

Time-Series Cross-Sectional Data: Möglichkeiten und Probleme

Regressionsmodelle

Sommersemester 2009

TSCS (Beck-Katz-Ansatz): Möglichkeiten und Grenzen

Beck/Katz 1995: „What to do (and not to do) with Time-Series Cross-Section Data“. APSR Vol. 89, p. 634-647.

Wilson/Butler 2007: „A Lot More to Do. The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications“. Political Analysis Vol. 15, p. 101-123.

Wiederholung

$N < 50$: Ein Problem und ein Lösungsvorschlag

Die Beck-Katz-Methode

Erhebungsdesign und Schätzverfahren

Die Technik

Grenzen und Alternativen

Fazit



N. Beck



J. Katz

Wie war das mit den UN-Missionen?

- ▶ Aufgabe
 - ▶ Vorzeichen der Koeffizienten?
 - ▶ Form der Hazard-Funktion?

Wie war das mit den UN-Missionen?

- ▶ Aufgabe
 - ▶ Vorzeichen der Koeffizienten?
 - ▶ Form der Hazard-Funktion?
- ▶ Civil War hat negatives Vorzeichen → dauert weniger lang als internationalised conflict
- ▶ Interstate Conflict hat positives Vorzeichen → dauert länger

Wie war das mit den UN-Missionen?

- ▶ Aufgabe
 - ▶ Vorzeichen der Koeffizienten?
 - ▶ Form der Hazard-Funktion?
- ▶ Civil War hat negatives Vorzeichen → dauert weniger lang als internationalised conflict
- ▶ Interstate Conflict hat positives Vorzeichen → dauert länger
- ▶ Hazard *muß* flach sein, da Exponentialmodell (konstanter Hazard)

Warum Time-Series Cross-Sectional Analysis?

- ▶ Häufigste Untersuchungseinheit in der Vergleichenden Politikwissenschaft: Politische Systeme

Warum Time-Series Cross-Sectional Analysis?

- ▶ Häufigste Untersuchungseinheit in der Vergleichenden Politikwissenschaft: Politische Systeme
- ▶ Aber
 - ▶ 30 OECD-Länder
 - ▶ 15 „alte“ EU-Mitgliedsstaaten

Warum Time-Series Cross-Sectional Analysis?

- ▶ Häufigste Untersuchungseinheit in der Vergleichenden Politikwissenschaft: Politische Systeme
- ▶ Aber
 - ▶ 30 OECD-Länder
 - ▶ 15 „alte“ EU-Mitgliedsstaaten
 - ▶ 7 ex-jugoslawische Republiken, 6 EU-Gründerstaaten, 5 ständige Mitglieder im UN-Sicherheitsrat, 4 skandinavische Staaten, ...

Warum Time-Series Cross-Sectional Analysis?

- ▶ Häufigste Untersuchungseinheit in der Vergleichenden Politikwissenschaft: Politische Systeme
- ▶ Aber
 - ▶ 30 OECD-Länder
 - ▶ 15 „alte“ EU-Mitgliedsstaaten
 - ▶ 7 ex-jugoslawische Republiken, 6 EU-Gründerstaaten, 5 ständige Mitglieder im UN-Sicherheitsrat, 4 skandinavische Staaten, ...
- ▶ (Viel) zu wenige Fälle, um statistische Modelle (sinnvoll) einsetzen zu können

Beck/Katz: Anwendungsbeispiele

Beck/Katz 1995

1. „Time-Series Cross-Sectional“ Design



Beck/Katz: Anwendungsbeispiele

Beck/Katz 1995

1. „Time-Series Cross-Sectional“ Design

- ▶ Vergleich über Raum *und* Zeit
- ▶ Vervielfacht die Zahl der Beobachtungen



Beck/Katz: Anwendungsbeispiele

Beck/Katz 1995

1. „Time-Series Cross-Sectional“ Design
 - ▶ Vergleich über Raum *und* Zeit
 - ▶ Vervielfacht die Zahl der Beobachtungen
2. Vereinfachtes Schätzverfahren
 - ▶ Einfache Regression („OLS“) zu optimistisch
 - ▶ Standardschätzverfahren (GLS) für politikwissenschaftliche Daten oft ungeeignet
 - ▶ „Kochrezept“: OLS + korrigierte Standardfehler



Beck/Katz: Anwendungsbeispiele

- ▶ Sind heterogene Gesellschaften repressiver? (Walker 2007)
- ▶ Reduziert soziale Ungleichheit die Wahlbeteiligung? (Lister 2007)
- ▶ Reduzieren freie Medien die Wahrscheinlichkeit von Kriegen? (Choi/James 2007)

Beck/Katz: Anwendungsbeispiele

- ▶ Sind heterogene Gesellschaften repressiver? (Walker 2007)
- ▶ Reduziert soziale Ungleichheit die Wahlbeteiligung? (Lister 2007)
- ▶ Reduzieren freie Medien die Wahrscheinlichkeit von Kriegen? (Choi/James 2007)
- ▶ Einfluß des israelisch-arabischen Konflikts auf das Repressionsniveau innerhalb anderer Staaten im Nahen Osten? (Lebovic/Thompson 2006)
- ▶ Einfluß von „third parties“ auf die Verwaltungen US-amerikanischer Bundesstaaten (Kelleher/Yackee 2006)

Was ist ein Time-Series Cross-Sectional Design?

