

Stochastische Überlagerung mit Hilfe der Mischungsverteilung

**Schätzung linearer (Panel-)Modelle
auf Basis anonymisierter Daten**

Version 49 (18. März 2009)

Gerd Ronning

Institut für Angewandte Wirtschaftsforschung e.V.
Ob dem Himmelreich 1 | 72074 Tübingen | Germany
Tel.: +49 7071 98960 | Fax: +49 7071 989699

ISSN: 1617-5654

IAW-Diskussionspapiere

Das Institut für Angewandte Wirtschaftsforschung (IAW) Tübingen ist ein unabhängiges außeruniversitäres Forschungsinstitut, das am 17. Juli 1957 auf Initiative von Professor Dr. Hans Peter gegründet wurde. Es hat die Aufgabe, Forschungsergebnisse aus dem Gebiet der Wirtschafts- und Sozialwissenschaften auf Fragen der Wirtschaft anzuwenden. Die Tätigkeit des Instituts konzentriert sich auf empirische Wirtschaftsforschung und Politikberatung.

Dieses IAW-Diskussionspapier können Sie auch von unserer IAW-Homepage als pdf-Datei herunterladen:

<http://www.iaw.edu/Publikationen/IAW-Diskussionspapiere>

ISSN 1617-5654

Weitere Publikationen des IAW:

- IAW-News (erscheinen 4x jährlich)
- IAW-Forschungsberichte

Möchten Sie regelmäßig eine unserer Publikationen erhalten, dann wenden Sie sich bitte an uns:

IAW Tübingen, Ob dem Himmelreich 1, 72074 Tübingen,
Telefon 07071 / 98 96-0
Fax 07071 / 98 96-99
E-Mail: iaw@iaw.edu

Aktuelle Informationen finden Sie auch im Internet unter:

<http://www.iaw.edu>

Der Inhalt der Beiträge in den IAW-Diskussionspapieren liegt in alleiniger Verantwortung der Autorinnen und Autoren und stellt nicht notwendigerweise die Meinung des IAW dar.

Stochastische Überlagerung mit Hilfe der Mischungsverteilung.

Schätzung linearer (Panel-)Modelle auf Basis anonymisierter Daten

GERD RONNING¹

(Stand: 18. März 2009 - Version 49)

Zusammenfassung

Die Arbeit betrachtet die Auswirkungen von additiver und multiplikativer Überlagerung mit Meßfehlern auf die Schätzung linearer Modelle. Dabei erfolgt diese Überlagerung geplant mit dem Ziel der Anonymisierung von Mikrodaten. Im Mittelpunkt steht die Überlagerung mittels einer (bimodalen) Mischungsverteilung. Die Arbeit analysiert sowohl den Fall von anonymisierten Querschnittsdaten als auch von anonymisierten Paneldaten. Die Auswirkungen der Autokorrelation des Regressors werden ausführlich dargestellt, insbesondere für den Fall "kurzer" Panels. Neben dem naiven Panelschätzer ("Within-Schätzer") werden auch Varianten der Instrumentvariablen(IV)-Schätzung betrachtet.

Summary

The paper considers the effect of additive and multiplicative measurement errors on the estimation of linear models. We assume that such measurement errors have been applied to the micro data by purpose in order to protect them against re-identification. In particular measurement errors with a bimodal mixture distribution are analyzed. First the case of cross-section data is assumed. Then for panel data both the "naive" estimator ("within estimator", fixed effects estimator) and IV estimators are considered. In particular the effect of autocorrelation of regressors in short panels is discussed.

¹Wirtschaftswissenschaftliche Fakultät, Universität Tübingen, D-72074 Tübingen. email: gerd.ronning@uni-tuebingen.de. Ausarbeitung im Rahmen des Projektes "Wirtschaftsstatistische Paneldaten und faktische Anonymisierung", das vom Bundesministerium für Bildung und Forschung finanziell gefördert wurde. Ich danke Elena Biewen, Robert Jung, Sandra Nolte, Winfried Pohlmeier, Martin Rosemann, Hans Schneeweiß und Andy Tremayne für hilfreiche Kommentare und Hinweise.

Inhaltsverzeichnis

1	Einleitung	1
2	Eindimensionale Mischungsverteilungen	1
2.1	Stetige Verteilung	1
2.2	Diskrete Verteilung	2
2.3	Beliebig viele stetige (oder diskrete) Zufallsvariable	2
3	Mehrdimensionale Mischungsverteilungen	3
4	Die Logarithmische Normalverteilung als Mischungsverteilung	6
4.1	Eindimensionale Verteilung	6
4.2	Multivariate Verteilung	7
4.3	Eindimensionale Mischungsverteilung basierend auf der Lognormalverteilung	8
4.4	Bezug zur Mischungsverteilung basierend auf Normalverteilungen	10
4.4.1	Inkonsistenz der additiven Mischung mit der multiplikativen Mischung	11
5	Stochastische Überlagerung durch Mischungsverteilungen	13
5.1	Eindimensionaler Fall	13
5.1.1	Additive Überlagerung	14
5.1.2	Multiplikative Überlagerung	14
5.2	Mehrdimensionaler Fall	14
5.2.1	Additive Überlagerung	15
5.2.2	Multiplikative Überlagerung	15
6	Das Höhne-Verfahren	16
6.1	Ein einziges Merkmal	16
6.1.1	Additiver Fall	16
6.1.2	Multiplikativer Fall	17

6.1.3	Dichtefunktion im Höhne-Ansatz	17
6.2	Mehrere Merkmale gemeinsam	18
6.2.1	Formale Darstellung des Verfahrens	19
6.2.2	Ableitung der Dichtefunktion	20
6.3	Die (spezielle) Höhne-Spezifikation im additiven Fall	21
6.4	Flexible Höhne-Spezifikation	22
6.4.1	Multiplikativer Fall	22
6.4.2	Additiver Fall	24
7	Anonymisierung mittels stochastischer Überlagerung	24
7.1	Additive Überlagerung	25
7.2	Multiplikative Überlagerung in univariaten Verteilungen	25
7.3	Multiplikative Überlagerung in multivariaten Verteilungen	26
7.3.1	Überlagerung mit Hilfe eines konstanten Faktors	27
7.3.2	Überlagerung mit Hilfe unkorrelierter Störvariablen	28
7.3.3	Überlagerung mit Hilfe von korrelierten Störvariablen	29
7.3.4	Überlagerung mit Hilfe der Mischungsverteilung	30
7.3.5	Zusammenfassung	31
7.4	Korrelation bei multiplikativer Überlagerung	31
8	Einfluß der Überlagerung auf Quotienten	33
8.1	Einleitende Bemerkungen	33
8.2	Das Verhältnis von Y_1 zu Y_2 bei multiplikativer Überlagerung	34
8.3	Simulationsergebnisse	36
9	Schätzung linearer Modelle aus Querschnittsdaten	38
9.1	Einleitung	38
9.2	Additive Überlagerung im linearen Regressionsmodell	39
9.2.1	Allgemeine Bemerkungen	39

9.2.2	Ausschließliche Überlagerung der Regressoren	40
9.2.3	Gemeinsame Überlagerung aller Variablen	41
9.3	Multiplikative Überlagerung im linearen Regressionsmodell	42
9.3.1	Allgemeine Bemerkungen	42
9.3.2	Ausschließliche Überlagerung der Regressoren	42
9.3.3	Gemeinsame Überlagerung aller Variablen	43
9.3.4	Alternativer Beweis für (9-18)	45
9.4	Korrekturschätzer	46
10	Schätzung linearer Modelle aus Paneldaten	47
10.1	Allgemeines	47
10.1.1	Überlagerungsstrategien	47
10.1.2	Das einfache lineare Panelmodell	48
10.2	Überlagerung von Paneldaten	48
10.2.1	Symbolik	48
10.2.2	Additive Überlagerung	49
10.2.3	Multiplikative Überlagerung	50
10.3	Das einfache lineare Panelmodell mit Individualeffekten	50
10.3.1	Feste Individualeffekte	50
10.3.2	Stochastische Individualeffekte	51
10.4	Schätzung des linearen Panelmodells aus anonymisierten Daten	52
10.4.1	Der naive Panelschätzer	52
10.4.2	Additive Meßfehler allgemein	52
10.4.3	Allgemeine multiplikative Überlagerung des Regressors	53
10.4.4	Allgemeine multiplikative Überlagerung beider (aller) Variablen	55
10.4.5	Meßfehler mit Faktorstruktur	57
10.4.6	Additive Überlagerung à la Hönne	57
10.4.7	Multiplikative Hönne-Überlagerung des Regressors	58

10.4.8	Multiplikative Höhne-Überlagerung beider (aller) Variablen	61
10.4.9	Eine allgemeinere Form der Höhne-Überlagerung	63
10.4.10	Multiplikative Überlagerung der abhängigen Variablen	63
10.5	Korrekturschätzer	64
10.5.1	Additive Höhne-Überlagerung	64
10.5.2	Multiplikative Höhne-Überlagerung	64
10.6	Analyse für mehrere Regressoren	65
10.6.1	Allgemeines	65
10.6.2	Der Schätzer	66
10.6.3	Eine rechentechnisch attraktivere Darstellung des Schätzers	67
10.6.4	Kovarianzmatrix des Schätzers	67
10.6.5	Eine alternative Ableitung des 'Within'-Schätzers	68
10.7	Ausschließliche Überlagerung der Regressoren	70
10.8	Gemeinsame Überlagerung aller Variablen	74
10.9	Eine alternative Schreibweise	75
10.10	Korrekturschätzer	76
10.11	Korrelation des Regressors mit dem individuenspezifischen Effekt	78
10.11.1	Additiver Fall	78
10.11.2	Multiplikativer Fall	79
11	Schätzung linearer Panelmodelle bei "kleinem" T	81
11.1	Einleitung	81
11.2	Autoregressiver Prozeß erster Ordnung	83
11.2.1	Terminologie	83
11.2.2	Die Schätzfunktion S^2 bei endlichem (kleinem) T	84
11.3	Der AR(1)-Prozeß bei additiver und multiplikativer Überlagerung	86
11.3.1	Allgemeines	86
11.3.2	Additive Überlagerung	87

11.3.3	Multiplikative Überlagerung	87
11.3.4	Multiplikative Überlagerung mit Faktorstruktur	88
11.4	Schätzung der Varianz bei additiver und multiplikativer Überlagerung . . .	90
11.4.1	Additive Überlagerung	90
11.4.2	Multiplikative Überlagerung	91
11.4.3	Multiplikative Überlagerung mit Faktorstruktur	94
11.5	Relevanz der Ergebnisse für das Regressionsmodell	95
11.5.1	Parameterschätzung	95
11.5.2	Simulationsrechnungen für additive Überlagerung	96
11.5.3	Simulationsrechnungen für allgemeine multiplikative Überlagerung .	99
11.5.4	Simulationsergebnisse für Überlagerung à la Höhne	99
11.6	Korrekturschätzer für endliches T	103
11.6.1	Additive Überlagerung	103
11.6.2	Multiplikative Überlagerung	104
11.7	Ein Ausblick auf die Situation bei Paneldaten	106
11.7.1	Additive Überlagerung	106
11.7.2	Allgemeine multiplikative Überlagerung	107
11.7.3	Multiplikative Höhne-Überlagerung	107
11.7.4	Korrekturschätzer	107
12	Instrumentvariablen-Schätzer (IV-Schätzer)	108
12.1	Allgemeine Vorbemerkungen	108
12.2	Die verzögerte Regressorvariable als Instrument	109
12.2.1	Additive Überlagerung:	111
12.2.2	Multiplikative Überlagerung:	112
12.2.3	Multiplikative Überlagerung mit Faktorstruktur:	113
12.2.4	Überprüfung der Eigenschaften (E1), (E2) und (E3)	113
12.3	Schätzung der Autokovarianz bei additiver und multiplikativer Überlagerung	115

12.3.1	Allgemeine Bemerkungen	115
12.3.2	Die Schätzfunktion C bei kleinem T	116
12.3.3	Die Schätzfunktion C_a bei kleinem T : Additive Überlagerung	118
12.3.4	Die Schätzfunktion C_a bei kleinem T : Multiplikative Überlagerung	119
12.3.5	Die Schätzfunktion C_a bei kleinem T : Multiplikative Überlagerung mit Faktorstruktur	120
12.4	Relevanz für den IV-Schätzer	121
12.4.1	Allgemeine Ausführungen	121
12.4.2	Additive Überlagerung	123
12.4.3	Allgemeine multiplikative Überlagerung	125
12.4.4	Multiplikative Höhne-Überlagerung	125
12.4.5	Korrektur-IV-Schätzer bei multiplikativer Überlagerung	126
12.5	Simulationsergebnisse für den IV-Schätzer	127
12.5.1	IV-Schätzung bei additiver Überlagerung	127
12.5.2	IV-Schätzung bei allgemeiner multiplikativer Überlagerung	129
12.5.3	IV-Schätzung bei multiplikativer Überlagerung mit Faktorstruktur	130
12.6	Ein (fast) optimaler IV-Schätzer	130
13	Literatur	134
A	Einige nützliche Matrizen-Resultate	136
B	Beweis von Resultaten	138
B.1	Drei Resultate	138
B.2	Hilfsresultate	138
B.3	Beweise	140
C	Ergebnisse für für symmetrischen IV-Schätzer	142
C.1	Einleitung	142
C.2	Die Schätzfunktion C bei kleinem T	143

C.3	Die Schätzfunktion C_a bei kleinem T	145
C.3.1	Additive Überlagerung	145
C.3.2	Allgemeine multiplikative Überlagerung	146
C.3.3	Multiplikative Überlagerung mit Faktorstruktur	147
C.4	Relevanz für den IV-Schätzer	149
C.4.1	Ergebnisse für den Panel-IV-Schätzer	149

1 Einleitung

Diese Arbeit beschäftigt sich mit speziellen Aspekten der stochastischen Überlagerung, die als Anonymisierungsmethode zur Sicherstellung der faktischen Anonymität von Mikrodaten eingesetzt wird.² Insbesondere soll der Zusammenhang zwischen der Überlagerung mittels einer Mischungsverteilung einerseits und der Überlagerung nach dem sogenannten "Höhne-Verfahren" im Einzelnen dargestellt werden. Wir gehen dabei auch auf die Auswirkungen auf die Beziehung zwischen verschiedenen Merkmalen sowie auf die Analyse von Verhältniszahlen ein. Die Auswirkungen auf die Schätzung linearer Modelle wird eingehend untersucht. Dabei wird zunächst nur der Fall von Querschnittsdaten betrachtet. In den letzten Kapiteln werden ergänzende Überlegungen für Paneldaten angestellt. Insbesondere werden die Auswirkungen von "kurzen Panels" untersucht.

2 Eindimensionale Mischungsverteilungen

2.1 Stetige Verteilung

Es seien V_1 und V_2 zwei stetige Zufallsvariable mit Dichtefunktionen $f_1(v)$ und $f_2(v)$ sowie Erwartungswerten μ_i und Varianzen σ_i^2 für $i = 1, 2$. Wir sagen, eine Zufallsvariable U entstamme einer (diskreten) Mischung der Verteilungen von V_1 und V_2 , falls ihre Dichtefunktion durch

$$g(u) = \alpha f_1(u) + (1 - \alpha) f_2(u) \quad , \quad 0 < \alpha < 1 \quad , \quad (2-1)$$

gegeben ist, wobei bis auf weiteres unerheblich ist, ob die beiden Ausgangsvariablen V_1 und V_2 stochastisch unabhängig sind oder nicht (siehe unten).

Gedanklich kann man sich die Mischungsverteilung auch wie folgt vorstellen: Es gibt zwei (bzw. allgemeiner k - siehe unten) Zustände, die jeweils mit Wahrscheinlichkeit α_i auftreten und die sich gegenseitig ausschließende Ereignisse darstellen. Für jeden Zustand gibt es eine Verteilung der Zufallsvariablen U . Je nachdem welcher Zustand eintritt, wird der Wert der Zufallsvariablen U aus der betreffenden Verteilung generiert.³ Diese Modellvorstellung nutzt man aus, um Zufallszahlen aus Mischungsverteilungen zu erzeugen.

Symbolisch schreiben wir die "Mischung" auch wie folgt:

$$U \sim \alpha F_1(\boldsymbol{\theta}_1) + (1 - \alpha) F_2(\boldsymbol{\theta}_2) \quad ,$$

wobei $F_i(\boldsymbol{\theta}_i)$ die Verteilungsfunktion von V_i bezeichnet und $\boldsymbol{\theta}_i$ deren Parametervektor. Im Fall der Normalverteilung schreiben wir

$$U \sim \alpha N(\mu_1, \sigma_1^2) + (1 - \alpha) N(\mu_2, \sigma_2^2) \quad . \quad (2-2)$$

Für den Erwartungswert von U ergibt sich unter Verwendung von (2-1) direkt

$$E[U] = \alpha \mu_1 + (1 - \alpha) \mu_2 \quad . \quad (2-3)$$

²Siehe Ronning et al (2005).

³Siehe z.B. McLachlan und Peel (2000 section 1.4 "Interpretation of Mixture Models").

Ferner ergibt sich für den Erwartungswert der quadrierten Zufallsvariablen

$$E[U^2] = \alpha E[V_1^2] + (1 - \alpha) E[V_2^2] .$$

Daraus folgt für die Varianz von U^4

$$\begin{aligned} V[U] &= [E[U^2] - (E[U])^2] \\ &= \alpha E[V_1^2] + (1 - \alpha) E[V_2^2] - [\alpha \mu_1 + (1 - \alpha) \mu_2]^2 \\ &= \alpha (\sigma_1^2 + \mu_1^2) + (1 - \alpha) (\sigma_2^2 + \mu_2^2) - [\alpha \mu_1 + (1 - \alpha) \mu_2]^2 \quad (2-4) \\ &= \alpha \sigma_1^2 + (1 - \alpha) \sigma_2^2 + \alpha(1 - \alpha) \mu_1^2 + (1 - \alpha) \alpha \mu_2^2 - 2\alpha(1 - \alpha) \mu_1 \mu_2 \\ &= \alpha \sigma_1^2 + (1 - \alpha) \sigma_2^2 + \alpha(1 - \alpha) (\mu_1 - \mu_2)^2 . \end{aligned}$$

Die letzte Zeile zeigt, daß nicht die absolute Größe der beiden Parameter μ_1 und μ_2 , sondern nur deren Differenz relevant ist. Im übrigen unterscheidet sich dies Ergebnis von dem, das man bekommt, wenn man die **Varianz der Summe aus V_1 und V_2** bestimmt.⁵ Dort würde auch von Bedeutung sein, ob die beiden Zufallsvariablen stochastisch unabhängig sind oder nicht !

2.2 Diskrete Verteilung

Für diesen Fall gelten dieselben Formeln wie im Fall von zwei stetigen Zufallsvariablen. Allerdings ist die Interpretation jetzt etwas anders. Dazu unterstellen wir, daß für V_1 und V_2 folgende Verteilungen gegeben sein sollen:

$$V_1 = \begin{cases} 1 & \text{mit Wahrscheinlichkeit } 0.5 \\ 3 & \text{mit Wahrscheinlichkeit } 0.5 \end{cases}, \quad V_2 = \begin{cases} 2 & \text{mit Wahrscheinlichkeit } 0.5 \\ 4 & \text{mit Wahrscheinlichkeit } 0.5 \end{cases}$$

Man beachte, daß die aus der Mischung von zwei **binären** Zufallsvariablen resultierende Zufallsvariable Y nun die vier verschiedenen Werte 1, 2, 3 und 4 annehmen kann. Alle vier Werte werden jeweils mit Wahrscheinlichkeit 1/4 realisiert, wenn $\alpha = 0.5$ gilt.

2.3 Beliebige viele stetige (oder diskrete) Zufallsvariable

Für eine beliebig große Zahl k von Zufallsvariablen lautet die Formel für die Dichtefunktion

$$g(u) = \sum_{i=1}^k \alpha_i f_i(u) \quad , \quad 0 < \alpha_i < 1 \text{ und } \sum_i \alpha_i = 1 \quad , \quad (2-5)$$

und für Erwartungswert und Varianz erhalten wir

$$E[U] = \sum_{i=1}^k \alpha_i \mu_i \quad , \quad (2-6)$$

⁴Im folgenden wird $V[U]$ und σ_u^2 gleichgesetzt. Davon zu unterscheiden ist der später eingeführte Mischungsverteilungs-Parameter σ_m^2 .

⁵ $Var[V_1 + V_2] = Var[V_1] + Var[V_2] + 2 cov[V_1, V_2]$.

sowie⁶

$$V[U] = \sum_{i=1}^k \alpha_i (\sigma_i^2 + \mu_i^2) - \left(\sum_i \alpha_i \mu_i \right)^2 . \quad (2-7)$$

3 Mehrdimensionale Mischungsverteilungen

Es sei \mathbf{V}_i ein r -dimensionaler Zufallsvektor mit (gemeinsamer) Dichtefunktion

$$f_i(v_{1i}, \dots, v_{ri})$$

sowie Erwartungsvektor

$$\boldsymbol{\mu}_i = \begin{pmatrix} \mu_{1i} \\ \mu_{2i} \\ \vdots \\ \mu_{r-1,i} \\ \mu_{r,i} \end{pmatrix}$$

und Kovarianzmatrix

$$\boldsymbol{\Sigma}_i = \begin{pmatrix} \sigma_{11i} & \sigma_{12i} & & \sigma_{1,r-1,i} & \sigma_{1ri} \\ \sigma_{21i} & \sigma_{22i} & & \sigma_{2,r-1,i} & \sigma_{2ri} \\ & & \ddots & & \\ \sigma_{r-1,1i} & \sigma_{r-1,2i} & & \sigma_{r-1,r-1,i} & \sigma_{r-1,ri} \\ \sigma_{r1i} & \sigma_{r2i} & & \sigma_{r,r-1,i} & \sigma_{rri} \end{pmatrix}$$

Wir sagen, der r -dimensionale Zufallsvektor \mathbf{U} sei durch eine diskrete Mischung von k verschiedenen r -dimensionalen Verteilungen erzeugt worden, wenn die (gemeinsame) Dichtefunktion für diesen Zufallsvektor durch

$$g(u_1, u_2, \dots, u_r) = \sum_{i=1}^k \alpha_i f_i(u_1, u_2, \dots, u_r) \quad (3-1)$$

gegeben ist.

Symbolisch schreiben wir - speziell im Fall der Normalverteilung - für die Mischung:

$$\mathbf{U} \sim \sum_{i=1}^k \alpha_i N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) .$$

Für den r -dimensionalen Vektor der Erwartungswerte des Zufallsvektors \mathbf{U} ist für jedes j das Integral

$$E[U_j] = \int \dots \int u_j g(u_1, \dots, u_j, \dots, u_r) du_r \dots du_j \dots du_1$$

⁶Eine Vereinfachung wie in Zeile 5 von Formel (2-4) für den Spezialfall von zwei Mischungskomponenten ist nicht möglich.

zu bestimmen.⁷ Der Einfachheit halber betrachten wir die erste Komponente, d.h. $j = 1$ und schreiben für das Integral

$$\begin{aligned} E[U_1] &= \int u_1 \left\{ \underbrace{\int \int \dots \int g(u_1, \dots, u_j, \dots, u_r) du_r du_{r-1} \dots du_2}_{g_1(u_1)} \right\} du_1 \\ &= \int u_1 g_1(u_1) du_1 \quad . \end{aligned}$$

Andererseits läßt sich unter Verwendung der Definition (3-1) für die gemeinsame Dichtefunktion schreiben:

$$\begin{aligned} E[U_1] &= \sum_{i=1}^k \alpha_i \int u_1 \left\{ \underbrace{\int \int \dots \int f_i(u_1, \dots, u_r) du_r du_{r-1} \dots du_2}_{f_{i1}(u_1)} \right\} du_1 \\ &= \sum_{i=1}^k \alpha_i \int u_1 f_{i1}(u_1) du_1 \\ &= \sum_{i=1}^k \alpha_i \mu_{1i} \quad . \end{aligned}$$

Dabei bezeichnet f_{i1} die Randdichte der Zufallsvariablen V_{i1} , d.h. der jeweils ersten Komponente bei den Zufallsvektoren \mathbf{V}_i . Daraus folgt

$$E[\mathbf{U}] = \sum_{i=1}^k \alpha_i \boldsymbol{\mu}_i \quad (3-2)$$

Für die Kovarianzmatrix des Zufallsvektors \mathbf{U} läßt sich unter Verwendung der Definition

$$\text{cov}[\mathbf{U}] = E[\mathbf{U}\mathbf{U}'] - E[\mathbf{U}]E[\mathbf{U}]'$$

außerdem das folgende Ergebnis herleiten:^{8 9}

$$\text{cov}[\mathbf{U}] = \sum_{i=1}^k \alpha_i (\boldsymbol{\Sigma}_i + \boldsymbol{\mu}_i \boldsymbol{\mu}_i') - \left(\sum_{i=1}^k \alpha_i \boldsymbol{\mu}_i \right) \left(\sum_{i=1}^k \alpha_i \boldsymbol{\mu}_i \right)' \quad (3-3)$$

⁷Siehe z.B. die Definition bei Graybill, Introduction to Matrices with Applications in Statistics. 1969. Kap. 10.4.

⁸Man beachte, daß

$$E[\mathbf{U}\mathbf{U}'] = \begin{pmatrix} E[U_1^2] & E[U_1 U_2] & \dots & E[U_1 U_k] \\ E[U_1 U_2] & E[U_2^2] & \dots & E[U_2 U_k] \\ \vdots & \vdots & \ddots & \vdots \\ E[U_k U_1] & \dots & \dots & E[U_k^2] \end{pmatrix}$$

und beispielsweise

$$E[U_1 U_2] = \sigma_{12} + \mu_1 \mu_2$$

gilt.

⁹Aus (3-3) folgt, daß das (t, s) -Element der Kovarianz-Matrix (3-3) wie folgt aussieht:

$$\text{cov}(u_t, u_s) = \sum_i \alpha_i (\sigma_i(t, s) + \mu_i(t) \mu_i(s)) - \left(\sum_i \alpha_i \mu_i(t) \right) \left(\sum_i \alpha_i \mu_i(s) \right).$$

Dabei ist $\sigma_i(t, s)$ das (t, s) -Element aus Matrix $\boldsymbol{\Sigma}_i$ und $\mu_i(t)$ das t -Element aus dem Vektor $\boldsymbol{\mu}_i$. In der Überlagerung zwecks Anonymisierung dürften die $\sigma_i(t, s)$ alle gleich Null sein, ausserdem ist $\alpha_i = 0.5, i = 1, 2$ und $\mu_i(t) = \pm \mu$ in der additiven Überlagerung bzw. $\mu_i(t) = 1 \pm \delta$ in der multiplikativen Überlagerung.

Man beachte, daß auch bei Nullkorrelation in den Ausgangsverteilungen eine von Null verschiedene Korrelation erzeugt wird, sofern nicht alle Elemente des Mittelwertvektors für alle r Merkmale gleich Null ist. Daß bei unterschiedlichen Mittelwerten eine von Null verschiedene Korrelation erzeugt wird, zeigt auch die Graphik 3/1. Man könnte dieses Bild auch als Streudiagramm mit zwei "schwergewichtigen" Ausreißern interpretieren, die die Korrelation erzeugen.

GAUSS Sun Aug 26 17:46:21 2007

MISCHUNG VON ZWEI BIVARIATEN NORMALVERTEILUNGEN

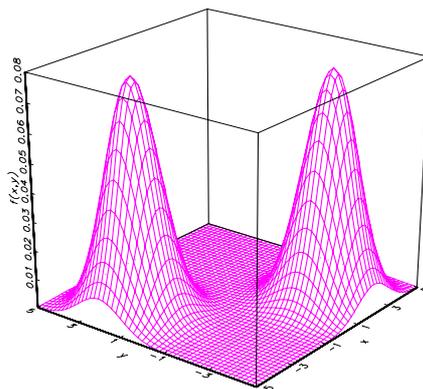
$$\mu_{-1}=(2,-3), \mu_{-2}=(-3,2), \text{sig}_{-1}=\text{sig}_{-2}=(1,1), \text{rho}_{-1}=\text{rho}_{-2}=0$$


Abbildung 3/1: Mischung von zwei bivariaten Normalverteilungen

Falls man unterstellt, daß alle k Komponenten der Mischungsverteilung von \mathbf{U} die gleiche Korrelation aufweisen, ist - beispielsweise für den Fall $r = 4$ - die Kovarianzmatrix durch

$$\Sigma_j = \sigma_j^2 \begin{pmatrix} 1 & \varrho_j & \varrho_j & \varrho_j \\ \varrho_j & 1 & \varrho_j & \varrho_j \\ \varrho_j & \varrho_j & 1 & \varrho_j \\ \rho & \varrho_j & \varrho_j & 1 \end{pmatrix} = \sigma_j^2 [(1 - \varrho_j) \mathbf{I} + \varrho_j \mathbf{u} \mathbf{u}'] \quad (3-4)$$

gegeben, wobei ϱ_j der Korrelationsparameter ist und \mathbf{u} den Einsvektor bezeichnet. Man beachte, daß dieser Parameter nicht beliebig negativ werden kann. Es muß

$$-\frac{1}{r-1} < \varrho_j \leq 1$$

gelten, damit die obige Matrix positiv definit ist.

Dann ergibt sich für die additive Überlagerung

$$\begin{aligned} \text{cov}(u_t, u_s) &= \sum_i \alpha_i (\mu_i(t) \mu_i(s)) - (\sum_i \alpha_i \mu_i(t)) (\sum_i \alpha_i \mu_i(s)) \\ &= 0.5(\mu^2 + (-\mu)^2) - (0.5\mu - 0.5\mu)^2 \\ &= \mu^2 \end{aligned}$$

und für die multiplikative Überlagerung

$$\begin{aligned} \text{cov}(u_t, u_s) &= \sum_i \alpha_i (\mu_i(t) \mu_i(s)) - (\sum_i \alpha_i \mu_i(t)) (\sum_i \alpha_i \mu_i(s)) \\ &= 0.5((1+\delta)^2 + (1-\delta)^2) - (0.5(1+\delta) + 0.5(1-\delta))^2 \\ &= \delta^2 \end{aligned}$$

Dies bedeutet, daß auch in diesem speziellen Fall durch die Mischungsverteilung eine **positive** Korrelation erzeugt wird.

Falls wir unterstellen, daß alle involvierten Mischungsverteilungen von der Struktur (3-4) sind, dann ergibt sich für die Kovarianzmatrix von \mathbf{U}

$$\text{cov}[\mathbf{U}] = \sum_{i=1}^k \alpha_i (\sigma_i^2 [(1 - \varrho_i) \mathbf{I} + \varrho_i \boldsymbol{\mu} \boldsymbol{\mu}'] + \boldsymbol{\mu}_i \boldsymbol{\mu}_i') - \left(\sum_{i=1}^k \alpha_i \boldsymbol{\mu}_i \right) \left(\sum_{i=1}^k \alpha_i \boldsymbol{\mu}_i \right)' \quad (3-5)$$

4 Die Logarithmische Normalverteilung als Mischungsverteilung

4.1 Eindimensionale Verteilung

Es sei Y eine normalverteilte Zufallsvariable mit $E[Y] = \mu_y$ und $V[Y] = \sigma_y^2$, d.h. $Y \sim N(\mu_y, \sigma_y^2)$. Dann sagen wir, $X := \exp(Y)$ sei lognormalverteilt und schreiben

$$X \sim L(\mu_y, \sigma_y^2) \quad . \quad (4-1)$$

Die Dichtefunktion ist gegeben durch

$$f(x) = \begin{cases} \frac{1}{x \sigma_y \sqrt{2\pi}} \cdot \exp \left\{ -\frac{1}{2\sigma_y^2} [\log_e(x) - \mu_y]^2 \right\} & , x > 0 \\ 0 & , \text{sonst} \end{cases} \quad (4-2)$$

und für Erwartungswert und Varianz gilt:

$$E[X] = \exp \left(\mu_y + \frac{\sigma_y^2}{2} \right) \quad (4-3)$$

$$V[X] = \exp(2\mu_y + \sigma_y^2) (\exp(\sigma_y^2) - 1) = \exp(2\mu_y + 2\sigma_y^2) - \exp(2\mu_y + \sigma_y^2) \quad (4-4)$$

Man beachte bei der Schreibweise in (4-1), daß μ_y und σ_y^2 *nicht* Erwartungswert und Varianz der Verteilung sind, sondern daß diese durch (4-3) und (4-4) gegeben sind. Ferner kann man zeigen, daß der Modalwert, also das Maximum der Dichtefunktion durch

$$h = \exp(\mu_y - \sigma_y^2) \quad (4-5)$$

gegeben ist, und daß für den Median bzw. Zentralwert

$$z = \exp(\mu_y) \quad (4-6)$$

gilt.

Bei der multiplikativen Überlagerung, die in dieser Arbeit eine wichtige Rolle spielt (siehe insbesondere Kapitel 7), wird gefordert, daß die betrachtete Überlagerungsvariable Erwartungswert 1 besitzen soll. Hier soll kurz die sich daraus ergebende Parametrisierung dargestellt werden. Dabei verwenden wir eine etwas andere Symbolik.

Die erwähnte Forderung bedeutet, daß

$$\mu_x = \exp \left\{ \mu_y + \frac{\sigma_y^2}{2} \right\} = 1$$

oder auch

$$\mu_y = -\frac{\sigma_y^2}{2}$$

gelten muß. Da andererseits die Varianz als

$$\sigma_x^2 = \mu_x^2 (\exp\{\sigma_y^2\} - 1)$$

geschrieben werden kann, erhalten wir wegen der hier geforderten Eigenschaft des Erwartungswertes

$$\exp\{\sigma_y^2\} = 1 + \sigma_x^2$$

oder

$$\sigma_y^2 = \ln(1 + \sigma_x^2) \quad (4-7)$$

und damit auch wegen der Beziehung zwischen der oben angegebenen Beziehung zwischen μ_y und σ_y^2

$$\mu_y = -\frac{1}{2} \ln(1 + \sigma_x^2) \quad (4-8)$$

Diese beiden Formeln können benutzt werden, um für gegebene Varianz der Überlagerungsvariablen X die Parameter der zugrundeliegenden Normalverteilung für Y zu bestimmen.

4.2 Multivariate Verteilung

Die multivariate Logarithmische Normalverteilung (Lognormalverteilung) wird beispielsweise bei S.J. Press (Applied Multivariate Analysis, 1972, Kap. 6.4) behandelt. Sie wird durch den Vektor $\boldsymbol{\mu}$ und die (positiv definite) Matrix $\boldsymbol{\Sigma}$ charakterisiert.

Für die gemeinsame Dichtefunktion des r -dimensionalen Vektors ergibt sich

$$f(v_1, \dots, v_r) = \frac{1}{(2\pi)^{r/2} (\det(\boldsymbol{\Sigma}))^{1/2} \prod_{s=1}^r v_s} \exp \left\{ -\frac{1}{2} (\overline{\log \mathbf{v}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\overline{\log \mathbf{v}} - \boldsymbol{\mu}) \right\}$$

Dabei ist

$$\overline{\log \mathbf{v}} \equiv \begin{pmatrix} \log(v_1) \\ \vdots \\ \log(v_r) \end{pmatrix}$$

Für die Erwartungswerte gilt

$$E[V_s] = \exp \left\{ \mu_s + \frac{\sigma_s^2}{2} \right\}$$

und für die Kovarianzen ergibt sich

$$\text{cov}[V_s, V_t] = \exp \left\{ \mu_s + \mu_t + \frac{\sigma_s^2 + \sigma_t^2}{2} + \rho_{st} \sigma_s \sigma_t \right\} - \exp \left\{ \mu_s + \mu_t + \frac{\sigma_s^2 + \sigma_t^2}{2} \right\}$$

Daraus folgt speziell für die Varianzen ($t = s$)

$$\begin{aligned} \text{var}[V_s] &= \exp \{2\mu_s + 2\sigma_s^2\} - \exp \{2\mu_s + \sigma_s^2\} \\ &= \exp \{2\mu_s + \sigma_s^2\} (\exp \{\sigma_s^2\} - 1) \end{aligned}$$

Demnach haben die Erwartungswerte und Varianzen im multivariaten Fall dieselbe Struktur wie im univariaten Fall.

Außerdem ist erwähnenswert, daß die Kovarianzen und damit die Korrelationskoeffizienten gleich Null sind, wenn die Parameter $\rho_{ts} = 0$ gesetzt werden. Andererseits impliziert die Parameterkonstellation $\rho_{ts} = 1$ nicht, daß die lognormalverteilten Zufallsvariablen 'perfekte' Korrelation aufweisen. Denn für $\rho_{ts} = 1$ ergibt sich

$$\begin{aligned} \text{cov}[V_s, V_t] &= \exp\left\{\mu_s + \mu_t + \frac{\sigma_s^2 + \sigma_t^2}{2} + \sigma_s \sigma_t\right\} - \exp\left\{\mu_s + \mu_t + \frac{\sigma_s^2 + \sigma_t^2}{2}\right\} \\ &= \exp\left\{\mu_s + \frac{\sigma_s^2}{2}\right\} \exp\left\{\mu_t + \frac{\sigma_t^2}{2}\right\} (\exp\{\sigma_s \sigma_t\} - 1) \end{aligned}$$

Die Korrelation wäre gleich 1, wenn dieser Ausdruck gleich dem Produkt der Standardabweichungen wäre. Jedoch ist dieser durch einen zwar sehr ähnlich aussehenden, aber abweichenden Ausdruck gegeben:

$$\sqrt{\text{var}[V_s, V_t]} = \exp\left\{\mu_s + \frac{\sigma_s^2}{2}\right\} \exp\left\{\mu_t + \frac{\sigma_t^2}{2}\right\} \sqrt{(\exp\{\sigma_s\} - 1)(\exp\{\sigma_t\} - 1)}$$

Damit die Korrelation stets kleiner oder gleich 1 ist, muß demnach der zuerst gegebene Ausdruck größer sein als der zweite, bzw. es muß

$$(\exp\{\sigma_s \sigma_t\} - 1) > \sqrt{(\exp\{\sigma_s\} - 1)(\exp\{\sigma_t\} - 1)}$$

gelten. Für den Fall $\sigma_s = \sigma_t$ ist dies offensichtlich der Fall.¹⁰ Anders gesagt bedeutet dies, daß die Korrelation **nie** gleich 1 sein kann!!

4.3 Eindimensionale Mischungsverteilung basierend auf der Lognormalverteilung

Eine eindimensionale Mischungsverteilung, deren **zwei** Komponenten lognormalverteilt sind, ergibt sich aus der allgemeinen Formel (2-1) unter Verwendung der speziellen Formel (4-2) für die Dichtefunktion der Lognormalverteilung. Ein Beispiel findet sich in Abbildung 4/2, rechte Seite. Dabei sind die Parameter wie folgt gewählt:¹¹

$$\mu_1 = -0.2957, \sigma_1^2 = 0,1452, \mu_2 = 0,6688, \sigma_2^2 = 0,0488,$$

d.h. die beiden Komponenten haben die folgenden Erwartungswerte:

$$E[V_1] = \exp\{-0,2231\} = 0,80 \quad , \quad E[V_2] = \exp\{0,6932\} = 2,00$$

Dies läßt sich optisch aus der Abbildung verifizieren. Man beachte, daß diese Mischungsverteilung stets nur positive Werte annimmt!

Für die multiplikative Überlagerung, bei der diese Mischungsverteilung eingesetzt werden soll, wird verlangt, daß der Erwartungswert dieser Verteilung gleich 1 ist. Formaler soll gelten:

$$0.5 \exp\left\{\mu_1 + \frac{\sigma_1^2}{2}\right\} + 0.5 \exp\left\{\mu_2 + \frac{\sigma_2^2}{2}\right\} = 1 \quad .$$

¹⁰Beweis für den allgemeinen Fall ???

¹¹Dieses Beispiel verwendet auch Ronning (2005). Siehe Fußnote Seite 77.

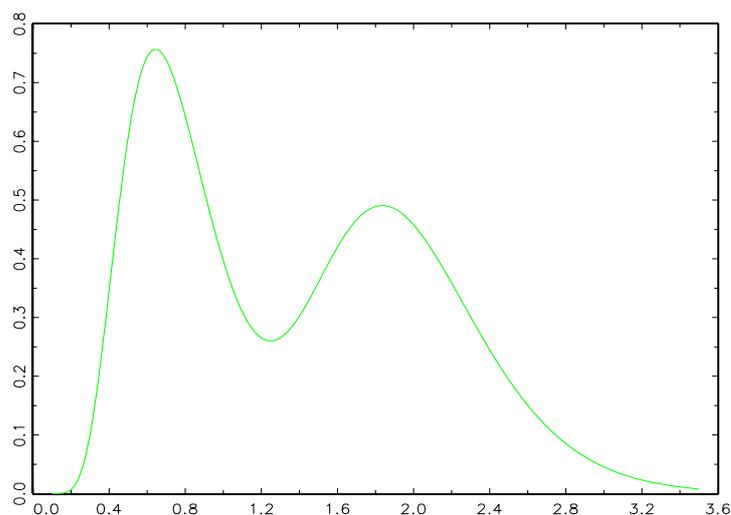


Abbildung 4/2: Mischung von zwei Logarithmischen Normalverteilungen

Dies ließe sich am einfachsten dadurch erreichen, daß man

$$\mu_1 = -\sigma_1^2/2 = \quad \text{und} \quad \mu_2 = -\sigma_2^2/2$$

setzt. Allerdings ist die resultierende Mischungsverteilung **eingipflig** und damit für unsere Zwecke wertlos.

Eine unter vielen Varianten, die zu einer bimodalen Verteilung führen, ist wie folgt: Wir setzen

$$\sigma_1^2 = \sigma_2^2 = \sigma^2, \quad \sigma^2 \text{ ein beliebiger positiver Wert}$$

und schreiben die obige Bedingung als

$$0.5 \cdot \exp\{\sigma^2/2\} (\exp\{\mu_1\} + \exp\{\mu_2\}) = 1$$

Sodann wählen wir zwei positive Zahlen c und d mit der Restriktion $c + d < 2$ und setzen

$$\mu_1 = \log(c) \quad \text{und} \quad \mu_2 = \log(d).$$

Dann läßt sich die Bedingung wie folgt schreiben:

$$c + d = 2 \cdot \exp\{-\sigma^2/2\}$$

Daraus bestimmen wir den Wert für σ^2 wie folgt:

$$\sigma^2 = 2 (\log(2) - \log(c + d))$$

Allerdings ist die Restriktion, daß c und d positiv und in der Summe kleiner als 2 sein sollen, vermutlich eine starke Restriktion. Numerische Beispiele dazu sind geplant.

4.4 Bezug zur Mischungsverteilung basierend auf Normalverteilungen

Für die Mischungsverteilung mit lognormalverteilten Komponenten gilt gemäß (2-1) und (4-2)

$$g(x) = \alpha \frac{1}{x \sigma_1 \sqrt{2\pi}} \exp \left\{ -\frac{(\log(x) - \mu_1)^2}{2\sigma_1^2} \right\} + (1-\alpha) \frac{1}{x \sigma_2 \sqrt{2\pi}} \exp \left\{ -\frac{(\log(x) - \mu_2)^2}{2\sigma_2^2} \right\} \quad (4-9)$$

Es soll die daraus resultierende Verteilung für die Zufallsvariable $Y = \log(X)$ bestimmt werden. Wir wenden den Transformationssatz für stetige Verteilungen an, der in allgemeiner Form wie folgt lautet:

Für die Dichte der Zufallsvariablen $Y = h(X)$ gilt

$$\phi_y(y) = \psi_x(v(y)) |v'(y)| \quad ,$$

wobei $v(y) = h^{-1}(y)$ die Umkehrfunktion von h und $\psi_x(x)$ die Dichtefunktion von X ist.

In unserem Fall gilt $Y = \log(X)$, d.h. $h(x) = \log(x)$ und damit $v(y) = \exp(y)$ bzw. $v'(y) = \exp(y) = \exp(\log(x)) = x$. Wir erhalten also

$$\phi(y) = \psi(\exp(x)) x$$

Anders gesagt: Die Dichtefunktion von X ist mit x zu multiplizieren, außerdem tritt an die Stelle von x jetzt $\exp(y)$. Damit erhalten wir für die Zufallsvariable $Y = \log(x)$ den folgenden Ausdruck:

$$\begin{aligned} \phi(y) &= x \left(\alpha \frac{1}{x \sigma_1 \sqrt{2\pi}} \exp \left\{ -\frac{(\log(\exp(y)) - \mu_1)^2}{2\sigma_1^2} \right\} + (1-\alpha) \frac{1}{x \sigma_2 \sqrt{2\pi}} \exp \left\{ -\frac{(\log(\exp(y)) - \mu_2)^2}{2\sigma_2^2} \right\} \right) \\ &= \alpha \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left\{ -\frac{(y - \mu_1)^2}{2\sigma_1^2} \right\} + (1-\alpha) \frac{1}{\sigma_2 \sqrt{2\pi}} \exp \left\{ -\frac{(y - \mu_2)^2}{2\sigma_2^2} \right\} \end{aligned} \quad (4-10)$$

Dies ist jedoch exakt die Formel für die Mischung von zwei Normalverteilungen mit Erwartungswert

$$E[Y] = \alpha \mu_1 + (1-\alpha) \mu_2 \quad . \quad (4-11)$$

Demnach läßt sich die Mischung von logarithmischen Normalverteilungen durch Logarithmierung in eine Mischung von Normalverteilungen überführen, allerdings mit einem Erwartungswert, der von der jeweiligen Parameterkonstellation abhängt (und **nicht automatisch den Wert Null annimmt**¹², wenn die Mischung der logarithmischen Normalverteilungen den Erwartungswert Eins hat!) Beispielsweise ergibt sich bei der oben gewählten Methode zur Festlegung der Parameter als Erwartungswert

$$E[Y] = \alpha \log(c) + (1-\alpha) \log(d) \quad .$$

Der Erwartungswert ist nur dann Null, wenn ($\alpha = 0.5$ und)

$$d = 1/c$$

¹²Das gilt übrigens auch bereits für die Beziehung zwischen Normalverteilung und logarithmischer Normalverteilung selbst (ohne Mischung)!

gilt, also z.B. $c = 3/2$ und $d = 2/3$.

Wählt man andererseits - ausgehend von der Mischung von zwei Normalverteilungen - die (multiplikative) Spezifikation à la Höhne, d.h. ($\alpha = 0.5$ und)

$$\mu_1 = 1 + \delta \quad , \quad \mu_2 = 1 - \delta \quad , \quad \sigma^2 = \delta^2 + \sigma_\varepsilon^2 \quad ,$$

dann ergibt sich für die Mischung der logarithmischen Normalverteilungen

$$E[X] = 0.5 \exp \left\{ 1 + \delta + \frac{\delta^2 + \sigma_\varepsilon^2}{2} \right\} + 0.5 \exp \left\{ 1 - \delta + \frac{\delta^2 + \sigma_\varepsilon^2}{2} \right\}$$

Dieser Ausdruck kann, weil die Varianz σ_ε^2 beliebig gewählt werden kann, auch beliebige Werte annehmen!

Entsprechend erhalten wir für die additive Spezifikation, d.h. ($\alpha = 0.5$ und)

$$\mu_1 = \mu \quad , \quad \mu_2 = -\mu \quad , \quad \sigma^2 = \mu^2 + \sigma_\varepsilon^2 \quad ,$$

als Erwartungswert

$$E[X] = 0.5 \exp \left\{ \mu + \frac{\mu^2 + \sigma_\varepsilon^2}{2} \right\} + 0.5 \exp \left\{ -\mu + \frac{\mu^2 + \sigma_\varepsilon^2}{2} \right\}.$$

4.4.1 Inkonsistenz der additiven Mischung mit der multiplikativen Mischung

Ich möchte kurz den Effekt der (additiven) Überlagerung der logarithmierten Werte auf die Situation bei den antilogarithmierten Werten beschreiben und zeigen, daß **nicht** auf eine multiplikative Mischung der antilogarithmierten Variablen geschlossen werden kann.

Ich unterstelle, daß ein Merkmal y logarithmiert wird, d.h. ich betrachte

$$Y = \log X \quad ,$$

und daß dieses (logarithmierte) Merkmal additiv mit einer Mischungsverteilung (Höhneverfahren) überlagert wird. Das bedeutet, daß die Dichte für die Überlagerungsvariable U durch

$$f(u) = 0.5 N(\mu, \sigma^2) + 0.5 N(-\mu, \sigma^2)$$

gegeben ist, also die Mischung von zwei Normalverteilungen.

Für die antilogarithmierte Überlagerungsvariable

$$W = \exp(U)$$

gilt, daß sie logarithmisch normalverteilt ist mit Erwartungswert einer Mischung von zwei Lognormalverteilungen, d.h. mit

$$E[W] = 0.5 \exp \left(\mu + \frac{\sigma^2}{2} \right) + 0.5 \exp \left(-\mu + \frac{\sigma^2}{2} \right)$$

Man beachte, daß die beiden Komponenten der Mischungsverteilung so gewählt sind, daß der Erwartungswert von U gleich Null ist. Was aber folgt für den Erwartungswert von W ?

Wir nehmen konkret an, daß wie im Projekt $\mu = 0.11$ und $\sigma = 0.03$ gewählt wurde. Dann ergibt sich aus obiger Formel

$$E[W] = 0.5 \exp\left(0.11 + \frac{0.09}{2}\right) + 0.5 \exp\left(-0.11 + \frac{0.09}{2}\right) = \frac{2.5472}{2} + \frac{3.1740}{2} = 2.8606 .$$

Wenn man also statt

$$Y^a = Y + U$$

die multiplikative Variante $X^a = \exp(Y^a)$ bzw.

$$X^a = X \cdot W$$

betrachtet, dann ist bei obigem Verfahren nicht gewährleistet, daß der Erwartungswert von W gleich 1 ist, wie dies für eine (unverzerrte) multiplikative Überlagerung verlangt wird. Anders gesagt, es ist nicht sinnvoll, bei diesem Verfahren auch die Originalvariablen (= antilogarithmierten Werte) als sinnvolle multiplikative Überlagerung zu interpretieren.

Auf das Problem, sowohl für additive als auch für multiplikative Überlagerung eine sinnvolle Parametrisierung anzugeben, gehe ich in Abschnitt 4.3 ausführlich ein. Dort wird eine Methode vorgeschlagen, die identische Varianz in beiden Mischungskomponenten (wie oben in der Methode) unterstellt. Dann ergibt sich folgende Gleichung:

$$\frac{\sigma^2}{2} = \ln(2) - \ln(\exp(\mu_1) + \exp(\mu_2))$$

aus der die beiden Erwartungswerte iterativ zu bestimmen wären. Dabei müßte noch zusätzlich unterstellt werden, daß

$$\mu_1 = -\mu_2$$

gilt, d.h. die Gleichung, die für μ_1 (bei gegebenem σ^2) oder für σ^2 (bei gegebenem μ_1) zu lösen ist, lautet

$$\frac{\sigma^2}{2} = \ln(2) - \ln(\exp(\mu_1) + \exp(-\mu_1))$$

bzw.

$$\sigma^2 = 2 [\ln(2) - \ln \{\exp(\mu_1) + \exp(-\mu_1)\}]$$

Um zu untersuchen, ob dafür immer eine Lösung existiert, betrachten wir zunächst die Ableitung von $\ln \{\exp(\mu_1) + \exp(-\mu_1)\}$ nach μ_1 , d.h.¹³

$$\frac{d\sigma^2}{d\mu_1} \ln \{\exp(\mu_1) + \exp(-\mu_1)\} = \frac{1}{\{\exp(\mu_1) + \exp(-\mu_1)\}} \{\exp(\mu_1) - \exp(-\mu_1)\} \begin{cases} > 0 & \mu_1 > 0 \\ = 0 & \mu_1 = 0 \\ < 0 & \mu_1 < 0 \end{cases}$$

d.h. diese Funktion hat ein Minimum bei $\mu_1 = 0$. Dieses Minimum ergibt sich als

$$\lim_{\mu_1 \rightarrow 0} \ln \{\exp(\mu_1) + \exp(-\mu_1)\} = \ln\{1 + 1\} = \ln(2)$$

Demnach ergibt sich für die obige Funktion σ^2 stets ein **negativer** Wert oder, anders gesagt, es gibt keine zulässige Lösung.

¹³Es wird ausgenutzt, daß

$$\{\exp(\mu_1) - \exp(-\mu_1)\} = \left\{ \exp(\mu_1) - \frac{1}{\exp(\mu_1)} \right\}$$

größer/kleiner als Null ist, wenn μ_1 größer/kleiner als Null ist.

5 Stochastische Überlagerung durch Mischungsverteilungen

Bei der stochastischen Überlagerung hat die Mischungsverteilung die Aufgabe, die Anonymisierung in der Weise zu verstärken, daß möglichst wenige Realisationen den Originalwert - annähernd - unverändert lassen. Anders gesagt: Die anonymisierten Werte sollten nicht in der Nähe der Originalwerte liegen. Deshalb sollten im Fall der additiven Überlagerung die Realisationen der Überlagerungsvariablen möglichst weit entfernt vom Wert Null liegen (siehe auch Abbildung 5/1), im Fall der multiplikativen Überlagerung möglichst weit entfernt vom Wert Eins.¹⁴ Dies läßt sich am besten bzw. einfachsten mit dem in Abschnitt 2.1 betrachteten Spezialfall von **zwei** (stetigen) Mischungskomponenten (d.h. $k = 2$) erreichen. Andererseits wird man zum Zwecke der Anonymisierung beide Mischungskomponenten symmetrisch wählen, insbesondere die Gewichte und die Varianzen gleich groß wählen.

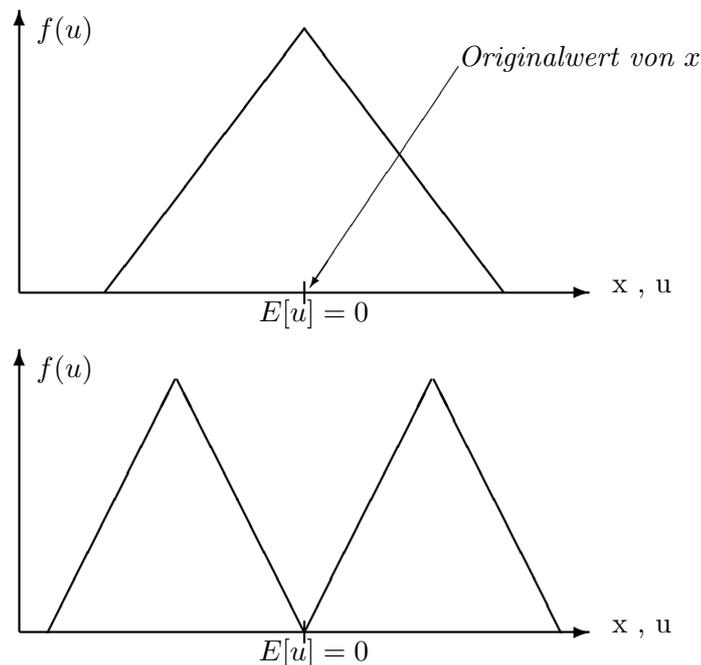


Abbildung 5/1: (Additive) Überlagerung mittels Mischungsverteilung

Wir behandeln zunächst - ausführlich - den eindimensionalen Fall und dann eher cursorisch den mehrdimensionalen Fall. In beiden Fällen wird zunächst die additive und dann die multiplikative Überlagerung betrachtet.

