

Läufer, Nikolaus K. A.

Working Paper

Anmerkungen zum ED-Ritual in der Ökonometrie

Diskussionsbeiträge - Serie A, No. 174

Provided in Cooperation with:

Department of Economics, University of Konstanz

Suggested Citation: Läufer, Nikolaus K. A. (1983) : Anmerkungen zum ED-Ritual in der Ökonometrie, Diskussionsbeiträge - Serie A, No. 174, Universität Konstanz, Fakultät für Wirtschaftswissenschaften und Statistik, Konstanz

This Version is available at:

<https://hdl.handle.net/10419/75109>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

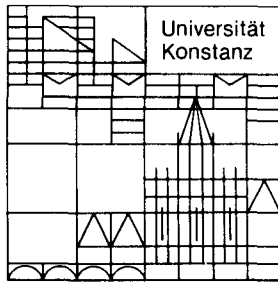
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



671 676 040

**Fakultät für
Wirtschaftswissenschaften
und Statistik**

Nikolaus K. A. Läufer

**Anmerkungen zum ED-Ritual
in der Ökonometrie**

26. NOV. 1984
Wirtschaftswissenschaften
Kiel

Diskussionsbeiträge

Postfach 5560
D-7750 Konstanz

Serie A — Nr. 174
Februar 1983

31. MRZ 1983
Weltwirtschaft
Kiel

Nr 176
nicht besetzt.

ANMERKUNGEN ZUM ED-RITUAL

IN DER ÖKONOMETRIE

Nikolaus K. A. Läufer

Serie A - Nr. 174

Februar 1983

Serie A: Volkswirtschaftliche Beiträge

Serie B: Finanzwissenschaftliche Arbeitspapiere

Serie C: Betriebswirtschaftliche Beiträge

Eine Anmerkung zur Variablentransformation mit Hilfe erster Differenzen (ED) in der gewöhnlichen Regressionsanalyse (OLS)

1. Fragestellung

Die Bildung erster Differenzen (BED) von Regressionsvariablen gilt bei Anwendung der gewöhnlichen Regressionsanalyse (OLS) noch als probates Mittel zur Beseitigung von Autokorrelation und Heteroskedastizität der Residuen, zur Vermeidung von Multikollinearität und zur Ausschaltung von Trendbewegungen in den Regressionsvariablen. Einerseits gilt - für die Zwecke der Koeffizientenschätzung mit OLS - die erste Differenz (ED) einer Regressionsgleichung als der Ausgangsgleichung "äquivalent", andererseits erlaubt die Transformation von Variablen durch BED unter Umständen eine Elimination der eingangs genannten schätztechnischen Störfaktoren mit der Folge einer qualitativen Verbesserung der Teststatistiken (t-Werte, DW-Statistik etc.). Durch BED können sich in dieser Sicht die Voraussetzungen zur regressionsanalytischen Überprüfung von ökonomischen Hypothesen nur verbessern. In der praktischen ökonometrischen Arbeit ist das Ergebnis der Anwendung von ED jedoch meist enttäuschend. Während die zu prüfende Hypothese bei Schätzung der Ausgangsgleichung oft noch bestätigt wird, beobachtet man häufig, daß nach BED der statistische Zusammenhang zwischen den untersuchten Variablen fast völlig verschwindet. Vielfach wird deshalb - wie ich meine: vorschnell - die zu prüfende Theorie verworfen.

Das geschilderte Verfahren wirft verschiedene Fragen auf, u. a. die folgenden:

1. Wenn für den Zweck der Koeffizientenschätzung die Ausgangsgleichung und ihre ED als äquivalent gelten, wie kommt es dann überhaupt zu so unterschiedlichen Koeffizientenschätzungen, je nachdem ob bei OLS vor oder nach BED geschätzt wird?

2. Ist es korrekt, wenn in der ökonometrischen Praxis mit OLS die Richtigkeit (Bestätigung) einer theoretischen Hypothese deswegen in Zweifel gezogen wird, weil sich die geschätzten Koeffizienten beim Übergang zu ED-transformierten Variablen verändern, insbesondere wenn der Absolutbetrag der geschätzten Koeffizienten nennenswert abnimmt?
3. Ist es in der Regressionsanalyse überhaupt sinnvoll, den Trend der ökonomischen Variablen durch BED auszuschalten?