Beobachtungen: $N = 4$

Land = 1, t=1

Land = 2, t=1

Land = ..., t=1

Land = N, t=1

Was ist ein Time-Series Cross-Sectional Design?

Beobachtungen: $N \times T = 16$

Land = 1, t=1

Land = 2, t=1

Land = ..., t=1

Land = N, t=1

Land = 1, t=2

Land = 2, t=2

Land = ..., t=2

Land = N, t=2

Land = 1, t = ...

Land = 2, t = ...

Land = ..., t = ...

Land = N, t = ...

Land = 1, t=T

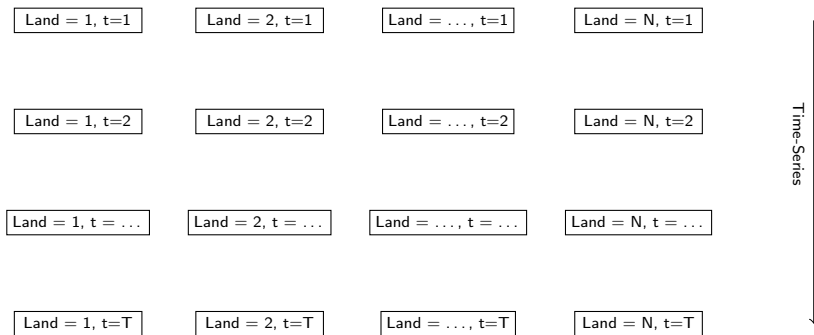
Land = 2, t=T

Land = ..., t=T

Land = N, t=T

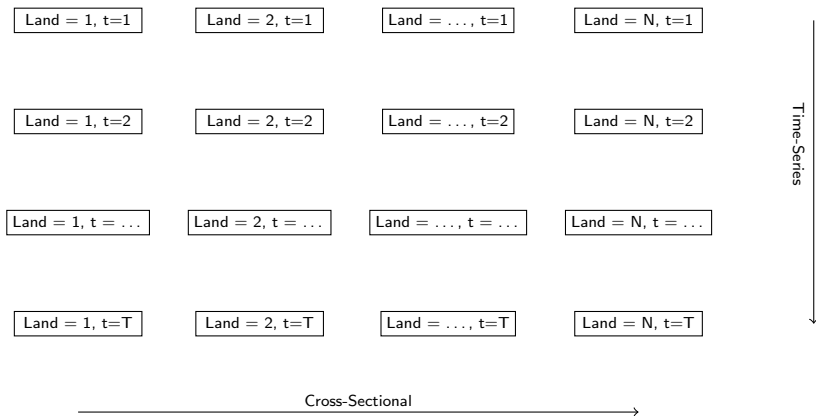
Was ist ein Time-Series Cross-Sectional Design?

Beobachtungen: $N \times T = 16$



Was ist ein Time-Series Cross-Sectional Design?

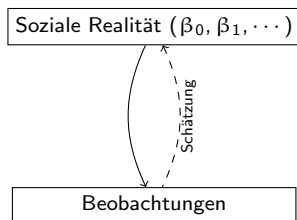
Beobachtungen: $N \times T$



„Panel“ mit wenigen Teilnehmern (Ländern) und sehr vielen Zeitpunkten

Was ist und wozu braucht man ein Schätzverfahren?

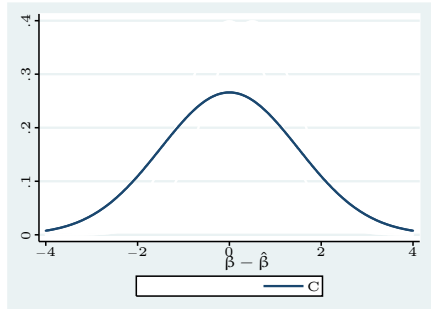
- ▶ Grundannahme: politische Realität läßt sich durch ein mathematisches Modell beschreiben
- ▶ $y = \beta_0 + \beta_1 x_1 + \dots + \epsilon$
- ▶ Z.B. Regression = Konstante + unabhängige Variablen + zufällige Einflüsse
- ▶ Wahre Parameter sind unbekannt
→ *Schätzung* der Parameter $\hat{\beta}_0, \hat{\beta}_1$
- ▶ Eine Konsequenz der zufälligen Einflüsse



Was ist und wozu braucht man ein Schätzverfahren?