5.1 Eindimensionaler Fall

Im folgenden beschreiben wir die speziellen Parametrisierungen im Fall der Überlagerung und wenden dann die Ergebnisse aus Abschnitt 2.1 an, um die ersten und zweiten Momente

¹⁴Die Idee der Überlagerung mittels Mischungsverteilung und insbesondere der multiplikativen Überlagerung wurde ebenfalls von Massell, Zayatz und Funk (2006) vorgeschlagen. Diese Autoren verwenden eine "gesplittete Dreiecksverteilung", was einer Mischungsverteilung entspricht, bei der zwei deutlich separierte, symmetrisch um 1 positionierte, Intervalle positive Wahrscheinlichkeits-Masse besitzen.

zu bestimmen.

5.1.1 Additive Überlagerung

Im Fall der additiven Überlagerung wählen wir¹⁵

$$\mu_1 = \mu_m, \mu_2 = -\mu_m \quad (5-1)$$

und

$$\sigma_1^2 = \sigma_2^2 = \sigma_m^2 \quad (5-2)$$

sowie

$$\alpha_1 = \alpha_2 = \frac{1}{2} \quad (5-3)$$

Dann ergibt sich aus den allgemeinen Formeln (2-3) und (2-4)

$$E[U] = 0$$

und

$$V[U] = \sigma_m^2 + \mu_m^2,$$

d.h. $\sigma_u^2 = \sigma_m^2 + \mu_m^2$ in der alternativen Schreibweise.

5.1.2 Multiplikative Überlagerung

Im Fall der multiplikativen Überlagerung wählen wir die alternative Spezifikation

$$\mu_1 = 1 + \delta_m, \mu_2 = 1 - \delta_m \quad (5-4)$$

für die Erwartungswerte. Im Übrigen werden die Spezifikationen des additiven Falls beibehalten: Für die Varianzen gilt (5-2) und für die Gewichte gilt (5-3).

Dann ergibt sich aus den allgemeinen Formeln (2-3) und (2-4)

$$E[U] = 1$$

und

$$V[U] = \sigma_m^2 + \delta_m^2,$$

d.h. $\sigma_u^2 = \sigma_m^2 + \delta_m^2$ in der alternativen Schreibweise.

5.2 Mehrdimensionaler Fall

Entsprechend dem Vorgehen im Abschnitt 5.1 beschreiben wir auch für mehrdimensionale Verteilungen die speziellen Parametrisierungen im Fall der Überlagerung und wenden dann die Ergebnisse aus Abschnitt 3 an, um die ersten und zweiten Momente zu bestimmen. Auch hier beschränken wir uns auf den Fall von $k = 2$ Mischungskomponenten.

¹⁵Der Index "m" steht für Mischung und bezeichnet Parameter, die die einzelnen Komponenten der Mischungsverteilung charakterisieren, wenn sie für alle Mischungskomponenten identisch gewählt werden.

5.2.1 Additive Überlagerung

Im Fall der additiven Überlagerung wählen wir

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu}_m, \boldsymbol{\mu}_2 = -\boldsymbol{\mu}_m \quad (5-5)$$

wobei $\boldsymbol{\mu}$ nicht notwendigerweise identische Erwartungswerte für alle r Komponenten besitzt. Ferner soll

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_m \quad (5-6)$$

sowie

$$\alpha_1 = \alpha_2 = \frac{1}{2} \quad (5-7)$$

gelten.

Dann ergibt sich aus den allgemeinen Formeln (3-2) und (3-3)

$$E[\mathbf{U}] = \mathbf{0}$$

und

$$\text{cov}[\mathbf{U}] = \boldsymbol{\Sigma}_m + \boldsymbol{\mu}_m \boldsymbol{\mu}_m' \quad (5-8)$$

Man beachte, daß die spezielle Parametrisierung der Mischungsverteilung mit $\boldsymbol{\mu}_1 = -\boldsymbol{\mu}_2 = \boldsymbol{\mu}_m$ eine "Vergrößerung" der Kovarianzmatrix bewirkt!

5.2.2 Multiplikative Überlagerung

Im Fall der multiplikativen Überlagerung wählen wir die alternative Spezifikation

$$\boldsymbol{\mu}_1 = \boldsymbol{\nu} + \boldsymbol{\delta}_m, \boldsymbol{\mu}_2 = \boldsymbol{\nu} - \boldsymbol{\delta}_m \quad (5-9)$$

für die Erwartungswerte, wobei $\boldsymbol{\delta}_m$ nicht notwendigerweise identische Werte für alle r Komponenten besitzt. Im Übrigen werden die Spezifikationen des additiven Falls beibehalten: Für die Varianzen gilt (5-2) und für die Gewichte gilt (5-3).

Dann ergibt sich aus den allgemeinen Formeln (2-3) und (2-4)

$$E[\mathbf{U}] = \boldsymbol{\nu}$$

und

$$\text{cov}[\mathbf{U}] = \boldsymbol{\Sigma}_m + \boldsymbol{\delta}_m \boldsymbol{\delta}_m' \quad (5-10)$$

Falls alle Komponenten von $\boldsymbol{\delta}_m$ identisch sind, d.h. falls

$$\boldsymbol{\delta}_m = \delta_m \boldsymbol{\nu} \quad (5-11)$$

gilt, ergibt sich für die Kovarianzmatrix die spezielle Form

$$\text{cov}[\mathbf{U}] = \boldsymbol{\Sigma}_m + \delta_m^2 \boldsymbol{\nu} \boldsymbol{\nu}' \quad (5-11)$$

Falls wir weiterhin noch

$$\boldsymbol{\Sigma}_m = \sigma_m^2 \mathbf{I} \quad (5-12)$$

verlangen, erhalten wir

$$\text{cov}[\mathbf{U}] = \sigma_m^2 \mathbf{I} + \delta_m^2 \boldsymbol{\nu} \boldsymbol{\nu}' \quad (5-12)$$

Dieses spezielle Ergebnis wird uns später bei der Betrachtung des Höhne-Verfahrens wieder begegnen. Siehe insbesondere Abschnitt 6.2. Aber auch hier sollte bereits die Bemerkung gemacht werden, daß die Mischungsverteilung mit zwei Komponenten, die mittels des Parameters δ_m symmetrische Zu- und Abschläge erzeugt, dazu führt, daß die r Komponenten miteinander korreliert sind. Denn man kann die Kovarianzmatrix auch wie folgt schreiben:

$$\text{cov}[\mathbf{U}] = (\sigma_m^2 + \delta_m^2) ((1 - \rho) \mathbf{I} + \rho \boldsymbol{\mu} \boldsymbol{\mu}')$$

Dabei steht ρ für den zuvor abgeleiteten Korrelationskoeffizienten

$$\rho = \frac{\delta_m^2}{\sigma_m^2 + \delta_m^2} .$$

Wohlgermerkt: Obwohl die Verteilungen der einzelnen Mischungskomponenten skalare Kovarianzmatrix besitzen, weist die Mischungsverteilung selbst positive Korrelation aus. Siehe (5-12) in Verbindung mit (5-2).

6 Das Höhne-Verfahren

Das Verfahren, das Jörg Höhne vorgeschlagen hat, ist vor allem mit der Absicht entwickelt worden, bei der Überlagerung von **verschiedenen** Merkmalen deren Beziehung zueinander, insbesondere das proportionale Verhältnis, einigermaßen zu erhalten. Hier geht es vor allem um die formale Darstellung des Verfahrens und um die Frage, ob dies Verfahren mit einer Mischungsverteilung äquivalent ist bzw. inwiefern es sich davon unterscheidet.

Wir betrachten zunächst das "Höhne-Verfahren" für ein einziges Merkmal, dann aber auch für mehrere Merkmale gemeinsam.

6.1 Ein einziges Merkmal

6.1.1 Additiver Fall

Wir betrachten das Modell

$$U = \mu D + \varepsilon , \quad (6-1)$$

wobei μ ein Parameter mit beliebigem Wert ist. Für die diskrete Zufallsvariable D gilt

$$D = \begin{cases} +1 & \text{mit Wahrscheinlichkeit } \alpha \\ -1 & \text{mit Wahrscheinlichkeit } 1 - \alpha \end{cases} \quad (6-2)$$

und die stetige Zufallsvariable ε ist wie folgt spezifiziert:

$$E[\varepsilon] = 0 , \quad V[\varepsilon] = \sigma_\varepsilon^2 \quad (6-3)$$

wobei die Annahme der Normalverteilung üblicherweise hinzutritt. Ferner wird angenommen, daß D und ε stochastisch unabhängig sind.

Erwartungswert und Varianz von D sind durch

$$E[D] = 2\alpha - 1 , \quad V[D] = 4\alpha(1 - \alpha)$$

gegeben. Wenn $\alpha = 1/2$ gilt, ist der Erwartungswert gleich 0 und die Varianz gleich 1. Außerdem ist die Wahrscheinlichkeitsdichtefunktion von D durch

$$h(d) = \alpha^{\frac{1+d}{2}} (1 - \alpha)^{\frac{1-d}{2}} \quad (6-4)$$

für $d \in \{+1, -1\}$ gegeben.

Für Erwartungswert und Varianz von U gemäß (6-1) ergibt sich dann

$$E[U] = \mu(2\alpha - 1) \quad (6-5)$$

sowie

$$V[U] = \mu^2 V[D] + V[U] = 4\alpha(1 - \alpha)\mu^2 + \sigma_\varepsilon^2 \quad (6-6)$$

Insbesondere erhalten wir für den "symmetrischen" Fall, d.h. wenn $\alpha = 1/2$ gilt,

$$E[U] = 0 \quad , \quad V[U] = \mu^2 + \sigma_\varepsilon^2 \quad ,$$

was dem Resultat für die Mischungsverteilung entspricht, wenn dort ebenfalls der "symmetrische" Fall betrachtet wird.

6.1.2 Multiplikativer Fall

Im Fall der multiplikativen Überlagerung verwenden wir statt (6-1) den Ansatz

$$U = (1 + \delta D) + \varepsilon \quad (6-7)$$

und erhalten in diesem Fall für Erwartungswert und Varianz

$$E[U] = 1 + \delta(2\alpha - 1) \quad (6-8)$$

sowie

$$V[U] = \delta^2 V[D] + V[U] = 4\alpha(1 - \alpha)\delta^2 + \sigma_\varepsilon^2 \quad (6-9)$$

Für $\alpha = 1 - \alpha = 1/2$ erhalten wir das Resultat

$$E[U] = 1 \quad , \quad V[U] = \delta^2 + \sigma_\varepsilon^2 \quad ,$$

was dem Resultat für die Mischungsverteilung (im multiplikativen Fall) entspricht, wenn dort ebenfalls der "symmetrische" Fall betrachtet wird.

6.1.3 Dichtefunktion im Höhe-Ansatz

Obwohl die Gleichheit der ersten und zweiten Momente dafür spricht, daß Mischungsverteilung und Höhe-Verfahren äquivalent sind, ist dies erst nachgewiesen, wenn die Gleichheit der Dichtefunktionen gezeigt ist. Dazu gehen wir wie folgt vor: Wir betrachten zunächst die bedingte Dichte von U gegeben D , multiplizieren diese Dichte mit der Randdichte von D , um die gemeinsame Dichte zu bestimmen und "integrieren" dann die Zufallsvariable D aus, um die Randdichte von U zu erhalten. Dabei ist in diesem Fall allerdings wegen der Diskretheit von D eine Summation vorzunehmen. Wir unterstellen im folgenden zusätzlich

$$\varepsilon \sim N(0, \sigma_\varepsilon^2) \quad . \quad (6-10)$$

Additiver Fall: Für gegebenes D ist U in (6-1) normalverteilt, d.h.

$$U|D = d \sim N(\mu d, \sigma_\varepsilon^2) \quad ,$$

wobei N die Dichtefunktion der Normalverteilung symbolisieren soll. Für die gemeinsame Dichte von U und D erhalten wir dann

$$g(u, d) = h(d) \cdot N(\mu d, \sigma_\varepsilon^2)$$

mit $h(d)$ aus (6-4). Dann erhalten wir die Randdichte von U durch folgende Formel:

$$f(u) = \sum_{d \in \{+1, -1\}} g(u, d) = \sum_{d \in \{+1, -1\}} \alpha^{\frac{1+D}{2}} (1 - \alpha)^{\frac{1-D}{2}} \cdot N(\mu d, \sigma_\varepsilon^2)$$

oder

$$f(u) = \alpha N(\mu_1, \sigma_\varepsilon^2) + (1 - \alpha) N(\mu_2, \sigma_\varepsilon^2) \quad . \quad (6-11)$$

Dies entspricht exakt der Dichte der Mischungsverteilung in (2-2), allerdings für den Spezialfall identischer Varianzen.

Multiplikativer Fall: Für gegebenes D ist U in (6-7) normalverteilt, d.h.

$$U|D = d \sim N(1 + \delta d, \sigma_\varepsilon^2) \quad .$$

Wie im additiven Fall erhalten wir für die gemeinsame Dichte von U und D dann

$$g(u, d) = h(d) \cdot N(1 + \delta d, \sigma_\varepsilon^2)$$

mit $h(d)$ aus (6-4) und für die Randdichte von U ergibt sich

$$f(u) = \sum_{d \in \{+1, -1\}} g(u, d) = \sum_{d \in \{+1, -1\}} \alpha^{\frac{1+D}{2}} (1 - \alpha)^{\frac{1-D}{2}} \cdot N(1 + \delta d, \sigma_\varepsilon^2)$$

oder

$$f(u) = \alpha N(1 + \delta, \sigma_\varepsilon^2) + (1 - \alpha) N(1 - \delta, \sigma_\varepsilon^2) \quad . \quad (6-12)$$

Dies entspricht exakt der Dichte der Mischungsverteilung in (2-2), allerdings für den Spezialfall identischer Varianzen.

6.2 Mehrere Merkmale gemeinsam

Höhne hat vorgeschlagen, alle Merkmale mit demselben (multiplikativen) Zu- bzw. Abschlag δ zu versehen und dann bei den einzelnen Merkmalen noch eine (merkmals-spezifische) Überlagerung durchzuführen. Dahinter steht die Idee, daß damit die proportionalen Beziehungen zwischen den einzelnen Merkmalen bzw. die daraus erzeugten Verhältniszahlen annähernd gleich bleiben.¹⁶ Dies soll im Folgenden dargestellt werden, sodann soll nachgewiesen werden, daß auch in diesem Fall Äquivalenz zum statistischen Modell der Mischungsverteilung besteht.

Da Höhnes Vorschlag nur die **multiplikative** Alternative betrachtet, werden wir uns im Text auf diese Variante beschränken. AAbschnitt 6.3 stellt ergänzend auch die additive Variante kurz dar, weil sie in Abschnitt 9 ebenfalls benötigt wird.

¹⁶Darauf werden wir in Abschnitt 8 näher eingehen.

6.2.1 Formale Darstellung des Verfahrens

Es wird der r -dimensionale Vektor \mathbf{U} betrachtet, für dessen j -te Komponente gelten soll:

$$U_j = (1 + \delta D) + \varepsilon_j, \quad j = 1, \dots, r, \quad (6-13)$$

Für alle r Komponenten gemeinsam kann man das auch schreiben als

$$\mathbf{U} = (1 + \delta D)\boldsymbol{\iota} + \boldsymbol{\varepsilon}, \quad (6-14)$$

wobei δ ein Parameter mit beliebigem Wert ist. Genau wie in Abschnitt 6.1.1 ist die diskrete Zufallsvariable D durch

$$D = \begin{cases} +1 & \text{mit Wahrscheinlichkeit } \alpha \\ -1 & \text{mit Wahrscheinlichkeit } 1 - \alpha \end{cases}$$

gegeben (siehe (6-2)) und der r -dimensionale stetige Zufallsvektor $\boldsymbol{\varepsilon}$ ist entsprechend (6-3) spezifiziert:

$$E[\boldsymbol{\varepsilon}] = \mathbf{0}, \quad \text{cov}[\boldsymbol{\varepsilon}] = \sigma_\varepsilon^2 \mathbf{I} \quad (6-15)$$

wobei die Annahme der Normalverteilung üblicherweise hinzutritt. Genau wie in Abschnitt 6.1.1 wird die stochastische Unabhängigkeit von D und $\boldsymbol{\varepsilon}$ unterstellt.

Für den Erwartungswert der Zufallskomponente U_j aus dem Zufallsvektors \mathbf{U} ergibt sich (wie im Fall der Betrachtung eines einzigen Merkmals)

$$E[U_j] = 1 + \delta(2\alpha - 1), \quad j = 1, \dots, r \quad (6-16)$$

oder auch kompakter

$$E[\mathbf{U}] = (1 + \delta(2\alpha - 1))\boldsymbol{\iota}. \quad (6-17)$$

Im 'symmetrischen' Fall ($\alpha = 1/2$) ergibt sich

$$E[\mathbf{U}] = \boldsymbol{\iota}.$$

Bei der Bestimmung der Kovarianzmatrix ist folgende Beobachtung wesentlich: Da alle r Komponenten die Zufallsvariable D enthalten, sind die Komponenten von \mathbf{U} miteinander korreliert, selbst wenn die Komponenten $varepsilon_j$ des Zufallsvektors $\boldsymbol{\varepsilon}$ unkorreliert sind!! Um die Kovarianz zu bestimmen, betrachten wir zunächst den Erwartungswert des Produktes $U_i U_j$. Dafür erhalten wir¹⁷

$$E_D[E_{\boldsymbol{\varepsilon}|D}[U_i U_j] | D] = E_D[(1 + \delta D)^2] = E_D[1 + 2\delta D + \delta^2 D^2] = 1 + 2\delta(2\alpha - 1) + \delta^2$$

und für die Kovarianz selbst erhalten wir

$$\text{cov}[U_i U_j] = 1 + 2\delta(2\alpha - 1) + \delta^2 - (1 + \delta(2\alpha - 1))^2 = \delta^2 - \delta^2(2\alpha - 1)^2 \geq 0.$$

Diese Kovarianz ist also gleich Null nur für $\alpha = 0$ bzw. $\alpha = 1$, d.h. wenn nur eine einzige Mischungskomponente verwendet wird. Ansonsten ist diese Kovarianz (und damit auch die entsprechende Korrelation) **stets positiv!** Im hier besonders interessierenden 'symmetrischen' Fall ($\alpha = 1/2$) ist die Kovarianz gleich δ^2 .

¹⁷Wir verwenden

$$E[D^2] = V[D] + (E[D])^2 = 4\alpha(1 - \alpha) + (2\alpha - 1)^2 = 1.$$

Für die Varianz von U_j gilt ferner (siehe die Ergebnisse in Abschnitt 6.1.2)

$$V[U_j] = 4\alpha(1-\alpha)\delta^2 + \sigma_\varepsilon^2, \quad ,$$

und speziell für den symmetrischen Fall ($\alpha = 1 - \alpha = 1/2$) erhalten wir

$$V[U] = \delta^2 + \sigma_\varepsilon^2 \quad .$$

Damit erhalten wir in diesem speziellen Fall für den Korrelationskoeffizienten eine besonders einfache Struktur:

$$\text{corr}[U_i U_j] = \frac{\delta^2}{\delta^2 + \sigma_\varepsilon^2} \quad \text{falls } \alpha = \frac{1}{2} \quad .$$

Die durch das Verfahren erzeugte Korrelation ist also umso größer, je größer der gemeinsame Zuschlag δ und je kleiner die Überlagerungsvarianz σ_ε^2 ist!

Als Kovarianzmatrix erhalten wir aus obigen Ergebnissen (beispielsweise für $r = 4$ Merkmale)

$$\text{cov}[\mathbf{U}] = \begin{pmatrix} 4\alpha(1-\alpha)\delta^2 + \sigma_\varepsilon^2 & \delta^2 - \delta^2(2\alpha-1)^2 & \delta^2 - \delta^2(2\alpha-1)^2 & \delta^2 - \delta^2(2\alpha-1)^2 \\ \delta^2 - \delta^2(2\alpha-1)^2 & 4\alpha(1-\alpha)\delta^2 + \sigma_\varepsilon^2 & \delta^2 - \delta^2(2\alpha-1)^2 & \delta^2 - \delta^2(2\alpha-1)^2 \\ \delta^2 - \delta^2(2\alpha-1)^2 & \delta^2 - \delta^2(2\alpha-1)^2 & 4\alpha(1-\alpha)\delta^2 + \sigma_\varepsilon^2 & \delta^2 - \delta^2(2\alpha-1)^2 \\ \delta^2 - \delta^2(2\alpha-1)^2 & \delta^2 - \delta^2(2\alpha-1)^2 & \delta^2 - \delta^2(2\alpha-1)^2 & 4\alpha(1-\alpha)\delta^2 + \sigma_\varepsilon^2 \end{pmatrix},$$

die sich im 'symmetrischen' Fall auf die einfache Form

$$\text{cov}[\mathbf{U}] = \sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\nu} \boldsymbol{\nu}' = (\sigma_\varepsilon^2 + \delta^2) ((1-\rho) \mathbf{I} + \rho \boldsymbol{\nu} \boldsymbol{\nu}') \quad (6-18)$$

reduziert, wobei ρ für den zuvor abgeleiteten Korrelationskoeffizienten

$$\rho = \frac{\delta^2}{\delta^2 + \sigma_\varepsilon^2}$$

steht. Man beachte, daß diese Korrelation stets positiv ist!

6.2.2 Ableitung der Dichtefunktion

Wir können festhalten, daß sich für Erwartungswert und Varianz unter bestimmten Bedingungen dieselben Resultate ergeben wie bei der Mischungsverteilung: In beiden Fällen gilt

$$E[\mathbf{U}] = \boldsymbol{\nu} \quad \text{und} \quad \text{cov}[\mathbf{U}] = \sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\nu} \boldsymbol{\nu}' \quad (6-19)$$

Dabei wird dieses Ergebnis in der Mischungsverteilung erreicht, wenn für die Mittelwertvektoren und die Kovarianzmatrizen in allen Mischungskomponenten folgende sehr spezielle Annahme getroffen wird:

$$\boldsymbol{\mu}_j = \delta \boldsymbol{\nu} \quad \text{und} \quad \boldsymbol{\Sigma}_j = \sigma_\varepsilon^2 \mathbf{I} \quad (6-20)$$

und der 'symmetrische Fall'

$$\alpha = \frac{1}{2}$$

sowohl bei der Mischungsverteilung als auch im Höhneverfahren unterstellt wird.

Es bleibt zu prüfen, ob auch im mehrdimensionalen Fall die Dichtefunktion des Höhneverfahrens mit der Definition der Mischungsverteilung gemäß (3-1) übereinstimmt. Bei dieser Analyse gehen wir wie im eindimensionalen Fall vor. Siehe dazu Abschnitt 6.1.3. Wieder beschränken wir uns auf den hier relevanten Fall der **multiplikativen** Überlagerung.

Wir betrachten zunächst die bedingte Dichte von \mathbf{U} gegeben D , multiplizieren diese Dichte mit der Randdichte von D , um die gemeinsame Dichte zu bestimmen und "integrieren" dann die Zufallsvariable D aus, um die Randdichte von U zu erhalten. Dabei unterstellen wir im folgenden wieder zusätzlich Normalverteilung:

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}) \quad . \quad (6-21)$$

Für gegebenes D ist \mathbf{U} in (6-14) normalverteilt, d.h.

$$\mathbf{U} | (D = d) \sim N((1 + \delta d) \boldsymbol{\nu}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}) \quad .$$

Für die gemeinsame Dichte von \mathbf{U} und D ergibt sich

$$g(\mathbf{u}, d) = h(d) \cdot N((1 + \delta d) \boldsymbol{\nu}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I})$$

mit $h(d)$ aus (6-4) und für die Randdichte von \mathbf{U} ergibt sich

$$f(\mathbf{u}) = \sum_{d \in \{+1, -1\}} g(\mathbf{u}, d) = \sum_{d \in \{+1, -1\}} \alpha^{\frac{1+d}{2}} (1 - \alpha)^{\frac{1-d}{2}} \cdot N((1 + \delta d) \boldsymbol{\nu}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I})$$

oder

$$f(\mathbf{u}) = \alpha N((1 + \delta) \boldsymbol{\nu}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}) + (1 - \alpha) N((1 - \delta) \boldsymbol{\nu}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}) \quad . \quad (6-22)$$

Dies entspricht exakt der Dichte der Mischungsverteilung in (3-1), allerdings für den Spezialfall identischer (skalarer) Kovarianzmatrizen.

6.3 Die (spezielle) Höhne-Spezifikation im additiven Fall

In Abschnitt 6.2 haben wir nur die multiplikative Variante des "speziellen" Höhne-Verfahrens betrachtet. Hier ergänzen wir diese Ergebnisse für den "additiven" Fall:

$$\mathbf{U} = \mu D \boldsymbol{\nu} + \boldsymbol{\varepsilon} \quad , \quad (6-23)$$

wobei μ ein Parameter mit beliebigem Wert ist. Genau wie in Abschnitt 6.2 ist die diskrete Zufallsvariable D durch

$$D = \begin{cases} +1 & \text{mit Wahrscheinlichkeit } \alpha \\ -1 & \text{mit Wahrscheinlichkeit } 1 - \alpha \end{cases}$$

gegeben (siehe (6-2)) und der r -dimensionale stetige Zufallsvektor $\boldsymbol{\varepsilon}$ ist entsprechend (6-3) spezifiziert:

$$E[\boldsymbol{\varepsilon}] = \mathbf{0} \quad , \quad cov[\boldsymbol{\varepsilon}] = \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I} \quad (6-24)$$

wobei die Annahme der Normalverteilung üblicherweise hinzutritt. Genau wie in Abschnitt 6.2 wird die stochastische Unabhängigkeit von D und $\boldsymbol{\varepsilon}$ unterstellt.

Im symmetrischen Fall ergibt sich für den additiven Fall

$$E[\mathbf{U}] = \mathbf{0}$$

und

$$\text{cov}[\mathbf{U}] = \sigma_\varepsilon^2 \mathbf{I} + \mu^2 \boldsymbol{\nu}' = (\sigma_\varepsilon^2 + \mu^2) ((1 - \rho) \mathbf{I} + \rho \boldsymbol{\nu}' \boldsymbol{\nu}) \quad (6-25)$$

wobei

$$\rho = \frac{\mu^2}{\mu^2 + \sigma_\varepsilon^2}$$

gilt. Man beachte, daß diese Korrelation stets positiv ist!

6.4 Flexible Höhne-Spezifikation

Im den folgenden Kapiteln werden wir die spezielle Höhne-Spezifikation (6-13) bzw. (6-14), die einen identischen Zuschlag δ für alle $r = K$ Merkmale (bei ausschließlicher Überlagerung der Regressoren bzw. $r = K + 1$ bei zusätzlicher Überlagerung auch der abhängigen Variablen) fordert, betrachten, die wir hier für einen bestimmten Beobachtungspunkt i wie folgt schreiben:

$$\mathbf{u}\langle i \rangle = (1 + \delta D\langle i \rangle) \boldsymbol{\nu} + \boldsymbol{\varepsilon}\langle i \rangle, \quad i = 1, \dots, n. \quad (6-26)$$

wobei die stochastischen Spezifikationen der Zufallsvariablen D und des Zufallsvektors $\boldsymbol{\varepsilon}$ für jedes i durch (6-2) und (6-3) gegeben sind.

Im Gegensatz dazu geht die "flexiblere" Spezifikation davon aus, daß für jedes der $r = K$ (bzw. $r = K + 1$) Merkmale das Vorzeichen des Zuschlags, nicht aber der Zuschlag selbst über die Merkmale variiert wird. Dies schreiben wir als

$$\mathbf{u}\langle i \rangle = (\boldsymbol{\nu} + \delta \mathbf{D}\langle i \rangle) + \boldsymbol{\varepsilon}\langle i \rangle, \quad i = 1, \dots, n. \quad (6-27)$$

wobei \mathbf{D} jetzt ein r -dimensionaler Zufallsvektor ist, dessen Komponenten jeweils die Spezifikation (6-2) erfüllen und die stochastisch voneinander unabhängig sind. Man kann zeigen, daß Erwartungswerte und Varianzen dieser Spezifikation denen von (6-26) bzw. (6-14) entsprechen. Dagegen sind alle Kovarianzen gleich Null.: Insbesondere gilt für den "symmetrischen" Fall

$$E[u_j] = 1, \quad V[u_j] = \delta^2 + \sigma_\varepsilon^2, \quad \text{cov}[u_j, u_k] = 0 \text{ für alle } j, k, j \neq k.$$

Die stochastischen Eigenschaften dieser Spezifikation werden in den beiden folgenden Unterabschnitten beschrieben! Insbesondere wird gezeigt, daß in diesem Fall keine Äquivalenz zu einer Mischungsverteilung besteht.

6.4.1 Multiplikativer Fall

Die "flexible" Variante von (6-26) wird hier ohne den Beobachtungs-Index $\langle i \rangle$ geschrieben:

$$\mathbf{U} = (\boldsymbol{\nu} + \delta \mathbf{D}) + \boldsymbol{\varepsilon} \quad . \quad i = 1, \dots, n. \quad (6-28)$$

Dabei ist \mathbf{D} jetzt ein r -dimensionaler Zufallsvektor, dessen Komponenten jeweils die Spezifikation (6-2) erfüllen und die stochastisch voneinander unabhängig sind. Um die Eigenschaften der dadurch implizierten Verteilung sowie der ersten und zweiten Momente abzuleiten, gehe ich wie in den Abschnitten 6.2.1 und 6.2.2 vor.

Da die einzelnen Komponenten von \mathbf{U} wegen der Struktur

$$U_j = 1 + \delta D_j + \varepsilon_j$$

voneinander stochastisch unabhängig sind und für die Varianz von U_j

$$\text{Var}[U_j] = \delta^2 \text{Var}[D_j] + \sigma_\varepsilon^2$$

gilt, ergibt sich für Erwartungswert und Kovarianzmatrix

$$E[\mathbf{U}] = (1 + \delta(2\alpha - 1))\boldsymbol{\iota} \quad , \quad \text{cov}[\mathbf{U}] = (4\alpha(1 - \alpha)\delta^2 + \sigma_\varepsilon^2) \mathbf{I} . \quad (6-29)$$

Im "symmetrischen" Fall, also für $\alpha = 0,5$, erhalten wir

$$E[\mathbf{U}] = \boldsymbol{\iota} \quad , \quad \text{cov}[\mathbf{U}] = (\delta^2 + \sigma_\varepsilon^2) \mathbf{I} .$$

Diese Ergebnisse entsprechen für Erwartungswerte und Varianzen, aber **nicht für Kovarianzen** den Ergebnissen der spezielleren Höhne-Spezifikation (6-26). Der grundlegende Unterschied ist, daß bei der "flexiblen" Spezifikation alle Komponenten unkorreliert sind.

Für die Ableitung der Dichtefunktion von \mathbf{U} gehen wir wie folgt vor: Wir betrachten zunächst die bedingte Dichte von \mathbf{U} gegeben der Vektor (!!!) \mathbf{D} , multiplizieren diese Dichte mit der Randdichte von \mathbf{D} , um die gemeinsame Dichte zu bestimmen und "integrieren" dann den Zufallsvektor \mathbf{D} aus, um die Randdichte von \mathbf{U} zu erhalten. Dabei unterstellen wir im folgenden wieder zusätzlich Normalverteilung:

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}) \quad . \quad (6-30)$$

Für gegebenen Vektor \mathbf{D} ist jede Komponente von \mathbf{U} in (6-28) normalverteilt. Zunächst gilt:

$$U_j | D_j \sim N(1 + \delta d_j, \sigma_\varepsilon^2) \quad .$$

Allerdings gilt wegen der stochastischen Unabhängigkeit der einzelnen D_j auch

$$U_j | D_j, D_k, k \neq j \sim N(1 + \delta d_j, \sigma_\varepsilon^2) \quad .$$

Außerdem sind wegen der stochastischen Unabhängigkeit der ε_j die einzelnen U_j (gegeben D_j) stochastisch voneinander unabhängig (bedingte Unabhängigkeit). Dann können wir für die multivariate bedingte Verteilung von \mathbf{U} gegeben \mathbf{D} schreiben:

$$\begin{pmatrix} U_1 | D_1 \\ \vdots \\ U_r | D_r \end{pmatrix} \sim N(\boldsymbol{\iota} + \delta \mathbf{D}, \sigma_\varepsilon^2 \mathbf{I}) = \prod_{j=1}^r N_j(1 + \delta D_j, \sigma_\varepsilon^2) \quad .$$

Dabei bezeichnet N_j die eindimensionale Dichte der Normalverteilung.

Außerdem gilt wegen der stochastischen Unabhängigkeit der Komponenten von \mathbf{D}

$$h(\mathbf{d}) = h(d_1) \cdot \dots \cdot h(d_r) \quad .$$

Demnach ergibt sich für die gemeinsame Dichte von \mathbf{U} und \mathbf{D}

$$g(\mathbf{u}, \mathbf{d}) = h(\mathbf{d}) \cdot N(\boldsymbol{\iota} + \delta \mathbf{d}, \sigma_\varepsilon^2 \mathbf{I}) = \prod_{j=1}^r \alpha^{\frac{1+d_j}{2}} (1 - \alpha)^{\frac{1-d_j}{2}} \cdot N(1 + \delta d_j, \sigma_\varepsilon^2) \quad .$$

Um die Randdichte des Vektors \mathbf{U} zu bestimmen, müssen die einzelnen Komponenten von \mathbf{D} 'heraussummiert' werden. Allerdings bedeutet dies, daß alle möglichen Ereignisse von \mathbf{D} betrachtet werden müssen. Ich illustriere dies für den einfachsten Fall von $r = 2$ Merkmalen: In diesem Fall sind für den Vektor (D_1, D_2) folgende Ereignisse möglich:

$$\{(+1, +1), (+1, -1), (-1, +1), (-1, -1)\} \quad .$$

Dann erhalten wir bei Summation über diese Ereignisse

$$\begin{aligned} f(\mathbf{u}) &= \sum_{d_1, d_2 \in \{+1, -1\}} \prod_{j=1}^r \alpha^{\frac{1+d_j}{2}} (1-\alpha)^{\frac{1-d_j}{2}} \cdot N(1 + \delta d_j, \sigma_\varepsilon^2) \\ &= \{\alpha N(1 + \delta, \sigma_\varepsilon^2) + \alpha N(1 + \delta, \sigma_\varepsilon^2)\} \{\alpha N(1 + \delta, \sigma_\varepsilon^2) + (1-\alpha) N(1 - \delta, \sigma_\varepsilon^2)\} \times \\ &\quad \times \{(1-\alpha) N(1 - \delta, \sigma_\varepsilon^2) + \alpha N(1 + \delta, \sigma_\varepsilon^2)\} \{(1-\alpha) N(1 - \delta, \sigma_\varepsilon^2) + (1-\alpha) N(1 - \delta, \sigma_\varepsilon^2)\} \end{aligned}$$

Es ist evident, daß dies nicht der allgemeinen Formel der Mischungsverteilung in Abschnitt 3 entspricht.

Die allgemeine Formulierung der Dichtefunktion lautet demnach

$$f(\mathbf{u}) = \sum_{d_1, d_2, \dots, d_r \in \{+1, -1\}} \prod_{j=1}^r \alpha^{\frac{1+d_j}{2}} (1-\alpha)^{\frac{1-d_j}{2}} \cdot N(1 + \delta d_j, \sigma_\varepsilon^2) \quad . \quad (6-31)$$

6.4.2 Additiver Fall

Im additiven Fall schreiben wir die "flexible" Spezifikation als

$$\mathbf{u} = \mu \mathbf{D} + \varepsilon \quad . \quad i = 1, \dots, n. \quad (6-32)$$

Dabei ist \mathbf{D} jetzt ein r -dimensionaler Zufallsvektor, dessen Komponenten jeweils die Spezifikation (6-2) erfüllen und die stochastisch voneinander unabhängig sind.

Im symmetrischen Fall ergibt sich für den additiven Fall

$$E[\mathbf{U}] = \mathbf{0}$$

und

$$cov[\mathbf{U}] = (\mu^2 + \sigma_\varepsilon^2) \mathbf{I}. \quad (6-33)$$

Auch hier wird wieder aus der skalaren Kovarianzmatrix deutlich, daß die flexible Spezifikation nicht mit einer Mischungsverteilung vereinbar ist, weil bei letzterer stets die Kovarianzen ungleich Null sind.

7 Anonymisierung mittels stochastischer Überlagerung

Im folgenden bezeichnet Y die originale Variable und Y^a die anonymisierte Variable, entsprechend bezeichnen \mathbf{Y} und \mathbf{Y}^a die betreffenden Vektoren im multivariaten Fall. Im folgenden gehen wir davon aus, daß alle Zufallsvariablen stetig sind.

Es soll gelten

$$E[Y] = \mu_y \text{ sowie } V[Y] = \sigma_y^2 \quad (7-1)$$

bzw. im multivariaten Fall für k -dimensionalen Vektor \mathbf{Y}

$$E[\mathbf{Y}] = \boldsymbol{\mu}_y \text{ sowie } cov[\mathbf{Y}] = \boldsymbol{\Sigma}_{yy} \quad (7-2)$$

7.1 Additive Überlagerung

Im eindimensionalen Fall gilt

$$Y^a = Y + U \quad (7-3)$$

wobei U eine stetige Zufallsvariable mit Erwartungswert 0 und Varianz $\sigma_u^2 > 0$ ist. Im multivariaten Fall erhalten wir

$$\mathbf{Y}^a = \mathbf{Y} + \mathbf{U} \quad (7-4)$$

wobei der Zufallsvektor \mathbf{U} Erwartungswert $\mathbf{0}$ und Kovarianzmatrix $cov[\mathbf{U}]$ besitzt.

Für Erwartungswert und Varianz bzw. Kovarianzmatrix ergibt sich

$$E[Y^a] = \mu_y \text{ und } V[Y^a] = \sigma_y^2 + \sigma_u^2 \quad (\text{eindimensional})$$

und

$$E[\mathbf{Y}^a] = \boldsymbol{\mu}_y \text{ und } cov[\mathbf{Y}^a] = cov[\mathbf{Y}] + cov[\mathbf{U}] \quad (\text{mehrdimensional}) \quad .$$

Falls die Überlagerung mittels der Mischungsverteilung aus Abschnitt 5.1.1 (bzw. mit dem Höhne-Verfahren) erfolgt, ergibt sich - im eindimensionalen Fall - für die Varianz der überlagerten Variablen

$$V[Y^a] = \sigma_y^2 + \sigma_m^2 + \mu_m^2 \quad (7-5)$$

wobei μ_m und σ_m^2 die Parameter aus der Mischungsverteilung sind, siehe (5-1).

Im multivariaten Fall ergibt sich aus (5-8) als Kovarianzmatrix der überlagerten Variablen

$$cov[\mathbf{Y}^a] = cov[\mathbf{Y}] + \boldsymbol{\Sigma}_m + \boldsymbol{\mu}_m \boldsymbol{\mu}_m' \quad , \quad (7-6)$$

wobei $\boldsymbol{\mu}_m$ und $\boldsymbol{\Sigma}_m$ die Parameter aus der Mischungsverteilung bezeichnen. Siehe Abschnitt 5.2.1.

7.2 Multiplikative Überlagerung in univariaten Verteilungen

Üblicherweise wird man bei der multiplikativen Überlagerung unterstellen, daß sowohl die zu überlagernde Variable Y als auch die Überlagerungsvariable U nichtnegativ sind, um zu garantieren, daß die anonymisierte Variable Y^a ebenfalls nichtnegativ ist. Dieser Aspekt wird hier nicht im Einzelnen berücksichtigt.

Für den eindimensionalen Fall ergibt sich hier

$$Y^a = Y U \quad (7-7)$$

wobei U eine **positive** stetige Zufallsvariable mit Erwartungswert 1 und Varianz $\sigma_u^2 > 0$ ist, die stochastisch unabhängig von Y erzeugt wird. Alternativ könnte man schreiben:

$$E[U|Y] = 1 \quad \text{und} \quad V[U|Y] = \sigma_u^2 .$$

Es ergibt sich aus (7-7) für den bedingten Erwartungswert

$$E[Y^a|Y] = 1 \cdot Y$$

und damit für den unbedingten Erwartungswert

$$E[Y^a] = \mu_y$$

Ferner erhalten wir für die bedingte Varianz

$$V[Y^a|Y] = Y^2 V[U|Y] = Y^2 \sigma_u^2$$

und damit für die unbedingte Varianz¹⁸

$$V[Y^a] = \sigma_y^2 + \sigma_u^2 (\mu_y^2 + \sigma_y^2) . \quad (7-8)$$

Falls die Überlagerung mittels der Mischungsverteilung aus Abschnitt 5.1.2 (bzw. mit dem Höhe-Verfahren) erfolgt, ergibt sich - im eindimensionalen Fall - für die Varianz der überlagerten Variablen

$$V[Y^a] = \sigma_y^2 + (\sigma_m^2 + \delta_m^2) (\mu_y^2 + \sigma_y^2) \quad (7-9)$$

wobei δ_m der Mittelwertparameter und σ_m^2 der Varianzparameter aus der Mischungsverteilung sind, siehe (5-4) und (5-2).

7.3 Multiplikative Überlagerung in multivariaten Verteilungen

Im multivariaten Fall sind der sogenannte "skalare" bzw. "konstante" Fall und der Fall der "eigentlichen multiplikativen" Überlagerung zu unterscheiden. Im ersten Fall werden alle Komponenten des Merkmals-Vektors mit **demselden Faktor** multiplikativ überlagert. Dies hat den Vorteil, daß bei der Bildung von Quotienten dieser Faktor tendenziell (siehe jedoch Genaueres in Abschnitt 8) eliminiert wird, d.h. daß die Quotienten durch diese Art der multiplikativen Überlagerung nicht verändert werden!! Formal ergibt sich in diesem Fall

$$\mathbf{Y}^a = U \mathbf{Y} \quad (\text{konstante multiplikative Überlagerung}) \quad (7-10)$$

Dabei soll für die (nichtnegative) Störvariable U gelten:

$$E[U|\mathbf{Y}] = 1 \quad \text{und} \quad V[U|\mathbf{Y}] = \sigma_u^2 .$$

¹⁸Es wird die Formel für bedingte und unbedingte Varianzen verwendet:

$$V[Y^a] = E_y\{V[Y^a|Y]\} + V_y\{E[Y^a|Y]\} .$$

Im Fall der eigentlichen multiplikativen Überlagerung erhalten wir dagegen

$$\mathbf{Y}^a = \mathbf{U} \odot \mathbf{Y} \quad (\text{eigentliche multiplikative Überlagerung}) \quad (7-11)$$

wobei \mathbf{U} ein **positiver** Zufalls-Vektor mit

$$E[\mathbf{U}] = \boldsymbol{\iota} \quad \text{und} \quad \text{cov}[\mathbf{U}] \text{ eine beliebige positiv definite Matrix}$$

ist und \odot das Hadamard-Produkt bezeichnet.¹⁹

Es sind aber auch allgemeine Spezifikationen denkbar.

- Beispielsweise lassen sich anonymisierte Variablen als Linearkombinationen von überlagerten Originalvariablen bilden:

$$\mathcal{U} \mathbf{Y} = \begin{pmatrix} U_{11}Y_1 + U_{12}Y_2 + \dots + U_{1r}Y_r \\ U_{21}Y_1 + U_{22}Y_2 + \dots + U_{2r}Y_r \\ \vdots \\ U_{r1}Y_1 + U_{r2}Y_2 + \dots + U_{rr}Y_r \end{pmatrix}$$

betrachten, wobei \mathcal{U} eine nichtnegative ($r \times r$) Zufalls-Matrix ist.

- Schließlich kann man auch noch

$$\mathcal{Y}^a = \mathcal{U} \odot \mathcal{Y} \quad (7-12)$$

betrachten, wobei die Zufalls-Matrizen \mathcal{Y} und \mathcal{U} identische Dimensionen besitzen. Wir bezeichnen dies als variablenspezifische multiplikative Überlagerung bei Matrizen.

7.3.1 Überlagerung mit Hilfe eines konstanten Faktors

Im "konstanten" Fall ergibt sich aus (7-10) für den bedingten Erwartungswert

$$E[\mathbf{Y}^a | \mathbf{Y}] = \mathbf{1} \cdot \mathbf{Y}$$

und damit für den unbedingten Erwartungswert

$$E[\mathbf{Y}^a] = \boldsymbol{\mu}_y$$

Ferner erhalten wir für die bedingte Varianz-Kovarianz-Matrix folgendes: Für die Varianz der einzelnen Komponenten ergibt sich entsprechend dem univariaten Fall

$$V[Y_k^a | \mathbf{Y}] = \sigma_u^2 Y_k^2 \text{ für alle } k$$

sowie

$$\text{cov}[Y_j^a, Y_k^a | \mathbf{Y}] = \sigma_u^2 Y_j Y_k \text{ für alle } j, k \text{ } j \neq k .$$

Insgesamt gilt also für die bedingte Kovarianzmatrix

$$\text{cov}[\mathbf{Y}^a | \mathbf{Y}] = \sigma_u^2 \mathbf{Y} \mathbf{Y}' .$$

¹⁹Siehe Marshall und Olkin (1979 Seite 257-8) für einige Eigenschaften des Hadamard-Produktes.

Zur Berechnung der unbedingten Kovarianzmatrix verwenden wir die Formel für bedingte Erwartungswerte und Kovarianzen und erhalten

$$\begin{aligned} \text{cov}[\mathbf{Y}^a] &= E[\text{cov}[\mathbf{Y}^a|\mathbf{Y}]] + \text{cov}[E[\mathbf{Y}^a|\mathbf{Y}]] = \sigma_u^2 E[\mathbf{Y} \mathbf{Y}'] + \text{cov}[\mathbf{Y}] \\ &= \sigma_u^2 \{ \text{cov}[\mathbf{Y}] + \boldsymbol{\mu}_y \boldsymbol{\mu}_y' \} + \text{cov}[\mathbf{Y}] \end{aligned} \quad (7-13)$$

mit

$$\sigma_u^2 = \sigma_m^2 + \delta_m^2 \quad .$$

Man beachte die Entsprechung zu (7-8) für den eindimensionalen Fall.

7.3.2 Überlagerung mit Hilfe unkorrelierter Störvariablen

Eine komplexere Formel ergibt sich für den Fall der "eigentlichen multiplikativen" Überlagerung. Dabei spielt offensichtlich auch die Annahme über die Varianzen und Kovarianzen von \mathbf{U} eine Rolle. Wir untersuchen zunächst den einfachsten Fall, in dem alle U_k identische Varianz σ_u^2 besitzen und miteinander unkorreliert sind.

Aus (7-11) ergibt sich für den bedingten Erwartungswert

$$E[\mathbf{Y}^a|\mathbf{Y}] = \mathbf{1} \cdot \mathbf{Y}$$

und damit für den unbedingten Erwartungswert

$$E[\mathbf{Y}^a] = \boldsymbol{\mu}_y$$

Ferner erhalten wir für die bedingte Varianz-Kovarianz-Matrix folgendes: Für die Varianz der einzelnen Komponenten ergibt sich entsprechend dem univariaten Fall

$$V[Y_k|\mathbf{Y}] = \sigma_u^2 Y_k^2 \text{ für alle } k$$

sowie

$$\text{cov}[Y_j^a, Y_k^a|\mathbf{Y}] = \text{cov}[U_j Y_j, U_k Y_k|\mathbf{Y}] = \text{cov}[U_j, U_k] Y_j Y_k = 0 \text{ für alle } j, k \ j \neq k,$$

weil die U_j unkorreliert sind. Insgesamt gilt also für die bedingte Kovarianzmatrix eine Diagonalmatrix (im Unterschied zum konstanten Fall!!):

$$\text{cov}[\mathbf{Y}^a|\mathbf{Y}] = \sigma_u^2 \begin{pmatrix} Y_1^2 & & & & \\ & Y_2^2 & & & \\ & & \ddots & & \\ & & & Y_{k-1}^2 & \\ & & & & Y_k^2 \end{pmatrix} .$$

Zur Berechnung der unbedingten Kovarianzmatrix verwenden wir wieder

$$\text{cov}[\mathbf{Y}^a] = E[\text{cov}[\mathbf{Y}^a|\mathbf{Y}]] + \text{cov}[E[\mathbf{Y}^a|\mathbf{Y}]]$$

und erhalten in diesem Fall

$$\begin{aligned}
 cov[\mathbf{Y}^a] &= \sigma_u^2 E \left[\begin{pmatrix} Y_1^2 & & & \\ & Y_2^2 & & \\ & & \ddots & \\ & & & Y_{k-1}^2 & \\ & & & & Y_k^2 \end{pmatrix} \right] + cov[\mathbf{Y}] \\
 &= \sigma_u^2 \begin{pmatrix} \sigma_{y(1)}^2 + \mu_{y(1)}^2 & & & & \\ & \sigma_{y(2)}^2 + \mu_{y(2)}^2 & & & \\ & & \ddots & & \\ & & & \sigma_{y(k-1)}^2 + \mu_{y(k-1)}^2 & \\ & & & & \sigma_{y(k)}^2 + \mu_{y(k)}^2 \end{pmatrix} \\
 &\quad + cov[\mathbf{Y}] \qquad \qquad \qquad (7-14)
 \end{aligned}$$

mit

$$\sigma_u^2 = \sigma_m^2 + \delta_m^2 \quad .$$

Man beachte auch hier die Entsprechung zu (7-8) im eindimensionalen Fall.

Demnach stimmen die Varianzen des "eigentlichen" Falls mit den Varianzen des "konstanten" Falls überein, nicht aber die Kovarianzen. Für diese läßt sich eine Abschätzung wie folgt vornehmen: Falls alle r Merkmale positiv miteinander korreliert sind und alle Erwartungswerte positiv sind (was bei nichtnegativen Merkmalen automatisch der Fall ist), ergeben sich für den eigentlichen Fall **kleinere** Kovarianzen als im konstanten Fall. Man vergleiche (7-13) mit (7-14). Wegen der identischen Varianzen bedeutet dies eine geringere Korrelation im Fall der "eigentlichen" Überlagerung.

7.3.3 Überlagerung mit Hilfe von korrelierten Störvariablen

Wir untersuchen nun den später interessierenden Fall, daß die Störvariablen U_j des Zufallsvektors \mathbf{U} miteinander korreliert sind. Die Kovarianzmatrix insgesamt soll jetzt allgemein durch $cov[\mathbf{U}]$ gegeben sein. Dies betrifft insbesondere die Überlagerung mit Hilfe einer Mischungsverteilung, auf die wir im folgenden Abschnitt 7.3.4 eingehen, was dann jedoch eine spezielle Struktur der Kovarianzmatrix zur Folge hat.

Wieder gilt für den unbedingten Erwartungswert

$$E[\mathbf{Y}^a] = \boldsymbol{\mu}_y$$

Ferner erhalten wir für die bedingte Varianz-Kovarianz-Matrix wegen

$$V[Y_k|\mathbf{Y}] = \sigma_u^2 Y_k^2 \text{ für alle } k$$

sowie

$$cov[Y_j^a, Y_k^a|\mathbf{Y}] = cov[U_j Y_j, U_k Y_k|\mathbf{Y}] = cov[U_j, U_k] Y_j Y_k \text{ für alle } j, k \ j \neq k ,$$

und somit für die bedingte Kovarianzmatrix insgesamt:

$$cov[\mathbf{Y}^a|\mathbf{Y}] = cov[\mathbf{U}] \odot \mathbf{Y} \mathbf{Y}' \quad .$$

Zur Berechnung der unbedingten Kovarianzmatrix verwenden wir wieder

$$\text{cov}[\mathbf{Y}^a] = E[\text{cov}[\mathbf{Y}^a|\mathbf{Y}]] + \text{cov}[E[\mathbf{Y}^a|\mathbf{Y}]]$$

und erhalten in diesem Fall

$$\begin{aligned} \text{cov}[\mathbf{Y}^a] &= \text{cov}[\mathbf{U}] \odot E[\mathbf{Y}\mathbf{Y}'] + \text{cov}[\mathbf{Y}] \\ &= \text{cov}[\mathbf{U}] \odot (\text{cov}[\mathbf{Y}] + \boldsymbol{\mu}_y\boldsymbol{\mu}_y') + \text{cov}[\mathbf{Y}] \quad . \end{aligned} \quad (7-15)$$

7.3.4 Überlagerung mit Hilfe der Mischungsverteilung

Wir haben in Abschnitt 5.2 und speziell für multiplikative Überlagerung im Unterabschnitt 5.2.2 festgestellt, daß die Annahme der Unkorreliertheit nicht mit der Mischungsverteilung verträglich ist, wenn man sie wie dort beschrieben für die Überlagerung einsetzt. Deshalb müssen wir, wenn wir eine Mischungsverteilung für die stochastische Überlagerung unterstellen, die Annahme aufgeben, daß die Überlagerungsfaktoren miteinander unkorreliert sind. Andererseits interessiert uns vor allem die Auswirkung der speziellen Parametrisierung, die wir als Höhenverfahren bezeichnen: Alle Merkmale werden mit demselben Zuschlag versehen, d.h. gemäß (6-13) soll gelten:

$$U_j = 1 + \delta D + \varepsilon_j \quad , j = 1, \dots, r ,$$

Ferner sollen die Störterme ε_j den "klassischen" Annahmen genügen:

$$E[\boldsymbol{\varepsilon}] = \mathbf{0} \quad , \quad \text{cov}[\boldsymbol{\varepsilon}] = \sigma_\varepsilon^2 \mathbf{I}$$

Siehe dazu (6-3). Dies führt zu folgender Kovarianzmatrix:

$$\text{cov}[\mathbf{U}] = \sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu}\boldsymbol{\mu}' = \frac{1}{\sigma_\varepsilon^2 + \delta^2} ((1 - \rho) \mathbf{I} + \rho \boldsymbol{\mu}\boldsymbol{\mu}') ,$$

wobei ρ durch

$$\rho = \frac{\delta^2}{\delta^2 + \sigma_\varepsilon^2}$$

gegeben ist. Siehe (6-18). Demnach gilt für Varianzen und Kovarianzen:

$$V[U_j] = \sigma_\varepsilon^2 + \delta^2 \quad \text{und} \quad \text{cov}[U_j, U_k] = \delta^2 \quad . \quad (7-16)$$

Dies entspricht der speziellen multiplikativen Überlagerung, wie sie in Abschnitt 5.2.2 beschrieben wurde.²⁰ Für diesen speziellen Fall wollen wir deshalb im folgenden Erwartungsvektor und Kovarianzmatrix des Vektors \mathbf{Y}^a bestimmen.