2. Analyse

- a) Die in der ersten Frage postulierte regressionsanalytische Äquivalenz einer Gleichung mit ihrer Form in ED gilt im allgemeinen nur für den Fall der Meßfehlerfreiheit der Beobachtungen für die Regressionsvariablen. Das läßt sich leicht zeigen im einfachsten Fehler-in-den-Variablen-Modell (FVM):

$$Y = Y^W + v \quad (w = \text{wahrer Wert})$$

$$X = X^W + u \quad (u, v = \text{Meßfehler})$$

$$Y^W = a + b X^W + \varepsilon$$

$$\Delta Y^W = b \Delta X^W + \Delta \varepsilon$$

Wie üblich, nehmen wir hier an, daß u und v gegenseitig und seriell unabhängig sind, konstante Varianzen besitzen und ebenfalls unabhängig sind von X^W und Y^W .

Die OLS-Schätzung \hat{b}_n von b und mit nicht transformierten Variablen

$$\hat{b}_n = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

hat bekanntlich den Wahrscheinlichkeitslimes:

$$l_1 = \lim_{n \rightarrow \infty} \hat{b}_n = \frac{b}{1 + \frac{\text{var}(u)}{\text{var}(X^W)}}$$

Wiederholt man die Schätzung nach BED, dann erhält man

$$\hat{b}_n = \frac{\sum_{i=2}^n (\Delta Y_i - \overline{\Delta Y}) (\Delta X_i - \overline{\Delta X})}{\sum_{i=2}^n (\Delta X_i - \overline{\Delta X})^2}$$

mit dem Wahrscheinlichkeitslimes

$$l_2 = \lim_{n \rightarrow \infty} \hat{b}_n = \frac{b}{1 + \frac{\text{var}(u)}{\frac{1}{2} \text{var}(\Delta X^W)}}$$

Die Wahrscheinlichkeitslimes der beiden Schätzungen sind nur dann gleich, wenn entweder die Regressorvariable fehlerfrei gemessen wird ($\text{var}(u) = 0$) oder wenn $\text{var}(X^W) = \frac{1}{2} \text{var}(\Delta X^W)$. Die Wahrscheinlichkeitslimes der beiden Schätzungen sind bei $\text{var}(v) > 0$ umso verschiedener, je stärker $\text{var}(X^W)$ und $\frac{1}{2} \text{var}(\Delta X^W)$ voneinander abweichen. Nimmt also bei hinreichend großem Stichprobenumfang der Absolutbetrag der Koeffizientenschätzung ab (zu), wenn man von einer einfachen Regression ohne ED zu einer solchen mit ED übergeht, dann muß man diesen Unterschied in unserem FVM durch die Ungleichung

$$\text{var}(X^W) > \frac{1}{2} \text{var}(\Delta X^W)$$

(<)

erklären.

Damit läßt sich die erste Frage wie folgt beantworten:
Stochastische Meßfehler beseitigen die regressionsanalytische Äquivalenz einer Gleichung und ihrer Form in ED. Die Formeln zeigen, daß Koeffizientenschätzungen durch Meßfehler in einem Ausmaß verzerrt werden, das von der Größe der Meßfehlervarianz relativ zur Varianz der wahren (systematischen) Komponente der Regressorvariablen abhängt. Durch BED erster Differenzen wird im allgemeinen die relative Bedeutung der Meßfehler und daher die Verzerrung der Koeffizientenschätzung geändert.

Der Fall der Abnahme des Absolutbetrages der Koeffizientenschätzung beim Übergang zu ED ist praktisch besonders häufig und läßt sich eindrucksvoll durch ein einfaches Beispiel illustrieren. Die wahre unbekannte Gleichung sei eine einfache Trendgleichung

$$Y_t^W = a + b X_t^W + \epsilon \quad b \neq 0$$

$$X_t^W = t \quad t = \text{Zeitindex}$$

Es sei aus irgendeinem Grunde t bzw. X_t^W nicht exakt meßbar: $X_t = t + u_t$. In diesem Falle ist $\Delta X_t^W = 1$ ($\forall t$), $\text{var}(\Delta X_t^W) = 0$ und $l_2 = 0$. Nach BED wird hier der Koeffizient b auf Null geschätzt, gleichgültig welchen Wert b in Wahrheit hat. Vor allen Dingen ist diese Verzerrung auch unabhängig von der Größe der Meßfehlervarianz $\text{var}(u)$. Die Meßfehlervarianz muß lediglich von Null verschieden sein, ein bißchen genügt.

Um nun die schädlichen Wirkungen der Trendelimination durch BED allgemeiner zu demonstrieren, führen wir zunächst ein Maß für die Stärke ein, mit der eine Variable trendbehaftet ist, und formulieren anschließend den Limes l_2 als Funktion dieses Maßes.