Konsequenzen der zufälligen Einflüsse:

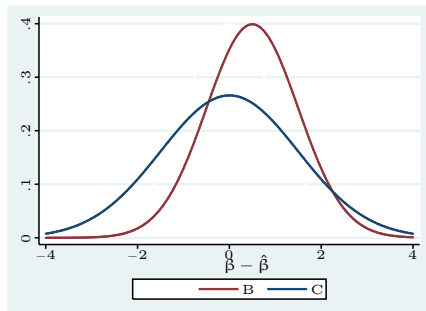
- ▶ Schätzungen sind probabilistisch
- ▶ Schätzungen für die Parameter ($\hat{\beta}$) und Schätzungen für deren zufällige Streuung um den wahren Wert (Standardfehler, $\sigma_{\hat{\beta}}$)
- ▶ Vergleich von Schätzverfahren
- ▶ Optimales Schätzverfahren
 - ▶ Unverzerrt
 - ▶ Effizient



Was ist und wozu braucht man ein Schätzverfahren?

Konsequenzen der zufälligen Einflüsse:

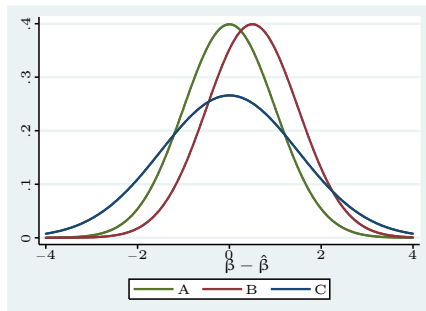
- ▶ Schätzungen sind probabilistisch
- ▶ Schätzungen für die Parameter ($\hat{\beta}$) und Schätzungen für deren zufällige Streuung um den wahren Wert (Standardfehler, $\sigma_{\hat{\beta}}$)
- ▶ Vergleich von Schätzverfahren
- ▶ Optimales Schätzverfahren
 - ▶ Unverzerrt
 - ▶ Effizient



Was ist und wozu braucht man ein Schätzverfahren?

Konsequenzen der zufälligen Einflüsse:

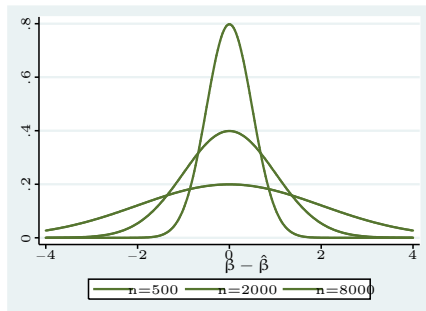
- ▶ Schätzungen sind probabilistisch
- ▶ Schätzungen für die Parameter ($\hat{\beta}$) und Schätzungen für deren zufällige Streuung um den wahren Wert (Standardfehler, $\sigma_{\hat{\beta}}$)
- ▶ Vergleich von Schätzverfahren
- ▶ Optimales Schätzverfahren
 - ▶ Unverzerrt
 - ▶ Effizient



Was ist und wozu braucht man ein Schätzverfahren?

Konsequenzen der zufälligen Einflüsse:

- ▶ Schätzungen sind probabilistisch
- ▶ Schätzungen für die Parameter ($\hat{\beta}$) und Schätzungen für deren zufällige Streuung um den wahren Wert (Standardfehler, $\sigma_{\hat{\beta}}$)
- ▶ Vergleich von Schätzverfahren
- ▶ Optimales Schätzverfahren
 - ▶ Unverzerrt
 - ▶ Effizient
 - ▶ Konsistent



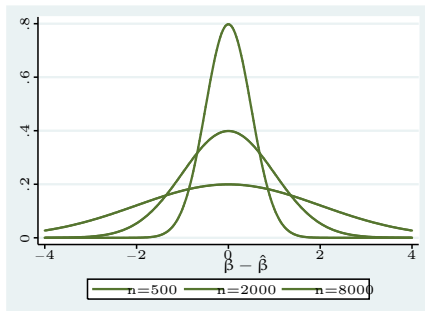
Was ist und wozu braucht man ein Schätzverfahren?

Konsequenzen der zufälligen Einflüsse:

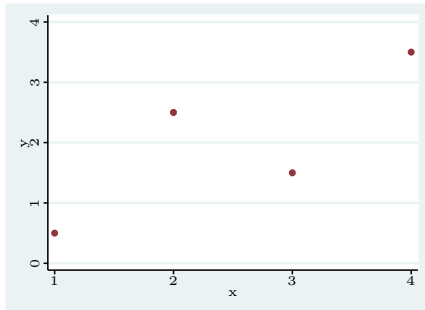
- ▶ Schätzungen sind probabilistisch
- ▶ Schätzungen für die Parameter ($\hat{\beta}$) und Schätzungen für deren zufällige Streuung um den wahren Wert (Standardfehler, $\sigma_{\hat{\beta}}$)
- ▶ Vergleich von Schätzverfahren
- ▶ Optimales Schätzverfahren
 - ▶ Unverzerrt
 - ▶ Effizient
 - ▶ Konsistent
- ▶ Plus realistische Einschätzung der Präzision (Standardfehler)

Standard in der Politikwissenschaft: Ordinary Least Squares = Methode der minimalen quadrierten Abweichungssummen

- ▶ Annahmen über die Struktur der zufälligen Einflüsse erforderlich

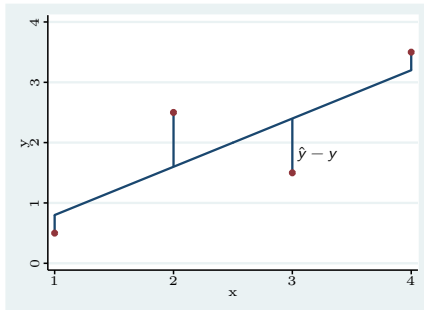


Wann ist OLS unverzerrt, effizient und konsistent?