Es gilt wieder

$$Y_j^a = U_j Y_j \quad , j = 1, \dots, r ,$$

Deshalb ergibt sich aus (7-11) für den bedingten Erwartungswert

$$E[\mathbf{Y}^a|\mathbf{Y}] = \mathbf{1} \cdot \mathbf{Y}$$

und damit für den unbedingten Erwartungswert

$$E[\mathbf{Y}^a] = \boldsymbol{\mu}_y$$

²⁰Dort wurden die etwas abweichenden Bezeichnungen σ_m^2 und δ_m gewählt.

Ferner erhalten wir für die bedingte Varianz-Kovarianz-Matrix folgendes: Für die bedingte Varianz der einzelnen Komponenten ergibt sich

$$V[Y_j|\mathbf{Y}] = V[U_j] \cdot Y_j^2 \text{ für alle } j$$

und für die bedingten Kovarianzen erhalten wir

$$\text{cov}[Y_j^a, Y_k^a|\mathbf{Y}] = \text{cov}[U_j Y_j, U_k Y_k|\mathbf{Y}] = \text{cov}[U_j, U_k] \cdot Y_j Y_k \text{ für alle } j, k \text{ } j \neq k,$$

Insgesamt ergibt sich also für die bedingte Kovarianzmatrix :

$$\text{cov}[\mathbf{Y}^a|\mathbf{Y}] = \text{cov}[\mathbf{U}] \odot \mathbf{Y} \mathbf{Y}' = (\sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu} \boldsymbol{\mu}') \odot \mathbf{Y} \mathbf{Y}'$$

Zur Berechnung der unbedingten Kovarianzmatrix verwenden wir wieder

$$\text{cov}[\mathbf{Y}^a] = E[\text{cov}[\mathbf{Y}^a|\mathbf{Y}]] + \text{cov}[E[\mathbf{Y}^a|\mathbf{Y}]]$$

und erhalten in diesem Fall

$$\text{cov}[\mathbf{Y}^a] = (\sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu} \boldsymbol{\mu}') \odot (\text{cov}[\mathbf{Y}] + \boldsymbol{\mu}_y \boldsymbol{\mu}_y') + \text{cov}[\mathbf{Y}] \quad (7-17)$$

Alternativ könnte man in dieser Formel die Parameter σ_m^2 und δ_m verwenden. Die Entsprechung zu (7-8) im eindimensionalen Fall ist hier nur noch bedingt gegeben.

7.3.5 Zusammenfassung

Die Ergebnisse der Unterabschnitte von Abschnitt 7.3 lassen sich wie folgt zusammenfassen: In allen drei Fällen wird auch mittels \mathbf{Y}^a der Mittelwertvektor erwartungstreu geschätzt. Für die Kovarianzmatrix von \mathbf{Y}^a ergibt sich in allen Fällen eine Kovarianzmatrix mit additiver Struktur: Zu der Kovarianzmatrix $\text{cov}[\mathbf{Y}]$ wird eine zweite Matrix addiert, in der die ersten und zweiten Momente von \mathbf{Y} sowie die Parameter der Überlagerungsvariablen eine Rolle spielen, d.h. es tritt eine zweite positiv semi-definite Matrix zu der Matrix $\text{cov}[\mathbf{Y}]$ hinzu, was den (unterschiedlich großen) Effizienzverlust charakterisiert.

7.4 Korrelation bei multiplikativer Überlagerung

Es soll nun die Auswirkung der Überlagerung auf den Korrelationskoeffizienten eingehender untersucht werden. Dabei konzentrieren wir uns auf die Resultate aus Abschnitt 7.3.4, in dem Mischungsverteilungen vom Höhne-Typ verwendet werden. Insbesondere ist hier die daraus resultierende Formel (7-17) für die Kovarianzmatrix relevant. Aus dieser Formel ergibt sich (unter der Annahme konstanter Korrelation ρ für alle Paare)

$$\begin{aligned} V[Y_j^a] &= (\delta^2 + \sigma_\varepsilon^2) (\sigma_j^2 + \mu_j^2) + \sigma_j^2 \\ \text{cov}[Y_j^a, Y_k^a] &= \delta^2 (\rho \sigma_j \sigma_k + \mu_j \mu_k) + \rho \sigma_j \sigma_k \end{aligned} \quad (7-18)$$

sowie

$$\text{corr}[Y_j^a, Y_k^a] = \frac{\rho (1 + \delta^2) \sigma_j \sigma_k + \delta^2 \mu_j \mu_k}{\sqrt{\{(\delta^2 + \sigma_\varepsilon^2) (\sigma_j^2 + \mu_j^2) + \sigma_j^2\} \{(\delta^2 + \sigma_\varepsilon^2) (\sigma_k^2 + \mu_k^2) + \sigma_k^2\}}} \quad (7-19)$$

Aus dieser Formel folgt, daß die Korrelation von Y_j^a und Y_k^a eine monotone Beziehung zur Korrelation der entsprechenden Originalvariablen aufweist: Je größer der Wert des Korrelationskoeffizienten ρ auf der Skala zwischen -1 und +1 ist, desto größer ist auch der Wert des Korrelationskoeffizienten für die überlagerten Variablen. Allerdings ist im Fall $\rho = 0$ (Originalvariablen unkorreliert) die Korrelation der überlagerten Variablen im Allgemeinen ungleich Null und hängt von den Vorzeichen der beiden Erwartungswerte μ_j und μ_k ab. Setzt man andererseits den Zählerausdruck von (7-19) gleich Null, so ergibt sich daraus

$$\rho = - \frac{\delta^2 \mu_j \mu_k}{(1 + \delta^2) \sigma_j \sigma_k} ,$$

d.h. die Originalvariablen werden nur dann ebenfalls Null-Korrelation aufweisen, wenn mindestens einer der Erwartungswerte identisch Null ist, was sicher ein singuläres Ereignis ist.

Aus der Formel (7-19) sieht man auch, daß nur die Veränderung des **Verhältnisses** der Erwartungswerte und Varianzen für die Korrelation der anonymisierten Variablen eine Rolle spielt. Denn man kann dafür auch schreiben:

$$\text{corr}[Y_j^a, Y_k^a] = \frac{\rho(1 + \delta^2) \gamma_j \gamma_k + \delta^2}{\sqrt{\left\{(\delta^2 + \sigma_\varepsilon^2) (\gamma_j^2 + 1) + \gamma_j^2\right\} \left\{(\delta^2 + \sigma_\varepsilon^2) (\gamma_k^2 + 1) + \gamma_k^2\right\}}} \quad (7-20)$$

mit

$$\gamma_j \equiv \frac{\sigma_j}{\mu_j} .$$

Falls $\mu_j > 0$ gilt, bezeichnet man γ_j als Variationskoeffizienten.

Da es für das Vorzeichen der Korrelation nur auf den Zähler ankommt, können wir schreiben:

$$\rho > - \frac{\delta^2}{(1 + \delta^2)} \frac{\mu_j}{\sigma_j} \frac{\mu_k}{\sigma_k} \implies \rho^a > 0 ,$$

d.h. es hängt nur von δ sowie γ_j und γ_k ab, ob beide Korrelationskoeffizienten das gleiche oder unterschiedliche Vorzeichen besitzen. Da δ in der Anonymisierungspraxis stets kleiner als 1 und üblicherweise einen Wert von etwa 0,10 aufweist (und damit $\delta^2 = 0,01$ gilt), ist die Abweichung im Vorzeichen zwischen den beiden Korrelationskoeffizienten allerdings in der Praxis nicht sehr häufig!

Ferner ergeben sich interessante Resultate für die Fälle, daß die Originalvariablen exakt positiv bzw. exakt negativ korreliert sind. Im Fall $\rho = 1$ erhalten wir aus (7-19)

$$\text{corr}[Y_j^a, Y_k^a]_{\rho=1} = \frac{(1 + \delta^2) \sigma_j \sigma_k + \delta^2 \mu_j \mu_k}{\sqrt{\left\{(\delta^2 + \sigma_\varepsilon^2) (\sigma_j^2 + \mu_j^2) + \sigma_j^2\right\} \left\{(\delta^2 + \sigma_\varepsilon^2) (\sigma_k^2 + \mu_k^2) + \sigma_k^2\right\}}} \quad (7-21)$$

und im Fall $\rho = -1$

$$\text{corr}[Y_j^a, Y_k^a]_{\rho=-1} = \frac{\delta^2 \mu_j \mu_k - (1 + \delta^2) \sigma_j \sigma_k}{\sqrt{\left\{(\delta^2 + \sigma_\varepsilon^2) (\sigma_j^2 + \mu_j^2) + \sigma_j^2\right\} \left\{(\delta^2 + \sigma_\varepsilon^2) (\sigma_k^2 + \mu_k^2) + \sigma_k^2\right\}}} \quad (7-22)$$

Vor allem die Formel (7-22) besagt, daß die Korrelation der überlagerten Variablen durchaus positiv sein kann, obwohl die Originalvariablen eine exakte negative Korrelation aufweisen! Dies wird tendenziell dann der Fall sein, wenn die Erwartungswerte im Verhältnis zu den Standardabweichungen groß sind.

Ferner wird bei exakter **positiver** Korrelation der Originalvariablen die Korrelation der überlagerten Variablen stets kleiner als 1 sein. Dies sieht man deutlicher, wenn man

$$\mu_k = \alpha \mu_j \quad \text{und} \quad \sigma_k = \beta \sigma_j$$

setzt und dann die rechte Seite von (7-21) wie folgt schreibt:

$$\frac{(1 + \delta^2) \beta \sigma_j^2 + \delta^2 \alpha \mu_j^2}{\sqrt{\left\{ (1 + \delta^2 + \sigma_\varepsilon^2) \sigma_j^2 + (\delta^2 + \sigma_\varepsilon^2) \mu_j^2 \right\} \left\{ (1 + \delta^2 + \sigma_\varepsilon^2) \beta^2 \sigma_j^2 + (\delta^2 + \sigma_\varepsilon^2) \alpha^2 \mu_j^2 \right\}}}$$

Für $\alpha = \beta = 1$ ist unmittelbar zu sehen, daß für $\sigma_\varepsilon^2 > 0$ der Ausdruck kleiner als 1 ist.²¹

8 Einfluß der Überlagerung auf Quotienten

8.1 Einleitende Bemerkungen

Das Verfahren, das Höhne vorschlug, soll die - sinnvolle - gemeinsame Anonymisierung mehrerer Merkmale ermöglichen. Das bedeutet insbesondere, daß die (proportionale) Beziehung zwischen verschiedenen Merkmalen nicht zu sehr verändert wird. Statistisch gesehen bedeutet dies, daß der Erwartungswert der Zufallsvariablen

$$Z = \frac{Y_1}{Y_2} \quad \text{mit} \quad W\{Y_2 > 0\} = 1$$

nicht zu sehr vom Erwartungswert der Zufallsvariablen

$$Z^a = \frac{Y_1^a}{Y_2^a} \quad \text{mit} \quad W\{Y_2^a > 0\} = 1$$

abweicht.²²

Dabei ist zu beachten, daß ganz allgemein der Erwartungswert von Z nicht gleich dem Verhältnis der Erwartungswerte der beiden Zufallsvariablen ist, d.h. es gilt ganz allgemein

$$E[Z] \neq \frac{E[Y_1]}{E[Y_2]}$$

und nur im Fall der **stochastischen Unabhängigkeit** von Y_1 und Y_2 läßt sich der Erwartungswert wie folgt schreiben:

$$E[Z] = E[Y_1] E\left[\frac{1}{Y_2}\right],$$

²¹Beweis für den allgemeinen Fall steht noch aus!

²²Man könnte auch die stärkere Forderung aufstellen, daß die Verteilung von Z nicht zu sehr von der Verteilung von Z^a abweichen soll.

was allerdings meist auch nicht sehr gut analysierbar ist.²³

Allerdings läßt sich - für den allgemeinen Fall - folgende Approximation angeben:²⁴

$$E\left[\frac{Y_1}{Y_2}\right] \approx \frac{E[Y_1]}{E[Y_2]} - \frac{1}{(E[Y_2])^2} \text{cov}[Y_1, Y_2] + \frac{E[Y_1]}{(E[Y_2])^3} V[Y_2] \quad (8-1)$$

Demnach ist der Erwartungswert von Z nur annähernd gleich dem Verhältnis der beiden merkmals-spezifischen Erwartungswerte! Die Abweichung hängt einerseits davon ab, ob die beiden Merkmale miteinander korreliert sind. Man beachte, daß dieser Term mit negativem Vorzeichen eingeht. Das bedeutet, daß eine positive Korrelation sich reduzierend auf den Erwartungswert von Z auswirkt. Falls die Korrelation Null ist, fällt der zweite Term auf der rechten Seite weg! Außerdem sind Größenordnung von Erwartungswert und Varianz der Nenner-Variablen Y_2 für das Ausmaß der Abweichung verantwortlich!

Wir haben in Abschnitt 6 gezeigt, daß der Ansatz von Höhne mit der Verwendung der Mischungsverteilung äquivalent ist. Deshalb werde ich im folgenden die Formulierung aus dem genannten Abschnitt bei der Beantwortung der Frage verwenden, welche Beziehung zwischen Z und Z^a bei stochastischer Überlagerung mit einer Mischungsverteilung besteht, weil sie anschaulicher ist als die direkte Darstellung der Mischungsverteilung in den Abschnitten 2 und 3. Dabei betrachten wir im folgenden nur den für unser Projekt relevanten Fall der **multiplikativen Überlagerung**.

8.2 Das Verhältnis von Y_1 zu Y_2 bei multiplikativer Überlagerung

Im Folgenden betrachten wir die beiden Zufallsvariablen bzw. Merkmale Y_1 und Y_2 , die jeweils multiplikativ mit den Faktoren

$$U_j = (1 + \delta_j D_j) + \varepsilon_j, \quad j = 1, 2, \quad (8-2)$$

überlagert werden, d.h. die anonymisierten Variablen sind durch

$$Y_j^a = Y_j \cdot U_j = Y_j \cdot (1 + \delta_j D_j + \varepsilon_j), \quad j = 1, 2, \quad (8-3)$$

gegeben.

Der Vorschlag von Höhne, der bereits im ersten Anonymisierungsprojekt realisiert wurde²⁵, läuft darauf hinaus, daß die folgenden Restriktionen eingeführt werden:

$$\delta_1 = \delta_2 = \delta \quad (8-4)$$

und

$$D_1 = D_2 \quad \text{mit Wahrscheinlichkeit } 1. \quad (8-5)$$

²³Von Nutzen ist oft die Jensensche Ungleichung, nach der ganz allgemein für eine nichtnegative Zufallsvariable W gilt:

$$E\left[\frac{1}{W}\right] \geq \frac{1}{E[W]}.$$

Siehe z.B. Mood Graybill Boes 1974 S. 72.

²⁴Mood Graybill Boes (1974 S. 181). Eine umfassende Analyse der Verteilung von Quotienten liefert das Buch von Springer (1979). Diesen Hinweis verdanke ich Robert Jung.

²⁵Siehe Ronning et al (2005).

Dann ergibt sich für das Verhältnis der anonymisierten Variablen

$$Z^a = \frac{Y_1 \cdot ((1 + \delta D) + \varepsilon_1)}{Y_2 \cdot ((1 + \delta D) + \varepsilon_2)} \quad (8-6)$$

Für den Erwartungswert eines Verhältnisses von Zufallsvariablen, in dem Fall von

$$E \left[\frac{Y_1 \cdot ((1 + \delta D) + \varepsilon_1)}{Y_2 \cdot ((1 + \delta D) + \varepsilon_2)} \right],$$

ist, wie oben gezeigt, die Analyse komplex. Wir beschränken uns im folgenden auf den Fall, daß Y_1 und Y_2 stochastisch unabhängig sind. Da jedoch die Zufallsvariable D im Zähler und Nenner von (8-6) steht, sind auch unter dieser Unabhängigkeitsannahme die beiden Terme nicht stochastisch unabhängig. Jedoch können wir den **bedingten Erwartungswert** für gegebenes D wie folgt bestimmen:

$$\begin{aligned} E[Z^a|D] &= E[Y_1] \cdot ((1 + \delta D)) E \left[\frac{1}{Y_2 \cdot ((1 + \delta D) + \varepsilon_2)} \right] \\ &= E[Y_1] \cdot ((1 + \delta D)) E \left[\frac{1}{Y_2} \right] \cdot E \left[\frac{1}{((1 + \delta D) + \varepsilon_2)} \right] \quad (8-7) \end{aligned}$$

Für den zweiten und dritten Erwartungswert auf der rechten Seite kann man folgende Approximation angeben²⁶:

$$E \left[\frac{1}{Y_2} \right] \approx \frac{1}{E[Y_2]} + \frac{1}{(E[Y_2])^3} V[Y_2]$$

sowie

$$E \left[\frac{1}{((1 + \delta D) + \varepsilon_2)} \right] \approx \frac{1}{(1 + \delta D)} + \frac{1}{(1 + \delta D)^3} \sigma_\varepsilon^2$$

Dabei hängt die Verzerrung vom Erwartungswert sowie von der Varianz der jeweiligen Zufallsvariablen ab! Bei der Variablen Y_2 ist dies eine empirische Frage; die Störvariable ε wird im allgemeinen bei der Anonymisierung nur eine kleine Varianz aufweisen, was eine nur unbedeutende Verzerrung impliziert.

Für den Erwartungswert in (8-7) ergibt sich damit

$$E[Z^a|D] \approx (1 + \delta D) E[Y_1] \left(\frac{1}{E[Y_2]} + \frac{1}{(E[Y_2])^3} V[Y_2] \right) \left(\frac{1}{(1 + \delta D)} + \frac{1}{(1 + \delta D)^3} \sigma_\varepsilon^2 \right) \quad (8-8)$$

Dafür können wir in grober Approximation auch schreiben:

$$\begin{aligned} E[Z^a|D] &\approx (1 + \delta D) E[Y_1] \left(\frac{1}{E[Y_2]} \right) \left(\frac{1}{(1 + \delta D)} \right) + \mathbf{R} \\ &= \frac{E[Y_1]}{E[Y_2]} + \mathbf{R}, \quad (8-9) \end{aligned}$$

wobei im allgemeinen für den Restterm

$$\mathbf{R} = (1 + \delta D) E[Y_1] \left(\frac{1}{E[Y_2]} \frac{1}{(1 + \delta D)^3} \sigma_\varepsilon^2 + \frac{1}{(E[Y_2])^3} V[Y_2] \frac{1}{(1 + \delta D)} + \frac{1}{(E[Y_2])^3} V[Y_2] \frac{1}{(1 + \delta D)^3} \sigma_\varepsilon^2 \right)$$

²⁶Siehe Mood Graybill Boes 1974 S. 181 sowie Formel (8-1) weiter oben. Man beachte, daß wir im Zähler des Ausdrucks die Konstante 1 stehen haben. Somit ist die Kovarianz zwischen Y_2 und 1 gleich Null!

unabhängig davon, ob $D = 1$ oder $D = -1$ gilt, die Ungleichung

$$\mathbf{R} > 0$$

gelten wird, d.h. der berechnete bedingte Erwartungswert von Z^a wird im allgemeinen das Verhältnis

$$\frac{E[Y_1]}{E[Y_2]}$$

überschätzen. Andererseits kürzt sich der Faktor $(1 + \delta D)$ umso exakter heraus (und spielt in \mathbf{R} keine Rolle), je kleiner die Überlagerungs(rest)varianz σ_ε^2 gewählt wird. Siehe (8-7).

Wenn sich der Faktor $(1 + \delta D)$ herauskürzt, dann ist der bedingte Erwartungswert $E[Z^a|D]$ nicht mehr von D abhängig, d.h. dann ist der bedingte Erwartungswert mit dem unbedingten Erwartungswert $E[Z^a]$ identisch. Die hier präsentierten Überlegungen gelten jedoch nur für den Fall, daß Y_1 und Y_2 stochastisch unabhängig sind. In der Simulationsstudie, die im folgenden Unterabschnitt präsentiert wird, werden wir auch den Fall betrachten, daß die beiden Variablen korreliert sind.

8.3 Simulationsergebnisse

In diesem Abschnitt sollen die zuvor theoretisch analysierten Sachverhalte durch einige numerische - simulierte - Beispiele illustriert werden. Dazu betrachte ich zwei Zufallsvariable X und Y sowie deren Verhältnis Z . Ferner betrachte ich die multiplikativ überlagerten Variablen X^a und Y^a sowie das daraus resultierende Verhältnis Z^a . Siehe dazu die Ausführungen in Abschnitt 8.1.

Für die beiden Ausgangsvariablen nehme ich eine gemeinsame Normalverteilung wie folgt an:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix} \right]$$

Zwecks Anonymisierung werden diese beiden Variablen wie folgt überlagert:

$$\begin{aligned} X^a &= (1 + \delta D + \varepsilon_x) X \\ Y^a &= (1 + \delta D + \varepsilon_y) Y \end{aligned}$$

mit

$$\begin{pmatrix} \varepsilon_x \\ \varepsilon_y \end{pmatrix} \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I})$$

Die beiden Annahmen der Normalverteilung sind insofern problematisch, als dadurch nicht garantiert ist, daß (a) die Ausgangsvariablen stets nichtnegativ und (b) die Überlagerungen stets positiv sind. Ich habe dies dadurch vermieden, daß ich die Erwartungswerte μ_x und μ_y relativ groß gewählt habe.²⁷

In den folgenden Beispielen wurde stets

$$\delta = 0.10 \quad \text{und} \quad \sigma_\varepsilon = 0.03$$

gewählt. Dabei werden drei verschiedene Ausgangsszenarios zugrunde gelegt:

²⁷Eine formal bessere Lösung wäre, sowohl für die Ausgangsvariablen als auch für die Überlagerungsreste Lognormalverteilungen zu verwenden. Dabei sollten die Ausgangsvariablen im Intervall $[0, \infty)$ und die Überlagerungen im Intervall $[-(1 + \delta D), \infty)$ variieren.

Tabelle 8.1: Simulationsergebnisse

μ_x	μ_y	σ_x	σ_y	ρ	ρ^a	r^a	$\frac{\mu_x}{\mu_y}$	\bar{z}	$\sigma_{\bar{z}}$	\bar{z}^a	$\sigma_{\bar{z}^a}$	bias
Szenario A												
10.00	10.00	2.00	2.00	-0.999	-0.5914	-0.5836	1.0000	1.0961	0.0454	1.0959	0.0453	0.0780
10.00	10.00	2.00	2.00	0.000	0.1948	0.1933	1.0000	1.0385	0.0335	1.0391	0.0337	0.0400
10.00	10.00	2.00	2.00	0.999	0.9810	0.9806	1.0000	1.0001	0.0010	1.0010	0.0039	0.0000
Szenario B												
25.00	25.00	2.00	2.00	-0.999	0.2039	0.2212	1.0000	1.0152	0.0130	1.0177	0.0142	0.0128
25.00	25.00	2.00	2.00	0.000	0.5757	0.5732	1.0000	1.0071	0.0111	1.0074	0.0123	0.0064
25.00	25.00	2.00	2.00	0.999	0.9475	0.9464	1.0000	1.0000	0.0004	1.0021	0.0045	0.0000
Szenario C												
10.00	20.00	2.00	4.00	-0.999	-0.5914	-0.6022	0.5000	0.5461	0.0266	0.5461	0.0264	0.0340
10.00	20.00	2.00	4.00	0.000	0.1948	0.2072	0.5000	0.5257	0.0160	0.5268	0.0160	0.0200
10.00	20.00	2.00	4.00	0.999	0.9810	0.9804	0.5000	0.5001	0.0006	0.5011	0.0020	0.0000

Bemerkungen

ρ^a gibt den theoretischen Korrelationskoeffizienten gemäß (7-19) an.

r^a gibt den empirischen Korrelationskoeffizienten über alle Wiederholungen gemittelt an.

\bar{z} gibt den geschätzten Wert von Z über alle Wiederholungen gemittelt an.

$\sigma_{\bar{z}}$ gibt die entsprechende Standardabweichung an.

\bar{z}^a gibt den geschätzten Wert von Z^a über alle Wiederholungen gemittelt an.

$\sigma_{\bar{z}^a}$ gibt die entsprechende Standardabweichung an.

"bias" gibt die Summe aus dem zweiten und dritten Summanden auf der rechten Seite von (8-1) an.

A $\mu_x = \mu_y$; $\sigma_x^2 = \sigma_y^2$, Variationskoeffizienten klein

B $\mu_x = \mu_y$; $\sigma_x^2 = \sigma_y^2$, Variationskoeffizienten sehr klein

C $\mu_x < \mu_y$; $\sigma_x^2 < \sigma_y^2$,

Für alle drei Szenarios wurde die Korrelation wie folgt variiert: $\rho \in \{-0.999, 0, +0.999\}$
 Alle Beispiele wurden mit $n = 100$ Beobachtungen und 50 Wiederholungen simuliert. Die Ergebnisse sind in der Tabelle 8.1 zusammengefaßt.²⁸

Als erstes fällt auf, daß bei allen drei Szenarios die Korrelation ρ^a der überlagerten Variablen stark von derjenigen für die Originalvariablen abweicht. Dies ist in Szenario B besonders deutlich; dort ergibt sich ein deutlich positiver Wert für ρ^a , während für die Originalvariablen $\rho = -0.999$ gilt! Wenn ρ positiv ist, nimmt die Diskrepanz ab, ist aber immer noch deutlich sichtbar. Im übrigen werden die Ergebnisse für ρ^a durch das empirische Maß r^a bestätigt. Bemerkenswert ist die Gleichheit von ρ^a in den Szenarios A und C. Hier wirkt sich aus, daß die Variationskoeffizienten für X und Y in beiden Fällen identisch sind. Aus (7-19) ergibt sich, daß dann das Maß ρ^a konstant bleibt.

Wesentlich ist ferner die Beobachtung, daß bereits der aus den Originalvariablen gebildete Quotient Z eine Verzerrung aufweist, die dann stärker ist, wenn die Originalvariablen sehr hohe **negative** Korrelation aufweisen. Diese Verzerrung wird durch die Bias-Formel (8-1) sehr gut abgebildet. Bezüglich des aus den überlagerten Variablen erzeugten Quotienten ergibt sich, daß dieser kaum von dem Quotienten, der aus den Originalvariablen gebildet wird, abweicht. Dies gilt praktisch unabhängig davon, wie groß die Diskrepanz zwischen den Korrelationskoeffizienten ρ und ρ^a ist. Die ursprüngliche Motivation für die Verwendung des "Höhne"-Verfahrens wird dadurch insoweit "legitimiert", daß der Quotienten-Bias

²⁸ Alle Ergebnisse wurden mit den GAUSS-Programmen QUOT01.prg, QUOT02.prg und QUOT03.prg (Laptop Ronning, Februar 2007) erzeugt.

der anonymisierten Variablen mit dem der Originalvariablen harmoniert. Andere Auswirkungen, vor allem die stark abweichende Korrelation sowie die Konsequenzen für die Schätzung von linearen Modellen (siehe dazu den folgenden Abschnitt), sind allerdings ebenfalls zu bewerten.

9 Schätzung linearer Modelle aus Querschnittsdaten

9.1 Einleitung

Im Folgenden soll untersucht werden, wie sich fehlerbehaftete Variablen auf die Schätzung von linearen Modellen auswirken. Gegenüber der bisherigen Literatur neu ist die Annahme von kontemporär korrelierten Fehler- bzw. Überlagerungsvariablen.

Wir betrachten das (klassische) multiple lineare Regressionsmodell, das wir wie folgt schreiben:

$$\mathbf{y} = \beta_0 \boldsymbol{\iota} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta} \quad (9-1)$$

Dabei ist $\boldsymbol{\iota}$ ein Einsvektor und \mathbf{X} eine $(n \times K)$ -Matrix, d.h. wir haben n Beobachtungen (im Querschnitt) und K (echte) Regressoren. Für die ersten und zweiten Momente von $\boldsymbol{\eta}$ soll gelten:

$$E[\boldsymbol{\eta}] = \mathbf{0} \quad \text{und} \quad \text{cov}[\boldsymbol{\eta}] = \sigma_\eta^2 \mathbf{I}. \quad (9-2)$$

Der OLS-Schätzer des Vektors $\boldsymbol{\beta}$ lautet

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{M}_\iota\mathbf{X})^{-1} \mathbf{X}'\mathbf{M}_\iota\mathbf{y} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{M}_\iota\mathbf{X})^{-1} \mathbf{X}'\mathbf{M}_\iota\boldsymbol{\eta} \quad (9-3)$$

und seine Kovarianzmatrix ist durch

$$\text{cov}[\hat{\boldsymbol{\beta}}] = \sigma_\eta^2 (\mathbf{X}'\mathbf{M}_\iota\mathbf{X})^{-1} \quad (9-4)$$

gegeben. Dabei ist die idempotente Matrix \mathbf{M}_ι durch

$$\mathbf{M}_\iota = \mathbf{I}_n - \frac{1}{n} \boldsymbol{\iota}\boldsymbol{\iota}'$$

gegeben.

Im folgenden sollen die Auswirkungen einer additiven bzw. multiplikativen Überlagerung der Regressormatrix \mathbf{X} sowie des Vektors \mathbf{y} auf die Schätzung des Parametervektors $\boldsymbol{\beta}$ untersucht werden. Dabei beschränken wir uns auf die Auswirkungen auf den naiven Schätzer

$$\hat{\boldsymbol{\beta}}^a = (\mathbf{X}^a\mathbf{M}_\iota\mathbf{X}^a)^{-1} \mathbf{X}^a\mathbf{M}_\iota\mathbf{y}^a \quad (9-5)$$

Für den Fall **unkorrelierter** Fehlervariablen können wir auf Ergebnisse aus Rosemann (2006) zurückgreifen. Dies gilt vor allem für den bisher wenig behandelten multiplikativen Fall. Für additive Überlagerung sind die Ergebnisse ohnehin seit längerem bekannt und können in jedem Ökonometrie-Lehrbuch nachgelesen werden. Beim Fall der korrelierten Fehlervariablen interessiert vor allem die spezielle Überlagerungsstruktur, die wir im Höhne-Verfahren betrachtet haben (siehe Abschnitt 6.2) und die wir im folgenden mit einer flexibleren Spezifikation vergleichen. Allerdings ist diese Spezifikation, die im einzelnen weiter unten dargestellt wird, nicht äquivalent zu einer Mischungsverteilung im Sinne von Abschnitt 3.

Im Fall der additiven Überlagerung gilt

$$\mathbf{y}^a = \mathbf{y} + \mathbf{u} \quad \text{und} \quad \mathbf{X}^a = \mathbf{X} + \mathbf{U} \quad (9-6)$$

im Fall der multiplikativen Überlagerung gilt

$$\mathbf{y}^a = \mathbf{y} \odot \mathbf{u} \quad \text{und} \quad \mathbf{X}^a = \mathbf{X} \odot \mathbf{U} \quad (9-7)$$

Dabei ist \mathbf{u} ein Zufallsvektor und \mathbf{U} eine Zufallsmatrix. Man beachte, daß im Gegensatz zu den Abschnitten 1 bis 7 hier eine andere Symbolik verwendet wird.²⁹ Für (9-7) können wir auch schreiben:

$$\mathbf{y}^a = \mathbf{y} \odot \mathbf{u}_y \quad \text{und} \quad (\mathbf{x}_1^a, \dots, \mathbf{x}_K^a) = (\mathbf{x}_1 \odot \mathbf{u}_1, \dots, \mathbf{x}_K \odot \mathbf{u}_K), \quad (9-8)$$

womit die separate Überlagerung jedes Regressor \mathbf{x}_k besonders gut analysiert werden kann. Es wird sich allerdings herausstellen, daß eine zeilenweise d.h. beobachtungspunkt-spezifische Darstellung mehr weiterhilft, indem wir (9-6) wie folgt schreiben:

$$\begin{aligned} (\mathbf{y}^a)' &= (y\langle 1 \rangle \cdot u\langle 1 \rangle, y\langle 2 \rangle \cdot u\langle 2 \rangle, \dots, y\langle n-1 \rangle \cdot u\langle n-1 \rangle, y\langle n \rangle \cdot u\langle n \rangle) \\ &\quad \text{und} \\ (\mathbf{X}^a)' &= (\mathbf{x}^a\langle 1 \rangle, \dots, \mathbf{x}^a\langle K \rangle) = (\mathbf{x}\langle 1 \rangle \odot \mathbf{u}\langle 1 \rangle, \dots, \mathbf{x}\langle n \rangle \odot \mathbf{u}\langle n \rangle), \end{aligned} \quad (9-9)$$

Dabei bezeichnet der Index in eckigen Klammern einen bestimmten Beobachtungspunkt. Der Vektor $\mathbf{x}\langle i \rangle$ ist K -dimensional und steht - gestürzt - für die i -te **Zeile (!)** der Matrix \mathbf{X} . Auch die Vektoren $\mathbf{u}\langle i \rangle$ sind K -dimensional und stehen - gestürzt - für die i -te Zeile der Matrix \mathbf{U} .

In bestimmten Fällen benötigen wir noch eine zusätzliche Darstellung für die "gemeinsame" (multiplikative) Überlagerung von Regressoren und Regressand: Dabei enthalten die einzelnen Zufallsvektoren K Elemente, wenn nur die Regressoren überlagert werden und $K + 1$ Elemente, wenn zusätzlich die abhängige Variable überlagert wird. Dies soll durch die Bezeichnungen \mathbf{u}_x und u_y kenntlich machen, die im letzteren Fall gemeinsam den $(K + 1)$ -dimensionalen Vektor \mathbf{u} bilden:

$$\mathbf{u} = \begin{pmatrix} \mathbf{u}_x \\ u_y \end{pmatrix}$$

9.2 Additive Überlagerung im linearen Regressionsmodell

9.2.1 Allgemeine Bemerkungen

Obwohl die Ergebnisse für den Fall der additiven Überlagerung seit langem bekannt sind, sollen sie hier der Vollständigkeit und auch zum Vergleich mit den Ergebnissen für den multiplikativen Fall kurz vorgestellt werden. Dabei nehme ich wiederholt Bezug auf Ergebnisse in Rosemann (2006), ohne dies an allen Stellen ausdrücklich kenntlich zu machen. Auch übergehe ich die einzelnen Annahmen, unter denen diese (asymptotischen) Ergebnisse abgeleitet werden. Des öfteren wird von der Annahme Gebrauch gemacht, daß

$$\text{plim} \frac{1}{n} \mathbf{X}' \mathbf{M}_l \mathbf{X} \equiv \mathbf{Q} \quad (9-10)$$

²⁹In Abschnitt 3 haben wir \mathbf{U} als Zufallsvektor eingeführt. Jetzt müssen wir die Überlagerung für alle n Beobachtungen modellieren, deshalb die geänderte Symbolik.

eine feste und reguläre d.h. invertierbare Matrix ist. Damit weiche ich von der üblichen Annahme ab, die die Momentenmatrix $\mathbf{X}'\mathbf{X}$ betrachtet, wie dies auch Rosemann (2006) tut. Der Unterschied besteht allerdings nur darin, daß ich statt der Originalwerte die Abweichungen vom Mittel betrachte, also zentrierte Momente statt der unzentrierten Momente.

9.2.2 Ausschließliche Überlagerung der Regressoren

Im Fall der ausschließlichen (additiven) Überlagerung der Regressoren (d.h. $\mathbf{y}^a = \mathbf{y}$) erhalten wir

$$\text{plim}\hat{\boldsymbol{\beta}}^a = (\text{cov}[\mathbf{u}] + \mathbf{Q})^{-1} \mathbf{Q}\boldsymbol{\beta} \quad (9-11)$$

wobei $\text{cov}[\mathbf{u}]$ durch (7-6) gegeben ist. Der "naive" Schätzer "unter"schätzt also in diesem all den wahren Wert derart, daß die Norm von $\text{plim}\hat{\boldsymbol{\beta}}^a$ kleiner ist als die von $\boldsymbol{\beta}$.³⁰ Dies entspricht dem bekannten Ergebnis im Fall der Einfachregression.

Vorausgesetzt daß \mathbf{Q} und $\text{cov}[\mathbf{u}]$ bekannt sind oder konsistent geschätzt werden können, läßt sich aus diesem Ergebnis ein konsistenter "Korrektur-Schätzer" wie folgt definieren:

$$\hat{\boldsymbol{\beta}}^{a, \text{korr}} = \mathbf{Q}^{-1} (\text{cov}[\mathbf{u}] + \mathbf{Q}) \hat{\boldsymbol{\beta}}^a .$$

Dabei ist $\hat{\boldsymbol{\beta}}^a$ durch (9-5) gegeben. Im Fall der Anonymisierung ist allerdings nur die Regressormatrix \mathbf{X}^a bekannt, die Matrix \mathbf{Q} also unbekannt. Da jedoch

$$\text{plim} \frac{1}{n} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a = \text{cov}[\mathbf{x}] + \text{cov}[\mathbf{u}] = \mathbf{Q} + \text{cov}[\mathbf{u}]$$

gilt, kann man - für bekannte Kovarianzmatrix $\text{cov}[\mathbf{u}]$ - folgende operationale Form des Korrektorschätzers verwenden:

$$\hat{\boldsymbol{\beta}}^{a, \text{korr}} = \left(\frac{1}{n} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a - \text{cov}[\mathbf{u}] \right)^{-1} \left(\frac{1}{n} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a \right) \hat{\boldsymbol{\beta}}^a . \quad (9-12)$$

was der Formel (16.77) bei Rosemann (2006) entspricht.

Es wird im Einzelnen noch zu untersuchen sein, wie sich die beiden unterschiedlichen Kovarianzmatrizen, die aus der "speziellen" bzw. "flexiblen" Spezifikation folgen, auf die Inkonsistenz des "naiven" Schätzers auswirken. Für den speziellen Fall erhalten wir aus Abschnitt 6.3

$$\text{cov}[\mathbf{u}] = \sigma_\varepsilon^2 \mathbf{I} + \mu^2 \boldsymbol{\omega}' \boldsymbol{\omega} = \frac{1}{\sigma_\varepsilon^2 + \mu^2} ((1 - \rho) \mathbf{I} + \rho \boldsymbol{\omega}' \boldsymbol{\omega})$$

wobei ρ durch

$$\rho = \frac{\mu^2}{\mu^2 + \sigma_\varepsilon^2}$$

gegeben ist. Man beachte, daß diese Korrelation stets positiv sein wird!

Dagegen gilt im Fall der "flexiblen" Spezifikation

$$\text{cov}[\mathbf{u}] = (\mu^2 + \sigma_\varepsilon^2) \mathbf{I} .$$

Siehe (6-29) für den "symmetrischen" Spezialfall ($\alpha = 0,5$). In diesem Fall besteht also keine Korrelation zwischen den einzelnen Komponenten!

³⁰Man kann zeigen, daß die Eigenwerte von $(\mathbf{A} + \mathbf{B})^{-1} \mathbf{A}$, \mathbf{A} und \mathbf{B} positiv definit, stets kleiner als 1 sind, was zu einer "Schrumpfung" (englisch "shrinkage") des Vektors \mathbf{x} führt, d.h. $\|(\mathbf{A} + \mathbf{B})^{-1} \mathbf{A} \mathbf{x}\| \leq \|\mathbf{x}\|$. Siehe z.B. Ronning(1977). Für entsprechende Ergebnisse bezüglich der Matrix $(\mathbf{A} + \mathbf{B})^{-1} \odot \mathbf{A}$ siehe Marshall und Olkin (1979 S.).

9.2.3 Gemeinsame Überlagerung aller Variablen

Derselbe Wahrscheinlichkeitsgrenzwert ergibt sich, wenn auch die abhängige Variable (additiv) überlagert wird und zwischen den Überlagerungsvariablen keine Korrelation besteht.³¹ Rosemann (2006) betrachtet jedoch auch den Fall, in dem Korrelation zwischen dem Vektor \mathbf{u}_x und der Überlagerung u_y für die abhängige Variable besteht. Dieser Fall wird in der Lehrbuchliteratur als nicht sehr bedeutsam angesehen. Im Fall der Überlagerung durch multivariate Mischungsverteilungen ist er allerdings wichtig, weil gerade im 'speziellen' Höhneverfahren diese Korrelation besteht, wie bereits oben demonstriert. In diesem Fall ergibt sich folgender Wahrscheinlichkeitsgrenzwert:

$$\text{plim} \hat{\beta}^a = (\text{cov}[\mathbf{u}] + \mathbf{Q})^{-1} (\mathbf{Q}\beta + \text{cov}[\mathbf{u}_x, u_y]) \quad (9-13)$$

Dabei ist $\text{cov}[\mathbf{u}_x, u_y]$ ein **Vektor**, der die Kovarianzen von \mathbf{u}_x mit u_y enthält. Im "speziellen" Fall hat dieser Vektor folgende Struktur:

$$\text{cov}[\mathbf{u}_x, u_y] = \mu^2 \boldsymbol{\iota} \quad . \quad (9-14)$$

Siehe Abschnitt 6.3. Dagegen ist im Fall der "Flexiblen" Spezifikation dieser Vektor gleich dem Nullvektor. Konsequenz dieses Ergebnisses ist, daß sich unterschiedliche Verzerrungen für beide Spezifikationen ergeben.

Für den Fall, daß Korrelation zwischen den Meßfehlern der Regressoren und des Regressanden bestehen, ergibt sich aus (9-13) folgender Korrektorschätzer:

$$\hat{\beta}^{a, \text{korr}} = \mathbf{Q}^{-1} \left\{ (\text{cov}[\mathbf{u}] + \mathbf{Q}) \hat{\beta}^a - \text{cov}[\mathbf{u}_x, u_y] \right\} \quad .$$

Dabei ist $\hat{\beta}^a$ durch (9-5) gegeben. Entsprechend (9-12) erhalten wir in diesem Fall als operationalen Korrektorschätzer

$$\hat{\beta}^{a, \text{korr}} = \left(\frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{X}^{\mathbf{a}} - \text{cov}[\mathbf{u}] \right)^{-1} \left(\frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{X}^{\mathbf{a}} \hat{\beta}^a + \text{cov}[\mathbf{u}_x, u_y] \right) \quad , \quad (9-15)$$

was Formel (16.81) bei Rosemann (2006) entspricht, wobei dieser noch ausnutzt, daß

$$\frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{X}^{\mathbf{a}} \hat{\beta}^a = \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{y}^{\mathbf{a}}$$

gilt.

Abschließend sei zur Vollständigkeit vermerkt, daß sich nur im Fall, daß **ausschließlich die abhängige Variable** (additiv) überlagert wird, eine konsistente Schätzung ergibt! Siehe z.B. Rosemann (2006 Seite 158-163). Auf Spezialfälle, in denen einzelne Regressoren überlagert werden, gehe ich nicht ein. Auch wird die nicht unwichtige Schätzung der Restvarianz bzw. der Kovarianzmatrix des naiven Schätzers nicht behandelt.

9.3 Multiplikative Überlagerung im linearen Regressionsmodell

9.3.1 Allgemeine Bemerkungen

Wir unterstellen, daß die einzelnen n Beobachtungspunkte stochastisch unabhängig voneinander überlagert werden, bzw. daß die Meßfehler zwar "kontemporär" (= beobachtungspunktspezifisch) korreliert sein können, nicht aber über verschiedene Beobachtungspunkte

³¹Allerdings steigt die Varianz des Schätzers.

hinweg.³² In diesem Fall ist die spezielle Formulierung der multiplikativen Überlagerung der Regressormatrix aus (9-8) besonders geeignet:

$$(\mathbf{X}^a)' = (\mathbf{x}^a\langle 1 \rangle, \dots, \mathbf{x}^a\langle K \rangle) = (\mathbf{x}\langle 1 \rangle \odot \mathbf{u}\langle 1 \rangle, \dots, \mathbf{x}\langle n \rangle \odot \mathbf{u}\langle n \rangle),$$

Die obige Annahme kann dann wie folgt präzisiert werden: Die n Zufallsvektoren

$$\mathbf{u}\langle 1 \rangle, \mathbf{u}\langle 2 \rangle, \dots, \mathbf{u}\langle n \rangle,$$

sind stochastisch voneinander unabhängig. Dabei enthalten die einzelnen Zufallsvektoren K Elemente, wenn nur die Regressoren überlagert werden und $K + 1$ Elemente, wenn zusätzlich die abhängige Variable überlagert wird. Dies werden wir durch die Symbolik \mathbf{u}_x und u_y kenntlich machen, die im letzteren Fall gemeinsam den $(K + 1)$ -dimensionalen Vektor \mathbf{u} bilden:

$$\mathbf{u} = \begin{pmatrix} \mathbf{u}_x \\ u_y \end{pmatrix} .$$

Bei Rosemann (2006) wird im Einzelnen ausgeführt, wie die Wahrscheinlichkeitsgrenzwerte im Fall der multiplikativen Überlagerung abgeleitet werden. Allerdings ist zu beachten, daß im Gegensatz zu Rosemann (2006) **in dieser Arbeit ein Absolutglied (siehe (9-1)) explizit berücksichtigt** wird, wodurch sich die Ergebnisse deutlich verändern. Deshalb sollen die - modifizierten - Beweise im Folgende ebenfalls geliefert werden.³³

9.3.2 Ausschließliche Überlagerung der Regressoren

Im Fall der ausschließlichen Überlagerung der Regressoren betrachten wir

$$\hat{\beta}^a = (\mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a)^{-1} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{y} .$$

Die Matrix $\mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a / n$ enthält die empirischen **zentrierten zweiten** Momente der multiplikativ überlagerten Regressoren und tendiert deshalb gegen die Kovarianzmatrix einer bestimmten Zeile von \mathbf{X}^a tendiert, die wir oben mit

$$\mathbf{x}^a\langle i \rangle = \mathbf{x}\langle i \rangle \odot \mathbf{u}\langle i \rangle \quad , \quad i = 1, \dots, n,$$

bezeichnet haben.

Wir verwenden nun die Formel (7-15), die in der hier verwendeten Symbolik (mit $\mathbf{x}\langle i \rangle \odot \mathbf{u}\langle i \rangle$ statt $\mathbf{Y}^a = \mathbf{Y} \odot \mathbf{U}$) zu folgendem Ergebnis führt:

$$cov[\mathbf{x}^a\langle i \rangle] = cov[\mathbf{u}\langle i \rangle] \odot (cov[\mathbf{x}\langle i \rangle] + E[\mathbf{x}\langle i \rangle] E[\mathbf{x}\langle i \rangle]') + cov[\mathbf{x}\langle i \rangle] .$$

Wegen der identischen Verteilung über alle n Stichprobenpunkte kann man dann auch schreiben:

$$cov[\mathbf{x}^a] = cov[\mathbf{u}] \odot (cov[\mathbf{x}] + \boldsymbol{\mu}_x \boldsymbol{\mu}_x') + cov[\mathbf{x}] . \quad (9-16)$$

³²Diese Annahme wird natürlich auch bei der additiven Überlagerung gemacht, wurde dort aber nicht speziell problematisiert.

³³Siehe Lin (1986) und Hwang(1989) für den Beweis im Fall nicht zentrierter Momente. Die vor allem bei Hwang (1989) verwendete Beweistechnik ist hier nicht anwendbar, weil der Operator \mathbf{M}_l das arithmetische Mittel über alle n Stichprobenpunkte fordert. Siehe jedoch Abschnitt 9.3.4, der unter expliziter Beachtung des Absolutglieds den Wahrscheinlichkeitsgrenzwert für den Gesamtvektor $(\beta_0, \beta')'$ bestimmt.

Unter Beachtung von (9-10) ergibt sich daraus

$$\text{plim} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{X}^{\mathbf{a}} = \text{cov}[\mathbf{u}] \odot (\mathbf{Q} + \boldsymbol{\mu}_x \boldsymbol{\mu}_x') + \mathbf{Q}$$

Wir benötigen ferner die Matrix

$$\frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{y} = \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l (\beta_0 \boldsymbol{\iota} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\eta}) = \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\eta})$$

Wegen

$$\text{plim} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{X} = \mathbf{Q} \quad \text{und} \quad \text{plim} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \boldsymbol{\eta} = \mathbf{0}$$

erhalten wir

$$\text{plim} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{y} = \text{plim} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l (\beta_0 \boldsymbol{\iota} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\eta}) = \mathbf{Q} \boldsymbol{\beta}$$

und damit

$$\text{plim} \hat{\boldsymbol{\beta}}^{\mathbf{a}} = (\text{cov}[\mathbf{u}] \odot (\mathbf{Q} + \boldsymbol{\mu}_x \boldsymbol{\mu}_x') + \mathbf{Q})^{-1} \mathbf{Q} \boldsymbol{\beta}, \quad (9-17)$$

wobei \mathbf{Q} durch (9-10) gegeben ist.

9.3.3 Gemeinsame Überlagerung aller Variablen

Im Fall der zusätzlichen Überlagerung auch der abhängigen Variablen lautet der Schätzer

$$\hat{\boldsymbol{\beta}}^{\mathbf{a}} = (\mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{X}^{\mathbf{a}})^{-1} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{y}^{\mathbf{a}}.$$

Somit ist der Wahrscheinlichkeits-Grenzwert des Vektors

$$\frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_l \mathbf{y}^{\mathbf{a}}$$

zu bestimmen. Da dieser Vektor die empirischen Kovarianzen zwischen $x_{ik}^{\mathbf{a}} = x_{ik} \cdot u_{ik}$ und $y_i^{\mathbf{a}} = y_i \cdot u_{iy}$ enthält, tendieren die einzelnen Elemente dieses Vektors gegen die entsprechenden theoretischen Kovarianzen, die im folgenden zu bestimmen sind. Dabei verwenden wir die Modellannahme

$$y_i = \beta_0 + \sum_{\ell=1}^K \beta_{\ell} x_{i\ell} + \eta_i, \quad i = 1, \dots, n,$$

die äquivalent zu (9-1) ist.

Wir verwenden wieder die Formel (7-15) für die Kovarianz, die wir hier - unter Vernachlässigung des Beobachtungsindex i - wie folgt schreiben:

$$\text{cov} \left[\begin{pmatrix} x_k^{\mathbf{a}} \\ y^{\mathbf{a}} \end{pmatrix} \right] = E_{xy} \left[\text{cov} \left[\begin{pmatrix} x_k^{\mathbf{a}} \\ y^{\mathbf{a}} \end{pmatrix} \mid \begin{pmatrix} x_k \\ y \end{pmatrix} \right] + \text{cov}_{xy} \left[E \left[\begin{pmatrix} x_k^{\mathbf{a}} \\ y^{\mathbf{a}} \end{pmatrix} \mid \begin{pmatrix} x_k \\ y \end{pmatrix} \right] \right] \right].$$

Für die bedingte Kovarianz ergibt sich

$$\text{cov} \left[\begin{pmatrix} x_k^{\mathbf{a}} \\ y^{\mathbf{a}} \end{pmatrix} \mid \begin{pmatrix} x_k \\ y \end{pmatrix} \right] = x_k \cdot y \cdot \text{cov}[u_k, u_y] = x_k (\beta_0 + \sum_{\ell=1}^K \beta_{\ell} x_{\ell} + \eta) \text{cov}[u_k, u_y]$$

und für den Vektor der bedingten Erwartungswerte erhalten wir

$$E \left[\begin{pmatrix} x_k^a \\ y^a \end{pmatrix} \mid \begin{pmatrix} x_k \\ y \end{pmatrix} \right] = \begin{pmatrix} x_k \\ y \end{pmatrix}$$

Der Erwartungswert der bedingten Kovarianzmatrix ergibt

$$E[x_k (\beta_0 + \sum_{\ell=1}^K \beta_\ell x_\ell + \eta) cov[u_k, u_y]] = (\beta_0 \mu_k + \sum_{\ell=1}^K \beta_\ell (\sigma_{k\ell} + \mu_k \mu_\ell)) cov[u_k, u_y]$$

und für die Kovarianz des Vektors der bedingten Erwartungswerte erhalten wir

$$cov[x_k, y] = cov[x_k, \beta_0] + cov[x_k, \sum_{\ell=1}^K \beta_\ell x_\ell] + cov[x_k, \eta] = \sum_{\ell=1}^K \beta_\ell \sigma_{k\ell}$$

Wenn wir die Ergebnisse für alle K Elemente gemeinsam schreiben, so erhalten wir für den Vektor der Ergebnisse aus dem ersten Ausdruck

$$\{\beta_0 \boldsymbol{\mu}_x + (\mathbf{Q} + \boldsymbol{\mu}_x \boldsymbol{\mu}_x') \boldsymbol{\beta}\} \odot cov[\mathbf{u}_x, u_y]$$

und der Vektor, der aus dem zweiten Ergebnis folgt, lautet

$$\mathbf{Q} \boldsymbol{\beta}$$

Damit erhalten wir folgenden Wahrscheinlichkeitsgrenzwert für den "naiven" Schätzer im Fall der gemeinsamen Überlagerung aller Variablen:³⁴

$$\text{plim } \hat{\boldsymbol{\beta}}^a = (cov[\mathbf{u}_x] \odot \{\mathbf{Q} + \boldsymbol{\mu}_x \boldsymbol{\mu}_x'\} + \mathbf{Q})^{-1} (\mathbf{Q} \boldsymbol{\beta} + \{\beta_0 \boldsymbol{\mu}_x + (\mathbf{Q} + \boldsymbol{\mu}_x \boldsymbol{\mu}_x') \boldsymbol{\beta}\} \odot cov[\mathbf{u}_x, u_y]) \quad (9-18)$$

wobei $cov[\mathbf{u}_x]$ die $(n \times n)$ -Kovarianzmatrix der Überlagerungsvariablen aus \mathbf{u}_x und $cov[\mathbf{u}_x, u_y]$ den n -dimensionalen Vektor der Kovarianzen zwischen den Elementen aus \mathbf{u}_x und u_y bezeichnet. Konsistenz ist demnach nur dann gegeben, wenn beide gerade erwähnten Ausdrücke gleich der Nullmatrix bzw. gleich dem Nullvektor sind. Insbesondere muß die Kovarianzmatrix $cov[\mathbf{u}_x]$ eine Nullmatrix sein, d.h. auch die Diagonalelemente (= Varianzen) müssen gleich Null sein!

Im Fall der Höhenüberlagerung gilt

$$cov[\mathbf{u}_x] = \sigma_\eta^2 \mathbf{I} + \delta^2 \boldsymbol{\nu} \boldsymbol{\nu}' \quad (9-19)$$

und

$$cov[\mathbf{u}_x, u_y] = \begin{cases} \delta^2 \boldsymbol{\nu} & \text{für die spezielle Spezifikation} \\ \mathbf{0} & \text{für die flexible Spezifikation} \end{cases} \quad (9-20)$$

9.3.4 Alternativer Beweis für (9-18)

Es soll gezeigt werden, daß auch mit der Beweis-Methode von Lin(1989) der Wahrscheinlichkeitsgrenzwert für den "naiven" Schätzer bei gemeinsamer multiplikativer Überlagerung aller Variablen abgeleitet werden kann. Siehe dazu Abschnitt 9.3.3.