Definition: In Übereinstimmung mit der Linearität des zugrundeliegenden Regressionsmodells nennen wir eine Regressionsvariable X umso stärker trendbehaftet, je höher die Korrelation $(\rho_{X^W, t})$ ihrer wahren Werte X^W mit einem linearen Zeitindex t ist.

Als nächstes stellen wir den Limes l_2 als Funktion von $\rho_{X^W, t}$ und m_1 def. $\frac{\text{var}(u)}{\text{var}(X^W)}$ dar.

Für die lineare Regression

$$X_t^W = \alpha + \beta t + \eta \quad \left[\text{cov}(t, \eta) = 0 \right]$$

gilt bekanntlich

$$\rho_{X^W, t}^2 = 1 - \frac{\text{var}(\eta)}{\text{var}(X^W)} .$$

Wegen

$$\Delta X^W \stackrel{\text{def.}}{=} X^W - X_{-1}^W = \beta + (\eta - \eta_{-1})$$

gilt ferner

$$\begin{aligned} \text{var}(\Delta X^W) &= \text{var}(X^W - X_{-1}^W) = \text{var}(\eta - \eta_{-1}) \\ &= 2(1 - \rho_{\eta, \eta_{-1}}) \text{var}(\eta) . \end{aligned}$$

Daraus folgt

$$\frac{\text{var}(X^W)}{\frac{1}{2} \text{var}(\Delta X^W)} = \frac{\text{var}(X^W)}{(1 - \rho_{\eta, \eta_{-1}}) \text{var}(\eta)} = \frac{1}{(1 - \rho_{\eta, \eta_{-1}}) (1 - \rho_{X^W, t}^2)}$$

$$m_2 \stackrel{\text{def.}}{=} \frac{\text{var}(u)}{\frac{1}{2} \text{var}(\Delta X^W)} = m_1 \frac{\text{var}(X^W)}{\frac{1}{2} \text{var}(\Delta X^W)} = m_1 \frac{1}{(1-\rho_{\eta, \eta-1})(1-\rho_{X^W, t}^2)}$$

Für l_1 und l_2 erhalten wir damit:

$$l_1 = \frac{b}{1+m_1}, \quad l_2 = \frac{b}{1+m_2} = \frac{b}{1 + \frac{m_1}{(1-\rho_{\eta, \eta-1})(1-\rho_{X^W, t}^2)}}$$

Bei endlichem $\text{var}(X^W)$ ist $\lim_{\rho_{X^W, t} \rightarrow 1} \text{var}(\Delta X_i^W) = \lim_{\rho_{X^W, t} \rightarrow 1} \rho_{\eta, \eta-1} = 0$

Für gegebenes $m_1 > 0$ ist daher

$$\lim_{\rho_{X^W, t} \rightarrow 1} l_2 = 0$$

Interpretation:

Die zusätzliche Verzerrung der Koeffizientenschätzung durch BED ist umso größer, je stärker die zugehörige Regressorvariable trendbehaftet ist, d. h. je größer die Korrelation der wahren Werte X^W mit dem Zeitindex t ist.

Damit ist auch eine Antwort auf die dritte Frage gegeben. Je mehr die wahren Werte der Regressionsvariablen mit der Variablen "Zeit" (t) korrelieren, je zeitähnlicher also die Trendbewegung einer Regressorvariablen ist, desto weniger ist es regressionsanalytisch sinnvoll, die meßfehlerbehafteten Variablen durch BED vom Trend zu befreien.

b) Zur Verallgemeinerung der Ergebnisse betrachteten wir als nächstes ein FVM mit drei Variablen:

$$\begin{aligned}
 Y &= Y^W + v \\
 X_1 &= X_1^W + u_1 \\
 X_2 &= X_2^W + u_2 \\
 Y^W &= a + b_1 X_1^W + b_2 X_2^W + \varepsilon \\
 \Delta Y^W &= b_1 \Delta X_1^W + b_2 \Delta X_2^W + \Delta \varepsilon
 \end{aligned}$$