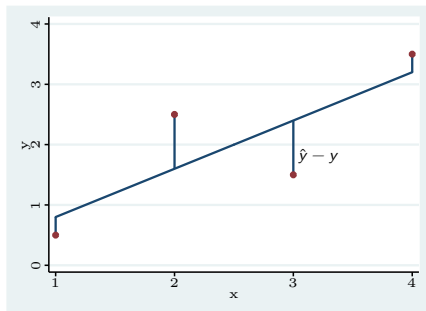
- ▶ OLS = Ordinary Least Squares

Wann ist OLS unverzerrt, effizient und konsistent?



- ▶ OLS = Ordinary Least Squares
- ▶ Minimiere quadrierte Distanz zwischen beobachteten und prognostizierten Werten

Wann ist OLS unverzerrt, effizient und konsistent?



- ▶ OLS = Ordinary Least Squares
- ▶ Minimiere quadrierte Distanz zwischen beobachteten und prognostizierten Werten

Optimal, wenn die zufälligen Einflüsse ϵ unabhängig voneinander sind (zufällige Auswahl unabhängiger Fälle)

OLS optimal wenn ...

Annahmen über ϵ , u. a.:

1. Für jede Beobachtung ist ϵ eine Ziehung aus einer Zufallsverteilung

OLS optimal wenn ...

Annahmen über ϵ , u. a.:

1. Für jede Beobachtung ist ϵ eine Ziehung aus einer Zufallsverteilung
2. Diese Verteilungen haben identische Varianz und sind voneinander unabhängig

OLS optimal wenn ...

Annahmen über ϵ , u. a.:

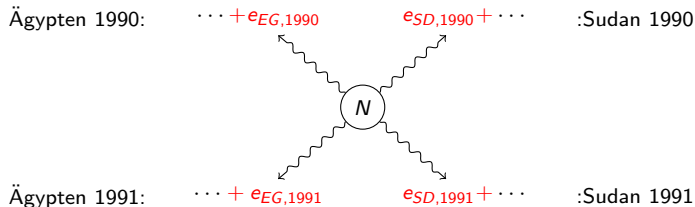
1. Für jede Beobachtung ist ϵ eine Ziehung aus einer Zufallsverteilung
2. Diese Verteilungen haben identische Varianz und sind voneinander unabhängig
3. Die Ziehungen/zufälligen Einflüsse sind voneinander unabhängig

OLS optimal wenn ...

Annahmen über ϵ , u. a.:

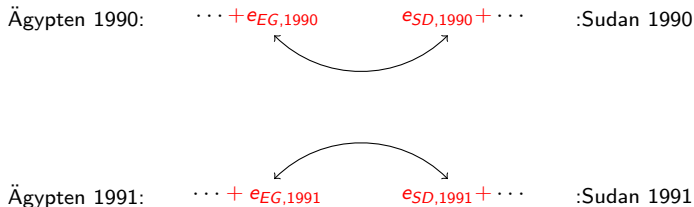
1. Für jede Beobachtung ist ϵ eine Ziehung aus einer Zufallsverteilung
2. Diese Verteilungen haben identische Varianz und sind voneinander unabhängig
3. Die Ziehungen/zufälligen Einflüsse sind voneinander unabhängig
4. Die Ziehungen sind von den systematischen Einflüssen (x) unabhängig

Warum macht das TSCS-Design Probleme?



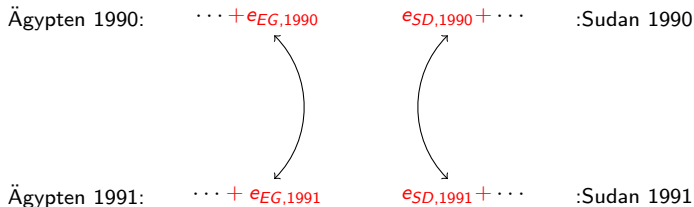
Warum macht das TSCS-Design Probleme?

1. Räumliche Korrelation



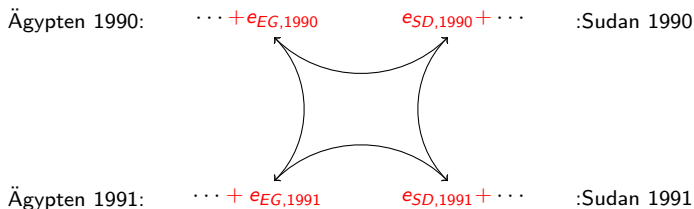
Warum macht das TSCS-Design Probleme?

1. Räumliche Korrelation
2. Zeitliche Korrelation (Autokorrelation)



Warum macht das TSCS-Design Probleme?

