



Faktorenanalyse und Clusteranalyse

illustriert mithilfe der Fragestellung
„Die Bürger und ihre Europäische Union. Hat die EU
ein Demokratie- und Legitimitätsdefizit?“

Ausführliche Variante der Vorlesung mit Screenshots zur Durchführung der Analysen
mithilfe von SPSS

Einstellungen der Bürger zur EU

Nicht unproblematisch

- Euroskeptizismus
- Unkenntnis
- „Demokratiedefizit“
 - Repräsentation der Bürger, Funktion der Institutionen
- „Legitimitätsdefizit“
 - Partizipation besonders an Wahlen zum EP, Willensbildungsprozess

Fragestellungen

- Welche Einstellungsmuster zeigt die Europäische Bevölkerung?
→ Faktorenanalyse
- Welche Länder sind besonders betroffen? Gibt es Ländergruppen?
→ Clusteranalyse

Forschungsdesign

Trennung der Begriffe

Theorie

Definition „Demokratiedefizit“
Definition „Legitimitätsdefizit“

Operationalisierung

Umfrage:
Eurobarometer 2008

Fragestellungen/Variablen
zu Demokratie und
Legitimität

Faktoren-
analyse

Einstel-
lungs-
dimen-
sionen

Demokratie

Legitimität

Länderunterschiede ?

Kennzeichnungsvariablen
aus der Faktorenanalyse

Cluster-
analyse

Ländergruppen mit ähnlicher
Einstellungsausprägung

Demokratie

Indikatoren ...	Variablen ... der demokratische Herrschaftsweise der EU
Herrschaft nach demokratischen Prinzipien/ Herrschaft für das Volk	<p>EU-Image: Ganz allgemein gesprochen, ruft die Europäische Union bei Ihnen ein sehr positives, ziemlich positives, weder positives noch negatives, ziemlich negatives oder sehr negatives Bild hervor?</p> <p>Stimme_EU: Meine Stimme zählt in der Europäischen Union.</p> <p>Stimme_MEP: Meine Meinung zu europäischen Themen wird von den Mitgliedern des Europäischen Parlaments berücksichtigt.</p> <p>Stimme_ER: Meine Meinung zu europäischen Themen wird vom Europäischen Rat berücksichtigt.</p>
Institutionen der Herrschaft und ihre Funktionalität	<p>Vertrauen_EU: Ich möchte nun gerne von Ihnen wissen, wie viel Vertrauen Sie in bestimmte Institutionen haben. Sagen Sie mir bitte für jede der folgenden Institutionen, ob Sie ihr eher vertrauen oder eher nicht vertrauen. Wie ist es mit...? EU</p>
EU funktioniert so wie sie nach den demokratischen Regeln funktionieren soll	<p>Welche der folgenden sind die Hauptgründe dafür, dass Sie den europäischen Institutionen vertrauen?</p> <p>VGrund_EP_Demo: Die Entscheidungen des Europäischen Parlaments werden auf demokratische Weise getroffen.</p> <p>VGrund_Komm_Demo: Die Entscheidungen der Europäischen Kommission werden auf demokratische Weise getroffen.</p> <p>VGrund_ER_Demo: Die Entscheidungen des Rates der EU werden auf demokratische Weise getroffen</p>

Operationalisierung: Legitimität aus politischer Unterstützung

Indikatoren politischer Unterstützung der Europäischen Union

Legitimität	<ul style="list-style-type: none"> • Mitgliedschaft in der EU als gute oder schlechte Sache (= Anerkennung an sich)
①	<p>MTG_EU: Ist die Mitgliedschaft (UNSER LAND)s in der Europäischen Union Ihrer Meinung nach eine gute oder eine schlechte Sache?</p>
Vertrauen	<ul style="list-style-type: none"> • Vertrauen in die einzelnen Institutionen der EU <p>Sagen Sie mir bitte für jede dieser europäischen Einrichtungen, ob Sie ihr eher vertrauen oder eher nicht vertrauen.</p> <p>Vertrauen_EP: Europäischen Parlament</p> <p>Vertrauen_Komm: Europäischen Kommission</p> <p>Vertrauen_ER: Rat der Europäischen Union</p> <p>Vertrauen_EZB: Europäischen Zentralbank</p>
Effektivität	<ul style="list-style-type: none"> • Zufriedenheit mit der EU
②	<p>zufr_RichtungEU: Würden Sie sagen, dass sich die Dinge in ... derzeit ganz allgemein in die richtige Richtung oder in die falsche Richtung entwickeln? Der Europäische Union</p>
OUTPUT	<ul style="list-style-type: none"> • Nutzen der EU-Mitgliedschaft für das eigene Land <p>mtgEU_Nutzen: Hat Ihrer Meinung nach (UNSER LAND) insgesamt gesehen durch die Mitgliedschaft in der Europäischen Union Vorteile, oder ist das nicht der Fall?</p>

Fallauswahl und Fallzahl

- Alle europäischen Bürger im Eurobarometer 69.2 März-Mai 2008: N= 30.170
→ Faktorenanalyse
- Alle EU-Mitgliedsländer und weitere Staaten: N=29
→ Clusteranalyse

Faktoren- und Clusteranalyse

- Hauptverfahren der Politikwissenschaft, aber auch der Markt- und Medienforschung
- Reduktion von (zu) großen und damit unübersichtlichen Datenmengen
- Multivariate Verfahren ohne Kausalstruktur
- Wissensgenerierend → Plausibilität, Validität, Reliabilität, Stabilität

Dimensionsreduktion – Faktorenanalyse

- Voraussetzungen:
 - Metrisches Skalenniveau → Skalenniveaus
 - Fallzahl = Variablenzahl *besser* Variablenzahl * 3
 - Strukturierungsfähige Daten (z.B. Skalen, item-Batterien)

Datenreduktion – Faktorenanalyse I

- Datenreduktionistisches Verfahren: viele Variablen werden zu einer inhaltlichen Dimension (= latente Variable) zusammengefasst, die das Variablenbündel ausreichend beschreibt
- Zweck: Ähnliche Variablen zu „Dimensionen“ zusammenfassen → Variablenreduktion mittels Korrelation der Variablen mit einer dritten, unbekannten Variable im „Hintergrund“
- im Beispiel:
8 Indikatoren für Demokratie und 7 Indikatoren für Legitimität können in 2 Dimensionen gebündelt werden

Datenreduktion – Faktorenanalyse II

explorativ:

- Aufdeckung unbekannter Strukturen der Variablen im Datensatz: vor der Analyse ist die Anzahl der zu extrahierenden Faktoren/Dimensionen und deren Bedeutung unbekannt
- Nur vage theoretische Vermutungen über die Zusammengehörigkeit von Indikatoren
- Herstellung eines Entdeckungszusammenhangs
- Hypothesengenerierung

Datenreduktion – Faktorenanalyse III

konfirmatorisch:

- Konkrete Vorstellung über mögliche hypothetische Faktoren hinter empirisch beobachteten Korrelationen von Variablen:
 1. nachzuweisende latente Faktoren werden definiert
 2. Festlegung der manifesten Variablen, die die Faktoren erfassen sollen
- Hypothesen über die Beziehung zwischen den beobachtbaren Variablen und die dahinterstehenden Dimensionen → **Demokratie und Legitimität**
- Hypothesentest
- Herstellung eines Begründungszusammenhangs

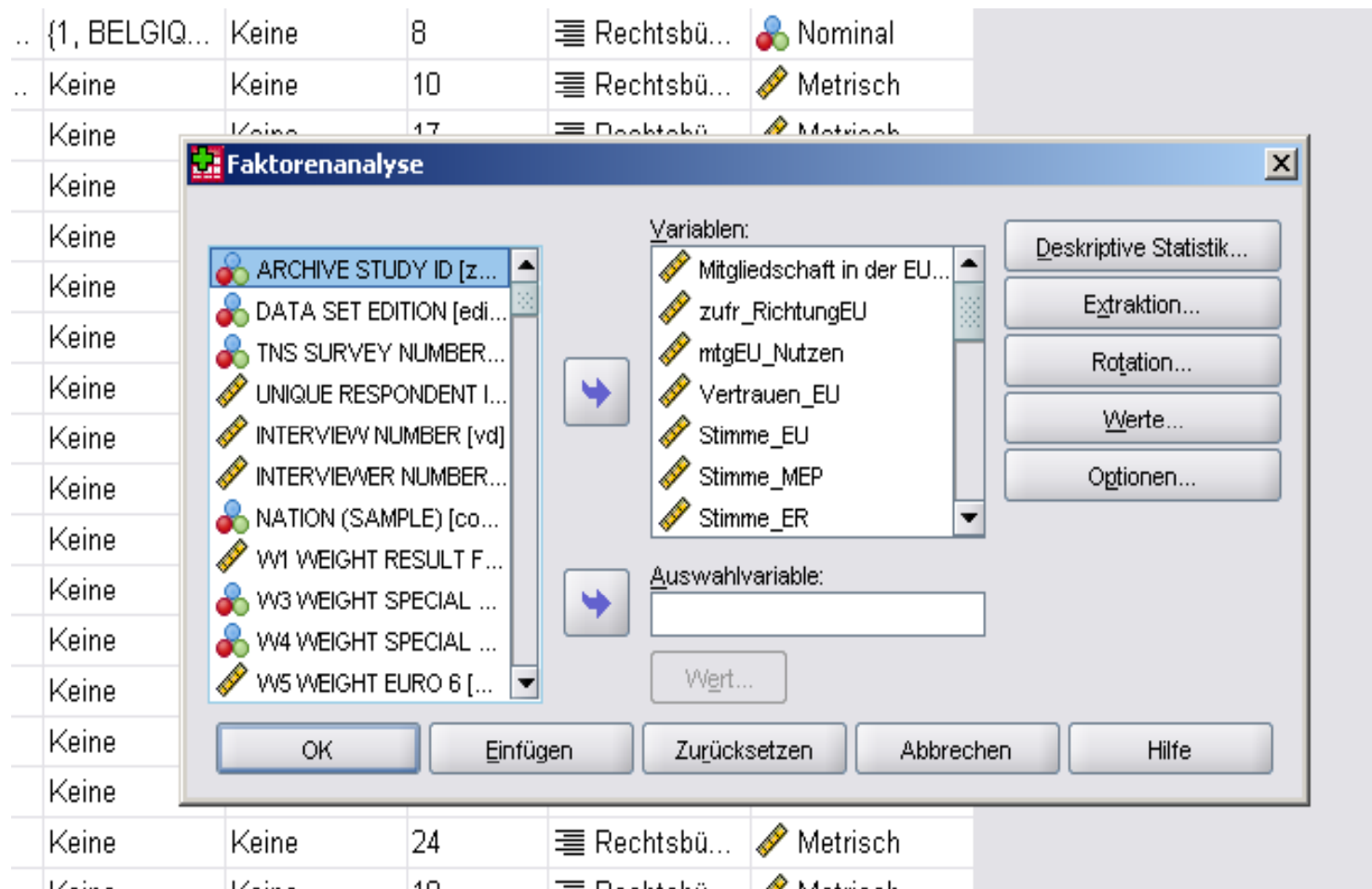
Datenreduktion – Analyseschritte

- 1) Datenmatrix - Dateneigenschaften:
Individualdaten des Eurobarometer
- 2) Korrelationsmatrix
- 3) Kommunalitätenschätzung
- 4) Faktorextraktion
- 5) Bestimmung der Faktorenzahl
- 6) Rotation
- 7) Interpretation
- 8) Bestimmung der Faktorwerte

Konfirmatorische Faktorenanalyse

– 1 Datenmatrix

→ Menü: Analysieren - Dimensionsreduzierung



Konfirmatorische Faktorenanalyse – 2

Korrelation

The screenshot shows the 'Faktorenanalyse: Deskriptive Statistiken' dialog box in SPSS. The 'Statistiken' section has 'Anfangslösung' checked. The 'Korrelationsmatrix' section has 'Koeffizienten', 'Signifikanzniveaus', and 'KMO und Bartlett-Test auf Sphärizität' checked. The 'Deskriptive Statistik...' button is circled in red.

Konfirmatorische Faktorenanalyse - 2

Korrelation

Korrelationsmatrix

	MTG_eu Mitgliedschaft in der EU - gut/schlecht für unser Land	zufr_Richtung EU	mtgEU_ Nutzen	Vertrauen_EU	Stimme EU	Stimme MEP	Stimme ER	Vertrauen EP	Vertrauen Komm
Korrelation	MTG_eu Mitgliedschaft in der EU - gut/schlecht für unser Land	1,000	,356	,627	,467	,309	,274	,262	,395
	zufr_RichtungEU	,356	1,000	,347	,342	,241	,237	,228	,290
	mtgEU_Nutzen	,627	,347	1,000	,465	,281	,260	,251	,374
	Vertrauen_EU	,467	,342	,465	1,000	,300	,280	,280	,541
	Stimme_EU	,309	,241	,281	,300	1,000	,584	,578	,288
	Stimme_MEP	,274	,237	,260	,280	,584	1,000	,776	,283
	Stimme_ER	,262	,228	,251	,280	,578	,776	1,000	,293
	Vertrauen_EP	,395	,290	,374	,541	,288	,283	,243	1,000
	Vertrauen_Komm	,386	,290	,364	,533	,295	,286	,305	,001
	Vertrauen_ER	,377	,285	,356	,522	,295	,288	,291	,763
	Vertrauen_EZB	,348	,277	,319	,456	,260	,247	,239	,654
	EU_Image	,525	,346	,478	,553	,307	,287	,292	,481
	VGrund_EP_Demo	,026	,027	,016	,023	,011	,016	-,003	,000
	VGrund_Komm_Demo	,016	,019	,006	,022	,020	,016	,005	,017
	VGrund_ER_demo	,018	,020	,010	,012	-,005	-,004	-,002	,016
Signifikanz (1-seitig)	MTG_eu Mitgliedschaft in der EU - gut/schlecht für unser Land		,000	,000	,000	,000	,000	,000	,000
	zufr_RichtungEU	,000		,000	,000	,000	,000	,000	,000
	mtgEU_Nutzen	,000	,000		,000	,000	,000	,000	,000
	Vertrauen_EU	,000	,000	,000		,000	,000	,000	,000
	Stimme_EU	,000	,000	,000	,000		,000	,000	,000
	Stimme_MEP	,000	,000	,000	,000	,000		,000	,000
	Stimme_ER	,000	,000	,000	,000	,000	,000		,000
	Vertrauen_EP	,000	,000	,000	,000	,000	,000	,000	
	Vertrauen_Komm	,000	,000	,000	,000	,000	,000	,000	,000
	Vertrauen_ER	,000	,000	,000	,000	,000	,000	,000	,000
	Vertrauen_EZB	,000	,000	,000	,000	,000	,000	,000	,000
	EU_Image	,000	,000	,000	,000	,000	,000	,000	,000
	VGrund_EP_Demo	,000	,000	,007	,000	,038	,007	,333	,500
	VGrund_Komm_Demo	,004	,001	,187	,000	,001	,007	,207	,003
	VGrund_ER_demo	,002	,000	,057	,024	,228	,280	,405	,006

Ein Faktor?