Wir nehmen wieder an, daß die v , u_1 und u_2 gegenseitig und seriell unabhängig sind, konstante Varianzen besitzen und ebenfalls unabhängig sind von X_1^W , X_2^W und Y^W . In diesem Fall erhält man für den Wahrscheinlichkeitslimes der OLS-Schätzung mit nichttransformierten Variablen

$$l_1' = \lim_{n \rightarrow \infty} \hat{b}_{1n} = \frac{\left[1 + \frac{\text{var}(u_2)}{(1-\rho^2)\text{var}(X_2^W)} \right] b_1 + \frac{\rho \text{var}(u_2)}{(1-\rho^2) \sqrt{\text{var}(X_1^W)\text{var}(X_2^W)}} b_2}{\left[1 + \frac{\text{var}(u_1)}{(1-\rho^2)\text{var}(X_1^W)} \right] \left[1 + \frac{\text{var}(u_2)}{(1-\rho^2)\text{var}(X_2^W)} \right] - \frac{\rho^2 \text{var}(u_1)\text{var}(u_2)}{(1-\rho^2)\text{var}(X_1^W)\text{var}(X_2^W)}}$$

Werden in dieser Formel die Indizes 1 und 2 vertauscht, dann hat man unmittelbar den Wahrscheinlichkeitslimes für \hat{b}_{2n} . Wir bezeichnen ihn mit l_2' . (ρ ist der Korrelationskoeffizient für die Variablen X_1^W und X_2^W). Ersetzt man in den Formeln für l_1' und l_2' wieder $\text{var}(u_i)$ durch $2\text{var}(u_i)$ und $\text{var}(X_i^W)$ durch $\text{var}(\Delta X_i^W)$, dann ergeben sich die Wahrscheinlichkeitslimes $l_1'(ED)$ und $l_2'(ED)$ für die OLS-Schätzung mit ED-transformierten Variablen. (ρ ist in $l_i'(ED)$ der Korrelationskoeffizient für die Variablen ΔX_1^W und ΔX_2^W).

Man sieht unmittelbar, daß die Bedingungen für die Gleichheit von l'_1 und $l'_1(ED)$ bzw. von l'_2 und $l'_2(ED)$ analog die gleichen sind wie im einfacheren FVM: $l'_j = l'_j(ED)$ ($j = 1, 2$), wenn entweder die Regressionsvariablen fehlerfrei gemessen werden, d. h. wenn $\text{var}(u_i) = 0$ ($i = 1, 2$) oder wenn sowohl $\text{var}(X_i^W) = \frac{1}{2} \text{var}(\Delta X_i^W)$ ($i = 1, 2$) als auch $\rho(x_1, x_2) = \rho(\Delta x_1, \Delta x_2)$.

Die Wahrscheinlichkeitslimites l'_j und $l'_j(ED)$ sind bei $\text{var}(u_i) > 0$ ($i = 1, 2$) umso verschiedener, je stärker, c. p., $\text{var}(X_i^W)$ und $\frac{1}{2} \text{var}(\Delta X_i^W)$ voneinander abweichen.

Insbesondere ist

$$\lim_{\text{var}(\Delta X_1^W) \rightarrow 0} l'_1(ED) = \lim_{\text{var}(\Delta X_2^W) \rightarrow 0} l'_2(ED) = 0$$

$$\lim_{\text{var}(\Delta X_2^W) \rightarrow 0} l'_1(ED) = \frac{b_1}{1 + \frac{2\text{var}(u_1)}{\text{var}(\Delta X_1^W)}}$$

$$\lim_{\text{var}(\Delta X_1^W) \rightarrow 0} l'_2(ED) = \frac{b_2}{1 + \frac{2\text{var}(u_2)}{\text{var}(\Delta X_2^W)}}$$

$$\lim_{\substack{\text{var}(\Delta X_1^W) \rightarrow 0 \\ \text{var}(\Delta X_2^W) \rightarrow 0}} l'_j(ED) = 0 \quad (j = 1, 2)$$

Wenn man sich erinnert, daß - bei endlichem $\text{var}(X_i^W)$ -

$$\lim_{\rho_{X^W, t} \rightarrow 1} \text{var}(\Delta X_i^W) = 0,$$

dann zeigen diese Formeln, daß sich die Analyse und Interpretation des einfachen FVM auf das FVM mit drei Variablen unmittelbar übertragen läßt. Insbesondere steht einer Übertragung des illustrierenden Beispiels mit reinen Trendvariablen ($X_{i,t}^W = t$) nichts im Wege. Die Ergebnisse sind vollständig analog.

- c) Wir wenden uns nun der Frage zu, warum eine Veränderung der Koeffizientenschätzung, insbesondere die Abnahme des Absolutbetrages einer Koeffizientenschätzung, durch Übergang zu ED die theoretische Hypothese, die hinter dem Regressionsansatz steht, eher bestätigt als widerlegt.