1. Räumliche Korrelation
2. Zeitliche Korrelation (Autokorrelation)
3. Unit-Effekte und länderspezifische Varianz



Warum macht das TSCS-Design Probleme?

1. Räumliche Korrelation
 2. Zeitliche Korrelation (Autokorrelation)
 3. Unit-Effekte und länderspezifische Varianz
- ▶ OLS immer noch unverzerrt und konsistent, aber nicht mehr effizient
 - ▶ Standardfehler viel zu optimistisch ($\approx 600\%$)

Warum macht das TSCS-Design Probleme?

1. Räumliche Korrelation
 2. Zeitliche Korrelation (Autokorrelation)
 3. Unit-Effekte und länderspezifische Varianz
- ▶ OLS immer noch unverzerrt und konsistent, aber nicht mehr effizient
 - ▶ Standardfehler viel zu optimistisch ($\approx 600\%$)
 - ▶ Ökonometrisches Standardverfahren: GLS

Warum macht das TSCS-Design Probleme?

1. Räumliche Korrelation
 2. Zeitliche Korrelation (Autokorrelation)
 3. Unit-Effekte und länderspezifische Varianz
- ▶ OLS immer noch unverzerrt und konsistent, aber nicht mehr effizient
 - ▶ **Standardfehler viel zu optimistisch ($\approx 600\%$)**
 - ▶ Ökonometrisches Standardverfahren: GLS
 - ▶ *Computersimulation*: GLS produziert ebenfalls zu optimistische Standardfehler wenn Zahl der Länder < 30 und die Zahl der Zeitpunkte nicht mindestens dreimal größer
 - ▶ Beck & Katz: „fix OLS“

(Unit-Effekte)

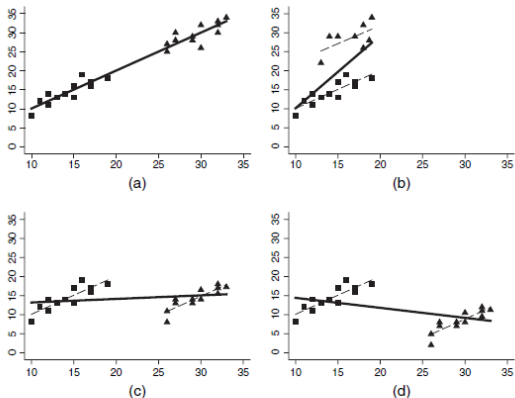


Fig. 1 The thick line in each panel is the estimated slope from the pooled regression. (a) Pooled regression correctly estimates slope; (b) pooled regression overestimates slope; (c) pooled regression underestimates slope; and (d) pooled regression estimates incorrect sign for slope.

Wie lautet das Argument von Beck und Katz?

$$e_{i,t} \sim \begin{pmatrix} \sigma_{1,1}^2 & \sigma_{2,1}^2 & \cdots & \sigma_{N,1}^2 \\ \sigma_{1,2}^2 & \sigma_{2,2}^2 & \cdots & \sigma_{N,2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,T}^2 & \sigma_{2,T}^2 & \cdots & \sigma_{N,T}^2 \end{pmatrix}$$

- ▶ GLS: Für jede der $(N \times T)$ Beobachtungen individuelle Varianz für $e_{i,t}$ + Kovarianzen zwischen diesen Verteilungen

Wie lautet das Argument von Beck und Katz?

$$e_{i,t} \sim \begin{pmatrix} \sigma_{1,1}^2 & \sigma_{2,1}^2 & \cdots & \sigma_{N,1}^2 \\ \sigma_{1,2}^2 & \sigma_{2,2}^2 & \cdots & \sigma_{N,2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,T}^2 & \sigma_{2,T}^2 & \cdots & \sigma_{N,T}^2 \end{pmatrix}$$

- ▶ GLS: Für jede der $(N \times T)$ Beobachtungen individuelle Varianz für $e_{i,t}$ + Kovarianzen zwischen diesen Verteilungen
- ▶ Für 20 Länder und 30 Jahre (=600 Beobachtungen) → 180 300 individuelle Varianzparameter → „infeasible“

Wie lautet das Argument von Beck und Katz?

$$e_{i,t} \sim \begin{pmatrix} \sigma_{1,1}^2 & \sigma_{2,1}^2 & \cdots & \sigma_{N,1}^2 \\ \sigma_{1,2}^2 & \sigma_{2,2}^2 & \cdots & \sigma_{N,2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,T}^2 & \sigma_{2,T}^2 & \cdots & \sigma_{N,T}^2 \end{pmatrix}$$

- ▶ GLS: Für jede der $(N \times T)$ Beobachtungen individuelle Varianz für $e_{i,t}$ + Kovarianzen zwischen diesen Verteilungen
- ▶ Für 20 Länder und 30 Jahre (=600 Beobachtungen) → 180 300 individuelle Varianzparameter → „infeasible“
- ▶ *Annahmen* über Struktur der zufälligen Einflüsse (räumliche/zeitliche Korrelation etc.) → „feasible“ GLS

Wie lautet das Argument von Beck und Katz?