Konfirmatorische Faktorenanalyse – 3

Kommunalitäten

Kommunalitäten

	Anfänglich	Extraktion
MTG_eu Mitgliedschaft in der EU - gut/schlecht für unser Land	1,000	,700
zufr_RichtungEU	1,000	,376
mtgEU_Nutzen	1,000	,687
Vertrauen_EU	1,000	,574
Stimme_EU	1,000	,649
Stimme_MEP	1,000	,828
Stimme_ER	1,000	,825
Vertrauen_EP	1,000	,805
Vertrauen_Komm	1,000	,845
Vertrauen_ER	1,000	,836
Vertrauen_EZB	1,000	,696
EU_Image	1,000	,587
VGrund_EP_Demo	1,000	,544
VGrund_Komm_Demo	1,000	,607
VGrund_ER_demo	1,000	,576

Varianz der Variable zufr_RichtungEU wird nur zu einem geringen Anteil erklärt → aus der Analyse entfernen?

Erklärte
Varianzanteile
der Variablen,
die durch die
Dimensionen
erklärt werden

Extraktionsmethode:
Hauptkomponentenanalyse.

Konfirmatorische Faktorenanalyse – 4

Komponentenextraktion

Hauptkomponentenanalyse

- Führt viele beobachtete Merkmale durch Transformation in wenige, unabhängige Komponenten über, die sukzessiv und in ihrer Gesamtheit die Varianz der beobachteten Variablen *vollständig* erklären.

$$z_{ij} = a_{j1}s_{1i} + a_{j2}s_{2i} + \dots + a_{jp}s_{jp}i$$

z_{ij} ist der Score von Person i in der manifesten Variablen j

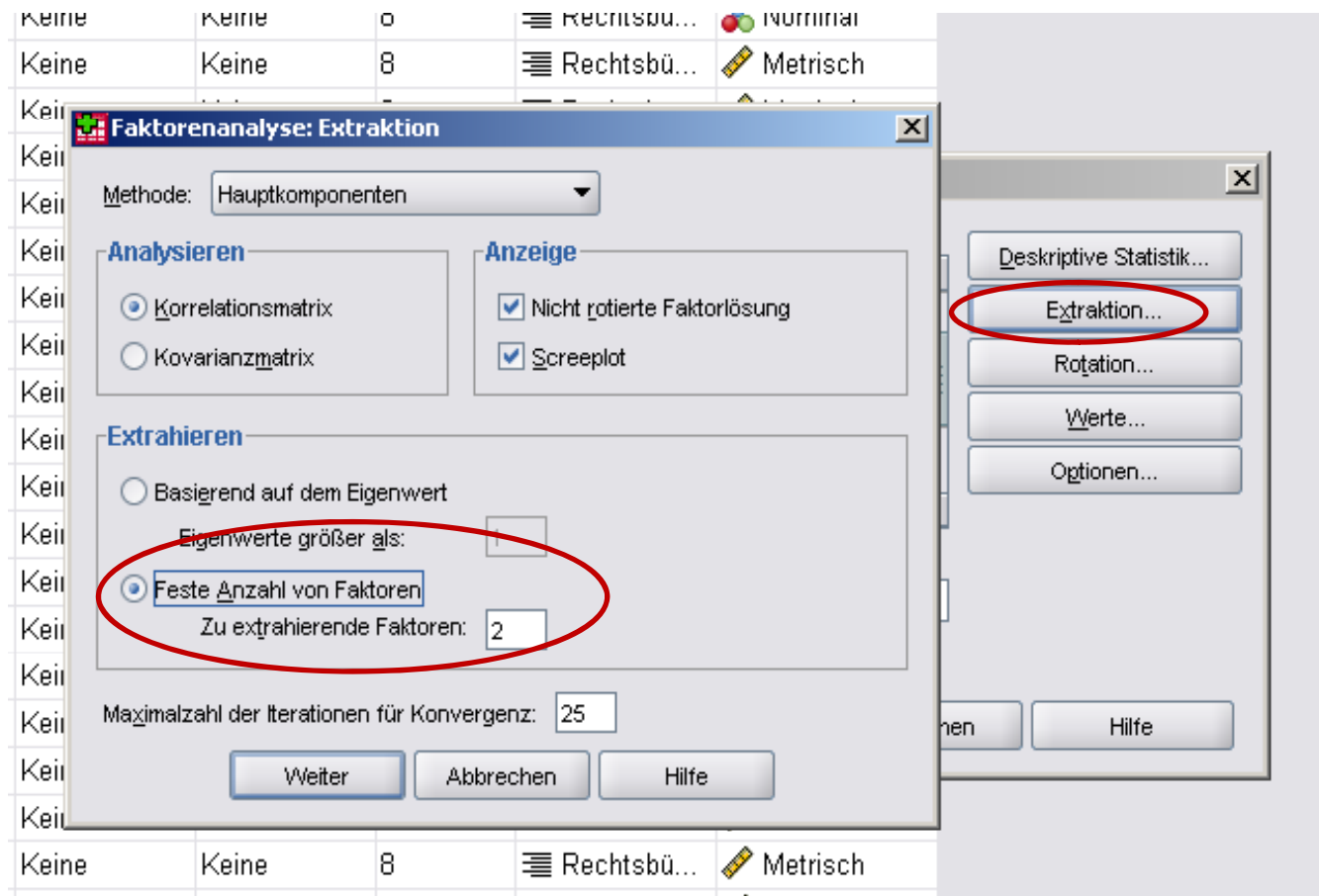
a_{jk} ist die Ladung der manifesten Variablen j auf der Komponente k ($k=1, \dots, p$)

s_{ki} ist der Score von Person i auf der Komponente k

Konfirmatorische Faktorenanalyse – 4

Komponentenextraktion

→ Extraktionsmethode und Faktorzahl



Konfirmatorische Faktorenanalyse

– 5 Bestimmung der Komponentenzahl

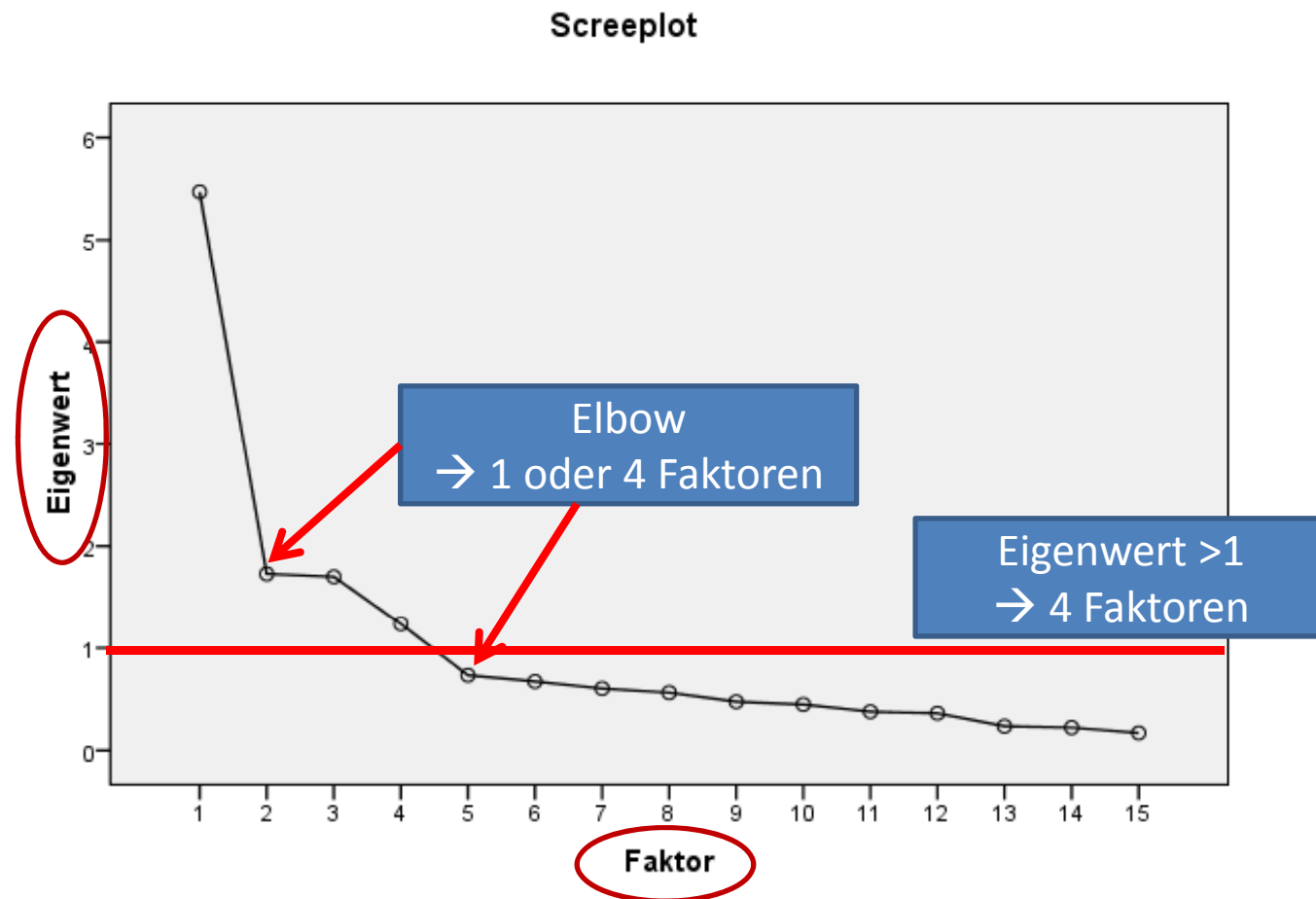
Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	5,471	36,474	36,474	5,471	36,474	36,474
2	1,728	11,523	47,997	1,728	11,523	47,997
3	1,701	11,339	59,336			
4	1,237	8,244	67,580			
5	,733	4,183	72,469			
6	,672	4,183	76,949			
7	,604	3,974	80,974			
8	,563	3,730	84,730			
9	,475	3,188	87,896			
10	,448	2,986	90,882			
11	,377	2,516	93,398			
12	,362	2,413	95,810			
13	,236	1,571	97,381			
14	,222	1,482	98,863			
15	,171	1,137	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Konfirmatorische Faktorenanalyse – 5

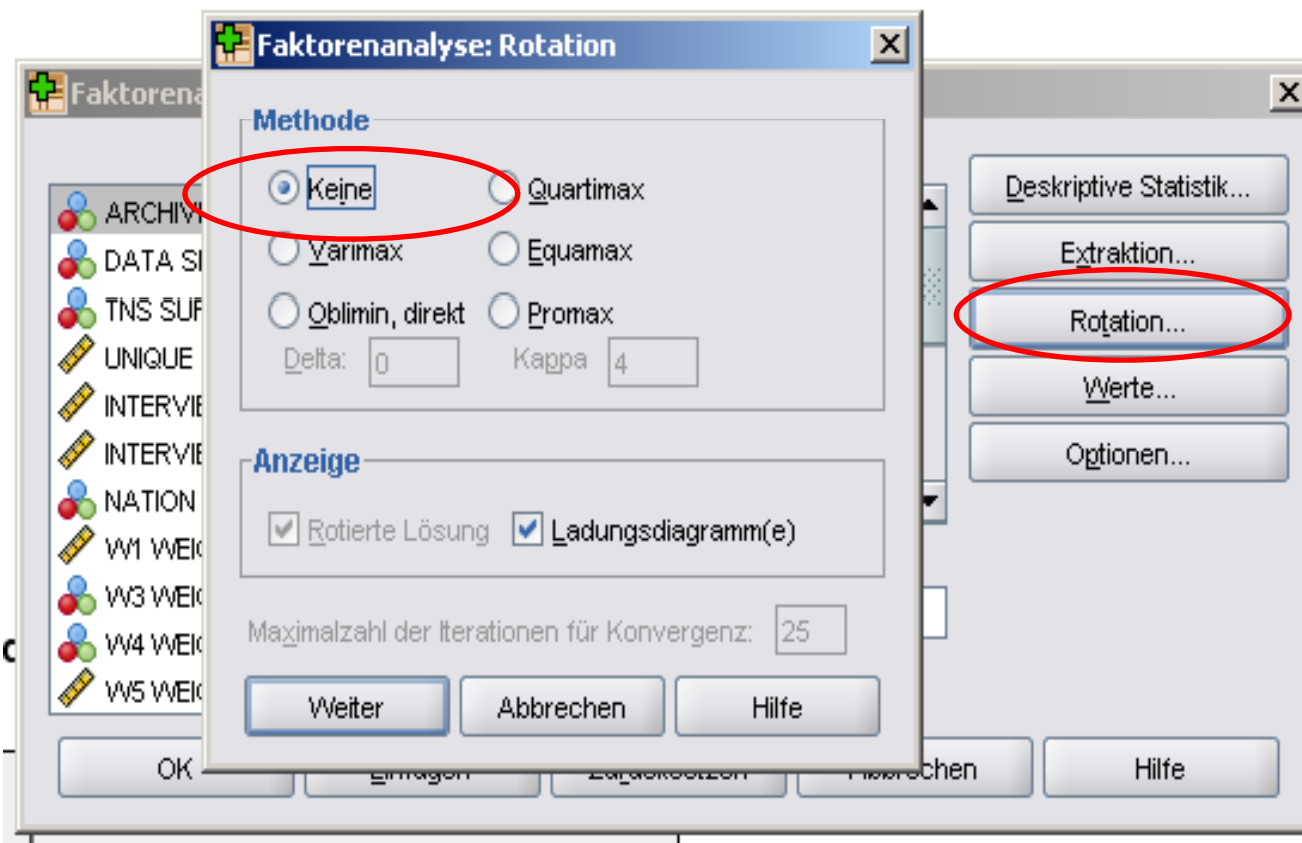
Bestimmung der Faktorenzahl

Theoretische Annahme: Die Variablen bilden die Dimensionen „Demokratie“ und „Legitimität“ ab → 2 Faktoren