³⁴Siehe auch den alternativen Beweis in Abschnitt 9.3.4.

Es wird unterstellt, daß die i -te Zeile der Regressormatrix \mathbf{X} bzw \mathbf{X}^a durch den Spaltenvektor

$$\mathbf{x}\langle i \rangle = \begin{pmatrix} 1 \\ x_{i2} \\ \vdots \\ x_{iK} \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{x}_2 \end{pmatrix} \quad \text{und} \quad \mathbf{x}^a\langle i \rangle = \begin{pmatrix} 1 \\ x_{i2}^a \\ \vdots \\ x_{iK}^a \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{x}_2^a \end{pmatrix}$$

gegeben ist und der korrespondierende Überlagerungsvektor \mathbf{u}_x folgende Gestalt hat:

$$\mathbf{u}_x\langle i \rangle = \begin{pmatrix} 1 \\ u_{2x} \\ \vdots \\ u_{Kx} \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{u}_{2x} \end{pmatrix}$$

Damit wird dem Absolutglied bzw. dem Einsvektor in der Regressormatrix \mathbf{X} Rechnung getragen.

Es gilt

$$\begin{aligned} 1/n \mathbf{X}^a \mathbf{X}^a &= 1/n \sum_{i=1}^n \mathbf{x}^a\langle i \rangle \mathbf{x}^a\langle i \rangle' \\ &= 1/n \sum_{i=1}^n (\mathbf{x}\langle i \rangle \odot \mathbf{u}_x\langle i \rangle) (\mathbf{x}\langle i \rangle \odot \mathbf{u}_x\langle i \rangle)' \\ &= 1/n \sum_{i=1}^n \mathbf{x}\langle i \rangle (\mathbf{x}\langle i \rangle)' \odot \mathbf{u}_x\langle i \rangle \mathbf{u}_x\langle i \rangle' \end{aligned}$$

und für den Erwartungswert ergibt sich

$$E[1/n \mathbf{X}^a \mathbf{X}^a] = \begin{pmatrix} 1 & \boldsymbol{\mu}'_x \\ \boldsymbol{\mu}_x & \mathbf{P} \odot \mathbf{M} \end{pmatrix}$$

mit

$$\mathbf{P} = E[1/n \sum_{i=1}^n \mathbf{x}_2\langle i \rangle (\mathbf{x}_2\langle i \rangle)'] = \text{cov}[\mathbf{x}] + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x$$

und

$$\mathbf{M} = E[\mathbf{u}_x\langle i \rangle \mathbf{u}_x\langle i \rangle'] = \text{cov}[\mathbf{u}_{2x}] + \boldsymbol{\nu} \boldsymbol{\nu}'$$

Unter Verwendung der Formel für die Invertierung einer zerlegten Matrix³⁵ erhalten wir ferner

$$\begin{pmatrix} 1 & \boldsymbol{\mu}'_x \\ \boldsymbol{\mu}_x & \mathbf{P} \odot \mathbf{M} \end{pmatrix}^{-1} = \begin{pmatrix} 1 + \boldsymbol{\mu}'_x (\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x)^{-1} \boldsymbol{\mu}_x & -\boldsymbol{\mu}'_x (\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x)^{-1} \\ -(\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x)^{-1} \boldsymbol{\mu}_x & (\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x)^{-1} \end{pmatrix}$$

³⁵Allgemein gilt für die reguläre Matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

die Formel

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} (\mathbf{I} + \mathbf{A}_{12} \mathbf{D}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1}) & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{D}^{-1} \\ -\mathbf{D}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{D}^{-1} \end{pmatrix}$$

mit $\mathbf{D} = \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}$.

Ferner gilt

$$\begin{aligned}
1/n \mathbf{X}^{\mathbf{a}'} \mathbf{y}^{\mathbf{a}} &= 1/n \sum_{i=1}^n (\mathbf{x}\langle i \rangle \odot \mathbf{u}_x\langle i \rangle) (y_i \cdot u_{yi}) \\
&= 1/n \sum_{i=1}^n \mathbf{x}\langle i \rangle y_i \odot \mathbf{u}_x\langle i \rangle u_{yi} \\
&= 1/n \sum_{i=1}^n \mathbf{x}\langle i \rangle \left\{ \mathbf{x}\langle i \rangle' \begin{pmatrix} \beta_0 \\ \boldsymbol{\beta} \end{pmatrix} + \eta_i \right\} \odot \mathbf{u}_x\langle i \rangle u_{yi} \\
&= 1/n \sum_{i=1}^n \begin{pmatrix} 1 \\ \mathbf{x}_2\langle i \rangle \end{pmatrix} \left\{ \begin{pmatrix} 1 \\ \mathbf{x}_2\langle i \rangle \end{pmatrix}' \begin{pmatrix} \beta_0 \\ \boldsymbol{\beta} \end{pmatrix} + \eta_i \right\} \odot \mathbf{u}_x\langle i \rangle u_{yi} \\
&= 1/n \sum_{i=1}^n \begin{pmatrix} \beta_0 + \mathbf{x}_2\langle i \rangle' \boldsymbol{\beta} + \eta_i \\ \beta_0 \mathbf{x}_2\langle i \rangle + \mathbf{x}_2\langle i \rangle \mathbf{x}_2\langle i \rangle' \boldsymbol{\beta} + \mathbf{x}_2\langle i \rangle \eta_i \end{pmatrix} \odot \begin{pmatrix} u_{yi} \\ \mathbf{u}_{2x}\langle i \rangle u_{yi} \end{pmatrix}
\end{aligned}$$

und für den Erwartungswert ergibt sich in diesem Fall

$$E[1/n \mathbf{X}^{\mathbf{a}'} \mathbf{y}^{\mathbf{a}}] = \begin{pmatrix} \beta_0 + \boldsymbol{\mu}'_x \boldsymbol{\beta} \\ \{\beta_0 \boldsymbol{\mu}_x + (\text{cov}[\mathbf{x}_2] + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) \boldsymbol{\beta}\} \odot \text{cov}[\mathbf{u}_{2x}, u_y] \end{pmatrix}$$

wobei

$$E[\mathbf{u}_{2x}\langle i \rangle u_{yi}] = \text{cov}[\mathbf{u}_{2x}, u_y] + \boldsymbol{\nu}$$

verwendet wurde.

Damit ergibt sich für den naiven Schätzer bei Beschränkung auf den $K - 1$ -dimensionalen Vektor $\boldsymbol{\beta}$:

$$\begin{aligned}
\text{plim} \hat{\boldsymbol{\beta}}^a &= -(\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x)^{-1} \boldsymbol{\mu}_x \cdot (\beta_0 + \boldsymbol{\mu}'_x \boldsymbol{\beta}) \\
&\quad + (\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x)^{-1} \{\beta_0 \boldsymbol{\mu}_x + (\text{cov}[\mathbf{x}_2] + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) \boldsymbol{\beta}\} \odot \{\text{cov}[\mathbf{u}_{2x}, u_y] + \boldsymbol{\nu}\} \\
&= (\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x)^{-1} (\text{cov}[\mathbf{x}_2] \boldsymbol{\beta} + \{\beta_0 \boldsymbol{\mu}_x + (\text{cov}[\mathbf{x}_2] + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) \boldsymbol{\beta}\} \odot \text{cov}[\mathbf{u}_{2x}, u_y])
\end{aligned}$$

Wenn wir weiter beachten, daß

$$\begin{aligned}
\mathbf{P} \odot \mathbf{M} - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x &= (\text{cov}[\mathbf{x}_2 + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x] \odot (\text{cov}[\mathbf{u}_{x2} + \boldsymbol{\nu} \boldsymbol{\nu}'] - \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) \\
&= \text{cov}[\mathbf{u}_{x2}] (\text{cov}[\mathbf{x}_2] + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) + \text{cov}[\mathbf{x}_2] \quad ,
\end{aligned}$$

dann können wir schreiben:

$$\text{plim} \hat{\boldsymbol{\beta}}^a = (\text{cov}[\mathbf{u}_{x2}] (\text{cov}[\mathbf{x}_2] + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) + \text{cov}[\mathbf{x}_2])^{-1} (\text{cov}[\mathbf{x}_2] \boldsymbol{\beta} + \{\beta_0 \boldsymbol{\mu}_x + (\text{cov}[\mathbf{x}_2] + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) \boldsymbol{\beta}\} \odot \text{cov}[\mathbf{u}_{2x}, u_y]) .$$

Dieser Ausdruck ist äquivalent mit der Formel (9-18) in Abschnitt 9.3. Dabei ist zu beachten, daß $\text{cov}[\mathbf{x}_2] = \mathbf{Q}$, $\text{cov}[\mathbf{u}_{2x}] = \text{cov}[\mathbf{u}_x]$ und $\text{cov}[\mathbf{u}_{2x}, u_y] = \text{cov}[\mathbf{u}_x, u_y]$ im Abschnitt 9.3.

9.4 Korrektorschätzer

Vorausgesetzt daß \mathbf{Q} und $\text{cov}[\mathbf{u}]$ bekannt sind oder konsistent geschätzt werden können, läßt sich aus dem Ergebnis (9-17) für den Fall der ausschließlichen (multiplikativer) Überlagerung der Regressoren ein konsistenter "Korrektur-Schätzer" wie folgt definieren:

$$\hat{\boldsymbol{\beta}}^{a, \text{korr}} = \mathbf{Q}^{-1} (\text{cov}[\mathbf{u}] \odot (\mathbf{Q} + \boldsymbol{\mu}_x \boldsymbol{\mu}'_x) + \mathbf{Q}) \hat{\boldsymbol{\beta}}^a .$$

Im Fall der Anonymisierung ist allerdings nur die Regressormatrix \mathbf{X}^a bekannt, die Matrix \mathbf{Q} also unbekannt. Da jedoch im multiplikativen Fall

$$\text{plim} \frac{1}{n} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a = \text{cov}[\mathbf{u}] \odot \mathbf{Q}$$

(siehe Hwang 1986) und somit³⁶

$$\text{plim} \left(\frac{1}{n} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a \div \text{cov}[\mathbf{u}] \right) = \mathbf{Q}$$

und außerdem

$$\text{plim} \frac{1}{n} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{y} = \mathbf{Q} \boldsymbol{\beta}$$

(siehe oben bzw. bei Hwang 1986) gilt, erhalten wir für den Fall, daß nur die Regressoren (multiplikativ) überlagert sind, einen konsistenten Korrektur-Schätzer als

$$\hat{\boldsymbol{\beta}}^{a, \text{korrr}} = (\mathbf{X}^{a'} \mathbf{M}_l \mathbf{X}^a \div \text{cov}[\mathbf{u}])^{-1} \mathbf{X}^{a'} \mathbf{M}_l \mathbf{y} \quad , \quad (9-21)$$

was der Formel (16.105) bei Rosemann (2006) entspricht.³⁷

Die Ableitung eines Korrektorschätzers **im Fall der gemeinsamen (multiplikativen) Überlagerung** ist komplizierter, weil eine direkte Auflösung nach $\boldsymbol{\beta}$ nicht möglich ist. Siehe (9-18). Allerdings läßt sich über den bei Rosemann (2006) für diesen Fall abgeleiteten Schätzer des Gesamtvektors (einschließlich Absolutglied) unter Verwendung der Methoden aus Abschnitt 9.3.4 eine explizite Lösung angeben! Dies wird hier nicht ausgeführt.

10 Schätzung linearer Modelle aus Paneldaten

10.1 Allgemeines

10.1.1 Überlagerungsstrategien

Im Fall von Paneldaten treten an die Stelle von r verschiedenen Merkmalen bzw. Zufallsvariablen

$$Y_1, Y_2, \dots, Y_{r-1}, Y_r$$

im Querschnitt die $T \cdot r$ Variablen Y_{jt} bzw. ausführlich

$$\begin{array}{cccccc} Y_{11} & Y_{12} & \dots & Y_{1,r-1} & Y_{1r} \\ Y_{21} & Y_{22} & \dots & Y_{2,r-1} & Y_{2r} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_{T1} & Y_{T2} & \dots & Y_{T,r-1} & Y_{Tr} \end{array}$$

für r Merkmale in T Zeitpunkten. Diese Zufallsvariablen lassen sich auch als T verschiedene r -dimensionale Zufallsvektoren

$$\mathbf{Y}_1 \mathbf{Y}_2 \dots, \mathbf{Y}_{T-1}, \mathbf{Y}_T$$

³⁶Das Symbol \div bezeichnet die "Hadamard-Division".

³⁷Dieser Korrektorschätzer ist auch gültig, wenn die abhängige Variable ebenfalls überlagert wird, deren Überlagerungsvariable aber nicht mit den Überlagerungen der Regressoren korreliert ist.

interpretieren.

Es stellt sich die Frage, in welcher Weise die Überlagerung nun vorgenommen werden soll. Wir beschränken uns hier auf die Spezifikation³⁸

$$U_{jt} = 1 + \delta D + \varepsilon_{jt} \quad (\text{"alle Merkmale und Zeitpunkte gemeinsam"}) \quad , \quad (10-1)$$

die den multiplikativen Fall betrifft. Wir werden jedoch auch kurz auf den entsprechenden additiven Fall eingehen.

10.1.2 Das einfache lineare Panelmodell

Entsprechend dem Vorgehen von Biørn (1996) (und abweichend von der Darstellung im Abschnitt 9.3 für Querschnittsdaten betrachten wir hier das **einfache** Regressionsmodell

$$y_{it} = \alpha + \beta x_{it} + \eta_{it} \quad i = 1, \dots, n \quad t = 1, \dots, T \quad (10-2)$$

mit einem einzigen Regressor.

Wir werden zunächst wieder die ausschließliche Überlagerung der Regressoren (in diesem Fall des einen Regressors) betrachten und später dann untersuchen, was sich ändert, wenn x und y gemeinsam überlagert werden oder nur y überlagert wird. Siehe dazu die Abschnitte 11-45, 10.4.7, 10.4.8 und 10.4.10. Wie wir in den vorhergehenden Kapiteln gesehen haben, hat dies **im Querschnitt** sowohl bei additiver als auch im multiplikativer Überlagerung Auswirkung auf die Verzerrung der Koeffizienten-Schätzung, im letzten Fall allerdings nur bei Korrelation der Überlagerungsvariablen. Allerdings müssen wir dafür zu allererst die Überlagerung von Paneldaten (Abschnitt 10.2) darstellen. Dabei werden wir uns wieder auf die multiplikative Überlagerung und dort auf die "spezielle" Höhe-Variante konzentrieren. Ferner werden wir in Abschnitt 10.3 das einfache lineare Panelmodell mit Individualeffekten einführen.

10.2 Überlagerung von Paneldaten

10.2.1 Symbolik

Bei gemeinsamer Betrachtung von r Merkmalen im Fall von Paneldaten benötigen wir folgende Symbolik:

$$y_{itj} \quad i = 1, \dots, n, \quad t = 1, \dots, T, \quad j = 1, \dots, r, \quad (10-3)$$

³⁸Denkbar sind folgende Varianten:

$U_{jt} = 1 + \delta_{jt} D_{jt} + \varepsilon_{jt}$	alle Merkmale und Zeitpunkte getrennt
$U_{jt} = 1 + \delta D_j + \varepsilon_{jt}$	alle Merkmale getrennt
$U_{jt} = 1 + \delta D_t + \varepsilon_{jt}$	alle Zeitpunkte getrennt
$U_{jt} = 1 + \delta D + \varepsilon_{jt}$	alle Merkmale und Zeitpunkte gemeinsam

Weitere Varianten, in denen δ und D nicht identisch behandelt werden, sind denkbar. Die Effekte der einzelnen Überlagerungsvarianten hängen auch davon ab, ob man eine Zeitreihenstruktur unterlegt, beispielsweise einen r -dimensionalen autoregressiven Prozeß.

bezeichnet das Merkmal³⁹ y_j für Beobachtungseinheit i im Zeitpunkt t , wobei insgesamt r Merkmale betrachtet werden.

10.2.2 Additive Überlagerung

Wenn wir die additive Entsprechung zu (10-1) für die Überlagerung verwenden, d.h.

$$u_{itj} = y_{itj} + \mu D_i + \varepsilon_{itj} \quad , \quad (10-4)$$

so ergibt sich für die anonymisierten Merkmale

$$y_{itj}^a = y_{itj} + u_{itj} = y_{itj} + \mu D_i + \varepsilon_{itj} \quad (10-5)$$

Dabei soll ε_{itj} konstante Varianz besitzen, ferner sollen alle (itj) -Kombinationen unkorreliert sein:

$$\begin{aligned} E[\varepsilon_{itj}] &= 0 \\ \text{var}[\varepsilon_{itj}] &= \sigma_\varepsilon^2 \\ \text{cov}[\varepsilon_{itj}, \varepsilon_{hsk}] &= 0, \quad i \neq h, t \neq s, j \neq k \end{aligned} \quad \left\{ \begin{array}{l} i = 1, \dots, n, \\ t = 1, \dots, T, \\ j = 1, \dots, r. \end{array} \right. \quad (10-6)$$

Gleichwohl wird über die Zufallsvariable D_i Korrelation sowohl über die Merkmale als auch über die Zeitpunkte erzeugt bzw. diese Korrelation wird verstärkt, soweit die Originalvariablen zeitliche Korrelation aufweisen. Entsprechend dem Vorgehen in Abschnitt 7 (Anonymisierung mittels stochastischer Überlagerung) für Querschnittsdaten läßt sich zeigen, daß

$$\text{var}[y_{itj}^a] = \sigma_{tj}^2 + \mu^2 + \sigma_\varepsilon^2 \quad (10-7)$$

und

$$\text{cov}[y_{itj}^a, y_{hsk}^a] = \begin{cases} \sigma_{jk} + \mu^2 & \text{falls } i = h, t = s, j \neq k \\ \sigma_{ts} + \mu^2 & \text{falls } i = h, t \neq s, j = k \\ 0 & \text{falls } i \neq h \end{cases} \quad (10-8)$$

für alle (j, k, t, s) -Kombinationen gilt. Dies entspricht (7-6) im Abschnitt 7 (Anonymisierung mittels stochastischer Überlagerung) für Querschnittsdaten. In (10-7) wird jetzt zugelassen, daß die Varianz für Merkmal j über die Zeit hin variiert und in (10-8) bezeichnet σ_{jk} die (kontemporäre) Kovarianz zwischen den Merkmalen j und k , während σ_{ts} die Kovarianz zwischen den Zeitpunkten t und s für das j -te Merkmal bezeichnet.

Durch die Art der Überlagerung ist jetzt zugelassen, daß – für eine bestimmte Beobachtungseinheit i – Korrelation auch über verschiedene Zeitpunkte ("Wellen") hinweg besteht, selbst wenn die zeitliche Korrelation der Originalvariablen gleich Null ist, d.h. falls $\sigma_{ts} = 0$. Formal gilt

$$\text{corr}[y_{itj}^a, y_{hsk}^a] = \begin{cases} \frac{\sigma_{jk} + \mu^2}{\sqrt{(\sigma_{tj}^2 + \mu^2 + \sigma_\varepsilon^2)(\sigma_{tk}^2 + \mu^2 + \sigma_\varepsilon^2)}} & \text{falls } i = h, t = s, j \neq k \\ \frac{\sigma_{ts} + \mu^2}{\sqrt{(\sigma_{tj}^2 + \mu^2 + \sigma_\varepsilon^2)(\sigma_{sj}^2 + \mu^2 + \sigma_\varepsilon^2)}} & \text{falls } i = h, t \neq s, j = k \\ 0 & \text{falls } i \neq h \end{cases} \quad (10-9)$$

Man beachte, daß dabei unterstellt wird, daß zwischen den einzelnen Beobachtungseinheiten keine Korrelation, beispielsweise in Form von Cluster-Effekten, besteht. Das wurde aber beim Querschnittsfall auch so unterstellt.

³⁹Das Symbol y steht hier für jedes beliebige Merkmal und nicht für die abhängige Variable. Dies entspricht dem Vorgehen im Abschnitt 7.

10.2.3 Multiplikative Überlagerung

Im multiplikativen Fall verwenden wir statt (10-5) die Spezifikation

$$y_{itj}^a = y_{itj} (1 + \delta D_i + \varepsilon_{itj}) \quad (10-10)$$

Dabei soll ε_{itj} wieder (10-6) erfüllen.

Wegen der Symmetrie bezüglich der Merkmalskomponente j einerseits und der Zeitkomponente t andererseits kann man auch hier die Ergebnisse der Auswirkung der stochastischen Überlagerung in Anlehnung an Abschnitt 7 (Anonymisierung mittels stochastischer Überlagerung) für Querschnittsdaten ableiten. Dabei beschränken wir uns auf die Überlagerung mithilfe der Mischungsverteilung, wie sie Abschnitt 7.3.4 (spezielle Höhne-Spezifikation) betrachtet wurde.

Es läßt sich zeigen, daß

$$\text{var}[y_{itj}^a] = \sigma_{tj}^2 + (\delta^2 + \sigma_\varepsilon^2) (\sigma_{tj}^2 + \mu_{tj}^2) \quad (10-11)$$

und

$$\text{cov}[y_{itj}^a, y_{hsk}^a] = \begin{cases} \sigma_{jk} + \delta^2 (\sigma_{jk} + \mu_{tj} \mu_{tk}) & \text{falls } i = h, t = s, j \neq k \\ \sigma_{ts} + \delta^2 (\sigma_{ts} + \mu_{tj} \mu_{sj}) & \text{falls } i = h, t \neq s, j = k \\ 0 & \text{falls } i \neq h \end{cases} \quad (10-12)$$

gilt. Dabei bezeichnet μ_{tj} den Erwartungswert für Merkmal j im Zeitpunkt t . Varianzen und Kovarianzen entsprechen exakt der Struktur, wie sie in (7-17) für Querschnittsdaten angegeben wurde. Neben die Merkmalskomponenten tritt allerdings nun die Zeitkomponente.

Für die Korrelation zwischen den Variablen als auch zwischen den einzelnen Zeitpunkten erhalten wir jetzt folgende Formel:

$$\text{corr}[y_{itj}^a, y_{hsk}^a] = \begin{cases} \frac{\sigma_{jk} + \delta^2 (\sigma_{jk} + \mu_{tj} \mu_{tk})}{\sqrt{\{\sigma_{tj}^2 + (\delta^2 + \sigma_\varepsilon^2) (\sigma_{tj}^2 + \mu_{tj}^2)\} \{\sigma_{tj}^2 + (\delta^2 + \sigma_\varepsilon^2) (\sigma_{tj}^2 + \mu_{tj}^2)\}}} & \text{falls } i = h, t = s, j \neq k \\ \frac{\sigma_{ts} + \delta^2 (\sigma_{ts} + \mu_{tj} \mu_{sj})}{\sqrt{\{\sigma_{tj}^2 + (\delta^2 + \sigma_\varepsilon^2) (\sigma_{tj}^2 + \mu_{tj}^2)\} \{\sigma_{tj}^2 + (\delta^2 + \sigma_\varepsilon^2) (\sigma_{tj}^2 + \mu_{tj}^2)\}}} & \text{falls } i = h, t \neq s, j = k \\ 0 & \text{falls } i \neq h \end{cases} \quad (10-13)$$

Ebenso wie beim additiven Fall wird auch hier unterstellt, daß zwischen den einzelnen Beobachtungseinheiten keine Korrelation, beispielsweise in Form von Cluster-Effekten, besteht. Das wurde aber auch so beim Querschnittsfall unterstellt. Es geht ja darum herauszuarbeiten, welche zusätzliche Wirkung bezüglich Korrelation die Überlagerung hat.

10.3 Das einfache lineare Panelmodell mit Individualeffekten

10.3.1 Feste Individualeffekte

Das einfache lineare Panelmodell mit festen Individualeffekten schreiben wir wie folgt:

$$y_{it} = \alpha + \beta x_{it} + \sum_{h=1}^n \tau_h W_{ht} + \eta_{it} \quad , \quad i = 1, \dots, n, \quad t = 1, \dots, T \quad (10-14)$$

wobei

$$W_{ht} = \begin{cases} 1 & \text{falls } i = h \\ 0 & \text{sonst} \end{cases}$$

die individuenspezifische Dummyvariable im Zeitpunkt t ist und τ_h den entsprechenden Effekt bezeichnet. Aus Gründen der Identifikation setzen wir

$$\alpha = 0 \quad .$$

In Matrixschreibweise erhalten wir dann

$$\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\boldsymbol{\tau} + \boldsymbol{\eta} \quad (10-15)$$

wobei \mathbf{x} ein Vektor und β ein Skalar ist. Die Matrix \mathbf{W} enthält die individuenspezifischen Dummyvariablen.⁴⁰ Als sogenannter "Within-Schätzer" (oder auch "fixed effects estimator") ergibt sich die Teilschätzung nach der Kleinstquadrat-Methode:

$$\hat{\beta}_W = \frac{\mathbf{x}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W})\mathbf{y}}{\mathbf{x}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W})\mathbf{x}} = \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})(y_{it} - \bar{y}_{i\bullet})}{\sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})^2} \quad (10-16)$$

mit

$$\bar{x}_{i\bullet} = \frac{1}{T} \sum_{t=1}^T x_{it} \quad , \quad \bar{y}_{i\bullet} = \frac{1}{T} \sum_{t=1}^T y_{it} \quad .$$

Dieser Schätzer hat die üblichen "guten" Eigenschaften des klassischen Regressionsmodells!

10.3.2 Stochastische Individualeffekte

Wenn wir statt der festen Effekte zufällige Effekte annehmen, dann erhalten wir

$$y_{it} = \alpha + \beta x_{it} + \tau_i + \eta_{it} \quad , \quad i = 1, \dots, n \quad , \quad t = 1, \dots, T \quad (10-17)$$

wobei die τ_i nun Zufallsvariable sind. Weil daraus

$$\bar{y}_{i\bullet} = \alpha + \bar{x}_{i\bullet} \beta + \tau_i + \bar{\eta}_{i\bullet}$$

und

$$y_{it} - \bar{y}_{i\bullet} = (x_{it} - \bar{x}_{i\bullet})\beta + \eta_{it} - \bar{\eta}_{i\bullet}$$

folgt, spielt der individuenspezifische (zufällige) Effekt bei der Schätzung mittels $\hat{\beta}_W$ keine Rolle:

$$\hat{\beta}_W = \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet}) \{ (x_{it} - \bar{x}_{i\bullet})\beta + (\eta_{it} - \bar{\eta}_{i\bullet}) \}}{\sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})^2} \quad (10-18)$$

d.h. $\hat{\beta}_W$ ist ein konsistenter Schätzer.⁴¹

Im Abschnitt 10.6 wird gezeigt, daß die Varianz dieses Schätzers durch

$$\text{var}[\hat{\beta}_W] = \frac{\sigma_\eta^2}{(\mathbf{x}'\mathbf{M}_W\mathbf{x})} = \frac{\sigma_\eta^2}{\sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})^2} \quad (10-19)$$

gegeben ist. Man beachte, daß die Varianz der Individualeffekte in dieser Formel nicht auftaucht.

⁴⁰Für eine exakte Beschreibung siehe die Ergebnisse für das multiple Regressionsmodell in Abschnitt 10.6.

⁴¹Siehe beispielsweise Wooldridge (2002, Kap. 10.5.5), wo der allgemeinere Fall einer verallgemeinerten Schätzung behandelt wird. Wooldridge (2002 Kap. 10.7.2) geht auch darauf ein, daß die "Robustheit gegenüber der Korrelation von τ_i und dem Regressor" dann problematisch ist, wenn man den Effekt von zeitlich nur wenig variierenden Regressoren schätzen will.

10.4 Schätzung des linearen Panelmodells aus anonymisierten Daten

10.4.1 Der naive Panelschätzer

Im Fall von anonymisierten Daten verwenden wir statt (10-16) den folgenden "naiven" Schätzer (Within-Schätzer, FE-Schätzer):

$$\hat{\beta}_W^a = \frac{\mathbf{x}^{a'} (\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}) \mathbf{y}^a}{\mathbf{x}^{a'} (\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}) \mathbf{x}^a} = \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a)}{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)^2} \quad (10-20)$$

mit

$$\bar{x}_{i\bullet}^a = \frac{1}{T} \sum_{t=1}^T x_{it}^a, \quad \bar{y}_{i\bullet}^a = \frac{1}{T} \sum_{t=1}^T y_{it}^a.$$

10.4.2 Additive Meßfehler allgemein

Biørn (1996 Kapitel 10.2.2) betrachtet ein lineares Panelmodell, in dem die Regressorvariable x latent ist und nur die meßfehlerbehaftete Variable x^a betrachtet werden kann⁴², und leitet den Wahrscheinlichkeitsgrenzwert des naiven Panelschätzers (10-20) bei additiver Überlagerung ab. Das Fehlermodell lautet⁴³

$$x_{it}^a = x_{it} + u_{it}.$$

Dann ergibt sich für den Wahrscheinlichkeitsgrenzwert⁴⁴

$$\begin{aligned} \text{plim}_{n \rightarrow \infty} \hat{\beta}^a &= \frac{\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a)}{\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)^2} \\ &= \frac{\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet} + u_{it} - \bar{u}_{i\bullet})((x_{it} - \bar{x}_{i\bullet})\beta + \eta_{it} - \bar{\eta}_{i\bullet})}{\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n ((x_{it} - \bar{x}_{i\bullet} + u_{it} - \bar{u}_{i\bullet}))^2} \end{aligned}$$

⁴²Aus Gründen der Einheitlichkeit bleibe ich auch hier bei der Symbolik des Anonymisierungsfalls, verwende also x statt x^* und x^a statt x , also ein "Fehler-in-den-Variablen-Modell" vorliegt.

⁴³Aus Gründen der Einheitlichkeit bleibe ich auch hier bei der Symbolik des Anonymisierungsfalls, verwende also x statt x^* und x^a statt x .

⁴⁴Ich betrachte ausschließlich den Grenzwert für $n \rightarrow \infty$. Deshalb schreibe ich im folgenden oftmals nur plim statt $\text{plim}_{n \rightarrow \infty}$.

Da x , u und η gemäß Annahme miteinander unkorreliert sind, ergibt sich⁴⁵

$$\begin{aligned} \text{plim}_{n \rightarrow \infty} \hat{\beta}^a &= \frac{\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})^2 \beta}{\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})^2 + \text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (u_{it} - \bar{u}_{i\bullet})^2} \\ &= \frac{Q_x}{Q_x + \left(1 - \frac{1}{T}\right) \sigma_u} \beta \end{aligned} \quad (10-21)$$

mit

$$Q_x = \text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})^2$$

Dabei läßt sich unter Zuhilfenahme der Ergebnisse aus Fußnote 45 auch für Q_x eine explizite Form angeben, sofern man einen bestimmten Prozeß für x_t unterstellt. Sofern alle x_t stochastisch voneinander unabhängig sind, ergibt sich demnach

$$Q_x = \left(1 - \frac{1}{T}\right) \sigma_x^2$$

und damit

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}^a = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2} \beta \quad ,$$

d.h. der Schätzer ist inkonsistent, genau wie im Fall der Querschnittsdaten.

In Abschnitt 11 wird der Fall betrachtet, daß die Regressorvariable durch einen autoregressiven Prozeß erzeugt wird. Insbesondere im Fall der multiplikativen Überlagerung ergeben sich dann für kleinen Beobachtungsumfang T Ergebnisse, die dramatisch von den asymptotischen Ergebnissen abweichen. Aus den dort abgeleiteten Ergebnissen ergeben sich im Panelfall deutlich andere Werte für Q_x , wenn der Regressor autokorreliert ist und T klein ist.

10.4.3 Allgemeine multiplikative Überlagerung des Regressors

Im Fall der allgemeinen multiplikativen Überlagerung des Regressors im Fall von Paneldaten verwenden wir

$$x_{it}^a = x_{it} u_{it} \quad .$$

Dann ergibt sich für die Berechnung der Abweichung vom Mittelwert

$$\bar{x}_{i\bullet}^a = \bar{x}_{i\bullet} \bar{u}_{i\bullet}$$

⁴⁵Siehe Biørn (1996 Formel (10-59), Symbolik leicht verändert)). Dabei wird ausgenutzt, daß

$$\sum_{t=1}^T \sum_{i=1}^n (u_{it} - \bar{u}_{i\bullet})^2 = \sum_{t=1}^T \sum_{i=1}^n (u_{it} - \bar{u})^2 - T \sum_{i=1}^n (\bar{u}_{i\bullet} - \bar{u})^2$$

und

$$\text{var}[\bar{u}_{i\bullet}] = \frac{\sigma_u^2}{T}$$

gilt. Damit erhalten wir

$$\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (u_{it} - \bar{u}_{i\bullet})^2 = \text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (u_{it} - \bar{u})^2 - \text{plim} \frac{1}{nT} T \sum_{i=1}^n (\bar{u}_{i\bullet} - \bar{u})^2 = \sigma_u^2 - \frac{\sigma_u^2}{T} .$$

und

$$x_{it}^a - \bar{x}_{i\bullet}^a = x_{it} u_{it} - \bar{x}u_{i\bullet}$$

wobei

$$\bar{x}u_{i\bullet} = \frac{1}{T} \sum_t x_{it} u_{it} .$$

Unter Verwendung der obigen Ergebnisse kann man den naiven Panelschätzer wie folgt schreiben:

$$\begin{aligned} \hat{\beta}_W^a &= \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it} - \bar{y}_{i\bullet})}{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)^2} \\ &= \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it} u_{it} - \bar{x}u_{i\bullet})(y_{it} - \bar{y}_{i\bullet})}{\sum_{t=1}^T \sum_{i=1}^n (x_{it} u_{it} - \bar{x}u_{i\bullet})^2} \\ &= \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it} u_{it} - \bar{x}u_{i\bullet})((x_{it} - \bar{x}_{i\bullet})\beta + \eta_{it} - \bar{\eta}_{i\bullet})}{\sum_{t=1}^T \sum_{i=1}^n (x_{it} u_{it} - \bar{x}u_{i\bullet})^2} \\ &= \frac{\sum_{t=1}^T \sum_{i=1}^n \{(x_{it} u_{it} - \bar{x}u_{i\bullet})(x_{it} - \bar{x}_{i\bullet})\beta + (x_{it} u_{it} - \bar{x}u_{i\bullet})(\eta_{it} - \bar{\eta}_{i\bullet})\}}{\sum_{t=1}^T \sum_{i=1}^n (x_{it} u_{it} - \bar{x}u_{i\bullet})^2} \end{aligned} \quad (10-22)$$

Zur Bestimmung des Wahrscheinlichkeitsgrenzwertes dieses Schätzers sind jeweils die den empirischen Momenten entsprechenden theoretischen Momente zu bestimmen, denn bei Berechnung des "plim" erhalten wir

$$\text{plim}(\hat{\beta}_W^a) = \frac{\frac{1}{T} \sum_{t=1}^T \{\Gamma_t + \Delta_t\}}{\frac{1}{T} \sum_{t=1}^T \{\Psi_t\}} \quad (10-23)$$

wobei

$$\Gamma_t = E[(X_t u_t - \bar{X}u)(X_t - \bar{X})\beta] \quad (10-24)$$

$$\Delta_t = E[(X_t u_t - \bar{X}u)(\eta_t - \bar{\eta})] \quad (10-25)$$

$$\Psi_t = E[(X_t u_t - \bar{X}u)^2] \quad (10-26)$$

gilt. Dabei sind die arithmetischen Mittel über alle T Zeitpunkte zu interpretieren. Beispielsweise gilt

$$\bar{X}u = \frac{1}{T} \sum_t X_t u_t .$$

Wegen der Grenzwertbildung $N \rightarrow \infty$ und der Annahme identischer Verteilung für alle N Beobachtungseinheiten wird der Beobachtungsindex i unterdrückt.

Wir erhalten

$$\Gamma_t = E[(X_t u_t - \bar{X}u)(X_t - \bar{X})] = \left(1 - \frac{1}{T}\right) \sigma_x^2 \beta .$$

sowie

$$\Delta_t = 0 ,$$

weil η_t stochastisch unabhängig von X_t und u_t ist. Außerdem erhalten wir

$$\Psi_t = \left(1 - \frac{1}{T}\right) (\sigma_x^2 + \sigma_u^2(\sigma_x^2 + \mu_x^2)) \quad .$$

Damit ergibt sich für den Wahrscheinlichkeitsgrenzwert von $\hat{\beta}^a$ bei ausschließlicher Überlagerung des Regressor mit der multiplikativen "speziellen" Höhne-Variante folgender Ausdruck:

$$\text{plim}(\hat{\beta}_W^a) = \frac{\sigma_x^2 \beta}{\sigma_x^2 + \sigma_u^2(\sigma_x^2 + \mu_x^2)} \quad . \quad (10-27)$$

Demnach verschwindet der Bias bei der allgemeinen multiplikativen Überlagerung nur dann, wenn

$$\sigma_u^2 = 0$$

gilt.

10.4.4 Allgemeine multiplikative Überlagerung beider (aller) Variablen

Falls auch die abhängige Variable y nach dem Höhneverfahren multiplikativ überlagert wird, gilt für diese überlagerte Variable⁴⁶

$$\begin{aligned} y_{it}^a &= y_{it} u_{ity} \\ &= u_{ity}(\alpha + \beta x_{it} + \tau_i + \eta_{it}) \\ &= (\alpha + \tau_i) u_{ity} + u_{ity}(\beta x_{it} + \eta_{it}) \end{aligned}$$

wobei jetzt u_{ity} anzeigt, daß es sich um die Störvariable für die Überlagerung von y handelt. Für den Mittelwert erhalten wir

$$\overline{y^a}_{i\bullet} = (\alpha + \tau_i) \overline{u}_{i\bullet y} + \beta \overline{xu}_{i\bullet y} + \overline{u\eta}_{i\bullet}$$

wobei

$$\overline{xu}_{i\bullet y} = \frac{1}{T} \sum_t x_{it} u_{itx}$$

und

$$\overline{u\eta}_{i\bullet y} = \frac{1}{T} \sum_t u_{ity} \eta_{it}$$

Dann erhalten wir für die Abweichung vom Mittelwert

$$y_{it}^a - \overline{y^a}_{i\bullet} = (\alpha + \tau_i) (u_{ity} - \overline{u}_{i\bullet y}) + (x_{it} u_{ity} - \overline{xu}_{i\bullet y}) \beta + (u_{ity} \eta_{it} - \overline{u\eta}_{i\bullet y})$$

Um den Wahrscheinlichkeitsgrenzwert von $\hat{\beta}_W^a$ zu bestimmen, ist der Grenzwert des Zählers von (10-20), d.h.

$$\text{plim}_{n \rightarrow \infty} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \overline{x^a}_{i\bullet}) (y_{it}^a - \overline{y^a}_{i\bullet}) \quad ,$$

⁴⁶Wieder unterstellen wir, daß das lineare Modell mit **stochastischen Effekten** gilt. Siehe Abschnitt 10.3.2. Dabei soll keine Korrelation zwischen diesem Effekt und dem Regressor x bestehen. Eine Verletzung dieser Annahme wird in Abschnitt 10.11 behandelt.

Tabelle 10.2: Erläuterungen zum Grenzwert (10-28)

Summanden von $y_{it}^a - \bar{y}_{i\bullet}^a$	$x_{it}^a - \bar{x}_{i\bullet}^a = x_{it} u_{itx} - \bar{x}u_{i\bullet x}$
$(\alpha + \tau_i) (u_{ity} - \bar{u}_{i\bullet y})$	0
$(x_{it} u_{ity} - \bar{x}u_{i\bullet y}) \beta$	0
$(u_{ity} \eta_{it} - \bar{u}\eta_{i\bullet y})$	0
Hinweis: In den einzelnen Zellen wird der jeweilige Grenzwert angegeben.	

neu zu bestimmen, während für den Nenner die Ergebnisse aus dem Abschnitt 10.4.3 übernommen werden können. Dabei verwenden wir jetzt für die Überlagerung von x ebenfalls die Symbolik u_{itx} für die hier relevante Störvariable und schreiben deshalb (10-35) und daraus folgende Ergebnis wie folgt:

$$x_{it}^a = x_{it} u_{itx} \quad ,$$

$$\bar{x}_{i\bullet}^a = \bar{x}u_{i\bullet x}$$

und

$$x_{it}^a - \bar{x}_{i\bullet}^a = x_{it} u_{itx} - \bar{x}u_{i\bullet x}$$

wobei

$$\bar{x}u_{i\bullet x} = \frac{1}{T} \sum_t x_{it} u_{itx}$$

gilt. Im Folgenden wird alternativ die Annahme benutzt, daß (a) u_{itx} und u_{ity} für alle i und t unkorreliert sind bzw. (b) daß die Korrelation ungleich Null ist.

Es werden nun die einzelnen Terme zur Berechnung des Grenzwertes des Zählers von (10-20), d.h. von

$$\text{plim}_{n \rightarrow \infty} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a) \quad , \quad (10-28)$$

systematisch dargestellt. Siehe dazu Tabelle 10.2.

Eine Analyse der einzelnen Kombinationen, die sich als Produkte der einzelnen Summanden ergeben, zeigt, daß nur die Kombination in der zweiten Zeile, die zum Ausdruck

$$((x_{it} u_{ity} - \bar{x}u_{i\bullet y}) (x_{it} u_{itx} - \bar{x}u_{i\bullet x}) \beta$$

führt, einen von Null verschiedenen Grenzwert ergibt.⁴⁷ Unter Verwendung der Formel für die Kovarianz von Produkten von Zufallsvariablen⁴⁸ erhalten wir

$$\text{plim}_{n \rightarrow \infty} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a) = \begin{cases} (1 + \frac{1}{T}) \sigma_x^2 \beta & \sigma_{u_x u_y} = 0, \\ (1 + \frac{1}{T}) (\sigma_x^2 + \sigma_{u_x u_y} (\sigma_x^2 + \mu_x^2)) \beta & \sigma_{u_x u_y} \neq 0 \end{cases}$$

⁴⁷Dabei wird unterstellt, daß der Effekt τ und der Regressor x unkorreliert sind.

⁴⁸ Die Kovarianz für die Produkte UX und WZ unter der Annahme, daß U und W stochastisch unabhängig voneinander sowie von X und Z sind und beide Zufallsvariable Erwartungswert 1 haben, lautet

$$\text{cov}[UX, WZ] = \text{cov}[X, Z] + \text{cov}[U, W] (\text{cov}[X, Z] + E[X]E[Z]) \quad .$$

Dabei gilt in diesem Fall $X = Z$ sowie $\text{cov}[U, W] = 0$ bzw. $\text{cov}[U, W] \neq 0$.

Damit ergibt sich für den Wahrscheinlichkeitsgrenzwert von $\hat{\beta}^a$ bei allgemeiner multiplikativer Überlagerung sowohl des Regressors als auch der abhängigen Variablen folgendes: Im Fall der Unkorreliertheit der Fehler gilt

$$\text{plim}(\hat{\beta}_W^a) = \frac{\sigma_x^2 \beta}{\sigma_x^2 + \sigma_u^2 (\sigma_x^2 + \mu_x^2)} \quad (10-29)$$

d.h. die zusätzliche Überlagerung der abhängigen Variablen verändert den Bias nicht gegenüber der alleinigen Überlagerung des Regressors. Dagegen ergibt sich im Fall der Korreliertheit der Fehler

$$\text{plim}(\hat{\beta}_W^a) = \frac{(\sigma_x^2 + \sigma_{u_x u_y} (\sigma_x^2 + \mu_x^2)) \beta}{\sigma_x^2 + \sigma_u^2 (\sigma_x^2 + \mu_x^2)} \quad (10-30)$$

10.4.5 Meßfehler mit Faktorstruktur

Biørn (1996 Kapitel 10.3) betrachtet außerdem ein lineares Panelmodell, in dem er für das Fehler-Modell eine Varianz-Komponenten-Struktur (error components structure) wie folgt unterstellt:

$$x_{it}^a = x_{it} + \tau_i + \phi_t + \varepsilon_{it} \quad (10-31)$$

Für den naiven Panelschätzer ergibt sich unter der Annahme, daß alle x_t identisch und unabhängig verteilt sind, bei dieser Fehler-Struktur der folgende Wahrscheinlichkeitsgrenzwert:

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}^a = \frac{(1 - \frac{1}{T}) \sigma_x^2}{(1 - \frac{1}{T}) \sigma_x^2 + \sigma_\phi^2 + (1 - \frac{1}{T}) \sigma_\varepsilon^2} \beta \quad (10-32)$$

Bemerkenswert ist, daß durch die Bildung von Abweichungen vom individuen-spezifischen Mittelwert der Effekt τ_i bei der Schätzung keine Rolle spielt, im Gegensatz zum zeit-spezifischen Effekt ϕ_t . Man beachte auch den engen Zusammenhang mit der additiven Überlagerung à la Hönhe in (10-5), wo der zeitspezifische Effekt entfällt. Wir werden später sehen, daß deshalb bei der additiven Überlagerung der Bias des naiven Panelschätzers nur von σ_ε^2 , nicht aber vom Parameter μ abhängt. Siehe dazu Abschnitt 11-45. Entsprechende Ergebnisse für den multiplikativen Fall werden in den Abschnitten 10.4.7 sowie 10.4.8 besprochen. Dort beeinflußt allerdings auch der Zuschlagsparameter (δ) den Bias der Schätzung.

10.4.6 Additive Überlagerung à la Hönhe

Bei der additiven Überlagerung der Regressoren à la Hönhe verwenden wir im Panelfall die Spezifikation (10-5), d.h.

$$x_{it}^a = x_{it} + \mu D_i + \varepsilon_{it}$$

und erhalten wegen

$$\bar{x}_{i\bullet}^a = \bar{x}_{i\bullet} + \mu D_i + \bar{\varepsilon}_{i\bullet}$$

und

$$x_{it}^a - \bar{x}_{i\bullet}^a = x_{it} - \bar{x}_{i\bullet} + \varepsilon_{it} - \bar{\varepsilon}_{i\bullet}$$

als Wahrscheinlichkeitsgrenzwert des "naiven" Panelschätzers

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}^a = \frac{(1 - \frac{1}{T})\sigma_x^2}{(1 - \frac{1}{T})\sigma_x^2 + (1 - \frac{1}{T})\sigma_\varepsilon^2} \beta = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2} \beta, \quad (10-33)$$

d.h. der Bias hängt nur von der Varianz des Restterms ε , nicht aber von der Größenordnung des Zuschlag-Parameters μ ab! Da der (quadrierte) Zuschlagsparameter deutlich größer ist als die Varianz σ_ε^2 , verringert sich der Bias gegenüber der Schätzung im Querschnittsfall (siehe Abschnitt 9.2) beträchtlich. Auch absolut gesehen wird der Bias äußerst gering sein, weil das Verhältnis $\sigma_x/\sigma_\varepsilon$ meist sehr groß sein wird.

Falls auch die abhängige Variable überlagert wird, erhalten wir entsprechend der obigen Analyse

$$x_{it}^a - \bar{x}_{i\bullet}^a = x_{it} - \bar{x}_{i\bullet} + \varepsilon_{itx} - \bar{\varepsilon}_{i\bullet x}$$

und

$$y_{it}^a - \bar{y}_{i\bullet}^a = y_{it} - \bar{y}_{i\bullet} + \varepsilon_{ity} - \bar{\varepsilon}_{i\bullet y}$$

wobei die beiden Überlagerungsvariablen nun angeben, ob sie zu x oder y gehören. Wie in Abschnitt 9.2 ausführlich dargestellt, gilt gemäß Annahme (!!)

$$\text{cov}[\varepsilon_x, \varepsilon_y] = 0.$$

Demnach erhalten wir für den naiven Panelschätzer⁴⁹

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}^a = \frac{(1 - \frac{1}{T})\sigma_x^2 \beta + (1 - \frac{1}{T})\text{cov}[\varepsilon_x, \varepsilon_y]}{(1 - \frac{1}{T})\sigma_x^2 + (1 - \frac{1}{T})\sigma_\varepsilon^2} = \frac{\sigma_x^2 \beta}{\sigma_x^2 + \sigma_\varepsilon^2}, \quad (10-34)$$

Dies entspricht der Formel (10-33) für den Fall, daß nur der Regressor (additiv) überlagert wird. Demnach ist derselbe Korrektorschätzer auch hier verwendbar. Man beachte, daß dieses Ergebnis deutlich von dem im Fall von Querschnittsdaten abweicht. Dort ergab sich eine zusätzliche Verzerrung bei Überlagerung beider Variablen, sofern die beiden Fehler - infolge Verwendung der Mischungsverteilung - miteinander korreliert sind! Siehe Abschnitt 9.2.

10.4.7 Multiplikative Höhle-Überlagerung des Regressors

Im Fall der multiplikativen Überlagerung des Regressors im Fall von Paneldaten verwenden wir statt (10-5) die Spezifikation (10-10), d.h.

$$x_{it}^a = x_{it} (1 + \delta D_i + \varepsilon_{it}) ,$$

⁴⁹Hier wird die folgende Zerlegung verwendet: Für beliebige Zufallsvariablen λ_{itj} mit $\text{cov}[\lambda_{itj}, \lambda_{itk}] = \kappa$, $j \neq k$ gilt

$$\sum_{t=1}^T \sum_{i=1}^n (\lambda_{itx} - \bar{\varepsilon}_{i\bullet x})(\lambda_{ity} - \bar{\varepsilon}_{i\bullet y}) = \sum_{t=1}^T \sum_{i=1}^n (\lambda_{itx} - \bar{\varepsilon}_x)(\lambda_{ity} - \bar{\varepsilon}_y) - T \sum_{i=1}^n (\bar{\varepsilon}_{i\bullet x} - \bar{\varepsilon}_x)(\bar{\varepsilon}_{i\bullet y} - \bar{\varepsilon}_y)$$

und

$$\text{cov}[\bar{\varepsilon}_{i\bullet x}, \bar{\varepsilon}_{i\bullet y}] = \frac{\kappa}{T}$$

gilt. Damit erhalten wir

$$\text{plim} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (\lambda_{itx} - \bar{\varepsilon}_{i\bullet x})(\lambda_{ity} - \bar{\varepsilon}_{i\bullet y}) = \kappa - \frac{\kappa}{T}.$$

Allerdings ist im obigen Fall die Kovarianz gleich Null!

die direkt aus (10-1) folgt. Dann ergibt sich für die Berechnung der Abweichung vom Mittelwert

$$\bar{x}_{i\bullet}^a = (1 + \delta D_i) \bar{x}_{i\bullet} + \bar{x}\bar{\varepsilon}_{i\bullet}$$

und

$$x_{it}^a - \bar{x}_{i\bullet}^a = (1 + \delta D_i) (x_{it} - \bar{x}_{i\bullet}) + x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet} \quad (10-35)$$

wobei

$$\bar{x}\bar{\varepsilon}_{i\bullet} = \frac{1}{T} \sum_t x_{it} \varepsilon_{it} \quad .$$

Im Gegensatz zur additiven Überlagerung verschwindet also der Zuschlagsparameter (δ) in diesem Fall nicht! Er wird also auch bei der Bestimmung des Bias eine Rolle spielen. Dies soll im folgenden abgeleitet werden.

Unter Verwendung der obigen Ergebnisse kann man den naiven Panelschätzer wie folgt schreiben:

$$\begin{aligned} \hat{\beta}_W^a &= \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it} - \bar{y}_{i\bullet})}{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)^2} \\ &= \frac{\sum_{t=1}^T \sum_{i=1}^n ((1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet}) + x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})(y_{it} - \bar{y}_{i\bullet})}{\sum_{t=1}^T \sum_{i=1}^n ((1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet}) + x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})^2} \\ &= \frac{\sum_{t=1}^T \sum_{i=1}^n ((1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet}) + x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})((x_{it} - \bar{x}_{i\bullet})\beta + \eta_{it} - \bar{\eta}_{i\bullet})}{\sum_{t=1}^T \sum_{i=1}^n ((1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet}) + x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})^2} \\ &= \frac{\sum_{t=1}^T \sum_{i=1}^n \left\{ (1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet})^2 \beta + (1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet})(\eta_{it} - \bar{\eta}_{i\bullet}) \right\}}{\sum_{t=1}^T \sum_{i=1}^n \left\{ (1 + \delta D_i)^2 (x_{it} - \bar{x}_{i\bullet})^2 + 2(1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet})(x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet}) + (x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})^2 \right\}} \\ &\quad + \frac{\sum_{t=1}^T \sum_{i=1}^n \left\{ (x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})(x_{it} - \bar{x}_{i\bullet})\beta + (x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})(\eta_{it} - \bar{\eta}_{i\bullet}) \right\}}{\sum_{t=1}^T \sum_{i=1}^n \left\{ (1 + \delta D_i)^2 (x_{it} - \bar{x}_{i\bullet})^2 + 2(1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet})(x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet}) + (x_{it} \varepsilon_{it} - \bar{x}\bar{\varepsilon}_{i\bullet})^2 \right\}} \end{aligned} \quad (10-36)$$

Zur Bestimmung des Wahrscheinlichkeitsgrenzwertes dieses Schätzers sind jeweils die den empirischen Momenten entsprechenden theoretischen Momente zu bestimmen, denn bei Berechnung des "plim" erhalten wir

$$\text{plim}(\hat{\beta}_W^a) = \frac{\frac{1}{T} \sum_{t=1}^T \{A_t + B_t\}}{\frac{1}{T} \sum_{t=1}^T \{\Phi_t + \Pi_t + \Psi_t\}} + \frac{\frac{1}{T} \sum_{t=1}^T \{\Gamma_t + \Delta_t\}}{\frac{1}{T} \sum_{t=1}^T \{\Phi_t + \Pi_t + \Psi_t\}} \quad (10-37)$$

wobei

$$A_t = E[(1 + \delta D)(X_t - \bar{X})^2 \beta] \quad (10-38)$$

$$B_t = E[(1 + \delta D)(X_t - \bar{X})(\eta_t - \bar{\eta})] \quad (10-39)$$

$$\Gamma_t = E[(X_t \varepsilon_t - \bar{X}\bar{\varepsilon})(X_t - \bar{X})\beta] \quad (10-40)$$

$$\Delta_t = E[(X_t \varepsilon_t - \bar{X}\bar{\varepsilon})(\eta_t - \bar{\eta})] \quad (10-41)$$

$$\Phi_t = E[(1 + \delta D)^2 (X_t - \bar{X})^2] \quad (10-42)$$

$$\Pi_t = 2E[(1 + \delta D)(X_t - \bar{X})(X_t \varepsilon_t - \bar{X}\bar{\varepsilon})] \quad (10-43)$$

$$\Psi_t = E[(X_t \varepsilon_t - \bar{X}\bar{\varepsilon})^2] \quad (10-44)$$

gilt. Dabei sind die arithmetischen Mittel über alle T Zeitpunkte zu interpretieren. Beispielsweise gilt

$$\overline{X\varepsilon} = \frac{1}{T} \sum_t X_t \varepsilon_t \quad .$$

Wegen der Grenzwertbildung $N \rightarrow \infty$ und der Annahme identischer Verteilung für alle N Beobachtungseinheiten wird der Beobachtungsindex i unterdrückt.