$$e_{i,t} \sim \begin{pmatrix} \sigma_{1,1}^2 & \sigma_{2,1}^2 & \cdots & \sigma_{N,1}^2 \\ \sigma_{1,2}^2 & \sigma_{2,2}^2 & \cdots & \sigma_{N,2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,T}^2 & \sigma_{2,T}^2 & \cdots & \sigma_{N,T}^2 \end{pmatrix}$$

- ▶ GLS: Für jede der $(N \times T)$ Beobachtungen individuelle Varianz für $e_{i,t}$ + Kovarianzen zwischen diesen Verteilungen
- ▶ Für 20 Länder und 30 Jahre (=600 Beobachtungen) → 180 300 individuelle Varianzparameter → „infeasible“
- ▶ *Annahmen* über Struktur der zufälligen Einflüsse (räumliche/zeitliche Korrelation etc.) → „feasible“ GLS
- ▶ *Computersimulation*: (F)GLS produziert ebenfalls zu optimistische Standardfehler wenn Zahl der Länder < 30 und die Zahl der Zeitpunkte nicht mindestens dreimal größer

Wie lautet das Argument von Beck und Katz?

- ▶ OLS läßt sich „reparieren“:

Wie lautet das Argument von Beck und Katz?

- ▶ OLS läßt sich „reparieren“:
 1. Homogenitätsannahme: Identische Zusammenhänge in allen Ländern
 2. Kompensation der zeitlichen Korrelation von ϵ durch dynamische Spezifikation („LDV“, Regression von y auf Vorjahreswert)
 3. Abschließende Korrektur der Standardfehler durch eine Gewichtungszprozedur \rightarrow „Panel Corrected Standard Errors“

Wie lautet das Argument von Beck und Katz?

- ▶ OLS läßt sich „reparieren“:
 1. Homogenitätsannahme: Identische Zusammenhänge in allen Ländern
 2. Kompensation der zeitlichen Korrelation von ϵ durch dynamische Spezifikation („LDV“, Regression von y auf Vorjahreswert)
 3. Abschließende Korrektur der Standardfehler durch eine Gewichtungszprozedur \rightarrow „Panel Corrected Standard Errors“
- ▶ Computersimulation: *besser* als GLS

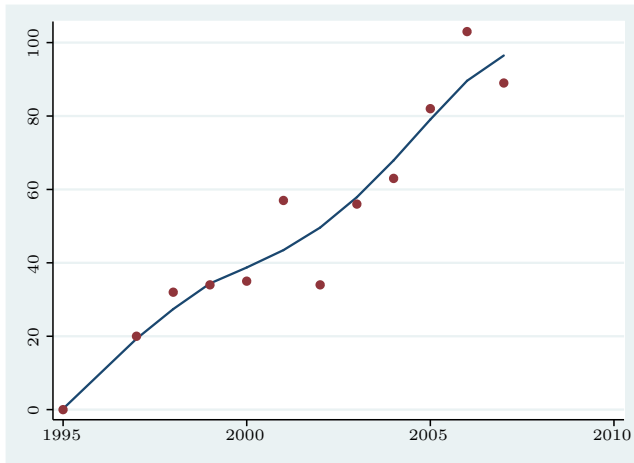
Wie lautet das Argument von Beck und Katz?

- ▶ OLS läßt sich „reparieren“:
 1. Homogenitätsannahme: Identische Zusammenhänge in allen Ländern
 2. Kompensation der zeitlichen Korrelation von ϵ durch dynamische Spezifikation („LDV“, Regression von y auf Vorjahreswert)
 3. Abschließende Korrektur der Standardfehler durch eine Gewichtungszprozedur \rightarrow „Panel Corrected Standard Errors“
- ▶ Computersimulation: *besser* als GLS
- ▶ „Beck-Katz-Ansatz“ seit Ende der 1990er Jahre dominanter Zugang zu TSCS-Analysen

Wie lautet das Argument von Beck und Katz?

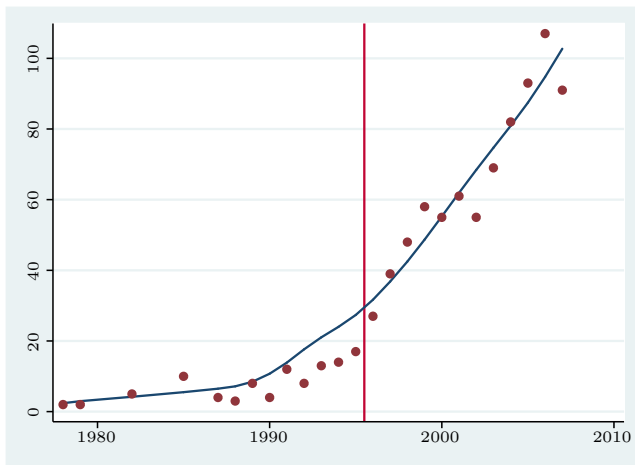
- ▶ OLS läßt sich „reparieren“:
 1. Homogenitätsannahme: Identische Zusammenhänge in allen Ländern
 2. Kompensation der zeitlichen Korrelation von ϵ durch dynamische Spezifikation („LDV“, Regression von y auf Vorjahreswert)
 3. Abschließende Korrektur der Standardfehler durch eine Gewichtungsspezifikation → „Panel Corrected Standard Errors“
- ▶ Computersimulation: *besser* als GLS
- ▶ „Beck-Katz-Ansatz“ seit Ende der 1990er Jahre dominanter Zugang zu TSCS-Analysen
- ▶ TSCS à la Beck/Katz als Standardverfahren in manchen Sub-Disziplinen (Political Economy)

Auswirkungen auf politikwissenschaftliche Forschung?