Konfirmatorische Faktorenanalyse – 6 Rotation der Faktoren

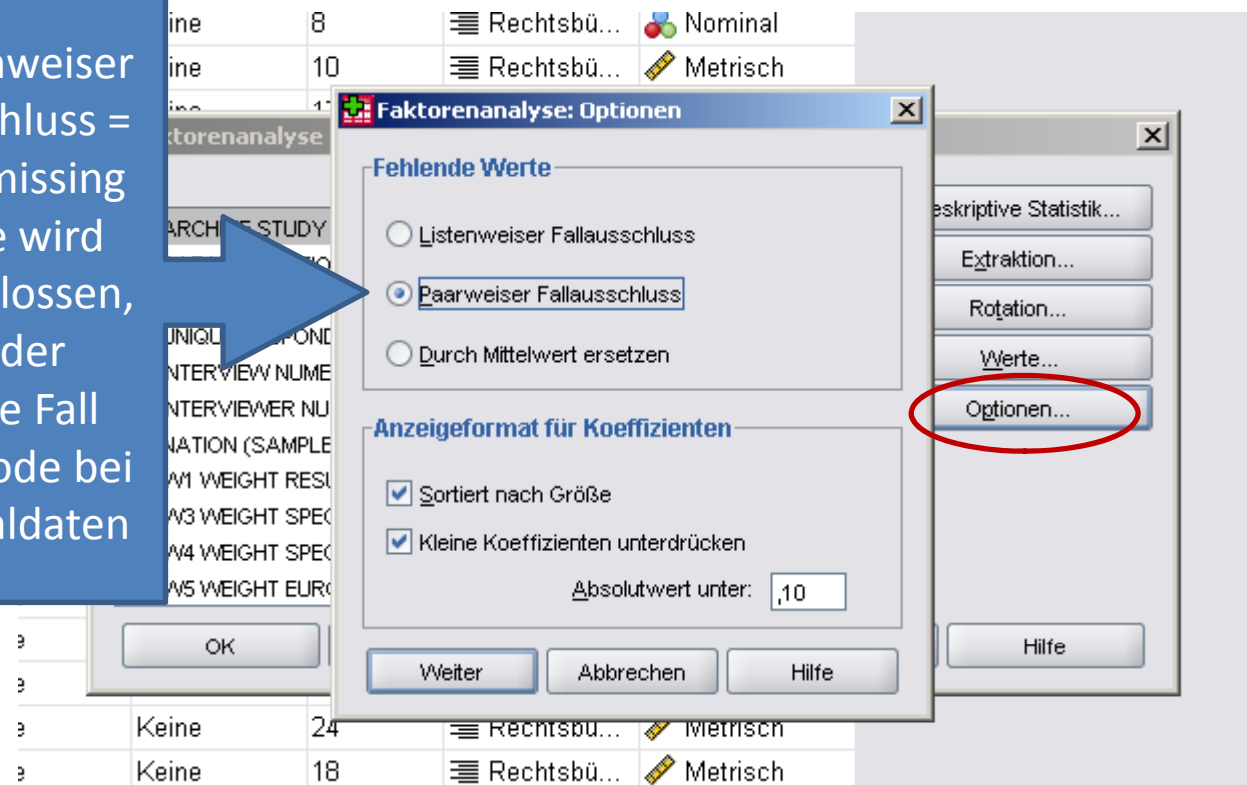
Hauptkomponentenanalyse benötigt keine Rotation der Faktoren



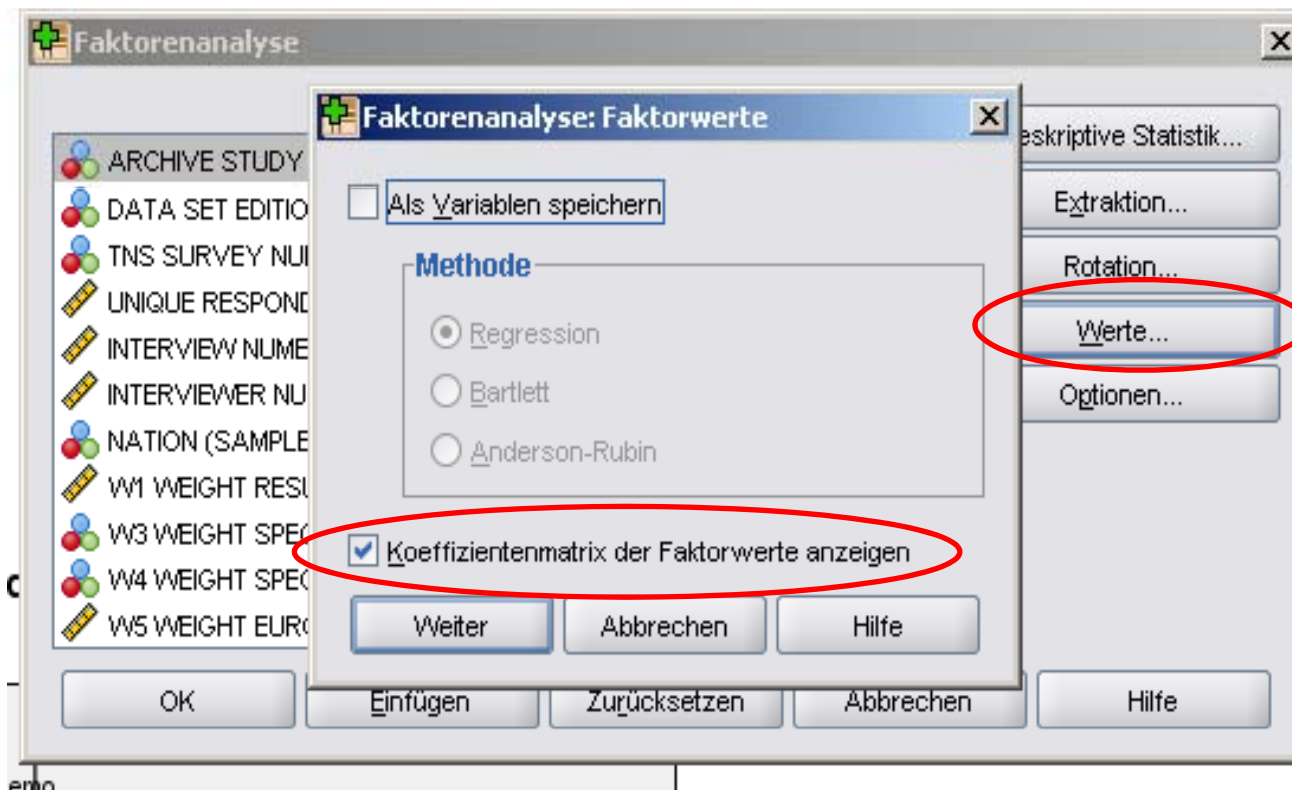
Konfirmatorische Faktorenanalyse – 7 Interpretation

Ziel: Erleichterung der Interpretation

Variablenweiser
Fallausschluss =
nur die missing
variable wird
ausgeschlossen,
nicht der
gesamte Fall
→ Methode bei
Individualdaten



Konfirmatorische Faktorenanalyse – 7 Interpretation der Komponenten



Konfirmatorische Faktorenanalyse – 7 Interpretation der Komponenten

„Gewichte“, mit denen die standardisierten Werte der Originalvariablen multipliziert werden können, um den jeweiligen Wert der Komponenten ermitteln zu können

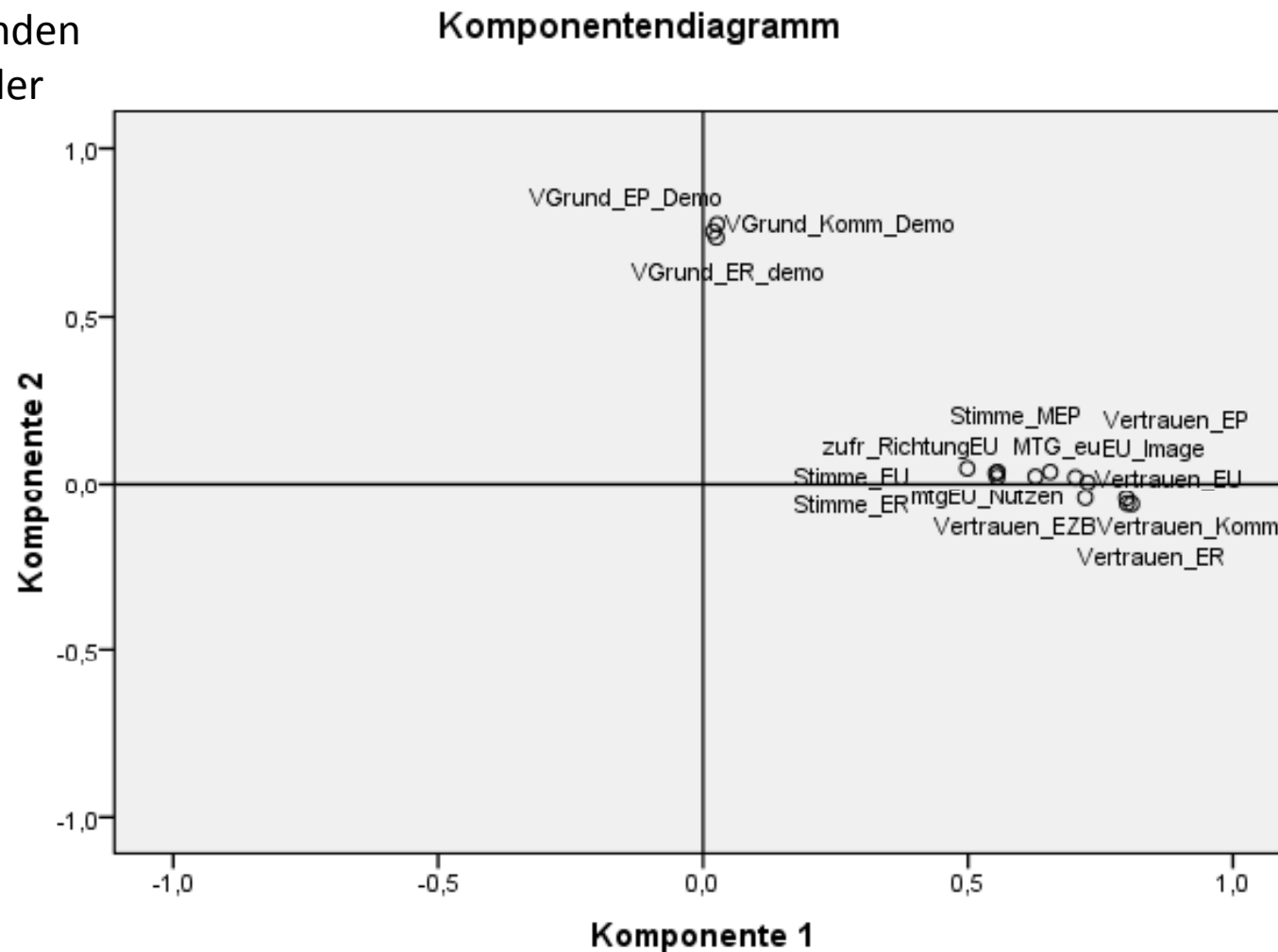
Koeffizientenmatrix der Komponentenwerte

	Komponente	
	1	2
MTG_eu Mitgliedschaft in der EU - gut/schlecht für unser Land	,120	,021
zufr_RichtungEU	,091	,027
mtgEU_Nutzen	,115	,013
Vertrauen_EU	,133	,003
Stimme_EU	,101	,019
Stimme_MEP	,102	,021
Stimme_ER	,101	,012
Vertrauen_EP	,146	-,029
Vertrauen_Komm	,148	-,039
Vertrauen_ER	,146	-,039
Vertrauen_EZB	,132	-,028
EU_Image	,128	,012
VGrund_EP_Demo	,005	,425
VGrund_Komm_Demo	,005	,448
VGrund_ER_demo	,004	,436

Extraktionsmethode:
Hauptkomponentenanalyse.

Konfirmatorische Faktorenanalyse – 7 Interpretation der Komponenten

Nähe der ladenden
Variablen auf der
jeweiligen
Komponente



Konfirmatorische Faktorenanalyse – 7 Interpretation der Komponenten

**Kovarianzmatrix des
Komponentenwerts**

Komponente	1	2
1	1,000	,000
2	,000	1,000

Extraktionsmethode:
Hauptkomponentenanalyse.