Für diesen Gedankengang setzen wir positive Meßfehlervarianzen und einen hinreichend großen Stichprobenumfang voraus. Wenn nun der Absolutbetrag einer Koeffizientenschätzung sich durch Übergang zu ED verändert, insbesondere wenn er abnimmt, dann ist das ein Hinweis dafür, daß der entsprechende Wahrscheinlichkeitslimes ein anderer geworden ist, insbesondere daß er abgenommen hat. Eine Veränderung des Wahrscheinlichkeitslimes tritt jedoch nur dann auf, wenn die wahren Koeffizienten (b_i) von Null verschieden sind. Sind die wahren Koeffizienten alle gleich Null, dann kann sich der Wahrscheinlichkeitslimes eines Koeffizienten beim Übergang von untransformierten zu ED-transformierten Variablen nicht verändern. Ist jedoch ein Koeffizient von Null verschieden, so ist das ein Zeichen für das Bestehen einer Abhängigkeit entsprechender Variablen und daher eine Bestätigung für eine Theorie, die diese Abhängigkeit postuliert.

3. Ergebnis

- a) Die Bildung erster Differenzen (BED) zum Zweck der Trend-elimination aus Regressionsvariablen lehnen wir mit folgender Begründung ab:
- α) Koeffizientenschätzungen werden bekanntlich durch stochastische Meßfehler in den Regressorvariablen in einem Ausmaß verzerrt, das von der Bedeutung der Meßfehlervarianz relativ zur Varianz der wahren Komponente abhängt. Wir haben gezeigt, daß und wie die relative Bedeutung der Meßfehlervarianz im allgemeinen durch BED verändert wird und daß daher durch stochastische Meßfehler die oft behauptete regressionsanalytische Äquivalenz einer Gleichung und ihrer Form in ED im allgemeinen aufgehoben wird. (Frage 1).

- β) Die meßfehlerbedingte Verzerrung in der Schätzung von Koeffizienten wird durch BED dann vergrößert, wenn sich durch BED die relative Meßfehlervarianz erhöht. Eine solche Erhöhung ist insbesondere bei trendbehafteten Variablen zu erwarten, und es gilt der folgende Zusammenhang: Je stärker eine Regressorvariable trendbehaftet ist, d. h. je höher die Korrelation der wahren Werte dieser Regressorvariablen mit der Variablen Zeit t ist, umso mehr wird die meßfehlerbedingte Verzerrung durch BED verstärkt. Je ausgeprägter also der Trend einer Regressionsvariablen und je mehr daher ein Ökonometriker versucht ist, den Trend auszuschalten, umso schädlicher ist es für die Koeffizientenschätzung, wenn er ED bildet. (Frage 3).
- b) Wenn sich bei der Regressionsanalyse durch den Übergang zu ED-transformierten Variablen die geschätzten Koeffizienten ändern, insbesondere wenn der Absolutbetrag der Koeffizienten-Schätzung abnimmt, dann ist das eher ein Zeichen dafür, daß die Hypothesen, die hinter dem Regressionsansatz stehen, bestätigt werden, als das Gegenteil. Diese Auffassung steht in deutlichem Gegensatz zur herrschenden ökonometrischen Praxis. Die Begründung, die wir für diese Auffassung gegeben haben, ist folgende: Wir werten die Veränderung der Koeffizientenschätzung als Zeichen für eine Veränderung des Wahrscheinlichkeitslimes der Koeffizientenschätzung. Dieser Limes kann sich jedoch nur verändern, wenn die wahren Koeffizienten der Regressorvariablen von Null verschieden sind. Andererseits sind die wahren Koeffizienten von Null verschieden, wenn die von der Theorie behaupteten Zusammenhänge bestehen. (Frage 2).
- c) Es war das Ziel dieser Anmerkungen, die herrschende ökonometrische Praxis bei der Transformation von Regressorvariablen durch Bildung erster Differenzen (ED) zum Zweck der Trendelimination in Frage zu stellen. Das ökonometrische ED-Ritual wurde dabei ausschließlich aus ökonometrischer Sicht betrachtet. Wie fragwürdig die oft rein mechanische Trendeliminierungspraxis in der Ökonometrie auch aus wirtschaftstheoretischer Sicht ist,

sei an dieser Stelle nur durch Hinweis auf die Permanent-Income-Theorie des Konsums angedeutet. Was bliebe wohl übrig zur Erklärung des Konsums, wenn man aus dem Konsum und dem Einkommen die dauerhaften Komponenten (permanent components), also praktisch den Trend eliminierte?