Zitationen von Beck/Katz 1995 im Social Science Citation Index, nur politikwissenschaftliche Zeitschriften

Effekt für die Autoren?



Zitationen von Neal Beck im Social Science Citation Index, nur politikwissenschaftliche Zeitschriften

Wie groß ist der Informationsgewinn durch TSCS?

20 Länder \times 30 Jahre \neq 600 unabhängige Beobachtungen

- ▶ Mehr Information durch Zeitdimension (jährliche, monatliche, wöchentliche Erhebung)?

Wie groß ist der Informationsgewinn durch TSCS?

20 Länder \times 30 Jahre \neq 600 unabhängige Beobachtungen

- ▶ Mehr Information durch Zeitdimension (jährliche, monatliche, wöchentliche Erhebung)?
- ▶ Häufig kaum Variation *innerhalb* der Länder (Föderalismus, Wahlsystem)
- ▶ Korrelationen zwischen Variablen (Föderalismus \times Wahlsystem in Westeuropa)

Ist die *Anwendung* des Beck/Katz-Ansatzes problematisch?

Wilson/Butler (2007): „A lot more to do“

- ▶ Untersuchen 195 publizierte Artikel (SSCI)

Ist die *Anwendung* des Beck/Katz-Ansatzes problematisch?

Wilson/Butler (2007): „A lot more to do“

- ▶ Untersuchen 195 publizierte Artikel (SSCI)
 1. Homogenität der Länder diskutiert?
 2. Werden andere dynamische Spezifikationen als LDV diskutiert?
Wird Autokorrelation getestet?
 3. (LDV führt bei Autokorrelation zu verzerrten/inkonsistenten Schätzungen)

Ist die *Anwendung* des Beck/Katz-Ansatzes problematisch?

Wilson/Butler (2007): „A lot more to do“

- ▶ Untersuchen 195 publizierte Artikel (SSCI)
 1. Homogenität der Länder diskutiert?
 2. Werden andere dynamische Spezifikationen als LDV diskutiert?
Wird Autokorrelation getestet?
 3. (LDV führt bei Autokorrelation zu verzerrten/inkonsistenten Schätzungen)
- ▶ Re-Analyse von acht Artikeln aus erstrangigen Zeitschriften
- ▶ Ergebnisse robust bei alternativen Spezifikationen?
- ▶ (Voraussetzungen des Beck/Katz Ansatzes (N, T))

Welche alternativen dynamischen Spezifikationen gibt es?

$$\text{Static model : } Y_t = \alpha + \beta_0 X_t + u_t, \quad (2)$$

$$\text{AR(1) model : } Y_t = \alpha + \beta_0 X_t + u_t; \quad u_t = \rho u_{t-1} + e_t, \quad (3)$$

$$\text{DL(1) model : } Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + u_t, \quad (4)$$

$$\text{LDV model : } Y_t = \alpha + \beta_0 X_t + \gamma_1 Y_{t-1} + u_t, \quad (5)$$

$$\text{ARDL(1, 1) model : } Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \gamma_1 Y_{t-1} + u_t, \quad (6)$$

$$\text{FD model : } Y_t - Y_{t-1} = \beta_0 (X_t - X_{t-1}) + u_t \quad (7)$$

(AR: autoregressive; DL: distributed lag; ARDL: autoregressive, distributed lag¹⁰;
FD: first difference).

Was zeigt die Untersuchung von Wilson/Butler?

- ▶ 61 Prozent setzen Homogenität ungeprüft voraus
- ▶ 75 Prozent ignorieren alternative dynamische Spezifikationen

Was zeigt die Untersuchung von Wilson/Butler?

- ▶ 61 Prozent setzen Homogenität ungeprüft voraus
- ▶ 75 Prozent ignorieren alternative dynamische Spezifikationen
- ▶ **5 Prozent** erfüllen minimale Voraussetzungen für problembewußte Analyse
- ▶ Researchers „are using B&K (1995) as a complete and authoritative guide to conducting TSCS analysis“

Was zeigt die Untersuchung von Wilson/Butler?

- ▶ 61 Prozent setzen Homogenität ungeprüft voraus
- ▶ 75 Prozent ignorieren alternative dynamische Spezifikationen
- ▶ **5 Prozent** erfüllen minimale Voraussetzungen für problembewußte Analyse
- ▶ Researchers „are using B&K (1995) as a complete and authoritative guide to conducting TSCS analysis“
- ▶ Re-Analyse von acht prominenten Artikeln mit
 - ▶ Fixed Unit Effects (individuelle Konstanten)
 - ▶ Alternativen dynamischen Spezifikationen

Was zeigt die Untersuchung von Wilson/Butler?