Komponenten werden so ermittelt, dass die nicht miteinander korrelieren
→ Unabhängigkeit der Komponenten

Konfirmatorische Faktorenanalyse – 7

Interpretation

Variable mit der
höchsten
Faktorladung =
Kennzeichnungs-
variable

2
Komponenten:
fassen die
Variablen
inhaltlich
zusammen

Vertrauen
Legitimität
Effektivität
=
Politische
Unterstützung

Demokratie

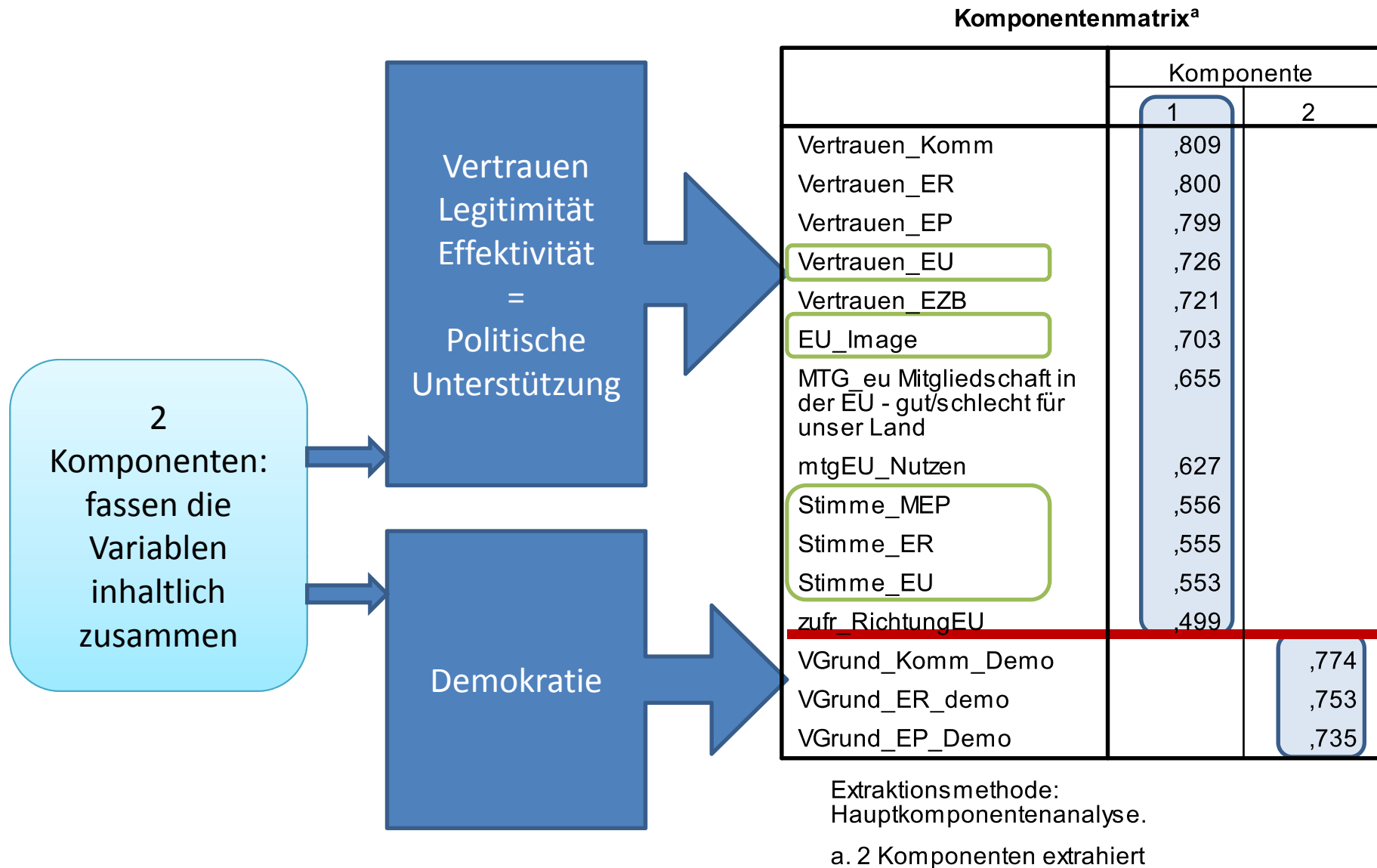
Komponentenmatrix^a

	Komponente	
	1	2
Vertrauen_Komm	,809	
Vertrauen_ER	,800	
Vertrauen_EP	,799	
Vertrauen_EU	,726	
Vertrauen_EZB	,721	
EU_Image	,703	
MTG_eu Mitgliedschaft in der EU - gut/schlecht für unser Land	,655	
mtgEU_Nutzen	,627	
Stimme_MEP	,556	
Stimme_ER	,555	
Stimme_EU	,553	
zufr_RichtungEU	,499	
VGrund_Komm_Demo		,774
VGrund_ER_demo		,753
VGrund_EP_Demo		,735

Extraktionsmethode:
Hauptkomponentenanalyse.

a. 2 Komponenten extrahiert

Konfirmatorische Faktorenanalyse III



Konfirmatorische Faktorenanalyse – Ergebnis

- Konfirmatorische Faktorenanalyse bestätigt die Existenz von zwei Faktoren, die mit politischer Unterstützung und Demokratie bezeichnet werden können
ABER: Indikatoren beider **theoretischen** Dimensionen werden gemischt
- Zwei Faktoren, geringere Varianzaufklärung → einfachere theoretische Erklärung (?)

Exkurs: explorative Faktorenanalyse

Veränderung im Bereich der

- Faktorextraktion
- Bestimmung der Faktorenzahl
- Interpretation

Explorative Faktorenanalyse

– 4 Faktorextraktion

→ Extraktionsmethode und Faktorzahl

Extraktion von Faktoren mit Eigenwerten > 1 (Kaiser-Kriterium)

The screenshot shows the 'Faktorenanalyse: Extraktion' dialog box in SPSS. The 'Methode' dropdown is set to 'Hauptkomponenten'. Under 'Analysieren', 'Korrelationsmatrix' is selected. Under 'Anzeige', 'Nicht rotierte Faktorenlösung' and 'Screeplot' are checked. Under 'Extrahieren', 'Basierend auf dem Eigenwert' is selected, and the 'Eigenwerte größer als:' field is set to 1. The 'Maximalzahl der Iterationen für Konvergenz' is set to 25. A blue callout box on the left points to the 'Extrahieren' section with the text 'Extraktion von Faktoren mit Eigenwerten > 1 (Kaiser-Kriterium)'. A red circle highlights the 'Methode' dropdown, and another red circle highlights the 'Basierend auf dem Eigenwert' option and the 'Eigenwerte größer als: 1' field. In the background, another dialog box is visible with the 'Extraktion...' button circled in red.

Keine Keine 8 Rechtsbü... Metrisch

Faktorenanalyse: Extraktion

Methode: Hauptkomponenten

Analysieren

☒ Korrelationsmatrix
☐ Kovarianzmatrix

Anzeige

☒ Nicht rotierte Faktorenlösung
☒ Screeplot

Extrahieren

☒ Basierend auf dem Eigenwert
Eigenwerte größer als: 1
☐ Feste Anzahl von Faktoren
Zu extrahierende Faktoren: 2

Maximalzahl der Iterationen für Konvergenz: 25

Weiter Abbrechen Hilfe

Deskriptive Statistik...
Extraktion...
Rotation...
Werte...
Optionen...
Hilfe

Explorative Faktorenanalyse

– 5 Bestimmung der Faktorenzahl

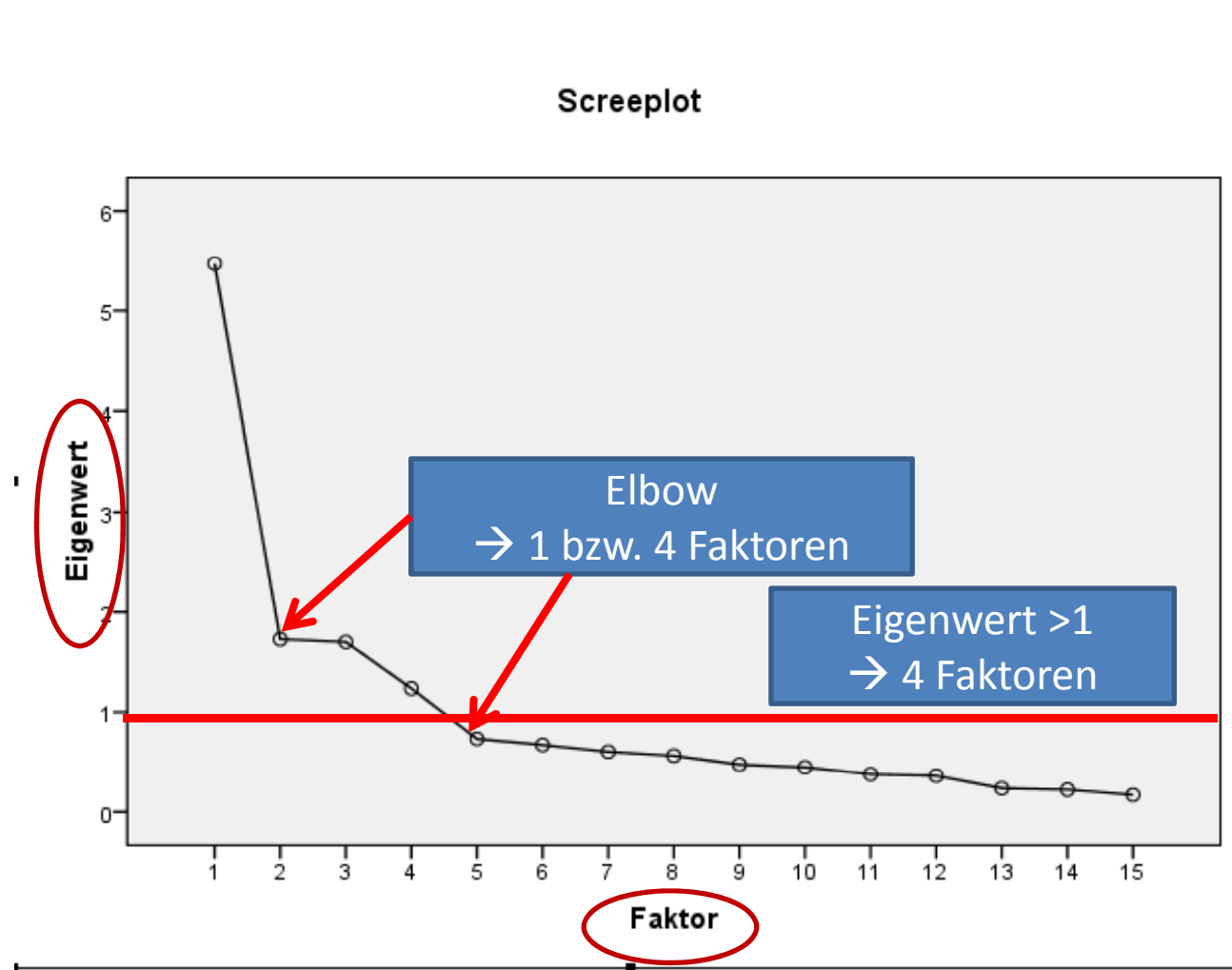
Anfängliche Eigenwerte				Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen ^a
Komponente	Gesamte Varianzaufklärung der Dimensionierung			Eigenwerte > 1			Gesamt
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	
1	5,471	36,474	36,474	5,471	36,474	36,474	4,531
2	1,728	11,523	47,997	1,728	11,523	47,997	1,731
3	1,701	11,339	59,336	1,701	11,339	59,336	3,188
4	1,237	8,244	67,580	1,237	8,244	67,580	3,987
5	,733	4,889	72,469				
6	,672	3,779	76,949				
7	,500	2,974	80,974				
8	,473	2,730	84,730				
9	,377	2,516	87,896				
10	,448	2,986	90,882				
11	,377	2,516	93,398				
12	,362	2,413	95,810				
13	,236	1,571	97,381				
14	,222	1,482	98,863				
15	,171	1,137	100,000				

Extraktionsmethode: Hauptkomponentenanalyse.

a. Wenn Komponenten korreliert sind, können die Summen der quadrierten Ladungen nicht addiert werden, um eine Gesamtvarianz zu erhalten.