Unter Beachtung der Bemerkungen in Unterabschnitt 10.4.2 ergibt sich für die einzelnen Ausdrücke folgendes: Wegen der Unabhängigkeit von D und X_t erhalten wir für (10-38)

$$A_t = \left(1 - \frac{1}{T}\right) \sigma_x^2 \beta \quad ,$$

wobei wir $E[D] = 0$ verwendet haben (symmetrischer Fall). Da X_t und η_t unabhängig sind, gilt für (10-39)

$$B_t = 0 \quad .$$

In (10-40) ist die Kovarianz zwischen $X_t \varepsilon_t$ und X_t zu bestimmen. Wegen der stochastischen Unabhängigkeit dieser beiden Zufallsvariablen sowie wegen $E[\varepsilon_t] = 0$ ergibt sich

$$E[(X_t \varepsilon_t - \overline{X\varepsilon})(X_t - \overline{X})] = 0 \quad .$$

Für (10-41) ergibt sich

$$\Delta_t = 0 \quad ,$$

weil η_t stochastisch unabhängig von X_t und ε_t ist. In (10-42) kann wieder die Unabhängigkeit zwischen D und X_t ausgenutzt werden. Weil

$$E[(1 + \delta D)^2] = 1 + \delta^2$$

gilt, erhalten wir

$$\Phi_t = (1 + \delta^2) \left(1 - \frac{1}{T}\right) \sigma_x^2 \quad .$$

Für (10-43) erhalten wir

$$\Pi_t = 0 \quad .$$

Vergleiche dazu die Berechnung zu (10-39) oben. Außerdem kann wieder die Unabhängigkeit zwischen D und X_t ausgenutzt werden. In (10-44) schließlich ist die Varianz des Produktes $X_t \varepsilon_t$ zu bestimmen. Deshalb erhalten wir

$$\Psi_t = \left(1 - \frac{1}{T}\right) \sigma_\varepsilon^2 (\sigma_x^2 + \mu_x^2) \quad .$$

Damit ergibt sich für den Wahrscheinlichkeitsgrenzwert von $\hat{\beta}^a$ bei ausschließlicher Überlagerung des Regressor mit der multiplikativen "speziellen" Höhne-Variante folgender Ausdruck:

$$\text{plim}(\hat{\beta}_W^a) = \frac{\sigma_x^2 \beta}{(1 + \delta^2) \sigma_x^2 + \sigma_\varepsilon^2 (\sigma_x^2 + \mu_x^2)} \quad . \quad (10-45)$$

Demnach verschwindet der Bias bei der multiplikativen Überlagerung nur dann, wenn sowohl der Zuschlag als auch die Restkomponente nicht auf die Variable X einwirken, d.h. wenn

$$\delta = 0 \quad \text{und} \quad \sigma_\varepsilon^2 = 0$$

gilt.

10.4.8 Multiplikative Höhne-Überlagerung beider (aller) Variablen

Falls auch die abhängige Variable y nach dem Höhneverfahren multiplikativ überlagert wird, gilt für diese überlagerte Variable⁵⁰

$$\begin{aligned} y_{it}^a &= y_{it} (1 + \delta D_i + \varepsilon_{ity}) \\ &= (1 + \delta D_i + \varepsilon_{ity})(\alpha + \beta x_{it} + \tau_i + \eta_{it}) \\ &= (1 + \delta D_i)(\alpha + \tau_i) + (1 + \delta D_i)(\beta x_{it} + \eta_{it}) + (\alpha + \tau_i)\varepsilon_{ity} + \varepsilon_{ity}(\beta x_{it} + \eta_{it}) \end{aligned}$$

wobei jetzt ε_{ity} anzeigt, daß es sich um die Störvariable für die Überlagerung von y handelt. Für den Mittelwert erhalten wir

$$\bar{y}_{i\bullet}^a = (1 + \delta D_i)(\alpha + \tau_i) + (1 + \delta D_i)(\beta \bar{x}_{i\bullet} + \bar{\eta}_{i\bullet}) + (\alpha + \tau_i)\bar{\varepsilon}_{i\bullet y} + \beta \bar{x}\bar{\varepsilon}_{i\bullet y} + \bar{\varepsilon}\bar{\eta}_{i\bullet}$$

wobei

$$\bar{x}\bar{\varepsilon}_{i\bullet y} = \frac{1}{T} \sum_t x_{it} \varepsilon_{itx}$$

und

$$\bar{\varepsilon}\bar{\eta}_{i\bullet y} = \frac{1}{T} \sum_t \varepsilon_{ity} \eta_{it}$$

Dann erhalten wir für die Abweichung vom Mittelwert

$$\begin{aligned} y_{it}^a - \bar{y}_{i\bullet}^a &= (1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet})\beta + (1 + \delta D_i)(\eta_{it} - \bar{\eta}_{i\bullet}) + (\alpha + \tau_i)(\varepsilon_{ity} - \bar{\varepsilon}_{i\bullet y}) \\ &\quad + (x_{it}\varepsilon_{ity} - \bar{x}\bar{\varepsilon}_{i\bullet y})\beta + (\varepsilon_{ity}\eta_{it} - \bar{\varepsilon}\bar{\eta}_{i\bullet y}) \end{aligned} \quad (10-46)$$

Um den Wahrscheinlichkeitsgrenzwert von $\hat{\beta}_W^a$ zu bestimmen, ist der Grenzwert des Zählers von (10-20), d.h.

$$\text{plim}_{n \rightarrow \infty} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a) \quad ,$$

unter Verwendung von (10-35) und (10-46) neu zu bestimmen, während für den Nenner die Ergebnisse aus dem Abschnitt 10.4.7 übernommen werden können. Dabei verwenden wir jetzt für die Überlagerung von x ebenfalls die Symbolik ε_{itx} für die hier relevante Störvariable und schreiben deshalb (10-35) und daraus folgende Ergebnis wie folgt:

$$\begin{aligned} x_{it}^a &= x_{it} (1 + \delta D_i + \varepsilon_{itx}) \quad , \\ \bar{x}_{i\bullet}^a &= (1 + \delta D_i)\bar{x}_{i\bullet} + \bar{x}\bar{\varepsilon}_{i\bullet x} \end{aligned}$$

und

$$x_{it}^a - \bar{x}_{i\bullet}^a = (1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet}) + x_{it}\varepsilon_{itx} - \bar{x}\bar{\varepsilon}_{i\bullet x} \quad (10-47)$$

wobei

$$\bar{x}\bar{\varepsilon}_{i\bullet x} = \frac{1}{T} \sum_t x_{it} \varepsilon_{itx}$$

⁵⁰Wieder unterstellen wir, daß das lineare Modell mit **stochastischen Effekten** gilt. Siehe Abschnitt 10.3.2. Dabei soll keine Korrelation zwischen diesem Effekt und dem Regressor x bestehen. Eine Verletzung dieser Annahme wird in Abschnitt 10.11 behandelt.

Tabelle 10.3: Erläuterungen zum Grenzwert (10-48)

Summanden von $y_{it}^a - \bar{y}_{i\bullet}^a$	Summanden von $x_{it}^a - \bar{x}_{i\bullet}^a$	
	$(1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet})$	$x_{it} \varepsilon_{itx} - \bar{x} \bar{\varepsilon}_{i\bullet x}$
$(1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet}) \beta$	$\Phi_t \beta$	0
$(1 + \delta D_i)(\eta_{ity} - \bar{\eta}_{i\bullet y})$	0	0
$(\alpha + \tau_i)(\varepsilon_{ity} - \bar{\varepsilon}_{i\bullet y})$	0	0
$(x_{it} \varepsilon_{it} - \bar{x} \bar{\varepsilon}_{i\bullet}) \beta$	0	0
$(\varepsilon_{ity} \eta_{it} - \bar{\varepsilon} \bar{\eta}_{i\bullet y})$	0	0
<u>Hinweis:</u> In den einzelnen Zellen wird der jeweilige Grenzwert angegeben.		

gilt. Im Folgenden wird die bisher nicht ausdrücklich formulierte Annahme benutzt, daß ε_{itx} und ε_{ity} für alle i und t unkorreliert sind. Siehe jedoch auch die ergänzenden Bemerkungen in Unterabschnitt 10.4.9 für den Fall der Korrelation.

Es werden jetzt die einzelnen Terme zur Berechnung des Grenzwertes des Zählers von (10-20), d.h. von

$$\text{plim}_{n \rightarrow \infty} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a) \quad , \quad (10-48)$$

unter Verwendung von (10-47) und (10-46) systematisch dargestellt. Siehe dazu Tabelle 10.3.

Eine Analyse der einzelnen Kombinationen, die sich als Produkte der einzelnen Summanden ergeben, zeigt, daß nur die Kombination in der linken oberen Zelle, die zum Ausdruck

$$(1 + \delta D_i)^2 (x_{it} - \bar{x}_{i\bullet})^2 \beta$$

führt, einen von Null verschiedenen Grenzwert ergibt. Wie bereits in Unterabschnitt 10.4.7 gezeigt, erhalten wir

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_i ((1 + \delta D_i)^2 (x_{it} - \bar{x}_{i\bullet}))^2 \beta = \Phi_t \beta = (1 + \delta^2) \left(1 + \frac{1}{T}\right) \sigma_x^2 \beta$$

Damit ergibt sich für den Wahrscheinlichkeitsgrenzwert von $\hat{\beta}^a$ bei Überlagerung sowohl des Regressors als auch der abhängigen Variablen mit der multiplikativen "speziellen" Höhne-Variante folgender Ausdruck:

$$\text{plim}(\hat{\beta}_W^a) = \frac{(1 + \delta^2) \sigma_x^2 \beta}{(1 + \delta^2) \sigma_x^2 + \sigma_\varepsilon^2 (\sigma_x^2 + \mu_x^2)} \quad . \quad (10-49)$$

Ein Vergleich mit dem in (10-45) gegebenen Grenzwert bei ausschließlicher Überlagerung des Regressors zeigt, daß die zusätzliche Überlagerung der abhängigen Variablen den Bias um den Faktor $1 + \delta^2$ erhöht. Auch in diesem Fall verschwindet der Bias nur dann, wenn sowohl der Zuschlag als auch die Restkomponente nicht auf die Variable X einwirken, d.h. wenn

$$\delta = 0 \quad \text{und} \quad \sigma_\varepsilon^2 = 0$$

gilt.

Tabelle 10.4: Erläuterungen zum Grenzwert (10-48 für den Fall der Korrelation von ε_{itx} und ε_{ity})

Summanden von $y_{it}^a - \bar{y}_{i\bullet}^a$	Summanden von $x_{it}^a - \bar{x}_{i\bullet}^a$	
	$(1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet})$	$x_{it} \varepsilon_{itx} - \bar{x}_{i\bullet} \varepsilon_{itx}$
$(1 + \delta D_i)(x_{it} - \bar{x}_{i\bullet}) \beta$	$\Phi_t \beta$	0
$(1 + \delta D_i)(\eta_{ity} - \bar{\eta}_{i\bullet y})$	0	0
$(\alpha + \tau_i)(\varepsilon_{ity} - \bar{\varepsilon}_{i\bullet y})$	0	0
$(x_{it} \varepsilon_{it} - \bar{x}_{i\bullet} \varepsilon_{it}) \beta$	0	$Q_t \beta$
$(\varepsilon_{ity} \eta_{it} - \bar{\varepsilon}_{i\bullet y} \eta_{it})$	0	0
Hinweis: In den einzelnen Zellen wird der jeweilige Grenzwert angegeben.		

10.4.9 Eine allgemeinere Form der Höhne-Überlagerung

Es soll hier kurz das Ergebnis aus dem vergangenen Unterabschnitt um die folgende Variante ergänzt werden:

Die Störvariablen ε_{itx} und ε_{ity} sind kontemporär miteinander korreliert.

Dazu betrachten wir nochmals die Tabelle reflatabelle-01 aus Abschnitt 10.4.8, die hier nochmals dargestellt wird.

Die Fehlervariablen aus der Höhne-Spezifikation spielen in den letzten drei Zeilen gemeinsam mit der zweiten Spalte eine Rolle. Allerdings ist die dritte Zeile wegen der unterstellten Unabhängigkeit des Effektes τ sowie die fünfte Zeile wegen der unterstellten Unabhängigkeit von η wiederum nicht relevant. Zu untersuchen ist allein der folgende Ausdrücke:

$$Q_t = E[(X_t \varepsilon_{tx} - \bar{X} \varepsilon_x)(X_t \varepsilon_{ty} - \bar{X} \varepsilon_y)] \beta \tag{10-50}$$

Weil die beiden Fehlervariablen den Erwartungswert Null besitzen, ergibt sich unmittelbar

$$Q_t = \left(1 - \frac{1}{T}\right) \lambda_{xy} (\sigma_x^2 + \mu_x^2) \beta \quad , \tag{10-51}$$

wobei λ_{xy} die Kovarianz zwischen ε_{itx} und ε_{ity} bezeichnet. Demnach erhalten wir als Wahrscheinlichkeitsgrenzwertes des naiven Panelschätzers statt (10-49) nun

$$\text{plim}(\hat{\beta}_W^a) = \frac{(1 + \delta^2) \sigma_x^2 + \lambda_{xy} (\sigma_x^2 + \mu_x^2)}{(1 + \delta^2) \sigma_x^2 + \sigma_\varepsilon^2 (\sigma_x^2 + \mu_x^2)} \beta \quad . \tag{10-52}$$

Wenn $\delta = 0$ gesetzt wird, entspricht dieses Ergebnis dem für die allgemeine Überlagerung in Abschnitt 10.4.4. Siehe insbesondere (10-30).

10.4.10 Multiplikative Überlagerung der abhängigen Variablen

Der Vollständigkeit halber sei noch angemerkt, daß sich bei ausschließlicher Überlagerung der abhängigen Variablen keine Verzerrung ergibt. Dies sieht man beispielsweise dadurch, daß man die Tabelle in Abschnitt 10.4.8 wie folgt modifiziert: Es gibt nur eine Spalte, und zwar mit dem Summanden $x_{it} - \bar{x}_{i\bullet}$, der jeweils mit den Zeilenausdrücken bezüglich der abhängigen Variablen zu kombinieren ist. Der einzige von Null verschiedene Grenzwert

ergibt sich für die oberste Zeile mit $A_t = \sigma_x^2 \beta$. Da auch der Nenner von $\hat{\beta}_W^a$ gegen σ_x^2 strebt, erhalten wir als Grenzwert insgesamt

$$\frac{\sigma_x^2}{\sigma_x^2} \beta = \beta \quad .$$

10.5 Korrekturschätzer

10.5.1 Additive Höhe-Überlagerung

Im Fall der additiven Höhe-Überlagerung gilt für den Wahrscheinlichkeitsgrenzwert des Panelschätzers

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}^a = \frac{(1 - \frac{1}{T})\sigma_x^2}{(1 - \frac{1}{T})\sigma_x^2 + (1 - \frac{1}{T})\sigma_\varepsilon^2} \beta = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2} \beta \quad ,$$

Siehe (10-33) in Abschnitt 10.4.6.

Ein Korrekturschätzer läßt sich wie folgt konstruieren: Wegen (7-5) gilt

$$\text{var}[x^a] = \sigma_x^2 + \mu^2 + \sigma_\varepsilon^2 \quad .$$

Weil außer σ_x^2 alle Größen bekannt sind bzw. geschätzt werden können, erhalten wir den Korrekturschätzer

$$\hat{\beta}^{a,korr} = \frac{\widehat{\text{var}}[x^a] - \hat{\mu}^2 - \frac{1}{T}\sigma_\varepsilon^2}{\widehat{\text{var}}[x^a] - \hat{\mu}^2 - \sigma_\varepsilon^2} \hat{\beta}_W^a \quad (10-53)$$

wobei $\widehat{\text{var}}[x^a]$ eine (konsistente) Schätzung von $\text{var}[x^a]$ ist. In der Korrekturschätzung taucht also auch der Parameter μ auf, der durch

$$\hat{\mu}_x = \bar{x}^a$$

geschätzt werden sollte.

10.5.2 Multiplikative Höhe-Überlagerung

Unter der Annahme, daß dem Datennutzer die Parameter der stochastischen Überlagerung, also δ und σ_ε^2 , bekannt sind, können aus den zuvor abgeleiteten Ergebnissen Korrekturschätzer für die einzelnen Situationen abgeleitet werden. Ich beschränke mich hier beispielhaft auf die Situation, in der sowohl x als auch y multiplikativ mit dem speziellen Höhe-Verfahren überlagert werden. Das entsprechende Ergebnis für den "naiven" Panelschätzer zeigt (10-49). Daraus erhalten wir bei Auflösung nach β folgenden Korrekturschätzer:

$$\hat{\beta}_W^{a,korr} = \frac{(1 + \delta^2) \widehat{\sigma}_x^2 + \sigma_\varepsilon^2 (\widehat{\sigma}_x^2 + \widehat{\mu}_x^2)}{(1 + \delta^2) \widehat{\sigma}_x^2} \hat{\beta}_W^a \quad . \quad (10-54)$$

Dabei ergibt sich unter Verwendung von (10-11) folgende Schätzung für σ_x^2 :

$$\widehat{\sigma}_x^2 = \frac{s_{x^a}^2 - (\delta^2 + \sigma_\varepsilon^2) \bar{x}^{a2}}{1 + \delta^2 + \sigma_\varepsilon^2} \quad , \quad (10-55)$$

wobei

$$\bar{x}^a = \frac{1}{nT} \sum_t \sum_i x_{it}^a$$

und

$$s_{x^a}^2 = \frac{1}{nT} \sum_t \sum_i (x_{it}^a - \bar{x}^a)^2$$

verwendet werden sollten. Ferner sollte

$$\widehat{\mu}_x = \bar{x}^a$$

benutzt werden.

Ein entsprechendes Ergebnis für den Korrektorschätzer im Fall mehrerer Regressoren findet sich in Abschnitt 10.10.

10.6 Analyse für mehrere Regressoren

10.6.1 Allgemeines

Wir betrachten hier das **multiple** Panelmodell mit stochastischen Individualeffekten und mit K Regressoren und einem Absolutglied⁵¹

$$y_{it} = \alpha + \sum_{k=1}^K \beta_k x_{itk} + \tau_i + \eta_{it} \quad i = 1, \dots, n \quad t = 1, \dots, T \quad , \quad (10-56)$$

das zum einfachen Modell aus Abschnitt 10.3.2 korrespondiert. Für Mittelwert und Abweichung vom Mittelwert ergibt sich in diesem Fall

$$\bar{y}_{i\bullet} = \alpha + \sum_{k=1}^K \bar{x}_{i\bullet k} \beta_k + \tau_i + \bar{\eta}_{i\bullet}$$

und

$$y_{it} - \bar{y}_{i\bullet} = \sum_{k=1}^K (x_{itk} - \bar{x}_{i\bullet k}) \beta_k + \eta_{it} - \bar{\eta}_{i\bullet}$$

Im folgenden verwenden wir auch die kompakte Schreibweise dieses Modells:

$$\mathbf{y} = \alpha \mathbf{1} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\tau} + \boldsymbol{\eta} \quad (10-57)$$

Dabei hat die Regressormatrix folgende Struktur:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}\langle 1 \rangle \\ \mathbf{X}\langle 2 \rangle \\ \vdots \\ \mathbf{X}\langle n-1 \rangle \\ \mathbf{X}\langle n \rangle \end{pmatrix}$$

⁵¹Dies entspricht der Formel (10-17) im Fall der Einfachregression.

Der Index in eckigen Klammern bezeichnet einen bestimmten Beobachtungspunkt i . Für jede der n verschiedenen $(T \times K)$ -Matrizen $\mathbf{X}\langle i \rangle$ gilt

$$\mathbf{X}\langle i \rangle = \left(\mathbf{x}_1\langle i \rangle \quad \mathbf{x}_2\langle i \rangle \quad \dots \quad \mathbf{x}_{K-1}\langle i \rangle \quad \mathbf{x}_K\langle i \rangle \right) \quad , \quad i = 1, \dots, n \quad ,$$

wobei der T -dimensionale Vektor $\mathbf{x}_k\langle i \rangle$ durch

$$\mathbf{x}_k\langle i \rangle = \begin{pmatrix} x_{i1k} \\ x_{i2k} \\ \vdots \\ x_{i,T-1,k} \\ x_{iT k} \end{pmatrix} \quad , \quad k = 1, \dots, K \quad ,$$

gegeben ist. Für die Matrix \mathbf{X}^a gilt die entsprechende Symbolik.

Entsprechend schreiben wir den nT -dimensionalen Vektor \mathbf{y} wie folgt:

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}\langle 1 \rangle \\ \mathbf{y}\langle 2 \rangle \\ \vdots \\ \mathbf{y}\langle n-1 \rangle \\ \mathbf{y}\langle n \rangle \end{pmatrix}$$

Dabei ist jeweils der n -dimensionale Vektor $\mathbf{y}\langle i \rangle$ durch

$$\mathbf{y}\langle i \rangle = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{i,T-1} \\ y_{iT} \end{pmatrix}$$

gegeben.

Da der individuenspezifische Effekt nur von i abhängt, ergibt sich für den Zufallsvektor $\boldsymbol{\tau}$ die folgende Form:

$$\boldsymbol{\tau} = \begin{pmatrix} \boldsymbol{\tau}\langle 1 \rangle \\ \boldsymbol{\tau}\langle 2 \rangle \\ \vdots \\ \boldsymbol{\tau}\langle n-1 \rangle \\ \boldsymbol{\tau}\langle n \rangle \end{pmatrix} \quad \text{mit} \quad \boldsymbol{\tau}\langle i \rangle = \tau_i \boldsymbol{\nu}_T$$

10.6.2 Der Schätzer

Die Matrix \mathbf{W} im Panelmodell mit fixen Effekten (siehe Abschnitt 10.3.1) hat folgende Struktur:

$$\mathbf{W} = \mathbf{I}_n \otimes \boldsymbol{\nu}_T$$

Daraus ergibt sich die idempotente Matrix

$$\mathbf{M}_W = \mathbf{I}_{nT} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}' = \mathbf{I}_n \otimes \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right)$$

und für den "within"-Schätzer der Originaldaten ergibt sich

$$\hat{\boldsymbol{\beta}}_W = (\mathbf{X}'\mathbf{M}_W\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}_W\mathbf{y} \quad . \quad (10-58)$$

Dieser Schätzer ist erwartungstreu!

10.6.3 Eine rechentechnisch attraktivere Darstellung des Schätzers

Da die Matrix $NT \times NT$ -Matrix \mathbf{M}_W relativ viel Speicherplatz benötigt, soll hier kurz eine alternative Darstellung präsentiert werden, die ohne diese Matrix auskommt. Zunächst ist zu beachten, daß für die Matrix $\mathbf{M}_W\mathbf{X}$ folgendes gilt:⁵²

$$\mathbf{M}_W\mathbf{X} = \begin{pmatrix} \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{X}\langle 1 \rangle \\ \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{X}\langle 2 \rangle \\ \vdots \\ \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{X}\langle n-1 \rangle \\ \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{X}\langle n \rangle \end{pmatrix}$$

Ferner ergibt sich für die quadratische Form von $\mathbf{X}'\mathbf{M}_W\mathbf{X}$ der Ausdruck

$$\mathbf{X}'\mathbf{M}_W\mathbf{X} = \sum_{i=1}^n \mathbf{X}'\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{X}\langle i \rangle$$

Entsprechend gilt für $\mathbf{X}'\mathbf{M}_W\mathbf{y}$ der Ausdruck

$$\mathbf{X}'\mathbf{M}_W\mathbf{y} = \sum_{i=1}^n \mathbf{X}'\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{y}\langle i \rangle .$$

Somit ergibt sich für den Schätzer alternativ zu (10-58) der Ausdruck

$$\hat{\boldsymbol{\beta}}_W = \left(\sum_{i=1}^n \mathbf{X}'\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{X}\langle i \rangle \right)^{-1} \sum_{i=1}^n \mathbf{X}'\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{y}\langle i \rangle . \quad (10-59)$$

Eine entsprechende Darstellung des "naiven" Panelschätzers bei anonymisierten Daten wird in den Abschnitten 10.7 und 10.8 verwendet.

10.6.4 Kovarianzmatrix des Schätzers

Um die Kovarianzmatrix des Schätzers angeben zu können, müssen wir zunächst die Kovarianzmatrix des Störterms $\boldsymbol{\tau} + \boldsymbol{\eta}$ aus (10-57) angeben:

$$\text{cov}[\boldsymbol{\tau} + \boldsymbol{\eta}] = \text{cov}[\boldsymbol{\eta}] + \text{cov}[\boldsymbol{\tau}] = \mathbf{I}_n \otimes (\sigma_\eta^2 \mathbf{I}_T + \sigma_\tau^2 \boldsymbol{\nu}_T \boldsymbol{\nu}_T')$$

⁵²Dieser Zusammenhang wird in den Abschnitten 10.7 und 10.8 ausgenutzt.

Ferner können wir für den Schätzer schreiben:

$$\hat{\beta}_W = \beta + (\mathbf{X}' \mathbf{M}_W \mathbf{X})^{-1} \mathbf{X}' \mathbf{M}_W (\boldsymbol{\tau} + \boldsymbol{\eta}) \quad ,$$

woraus die zuvor behauptete Erwartungstreue/Unverzerrtheit unmittelbar folgt. (Die Regressoren werden dabei als fix betrachtet; bedingte Betrachtung!)

Daraus erhalten wir die Kovarianzmatrix des Schätzers wie folgt:

$$\text{cov}[\hat{\beta}_W] = (\mathbf{X}' \mathbf{M}_W \mathbf{X})^{-1} \mathbf{X}' \mathbf{M}_W E[(\boldsymbol{\tau} + \boldsymbol{\eta})(\boldsymbol{\tau} + \boldsymbol{\eta})'] \mathbf{M}_W \mathbf{X} (\mathbf{X}' \mathbf{M}_W \mathbf{X})^{-1} \quad . \quad (10-60)$$

wobei $\text{cov}[\boldsymbol{\tau} + \boldsymbol{\eta}] = E[(\boldsymbol{\tau} + \boldsymbol{\eta})(\boldsymbol{\tau} + \boldsymbol{\eta})']$ gilt. Demnach erhalten wir für den "mittleren" Ausdruck $\mathbf{M}_W E[(\boldsymbol{\tau} + \boldsymbol{\eta})(\boldsymbol{\tau} + \boldsymbol{\eta})'] \mathbf{M}_W$:

$$\mathbf{M}_W (\mathbf{I}_n \otimes (\sigma_\eta^2 \mathbf{I}_T + \sigma_\tau^2 \boldsymbol{\nu}_T \boldsymbol{\nu}_T')) \mathbf{M}_W = \sigma_\eta^2 \mathbf{M}_W$$

und damit

$$\text{cov}[\hat{\beta}_W] = \sigma_\eta^2 (\mathbf{X}' \mathbf{M}_W \mathbf{X})^{-1} \quad . \quad (10-61)$$

Man beachte, daß die Varianz des individuenspezifischen Effektes in dieser Formel nicht mehr auftaucht.

Für den Spezialfall nur eines Regressors, den wir in Abschnitt 10.3 betrachten, erhalten wir daraus

$$\text{var}[\hat{\beta}_W] = \frac{\sigma_\eta^2}{(\mathbf{x}' \mathbf{M}_W \mathbf{x})} = \frac{\sigma_\eta^2}{\sum_{t=1}^T \sum_{i=1}^n (x_{it} - \bar{x}_{i\bullet})^2} \quad . \quad (10-62)$$

10.6.5 Eine alternative Ableitung des 'Within'-Schätzers

Das STATA-Handbuch zur statistischen Analyse von Paneldaten⁵³ schlägt folgendes Schätzverfahren für die 'Within'-Schätzung vor: Es soll eine Kleinstquadrateschätzung der Parameter β_k aus dem folgenden Ansatz bestimmt werden:

$$y_{it} - \bar{y}_{i\bullet} + \bar{y}_{\bullet\bullet} = \sum_{k=1}^K \beta_k (x_{itk} - \bar{x}_{i\bullet k} + \bar{x}_{\bullet\bullet k}) + \bar{\tau} + \eta_{it} - \bar{\eta}_{i\bullet k} + \bar{\eta}_{\bullet\bullet k} \quad i = 1, \dots, n, t = 1, \dots, T \quad , \quad (10-63)$$

Dabei gilt

$$\bar{y}_{\bullet\bullet} = \frac{1}{n} \frac{1}{T} \sum_i \sum_t y_{it} \quad ,$$

d.h. $\bar{y}_{\bullet\bullet}$ gibt das 'Gesamtmittel' (grand mean) für y an. Entsprechend sind die Gesamtmittelwerte für den Störterm η und die Regressoren x_k definiert. Ferner gilt

$$\bar{\tau} = \frac{1}{n} \sum_i \tau_i \quad .$$

Wir unterstellen ein Absolutglied und damit

$$x_{it1} = 1 \quad \text{für alle } i \text{ und } t$$

⁵³Siehe STATA-Handbuch (STATA release 8 Cross-Sectional Time Series. Routine XTREG, S. 207.

und damit ergibt sich für (10-63)

$$y_{it} - \bar{y}_{i\bullet} + \bar{y}_{\bullet\bullet} = \beta_1 + \sum_{k=2}^K \beta_k (x_{itk} - \bar{x}_{i\bullet k} + \bar{x}_{\bullet\bullet k}) + \bar{\tau} + (\eta_{it} - \bar{\eta}_{i\bullet k} + \bar{\eta}_{\bullet\bullet k}) \quad , \quad (10-64)$$

Im Folgenden wird gezeigt, daß dies äquivalent zur Schätzung der β -Koeffizienten gemäß (10-58) ist. Wir geben ferner an, wie daraus eine Schätzung für das Absolutglied β_1 gewonnen werden kann.

Wir betrachten zunächst die Matrix

$$\mathbf{A} = (\mathbf{I}_n \otimes \mathbf{I}_T) - \left(\mathbf{I}_n \otimes \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) + \left(\frac{1}{n} \boldsymbol{\nu}_n \boldsymbol{\nu}'_n \otimes \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) \quad (10-65)$$

Wie man nachprüfen kann, ist diese Matrix symmetrisch idempotent und erfüllt

$$\mathbf{A} (\boldsymbol{\nu}_n \otimes \boldsymbol{\nu}_T) = (\boldsymbol{\nu}_n \otimes \boldsymbol{\nu}_T) \quad (10-66)$$

sowie⁵⁴

$$\mathbf{M}_\nu \mathbf{A} = \mathbf{I}_n \otimes \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) = \mathbf{M}_W \quad (10-67)$$

Bei Prämultiplikation des Vektors \mathbf{y} mit \mathbf{A} ergibt sich

$$\mathbf{A} \mathbf{y} = \mathbf{y} - \begin{pmatrix} \bar{y}_{1\bullet} \boldsymbol{\nu}_T \\ \bar{y}_{2\bullet} \boldsymbol{\nu}_T \\ \vdots \\ \bar{y}_{n\bullet} \boldsymbol{\nu}_T \end{pmatrix} + \begin{pmatrix} \bar{\bar{y}}_{\bullet\bullet} \boldsymbol{\nu}_T \\ \bar{\bar{y}}_{\bullet\bullet} \boldsymbol{\nu}_T \\ \vdots \\ \bar{\bar{y}}_{\bullet\bullet} \boldsymbol{\nu}_T \end{pmatrix} \quad ,$$

d.h. es ergibt sich die linke Seite von (10-63) als Vektor.

Ich benutze deshalb nun die Matrix \mathbf{A} für eine zu (10-63) äquivalente kompakte Darstellung, die sich durch Prämultiplikation von (10-57) mit \mathbf{A} ergibt:

$$\mathbf{A} \mathbf{y} = \mathbf{A} \mathbf{X} \boldsymbol{\beta} + \mathbf{A} \boldsymbol{\tau} + \mathbf{A} \boldsymbol{\eta} \quad (10-68)$$

Wegen der Struktur der Matrix \mathbf{X} bei Vorhandensein eines Absolutglieds

$$\mathbf{X} = (\boldsymbol{\nu}_n \otimes \boldsymbol{\nu}_T, \mathbf{X}_2)$$

ergibt sich unter Beachtung der Ergebnisse für \mathbf{A} oben

$$\mathbf{A} \mathbf{y} = \beta_1 (\boldsymbol{\nu}_n \otimes \boldsymbol{\nu}_T) + \mathbf{A} \mathbf{X}_2 \boldsymbol{\beta}_2 + \mathbf{A} \boldsymbol{\tau} + \mathbf{A} \boldsymbol{\eta} \quad (10-69)$$

Der Kleinstquadrateschätzer für $\boldsymbol{\beta}_2$ läßt sich nun als Teilschätzung wie folgt schreiben:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_2 &= ((\mathbf{A} \mathbf{X}_2)' \mathbf{M}_\nu \mathbf{A} \mathbf{X}_2)^{-1} (\mathbf{A} \mathbf{X}_2)' \mathbf{M}_\nu \mathbf{A} \mathbf{y} \\ &= (\mathbf{X}'_2 \mathbf{M}_W \mathbf{X}_2)^{-1} \mathbf{X}'_2 \mathbf{M}_W \mathbf{y} \end{aligned}$$

⁵⁴Der Tn -dimensionale Vektor $\boldsymbol{\nu}$ ist durch

$$\boldsymbol{\nu} = (\boldsymbol{\nu}_n \otimes \boldsymbol{\nu}_T)$$

definiert.

wobei $\mathbf{M}_l \mathbf{A} = \mathbf{M}_W$ verwendet wurde. Dies ist aber genau der in (10-58) gegebene Schätzer.

Zur zusätzlichen Schätzung des Absolutglieds β_1 kann man die alternative Form der Teilschätzung verwenden.⁵⁵ Sie lautet

$$\begin{aligned}\hat{\beta}_1 &= ((\mathbf{A} \boldsymbol{\iota})' \mathbf{A} \boldsymbol{\iota})^{-1} (\mathbf{A} \boldsymbol{\iota})' (\mathbf{A} \mathbf{y} - \mathbf{A} \mathbf{X}_2 \hat{\beta}_2) \\ &= (\mathbf{X}_2' \mathbf{A} \mathbf{X}_2)^{-1} \mathbf{X}_2' (\mathbf{A} \mathbf{y} - \mathbf{A} \mathbf{X}_2 \hat{\beta}_2) \\ &= \frac{1}{nT} \boldsymbol{\iota}' (\mathbf{y} - \mathbf{X}_2 \hat{\beta}_2) \\ &= \bar{y} - \bar{\mathbf{x}}' \hat{\beta}_2\end{aligned}\tag{10-70}$$

wobei $\bar{\mathbf{x}}$ der $(K - 1)$ -dimensionale Vektor der Gesamtmittelwerte (grand means) für die einzelnen Regressoren ist.

10.7 Ausschließliche Überlagerung der Regressoren

Der "naive" Panelschätzer für anonymisierte Regressoren (und Originalwerte der abhängigen Variablen y) ist durch

$$\hat{\beta}_W^a = (\mathbf{X}^a' \mathbf{M}_W \mathbf{X}^a)^{-1} \mathbf{X}^a' \mathbf{M}_W \mathbf{y} \quad .\tag{10-71}$$

oder auch durch

$$\hat{\beta}_W^a = \left(\sum_{i=1}^n \mathbf{X}^{a' \langle i \rangle} \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right) \mathbf{X}^{a \langle i \rangle} \right)^{-1} \sum_{i=1}^n \mathbf{X}^{a' \langle i \rangle} \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right) \mathbf{y}^{a \langle i \rangle} .\tag{10-72}$$

gegeben, wobei die zweite Form in Abschnitt 10.6.3 erläutert wurde.

Wir untersuchen zunächst die Struktur von $\mathbf{M}_W \mathbf{X}^a$. Unter Verwendung der Ergebnisse von oben erhalten wir

$$\mathbf{M}_W \mathbf{X}^a = \begin{pmatrix} \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right) \mathbf{X}^a \langle 1 \rangle \\ \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right) \mathbf{X}^a \langle 2 \rangle \\ \vdots \\ \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right) \mathbf{X}^a \langle n-1 \rangle \\ \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right) \mathbf{X}^a \langle n \rangle \end{pmatrix}$$

⁵⁵Siehe beispielsweise Greene (3. Auflage) Formel (6-24). Beachte: Der Tn -dimensionale Vektor $\boldsymbol{\iota}$ ist durch

$$\boldsymbol{\iota} = (\boldsymbol{\iota}_n \otimes \boldsymbol{\iota}_T)$$

definiert.

und für die quadratische Form von $\mathbf{X}^{\mathbf{a}'\mathbf{M}_W\mathbf{X}^{\mathbf{a}}$ ergibt sich

$$\begin{aligned} \mathbf{X}^{\mathbf{a}'\mathbf{M}_W\mathbf{X}^{\mathbf{a}}} &= \\ &= \sum_{i=1}^n \mathbf{X}^{\mathbf{a}'\langle i \rangle} \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) \mathbf{X}^{\mathbf{a}\langle i \rangle} \\ &= \sum_{i=1}^n \begin{pmatrix} (\mathbf{x}_1^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_1^{\mathbf{a}\langle i \bullet \rangle})^2 & (\mathbf{x}_1^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_1^{\mathbf{a}\langle i \bullet \rangle})' (\mathbf{x}_2^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_2^{\mathbf{a}\langle i \bullet \rangle}) & \dots & (\mathbf{x}_K^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_K^{\mathbf{a}\langle i \bullet \rangle})' (\mathbf{x}_K^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_K^{\mathbf{a}\langle i \bullet \rangle}) \\ (\mathbf{x}_2^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_2^{\mathbf{a}\langle i \bullet \rangle})' (\mathbf{x}_1^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_1^{\mathbf{a}\langle i \bullet \rangle}) & (\mathbf{x}_2^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_2^{\mathbf{a}\langle i \bullet \rangle})^2 & \dots & (\mathbf{x}_2^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_2^{\mathbf{a}\langle i \bullet \rangle})' (\mathbf{x}_K^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_K^{\mathbf{a}\langle i \bullet \rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{x}_K^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_K^{\mathbf{a}\langle i \bullet \rangle})' (\mathbf{x}_1^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_1^{\mathbf{a}\langle i \bullet \rangle}) & (\mathbf{x}_K^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_K^{\mathbf{a}\langle i \bullet \rangle})' (\mathbf{x}_2^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_2^{\mathbf{a}\langle i \bullet \rangle}) & \dots & (\mathbf{x}_K^{\mathbf{a}\langle i \rangle} - \overline{\mathbf{x}}_K^{\mathbf{a}\langle i \bullet \rangle})^2 \end{pmatrix} \\ &= \sum_{i=1}^n \sum_{t=1}^T \begin{pmatrix} (x_{it1}^{\mathbf{a}} - \overline{x}_{i\bullet 1}^{\mathbf{a}})^2 & (x_{it1}^{\mathbf{a}} - \overline{x}_{i\bullet 1}^{\mathbf{a}})(x_{it2}^{\mathbf{a}} - \overline{x}_{i\bullet 2}^{\mathbf{a}}) & \dots & (x_{it1}^{\mathbf{a}} - \overline{x}_{i\bullet 1}^{\mathbf{a}})(x_{itK}^{\mathbf{a}} - \overline{x}_{i\bullet K}^{\mathbf{a}}) \\ (x_{it1}^{\mathbf{a}} - \overline{x}_{i\bullet 1}^{\mathbf{a}})(x_{it2}^{\mathbf{a}} - \overline{x}_{i\bullet 2}^{\mathbf{a}}) & (x_{it2}^{\mathbf{a}} - \overline{x}_{i\bullet 2}^{\mathbf{a}})^2 & \dots & (x_{it2}^{\mathbf{a}} - \overline{x}_{i\bullet 2}^{\mathbf{a}})(x_{itK}^{\mathbf{a}} - \overline{x}_{i\bullet K}^{\mathbf{a}}) \\ \vdots & \vdots & \ddots & \vdots \\ (x_{itK}^{\mathbf{a}} - \overline{x}_{i\bullet K}^{\mathbf{a}})(x_{it1}^{\mathbf{a}} - \overline{x}_{i\bullet 1}^{\mathbf{a}}) & (x_{itK}^{\mathbf{a}} - \overline{x}_{i\bullet K}^{\mathbf{a}})(x_{it2}^{\mathbf{a}} - \overline{x}_{i\bullet 2}^{\mathbf{a}}) & \dots & (x_{itK}^{\mathbf{a}} - \overline{x}_{i\bullet K}^{\mathbf{a}})^2 \end{pmatrix} \end{aligned}$$

wobei die K -dimensionalen Vektoren der individuellen-spezifischen Mittelwerte wie folgt definiert sind:

$$\overline{\mathbf{x}}_k^{\mathbf{a}\langle i \bullet \rangle} = \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \mathbf{x}_k^{\mathbf{a}\langle i \rangle} = \frac{1}{T} \sum_{t=1}^T \begin{pmatrix} x_{it1}^{\mathbf{a}} \\ x_{it2}^{\mathbf{a}} \\ \vdots \\ x_{itK}^{\mathbf{a}} \end{pmatrix}$$

Ich greife nun auf die Ergebnisse aus Abschnitt 10.4.7 zurück, in dem die Überlagerung eines einzigen Regressors betrachtet wurde. Daraus ergibt sich bei Überlagerung mittels (10-1) für den k -ten Regressor

$$x_{itk}^{\mathbf{a}} = x_{itk} (1 + \delta D_i + \varepsilon_{itk}) \quad ,$$

und

$$\overline{x}_{i\bullet k}^{\mathbf{a}} = (1 + \delta D_i) \overline{x}_{i\bullet k} + \overline{x\varepsilon}_{i\bullet k}$$

wobei

$$\overline{x\varepsilon}_{i\bullet k} = \frac{1}{T} \sum_t x_{itk} \varepsilon_{itk} \quad .$$

gilt. Für die Abweichung vom Mittelwert ergibt sich daraus

$$x_{itk}^{\mathbf{a}} - \overline{x}_{i\bullet k}^{\mathbf{a}} = (1 + \delta D_i)(x_{itk} - \overline{x}_{i\bullet k}) + x_{itk} \varepsilon_{itk} - \overline{x\varepsilon}_{i\bullet k} \quad (10-73)$$

Wie bereits in Abschnitt 10.4.7 bemerkt, verschwindet im Gegensatz zur additiven Überlagerung der Zuschlagsparameter (δ) in diesem Fall nicht! Ich habe im erwähnten Abschnitt auch bereits gezeigt, daß

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_{itk}^{\mathbf{a}} - \overline{x}_{i\bullet k}^{\mathbf{a}})^2 = \left(1 - \frac{1}{T}\right) ((1 + \delta^2) \sigma_k^2 + \sigma_\varepsilon^2 (\sigma_k^2 + \mu_k^2)) \quad (10-74)$$

gilt. Siehe die Ergebnisse zu (10-42) bis (10-44). Allerdings schreiben wir dieses Ergebnis jetzt für den k -ten Regressor mit Varianz σ_k^2 und Erwartungswert μ_k .

Zusätzlich benötigen wir jetzt ein ein Ergebnis für

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_{itk}^{\mathbf{a}} - \overline{x}_{i\bullet k}^{\mathbf{a}}) (x_{ith}^{\mathbf{a}} - \overline{x}_{i\bullet h}^{\mathbf{a}}) \quad , k \neq h \quad .$$

Entsprechend dem Vorgehen im erwähnten Abschnitt zerlegen wir wieder den Summenausdruck:

$$\begin{aligned}
& \sum_{i=1}^n (x_{itk}^a - \bar{x}_{i\bullet k}^a) (x_{ith}^a - \bar{x}_{i\bullet h}^a) \\
&= \sum_{i=1}^n \{(1 + \delta D_i) (x_{itk} - \bar{x}_{i\bullet k}) + (x_{itk}\varepsilon_{itk} - \bar{x}\varepsilon_{i\bullet k})\} \{(1 + \delta D_i) (x_{ith} - \bar{x}_{i\bullet h}) + (x_{ith}\varepsilon_{ith} - \bar{x}\varepsilon_{i\bullet h})\} \\
&= \sum_{i=1}^n \{(1 + \delta D_i)^2 (x_{itk} - \bar{x}_{i\bullet k})(x_{ith} - \bar{x}_{i\bullet h}) + (x_{itk}\varepsilon_{itk} - \bar{x}\varepsilon_{i\bullet k})(x_{ith}\varepsilon_{ith} - \bar{x}\varepsilon_{i\bullet h})\} \\
&\quad + \sum_{i=1}^n \{(1 + \delta D_i) (x_{itk} - \bar{x}_{i\bullet k})(x_{ith}\varepsilon_{ith} - \bar{x}\varepsilon_{i\bullet h}) + (1 + \delta D_i) (x_{itk}\varepsilon_{itk} - \bar{x}\varepsilon_{i\bullet k})(x_{ith} - \bar{x}_{i\bullet h})\}
\end{aligned}$$

Aus dieser Zerlegung erkennt man, daß

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_{itk}^a - \bar{x}_{i\bullet k}^a) (x_{ith}^a - \bar{x}_{i\bullet h}^a) = \left(1 - \frac{1}{T}\right) (1 + \delta^2) \sigma_{kh} \quad (10-75)$$

gilt, wobei σ_{kh} die Kovarianz der Regressoren k und h bezeichnet. Man beachte insbesondere, daß im Gegensatz zu (10-44) hier statt der Varianz von $X_{tk}\varepsilon_{tk}$ die Kovarianz zwischen $X_{tk}\varepsilon_{tk}$ und $X_{th}\varepsilon_{th}$ betrachtet wird, die Null ist.

Damit erhalten wir insgesamt

$$\begin{aligned}
n \xrightarrow{\text{plim}} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_W \mathbf{X}^{\mathbf{a}} &= n \xrightarrow{\text{plim}} \frac{1}{n} \sum_{i=1}^n \mathbf{X}^{\mathbf{a}'} \langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{X}^{\mathbf{a}} \langle i \rangle \\
&= \left(1 - \frac{1}{T}\right) \begin{pmatrix} (1 + \delta^2) \sigma_1^2 + \sigma_\varepsilon^2 (\sigma_1^2 + \mu_1^2) & (1 + \delta^2) \sigma_{12} & \dots & (1 + \delta^2) \sigma_{1K} \\ (1 + \delta^2) \sigma_{21} & (1 + \delta^2) \sigma_2^2 + \sigma_\varepsilon^2 (\sigma_2^2 + \mu_2^2) & \dots & (1 + \delta^2) \sigma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ (1 + \delta^2) \sigma_{K1} & (1 + \delta^2) \sigma_{K2} & \dots & (1 + \delta^2) \sigma_K^2 + \sigma_\varepsilon^2 (\sigma_K^2 + \mu_K^2) \end{pmatrix} \quad (10-76)
\end{aligned}$$

Im Gegensatz zur additiven Überlagerung stehen hier nicht zwei Matrizen additiv nebeneinander. Bemerkenswert ist, daß die Diagonalelemente stärker erhöht werden als die Nichtdiagonalelemente!

Um den Wahrscheinlichkeitsgrenzwert für den "naiven" Panelschätzer zu bestimmen, benötigen wir ferner

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_W \mathbf{y}$$

Dazu betrachten wir

$$\begin{aligned}
\mathbf{X}^{\mathbf{a}'} \mathbf{M}_W \mathbf{y} &= \sum_{i=1}^n \mathbf{X}^{\mathbf{a}'} \langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}_T' \right) \mathbf{y} \langle i \rangle \\
&= \sum_{i=1}^n \begin{pmatrix} (\mathbf{x}_1^{\mathbf{a}} \langle i \rangle - \bar{\mathbf{x}}_1^{\mathbf{a}} \langle i \bullet \rangle)' (\mathbf{y} \langle i \rangle - \bar{\mathbf{y}} \langle i \bullet \rangle) \\ (\mathbf{x}_1^{\mathbf{a}} \langle i \rangle - \bar{\mathbf{x}}_1^{\mathbf{a}} \langle i \bullet \rangle)' (\mathbf{y} \langle i \rangle - \bar{\mathbf{y}} \langle i \bullet \rangle) \\ \vdots \\ (\mathbf{x}_1^{\mathbf{a}} \langle i \rangle - \bar{\mathbf{x}}_1^{\mathbf{a}} \langle i \bullet \rangle)' (\mathbf{y} \langle i \rangle - \bar{\mathbf{y}} \langle i \bullet \rangle) \end{pmatrix} \\
&= \sum_{i=1}^n \sum_{t=1}^T \begin{pmatrix} (x_{it1}^a - \bar{x}_{i\bullet 1}^a) (y_{it} - \bar{y}_{i\bullet}) \\ (x_{it2}^a - \bar{x}_{i\bullet 2}^a) (y_{it} - \bar{y}_{i\bullet}) \\ \vdots \\ (x_{itK}^a - \bar{x}_{i\bullet K}^a) (y_{it} - \bar{y}_{i\bullet}) \end{pmatrix}
\end{aligned}$$

Auch hier können die Ergebnisse aus Abschnitt 10.4.7 sowie zusätzlich aus 10.6.1 verwendet werden. Für jeweils ein Element aus dem zuvor bestimmten Vektor ergibt sich

$$\begin{aligned}
& \sum_{i=1}^n \sum_{t=1}^T (x_{itk}^a - \bar{x}_{i\bullet k}) (y_{it} - \bar{y}_{i\bullet}) \\
&= \sum_{i=1}^n \sum_{t=1}^T \left\{ ((1 + \delta D_i) (x_{itk} - \bar{x}_{i\bullet k}) + x_{itk} \varepsilon_{itk} - \bar{x} \bar{\varepsilon}_{i\bullet k}) \left(\sum_{h=1}^K (x_{ith} - \bar{x}_{i\bullet h}) \beta_h + \eta_{it} - \bar{\eta}_{i\bullet} \right) \right\} \\
&= \sum_{t=1}^T \sum_{i=1}^n (1 + \delta D_i) \sum_{h=1}^K (x_{ith} - \bar{x}_{i\bullet h}) (x_{itk} - \bar{x}_{i\bullet k}) \beta_h \\
&\quad + \sum_{t=1}^T \sum_{i=1}^n (1 + \delta D_i) (x_{itk} - \bar{x}_{i\bullet k}) (\eta_{it} - \bar{\eta}_{i\bullet}) \\
&\quad + \sum_{t=1}^T \sum_{i=1}^n \left\{ \sum_{h=1}^K \{ (x_{itk} \varepsilon_{itk} - \bar{x} \bar{\varepsilon}_{i\bullet k}) (x_{ith} - \bar{x}_{i\bullet h}) \beta_h \} \right\} \\
&\quad + \sum_{t=1}^T \sum_{i=1}^n \{ (x_{itk} \varepsilon_{itk} - \bar{x} \bar{\varepsilon}_{i\bullet k}) (\eta_{it} - \bar{\eta}_{i\bullet}) \}
\end{aligned}$$

Eine Analyse der einzelnen Summanden entsprechend Abschnitt 10.4.7 ergibt, daß nur der erste Summand zu einem von Null verschiedenen Grenzwert führt, d.h. für bestimmtes k ergibt sich

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (1 + \delta D_i) \sum_{h=1}^K (x_{ith} - \bar{x}_{i\bullet h}) (x_{itk} - \bar{x}_{i\bullet k}) \beta_h = \left(1 - \frac{1}{T} \right) \sum_{h=1}^K \sigma_{kh} \beta_h$$

Demnach gilt

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^a \mathbf{M}_W \mathbf{y} = \left(1 - \frac{1}{T} \right) \begin{pmatrix} \sum_{h=1}^K \sigma_{1h} \beta_h \\ \sum_{h=1}^K \sigma_{2h} \beta_h \\ \vdots \\ \vdots \\ \sum_{h=1}^K \sigma_{Kh} \beta_h \end{pmatrix} \quad (10-77)$$

Für den Wahrscheinlichkeitsgrenzwert des naiven Panelschätzers (10-58) ergibt sich dann aus (10-76) und (10-77)

$$\begin{aligned}
\text{plim}_{n \rightarrow \infty} \hat{\beta}^a &= \left(\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^a \mathbf{M}_W \mathbf{X}^a \right)^{-1} \text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^a \mathbf{M}_W \mathbf{y} \\
&= \begin{pmatrix} (1 + \delta^2) \sigma_1^2 + \sigma_\varepsilon^2 (\sigma_1^2 + \mu_1^2) & (1 + \delta^2) \sigma_{12} & \dots & (1 + \delta^2) \sigma_{1k} \\ (1 + \delta^2) \sigma_{21} & (1 + \delta^2) \sigma_2^2 + \sigma_\varepsilon^2 (\sigma_2^2 + \mu_2^2) & \dots & (1 + \delta^2) \sigma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ (1 + \delta^2) \sigma_{K1} & (1 + \delta^2) \sigma_{K2} & \dots & (1 + \delta^2) \sigma_K^2 + \sigma_\varepsilon^2 (\sigma_K^2 + \mu_K^2) \end{pmatrix}^{-1} \begin{pmatrix} \sum_{h=1}^K \sigma_{1h} \beta_h \\ \sum_{h=1}^K \sigma_{2h} \beta_h \\ \vdots \\ \vdots \\ \sum_{h=1}^K \sigma_{Kh} \beta_h \end{pmatrix}
\end{aligned} \quad (10-78)$$

Wie auch im Fall eines einzigen Regressors ist dieser Schätzer nur dann konsistent, wenn sowohl $\delta = 0$ als auch $\sigma_\varepsilon^2 = 0$ gilt, d.h. wenn sowohl der Zuschlag δ als auch die Restüberlagerung durch ε entfällt.

