Least Squares“) bezeichnet

Was sind Standardfehler?

- Regressionskoeffizienten sind Schätzungen
- Für eine andere Stichprobe würden sich etwas andere Werte ergeben
- Wie stark sich die Koeffizienten über die Stichproben hinweg unterscheiden, kann mit Hilfe des Standardfehlers geschätzt werden
- Der Standardfehler ist eine Schätzung für die Streuung (=Standardabweichung) einer Normal- bzw. t-Verteilung der Koeffizienten um deren (unbekannten) wahren Wert

Was macht man mit Standardfehlern?

- Generell geben Standardfehler Auskunft darüber, wie sicher man sich über den Wert eines Koeffizienten sein kann
- Signifikanztests
 - Wie groß ist die Wahrscheinlichkeit p , den Stichprobenwert zu erhalten, wenn der wahre Parameter = 0 (Nullhypothese)
 - Wenn $p < 0,05$ od. $0,01$ ist der Wert „statistisch signifikant“ von null verschieden
 - Schließt nicht aus, daß Wert doch = 0
 - Statistische Signifikanz \neq inhaltliche Relevanz
 - Andere Nullhypothesen möglich, aber fast nie untersucht
- Konfidenzintervalle
 - Innerhalb welchen Bereiches liegt mit einer gewissen Wahrscheinlichkeit (95/99%) der wahre Parameter?
 - Inhaltliche Interpretation der Koeffizienten – „mit großer Sicherheit nimmt der Konservatismus-Wert mit jedem Lebensjahr um mindestens 0,01 und höchstens 0,33 Skalenpunkte ab“

Was passiert wenn...

- Das klassische lineare Modell geht von einigen Annahmen aus
- Gegenüber vielen Verletzungen dieser Annahmen ist das Modell relativ robust
- In Studien wie dem Eurobarometer werden zwei Annahmen besonders häufig verletzt:
 - Die Varianz der Fehler soll über verschiedene Ausprägungen der unabhängigen Variablen hinweg identisch sein (Homoskedasität)
 - Die Beobachtungen (und damit die Fehler) sollen voneinander unabhängig sein (keine Autokorrelation)

Was passiert wenn...

- Beispiele
 - Heteroskedasität
 - die Varianz der zufälligen Fehler nimmt mit dem politischen Interesse ab
 - die Varianz der Fehler ist manchen Ländern größer als in anderen
 - Autokorrelation
 - länderspezifische Faktoren werden nicht modelliert (räumliche Korrelation)
 - In Zeitreihen wirken nicht-spezifizierte Einflüsse zum Zeitpunkt t auch noch bei $t+1$ (serielle Korrelation)
- Konsequenzen
 - Schätzung der Koeffizienten i.d.R. weiterhin unverzerrt, aber nicht mehr effizient → evt. alternative Schätzverfahren
 - Schätzung der Standardfehler (nach unten) verzerrt → Konfidenzintervalle und Tests zu optimistisch → evtl. alternative Berechnung der Standardfehler notwendig

Wo finde ich Hilfe?

- Anwendungsorientierte Einführungen zu multivariaten Verfahren
 - Backhaus et al. (diverse Auflagen)
 - Hair et al. (diverse Auflagen)
- Regression
 - Berry 1993 (knappe Einführung in die Annahmen des klassischen Regressionsmodells und deren Konsequenzen)
 - Fox 1997 (sehr umfassende Darstellung der linearen Regression und zahlreicher verwandter Modelle; eindeutig für Fortgeschrittene)