Explorative Faktorenanalyse

– 5 Bestimmung der Faktorenzahl

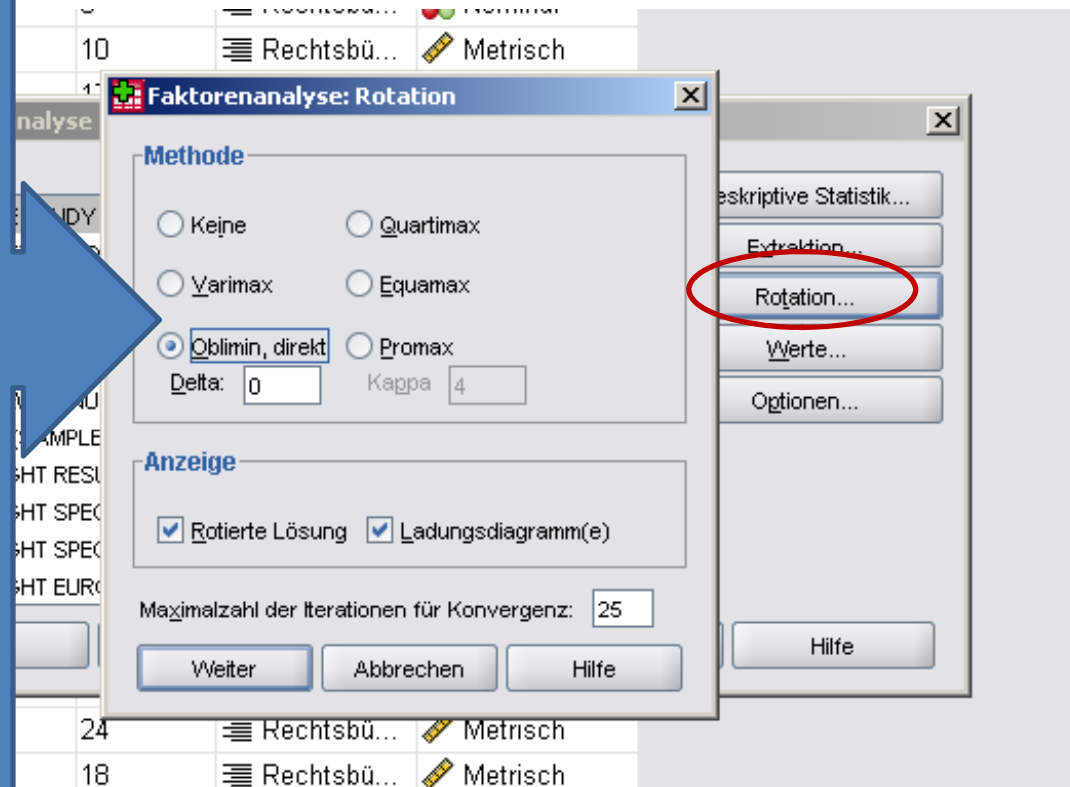


Explorative Faktorenanalyse – 6 Rotation der Faktoren

Varimax:
orthogonale
Rotation =
Faktorachsen im
90° Winkel →
Achsen bzw.
Faktoren
korrelieren nicht
untereinander

Oblimin:
oblique Rotation =
Faktoren können
korrelieren =
realitätsnäher bei
Einstellungen

Ziel: möglichst viele Punkte sollen auf den Achsen liegen



Explorative Faktorenanalyse – 6 Rotation der Faktoren

Rotationsvarianten

Kovarianzmatrix des Komponentenwerts

Komponente	1	2	3	4
1	1,727	,057	2,548	1,549
2	,057	1,002	,394	,062
3	2,548	,394	3,614	1,776
4	1,549	,062	1,776	1,727

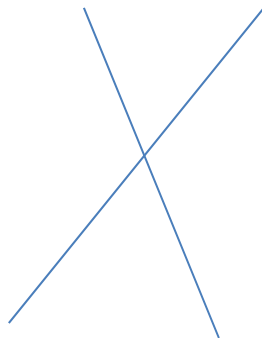
Extraktionsmethode: Hauptkomponentenanalyse.
Rotationsmethode: Oblimin mit Kaiser-Normalisierung.

Kovarianzmatrix des Komponentenwerts

Komponente	1	2	3	4
1	1,000	,000	,000	,000
2	,000	1,000	,000	,000
3	,000	,000	1,000	,000
4	,000	,000	,000	1,000

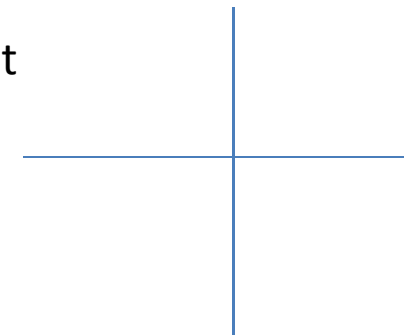
Extraktionsmethode: Hauptkomponentenanalyse.
Rotationsmethode: Varimax mit Kaiser-Normalisierung.

Schiefwinklige Rotation



Ziel:
Optimierung der
Anpassung der
Variablen an die
Faktoren/
Dimensionen

Rechtwinklige Rotation

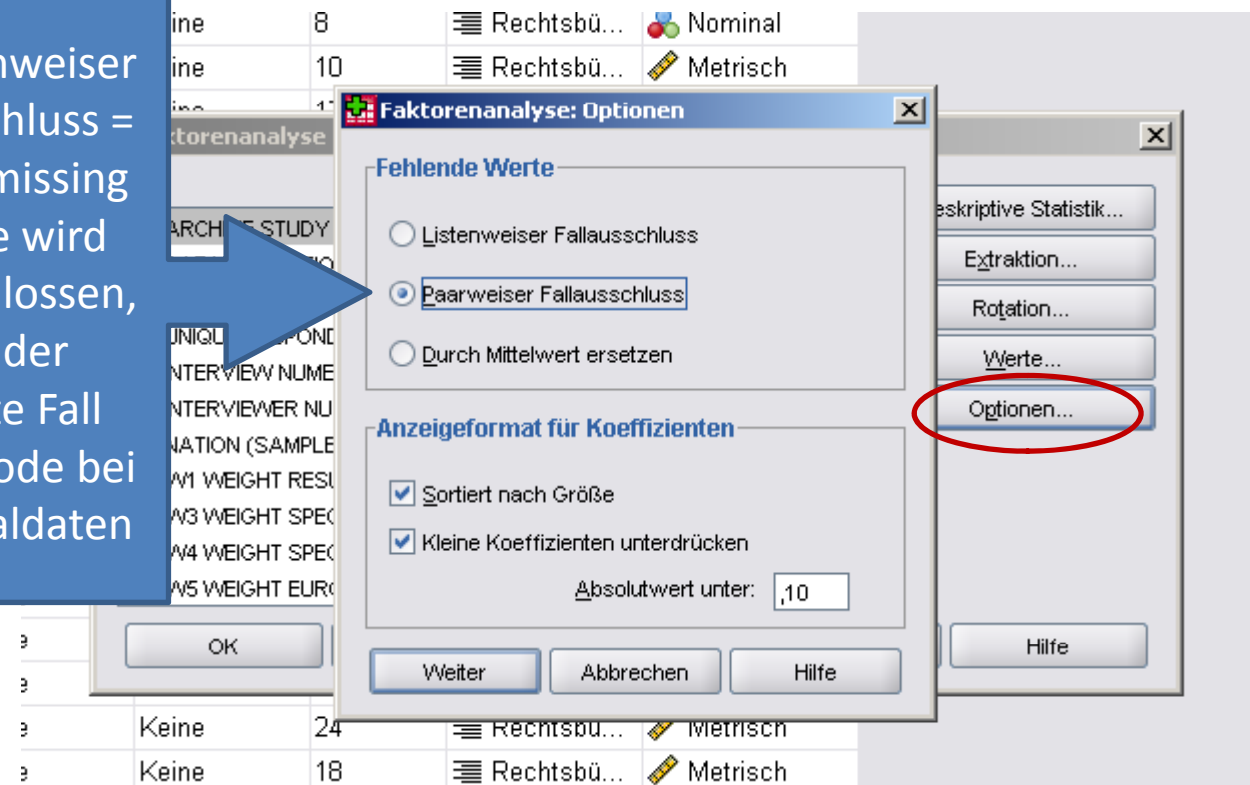


Ziel:
Unabhängigkeit
der Faktoren/
Dimensionen

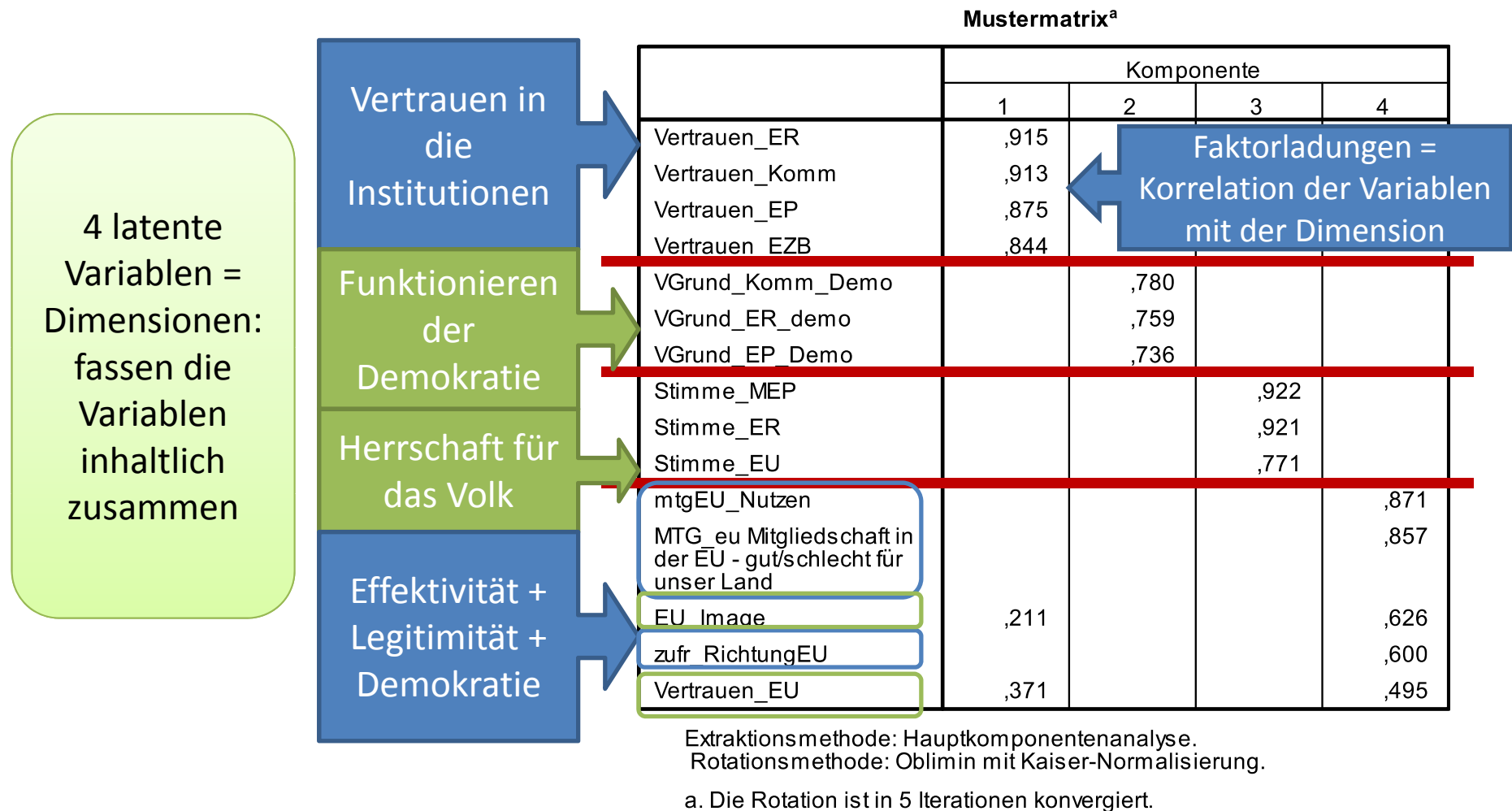
Explorative Faktorenanalyse – 7 Interpretation

Ziel: Erleichterung der Interpretation

Variablenweiser
Fallausschluss =
nur die missing
variable wird
ausgeschlossen,
nicht der
gesamte Fall
→ Methode bei
Individualdaten



Explorative Faktorenanalyse – 7 Interpretation



Explorative Faktorenanalyse - Ergebnis

- explorative Faktorenanalyse kann keine eindeutige **technische** Trennung zwischen Demokratie- und Legitimitätsdefizit herbeiführen
- Vier Faktoren, höhere Varianzaufklärung → komplexe theoretische Erklärung

Fallreduktion – Clusteranalyse I

- Voraussetzungen:
- Metrisches Skalenniveau
- Strukturiertheit der Daten

Fallreduktion – Clusteranalyse II

- Ziel und Zweck: Fälle mit ähnlichen Variablenausprägungen zu Gruppen zusammenfassen → Fallreduktion mittel Ähnlichkeitsmaßen

Fallreduktion – Clusteranalyse III

1. Konkretisierung der Problemstellung
2. Bestimmung der zu klassifizierenden Objekte/Fälle
3. Auswahl der Variablen
4. Festlegung des Ähnlichkeits-/Distanzmaßes
5. Auswahl des Algorithmus zur Gruppierung
6. Bestimmung der Gruppen(Cluster)zahl
7. Durchführung des Gruppierungsvorgangs
8. Analyse und Interpretation der Ergebnisse

Clusteranalyse – 1 Problemstellung

- Sind in allen Europäischen Ländern gleiche Einstellungsstrukturen zu finden oder variiert das Empfinden eines Demokratie- bzw. Legitimitätsdefizits zwischen den EU-Mitgliedsländern?

Clusteranalyse – 2 Objekte/Fälle

- Alle Mitgliedsstaaten der EU und Beitrittsanwärter sowie Territorien der Mitglieder
(Nordirland, Deutschland West-Ost)
→ 29 Fälle

Clusteranalyse – 3 Variablenauswahl



Kennzeichnungsvariablen

Mustermatrix^a

	Komponente			
	1	2	3	4
Vertrauen_ER	,915			
Vertrauen_Komm	,913			
Vertrauen_EP	,875			
Vertrauen_EZB	,844			
VGrund_Komm_Demo		,780		
VGrund_ER_demo		,759		
VGrund_EP_Demo		,736		
Stimme_MEP			,922	
Stimme_ER			,921	
Stimme_EU			,771	
mtgEU_Nutzen				,871
MTG_eu Mitgliedschaft in der EU - gut/schlecht für unser Land				,857
EU_Image	,211			,626
zufr_RichtungEU				,600
Vertrauen_EU	,371			,495

Extraktionsmethode: Hauptkomponentenanalyse.
Rotationsmethode: Oblimin mit Kaiser-Normalisierung.

a. Die Rotation ist in 5 Iterationen konvergiert.