10.8 Gemeinsame Überlagerung aller Variablen

Falls alle Variablen überlagert wird (und das wird der Normalfall sein), ist auch die abhängige Variable y in die Anonymisierung einbezogen. Wie in Abschnitt 10.4.8 bereits für einen einzigen Regressor ausgeführt, gilt dann für die anonymisierte abhängige Variable

$$\begin{aligned} y_{it}^a &= y_{it} (1 + \delta D_i + \varepsilon_{ity}) \\ &= (1 + \delta D_i + \varepsilon_{ity}) (\sum_{h=1}^K \beta_h x_{ith} + \tau_i + \eta_{it}) \\ &= (1 + \delta D_i) \tau_i + (1 + \delta D_i) (\sum_{h=1}^K \beta_h x_{ith} + \eta_{it}) + \tau_i \varepsilon_{ity} + \varepsilon_{ity} (\sum_{h=1}^K \beta_h x_{ith} + \eta_{it}) \end{aligned}$$

wobei jetzt ε_{ity} anzeigt, daß es sich um die Störvariable für die Überlagerung von y handelt. Für den Mittelwert erhalten wir

$$\bar{y}_{i\bullet}^a = (1 + \delta D_i) \tau_i + (1 + \delta D_i) \left(\sum_{h=1}^K \beta_h \bar{x}_{i\bullet h} + \bar{\eta}_{i\bullet} \right) + \tau_i \bar{\varepsilon}_{i\bullet y} + \sum_{h=1}^K \beta_h \bar{x}_{i\bullet h y} + \bar{\varepsilon}_{i\bullet y}$$

wobei

$$\bar{x}_{i\bullet ky} = \frac{1}{T} \sum_t x_{itk} \varepsilon_{ity}$$

und

$$\bar{\varepsilon}_{i\bullet y} = \frac{1}{T} \sum_t \varepsilon_{ity} \eta_{it}$$

Dann erhalten wir für die Abweichung vom Mittelwert

$$\begin{aligned} y_{it}^a - \bar{y}_{i\bullet}^a &= (1 + \delta D_i) \sum_{h=1}^K (x_{ith} - \bar{x}_{i\bullet h}) \beta_h + (1 + \delta D_i) (\eta_{it} - \bar{\eta}_{i\bullet}) + \tau_i (\varepsilon_{ity} - \bar{\varepsilon}_{i\bullet y}) \\ &\quad + \sum_{h=1}^K \beta_h (x_{ith} \varepsilon_{ity} - \bar{x}_{i\bullet h y}) + (\varepsilon_{ity} \eta_{it} - \bar{\varepsilon}_{i\bullet y}) \end{aligned} \tag{10-79}$$

Um den Wahrscheinlichkeitsgrenzwert für den "naiven" Panelschätzer zu bestimmen, benötigen wir bei "zusätzlicher" Überlagerung der abhängigen Variablen

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^a{}' \mathbf{M}_W \mathbf{y}^a$$

Dazu betrachten wir

$$\begin{aligned} \mathbf{X}^a{}' \mathbf{M}_W \mathbf{y}^a &= \sum_{i=1}^n \mathbf{X}^a{}' \langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \right) \mathbf{y}^a \langle i \rangle \\ &= \sum_{i=1}^n \begin{pmatrix} (\mathbf{x}_1^a \langle i \rangle - \bar{\mathbf{x}}_1^a \langle i \bullet \rangle)' (\mathbf{y}^a \langle i \rangle - \bar{\mathbf{y}} \langle i \bullet \rangle) \\ (\mathbf{x}_1^a \langle i \rangle - \bar{\mathbf{x}}_1^a \langle i \bullet \rangle)' (\mathbf{y}^a \langle i \rangle - \bar{\mathbf{y}}^a \langle i \bullet \rangle) \\ \vdots \\ (\mathbf{x}_1^a \langle i \rangle - \bar{\mathbf{x}}_1^a \langle i \bullet \rangle)' (\mathbf{y}^a \langle i \rangle - \bar{\mathbf{y}}^a \langle i \bullet \rangle) \end{pmatrix} \\ &= \sum_{i=1}^n \sum_{t=1}^T \begin{pmatrix} (x_{it1}^a - \bar{x}_{i\bullet 1}^a) (y_{it}^a - \bar{y}_{i\bullet}^a) \\ (x_{it2}^a - \bar{x}_{i\bullet 2}^a) (y_{it}^a - \bar{y}_{i\bullet}^a) \\ \vdots \\ (x_{itK}^a - \bar{x}_{i\bullet K}^a) (y_{it}^a - \bar{y}_{i\bullet}^a) \end{pmatrix} \end{aligned}$$

Im folgenden werden die einzelnen Terme zur Berechnung des Grenzwertes von

$$\text{plim}_{n \rightarrow \infty} \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_{itk}^a - \bar{x}_{i \bullet k}^a)(y_{it}^a - \bar{y}_{i \bullet}^a) \quad ,$$

unter Verwendung von (10-73) und (10-79) systematisch dargestellt.

Summanden von $y_{it}^a - \bar{y}_{i \bullet}^a$	Summanden von $x_{it}^a - \bar{x}_{i \bullet}^a$	
	$(1 + \delta D_i) (x_{itk} - \bar{x}_{i \bullet k})$	$x_{itk} \varepsilon_{itk} - \bar{x} \bar{\varepsilon}_{i \bullet k}$
$(1 + \delta D_i) \sum_{h=1}^K (x_{ith} - \bar{x}_{i \bullet h}) \beta_h$	$(1 - \frac{1}{T}) (1 + \delta^2) \sum_{h=1}^K \sigma_{kh} \beta_h$	0
$(1 + \delta D_i) (\eta_{it} - \bar{\eta}_{i \bullet})$	0	0
$\tau_i (\varepsilon_{ity} - \bar{\varepsilon}_{i \bullet y})$	0	0
$\sum_{h=1}^K \beta_h (x_{ith} \varepsilon_{ity} - \bar{x} \bar{\varepsilon}_{i \bullet hy})$	0	0
$(\varepsilon_{ity} \eta_{it} - \bar{\varepsilon} \bar{\eta}_{i \bullet y})$	0	0

Hinweis: In den einzelnen Zellen wird der jeweilige Grenzwert angegeben.

Eine Analyse analog zu der in Abschnitt 10.4.8 zeigt, daß ausschließlich das linke oberste Element einen von Null verschiedenen Grenzwert aufweist. Wir erhalten demnach

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^a \mathbf{M}_W \mathbf{y}^a = \left(1 - \frac{1}{T}\right) (1 + \delta^2) \begin{pmatrix} \sum_{h=1}^K \sigma_{1h} \beta_h \\ \sum_{h=1}^K \sigma_{2h} \beta_h \\ \vdots \\ \vdots \\ \sum_{h=1}^K \sigma_{Kh} \beta_h \end{pmatrix} \quad (10-80)$$

und für den naiven Panelschätzer (10-58) ergibt sich in diesem Fall

$$\begin{aligned} \text{plim}_{n \rightarrow \infty} \hat{\beta}^a &= \left(\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^a \mathbf{M}_W \mathbf{X}^a \right)^{-1} \text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^a \mathbf{M}_W \mathbf{y}^a \\ &= (1 + \delta^2) \begin{pmatrix} (1 + \delta^2) \sigma_1^2 + \sigma_\varepsilon^2 (\sigma_1^2 + \mu_1^2) & (1 + \delta^2) \sigma_{12} & \dots & (1 + \delta^2) \sigma_{1k} \\ (1 + \delta^2) \sigma_{21} & (1 + \delta^2) \sigma_2^2 + \sigma_\varepsilon^2 (\sigma_2^2 + \mu_2^2) & \dots & (1 + \delta^2) \sigma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ (1 + \delta^2) \sigma_{K1} & (1 + \delta^2) \sigma_{K2} & \dots & (1 + \delta^2) \sigma_K^2 + \sigma_\varepsilon^2 (\sigma_K^2 + \mu_K^2) \end{pmatrix}^{-1} \begin{pmatrix} \sum_{h=1}^K \sigma_{1h} \beta_h \\ \sum_{h=1}^K \sigma_{2h} \beta_h \\ \vdots \\ \vdots \\ \sum_{h=1}^K \sigma_{Kh} \beta_h \end{pmatrix} \end{aligned} \quad (10-81)$$

d.h. genau wie im Fall nur eines Regressors verändert sich bei zusätzlicher Überlagerung der abhängigen Variablen der Wahrscheinlichkeitsgrenzwert um den Faktor $(1 + \delta^2)$. Siehe Abschnitt 10.4.8.

10.9 Eine alternative Schreibweise

Zur Konstruktion eines Korrektorschätzers ist es nützlich, die zuvor abgeleiteten Grenzwerte anders darzustellen. Dazu definieren wir zunächst

$$\text{cov}[\mathbf{x}] = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1K} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{K1} & \sigma_{K2} & \dots & \sigma_K^2 \end{pmatrix}$$

Dann können wir für (10-76) schreiben:

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_W \mathbf{X}^{\mathbf{a}} = (1 + \delta^2) \text{cov}[\mathbf{x}] + \sigma_\varepsilon^2 \begin{pmatrix} \sigma_1^2 + \mu_1^2 & & & \\ & \sigma_2^2 + \mu_2^2 & & \\ & & \ddots & \\ & & & \sigma_K^2 + \mu_K^2 \end{pmatrix} \quad (10-82)$$

Entsprechend ergeben sich aus (10-78) und (10-81)

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_W \mathbf{y} = \text{cov}[\mathbf{x}] \boldsymbol{\beta} \quad (10-83)$$

bzw.

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \mathbf{X}^{\mathbf{a}'} \mathbf{M}_W \mathbf{y}^{\mathbf{a}} = (1 + \delta^2) \text{cov}[\mathbf{x}] \boldsymbol{\beta} \quad (10-84)$$

und damit für den naiven Schätzer im Fall der ausschließlichen Überlagerung der Regressoren

$${}_n \text{plim}_{\infty} \hat{\boldsymbol{\beta}}^{\mathbf{a}}_W = \left((1 + \delta^2) \text{cov}[\mathbf{x}] + \sigma_\varepsilon^2 \begin{pmatrix} \sigma_1^2 + \mu_1^2 & & & \\ & \sigma_2^2 + \mu_2^2 & & \\ & & \ddots & \\ & & & \sigma_K^2 + \mu_K^2 \end{pmatrix} \right)^{-1} \text{cov}[\mathbf{x}] \boldsymbol{\beta} \quad (10-85)$$

bzw. bei Überlagerung aller Variablen

$${}_n \text{plim}_{\infty} \hat{\boldsymbol{\beta}}^{\mathbf{a}}_W = \left(\text{cov}[\mathbf{x}] + \frac{\sigma_\varepsilon^2}{(1 + \delta^2)} \begin{pmatrix} \sigma_1^2 + \mu_1^2 & & & \\ & \sigma_2^2 + \mu_2^2 & & \\ & & \ddots & \\ & & & \sigma_K^2 + \mu_K^2 \end{pmatrix} \right)^{-1} \text{cov}[\mathbf{x}] \boldsymbol{\beta} \quad (10-86)$$

Aus diesen Formeln lassen sich Korrekturschätzer bestimmen. Dafür sind wieder geeignete Schätzer für die Mittelwerte und Varianzen sowie Kovarianzen der Regressorvariablen abzuleiten. Dies wird im folgenden Unterabschnitt für den Fall der gemeinsamen Überlagerung aller Variablen im Detail dargestellt.

10.10 Korrekturschätzer

Die Ergebnisse aus Abschnitt 10.9 lassen sich zur Konstruktion eines - konsistenten - Korrekturschätzers verwenden: Aus (10-86) ergibt sich direkt

$$\hat{\boldsymbol{\beta}}_W^{a, \text{korr}} = \text{cov}[\mathbf{x}]^{-1} \left(\text{cov}[\mathbf{x}] + \frac{\sigma_\varepsilon^2}{1 + \delta^2} \begin{pmatrix} \sigma_1^2 + \mu_1^2 & & & \\ & \sigma_2^2 + \mu_2^2 & & \\ & & \ddots & \\ & & & \sigma_K^2 + \mu_K^2 \end{pmatrix} \right) \hat{\boldsymbol{\beta}}_W^a \quad (10-87)$$

für den Fall der gemeinsamen Überlagerung aller Variablen.⁵⁶ Allerdings sind Erwartungswerte, Varianzen und Kovarianzen der Originalvariablen des Vektors \mathbf{x} nicht bekannt. Sie können aber wie folgt aus den Schätzungen der Originalvariablen bestimmt werden:⁵⁷

⁵⁶Entsprechend läßt sich aus (10-85) der Korrekturschätzer für den Fall der Überlagerung nur der Regressoren bestimmen. Wir gehen darauf hier nicht weiter ein.

⁵⁷Siehe auch die entsprechenden Ergebnisse für den Fall eines einzigen Regressors in Abschnitt 10.5.

Aus (7-17) ergibt sich der Zusammenhang zwischen der Kovarianzmatrix eines multiplikativ überlagerten Zufallsvektors und der Kovarianzmatrix des entsprechenden Vektors der Originalvariablen. Dies schreiben wir hier wie folgt:

$$\begin{aligned}\widehat{cov[\mathbf{x}^a]} &= (\sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu}'') \odot (\widehat{cov[\mathbf{x}]} + \widehat{\boldsymbol{\mu}}_x \widehat{\boldsymbol{\mu}}_x') + \widehat{cov[\mathbf{x}]} \\ &= (\sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu}'' + \boldsymbol{\mu}'') \odot \widehat{cov[\mathbf{x}]} + (\sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu}'') \odot \widehat{\boldsymbol{\mu}}_x \widehat{\boldsymbol{\mu}}_x' \quad (10-88)\end{aligned}$$

wobei das Symbol $\widehat{}$ andeutet, daß es sich hier um Schätzungen handelt.

Bei Auflösung nach $\widehat{cov[\mathbf{x}]}$ ergibt sich daraus

$$\begin{aligned}\widehat{cov[\mathbf{x}]} &= \left\{ \widehat{cov[\mathbf{x}^a]} - (\sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu}'') \odot \widehat{\boldsymbol{\mu}}_x \widehat{\boldsymbol{\mu}}_x' \right\} \div (\sigma_\varepsilon^2 \mathbf{I} + (1 + \delta^2) \boldsymbol{\mu}'') \\ &= \left\{ \widehat{cov[\mathbf{x}^a]} - (\sigma_\varepsilon^2 \mathbf{I} + \delta^2 \boldsymbol{\mu}'') \odot \widehat{\boldsymbol{\mu}}_x \widehat{\boldsymbol{\mu}}_x' \right\} \odot \left(\frac{1}{\sigma_\varepsilon^2 + 1 + \delta^2} \mathbf{I} + \frac{1}{1 + \delta^2} (\boldsymbol{\mu}'' - \mathbf{I}) \right), \quad (10-89)\end{aligned}$$

wobei \div in der ersten Zeile die Hadamard-Division bezeichnet. In der zweiten Zeile wurde das entsprechende Hadamard-**Produkt** gebildet.

Man beachte, daß sich daraus im Spezialfall nur eines einzigen Regressors x die Formel (10-54) aus Abschnitt 10.5 ergibt, d.h.

$$\widehat{\sigma}_x^2 = \frac{s_{x^a}^2 - (\delta^2 + \sigma_\varepsilon^2) \bar{x}^{a2}}{1 + \delta^2 + \sigma_\varepsilon^2}.$$

Dabei soll

$$\bar{x}^a = \frac{1}{nT} \sum_t \sum_i x_{it}^a$$

und

$$s_{x^a}^2 = \frac{1}{nT} \sum_t \sum_i (x_{it}^a - \bar{x}^a)^2$$

gelten.

Im obigen multiplen Regressionsmodell sind außerdem die Kovarianzen aus $\widehat{cov[\mathbf{x}^a]}$ nach der üblichen Formel für empirische Momente zu bestimmen. Der unbekannte Vektor $\boldsymbol{\mu}_x$ sollte durch den entsprechenden Vektor der arithmetischen Mittel der anonymisierten Variablen geschätzt werden, d.h.

$$\widehat{\boldsymbol{\mu}}_x = \begin{pmatrix} \bar{x}_1^a \\ \bar{x}_2^a \\ \vdots \\ \bar{x}_{K-1}^a \\ \bar{x}_K^a \end{pmatrix}.$$

Da bei Anonymisierung δ und σ_ε^2 bekannt sind, läßt sich der Korrekturschätzer (10-87) nun berechnen.

10.11 Korrelation des Regressors mit dem individuenpezifischen Effekt

Bekanntlich reagiert der in (10-16) gegebene Within-Schätzer nicht auf eine mögliche Korrelation zwischen Regressor x_{it} und Effekt τ_i , da der Effekt in der Schätzformel gar nicht auftaucht. Siehe dazu die äquivalente Formulierung des Schätzers in (10-18). Dies gilt allerdings natürlich nur für den Fall nicht anonymisierter Variablen. Deshalb ist im Folgenden zu untersuchen, ob sich dieses Ergebnis ändert, wenn stochastisch überlagerte Variablen in der "naiven" Within-Schätzung (10-20) verwendet werden. Dabei werden wir - für additive und multiplikative Variante - zunächst den "allgemeinen" Fall betrachten und dann Folgerungen für den speziellen Fall der Überlagerung mittels Mischungsverteilung (spezielle Höhe-Spezifikation) anstellen.

10.11.1 Additiver Fall

In Abschnitt 10.4.2 habe ich den Bias für den Fall abgeleitet, daß nur der Regressor additiv überlagert wird, d.h. daß

$$x_{it}^a = x_{it} + u_{itx} \quad .$$

und damit

$$\bar{x}_{i\bullet}^a = \bar{x}_{i\bullet} + \bar{u}_{i\bullet x}$$

und

$$x_{it}^a - \bar{x}_{i\bullet}^a = x_{it} - \bar{x}_{i\bullet} + u_{itx} - \bar{u}_{i\bullet x}$$

gilt. Zusätzlich soll nun y additiv überlagert sein, d.h. es soll ebenfalls

$$y_{it}^a = y_{it} + u_{ity} = \alpha + \beta x_{it} + \tau_i + \eta_{it} + u_{ity}$$

und damit

$$\bar{y}_{i\bullet}^a = \alpha + \beta \bar{x}_{i\bullet} + \tau_i + \bar{\eta}_{i\bullet} + \bar{u}_{i\bullet y}$$

sowie

$$y_{it}^a - \bar{y}_{i\bullet}^a = \beta(x_{it} - \bar{x}_{i\bullet}) + \eta_{it} - \bar{\eta}_{i\bullet} + u_{ity} - \bar{u}_{i\bullet y}$$

gelten. Damit ist klar, daß der Ausdruck

$$(x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a)$$

den stochastischen Effekt τ_i nicht enthält und damit die möglicherweise vorhandene Korrelation zwischen x und τ die Bestimmung des Bias nicht berührt. Daß sich die Reststreuung durch die Hinzufügung von u_{ity} erhöht hat, wirkt sich bekanntlich nicht auf die Bestimmung des Grenzwertes aus. Wir erhalten also denselben Wahrscheinlichkeitsgrenzwert wie in Abschnitt 10.4.2 (siehe (10-21)), den wir jetzt wie folgt schreiben:

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}^a = \frac{(1 - \frac{1}{T}) \sigma_x^2}{(1 - \frac{1}{T}) \sigma_x^2 + (1 - \frac{1}{T}) \gamma_x^2} \beta = \frac{\sigma_x^2}{\sigma_x^2 + \gamma_x^2} \beta \quad (10-90)$$

wobei wir jetzt für die Varianz von u_{itx} das Symbol γ_x^2 anstelle von σ_u^2 verwenden.

Falls die beiden Fehlervariablen korreliert sind, ergibt sich entsprechend Abschnitt 9.2 (siehe (9-13)) allerdings ein anderer Grenzwert:

$$\begin{aligned} \text{plim}_{n \rightarrow \infty} \hat{\beta}^a &= \frac{(1 - \frac{1}{T}) \sigma_x^2}{(1 - \frac{1}{T}) \sigma_x^2 + (1 - \frac{1}{T}) \gamma_x^2} \beta + \frac{(1 - \frac{1}{T}) \gamma_{xy}}{(1 - \frac{1}{T}) \sigma_x^2 + (1 - \frac{1}{T}) \gamma_x^2} \\ &= \frac{\sigma_x^2}{\sigma_x^2 + \gamma_x^2} \beta + \frac{\gamma_{xy}}{\sigma_x^2 + \gamma_x^2} \end{aligned}$$

Dabei bezeichnet γ_{xy} die Kovarianz zwischen u_x und u_y .

Allerdings ist diese Aussage nicht allgemein gültig: Wir wissen, daß die gemeinsame Überlagerung mittels der Mischungsverteilung gemäß der speziellen Höhne-Spezifikation Korrelation zwischen den Fehlern u_x und u_y bedingt, nicht aber zwischen den "Restfehlern" ε_x und ε_y und daß der gemeinsame Faktor D , der diese Korrelation bewirkt, durch die Differenzbildung aus den Ausdrücken $u_{itx} - \bar{u}_{i \bullet x}$ und $u_{ity} - \bar{u}_{i \bullet y}$ herausfällt. Deshalb ergibt sich in diesem speziellen Fall **kein** zusätzlicher Effekt auf den Bias durch die Korrelation der Fehlervariablen.

Zusammenfassend kann festgestellt werden, daß im additiven Fall der naive Panelschätzer (10-20) durch die mögliche Korrelation zwischen Regressor x und stochastischem (individuenspezifischen) Effekt τ_i (d.h. $\sigma_{\tau x} \neq 0$) niemals beeinflusst wird. Dagegen ergibt sich bei Korrelation der Fehlervariablen (d.h. $\gamma_{xy} \neq 0$) ein Effekt, sofern die "allgemeine" Überlagerung verwendet wird, nicht aber im speziellen Höhnefall.

10.11.2 Multiplikativer Fall

Wir beginnen auch hier mit der "allgemeinen" Betrachtung, d.h. wir unterstellen

$$x_{it}^a = x_{it} u_{itx} \quad \text{und} \quad y_{it}^a = y_{it} u_{ity}$$

mit

$$V[u_{itx}] = \gamma_x^2 \quad \text{und} \quad V[u_{ity}] = \gamma_y^2 \quad \text{sowie} \quad \text{cov}[u_x u_y] = \gamma_{xy}$$

und erhalten

$$x_{it}^a - \bar{x}_{i \bullet}^a = x_{it} u_{itx} - \bar{x} \bar{u}_{i \bullet x}$$

sowie

$$y_{it}^a - \bar{y}_{i \bullet}^a = (\alpha + \tau_i) (u_{ity} - \bar{u}_{i \bullet y}) + \beta (x_{it} u_{ity} - \bar{x} \bar{u}_{i \bullet y}) + (\eta_{it} u_{ity} - \bar{\eta} \bar{u}_{i \bullet y}) \quad .$$

Man beachte, daß für den Fall $\gamma_{xy} \neq 0$ die beiden Überlagerungsvariablen u_{itx} und u_{ity} miteinander korreliert sind. Im folgenden soll außerdem zugelassen sein, daß x und τ korreliert sind. Die entsprechende Kovarianz bezeichnen wir mit $\sigma_{\tau x}$.

Zur Bestimmung des Wahrscheinlichkeitsgrenzwert von

$$\hat{\beta}_W^a = \frac{\frac{1}{T} \frac{1}{n} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i \bullet}^a) (y_{it}^a - \bar{y}_{i \bullet}^a)}{\frac{1}{T} \frac{1}{n} \sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i \bullet}^a)^2}$$

benötigen wir die relevanten Ausdrücke für Zähler und Nenner, die im folgenden abgeleitet werden. Dabei wird der Faktor $1/T$ sowie die Summation über alle t unterdrückt, weil der Grenzwert für $n \rightarrow \infty$ betrachtet wird.

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \alpha (x_{it} u_{itx} - \bar{x} \bar{u}_{i \bullet x}) (u_{ity} - \bar{u}_{i \bullet y}) = \left(1 - \frac{1}{T}\right) \alpha \mu_x \gamma_{xy} \quad (10-91)$$

Beweis: Für das entsprechende theoretische Moment gilt (unter Vernachlässigung des Parameters α)

$$E[(X_t U_{xt} - \bar{X} \bar{U}_x) (U_{yt} - \bar{U}_y)] = \left(1 - \frac{1}{T}\right) \mu_x \gamma_{xy}$$

Zur Herleitung wird auf Abschnitt B.1 in Appendix B verwiesen. ■

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \tau_i (x_{it} u_{itx} - \bar{x} \bar{u}_{i \bullet x}) (u_{ity} - \bar{u}_{i \bullet y}) = \left(1 - \frac{1}{T}\right)^2 \sigma_{\tau x} \gamma_{xy} \quad (10-92)$$

Beweis: Für das entsprechende theoretische Moment gilt

$$E[\tau (X_t U_{xt} - \bar{X} \bar{U}_x) (U_{yt} - \bar{U}_y)] = \left(1 - \frac{1}{T}\right)^2 \sigma_{\tau x} \gamma_{xy}$$

Zur Herleitung wird auf Abschnitt B.1 in Appendix B verwiesen. ■

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \beta (x_{it} u_{itx} - \bar{x} \bar{u}_{i \bullet x}) (x_{it} u_{ity} - \bar{x} \bar{u}_{i \bullet y}) = \left(1 - \frac{1}{T}\right) \beta (\sigma_x^2 + \gamma_{xy} (\mu_x^2 + \sigma_x^2)) \quad (10-93)$$

Beweis: In diesem Fall gilt für das theoretische Moment (ohne Berücksichtigung von β)

$$E[(X_t U_{xt} - \bar{X} \bar{U}_x) (X_t U_{yt} - \bar{X} \bar{U}_y)] = \left(1 - \frac{1}{T}\right) \sigma_x^2 + \gamma_{xy} (\mu_x^2 + \sigma_x^2) .$$

Zur Herleitung wird auf Abschnitt B.1 in Appendix B verwiesen. ■

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_{it} u_{itx} - \bar{x} \bar{u}_{i \bullet x}) (\eta_{it} u_{ity} - \bar{\eta} \bar{u}_{i \bullet y}) = 0 \quad (10-94)$$

Beweis: In diesem Fall folgt das Resultat unmittelbar aus der Tatsache, daß

$$E[\eta | X, U_x, U_y] = 0$$

gilt. ■

Ferner gilt

$$\text{plim}_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_{it}^a - \bar{x}_{i \bullet}^a)^2 = \left(1 - \frac{1}{T}\right) (\sigma_x^2 + \gamma_x^2 (\sigma_x^2 + \mu_x^2)) , \quad (10-95)$$

was aus der allgemeinen Formel für die Varianz des Produktes der stochastisch unabhängigen Zufallsvariablen X und U_x folgt.

Demnach erhalten wir als Wahrscheinlichkeitsgrenzwert für den "naiven" Within-Schätzer mit $c = 1 - 1/n$:

$$\begin{aligned} \text{plim}_{n \rightarrow \infty} \hat{\beta}_W^a &= \frac{c \alpha \mu_x \gamma_{xy} + c^2 \sigma_{\tau x} \gamma_{xy} + c \beta (\sigma_x^2 + \gamma_{xy} (\mu_x^2 + \sigma_x^2)) + 0}{c (\sigma_x^2 + \gamma_x^2 (\sigma_x^2 + \mu_x^2))} \\ &= \frac{\sigma_x^2 \beta + \{\alpha \mu_x + c \sigma_{\tau x} + (\mu_x^2 + \sigma_x^2) \beta\} \gamma_{xy}}{\sigma_x^2 + \gamma_x^2 (\sigma_x^2 + \mu_x^2)} \end{aligned} \quad (10-96)$$

Formel (10-96) zeigt, daß im Fall einer multiplikativen Überlagerung die Schätzung von β durch die mögliche Korrelation zwischen Regressor x und stochastischem (individuen-spezifischen) Effekt τ_i (d.h. $\sigma_{\tau x} \neq 0$) überhaupt nur dann beeinflusst werden kann, wenn die beiden Fehlervariablen miteinander korreliert sind, d.h. wenn $\gamma_{xy} \neq 0$ gilt. Dann hat übrigens auch das Absolutglied α einen Effekt auf die Schätzung von β . Bemerkenswert ist auch der Effekt des Faktors $c = 1 - 1/T$ auf die Kovarianz $\sigma_{\tau x}$: Für (sehr) kleines T wird der Effekt der Kovarianz (deutlich) **reduziert!**⁵⁸

Allerdings muß genau wie im additiven Fall (Abschnitt 10.11.1) auch hier eine Ergänzung bezüglich der (multiplikativen) Überlagerung mittels der speziellen Höhne-Spezifikation (Mischungsverteilung) erfolgen: Aus dem Studium der Tabelle 10.3 erkennt man, daß der Effekt τ_i in den Ausdrücken

$$\begin{aligned} &(\alpha + \tau_i) (\varepsilon_{ity} - \bar{\varepsilon}_{i \bullet y}) (1 + \delta D_i) (x_{it} - \bar{x}_{i \bullet}) \\ &(\alpha + \tau_i) (\varepsilon_{ity} - \bar{\varepsilon}_{i \bullet y}) (x_{it} \varepsilon_{itx} - \bar{x} \bar{\varepsilon}_{i \bullet x}) \end{aligned}$$

auf den Regressor x "trifft". Da allerdings in beiden Fällen der Restfehler ε_{ity} allein bzw. gemeinsam mit ε_{itx} involviert ist und diese beiden Fehler gemäß Annahme unkorreliert sind, bleiben auch im Fall $\sigma_{\tau x} \neq 0$ die Ergebnisse aus Tabelle 10.3 gültig, d.h. es gilt weiterhin (10-49).

11 Schätzung linearer Panelmodelle bei "kleinem" T

11.1 Einleitung

Panel-Mikro-Daten sind üblicherweise zwar für viele Untersuchungseinheiten (Haushalte, Firmen) verfügbar, allerdings meistens nur für eine begrenzte Anzahl von Zeitpunkten/Perioden, die oftmals auch als "Wellen" bezeichnet werden. In diesem Kapitel sollen die Auswirkungen dieser Konstellation auf die Schätzung linearer Panelmodelle untersucht werden. Dabei ist von besonderer Relevanz, ob die Regressorvariablen autokorreliert sind. Die Effekte der Autokorrelation verstärken sich in gewisser Weise, wenn die Regressoren fehlerüberlagert/-behaftet sind. Die Darstellung in diesem Kapitel beschränkt sich auf das lineare Modell mit einem einzigen Regressor (Lineare Einfachregression) und betrachtet auch nur den Fall - wichtigen - Spezialfall, daß die Regressorvariable X_t einem autoregressiven Prozeß erster Ordnung (AR(1)-Prozeß) folgt.

Um die Auswirkungen im Detail zu studieren, hat es sich als vorteilhaft herausgestellt, zunächst nur den Fall einer eindimensionalen Zeitreihe bzw. ein Panelmodell für eine einzige Untersuchungseinheit ($N = 1$) zu untersuchen. Insbesondere soll die Schätzung der (unbedingten) Varianz dieses Prozesses analysiert werden, also

$$S^2 = \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2$$

sowie deren Äquivalent für fehlerüberlagerte Variablen,

$$S_a^2 = \frac{1}{T} \sum_{t=1}^T (x_t^a - \bar{x}^a)^2$$

⁵⁸Der Effekt einer kleinen Anzahl von Wellen T wird systematisch im folgenden Kapitel 11 behandelt.

wobei sowohl additive als auch multiplikative Überlagerung studiert werden sollen. Wie noch im einzelnen zu zeigen sein wird, spielt das Verhältnis dieser beiden Varianzen sowie das Verhältnis der betreffenden Erwartungswerte

$$\frac{E[S^2]}{E[S_a^2]}$$

eine wichtige Rolle bei der Ableitung der Verzerrung, die sich für kleines T (= Umfang der Zeitreihe bzw. im Panelfall Anzahl der Wellen) ergibt.

Falls man im Panelfall analog zu den obigen Varianzen die "individuellen" Varianzen mit

$$S_i^2 = \frac{1}{T} \sum_{t=1}^T (x_{it} - \bar{x}_{i\bullet})^2$$

sowie deren "anonymisiertes Äquivalent" mit

$$S_{i,a}^2 = \frac{1}{T} \sum_{t=1}^T (x_{it}^a - \bar{x}_{i\bullet}^a)^2$$

bezeichnet, dann sieht man, daß diese Varianzen im bereits ausführlich in Kapitel 10 studierten Within-Schätzer $\hat{\beta}_W$ (siehe (10-16)) und mehr noch im "naiven" Schätzer für anonymisierte Daten

$$\hat{\beta}_W^a = \frac{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)(y_{it}^a - \bar{y}_{i\bullet}^a)}{\sum_{t=1}^T \sum_{i=1}^n (x_{it}^a - \bar{x}_{i\bullet}^a)^2}$$

(siehe (10-20)) im folgenden Sinne eine Rolle: Falls N (= Anzahl der Untersuchungseinheiten im Panelfall) genügend groß ist, gilt für den Wahrscheinlichkeitsgrenzwert dieses Schätzers⁵⁹

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^a = \frac{E[S^2]}{E[S_a^2]} \beta$$

wobei vorausgesetzt ist, daß der Regressor x exogen ist und demnach die Kovarianz zwischen Regressor und Gleichungsfehler für $N \rightarrow \infty$ gegen Null strebt. Da $E[S^2]$ und $E[S_a^2]$ den Effekt kleiner Anzahl Wellen berücksichtigen, erhalten wir damit ein wichtiges Ergebnis bezüglich des "naiven" Panelschätzers $\hat{\beta}_W^a$ für "kleines" T . Allerdings werden dabei zwei wichtige Annahmen gemacht, die die Allgemeingültigkeit des Ergebnisses einschränken:

⁵⁹Dabei wird unterstellt, daß für alle i , $i = 1, \dots, N$, der Erwartungswert der Zufallsvariablen bzw. Schätzfunktion

$$S_i^2 = \frac{1}{T} \sum_{t=1}^T (X_{it} - \bar{X}_{i\bullet})^2$$

durch

$$E[S_i^2] = E[S^2]$$

gegeben ist und daß für die Varianz von

$$S_N^2 = \frac{1}{N} \sum_{i=1}^N S_i^2$$

gilt:

$$V[S_N^2] = \frac{V[S_i^2]}{N} = \frac{V[S^2]}{N}$$

In diesem Fall konvergiert S_N^2 im quadratischen Mittel gegen $E[S^2]$. Dabei ist die genaue Form von $V[S^2]$ unerheblich! Entsprechendes ist für die Schätzung der Varianz bezüglich der anonymisierten Daten zu unterstellen.

- Alle N Regressorvariablen folgen demselben AR(1)-Prozeß.
- Nur die Regressorvariable ist fehlerüberlagert oder die Fehlerüberlagerung der abhängigen Variablen ist unkorreliert mit den Überlagerungen der Regressoren.

Die erste Annahme wird in der Empirie stets nur approximativ erfüllt sein. Gleichwohl vermittelt dieses Ergebnis wichtige Einsichten in das Verhalten von Panelschätzern bei kleinem T . Wesentlicher sind die präsentierten Ergebnisse im Fall der fehlerbehafteten Regressoren: zumindest für den Fall der multiplikativen Überlagerung scheinen diese Ergebnisse bisher noch nicht in der Literatur publiziert worden sein.

Der Fall der Korrelation der Überlagerungsvariablen, der ebenfalls hier per Annahme ausgeschlossen ist, führt zu Ergebnissen, die bisher nicht analytisch abgeleitet wurden, sondern in dieser Arbeit nur durch Simulationsergebnisse illustriert werden.

Die Untersuchung der Eigenschaften des naiven Panelschätzers für kleines T diene vor allem dem Ziel, die Korrektorschätzer, die sich aus den asymptotischen Ergebnissen für den verzerrten naiven Schätzer ableiten lassen, zu verbessern. Eine Alternative stellt der **Instrumentvariablen-Schätzer** (IV-Schätzer) dar, der zumindest asymptotisch konsistent ist. Allerdings werden wir sehen, daß auch dieser Schätzer für kleines T (starke???) Verzerrungen aufweist und demnach diesem Schätzansatz (bzw. der Klasse der GMM-Schätzer insgesamt) nicht die Bedeutung bei der Analyse von Panel-**Mikro**daten zukommt, die ihr im allgemeinen in der Literatur zugemessen wird.

11.2 Autoregressiver Prozeß erster Ordnung

11.2.1 Terminologie

Wir unterstellen für die Zufallsvariable X_t einen autoregressiven Prozeß erster Ordnung (AR(1)-Prozeß, den wir wie folgt schreiben:

$$X_t = \rho_0 + \rho X_{t-1} + \omega_t \quad (11-1)$$

mit

$$E[\omega_t] = 0 \quad \text{und} \quad \text{Var}[\omega_t] = \sigma_\omega^2$$

und

$$|\rho| < 1 \quad .$$

Dann gilt⁶⁰

$$\begin{aligned} E[X_t] &= \frac{\rho_0}{1 - \rho} \\ \text{var}[X_t] &= \frac{\sigma_\omega^2}{1 - \rho^2} \\ \text{cov}[X_t, X_{t-1}] &= \rho \text{var}[X_t] \end{aligned}$$

und demnach

$$\text{corr}[X_t, X_{t-1}] = \rho \quad .$$

⁶⁰Siehe dazu beliebiges Buch über Zeitreihenanalyse.

Allgemeiner gilt

$$\text{cov}[X_t, X_s] = \rho^{|t-s|} \text{var}[X_t]$$

und demnach

$$\text{corr}[X_t, X_s] = \rho^{|t-s|} .$$

Die Auto-Kovarianzen werden wie üblich abkürzend mit

$$\gamma_\tau \equiv \text{cov}[X_t, X_{t+\tau}] = \text{cov}[X_t, X_{t-\tau}]$$

bezeichnet, d.h. γ_0 steht für die Varianz von X_t .

Im folgenden soll untersucht werden, wie sich die geschätzte Varianz

$$S^2 = \frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})^2 \quad (11-2)$$

verhält, wenn X_t einem AR(1)-Prozeß folgt. Dabei ist aus Sicht der Analyse von Paneldaten vor allem die Situation bei kleinem T von Interesse.

11.2.2 Die Schätzfunktion S^2 bei endlichem (kleinem) T

Um den Erwartungswert von S^2 zu bestimmen, schreiben wir diese zunächst wie folgt:

$$S^2 = \frac{1}{T} \sum_{t=1}^T \{ (X_t - \mu)^2 + (\bar{X} - \mu)^2 - 2(X_t - \mu)(\bar{X} - \mu) \} , \quad (11-3)$$

wobei $E[X_t] = \mu$ gilt. Für die einzelnen Summanden soll (unter Vernachlässigung des Faktors 2 und des Minuszeichens) der Erwartungswert bestimmt werden. Für den ersten Summanden erhalten wir unmittelbar

$$E \left[\frac{1}{T} \sum_{t=1}^T (X_t - \mu)^2 \right] = \frac{1}{T} T \text{var}[X_t] = \gamma_0 \quad (11-4)$$

Für den zweiten Summanden greifen wir auf das Ergebnis von Hamilton (1994 S. 187) zurück:⁶¹

$$\begin{aligned} & E \left[\frac{1}{T} \sum_{t=1}^T (\bar{X} - \mu)^2 \right] \\ &= \left(\frac{1}{T} \right)^2 \{ T\gamma_0 + 2(T-1)\gamma_1 + 2(T-2)\gamma_2 + 2(T-3)\gamma_3 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1} \} \end{aligned} \quad (11-5)$$

Für die Bestimmung des Erwartungswertes des dritten Summanden nutzen wir aus, daß

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T (X_t - \mu)(\bar{X} - \mu) &= \frac{1}{T} \sum_{t=1}^T (X_t - \mu) \left(\frac{1}{T} \sum_{s=1}^T X_s - \frac{1}{T} \sum_{s=1}^T \mu \right) \\ &= \left(\frac{1}{T} \right)^2 \sum_{t=1}^T \sum_{s=1}^T (X_t - \mu)(X_s - \mu) \end{aligned}$$

⁶¹Dieses Ergebnis findet sich auch bei Box, Jenkins und Reinsel (1994) Abschnitt 2.1.5, in der Form:

$$V[\bar{X}] = \frac{\gamma_0}{T} \left(1 + 2 \sum_{k=1}^{T-1} \left(1 - \frac{k}{T} \right) \rho_k \right),$$

wobei ρ_k den k -ten Autokorrelationskoeffizienten angibt.

gilt. Somit erhalten wir für den Erwartungswert die Summe aller Elemente aus der Auto-Kovarianzmatrix, multipliziert mit einem Faktor:

$$\begin{aligned} E \left[\frac{1}{T} \sum_{t=1}^T (\bar{X} - \mu)(X_t - \mu) \right] \\ = \left(\frac{1}{T} \right)^2 \{ T\gamma_0 + 2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1} \} \end{aligned} \quad (11-6)$$

Wegen der Gleichheit von (11-6) und (11-5) ergibt sich deshalb für den Erwartungswert von S^2 der folgende Ausdruck:

$$E[S^2] = \gamma_0 - \psi(\gamma, T) \quad (11-7)$$

mit

$$\begin{aligned} \psi(\gamma, T) &= \left(\frac{1}{T} \right)^2 \{ T\gamma_0 + 2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1} \} \\ &= \frac{\gamma_0}{T} \left(1 + 2 \sum_{k=1}^{T-1} \left(1 - \frac{k}{T} \right) \rho_k \right). \end{aligned} \quad (11-8)$$

Zu diesem Ergebnis sind folgende Bemerkungen zu machen:

1. Der Bias-Term ψ geht für $T \rightarrow \infty$ gegen Null, d.h. S^2 ist asymptotisch erwartungstreu.⁶²
2. Falls alle Auto-Kovarianzen **positiv** sind, weist die Schätzung für endliches T einen **negativen** Bias auf. Dagegen ist, falls die Autokovarianzen unterschiedliches Vorzeichen besitzen, über die Richtung des Bias keine allgemein-gültige Aussage möglich. Siehe jedoch die Aussagen weiter unten zum Spezialfall des AR(1)-Prozesses.
3. Im hier betrachteten AR(1)-Prozeß gilt

$$\gamma_\tau = \gamma_0 \rho^\tau, \quad \tau = 1, 2, \dots, T-2, T-1.$$

Demnach hat die Bias-Funktion in diesem speziellen Fall folgende Form:

$$\psi(\gamma, T) = \frac{\gamma_0}{T^2} \{ T + 2(T-1)\rho^1 + 2(T-2)\rho^2 + \dots + 2 \cdot 2\rho^{T-2} + 2 \cdot 1\rho^{T-1} \}. \quad (11-9)$$

4. Falls alle Autokovarianzen/-korrelationen gleich Null sind, d.h. im AR(1)-Prozeß $\rho = 0$ gesetzt wird, ergibt sich

$$E[S^2] = \left(1 - \frac{1}{T} \right) \sigma_\omega^2 < \sigma_\omega^2 = V[X_t] \text{ falls } \gamma_\tau = 0, \tau > 0.$$

5. Biørn (1996 Kap. 10.5.2 Formeln (10-195) bis (10-197)) erhält genau dieselbe Biasfunktion in einem Modell, in dem die Meßfehler autokorreliert sind und der stochastische Prozeß für die Regressorvariable x unspezifiziert bleibt.

⁶²Darüberhinaus ist S^2 auch ein konsistenter Schätzer. Siehe z.B. **Anderson (1971) Kapitel 8.

Tabelle 11.1: Werte der Biasfunktion $\psi(\gamma, T)$ für AR(1)-Prozeß.

ρ	Anzahl Wellen (T)							
	3	4	5	10	30	50	100	500
-0.90000	0.11333	0.02387	0.04225	0.00851	0.00228	0.00125	0.00058	0.00011
-0.50000	0.16667	0.10937	0.08500	0.03777	0.01160	0.00684	0.00338	0.00067
-0.10000	0.29111	0.21487	0.17025	0.08347	0.02746	0.01643	0.00820	0.00164
0.00000	0.33333	0.25000	0.20000	0.10000	0.03333	0.02000	0.01000	0.00200
0.10000	0.38000	0.29013	0.23457	0.11975	0.04047	0.02435	0.01220	0.00244
0.50000	0.61111	0.51562	0.44500	0.26004	0.09556	0.05840	0.02960	0.00598
0.90000	0.91333	0.88113	0.85153	0.72762	0.44181	0.30837	0.17200	0.03728
<u>Hinweis:</u>	$\sigma_\varepsilon^2 = 1 - \rho^2$ und damit $\gamma_0 = 1$ für jedes ρ .							

Die Tabelle 11.1 zeigt für typische Werte von T und ρ den resultierenden Wert der Biasfunktion $\psi(\gamma, T)$ für den speziellen aber auch besonders wichtigen Fall des AR(1)-Prozesses. Dabei wurde aus Gründen der Vergleichbarkeit $\sigma_\varepsilon^2 = 1 - \rho^2$ und damit $\gamma_0 = 1$ gesetzt.⁶³ Damit geben die Werte in der Tabelle jeweils den Bias bezüglich des wahren Parameterwertes $\gamma_0 = 1$ an. Deutlich zu erkennen ist die weit stärkere Verzerrung für großes **positives** ρ . Beachtlich ist auch, daß sich im Fall des hier betrachteten AR(1)-Prozesses auch für negatives ρ stets ein positiver Wert der Biasfunktion ergibt, d.h. auch in diesem Fall ergibt sich eine negative Verzerrung von S^2 . Allerdings ergibt sich laut Tabelle für den Fall des AR(1)-Prozesses stets $E[S^2] > 0$.⁶⁴

11.3 Der AR(1)-Prozeß bei additiver und multiplikativer Überlagerung

11.3.1 Allgemeines

Es sollen nun die Auswirkungen der additiven bzw. multiplikativen Überlagerung auf das stochastische Modell, d.h. den (beobachtbaren) AR(1)-Prozeß, wie auch auf die Schätzung der Parameter untersucht werden.⁶⁵ Wir betrachten demnach weiterhin das Modell (11-1) und zusätzlich die beiden Anonymisierungsvarianten

$$X_t^a = X_t + U_t, \quad E[U_t|X_t] = 0, \quad (11-10)$$

oder

$$X_t^a = X_t \cdot U_t, \quad E[U_t|X_t] = 1 \quad (11-11)$$

Dann erhalten wir für den (bedingten) Erwartungswert von X_t^a

$$E[X_t^a|X_t] = X_t$$

und für den unbedingten Erwartungswert unter Verwendung der Ergebnisse aus Abschnitt 11.2.1

$$E[X_t] = \frac{\rho_0}{1 - \rho}.$$

⁶³Um den Bias für ein beliebiges σ_ε^2 zu bestimmen, ist der jeweilige Wert der Tabelle mit dem Faktor $\sigma_\varepsilon^2/(1 - \rho^2)$ zu multiplizieren. Beispielsweise ergibt sich für $T = 4$ und $\rho = 0.90$ sowie $\sigma_\varepsilon^2 = 1.00$ ein Wert von $0.88113/0.19 = 4.63750$.

⁶⁴Ob dies für beliebigen autoregressiven bzw. linearen stochastischen Prozeß gilt, ist damit noch nicht gesagt!

⁶⁵Siehe dazu vor allem die Arbeit von Staudenmayer und Buonaccorsi(2005), die allerdings nur additive Meßfehler betrachten.

Dies gilt sowohl für den additiven als auch für den multiplikativen Fall. Dagegen ergeben sich Unterschiede bei den zweiten Momenten, die in den folgenden beiden Unterabschnitten behandelt werden sollen.

11.3.2 Additive Überlagerung

Für additive Überlagerung ergibt sich als Varianz von X_t^a wegen der unterstellten stochastischen Unabhängigkeit von X_t und U_t

$$\text{var}[X_t^a] = \text{var}[X_t] + \text{var}[U_t] \quad , \quad (11-12)$$

mit $\text{var}[X_t]$ aus Abschnitt 11.2.1. Bedeutsam ist im Zeitreihenzusammenhang auch die Betrachtung der Auto-Kovarianzen. Es gilt

$$\text{cov}[X_t^a, X_s^a] = \text{cov}[X_t, X_s] + \text{cov}[U_t, U_s] = \text{cov}[X_t, X_s] \quad , \quad t \neq s \quad , \quad (11-13)$$

weil bzw. sofern die U_t keine zeitliche Korrelation aufweisen. Damit erhalten wir unter Verwendung der Ergebnisse aus Abschnitt 11.2.1 für die Autokorrelation der additiv überlagerten Variablen X

$$\text{corr}[X_t^a, X_s^a] = \frac{\text{cov}[X_t, X_s]}{\text{var}[X_t^a]} = \frac{\gamma_0}{\gamma_0 + \sigma_u^2} \rho^{|t-s|} \quad , \quad t \neq s \quad , \quad (11-14)$$

Das bedeutet, daß die Autokorrelation umso stärker gedämpft (d.h. gegen den Wert Null verzerrt) wird, je größer die Fehlervarianz σ_u^2 gewählt wird.

11.3.3 Multiplikative Überlagerung

Im multiplikativen Fall erhalten wir für die Varianz das bekannte Ergebnis (siehe (7-8))

$$\text{var}[X_t^a] = \text{var}[X_t] + \text{var}[U_t] (\text{var}[X_t] + (E[X_t])^2) \quad , \quad (11-15)$$

mit $E[X_t]$ und $\text{var}[X_t]$ aus Abschnitt 11.2.1, d.h. wir erhalten

$$\text{var}[X_t^a] = \gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \quad , \quad \gamma_0 = \frac{\sigma_\omega^2}{1-\rho^2}$$

Für die Autokovarianzen gilt wie im additiven Fall⁶⁶

$$\text{cov}[X_t^a, X_s^a] = \text{cov}[X_t, X_s] \quad (11-16)$$

sowie⁶⁷

$$\text{cov}[X_t, X_s^a] = \text{cov}[X_t, X_s] \quad . \quad (11-17)$$

⁶⁶Wegen der stochastischen Unabhängigkeit von X_t und U_s für alle t und s gilt beispielsweise für $s = t-1$:

$$E[X_t U_t X_{t-1} U_{t-1}] - E[X_t U_t] E[X_{t-1} U_{t-1}] = E[X_t X_{t-1}] - E[X_t] E[X_{t-1}].$$

⁶⁷Die folgende Formel wird bei der Diskussion der Instrumentvariablen-Schätzung eine Rolle spielen.

Demnach erhalten wir ähnlich wie im additiven Fall das Ergebnis, daß die Autokorrelation in der anonymisierten Zeitreihe gegen Null verzerrt ist:

$$\text{corr}[X_t^a, X_s^a] = \frac{\text{cov}[X_t, X_s]}{\text{var}[X_t^a]} = \frac{\gamma_0}{\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\mu}{1-\rho} \right)^2 \right)} \rho^{|t-s|}, \quad t \neq s. \quad (11-18)$$

Speziell für die Autokorrelation erster Ordnung ergibt sich

$$\text{corr}[X_t^a, X_{t-1}^a] = \frac{\text{cov}[X_t, X_{t-1}]}{\text{var}[X_t^a]} = \frac{\gamma_0}{\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\mu}{1-\rho} \right)^2 \right)} \rho. \quad (11-19)$$

11.3.4 Multiplikative Überlagerung mit Faktorstruktur

Falls die Überlagerung von X_t nach der Idee von Höhne durchgeführt wird, lautet das Fehlermodell (siehe Abschnitt 6.1.2)

$$U_t = 1 + \delta D + \varepsilon_t \quad (11-20)$$

mit

$$D = \begin{cases} +1 & \text{mit Wahrscheinlichkeit } 1/2 \\ -1 & \text{mit Wahrscheinlichkeit } 1/2 \end{cases}$$

und

$$E[\varepsilon_t] = 0, \quad V[\varepsilon_t] = \sigma_\varepsilon^2.$$

Man beachte, daß diese Formulierung für eine einzige Zeitreihe gilt, der Panelaspekt also ausgeblendet ist, bzw. $N = 1$ gesetzt ist und daß die Zufallsvariable D für die gesamte Zeitreihe, also für alle T Beobachtungspunkte, nur ein einziges Mal "gezogen" wird!

Für Erwartungswert und Varianz gilt

$$E[U_t] = 1 \quad \text{und} \quad V[U_t] = \delta^2 + \sigma_\varepsilon^2$$

Ferner ergibt sich für die (Auto-)Kovarianz

$$\text{cov}[U_t, U_s] = V[D] = \delta^2 \neq 0, \quad t \neq s, \quad (11-21)$$

d.h. die die Überlagerungsvariablen sind nun zeitlich korreliert, wobei die Autokorrelation **konstant** ist.

Damit erhalten wir für die Varianz und (Auto-)Kovarianz⁶⁸ der multiplikativ überlagerten Variablen X_t^a

$$V[X_t^a] = \gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right), \quad \gamma_0 = \frac{\sigma_\omega^2}{1-\rho^2} \quad (11-22)$$

⁶⁸Für die Berechnung der Kovarianz verwenden wir:

$$E[X_t U_t X_s U_s] = E_x[X_t X_s E_{u|x}[U_t U_s]] = (1 + \delta^2) E_x[X_t X_s], \quad t \neq s,$$

sowie

$$E[X_t U_t] E[X_s U_s] = E[X_t] E[X_s].$$

sowie

$$\begin{aligned}
\text{cov}[X_t^a, X_s^a] &= (1 + \delta^2)E[X_t X_s] - E[X_t] E[X_s] \\
&= \text{cov}[X_t, X_{t-1}] + \delta^2 E_x[X_t X_s] \\
&= \gamma_0 \rho^{|t-s|} + \delta^2 \left\{ (1 + \rho + \dots + \rho^{t-s-1}) \rho_0 \mu_x + \rho^{(t-s)} (\sigma_x^2 + \mu_x^2) \right\} \\
&= \gamma_0 \rho^{|t-s|} + \delta^2 \left\{ (1 + \rho + \dots + \rho^{t-s-1}) \frac{\rho_0^2}{1-\rho} + \rho^{(t-s)} \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right\}
\end{aligned} \tag{11-23}$$

und für die Autokorrelation ergibt sich (für $t \neq s$) demnach

$$\begin{aligned}
\text{corr}[X_t^a, X_s^a] &= \frac{\text{cov}[X_t^a, X_s^a]}{\text{var}[X_t^a]} \\
&= \frac{\gamma_0 \rho^{|t-s|} + \delta^2 \left\{ (1 + \rho + \dots + \rho^{t-s-1}) \frac{\rho_0^2}{1-\rho} + \rho^{(t-s)} \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right\}}{\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right)}
\end{aligned} \tag{11-24}$$

mit $\gamma_0 = \sigma_\omega^2 / (1 - \rho^2)$. Speziell für die Autokorrelation erster Ordnung erhalten wir

$$\begin{aligned}
\text{corr}[X_t^a, X_{t-1}^a] &= \frac{\gamma_0 \rho + \delta^2 \left\{ \frac{\rho_0^2}{1-\rho} + \rho \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right\}}{\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right)} \\
&= \frac{(1 + \delta^2) \rho \gamma_0 + \delta^2 \left(\frac{\rho_0}{1-\rho} \right)^2}{(1 + \delta^2 + \sigma_\varepsilon^2) \gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\frac{\rho_0}{1-\rho} \right)^2}
\end{aligned} \tag{11-25}$$

Diese Formel zeigt, daß vor allem für klein gewählte Varianz σ_ε^2 die additive Konstante

$$\delta^2 \left(\frac{\rho_0}{1-\rho} \right)^2$$

in Zähler und Nenner dafür sorgt, daß die Autokorrelation der überlagerten Zeitreihe tendentiell gegen 1 strebt. Dies gilt besonders für großes ρ_0 und $\rho \rightarrow +1$.