- ▶ 61 Prozent setzen Homogenität ungeprüft voraus
- ▶ 75 Prozent ignorieren alternative dynamische Spezifikationen
- ▶ **5 Prozent** erfüllen minimale Voraussetzungen für problembewußte Analyse
- ▶ Researchers „are using B&K (1995) as a complete and authoritative guide to conducting TSCS analysis“
- ▶ Re-Analyse von acht prominenten Artikeln mit
 - ▶ Fixed Unit Effects (individuelle Konstanten)
 - ▶ Alternativen dynamischen Spezifikationen
- ▶ **In sechs von acht Fällen Ergebnisse nicht robust**

Was zeigt die Untersuchung von Wilson/Butler?

- ▶ 61 Prozent setzen Homogenität ungeprüft voraus
- ▶ 75 Prozent ignorieren alternative dynamische Spezifikationen
- ▶ **5 Prozent** erfüllen minimale Voraussetzungen für problembewußte Analyse
- ▶ Researchers „are using B&K (1995) as a complete and authoritative guide to conducting TSCS analysis“
- ▶ Re-Analyse von acht prominenten Artikeln mit
 - ▶ Fixed Unit Effects (individuelle Konstanten)
 - ▶ Alternativen dynamischen Spezifikationen
- ▶ **In sechs von acht Fällen Ergebnisse nicht robust**
 - ▶ Signifikanz von Effekten
 - ▶ Größenordnung von Effekten
 - ▶ **Vorzeichen von Effekten**

Was folgt für uns daraus?

Wilson/Butler 2007

„It is more than a little ironic that even though B&K's analysis focused on the danger of using estimators without fully understanding their properties, so many in the profession applied the B&K method without paying any attention to the simple textbook issues“

Was folgt für uns daraus?

Wilson/Butler 2007

„It is more than a little ironic that even though B&K's analysis focused on the danger of using estimators without fully understanding their properties, so many in the profession applied the B&K method without paying any attention to the simple textbook issues“

- ▶ Informationsgehalt von komparativen Datensätzen auch bei Berücksichtigung der Zeitdimension beschränkt

Was folgt für uns daraus?

Wilson/Butler 2007

„It is more than a little ironic that even though B&K's analysis focused on the danger of using estimators without fully understanding their properties, so many in the profession applied the B&K method without paying any attention to the simple textbook issues“

- ▶ Informationsgehalt von komparativen Datensätzen auch bei Berücksichtigung der Zeitdimension beschränkt
- ▶ TSCS Analysis ist ein mächtiges Werkzeug ...
- ▶ ... das aber reflektiert angewendet werden muß

Was folgt für uns daraus?

Wilson/Butler 2007

„It is more than a little ironic that even though B&K's analysis focused on the danger of using estimators without fully understanding their properties, so many in the profession applied the B&K method without paying any attention to the simple textbook issues“

- ▶ Informationsgehalt von komparativen Datensätzen auch bei Berücksichtigung der Zeitdimension beschränkt
- ▶ TSCS Analysis ist ein mächtiges Werkzeug ...
- ▶ ... das aber reflektiert angewendet werden muß
- ▶ Methodologisches Hintergrundwissen *und* fachwissenschaftliches Verständnis des Forschungsproblems statt blinder Anwendung von „Kochrezepten“

Warum überhaupt Standardfehler?

- ▶ Zugang 1: Superpopulation
- ▶ Zugang 2: Relative Stärke/Bedeutsamkeit
- ▶ Zugang 3: Reliabilität der Schätzung / Inferenz über den Prozeß
- ▶ Zugang 4: Bayesianische (Subjektive) Interpretation

Wie macht man das in Stata?

- ▶ Daten in Panelformat bringen (ein Fall pro Land/Jahr)
- ▶ Panelstruktur definieren (Gruppenvariable und evtl. Zeitvariable setzen) mit `xtset`
- ▶ `xt`-Kommandos verwenden, z. B. `xtpcse` für den Beck-Katz-Ansatz

Übung für heute: Accountability und Korruption

- ▶ Hat accountability einen Einfluß auf Korruption?
- ▶ Weltbank Daten für 209 Länder (Afghanistan – Zimbabwe) zu fünf Zeitpunkten (1996-2004)
- ▶ Laden Sie den Datensatz von der Homepage herunter :
`http://www.kai-arzheimer.com/Lehre-Regression/wbgovernance.dta`
- ▶ Schauen Sie sich die Daten und die Dokumentation von `xtset` an. Wie können Sie die TSCS-Struktur definieren?
- ▶ Rechnen Sie eine normale Regression von Korruption auf Accountability
- ▶ Lesen Sie die Dokumentation zu `xtpcse`. Versuchen Sie mit `xtpcse` eine panel-korrigierte Regression von Korruption auf Accountability zu rechnen. Vergleichen Sie die Ergebnisse.