Clusteranalyse – 4 Ähnlichkeits-/ Distanzmaß

Quadrierte Euklidische Distanz

- Aus Differenzwerten bei jeder Variable für ein Fallpaar

$$d(x, y) = \|x - y\|_2 = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- Addition der Differenzwerte \Rightarrow Quadratwurzel aus der Summe

Quadrierung

\Rightarrow große Differenzwerte werden stärker berücksichtigt als kleine

Basis für Ward-Methode (Gruppierungsalgorithmus)

Clusteranalyse – 5 Algorithmus

Hierarchische
Verfahren

agglomerativ

Ausgangspunkt:
feinste Partition →
Anzahl der Gruppen
= Anzahl der Fälle
→ Zusammenfassung

Ward-Methode

divisiv

Ausgangspunkt:
größte
Partition →
Anzahl der
Gruppen = 1
→ Aufteilung

Partitionierende
Verfahren

Ausgangspunkt:
Struktur, die
durch Anpassung
und Austausch
zwischen den
Gruppen
optimiert wird

Clusteranalyse – 5 Algorithmen

EB69_4_aggr.sav [DatenSet1] - SPSS Statistics Daten-Editor

Datei Bearbeiten Ansicht Daten Transformieren **Analysieren** Diagramme Extras Add-Ons Fenster Hilfe

1 : country 1,0

	country	Vertrauen_ER_pg	Stimme_MEP_pgt
1	1		
2	2		
3	3		
4	4		
5	5		
6	6		
7	7		
8	8		
9	9		
10	10		
11	11		
12	12		
13	13		
14	14		
15	15		
16	16		
17	17		
18	18		
19	19		

Analysieren

- Berichte
- Deskriptive Statistiken
- Tabellen
- Mittelwerte vergleichen
- Allgemeines lineares Modell
- Verallgemeinerte lineare Modelle
- Gemischte Modelle
- Korrelation
- Regression
- Loglinear
- Klassifizieren**
 - Iwo-Step-Clusteranalyse...
 - Clusterzentrenanalyse...
 - Hierarchische Cluster...**
 - Diskriminanzanalyse
 - Nächstgelegener Nachbar...
- Dimensionsreduzierung
- Skalierung
- Nichtparametrische Tests
- Vorhersage
- Überleben
- Mehrfachantworten
- Analyse fehlender Werte...
- Multiple Imputation
- Qualitätskontrolle
- ROC-Kurve...

Clusteranalyse – 5 Algorithmus

hierarchisch-agglomerative Clusteranalyse

1. 1 Fall = 1 Cluster/Gruppe \Rightarrow 29 Gruppen
2. Quadrierte Euklidische Distanz für alle Fälle
3. Fusion der 2 Cluster mit der geringsten Distanz
4. Berechnung der Abstände zwischen den neuen und den übrigen Clustern
5. Fusion zu neuen Clustern

Clusteranalyse – 5 Algorithmus

Ward-Methode

1. Quadrierte Euklidische Distanzen
2. „Vereinige diejenigen Cluster, die die Fehlerquadratsumme am wenigsten erhöhen.“
 - Zusammenfassung von Gruppen, die ein vorgegebenes **Heterogenitätsmaß** am wenigsten vergrößern; d.h. der Fälle mit der kleinsten quadrierten Euklidischen Distanz
 - Fälle/Gruppen zusammenfassen, die die **Varianz** *innerhalb* einer Gruppe möglichst wenig erhöhen
 - ⇒ Varianzkriterium oder Fehlerquadratsumme

Clusteranalyse – 5 Algorithmus

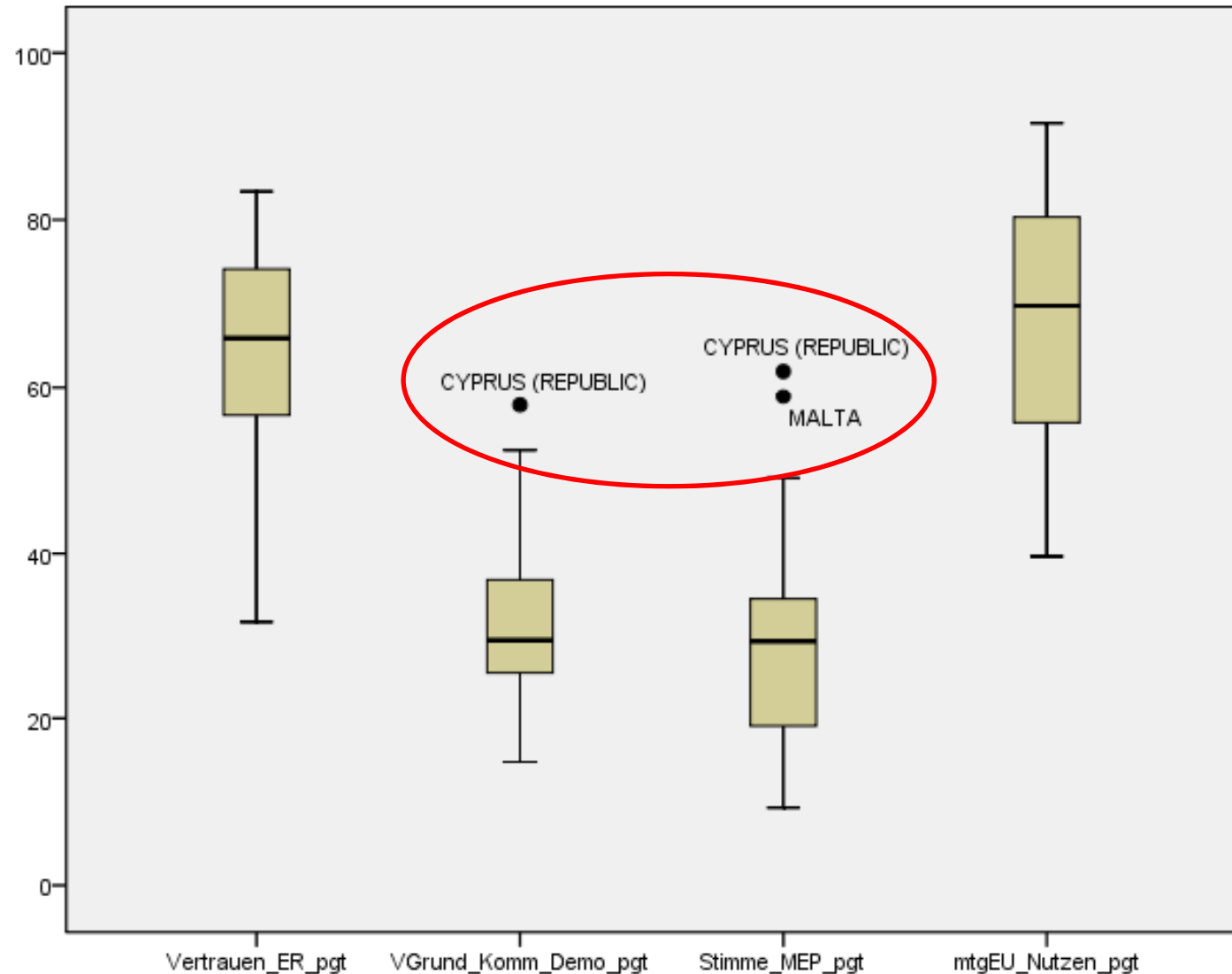
Ward-Methode - Voraussetzungen

- Distanz der Ländergruppen
 - Metrisches Skalenniveau
 - Keine Ausreißer in der Fallauswahl
 - Variablen sind weitgehend unkorreliert
 - Etwa gleiche Ausdehnung (Spannweite) der Gruppen
- Test: Boxplott und Korrelation

Clusteranalyse – 5 Algorithmus

Boxplott:

- 2 Ausreißer
- ähnliche Spannweiten



Clusteranalyse – 5 Algorithmus

Korreation:

Variable mtgEU_nutzen kann/muss entfernt werden

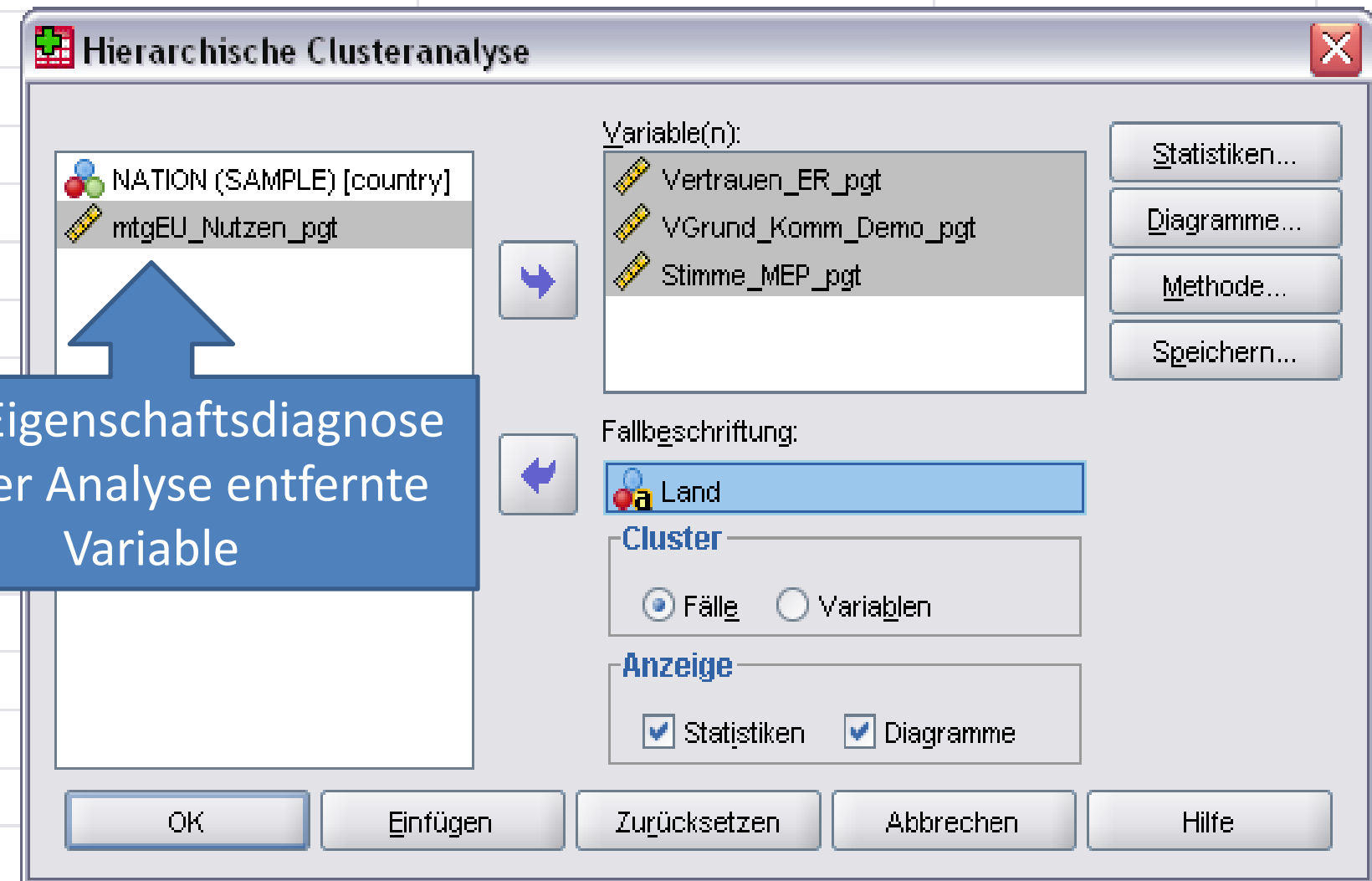
Korrelationen

		Vertrauen_ ER_pgt	VGrund_ Komm_ Demo_pgt	Stimme_ MEP_pgt	mtgEU_ Nutzen_pgt
Vertrauen_ER_pgt	Korrelation nach Pearson	1	,277	,407*	,746**
	Signifikanz (2-seitig)		,146	,029	,000
	N	33	29	29	29
VGrund_Komm_Demo_pgt	Korrelation nach Pearson	,277	1	,362	,184
	Signifikanz (2-seitig)	,146		,054	,339
	N	29	29	29	29
Stimme_MEP_pgt	Korrelation nach Pearson	,407*	,362	1	,357
	Signifikanz (2-seitig)	,029	,054		,058
	N	29	29	29	29
mtgEU_Nutzen_pgt	Korrelation nach Pearson	,746**	,184	,357	1
	Signifikanz (2-seitig)	,000	,339	,058	
	N	29	29	29	29

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Clusteranalyse – 5 Algorithmus



Clusteranalyse – 5 Algorithmus

Hierarchische Clusteranalyse: Methode

Cluster-Methode: Ward-Methode

Messniveau

☒ Intervall: Quadrierter Euklidischer Abstand
Exponent: 2 Wurzel: 2

☐ Häufigkeiten: Chi-Quadrat-Maß

☐ Binär: Quadrierter Euklidischer Abstand
Vorhanden: 1 Nicht vorhanden: 0

Werte transformieren

Standardisieren: Keine
☒ Nach Variablen
☐ Nach Fällen:

Maße transformieren

☐ Absolutwerte
☐ Vorzeichen ändern
☐ Auf Bereich 0-1 skalieren

Weiter Abbrechen Hilfe

Statistiken...
Diagramme...
Methode...
Speichern...
Hilfe

19,1
44,5
30,0
12,9

Clusteranalyse – 6 Clusterzahl

- Keine theoretisch oder sachlogisch begründeten Vorstellungen zur Gruppierung der Fälle \Rightarrow Aufdecken von Gruppierungen der Fälle (explorativ)
- Bestimmung der Clusterzahl durch statistische Kriterien \Rightarrow Entwicklung des Heterogenitätsmaßes