Für X_t und $t > s$ läßt sich unter Verwendung von (11-1) schreiben:

$$\begin{aligned}
X_t &= \rho_0 + \rho X_{t-1} + \omega_t \\
&= \rho_0 + \rho (\rho_0 + \rho X_{t-2} + \omega_{t-1}) + \omega_t \\
&= \rho_0 + \rho (\rho_0 + \rho (\rho_0 + \rho X_{t-3} + \omega_{t-2}) + \omega_{t-1}) + \omega_t \\
&\quad \vdots \\
&= (1 + \rho + \dots + \rho^{t-s-1}) \rho_0 + \rho^{(t-s)} X_s + \sum_{j=0}^{t-s-1} \rho^j \omega_{t-j} .
\end{aligned}$$

Wegen der unterstellten Unabhängigkeit von ω_ℓ und X_k für $k \leq \ell$ erhalten wir (für $t \neq s$)

$$\begin{aligned}
E[X_t X_s] &= (1 + \rho + \dots + \rho^{t-s-1}) \rho_0 E[X_s] + \rho^{(t-s)} E[X_s^2] \\
&= (1 + \rho + \dots + \rho^{t-s-1}) \rho_0 \mu_x + \rho^{(t-s)} (\sigma_x^2 + \mu_x^2) \\
&= (1 + \rho + \dots + \rho^{t-s-1}) \frac{\rho_0^2}{1-\rho} + \rho^{(t-s)} \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right)
\end{aligned}$$

Im Gegensatz zur allgemeinen Überlagerung ist die Autokorrelation auch für den Fall $\rho = 0$ nicht gleich Null, denn wegen den zeitkonstanten Faktors ergibt sich insbesondere für die Autokorrelation erster Ordnung

$$\text{corr}[X_t^a, X_{t-1}^a] = \frac{\delta^2 \rho_0^2}{(1 + \delta^2 + \sigma_\varepsilon^2)\sigma_\omega^2 + (\delta^2 + \sigma_\varepsilon^2)\rho_0^2} \neq 0 \quad \text{für } \rho = 0.$$

Andererseits gilt ganz allgemein im Fall der multiplikativen Höhne-Überlagerung⁶⁹

$$\text{cov}[X_t, X_s^a] = \text{cov}[X_t, X_s] \quad . \quad (11-26)$$

11.4 Schätzung der Varianz bei additiver und multiplikativer Überlagerung

Wir nehmen nun an, daß nur anonymisierte Daten zur Verfügung stehen und verwenden an Stelle der Schätzfunktion S^2 nun

$$S_a^2 = \frac{1}{T} \sum_{t=1}^T (X_t^a - \bar{X}^a)^2 \quad (11-27)$$

Weiterhin soll gelten, daß die Variable X_t durch einen AR(1)-Prozeß erzeugt wird.

11.4.1 Additive Überlagerung

Im additiven Fall erhalten wir⁷⁰

$$\begin{aligned} E[S_a^2] &= \frac{1}{T} \sum_{t=1}^T E[(X_t^a - \bar{X}^a)^2] \\ &= E[S^2] + (1 - \frac{1}{T})\sigma_u^2 \\ &= \gamma_0 - \psi(\gamma, T) + (1 - \frac{1}{T})\sigma_u^2. \end{aligned} \quad (11-28)$$

Dies Ergebnis ist wie folgt zu interpretieren: Mit der "naiven" Varianzschätzung S_a^2 wird der **Erwartungswert von S^2** systematisch **überschätzt**. Andererseits wissen wir aus obiger Analyse, daß zumindest bei positiver Autokorrelation die wahre Varianz γ_0 systematisch **unterschätzt** wird. Demnach läßt sich über die Richtung der Verzerrung nichts allgemein Gültiges sagen. Man beachte jedoch, daß für bestimmte Parameter-Konstellationen die Bias-Funktion $\psi(\gamma, T)$ in Tabelle 11.1 dargestellt wird.

Im Fall $\rho = 0$ ergibt sich aus obiger Formel

$$E[S_a^2] = \left(1 - \frac{1}{T}\right) (\sigma_\omega^2 + \sigma_u^2)$$

⁶⁹Die folgende Formel wird bei der Diskussion der Instrumentvariablen-Schätzung eine Rolle spielen.

⁷⁰Es gilt

$$\begin{aligned} E\left[\frac{1}{T} \sum_{t=1}^T (U_t - \bar{U})^2\right] &= E\left[\frac{1}{T} \sum_{t=1}^T (U_t - E[U_t])^2\right] + E\left[\frac{1}{T} \sum_{t=1}^T (\bar{U} - E[U_t])^2\right] \\ &\quad - 2E\left[\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T (U_t - E[U_t])(U_s - E[U_s])\right] \\ &= \sigma_u^2 + \frac{\sigma_u^2}{T} - 2\frac{\sigma_u^2}{T} \end{aligned}$$

11.4.2 Multiplikative Überlagerung

Im Fall der multiplikativen Überlagerung ist der Erwartungswert von

$$S_a^2 = \frac{1}{T} \sum_{t=1}^T (X_t^a - \overline{X^a})^2 = \frac{1}{T} \sum_{t=1}^T \left(X_t U_t - \frac{1}{T} \sum_{s=1}^T X_s U_s \right)^2$$

zu bestimmen. Weil $E[X_t U_t] = E[X_t] = \mu_x$ gilt, erweitern wir den obigen Ausdruck wie folgt:

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \left(X_t U_t - \frac{1}{T} \sum_{s=1}^T X_s U_s \right)^2 &= \frac{1}{T} \sum_{t=1}^T \left(X_t U_t - \mu_x + \mu_x - \frac{1}{T} \sum_{s=1}^T X_s U_s \right)^2 \\ &= \frac{1}{T} \sum_{t=1}^T (X_t U_t - \mu_x)^2 + \frac{1}{T} \sum_{t=1}^T \left(\frac{1}{T} \sum_{s=1}^T X_s U_s - \mu_x \right)^2 \\ &\quad - 2 \frac{1}{T} \sum_{t=1}^T (X_t U_t - \mu_x) \left(\frac{1}{T} \sum_{s=1}^T X_s U_s - \mu_x \right) \\ &= \frac{1}{T} \sum_{t=1}^T (X_t U_t - \mu_x)^2 + \frac{1}{T^2} \left(\sum_{s=1}^T (X_s U_s - \mu_x) \right)^2 \\ &\quad - 2 \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T (X_t U_t - \mu_x) (X_s U_s - \mu_x) \end{aligned}$$

Für den ersten Summanden erhalten wir

$$E\left[\frac{1}{T} \sum_{t=1}^T (X_t U_t - \mu_x)^2\right] = \frac{1}{T} T \operatorname{var}[X_t U_t] = \gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right)$$

Für den zweiten Summanden ergibt sich zunächst

$$\begin{aligned} &E\left[\frac{1}{T^2} \left(\sum_{s=1}^T (X_s U_s - \mu_x) \right)^2\right] \\ &= \frac{1}{T^2} \left(\sum_{s=1}^T \operatorname{var}[X_s U_s] + \sum_{s=1}^T \sum_{s'=1, s' \neq s}^T \operatorname{cov}[X_s U_s, X_{s'} U_{s'}] \right) \\ &= \frac{1}{T^2} \left(\sum_{s=1}^T \operatorname{var}[X_s U_s] + \sum_{s=1}^T \sum_{s'=1, s' \neq s}^T \operatorname{cov}[X_s, X_{s'}] \right) \end{aligned}$$

wobei das Ergebnis in der dritten Zeile ausnutzt, daß bei multiplikativer Überlagerung

$$\operatorname{cov}[X_s U_s, X_{s'} U_{s'}] = \operatorname{cov}[X_s, X_{s'}]$$

gilt. Siehe (11-16). Damit erhalten wir

$$\begin{aligned} &E\left[\frac{1}{T^2} \left(\sum_{s=1}^T (X_s U_s - \mu_x) \right)^2\right] \\ &= \frac{1}{T} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{T^2} (2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1}) \end{aligned}$$

Für den dritten Summanden ergibt sich (unter Vernachlässigung des Faktors -2)

$$\begin{aligned} &E\left[\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T (X_t U_t - \mu_x) (X_s U_s - \mu_x)\right] \\ &= \frac{1}{T^2} \sum_{t=1}^T \operatorname{var}[X_t U_t] + \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1, s \neq t}^T \operatorname{cov}[X_t, X_s] \\ &= \frac{1}{T} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{T^2} \{2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1}\} \end{aligned}$$

Da wie bei den Originalvariablen (siehe Abschnitt 11.2.2) die Erwartungswerte des zweiten und dritten Summanden (bis auf den Faktor -2) identisch sind, ergibt sich für den Erwartungswert der "naiven" Varianzschätzung bei multiplikativer Überlagerung:

$$E[S_a^2] = \gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) - \frac{1}{T} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) - \frac{1}{T^2} \{2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1}\} \quad (11-29)$$

oder auch

$$E[S_a^2] = \gamma_0 + \chi_1(\gamma_0, \rho_0, \rho, \sigma_u^2) - \chi_2(\gamma_0, \rho_0, \rho, \sigma_u^2, T) - \chi_3(\gamma_1, \gamma_2, \dots, \gamma_{T-1}, T) \quad (11-30)$$

mit

$$\begin{aligned} \chi_1(\gamma_0, \rho_0, \rho, \sigma_u^2) &= \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \\ \chi_2(\gamma_0, \rho_0, \rho, \sigma_u^2, T) &= \frac{1}{T} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) \\ \chi_3(\gamma_1, \gamma_2, \dots, \gamma_{T-1}, T) &= \frac{1}{T^2} \{2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1}\} \end{aligned}$$

Zu diesem Ergebnis sind folgende Bemerkungen zu machen:

1. Die Bias-Terme χ_2 und χ_3 gehen für $T \rightarrow \infty$ gegen Null. Da jedoch der Bias-Term χ_1 nicht diese Eigenschaft besitzt, ist S^2 **nicht** asymptotisch erwartungstreu, sondern wird wegen $\chi_1 > 0$ systematisch überschätzt. Diese Aussage läßt allerdings die Effekte der Bias-Terme χ_2 und χ_3 unberücksichtigt.
2. Der Bias-Term χ_2 ist stets positiv. Da er mit negativem Vorzeichen eingeht, wird dadurch also die von χ_1 verursachte Überschätzung gemildert. Wegen

$$\chi_1 - \chi_2 = \left(1 - \frac{1}{T} \right) \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) - \frac{\gamma_0}{T}$$

ist zu vermuten, daß der gemeinsame Effekt vermutlich einen positiven Bias erzeugt.

3. Falls alle Auto-Kovarianzen **positiv** sind, ergibt sich für endliches T ein negativer Effekt durch den Bias-Term χ_3 , der ebenfalls mit negativem Vorzeichen eingeht.
4. Über die Richtung des Bias insgesamt läßt sich also auch bei durchweg positiven Autokovarianzen keine allgemein-gültige Aussage machen, während sich für die Schätzfunktion S^2 in diesem Fall ein negativer Bias ergibt!
5. Im hier betrachteten AR(1)-Prozeß gilt

$$\gamma_\tau = \gamma_0 \rho^\tau \quad , \quad \tau = 1, 2, \dots, T-2, T-1 .$$

Demnach hat die Bias-Funktion χ_3 in diesem speziellen Fall folgende Form:

$$\chi_3(\gamma, T) = \frac{\gamma_0}{T^2} \{2(T-1)\rho^1 + 2(T-2)\rho^2 + \dots + 2 \cdot 2\rho^{T-2} + 2 \cdot 1\rho^{T-1}\} . \quad (11-31)$$

und für $\rho > 0$ gilt also stets $\chi_3 > 0$.

Tabelle 11.2: Werte der Biasfunktion $\chi(\gamma, T)$ für AR(1)-Prozeß.

ρ	Anzahl Wellen (T)							
	3	4	5	10	30	50	100	500
$\sigma_u = 0.10$ und $\rho_0 = 0$								
-0.90000	-0.10667	-0.01637	-0.03425	0.00049	0.00738	0.00855	0.00932	0.00987
-0.50000	-0.16000	-0.10187	-0.07700	-0.02877	-0.00194	0.00296	0.00652	0.00931
-0.10000	-0.28444	-0.20737	-0.16225	-0.07447	-0.01779	-0.00663	0.00170	0.00834
0.00000	-0.32667	-0.24250	-0.19200	-0.09100	-0.02367	-0.01020	-0.00010	0.00798
0.10000	-0.37333	-0.28263	-0.22657	-0.11075	-0.03080	-0.01455	-0.00230	0.00754
0.50000	-0.60444	-0.50812	-0.43700	-0.25104	-0.08589	-0.04860	-0.01970	0.00400
0.90000	-0.90667	-0.87363	-0.84353	-0.71862	-0.43214	-0.29857	-0.16210	-0.02730
$\sigma_u = 0.50$ und $\rho_0 = 0$								
-0.90000	0.05333	0.16363	0.15775	0.21649	0.23938	0.24375	0.24692	0.24939
-0.50000	0.00000	0.07812	0.11500	0.18723	0.23006	0.23816	0.24412	0.24883
-0.10000	-0.12444	-0.02737	0.02975	0.14153	0.21421	0.22857	0.23930	0.24786
0.00000	-0.16667	-0.06250	0.00000	0.12500	0.20833	0.22500	0.23750	0.24750
0.10000	-0.21333	-0.10263	-0.03457	0.10525	0.20120	0.22065	0.23530	0.24706
0.50000	-0.44444	-0.32813	-0.24500	-0.03504	0.14611	0.18660	0.21790	0.24352
0.90000	-0.74667	-0.69363	-0.65153	-0.50262	-0.20014	-0.06337	0.07550	0.21222
$\sigma_u = 1.00$ und $\rho_0 = 0$								
-0.90000	0.55333	0.72613	0.75775	0.89149	0.96438	0.97875	0.98942	0.99789
-0.50000	0.50000	0.64062	0.71500	0.86223	0.95506	0.97316	0.98662	0.99733
-0.10000	0.37556	0.53512	0.62975	0.81653	0.93921	0.96357	0.98180	0.99636
0.00000	0.33333	0.50000	0.60000	0.80000	0.93333	0.96000	0.98000	0.99600
0.10000	0.28667	0.45987	0.56543	0.78025	0.92620	0.95565	0.97780	0.99556
0.50000	0.05556	0.23437	0.35500	0.63996	0.87111	0.92160	0.96040	0.99202
0.90000	-0.24667	-0.13113	-0.05153	0.17238	0.52486	0.67163	0.81800	0.96072
<u>Hinweis:</u> $\sigma_\varepsilon^2 = 1 - \rho^2$ und damit $\gamma_0 = 1$ für jedes ρ .								

6. Für den Fall $\rho = 0$ erhalten wir aus (11-29)

$$E[S_a^2] = \left(1 - \frac{1}{T}\right) (\sigma_\omega^2 + \sigma_u^2 (\sigma_\omega^2 + \rho_0^2)).$$

7. Falls keine Überlagerung vorliegt, d.h. $\sigma_u^2 = 0$ gilt, ergibt sich wegen $\chi_1 = 0, \chi_2 + \chi_3 = \psi$ für die Schätzfunktion S_a^2 für beliebiges ρ

$$E[S_a^2] = E[S^2] .$$

Eine der Tabelle 11.1 entsprechende Tabellierung der Biasfunktion bezüglich S_a^2 ist problematisch, weil nun zusätzlich die Überlagerungsvarianz σ_u^2 und der Parameter ρ_0 des AR(1)-Prozesses berücksichtigt werden müssen. Die Tabelle 11.2 zeigt alternativ Werte für $\sigma_u = 0.10$, $\sigma_u = 0.50$ und $\sigma_u = 1.00$ sowie jeweils $\rho_0 = 0$. Wie bei den Originaldaten gilt $\sigma_\varepsilon^2 = 1 - \rho^2$ und damit $\gamma_0 = 1$. Siehe Tabelle 11.1. Man beachte, daß für kleine Überlagerungsvarianz die Verzerrung von S_a^2 der von S^2 sehr ähnlich ist, was natürlich nicht überraschend ist. Je größer die Überlagerungsvarianz gewählt wird, desto öfter ergibt sich ein **positiver** Bias, im Gegensatz zur Verzerrung von S^2 , die immer negativ ist.

11.4.3 Multiplikative Überlagerung mit Faktorstruktur

Wir wollen nun eine den vorhergehenden Unterabschnitten entsprechende Analyse des Bias bei kleinem T für die multiplikative Überlagerung à la Hönne vornehmen. Es wird sich allerdings kein ähnlich einfaches Resultat ergeben. Zunächst versammeln wir hier die Ergebnisse für erste und zweite Momente, die in Abschnitt 11.3.4 abgeleitet wurden, wobei hier ausschließlich der Fall des AR(1)-Prozesses unterstellt wird: Aus (11-22) und (11-23) erhalten wir

$$\begin{aligned} E[X_t^a] &= \frac{\varrho_0}{1 - \rho} \\ \text{var}[X_t^a] &= \gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right), \quad \gamma_0 = \frac{\sigma_\omega^2}{1 - \rho^2} \\ \text{cov}[X_t^a, X_s^a] &= \gamma_0 \rho^{|t-s|} + \delta^2 \left\{ (1 + \rho + \dots + \rho^{t-s-1}) \frac{\rho_0^2}{1 - \rho} + \rho^{(t-s)} \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right) \right\} \end{aligned} \quad (11-32)$$

Im folgenden verwenden wir für die (Auto-)Kovarianz in (11-32) die abkürzende Schreibweise

$$\gamma_{t-s}^H \equiv \text{cov}[X_t^a, X_s^a]$$

Genau wie in Abschnitt 11.4.2 untersuchen wir den Erwartungswert der einzelnen Summanden der Varianz S_a^2 : Für den ersten Summanden erhalten wir

$$E\left[\frac{1}{T} \sum_{t=1}^T (X_t U_t - \mu_x)^2\right] = \frac{1}{T} T \text{var}[X_t^a] = \gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right)$$

Für den zweiten Summanden ergibt sich zunächst

$$\begin{aligned} &E\left[\frac{1}{T^2} \left(\sum_{s=1}^T (X_s U_s - \mu_x) \right)^2\right] \\ &= \frac{1}{T^2} \left(\sum_{s=1}^T \text{var}[X_s^a] + \sum_{s=1}^T \sum_{s'=1, s' \neq s}^T \text{cov}[X_s^a, X_{s'}^a] \right) \end{aligned}$$

Damit erhalten wir

$$\begin{aligned} &E\left[\frac{1}{T^2} \left(\sum_{s=1}^T (X_s U_s - \mu_x) \right)^2\right] \\ &= \frac{1}{T} \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right) \right) \\ &\quad + \frac{1}{T^2} (2(T-1)\gamma_1^H + 2(T-2)\gamma_2^H + \dots + 2 \cdot 2\gamma_{T-2}^H + 2 \cdot 1\gamma_{T-1}^H) \end{aligned}$$

Da, wie in Abschnitt 11.3.4 gezeigt, der Erwartungswert des dritten Summanden bis auf den Faktor (-2) mit dem zweiten Summanden identisch ist, erhalten wir für den Fall der Hönneüberlagerung

$$\begin{aligned} E[S_a^2] &= \gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right) - \frac{1}{T} \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right) \right) \\ &\quad - \frac{1}{T^2} \{2(T-1)\gamma_1^H + 2(T-2)\gamma_2^H + \dots + 2 \cdot 2\gamma_{T-2}^H + 2 \cdot 1\gamma_{T-1}^H\} \end{aligned} \quad (11-33)$$

oder auch

$$E[S_a^2] = \gamma_0 + \chi_1^H(\gamma_0, \rho_0, \rho, \delta, \sigma_\varepsilon^2) - \chi_2^H(\gamma_0, \rho_0, \rho, \delta, \sigma_\varepsilon^2, T) - \chi_3^H(\gamma_1^H, \gamma_2^H, \dots, \gamma_{T-1}^H, T) \quad (11-34)$$

mit

$$\begin{aligned} \chi_1^H(\gamma_0, \rho_0, \rho, \delta, \sigma_\varepsilon^2) &= (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \\ \chi_2^H(\gamma_0, \rho_0, \rho, \delta, \sigma_\varepsilon^2, T) &= \frac{1}{T} \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) \\ \chi_3^H(\gamma_1^H, \gamma_2^H, \dots, \gamma_{T-1}^H, T) &= \frac{1}{T^2} \{ 2(T-1)\gamma_1^H + 2(T-2)\gamma_2^H + \dots + 2 \cdot 2\gamma_{T-2}^H + 2 \cdot 1\gamma_{T-1}^H \} \end{aligned}$$

Man beachte, daß wegen der komplexen Struktur der Autokovarianzen γ_τ^H (siehe (11-23)) jetzt vor allem der Biasterm χ_3^H eine komplizierte Struktur aufweist, die wegen der vielen eingehenden Parameter ($\gamma_0, \rho_0, \rho, \delta, T$) schwierig zu analysieren ist. Siehe jedoch die Simulationsergebnisse sowie die abgebildeten Autokorrelationsfunktionen für diese Anonymisierungsvariante in Abschnitt 11.5.4.

11.5 Relevanz der Ergebnisse für das Regressionsmodell

Die bisher in diesem Abschnitt gewonnenen Ergebnisse zeigen, daß die multiplikative Überlagerung einen starken "finite sample bias" aufweist, während sich im Fall der additiven Überlagerung diese Effekte nicht ergeben.⁷¹ Für die Anonymisierung spielt jedoch die multiplikative Überlagerung die deutlich wichtigere Rolle. Andererseits sollte die Relevanz multiplikativer Fehlerstruktur auch bei empirischen Daten nicht unterschätzt werden!

11.5.1 Parameterschätzung

Die in den vorhergehenden Unterabschnitten gewonnenen Ergebnisse erhalten zusätzliche Relevanz, wenn man sich verdeutlicht, daß die Schätzung des Steigungskoeffizienten β im linearen Regressionsmodell

$$y_t = \beta_0 + \beta x_t + \eta_t \quad (11-35)$$

sowohl bei additiver als auch bei multiplikativer Überlagerung asymptotisch maßgeblich durch den Faktor

$$\frac{\text{var}[X_t]}{\text{var}[X_t^a]}$$

beeinflusst wird und daß wegen $\text{var}[X_t] < \text{var}[X_t^a]$ der wahre Parameterwert systematisch gegen Null verzerrt ist.⁷²

Bei der Schätzung des Parameters kommen die beiden oben betrachteten Schätzfunktionen ins Spiel, denn sowohl für Querschnittsdaten (siehe Abschnitt 9.3.3) als auch für Panelda-

⁷¹Cheang und Reinsel (2000) betrachten ein Modell mit autokorrelierten Störterm statt mit einem autokorrelierten Regressor und leiten den Bias des KQ-Schätzers für kleines T ab.

⁷²Wir vernachlässigen hier den Fall, daß die Überlagerungsvariablen für abhängige und unabhängige Variablen miteinander korreliert sind. Dann gelten andere Ergebnisse. Siehe z.B. Abschnitt 9.3.3.

ten (siehe Abschnitt 10.4.8) läßt sich die "naive" Schätzfunktion als

$$\hat{\beta}^a \cong \frac{S^2}{S_a^2} \beta \quad (11-36)$$

darstellen.⁷³

Daraus könnte man für den Erwartungswert

$$E[\hat{\beta}^a] \cong \frac{E[S^2]}{E[S_a^2]} \beta$$

folgern. Zwar gilt für den Erwartungswert des Verhältnisses der beiden Varianzschätzer

$$E\left[\frac{S^2}{S_a^2}\right] \neq \frac{E[S^2]}{E[S_a^2]}, \quad (11-37)$$

doch dürften in der Tendenz die oben gewonnenen Ergebnisse für die Erwartungswerte der beiden Schätzfunktionen S^2 und S_a^2 den Effekt beschreiben, der sich bei sehr kleinem T ergibt. Dies wird in den Simulationen des folgenden Unterabschnitts bestätigt.

Für Schätzungen auf der Basis von Paneldaten, d.h. für $N > 1$, wurde bereits im einleitenden Abschnitt 11.1 darauf hingewiesen, daß für $N \rightarrow \infty$ das Verhältnis der beiden Varianzen exakt den Bias des "naiven" Panelschätzers bestimmt, sofern bestimmte Annahmen gemacht werden. Siehe dazu auch Abschnitt 11.7.

Wir präsentieren zunächst Ergebnisse für die **additive Überlagerung** und die **allgemeine multiplikative Überlagerung**, für die in den Abschnitten 11.4.1 und 11.4.2 ausführlich der Bias bei kleinem T hergeleitet wurde. In einem weiteren Unterabschnitt werden dann Simulationsergebnisse für die **multiplikative Überlagerung à la Höhne** vorgestellt.

11.5.2 Simulationsrechnungen für additive Überlagerung

Die in Tabelle 11.5.2 dargestellten Simulationsrechnungen zeigen, daß die durch die Fehlerüberlagerung bedingte **Unterschätzung** von β bei sehr kleinem T deutlich verstärkt wird, wenn die Autokorrelation positiv ist ($\rho > 0$), was durch das kleine T verursachte negative Verzerrung von S^2 widerspiegelt. Im Fall negativer Autokorrelation ergeben sich dagegen auch für sehr kleines T kaum Abweichungen von den asymptotischen Ergebnissen! Die Tabelle gibt für $\rho = 0.90$ und $\rho = -0.90$ jeweils Ergebnisse für unterschiedlich großes T an. Dabei wurde ein Simulations-Design gemäß Tabelle 11.3 vorgegeben. Der AR(1)-Prozeß wurde mit zufälligen Startwert aus der $N(\rho_0/(1-\rho_1), \sigma_\omega^2/(1-\rho_1^2))$ -Verteilung erzeugt. Dabei wurden die ersten 200 Werte des Prozesses nicht verwendet, um die Effekte des Startwertes auszuschalten.

⁷³Für den naiven Schätzer läßt sich schreiben:

$$\hat{\beta}^a = \frac{S^2}{S_a^2} \beta + \frac{S_{x\eta}}{S_a^2}$$

wobei $S_{x\eta}$ die empirische Kovarianz zwischen dem Gleichungsfehler η und dem Regressor x ist. Wegen der üblicherweise angenommenen Unabhängigkeit tendiert der zweite Summand auf der rechten Seite gegen Null.

Tabelle 11.3: Simulations-Design bei additiver und multiplikativer Überlagerung

ADDITIVE ÜBERLAGERUNG	
Anzahl Wiederholungen	200
Anzahl Beobachtungen (T)	10 , 30 , 50, 100, 500
AR(1) - Parameter ρ_0	4.35
AR(1) - Parameter ρ_1	+0.9 , -0.9
AR(1) - Parameter σ_ω^2	1.00
Überlagerungsvarianz σ_u^2	0.25
Regressionsparameter α	0.00
Regressionsparameter β	-2.50
Regressionsparameter σ_η^2	0.25
MULTIPLIKATIVE ÜBERLAGERUNG	
Anzahl Wiederholungen	200
Anzahl Beobachtungen (T)	10 , 30 , 50, 100, 500
AR(1) - Parameter ρ_0	4.35
AR(1) - Parameter ρ_1	+0.9 , -0.9
AR(1) - Parameter σ_ω^2	1.00
Überlagerungsvarianz σ_u^2	0.01
Überlagerungsparameter (Höhe) δ	0.11
Überlagerungsvarianz (Höhe) σ_ε^2	0.03 , 0.20
Regressionsparameter α	0.00
Regressionsparameter β	-2.50
Regressionsparameter σ_η^2	0.25

Die Tabelle 11.4 zeigt deutlich, daß die asymptotischen Ergebnisse gerade bei sehr kleinem T wenig zuverlässig sind, wenn positive Autokorrelation gegeben ist, während für negative Autokorrelation kaum Effekte des Beobachtungsumfangs zu beachten sind.⁷⁴ Die Approximation der empirischen Ergebnisse des "naiven" Schätzers $\hat{\beta}^a$ durch

$$E[S^2]/E[S_a^2] \cdot \beta$$

sind erstaunlich gut. Dabei besteht sowohl für positive als auch für negative Autokorrelation eine (leichte) Tendenz zur **Überschätzung**. Wir kommen auf dieses Ergebnis bei der Konstruktion von Korrektorschätzern in Abschnitt 11.6 zurück. Dort wird der Korrektorschätzer

$$\hat{\beta}^{a,korr} = \frac{S_a^2}{S^2} \hat{\beta}^a \quad (11-38)$$

betrachtet, der den Bias für kleines T berücksichtigt und deshalb dem üblichen Korrektorschätzer, der sich aus der Asymptotik ableiten läßt, vorzuziehen ist.

⁷⁴In der Tabelle 11.4 wird unter "plim $\hat{\beta}^a$ " folgender Wert ausgewiesen:

$$\text{plim } \hat{\beta} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2} \beta \quad .$$

Tabelle 11.4: Simulationsergebnisse bei additiver Überlagerung

POSITIVE AUTOKORRELATION ($\rho = +0.90$)					
	Beobachtungsumfang T				
	10	30	50	100	500
$var[X_t] = \gamma_0$	5.2632				
$var[X_t]$ exakt (siehe (11-6))	1.4336	2.9378	3.6402	4.3579	5.0669
s^2 empirisch	1.5120	3.0937	3.6687	4.4564	5.0509
$var[X_t^a] = \gamma_0 + \sigma_u^2$	5.5132				
$var[X_t^a]$ exakt (siehe (11-28))	1.6586	3.1795	3.8852	4.6054	5.3164
s_a^2 empirisch	1.7679	3.3389	3.9105	4.7053	5.2986
$var[X_t]/var[X_t^a]$ asymptotisch	0.9547				
$var[X_t]/var[X_t^a]$ exakt	0.8643	0.9240	0.9369	0.9463	0.9531
$plim \hat{\beta}^a$	-2.3866				
$\hat{\beta}$ empirisch	-2.0638	-2.2432	-2.3149	-2.3511	-2.3787
$(E[S^2]/E[S_a^2]) \cdot \beta$	-2.1608	-2.3100	-2.3423	-2.3658	-2.3828
NEGATIVE AUTOKORRELATION ($\rho = -0.90$)					
	Beobachtungsumfang T				
	10	30	50	100	500
$var[X_t] = \gamma_0$	5.2632				
$var[X_t]$ exakt (siehe (11-6))	5.2184	5.2511	5.2566	5.2601	5.2626
s^2 empirisch	5.5495	5.0489	5.0161	5.1936	5.3164
$var[X_t^a] = \gamma_0 + \sigma_u^2$	5.5132				
$var[X_t^a]$ exakt (siehe (11-28))	5.2184	5.2511	5.2566	5.2601	5.2626
s_a^2 empirisch	5.7523	5.3431	5.2778	5.4435	5.5558
$var[X_t]/var[X_t^a]$ asymptotisch	0.9804				
$var[X_t]/var[X_t^a]$ exakt	0.9822	0.9810	0.9808	0.9806	0.9805
$plim \hat{\beta}^a$	-2.3866				
$\hat{\beta}$ empirisch	-2.3222	-2.3315	-2.3492	-2.3715	-2.3865
$(E[S^2]/E[S_a^2]) \cdot \beta$	-2.4555	-2.4525	-2.4520	-2.4515	-2.4513

11.5.3 Simulationsrechnungen für allgemeine multiplikative Überlagerung

Die Simulationsrechnungen für multiplikative Überlagerung in Tabelle 11.5 zeigen, daß die **Unterschätzung** von β bei sehr kleinem T deutlich verstärkt wird, wenn die Autokorrelation positiv ist ($\rho > 0$), was durch das kleine T verursachte negative Verzerrung von S^2 widerspiegelt. Im Fall negativer Autokorrelation ergeben sich dagegen auch für sehr kleines T kaum Abweichungen von den asymptotischen Ergebnissen! Die Tabelle 11.5 gibt für $\rho = 0.90$ und $\rho = -0.90$ jeweils Ergebnisse für unterschiedlich großes T an. Dabei wurde ein Simulations-Design gemäß Tabelle 11.3 vorgegeben. Der AR(1)-Prozeß wurde mit zufälligen Startwert aus der $N(\rho_0/(1-\rho_1), \sigma_\omega^2/(1-\rho_1^2))$ -Verteilung erzeugt. Dabei wurden die ersten 200 Werte des Prozesses nicht verwendet, um die Effekte des Startwertes auszuschalten.

Die Tabelle 11.5 zeigt deutlich, daß die asymptotischen Ergebnisse gerade bei sehr kleinem T vollkommen irreführend sind, wenn positive Autokorrelation gegeben ist, während für negative Autokorrelation kaum Effekte des Beobachtungsumfangs zu beachten sind.⁷⁵ Die Approximation der empirischen Ergebnisse des "naiven" Schätzers $\hat{\beta}^a$ durch

$$E[S^2]/E[S_a^2] \cdot \beta$$

sind erstaunlich gut. Dabei besteht sowohl für positive als auch für negative Autokorrelation eine (leichte) Tendenz zur **Überschätzung**. Wir kommen auf dieses Ergebnis bei der Konstruktion von Korrekturschätzern in Abschnitt 11.6 zurück. Dort wird der Korrekturschätzer

$$\hat{\beta}^{a,korr} = \frac{S_a^2}{S^2} \hat{\beta}^a \quad (11-39)$$

betrachtet, der den Bias für kleines T berücksichtigt und deshalb dem üblichen Korrekturschätzer, der sich aus der Asymptotik ableiten läßt, vorzuziehen ist.

11.5.4 Simulationsergebnisse für Überlagerung à la Höhne

In Tabelle 11.5 sind für hohe positive Autokorrelation der Originalvariablen ($\rho = 0.90$) auch Ergebnisse für die multiplikative Überlagerung à la Höhne dargestellt. Dabei wurde für den Zuschlagsparameter $\delta = 0.11$ und für die Restvarianz alternativ $\sigma_\varepsilon = 0.03$ und $\sigma_\varepsilon = 0.20$ gesetzt.

Die im mittleren Teil der genannten Tabelle dargestellten Schätzergebnisse für den "naiven" Kleinstquadrateschätzer offenbaren, daß der Bias (für festes δ) von der Größenordnung der Restvarianz σ_ε^2 abhängt: Während für kleine Varianz die Schätzung deutlich weniger verzerrt ist als unter der allgemeinen multiplikativen Überlagerung (der wahre Parameterwert ist -2.50), ergibt sich für große Varianz eine noch deutlich schlechtere Schätzung selbst für sehr großes T . Der Wahrscheinlichkeitsgrenzwert für $T \rightarrow \infty$ des naiven Schätzers weist offenbar einen deutlichen Bias auf, der vor allem von der Varianz σ_ε^2 abhängt!⁷⁶ Dies Ergebnis

⁷⁵In den Tabellen wird unter "plim $\hat{\beta}^a$ " folgender Wert ausgewiesen:

$$\text{plim } \hat{\beta} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2(\sigma_x^2 + \mu_x^2)} \beta = \frac{\gamma_0}{\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right)} \beta$$

⁷⁶Für $T = 5000$ und $\sigma_\varepsilon^2 = 0.03^2$ ergibt sich bei 1000 Wiederholungen ein Schätzwert von -1.8849.

Tabelle 11.5: Simulationsergebnisse bei multiplikativer Überlagerung

POSITIVE AUTOKORRELATION ($\rho = +0.90$)					
	Beobachtungsumfang T				
	10	30	50	100	500
$var[X_t] = \gamma_0$	5.2632				
$var[X_t]$ exakt (siehe (11-6))	1.4336	2.9378	3.6402	4.3579	5.0669
s^2 empirisch	1.7041	3.1404	3.7368	4.2660	4.9888
$var[X_t^a] = \gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right)$	24.2383				
$var[X_t^a]$ exakt (siehe (11-30))	18.5112	21.2805	22.2358	23.1433	24.0041
s_a^2 empirisch	21.3175	22.5583	22.2223	23.3901	24.1272
$var[X_t]/var[X_t^a]$ asymptotisch	0.2171				
$var[X_t]/var[X_t^a]$ exakt	0.0774	0.1381	0.1637	0.1883	0.2111
$plim \hat{\beta}^a$	-0.5429				
$\hat{\beta}$ empirisch	-0.1887	-0.3346	-0.4000	-0.4434	-0.5192
$(E[S^2]/E[S_a^2]) \cdot \beta$	-0.1935	-0.3453	-0.4093	-0.4708	-0.5278
Multiplikative Höhe-Überlagerung ($\delta = 0.11$)					
$(\sigma_\varepsilon = 0.03)$					
s_a^2 empirisch	3.5887	4.8476	5.3324	6.0742	6.8562
$\hat{\beta}$ empirisch	-1.1144	-1.5138	-1.6090	-1.7592	-1.8504
$(\sigma_\varepsilon = 0.20)$					
s_a^2 empirisch	82.4701	80.2343	79.4670	79.5217	80.5935
$\hat{\beta}$ empirisch ($\sigma_\varepsilon^2 = 0.20$)	-0.0663	-0.0929	-0.1192	-0.1320	-0.1586
NEGATIVE AUTOKORRELATION ($\rho = -0.90$)					
	Beobachtungsumfang T				
	10	30	50	100	500
$var[X_t] = \gamma_0$	5.2632				
$var[X_t]$ exakt (siehe (11-6))	5.2184	5.2511	5.2566	5.2601	5.2626
s^2 empirisch	5.8541	5.1251	5.4220	5.5576	5.2712
$var[X_t^a] = \gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right)$	5.3682				
$var[X_t^a]$ exakt (siehe (11-30))	5.3129	5.3527	5.3595	5.3641	5.3674
s_a^2 empirisch	6.1085	5.2119	5.5255	5.6720	5.3784
$var[X_t]/var[X_t^a]$ asymptotisch	0.9804				
$var[X_t]/var[X_t^a]$ exakt	0.9822	0.9810	0.9808	0.9806	0.9805
$plim \hat{\beta}^a$	-2.4511				
$\hat{\beta}$ empirisch	-2.4184	-2.4471	-2.4506	-2.4474	-2.4501
$(E[S^2]/E[S_a^2]) \cdot \beta$	-2.4555	-2.4525	-2.4520	-2.4515	-2.4513

Tabelle 11.6: Theoretische Autokorrelation bei multiplikativer Überlagerung ($\rho = +0.90$)

Überlagerung	γ_1	Bemerkungen
Originaldaten	0.90	$\gamma_1 = \rho$
allgemein multiplikativ	0.1954	siehe (11-19)
multiplikativ Höhe	0.9251	siehe (11-25)($\sigma_\varepsilon = 0.03$)
multiplikativ Höhe	0.2659	siehe (11-25)($\sigma_\varepsilon = 0.20$)

ist allerdings nicht konsistent mit den Ergebnissen in Abschnitt 9.2.2 für Querschnittsdaten, ebenso wenig wie die empirische Varianz auch für großes T **nicht** den Ergebnissen in Abschnitt 7.2 für die Varianz einer nach Höhe multiplikativ überlagerten Variablen entspricht.

Denkbar ist, daß die unterschiedliche Autokorrelation die Ursache ist. Dazu bestimmen wir die Autokorrelation erster Ordnung aus den beiden Formeln (11-19) und (11-25). Unter Verwendung der Vorgaben aus Tabelle 11.3 erhalten wir die Ergebnisse, die Tabelle 11.6 präsentiert. Daraus ergibt sich, daß für kleine Höhe-Varianz die Autokorrelation der anonymisierten Zeitreihe noch **erhöht** wird, während sich für größere Höhe-Varianz eine mittelgroße Autokorrelation erster Ordnung ergibt ($\gamma_1 = 0.29$). Dieser Wert ist jedoch noch deutlich größer als derjenige für die allgemeine Überlagerung ($\gamma_1 = 0.19$) und kann somit nur unvollständig die unterschiedlich starke Verzerrung erklären, die ja tendentiell mit wachsender Autokorrelation zunimmt, sofern multiplikative Überlagerung vorliegt.

GAUSS Fri Nov 16 13:29:52 2007

AUTOKORRELATIONSFUNKTION ORIGINALDATEN (rho=0.90)

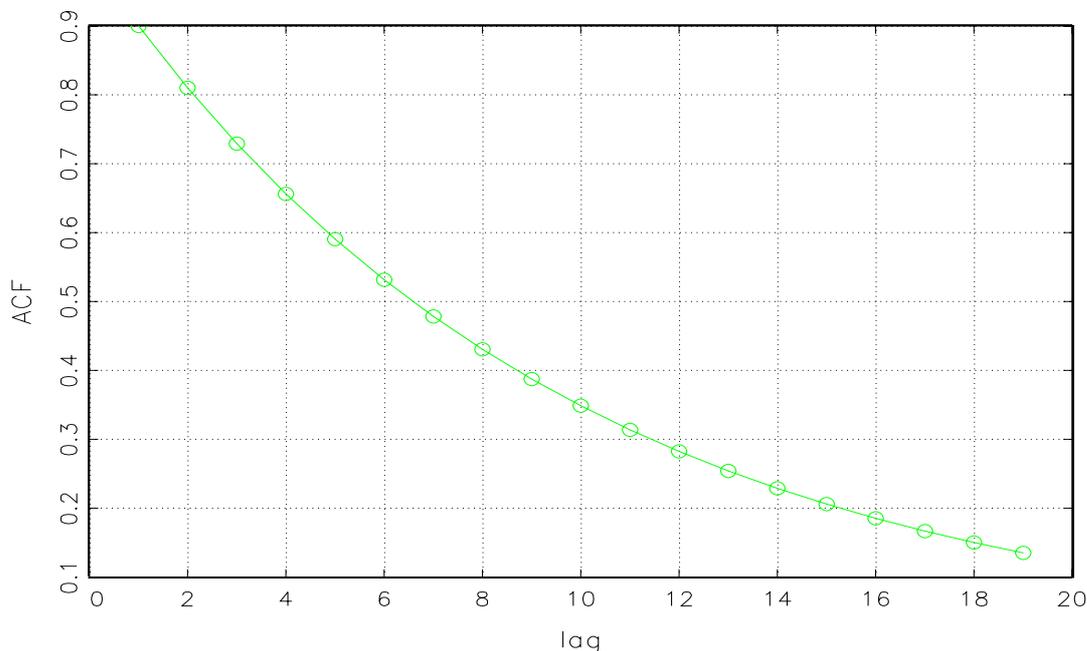


Abbildung 11/1: Autokorrelationsfunktion für die Originaldaten

Ein Blick auf die Formeln (11-18) und ((11-24)) für die Autokorrelation bei multiplikativer Überlagerung macht klar, daß die Autokorrelationsfunktion bei multiplikativer Überlage-

nung eine deutlich andere Gestalt als im (unterstellten bzw. unterliegenden) AR(1)-Prozeß hat, der ja exponentiell abklingt. Dies ist in Abbildung 11/1 für den hier speziell betrachteten Fall $\rho = 0.90$ dargestellt.

Abbildung 11/2 zeigt den Fall der allgemeinen Überlagerung, ebenfalls für $\rho = 0.90$. Daraus wird ersichtlich, daß die Autokorrelationsfunktion ab Lag 1 relativ "flach" verläuft. Die ungewöhnliche, "geknickte" Form der der Funktion bedarf noch eines Kommentars: Durch Simulationen wurde verifiziert, daß die dargestellte Funktion durch eine geschätzte Autokorrelationsfunktion mit großem T sehr gut approximiert werden kann!

GAUSS Mon Nov 26 08:44:37 2007

AUTOKORRELATIONSFUNKTION MULTIPLIKATIVE UEBERLAGERUNG (rho=0.90)

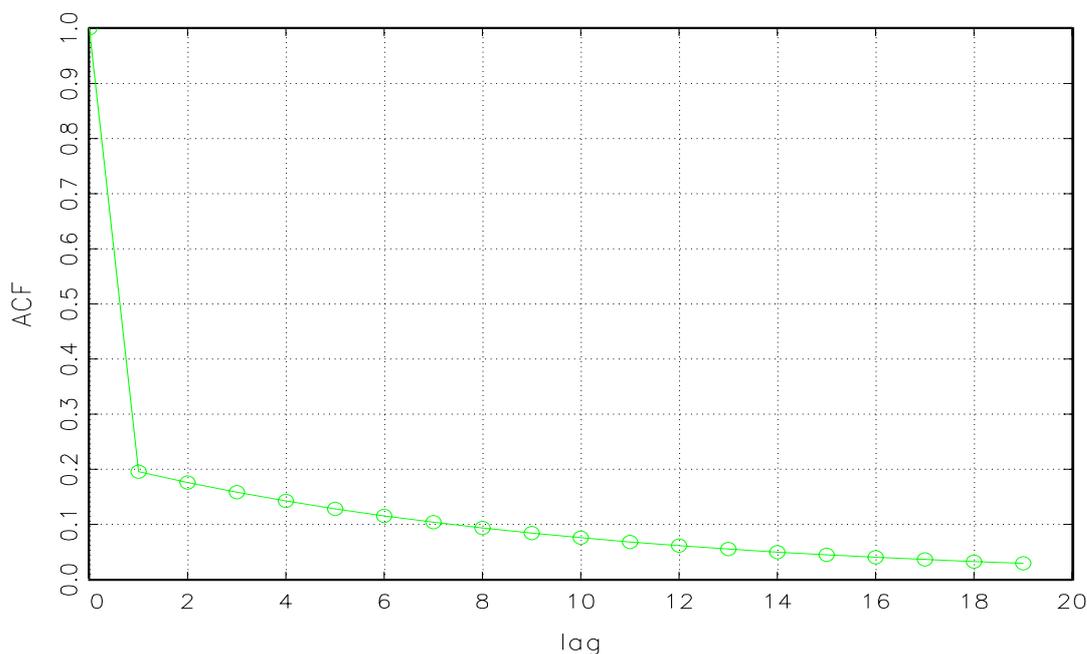


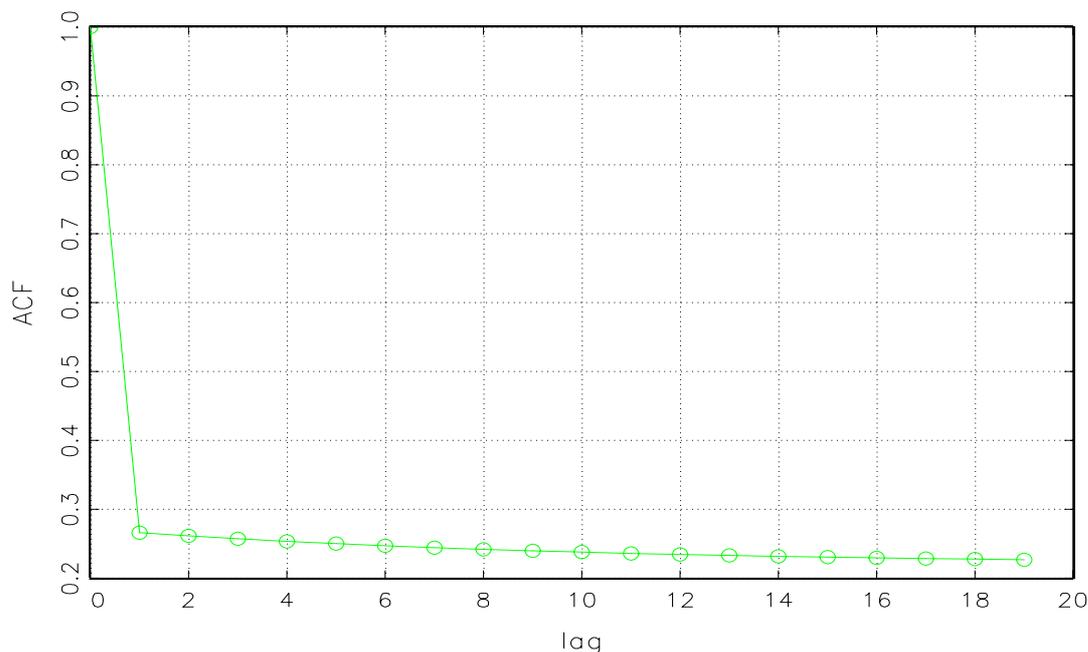
Abbildung 11/2: Autokorrelationsfunktion Allgemeine Überlagerung

Die beiden restlichen Abbildungen betrachten die multiplikative Höhe-Überlagerung, wiederum für den Fall $\rho = 0.90$. In Abbildung 11/3 wird der Fall der größeren Varianz und in Abbildung 11/4 der Fall einer besonders kleinen Varianz⁷⁷ dargestellt. Die unterschiedliche vertikale Skala sollte beachtet werden. Im Fall der kleinen Varianz zeigt die Autokorrelationsfunktion "persistentes" Verhalten, d.h. die Korrelation zwischen entfernter liegenden Punkten ist fast genau so groß wie die zwischen benachbarten Punkten. Dies dürfte die Ursache für die besonders starke Verzerrung des naiven Schätzers (siehe Tabelle 11.5) in diesem Fall sein. Allerdings ist dies bisher nur eine Vermutung, die näher analysiert werden muß.

⁷⁷Dies ist allerdings die im FAWE-Panel-Projekt gewählte Parameter-Konstellation!

GAUSS Mon Nov 26 08:38:04 2007

AUTOKORRELATIONSFUNKTION MULT.HOEHNE UEBERLAGERUNG (rho=0.90)

Abbildung 11/3: Autokorrelationsfunktion Höhe-Überlagerung ($\sigma_\varepsilon = 0.20$)

11.6 Korrektorschätzer für endliches T

Der in (11-38) bzw. (11-39) vorgeschlagene Schätzer ist nicht operational, weil S^2 Verfügbarkeit der Originaldaten unterstellt. Es können jedoch nur die anonymisierten Variablen verwendet werden. Die in dieser Arbeit sonst dazu ausgenutzten Beziehungen zwischen asymptotischen Ergebnissen (siehe die Abschnitte über Korrektorschätzer) sollten nicht eingesetzt werden, weil sie den endlichen Beobachtungsumfang unberücksichtigt lassen. Andererseits erfordert die Berücksichtigung des Bias für kleines T sowohl im additiven als auch im multiplikativen Fall die Kenntnis der Parameter, die in die Bias-Funktion eingehen. Im Fall des AR(1)-Prozesses ist dies der Parameter ρ . Außerdem ist die Varianz σ_ω^2 dieses Prozesses zu bestimmen.

11.6.1 Additive Überlagerung

Im additiven Fall ließe sich ein Schätzer für $\sigma_x^2 = \text{var}[X_t]$ wie folgt konstruieren: Weil

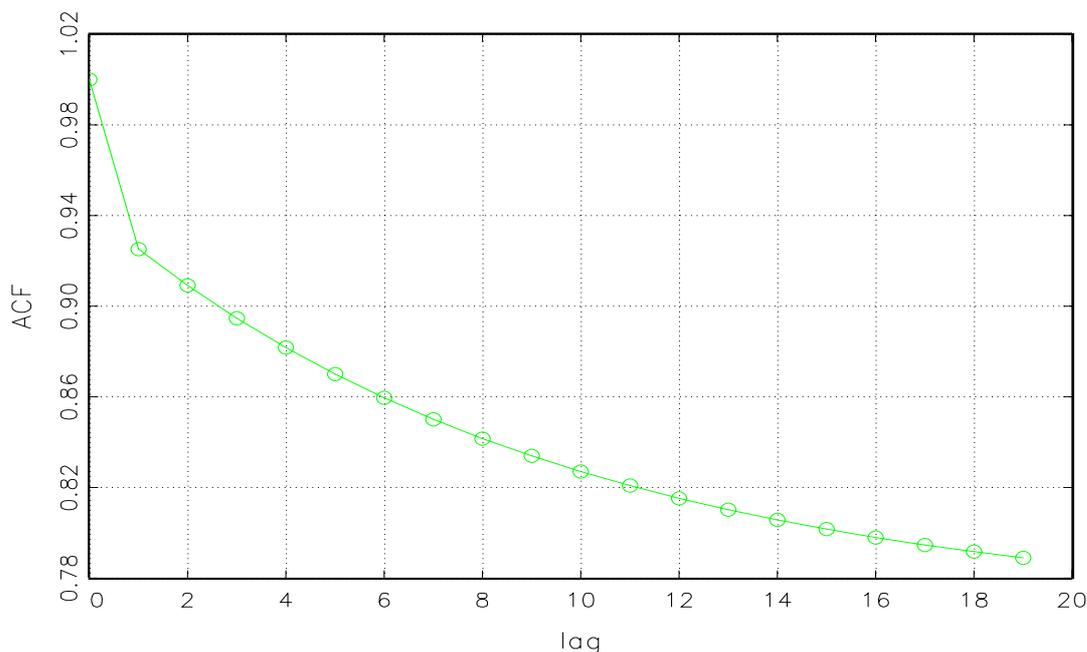
$$\text{var}[X_t^a] = \text{var}[X_t] + \sigma_u^2$$

gilt und die Fehlervarianz (annahmegemäß) bekannt ist, verwendet man

$$\hat{\sigma}_x^2 = S_a^2 - \sigma_u^2 \quad . \quad (11-40)$$

GAUSS Mon Nov 26 08:39:57 2007

AUTOKORRELATIONSFUNKTION MULT.HOEHNE UEBERLAGERUNG (rho=0.90)

Abbildung 11/4: Autokorrelationsfunktion Höhe-Überlagerung ($\sigma_\varepsilon = 0.03$)

Eine operationale Form des Korrektorschätzers erhält man also, wenn in (11-38) der Schätzer $\hat{\sigma}_x^2$ statt S^2 verwendet wird. Das ergibt dann

$$\widehat{\beta^{a,korr}} = \frac{S_a^2}{\hat{\sigma}_x^2} \hat{\beta}^a = \frac{S_a^2}{S_a^2 - \sigma_u^2} \hat{\beta}^a \quad (11-41)$$

Zur Bestimmung von σ_w^2 und ρ sei auf die folgenden Ausführungen für die multiplikative Überlagerung verwiesen. Daraus wird deutlich, daß die Konstruktion eines Korrektorschätzers im (weniger interessanten) Fall der additiven Überlagerung deutlich einfacher wäre.

11.6.2 Multiplikative Überlagerung

In diesem Fall läßt sich ein Schätzer für $\sigma_x^2 = \text{var}[X_t]$ wie folgt konstruieren: Weil

$$\text{var}[X_t^a] = \text{var}[X_t] + \sigma_u^2 (\text{var}[X_t] + (E[X_t])^2)$$

gilt, die Fehlervarianz (annahmegemäß) bekannt ist und $E[X_t]$ auch für endlichen Beobachtungsumfang T zuverlässig durch $\hat{\mu}_x = \overline{X^a}$ geschätzt werden kann, verwendet man

$$\hat{\sigma}_x^2 = \frac{S_a^2 - \sigma_u^2 \hat{\mu}_x^2}{1 + \sigma_u^2} \quad (11-42)$$

Eine operationale Form des Korrektorschätzers erhält man also, wenn in (11-39) der Schätzer $\hat{\sigma}_x^2$ statt S^2 verwendet wird. Das ergibt dann

$$\widehat{\beta^{a,korr}} = \frac{S_a^2}{\hat{\sigma}_x^2} \hat{\beta}^a = \frac{(1 + \sigma_u^2) S_a^2}{S_a^2 - \sigma_u^2 \overline{X^a}^2} \hat{\beta}^a \quad (11-43)$$

Wegen

$$S_a^2 = \overline{X_a^2} - \overline{X^a}^2$$

wäre

$$\widehat{\beta^{a,korr}} = \frac{(1 + \sigma_u^2) S_a^2}{\overline{X_a^2} - (1 + \sigma_u^2) \overline{X^a}^2} \hat{\beta}^a \quad (11-44)$$

eine Alternative.⁷⁸

Einige - hier nicht präsentierte - Simulationsrechnungen zeigen, daß beide Varianten keine zufriedenstellenden Ergebnisse produzieren. Dies liegt einerseits daran, daß nicht gewährleistet ist, daß der Nenner immer positiv geschätzt wird. Vor allem aber macht sich bemerkbar, daß die oben präsentierten Endlichkeitskorrekturen nicht berücksichtigt sind. Um diese Korrekturen verwenden zu können, müßten jedoch Schätzwerte für ρ_0 , ρ und σ_ω^2 zur Verfügung stehen.