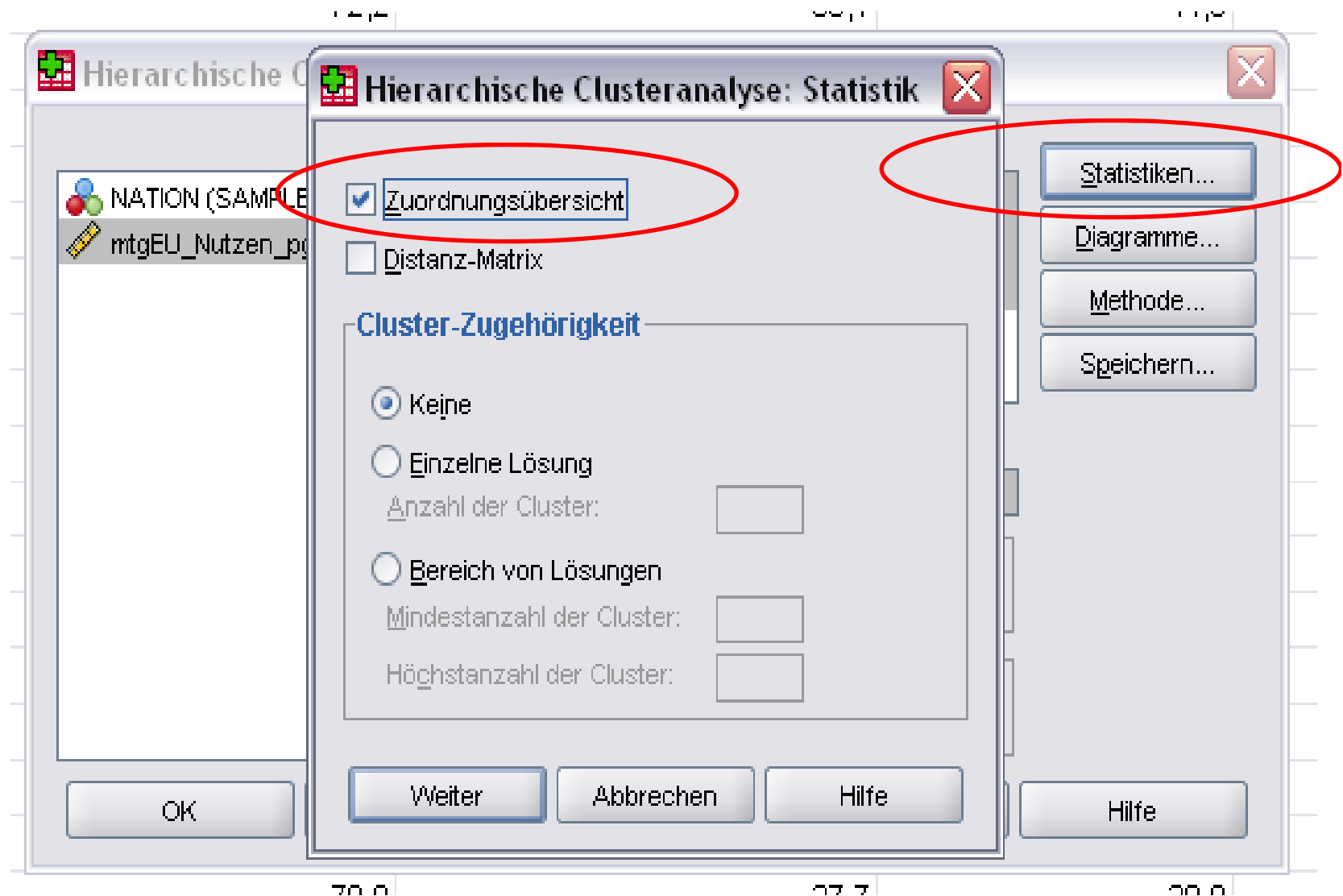
Anzahl der Cluster?

1. Tabelle „Zuordnungsübersicht“ \rightarrow Spalte „Koeffizienten“
 \rightarrow Elbow-Kriterium
2. Dendrogramm

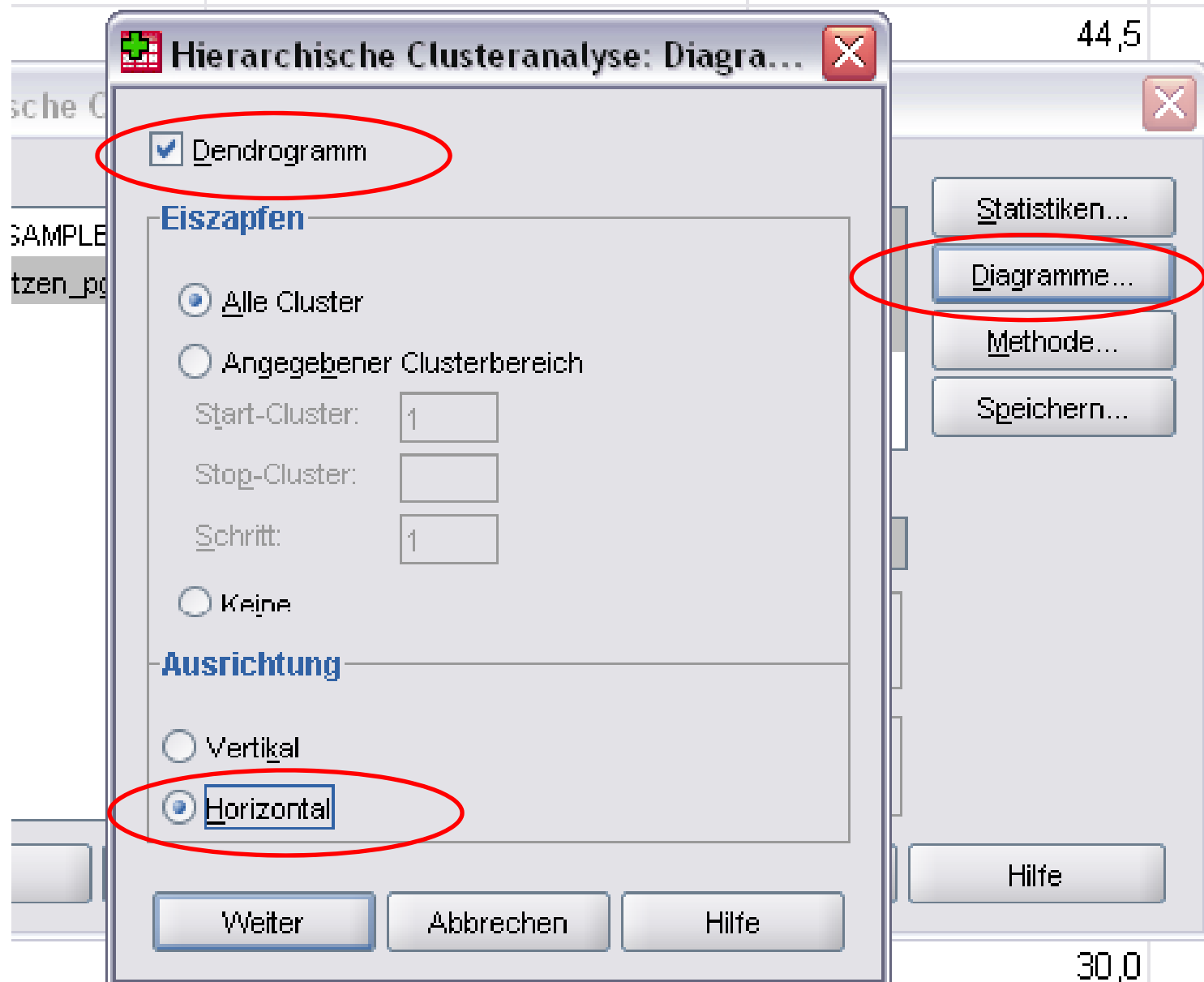
Grundproblem:

Entscheidung zwischen Homogenität in den Clustern und Handhabbarkeit der Clusterlösung

Clusteranalyse – 6 Clusterzahl



Clusteranalyse – 6 Clusterzahl



Zuordnungsübersicht

Fusionierungs-
schritt

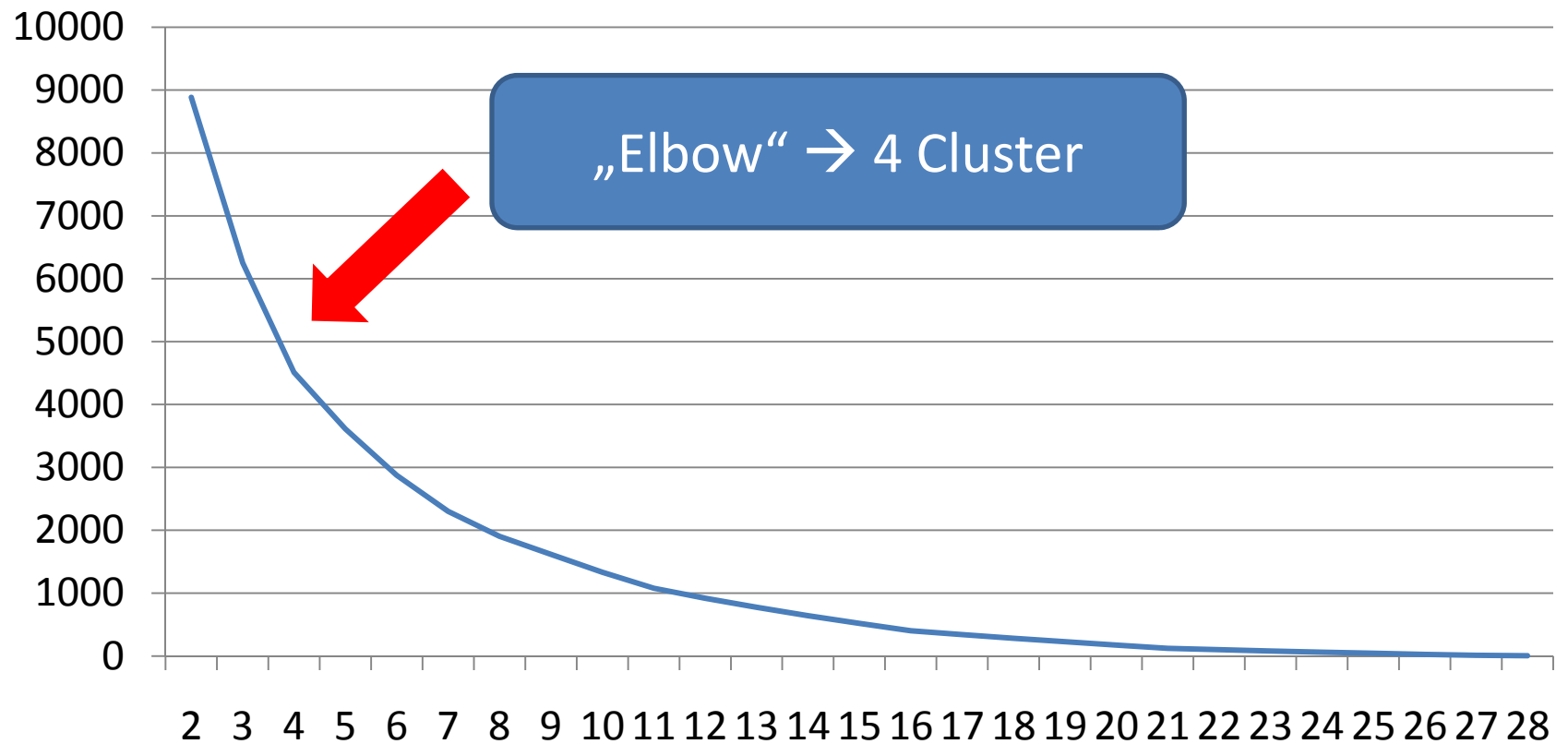
Nummer der
fusionierten
Objekte

Wert des
Heterogenitäts-
maßes

Schritt	Zusammengeführte Cluster		Koeffizienten	Erstes Vorkommen des Clusters		Nächster Schritt
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	7	8	4,711	0	0	5
2	10	19	12,517	0	0	19
3	4	13	27,292	0	0	17
4	26	28	43,694	0	0	12
5	3	7	60,987	0	1	18
6	25	27	80,611	0	0	7
7	12	25	100,815	0	6	15
8	14	20	123,062	0	0	19
9	16	17	174,716	0	0	20
10	9	29	228,458	0	0	22
11	2	15	282,760	0	0	25
12	21	26	340,397	0	4	14
13	21	23	400,608	0	0	18
14	21	23	518,763	12	0	22
15	11	12	641,609	0	7	21
16	18	24	774,692	0	0	27
17	4	22	919,333	3	0	20
18	1	3	1078,262	13	5	23
19	10	14	1333,067	2	8	24
20	4	16	1615,692	17	9	28
21	6	11	1903,484	0	15	23
22	9	21	2301,479	10	14	25
23	1	6	2872,649	18	21	24
24	1	10	3612,226	23	19	26
25	2	9	4511,917	11	22	26
26	1	2	6252,953	24	25	27
27	1	18	8883,937	26	16	28
28	1	4	12533,469	27	20	0

Clusteranalyse – 6 Clusterzahl

Fehlerquadratsumme

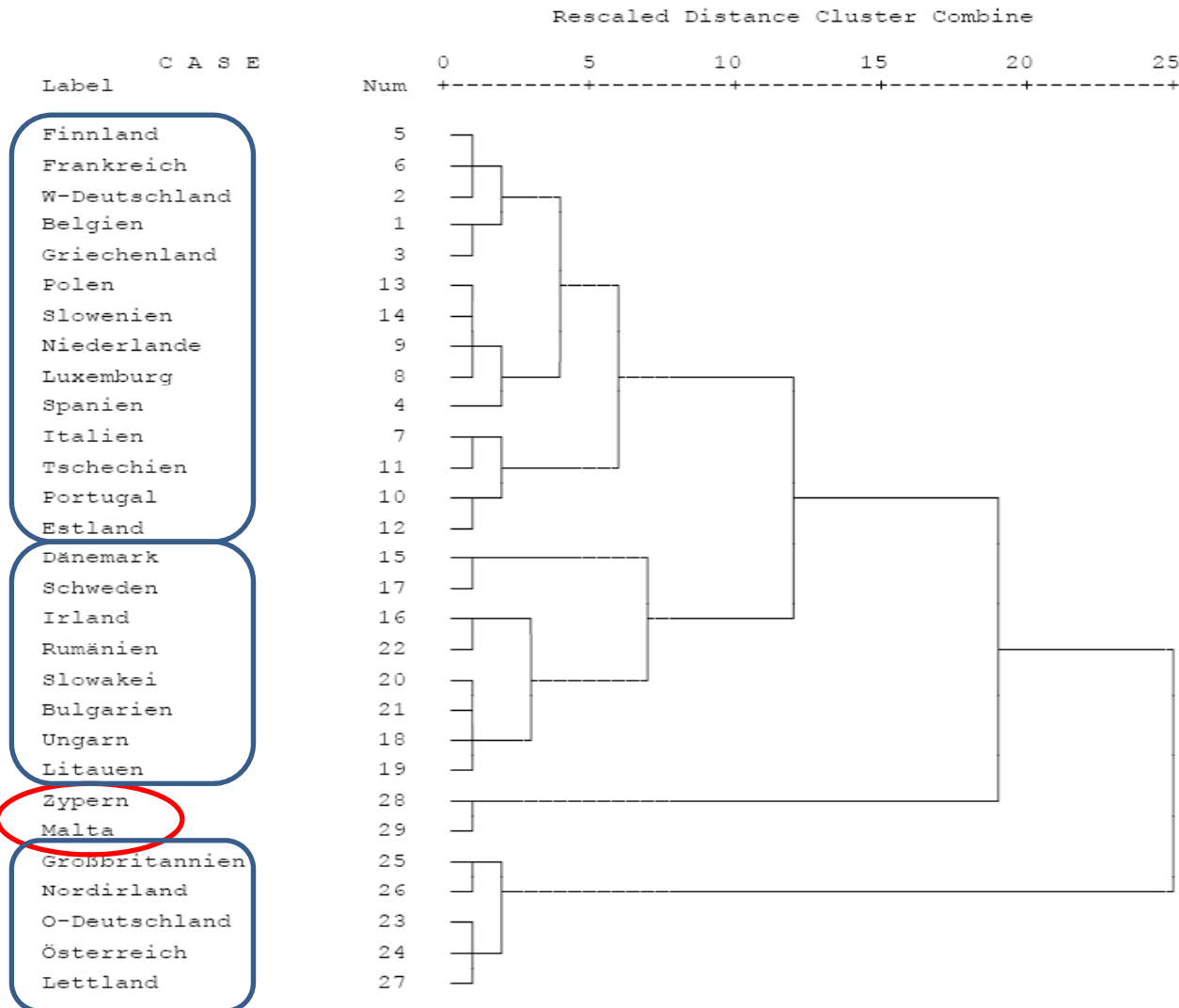


Clusteranzahl

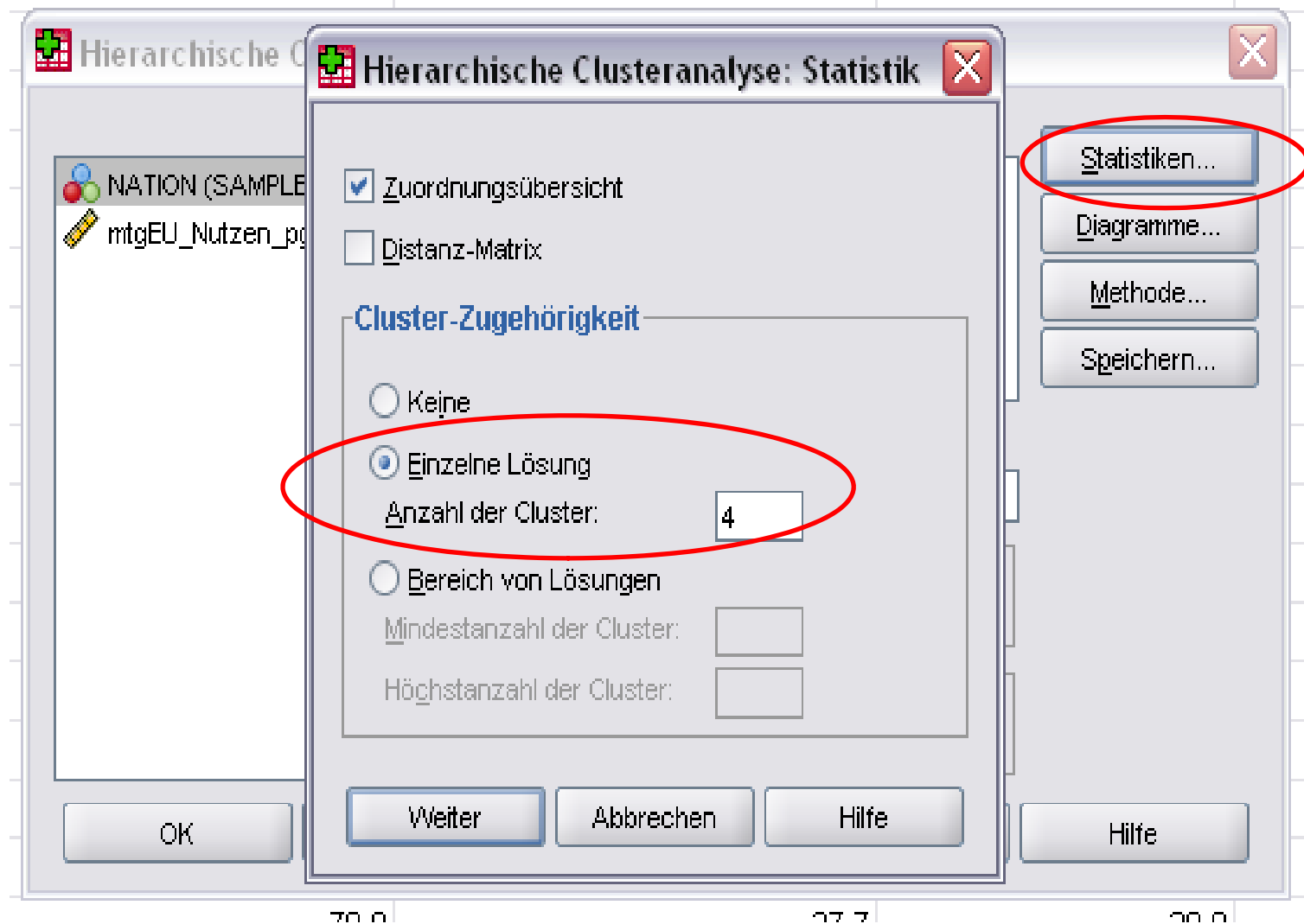
Clusteranalyse – 6 Clusterzahl

Dendrogram using Ward Method

Dendrogramm



Clusteranalyse – 6 Clusterzahl

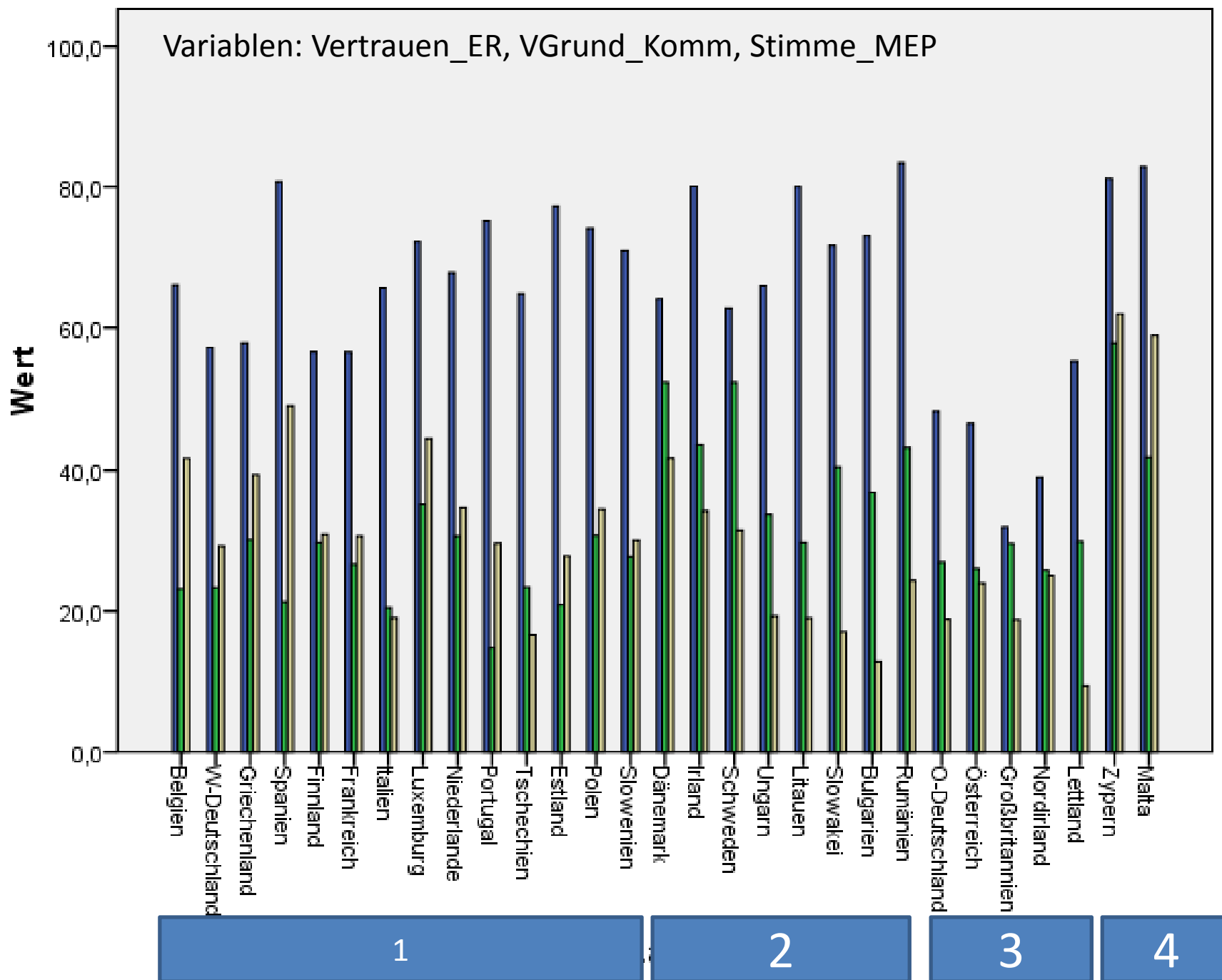


Cluster-Zugehörigkeit

Fall	4 Cluster
1:Belgien	1
2:Dänemark	2
3:W-Deutschland	1
4:O-Deutschland	3
5:Griechenland	1
6:Spanien	1
7:Finnland	1
8:Frankreich	1
9:Irland	2
10:Italien	1
11:Luxemburg	1
12:Niederlande	1
13:Österreich	3
14:Portugal	1
15:Schweden	2
16:Großbritannien	3
17:Nordirland	3
18:Zypern	4
19:Tschechien	1
20:Estland	1
21:Ungarn	2
22:Lettland	3
23:Litauen	2
24:Malta	4
25:Polen	1
26:Slowakei	2
27:Slowenien	1
28:Bulgarien	2
29:Rumänien	2

Clusteranalyse – 7 Interpretation

Land	Cluster	Land	Cluster
Belgien	1	Dänemark	2
W-D'land	1	Irland	2
Griechenland	1	Schweden	2
Spanien	1	Ungarn	2
Finnland	1	Litauen	2
Frankreich	1	Slowakei	2
Italien	1	Bulgarien	2
Luxemburg	1	Rumänien	2
Niederlande	1	O-D'land	3
Portugal	1	Österreich	3
Tschechien	1	GB	3
Estland	1	N'Irland	3
Polen	1	Lettland	3
Slowenien	1	Zypern	4
		Malta	4



Clusteranalyse – Ergebnis: Mittelwertvergleich

Bericht

Cluster		Vertrauen_ ER_pgt	VGrund_ Komm_ Demo_pgt	Stimme_ MEP_pgt
1	Mittelwert	67,304	25,568	32,671
	N	14	14	14
	Standardabweichung	8,1249	5,4305	8,9686
2	Mittelwert	72,601	41,534	24,989
	N	8	8	8
	Standardabweichung	7,9702	8,1934	9,8802
3	Mittelwert	44,229	27,626	19,205
	N	5	5	5
	Standardabweichung	9,0238	1,9170	6,1903
4	Mittelwert	82,002	49,845	60,386
	N	2	2	2
	Standardabweichung	1,1965	11,2679	2,1644
Insgesamt	Mittelwert	65,800	32,002	30,141
	N	29	29	29
	Standardabweichung	13,2060	10,4093	12,8403

Clusteranalyse – Ergebnis: Gruppenvergleich

Ländergruppe Durchschnittliche Zustimmung in %	Vertrauen	Funktionieren der Demokratie	Herrschaft für das Volk
Malta, Zypern ④	82	50	60
Dänemark, Irland, Schweden, Ungarn, Litauen, Slowakei, Bulgarien, Rumänien ②	73	42	25
Belgien, W-Deutschland, Griechenland, Spanien, Finnland, Frankreich, Italien, Luxemburg, Niederlande, Portugal, Tschechien, Estland, Polen, Slowenien ①	67	26	33
O-Deutschland, Österreich, GB, Nordirland, Lettland ③	44	28	19

Ergebnis

- Faktorenanalyse:
Die Einstellungen der EU-Bürger können zu den Dimensionen „Demokratie“ und „Legitimität“ zusammengefasst werden
→ Trennung von Legitimitäts- und Demokratiedefizit prinzipiell möglich
- Clusteranalyse:
Es gibt Ländergruppen mit voneinander unterscheidbaren Einstellungen:
EU-Euphoriker, kritische Europäer, kritische Demokraten und Euroskeptiker

Literaturempfehlung

- Klaus Backhaus u.a. 2003: Multivariate Analysemethoden. Berlin u.a. Springer Verlag
- Christian FG Schendera 2010: Clusteranalyse mit SPSS. München Oldenbourg Wissenschaftsverlag