Aus (11-16) und (11-18) ist bekannt, daß für die Autokovarianz und Autokorrelation der multiplikativ überlagerten Variablen

$$\text{cov}[X_t^a, X_s^a] = \text{cov}[X_t, X_s] = \gamma_0 \rho^{|t-s|}$$

sowie

$$\text{corr}[X_t^a, X_s^a] = \frac{\text{cov}[X_t, X_s]}{\text{var}[X_t^a]} = \frac{\gamma_0}{\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\mu}{1-\rho} \right)^2 \right)} \rho^{|t-s|}$$

gilt.

Aus der ersten Beziehung ergibt sich

$$\begin{aligned} \text{cov}[X_t^a, X_{t-1}^a] &= \gamma_0 \rho \\ \text{cov}[X_t^a, X_{t-2}^a] &= \gamma_0 \rho^2 \\ &\vdots \\ \text{cov}[X_t^a, X_{t-k}^a] &= \gamma_0 \rho^k \end{aligned}$$

sowie

$$\frac{\text{cov}[X_t^a, X_{t-2}^a]}{\text{cov}[X_t^a, X_{t-1}^a]} = \rho \quad ,$$

was unter der Annahme, daß die Regressorvariable x einem AR(1)-Prozeß folgt, zur Schätzung von γ_0 und ρ verwendet werden kann. Dabei kann nach Berechnung von ρ aus der ersten Gleichung auch γ_0 bestimmt werden.⁷⁹ Allerdings dürften auch hier wieder die Mißachtung der Endlichkeitskorrektur zu Problemen führen.

⁷⁸

$$\overline{X_a} = \frac{1}{T} \sum_t X_t^a \quad \text{und} \quad \overline{X_a^2} = \frac{1}{T} \sum_t (X_t^a)^2$$

⁷⁹Technisch gesehen sollte zur Schätzung von ρ die Gleichung

$$\tilde{\rho} = \frac{\hat{\gamma}_2}{\hat{\gamma}_1}$$

geschätzt werden, wobei $\hat{\gamma}_k$ die empirische Autokovarianz k -ter Ordnung ist. Sodann ergibt sich ein Wert für γ_0 aus

$$\tilde{\gamma}_0 = \frac{\hat{\gamma}_1}{\tilde{\rho}} .$$

11.7 Ein Ausblick auf die Situation bei Paneldaten

Falls mehrere Untersuchungseinheiten über die Zeit beobachtet werden und damit Panelmikrodaten verfügbar sind, treten weitere Probleme hinzu. Insbesondere ist nicht klar, wie die möglicherweise sehr heterogene Autokorrelation bei den verschiedenen Untersuchungseinheiten⁸⁰ geschätzt und - wichtiger - welches ρ insgesamt dann unterstellt werden soll, wenn ein Korrektorschätzer verwendet werden soll.

Andererseits zeigen die Ergebnisse der vorangehenden Abschnitte, daß der bisher wenig untersuchte Fall multiplikativer Meßfehler bei kleiner Anzahl von Panel-Wellen (T) zu stark verzerrten Schätzungen führt, was bisher in der Panel-Literatur nicht zur Kenntnis genommen wurde. Wie bereits im einleitenden Kapitel 11.1 erläutert wurde, kann mit den für den Fall $N = 1$ abgeleiteten Ergebnissen für die Schätzfunktionen S^2 und S_a^2 die asymptotische Verzerrung des "Panelschätzers" $\hat{\beta}_W$ durch

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^a = \frac{E[S^2]}{E[S_a^2]} \beta$$

bestimmt werden, sofern die folgenden Annahmen erfüllt sind:

- Die Anzahl T der Wellen ist eine feste Zahl.
- Alle N Regressorvariablen folgen demselben AR(1)-Prozeß (und sind damit identisch verteilt!).⁸¹
- Nur die Regressorvariable ist fehlerüberlagert oder die Fehlerüberlagerung der abhängigen Variablen ist unkorreliert mit Überlagerungen der Regressoren.⁸²

11.7.1 Additive Überlagerung

Bei additiver Überlagerung ergibt sich

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^a = \frac{\gamma_0 - \psi(\gamma, T)}{\gamma_0 - \psi(\gamma, T) + (1 - \frac{1}{T})\sigma_u^2} \beta \quad (11-45)$$

Diese Formel zeigt, daß die Verzerrung durch ψ dann harmlos ist, wenn keine Meßfehler vorliegen ($\sigma_u^2 = 0$). Solange die Verzerrungsfunktion die Bedingung $\gamma_0 - \psi(\gamma, T) > 0$ erfüllt⁸³, ergibt sich auch für kleines T eine **Unterschätzung** von β .

⁸⁰Der bei Biorn (1996) untersuchte Fall der Autokorrelation der Meßfehler muß sich streng genommen derselben Frage stellen, doch ist es vermutlich eher akzeptabel, einen identischen AR(1)-Prozeß für die Meßfehler aller Unternehmen als für die sie beschreibenden Merkmale zu unterstellen.

⁸¹Falls man dies als utopische Annahme ablehnt, bleibt immer noch die Tatsache, daß für das Modell der Zeitreihenregression von der multiplikativer Überlagerung starke Effekte auf die Schätzung ausgehen, sofern T klein ist.

⁸²Genauer: Bisher liegen für den Fall korrelierter Meßfehler noch keine Untersuchungsergebnisse vor.

⁸³Siehe dazu die Diskussion in Abschnitt 11.2.2.

11.7.2 Allgemeine multiplikative Überlagerung

Bei multiplikativer Überlagerung ergibt sich

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^a = \frac{\gamma_0 - \psi(\boldsymbol{\gamma}, T)}{\gamma_0 + \chi_1(\gamma_0, \rho_0, \rho, \sigma_u^2) - \chi_2(\gamma_0, \rho_0, \rho, \sigma_u^2, T) - \chi_3(\gamma_1, \gamma_2, \dots, \gamma_{T-1}, T)} \beta \quad (11-46)$$

Im Gegensatz zur additiven Überlagerung sind für diese Formel keine eindeutigen allgemeinen Aussagen über Richtung und Ausmaß der Verzerrung zu machen. Simulationsrechnungen zeigen jedoch, daß die Verzerrung vor allem für positive Autokorrelation und für kleines T , $T < 10$, erheblich ist! Allerdings ist auch in diesem - multiplikativen - Fall die durch kleines T bedingte Verzerrung der Erwartungswerte von S^2 und S_a^2 dann kein Problem für die Schätzung, wenn keine stochastische Überlagerung stattfindet ($\sigma_u^2 = 0$). Denn auch hier gilt:

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^a = \frac{E[S^2]}{E[S_a^2]} \beta = \frac{E[S^2]}{E[S^2]} \beta = \beta \quad \text{für beliebiges } \rho = 0 \text{ .}$$

11.7.3 Multiplikative Höhne-Überlagerung

Bei multiplikativer Überlagerung ergibt sich

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^a = \frac{\gamma_0 - \psi(\boldsymbol{\gamma}, T)}{\gamma_0 + \chi_1^H(\gamma_0, \rho_0, \rho, \sigma_u^2) - \chi_2^H(\gamma_0, \rho_0, \rho, \sigma_u^2, T) - \chi_3^H(\gamma_1^H, \gamma_2^H, \dots, \gamma_{T-1}^H, T)} \beta \quad (11-47)$$

mit

$$\begin{aligned} \chi_1^H(\gamma_0, \rho_0, \rho, \delta, \sigma_\varepsilon^2) &= (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \\ \chi_2^H(\gamma_0, \rho_0, \rho, \delta, \sigma_\varepsilon^2, T) &= \frac{1}{T} \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) \\ \chi_3^H(\gamma_1^H, \gamma_2^H, \dots, \gamma_{T-1}^H, T) &= \frac{1}{T^2} \{ 2(T-1)\gamma_1^H + 2(T-2)\gamma_2^H + \dots + 2 \cdot 2\gamma_{T-2}^H + 2 \cdot 1\gamma_{T-1}^H \} \end{aligned}$$

Siehe Abschnitt 11.4.3.

11.7.4 Korrektorschätzer

Aus den vorhergehenden Ausführungen ergibt sich, daß die in Abschnitt 11.6 gemachten Ausführungen hier wegen der Asymptotik $N \rightarrow \infty$ und festem T ebenfalls relevant sind, sofern die oben gemachten Annahmen über die identische Verteilung für alle N Unternehmen und T Zeitpunkte erfüllt sind

Insbesondere zeigt die obige Formel (11-45) für additive Überlagerung, daß ein einfacher Korrektorschätzer für den Panelfall durch

$$\widehat{\beta^{a, \text{korr}}} = \frac{S_a^2}{S_a^2 - (1 - \frac{1}{T})\sigma_u^2} \hat{\beta}_W^a \quad (11-48)$$

gegeben ist. Wohlgermerkt: Eine Kenntnis der Parameter des autoregressiven Prozesses für die Regressor-Variable(n) ist hier nicht erforderlich!

12 Instrumentvariablen-Schätzer (IV-Schätzer)

12.1 Allgemeine Vorbemerkungen

Eine Alternative zum Korrekturschätzer (und zum SIMEX-Schätzer) bietet der IV-Schätzer, der in diesem Unterabschnitt betrachtet werden soll. Dabei wird weiterhin das **einfache** lineare Regressionsmodell

$$y_t = \beta_0 + \beta x_t + \eta_t$$

(siehe (11-35)) mit einem Regressor x , der einem AR(1)-Prozeß folgt (siehe (11-1)), betrachtet.

Wir beschränken uns außerdem auf den einfachsten Fall der GMM/IV-Schätzung, in dem nur **eine einzige Instrumentvariable** verwendet wird, die fortan mit dem Symbol z belegt wird. Diese Instrumentvariable soll folgende Eigenschaften haben⁸⁴:

Eigenschaft (E1) Die Instrumentvariable z darf nicht stochastisch unabhängig von der (meßfehlerfreien) Variablen x sein.

Eigenschaft (E2) Die Instrumentvariable z muß unkorreliert mit der Meßfehlervariablen (Überlagerung) u sein.

Eigenschaft (E3) Die Instrumentvariable z muß unkorreliert sein mit dem Gleichungsfehler η .

Wir verwenden als IV-Schätzer für β die Form

$$\hat{\beta}^{IV} = \frac{\widehat{cov}[z, y]}{\widehat{cov}[z, x^a]} = \frac{\sum_t (z_t - \bar{z})(y_t - \bar{y})}{\sum_t ((z_t - \bar{z})(x_t^a - \bar{x}^a))}, \quad (12-1)$$

die dem "Within"-Schätzer nachempfunden ist.⁸⁵

Dieser Schätzer ist unter den Voraussetzungen (E1) bis (E3) konsistent: Es gilt (bei **additiver Überlagerung**)

$$\text{plim } \hat{\beta}^{IV} = \frac{cov[z, y]}{cov[z, x^a]} = \frac{cov[z, x]\beta + cov[z, \eta]}{cov[z, x] + cov[z, u]}$$

⁸⁴Siehe beispielsweise Carroll et al (2006 Kapitel 6, S. 129).

⁸⁵Im Fall der multiplen Regression mit insgesamt K Regressoren sowie K Instrumentvariablen lautet der Schätzer

$$\hat{\beta}^{IV} = (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{y}$$

wobei \mathbf{Z} und \mathbf{X} zwei $(T \times K)$ -Matrizen sind und \mathbf{y} ein T -dimensionaler Vektor ist. Einen breiten Überblick über IV-Schätzer für Panelmodelle mit (additiven) Meßfehlern geben Biørn und Krishnakumar (2008). Den hier verwendeten Schätzer untersuchen Sie allerdings nicht. Andererseits analysieren sie Schätzer, bei denen Niveaus und Differenzen von Regressorvariablen miteinander kombiniert werden. Die Autoren weisen darauf hin, daß dabei allerdings wegen der geringen Korrelation zwischen Regressor und Instrument des Problem der "weak instruments" zu beachten ist. Siehe Biørn und Krishnakumar (2008) S. 336.

Wegen (E2) gilt $cov[z, u] = 0$ und wegen (E3) gilt $cov[z, \eta] = 0$. Wenn jedoch (E1) nicht gilt und damit insbesondere $cov[z, x] = 0$ gilt, ergibt sich kein sinnvoller Wahrscheinlichkeitsgrenzwert. Darüberhinaus kann dieser Schätzer durchaus schlechtere Ergebnisse als der naive KQ-Schätzer liefern, falls nicht alle drei Voraussetzungen erfüllt sind.⁸⁶

Bei **multiplikativer** Überlagerung ergibt sich ein anderer Ausdruck für $cov[z, x^a]$ und damit für den Wahrscheinlichkeitsgrenzwert von $\hat{\beta}^{IV}$. Wir geben hier zunächst allgemein die Kovarianz für die Produkte $X^a = UX$ und dem dem Instrument WZ an, wobei unterstellt wird, daß U und W stochastisch unabhängig voneinander sowie von X und Z sind und beide Zufallsvariable (U und V) Erwartungswert 1 haben. Dann gilt:

$$cov[UX, WZ] = cov[X, Z] + cov[U, W] (cov[X, Z] + E[X]E[Z]) \quad .$$

Da im Fall des IV-Schätzers $X^a = XU$ sowie W identisch 1 ist und damit $cov[U, W] = 0$ gilt, erhalten wir bei multiplikativer Überlagerung

$$plim \hat{\beta}^{IV} = \frac{cov[z, y]}{cov[z, x^a]} = \frac{cov[z, x]\beta + cov[z, \eta]}{cov[z, x]} \quad .$$

Dieses Ergebnis suggeriert, daß die mögliche Korrelation zwischen Z und U keine Rolle spielt! Falls jedoch z ebenfalls eine fehlerbehaftete Variable ist und der betreffende Überlagerungsfehler mit dem Meß/Überlagerungsfehler in x^a korreliert ist, ergibt sich ein anderes Bild. Siehe dazu den folgenden Unterabschnitt.

In jüngerer Zeit hat vor allem der Begriff "schwacher" Instrumente eine Rolle in der wissenschaftlichen Diskussion gespielt.⁸⁷ Für den hier betrachteten Fall, in dem die Inkonsistenz der KQ-Schätzers durch die Fehler-Überlagerung verursacht wird, bedeutet "schwach", daß das Instrument z eine niedrige Korrelation mit der Regressorvariablen x aufweist.

12.2 Die verzögerte Regressorvariable als Instrument

In der Zeitreihenregression wird, falls die Regressorvariable nicht exogen ist, üblicherweise die verzögerte Regressorvariable als Instrumentvariable verwendet, d.h. es gilt⁸⁸

$$z_t = x_{t-1} \quad .$$

Diese Wahl ist bei fehlerbehafteten Regressoren nicht möglich; es kann nur

$$z_t = x_{t-1}^a \quad (12-2)$$

als Instrument verwendet werden. Der entsprechende IV-Schätzer lautet

$$\begin{aligned} \hat{\beta}^{a,IV} &= \frac{\sum_{t=2}^T (x_{t-1}^a - \bar{x}^a)(y_t - \bar{y})}{\sum_{t=2}^T (x_{t-1}^a - \bar{x}^a)(x_t^a - \bar{x}^a)} && \text{"symmetrischer IV-Schätzer"} \\ &= \frac{\sum_{t=2}^T (x_{t-1}^a - \bar{x}^a)(x_t - \bar{x})}{\sum_{t=2}^T (x_{t-1}^a - \bar{x}^a)(x_t^a - \bar{x}^a)} \beta + \frac{\sum_{t=2}^T (x_{t-1}^a - \bar{x}^a)(\eta_t - \bar{\eta})}{\sum_{t=2}^T (x_{t-1}^a - \bar{x}^a)(x_t^a - \bar{x}^a)} \quad (12-3) \end{aligned}$$

⁸⁶Siehe beispielsweise Carroll et al (2006 Kapitel 6).

⁸⁷Siehe beispielsweise Bound, Jaeger und Baker (1995), Staiger und Stock (1997) sowie Stock, Wright und Yogo (2002).

⁸⁸Siehe beispielsweise ****

d.h. dieser Schätzer entspricht dem "naiven" Schätzer $\hat{\beta}^a$, nur tritt an die Stelle von $(x_t^a - \bar{x}^a)$ nun $(x_{t-1}^a - \bar{x}^a)$. Dabei werden die arithmetischen Mittel jeweils mit allen T verfügbaren Beobachtungen gebildet:

$$\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t, \quad \bar{x}^a = \frac{1}{T} \sum_{t=1}^T x_t^a, \quad (12-4)$$

Wegen der Verwendung identischer Mittelwerte bezeichnen wir (12-3) als "symmetrische" Form des IV-Schätzers.

Allerdings ist diese Formulierung nicht gültig, wenn die übliche Konstruktion des IV-Schätzers verwendet wird. Dann ergibt sich ein Schätzer mit Mittelwerten, die jeweils aus $T - 1$ Beobachtungswerten gebildet werden. Der resultierende Schätzer lautet dann

$$\hat{\beta}^{a,IV} = \frac{\sum_{t=2}^T (x_{t-1}^a - \bar{x}_1^a)(y_t - \bar{y}_2)}{\sum_{t=2}^T (x_{t-1}^a - \bar{x}_1^a)(x_t^a - \bar{x}_2^a)} \quad \text{"asymmetrischer IV-Schätzer"} \quad (12-5)$$

mit

$$\bar{y}_2 = \frac{1}{T-1} \sum_{t=2}^T y_t, \quad \bar{x}_1^a = \frac{1}{T-1} \sum_{t=2}^T x_{t-1}^a \quad \text{und} \quad \bar{x}_2^a = \frac{1}{T-1} \sum_{t=2}^T x_t^a \quad . \quad (12-6)$$

Wegen der Verwendung unterschiedlicher Mittelwerte bezeichnen wir (12-5) als "asymmetrische" Form des IV-Schätzers.

Dies soll kurz erläutert werden: Wir betrachten zunächst den allgemeinen Fall der multiplen Regression mit insgesamt K Regressoren sowie K Instrumentvariablen. Dann lautet der naive IV-Schätzer

$$\hat{\beta}^{IV} = (\mathbf{Z}'\mathbf{X}^a)^{-1} \mathbf{Z}'\mathbf{y}$$

wobei \mathbf{Z} und \mathbf{X}^a zwei $((T-1) \times K)$ -Matrizen sind und \mathbf{y} ein $(T-1)$ -dimensionaler Vektor ist.⁸⁹ Wählt man für $K = 2$ die Matrizen \mathbf{Z} und \mathbf{X}^a sowie den Vektor \mathbf{y} als

$$\mathbf{X}^a = \begin{pmatrix} 1 & x_2^a \\ 1 & x_3^a \\ \vdots & \vdots \\ 1 & x_{T-1}^a \\ 1 & x_T^a \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} 1 & x_1^a \\ 1 & x_2^a \\ \vdots & \vdots \\ 1 & x_{T-2}^a \\ 1 & x_{T-1}^a \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_{T-1} \\ y_T \end{pmatrix},$$

so lautet der Schätzer

$$\hat{\beta}^{a,IV} = \frac{\sum_{t=2}^T \left(x_{t-1}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{s-1}^a \right) \right) \left(y_t - \left(\frac{1}{T-1} \sum_{s=2}^T y_s \right) \right)}{\sum_{t=2}^T \left(x_{t-1}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{s-1}^a \right) \right) \left(x_t^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_s^a \right) \right)}$$

und das entspricht der oben verwendeten Definition des Schätzers. Im folgenden werden wir stets die "asymmetrische" Form betrachten, sofern nichts anderes vermerkt wird.⁹⁰

⁸⁹Siehe beispielsweise *****.

⁹⁰Vor allem bei der Ableitung der Erwartungswerte der Stichprobenfunktionen C und C_a in Abschnitt 12.3 spielt dies eine Rolle. Die Ableitung für die symmetrische Form wird in einem Appendix dargestellt. Siehe Abschnitt C.

In Abschnitt 10.6.3 wurde der Panelschätzer ("Within"-Schätzer) in Matrixschreibweise vorgestellt (siehe (10-61)):

$$\hat{\beta}_W = \left(\sum_{i=1}^n \mathbf{X}'\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) \mathbf{X}\langle i \rangle \right)^{-1} \sum_{i=1}^n \mathbf{X}'\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) \mathbf{y}\langle i \rangle .$$

Der dazu korrespondierende "naive" Panelschätzer bei Verwendung fehlerüberlagerter (bzw. anonymisierter) Daten lautet

$$\hat{\beta}_W^a = \left(\sum_{i=1}^n \mathbf{X}^{a'}\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) \mathbf{X}^a\langle i \rangle \right)^{-1} \sum_{i=1}^n \mathbf{X}^{a'}\langle i \rangle \left(\mathbf{I}_T - \frac{1}{T} \boldsymbol{\nu}_T \boldsymbol{\nu}'_T \right) \mathbf{y}^a\langle i \rangle ,$$

wobei $\mathbf{y}\langle i \rangle$ an die Stelle von $\mathbf{y}^a\langle i \rangle$ tritt, sofern nur die Regressoren überlagert werden.

Anhand dieser Formel läßt sich nun auch der obige IV-Schätzer auf den Fall von Paneldaten erweitern. Dafür betrachten wir zunächst für jedes i die $((T-1) \times K)$ -Matrizen $\mathbf{X}^a\langle i \rangle$, $\mathbf{Z}\langle i \rangle$ sowie den $(T-1)$ -dimensionalen Vektor $\mathbf{y}\langle i \rangle$ und schreiben dann den "naiven" Panel-IV-Schätzer als

$$\hat{\beta}_P^a = \left(\sum_{i=1}^N \mathbf{Z}'\langle i \rangle \left(\mathbf{I}_{T-1} - \frac{1}{T-1} \boldsymbol{\nu}_{T-1} \boldsymbol{\nu}'_{T-1} \right) \mathbf{X}^a\langle i \rangle \right)^{-1} \sum_{i=1}^N \mathbf{Z}'\langle i \rangle \left(\mathbf{I}_{T-1} - \frac{1}{T-1} \boldsymbol{\nu}_{T-1} \boldsymbol{\nu}'_{T-1} \right) \mathbf{y}\langle i \rangle . \quad (12-7)$$

Für den Fall $K = 1$ (nur ein Regressor und kein (!) Absolutglied) läßt sich dafür explizit schreiben

$$\mathbf{X}^a\langle i \rangle = \begin{pmatrix} x_{i2}^a \\ x_{i3}^a \\ \vdots \\ x_{i,T-1}^a \\ x_{iT}^a \end{pmatrix}, \quad \mathbf{Z}\langle i \rangle = \begin{pmatrix} x_{i1}^a \\ x_{i2}^a \\ \vdots \\ x_{i,T-2}^a \\ x_{i,T-1}^a \end{pmatrix}, \quad \mathbf{y}\langle i \rangle = \begin{pmatrix} y_{i2} \\ y_{i3} \\ \vdots \\ y_{i,T-1} \\ y_{iT} \end{pmatrix},$$

und als Schätzer erhalten wir

$$\hat{\beta}_P^{a,IV} = \frac{\sum_{i=1}^N \sum_{t=2}^T \left(x_{i,t-1}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{i,s-1}^a \right) \right) \left(y_{it} - \left(\frac{1}{T-1} \sum_{s=2}^T y_{is} \right) \right)}{\sum_{i=1}^N \sum_{t=2}^T \left(x_{i,t-1}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{i,s-1}^a \right) \right) \left(x_{it}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{i,s}^a \right) \right)}, \quad (12-8)$$

wobei sich auch hier die "asymmetrische" Formulierung ergibt. Dabei wird hier unterstellt, daß nur der Regressor x überlagert wird, während y fehlerfrei beobachtet wird. Andernfalls tritt y_{it}^a an die Stelle von y_{it} .

12.2.1 Additive Überlagerung:

Im Fall der additiven Überlagerung gilt $z_t = x_{t-1}^a = x_{t-1} + u_{t-1}$. Für die Kovarianz zwischen Instrument und Regressor gilt dann

$$\text{cov}[X_t^a, X_{t-1}^a] \equiv \text{cov}[X_t + U_t, X_{t-1} + U_{t-1}] = \text{cov}[X_t, X_{t-1}] + \text{cov}[U_t, U_{t-1}] .$$

Außerdem ergibt sich für U_t identisch 1

$$\text{cov}[X_t, X_{t-1}^a] \equiv \text{cov}[X_t, X_{t-1}U_{t-1}] = \text{cov}[X_t, X_{t-1}] + 0 ,$$

so daß im additiven Fall für den Wahrscheinlichkeitsgrenzwert gilt:

$$\begin{aligned}
\text{plim } \hat{\beta}^{a,IV} &= \frac{\text{cov}[z, y]}{\text{cov}[z, x^a]} \\
&= \frac{\text{cov}[z, x]\beta + \text{cov}[z, \eta]}{\text{cov}[z, x^a]} \\
&= \frac{\text{cov}[X_t, X_{t-1} + U_{t-1}]\beta + \text{cov}[X_{t-1} + U_{t-1}, \eta]}{\text{cov}[X_t + U_t, X_{t-1} + U_{t-1}]} \\
&= \frac{\text{cov}[X_t, X_{t-1}]\beta + \text{cov}[X_{t-1} + U_{t-1}, \eta]}{\text{cov}[X_t, X_{t-1}] + \text{cov}[U_t, U_{t-1}]} \\
&= \frac{\text{cov}[X_t, X_{t-1}]\beta + \text{cov}[X_{t-1}, \eta] + \text{cov}[U_{t-1}, \eta]}{\text{cov}[X_t, X_{t-1}] + \text{cov}[U_t, U_{t-1}]} .
\end{aligned}$$

Konsistenz ergibt sich also nur für den Fall, daß

$$\text{cov}[U_t, U_{t-1}] = 0$$

sowie

$$\text{cov}[U_{t-1}, \eta] = 0 \quad \text{und} \quad \text{cov}[X_{t-1}, \eta] = 0$$

gilt, wobei die beiden letzten Eigenschaften 'harmlos' sind und insbesondere die letzte Bedingung gilt, sofern der Regressor streng exogen ist.

12.2.2 Multiplikative Überlagerung:

Im Fall der multiplikativen Überlagerung gilt in diesem Fall $z_t = x_{t-1}^a = x_{t-1}u_{t-1}$. Für die Kovarianz zwischen Instrument und Regressor gilt dann

$$\text{cov}[X_t^a, X_{t-1}^a] \equiv \text{cov}[U_t X_t, U_{t-1} X_{t-1}] = \text{cov}[X_t, X_{t-1}] + \text{cov}[U_t, U_{t-1}] (\text{cov}[X_t, X_{t-1}] + E[X_t]E[X_{t-1}]) .$$

Außerdem ergibt sich für U_t identisch 1

$$\text{cov}[X_t, X_{t-1}^a] \equiv \text{cov}[X_t, U_{t-1} X_{t-1}] = \text{cov}[X_t, X_{t-1}] + 0$$

so daß für den Wahrscheinlichkeitsgrenzwert gilt:

$$\begin{aligned}
\text{plim } \hat{\beta}^{a,IV} &= \frac{\text{cov}[z, y]}{\text{cov}[z, x^a]} \\
&= \frac{\text{cov}[z, x]\beta + \text{cov}[z, \eta]}{\text{cov}[z, x^a]} \\
&= \frac{\text{cov}[U_t X_t, U_{t-1} X_{t-1}]\beta + \text{cov}[z, \eta]}{\text{cov}[U_t X_t, U_{t-1} X_{t-1}]} \\
&= \frac{\text{cov}[X_t, X_{t-1}]\beta + \text{cov}[X_{t-1} + U_{t-1}, \eta]}{\text{cov}[X_t, X_{t-1}] + \text{cov}[U_t, U_{t-1}] (\text{cov}[X_t, X_{t-1}] + E[X_t]E[X_{t-1}])} \\
&= \frac{\text{cov}[X_t, X_{t-1}]\beta + \text{cov}[X_{t-1}, \eta] + \text{cov}[U_{t-1}, \eta]}{\text{cov}[X_t, X_{t-1}] + \text{cov}[U_t, U_{t-1}] (\text{cov}[X_t, X_{t-1}] + E[X_t]E[X_{t-1}])} .
\end{aligned}$$

Konsistenz ergibt sich also – genau wie im additiven Fall – nur dann, wenn

$$\text{cov}[U_t, U_{t-1}] = 0$$

sowie

$$\text{cov}[U_{t-1}, \eta] = 0 \quad \text{und} \quad \text{cov}[X_{t-1}, \eta] = 0$$

gilt, wobei die beiden letzten Eigenschaften 'harmlos' sind und insbesondere die letzte Bedingung gilt, sofern der Regressor streng exogen ist.

12.2.3 Multiplikative Überlagerung mit Faktorstruktur:

Da im Höhne-Fall ebenfalls

$$\text{cov}[X_t, U_{t-1}X_{t-1}] = \text{cov}[X_t, X_{t-1}]$$

gilt (siehe (11-26)), gelten die Ausführungen für die allgemeine Überlagerung ebenso in diesem Fall, mit dem Unterschied, daß für die Höhne-Überlagerung die Bedingung

$$\text{cov}[U_t, U_{t-1}] = 0$$

nicht erfüllt ist und damit die Konsistenz des IV-Schätzers in diesem Fall (zumindest für verzögerte Variable X_{t-1}^a als Instrument) ausgeschlossen ist .

12.2.4 Überprüfung der Eigenschaften (E1), (E2) und (E3)

(E1) Wir prüfen nun, ob die Eigenschaft (E1) für das durch (12-2) definierten Instrument gegeben ist, d.h. ob $z_t = x_{t-1}^a$ und x_t (also nicht x_t^a) miteinander korreliert sind.

1. Für die Korrelation zwischen x_t und $z_t = x_{t-1}^a$ erhalten wir im **additiven Fall**

$$\text{corr}[X_t, X_{t-1}^a] = \frac{\gamma_0}{\sqrt{\gamma_0(\gamma_0 + \sigma_u^2)}} \rho \quad ,$$

2. Im multiplikativen Fall ergibt sich

$$\text{corr}[X_t, X_{t-1}^a] = \frac{\gamma_0}{\sqrt{\gamma_0 \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\mu}{1-\rho} \right)^2 \right) \right)}} \rho \quad .$$

3. Schließlich erhalten wir für die Höhne-Überlagerung (siehe (11-25))

$$\text{corr}[X_t, X_{t-1}^a] = \frac{\gamma_0}{\sqrt{\gamma_0 \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right)}} \rho$$

In allen drei Fällen wird die Korrelation ρ der Originalvariablen zwar gedämpft, doch ist vor allem bei starker positiver Korrelation die Eigenschaft (E1) sicher gegeben.

(E2) Als nächstes prüfen wir, ob die Eigenschaft (E2) gegeben ist, d.h. ob Instrument $z_t = x_{t-1}^a$ und Überlagerungsvariable u_t unkorreliert sind. Es ist unmittelbar einsehbar, daß dies gegeben ist, solange u_t und u_{t-1} keine Korrelation miteinander aufweisen, also im Fall der additiven wie auch im Fall der "allgemeinen" multiplikativen Überlagerung. Dagegen kann im Fall der Höhne-Überlagerung das Instrument nicht durch (12-2) gegeben sein!!

Alternativ bietet sich im Fall der **additiven Höhne-Überlagerung**, die in (10-4) beschrieben wird, das Instrument

$$z_t = x_{t-1}^a - x_{t-2}^a = x_{t-1} - x_{t-2} + \varepsilon_{t-1} - \varepsilon_{t-2} \quad (12-9)$$

an, da durch die Differenzenbildung der Faktor D eliminiert wird und gemäß Annahme $\varepsilon_{t-\tau}$, $\tau > 0$, mit u_t unkorreliert ist. Allerdings "leidet" darunter die Forderung, daß Instrumentvariable und Regressorvariable hohe Korrelation aufweisen sollen. Denn es ergibt sich⁹¹

$$\text{cov}[x_{t-1}^a - x_{t-2}^a, x_t^a] = \rho(1 - \rho)\sigma_x^2,$$

sowie

$$\text{cov}[x_{t-1}^a - x_{t-2}^a, x_t] = \rho(1 - \rho)\sigma_x^2,$$

d.h. die Kovarianz des Instrumentes (sowohl mit dem fehlerbehafteten als auch mit dem wahren Regressor) tendiert gegen Null, wenn ρ positiv ist (das ist der ökonomisch relevante Fall!) und gegen +1 strebt! Sie hat ein (lokales) Maximum bei $\rho = +0.50$ und steigt für negatives ρ monoton an.

Für den Fall der **multiplikativen Höhne-Überlagerung** (siehe (10-10)) scheint dagegen kein Instrument zu existieren, das die Forderung (E2) erfüllt. Denn auch eine multiplikative Verknüpfung von x_{t-1}^a und x_{t-2}^a beispielsweise in Form eines Bruchs eliminiert den problematischen Faktor D nur unvollkommen. Siehe dazu die Diskussion in Abschnitt 8.

Im Fall der Anonymisierung (aber nicht im "normalen" Ökonometriebetrieb) bietet sich schließlich noch an, eine "zweite" fehlerbehaftete Variable $z_t = x_t^b$ zu erzeugen, die hoch mit der durch Höhne-Überlagerung erzeugten Variablen x_t^a korreliert ist.⁹² Sei beispielsweise

$$x_t^b = x_t w_t$$

wobei w_t eine Zufallsvariable mit Erwartungswert 1 und Varianz σ_w^2 ist, die mit allen anderen Fehlervariablen unkorreliert ist. Dann ergibt sich für multiplikativ Höhne-überlagertes $x_t^a = (1 + \mu D + \varepsilon_t) x_t$ die Kovarianz

$$\begin{aligned} \text{cov}[x_t^a, x_t^b] &= \text{cov}[(1 + \mu D + \varepsilon_t) x_t, x_t w_t] \\ &= \text{cov}[x_t \varepsilon_t, x_t w_t] \\ &= \sigma_x^2 \end{aligned}$$

⁹¹Für die Kovarianz ergibt sich im Einzelnen

$$\begin{aligned} \text{cov}[x_{t-1}^a - x_{t-2}^a, x_t^a] &= \text{cov}[(x_{t-1} - x_{t-2}) + (\varepsilon_{t-1} - \varepsilon_{t-2}), x_t + \mu D + \varepsilon_t] \\ &= \text{cov}[x_t, x_{t-1}] - \text{cov}[x_t, x_{t-2}] \\ &= \rho(1 - \rho)\sigma_x^2 \end{aligned}$$

⁹²Diesen Hinweis verdanke ich Martin Rosemann. Allerdings haben auch Lechner und Pohlmeier (2003) diese Idee bereits im Zusammenhang mit der Verwendung anonymisierter Daten in der Schätzung ökonomischer Modelle erwähnt/realisiert.

wobei die Formel für die Kovarianz von Produkten von Zufallsvariablen aus Abschnitt 12.1 verwendet wurde. Offensichtlich ist diese Instrument ein "starkes", denn die Korrelation zwischen x_t^a und x_t^b wird nur durch die Fehlervarianzen der beiden Überlagerungen reduziert! Genauer: Für die Korrelation ergibt sich

$$\text{corr}[x_t^a, x_t^b] = \frac{\sigma_x^2}{\sqrt{(\sigma_x^2 + (\delta^2 + \sigma_\varepsilon^2)(\sigma_x^2 + \mu_x^2))(\sigma_x^2 + \sigma_w^2(\sigma_x^2 + \mu_x^2))}}$$

Etwas ausführlicher geht Abschnitt 12.6 auf diesen Ansatz ein.

(E3) Schließlich ist noch die Eigenschaft (E3) für das Instrument $z_t = x_{t-1}^a$ zu überprüfen. Da x_{t-1} im Verhältnis zum Gleichungsfehler η_t eine vorherbestimmte Variable ist und u_{t-1} gemäß Annahme (bzw. im Fall der Anonymisierung gemäß Konstruktion) mit dem Gleichungsfehler unkorreliert ist, ist die Eigenschaft (E3) in allen drei Fällen (additive, allgemeine multiplikative und multiplikative Höhle-Überlagerung) gegeben.

12.3 Schätzung der Autokovarianz bei additiver und multiplikativer Überlagerung

12.3.1 Allgemeine Bemerkungen

Wenn der im vorhergehenden Unterabschnitt betrachtete IV-Schätzer $\hat{\beta}^{a,IV}$ in der **asymmetrischen Form** (siehe (12-3))⁹³ verwendet wird, dann wird ein kleines T ebenfalls Auswirkungen haben. Allerdings müssen wir jetzt anstelle der Varianzen S^2 und S_a^2 die (Auto-)Kovarianzen

$$C = \frac{1}{T-1} \sum_{t=2}^T (X_t - \bar{X}_2)(X_{t-1} - \bar{X}_1) \quad (12-10)$$

sowie

$$C_a = \frac{1}{T-1} \sum_{t=2}^T (X_t^a - \bar{X}_2^a)(X_{t-1}^a - \bar{X}_1^a) \quad (12-11)$$

mit

$$\bar{X}_1 = \frac{1}{T-1} \sum_{t=2}^T X_{t-1}, \quad \bar{X}_2 = \frac{1}{T-1} \sum_{t=2}^T X_t,$$

(und entsprechend für X^a und η) verwenden, denn für den "naiven" IV-Schätzer läßt sich schreiben:

$$\hat{\beta}^{a,IV} = \frac{\frac{1}{T-1} \sum_{t=2}^T (X_t - \bar{X}_2)(X_{t-1}^a - \bar{X}_1^a)}{\frac{1}{T-1} \sum_{t=2}^T (X_t^a - \bar{X}_2^a)(X_{t-1}^a - \bar{X}_1^a)} \beta + \frac{\sum_t (X_{t-1}^a - \bar{X}_1^a)(\eta_t - \bar{\eta}_2)}{\sum_t (X_{t-1}^a - \bar{X}_1^a)(X_t^a - \bar{X}_2^a)} \quad (12-12)$$

⁹³Die entsprechenden Ergebnisse für die **symmetrische** Form des IV-Schätzers werden in Abschnitt ?? behandelt. Eine ausführliche Analyse verschiedener Varianten, die hier als "symmetrisch" bzw. "asymmetrisch" bezeichnet werden, findet sich beispielsweise bei Anderson (1971) Abschnitt 8.2.1.

Man beachte, daß der Wahrscheinlichkeitsgrenzwert von

$$\frac{\sum_t (X_{t-1}^a - \bar{X}_1^a)(\eta_t - \bar{\eta})_2}{\sum_t (X_{t-1}^a - \bar{X}_1^a)(X_t^a - \bar{X}_2^a)}$$

gleich Null ist. Ferner kann man ausnutzen, daß

$$\text{cov}[X_t, X_{t-1}^a] = \text{cov}[X_t, X_{t-1}]$$

sowohl im additiven als auch im multiplikativen Fall gilt⁹⁴ und somit C anstelle der empirischen Kovarianz

$$\frac{1}{T-1} \sum_{t=2}^T (X_t - \bar{X}_2)(X_{t-1}^a - \bar{X}_1^a)$$

als Schätzer sinnvoll ist. Somit ergibt sich, daß

$$\text{plim} \hat{\beta}^{a,IV} = \text{plim} \frac{C}{C_a} \beta = \frac{E[C]}{E[C_a]} \beta$$

Im folgendes wird zunächst die Schätzfunktion C und dann die Schätzfunktion C_a bei additiver, multiplikativer und Höhne-Überlagerung untersucht.

12.3.2 Die Schätzfunktion C bei kleinem T

Ähnlich dem Vorgehen in Abschnitt 11.2.2 schreiben wir die Schätzfunktion als

$$\begin{aligned} C &= \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu + \mu - \bar{X}_2)(X_{t-1} - \mu + \mu - \bar{X}_1) \\ &= \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(X_{t-1} - \mu) + \frac{1}{T-1} \sum_{t=2}^T ((\bar{X}_2 - \mu)(\bar{X}_1 - \mu) \\ &\quad - \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(\bar{X}_1 - \mu) - \frac{1}{T-1} \sum_{t=2}^T (X_{t-1} - \mu)(\bar{X}_2 - \mu)) \\ &= \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(X_{t-1} - \mu) + \frac{1}{(T-1)^2} (\sum_{s=2}^T \sum_{s'=2}^T (X_s - \mu)(X_{s'-1} - \mu) \\ &\quad - \frac{1}{(T-1)^2} \sum_{t=2}^T \sum_{s=2}^T (X_t - \mu)(X_{s-1} - \mu) - \frac{1}{(T-1)^2} \sum_{t=2}^T \sum_{s=2}^T (X_{t-1} - \mu)(X_s - \mu)) \\ &= \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(X_{t-1} - \mu) - \frac{1}{(T-1)^2} (\sum_{s=2}^T \sum_{s'=2}^T (X_s - \mu)(X_{s'-1} - \mu) \end{aligned}$$

Erwartungswertbildung ergibt für den ersten Summanden

$$E[\frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(X_{t-1} - \mu)] = \gamma_1$$

für den zweiten Summanden (siehe dazu auch Abschnitt 11.2.2) erhalten wir⁹⁵

$$\begin{aligned} &E[\frac{1}{(T-1)^2} (\sum_{s=2}^T \sum_{s'=2}^T (X_s - \mu)(X_{s'-1} - \mu))] \\ &= \left(\frac{1}{T-1}\right)^2 [(T-2)\gamma_0 + 2\{(T-2)\gamma_1 + (T-3)\gamma_2 + \dots + \cdot 3\gamma_{T-4} + \cdot 2\gamma_{T-3} + \gamma_{T-2}\} + \gamma_{T-1}] \end{aligned}$$

⁹⁴. Siehe dazu die folgenden Ausführungen in diesem Unterabschnitt.

⁹⁵ Die Ergebnisse lassen sich am leichtesten verifizieren, wenn man (hier beispielsweise für den dritten

Damit erhalten wir für den Erwartungswert der Schätzfunktion C

$$E[C] = \gamma_1 - \zeta(\gamma, T) \tag{12-13}$$

mit

$$\zeta(\gamma, T) = \frac{T-2}{(T-1)^2} \gamma_0 + \frac{1}{(T-1)^2} \left(2 \sum_{k=1}^{T-2} (T-k-1) \gamma_k + \gamma_{T-1} \right) \tag{12-14}$$

Eine Aussage über die Richtung der Verzerrung für kleines T läßt sich direkt für positive Autokorrelation treffen:

$$E[C] < \gamma_1 \quad \text{falls } \gamma_k > 0 \text{ für } k = 1, 2, \dots, T-1, \tag{12-15}$$

d.h. die Autokovarianz γ_1 wird durch C unterschätzt, falls alle Autokorrelationen positiv sind.

Speziell für den Fall, daß keine Autokorrelation vorliegt ($\gamma_\tau = 0$ für $\tau > 0$) ergibt sich demnach

$$E[C] = -\frac{T-2}{(T-1)^2} \gamma_0 < 0$$

und demnach wird der wahre Autokorrelationskoeffizient γ_1 ($= 0$) für endliches (und insbesondere für "kleines") T unterschätzt.⁹⁶

Für den Spezialfall des AR(1)-Prozesses mit Parameter ϱ ergibt sich für die Biasfunktion

$$\zeta(\gamma_0, \varrho, T) = \frac{\gamma_0}{(T-1)^2} \left[T-2 + 2 \sum_{k=1}^{T-2} (T-k-1) \rho^k + \varrho^{T-1} \right] .$$

Die Tabelle 12.7 zeigt ausgewählte Werte für diese Biasfunktion $\zeta(\gamma_0, \varrho, T)$ für $\gamma_0 = 1$. Man erkennt, daß nur für extrem großes negatives ϱ die Verzerrung positiv ist. Überdies ist die Verzerrungen für negatives ρ deutlich geringer.

Ferner wird aus der Tabelle deutlich, daß die durch C geschätzte Autokovarianz γ_1 vor allem für kleines T sehr große Verzerrungen aufweist: Beispielsweise lautet für $T = 5$ und

Summanden) die folgende Tabelle bildet:

		s								
		2	3	4	5	...	T-3	T-2	T-1	T
s'	1	γ_1	γ_2	γ_3	γ_4	...	γ_{T-4}	γ_{T-3}	γ_{T-2}	γ_{T-1}
	2	γ_0	γ_1	γ_2	γ_3	...	γ_{T-5}	γ_{T-4}	γ_{T-3}	γ_{T-2}
	3	γ_1	γ_0	γ_1	γ_2	...	γ_{T-6}	γ_{T-5}	γ_{T-4}	γ_{T-3}
	4	γ_2	γ_1	γ_0	γ_1	...	γ_{T-7}	γ_{T-6}	γ_{T-5}	γ_{T-4}
	⋮					⋮				
	T-4	γ_{T-4}	γ_{T-5}	γ_{T-6}	γ_{T-7}	...	γ_1	γ_2	γ_3	γ_4
	T-3	γ_{T-5}	γ_{T-6}	γ_{T-7}	γ_{T-8}	...	γ_0	γ_1	γ_2	γ_3
	T-2	γ_{T-4}	γ_{T-5}	γ_{T-6}	γ_{T-7}	...	γ_1	γ_0	γ_1	γ_2
	T-1	γ_{T-3}	γ_{T-4}	γ_{T-5}	γ_{T-6}	...	γ_2	γ_1	γ_0	γ_1

Die Tabelle zeigt, daß beispielsweise γ_0 insgesamt $(T-2)$ -mal auftaucht, daß γ_1 $(T-3)$ -mal unterhalb der Diagonalen und $(T-1)$ -mal oberhalb der Diagonalen, insgesamt als $2(T-2)$ -mal auftaucht, schließlich daß γ_{T-3} ein und drei mal, insgesamt also vier mal in der Tabelle vorhanden ist.

⁹⁶Dieses Ergebnis ist seit langem bekannt; siehe beispielsweise Anderson (1971) S. 448.

Tabelle 12.7: Werte der Biasfunktion $\zeta(\gamma_0, \varrho, T)$ für AR(1)-Prozeß.

ρ	Anzahl Wellen (T)							
	3	4	5	10	30	50	100	500
-0.90000	0.00250	-0.07878	0.00238	-0.00274	0.00119	0.00086	0.00048	0.00010
-0.50000	0.06250	0.04167	0.05078	0.03016	0.01083	0.00657	0.00331	0.00067
-0.10000	0.20250	0.17989	0.15238	0.08060	0.02722	0.01635	0.00818	0.00164
0.00000	0.25000	0.22222	0.18750	0.09877	0.03329	0.01999	0.01000	0.00200
0.10000	0.30250	0.26900	0.22763	0.12041	0.04066	0.02442	0.01222	0.00244
0.50000	0.56250	0.51389	0.45703	0.27173	0.09750	0.05914	0.02979	0.00599
0.90000	0.90250	0.88322	0.85963	0.74226	0.45009	0.31280	0.17345	0.03735
<u>Hinweis:</u>	$\sigma_\varepsilon^2 = 1 - \rho^2$ und damit $\gamma_0 = 1$ sowie $\gamma_1 = \rho$ für jedes ρ .							

$\rho = 0.5$ der wahre Wert $\gamma_1 = \gamma_0\rho = 0.50$. Die ausgewiesene Verzerrung in der Tabelle zeigt dann, daß sich für den Erwartungswert

$$E[C] = 0.50 - 0.45703 = 0.04297$$

ergibt!

12.3.3 Die Schätzfunktion C_a bei kleinem T : Additive Überlagerung

Wir schreiben die asymmetrische Form der Schätzfunktion C_a als

$$C_a = \frac{1}{T-1} \sum_{t=2}^T (X_t + U_t - \bar{X}_2 - \bar{U}_2)(X_{t-1} + U_{t-1} - \bar{X}_1 - \bar{U}_1)$$

wobei unterstellt wird, daß X_t^a durch additive Überlagerung generiert wird.

Erwartungswertbildung ergibt wegen der Unabhängigkeit von X_t und U_s für alle t und s :

$$\begin{aligned} E[C_a] &= E\left[\frac{1}{T-1} \sum_{t=2}^T (X_t - \bar{X}_2)(X_{t-1} - \bar{X}_1)\right] + E\left[\frac{1}{T-1} \sum_{t=2}^T (U_t - \bar{U}_2)(U_{t-1} - \bar{U}_1)\right] \\ &= E[C] - \frac{T-2}{(T-1)^2} \sigma_u^2 \quad (\text{siehe (12-13)}) \\ &= \gamma_1 - \zeta(\gamma, T) - \frac{T-2}{(T-1)^2} \sigma_u^2, \end{aligned} \quad (12-16)$$

wobei in der zweiten Zeile ausgenutzt wurde, daß⁹⁷

$$\begin{aligned} &E\left[\frac{1}{T-1} \sum_{t=2}^T (U_t - \bar{U}_2)(U_{t-1} - \bar{U}_1)\right] \\ &= E\left[\frac{1}{T-1} \sum_{t=2}^T \left(U_t - \left(\frac{1}{T-1} \sum_{s=2}^T U_s\right)\right) \left(U_{t-1} - \left(\frac{1}{T-1} \sum_{s=2}^T U_{s-1}\right)\right)\right] \\ &= E\left[\frac{1}{T-1} \sum_{t=2}^T U_t U_{t-1}\right] - E\left[\frac{1}{(T-1)^2} \sum_{t=2}^T \sum_{s=2}^T U_t U_{s-1}\right] \\ &= 0 - \frac{T-2}{(T-1)^2} \sigma_u^2 \end{aligned}$$

⁹⁷Sowohl Hans Schneeweiß als auch Andy Tremayne haben - zeitgleich - auf das bisher verwendete fehlerhafte Ergebnis hingewiesen, wonach der folgende Erwartungswert gleich Null sein soll!

gilt. Demnach ist der Erwartungswert von C_a bei additiver Überlagerung und $T \geq 3$ stets kleiner als der Erwartungswert von C . Damit ist klar, daß der "naive" IV-Schätzer bei additiver Überlagerung nicht konsistent ist. Siehe dazu im Einzelnen Abschnitt 12.4.

Speziell für den Fall, daß keine Autokorrelation vorliegt, gilt

$$E[C_a] = - \frac{T-2}{(T-1)^2} (\gamma_0 + \sigma_u^2) < 0$$

und demnach wird der wahre Autokorrelationskoeffizient $\gamma_1 (= 0)$ stärker als im Fall der Originaldaten (siehe Abschnitt 12.3.2) unterschätzt.

12.3.4 Die Schätzfunktion C_a bei kleinem T : Multiplikative Überlagerung

Im Fall der "allgemeinen" multiplikativen Überlagerung schreiben wir die asymmetrische Form der Schätzfunktion als

$$\begin{aligned} C_a &= \frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \frac{1}{T-1} \sum_{s=2}^T X_s U_s) (X_{t-1} U_{t-1} - \frac{1}{T-1} \sum_{s=2}^T X_{s-1} U_{s-1}) \\ &= \frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu + \mu - \frac{1}{T-1} \sum_{s=2}^T X_s U_s) (X_{t-1} U_{t-1} - \mu + \mu - \frac{1}{T-1} \sum_{s=2}^T X_{s-1} U_{s-1}) \\ &= \frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu_x) (X_{t-1} U_{t-1} - \mu_x) \\ &\quad - \frac{1}{(T-1)^2} \sum_{s=2}^T \sum_{s'=2}^T (X_s U_s - \mu_x) (X_{s'-1} U_{s'-1} - \mu_x) \end{aligned}$$

Erwartungswertbildung für die zwei Summanden ergibt folgendes: Für den ersten Summanden erhalten wir wegen (11-16)

$$E \left[\frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu_x) (X_{t-1} U_{t-1} - \mu_x) \right] = \text{cov}[X_t^a, X_{t-1}^a] = \text{cov}[X_t, X_{t-1}] = \gamma_1 \quad .$$

Für den zweiten Summanden ergibt sich

$$\begin{aligned} &E \left[\frac{1}{(T-1)^2} \sum_{s=2}^T \sum_{s'=2}^T (X_s U_s - \mu_x) (X_{s'-1} U_{s'-1} - \mu_x) \right] \\ &= \frac{1}{(T-1)^2} \left(\sum_{s=2}^{T-1} \text{var}[X_s U_s] + \sum_{s=2}^T \sum_{s'=2, s \neq s'}^T \text{cov}[X_s U_s, X_{s'-1} U_{s'-1}] \right) \\ &= \frac{1}{(T-1)^2} \left(\sum_{s=2}^{T-1} \text{var}[X_s U_s] + \sum_{s=2}^T \sum_{s'=2, s \neq s'}^T \text{cov}[X_s, X_{s'-1}] \right) \end{aligned}$$

wobei das Ergebnis in der dritten Zeile wieder ausnutzt, daß bei allgemeiner multiplikativer Überlagerung

$$\text{cov}[X_s U_s, X_{s'} U_{s'}] = \text{cov}[X_s, X_{s'}] \quad \text{falls } s \neq s' \quad .$$

gilt. Siehe (11-16). Damit erhalten wir

$$\begin{aligned} &E \left[\frac{1}{(T-1)^2} \sum_{s=2}^T \sum_{s'=2}^T (X_s U_s - \mu_x) (X_{s'-1} U_{s'-1} - \mu_x) \right] \\ &= \frac{T-2}{(T-1)^2} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{(T-1)^2} \left(2 \sum_{k=1}^{T-2} (T-k-1) \gamma_k + \gamma_{T-1} \right) \\ &= \frac{T-2}{(T-1)^2} \left(\sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{T-2}{(T-1)^2} \gamma_0 + \frac{1}{(T-1)^2} \left(2 \sum_{k=1}^{T-2} (T-k-1) \gamma_k + \gamma_{T-1} \right) \\ &= \frac{T-2}{(T-1)^2} \left(\sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \zeta(\gamma, T) \end{aligned}$$

Diese Einzelergebnisse lassen sich wie folgt zusammenfassen: Für den Erwartungswert der asymmetrischen Form von C_a gilt bei allgemeiner multiplikativer Überlagerung

$$E[C_a] = \gamma_1 - \left\{ \frac{T-2}{(T-1)^2} \left(\sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \zeta(\gamma, T) \right\} \quad , \quad (12-17)$$

wobei $\zeta(\gamma, T)$ durch (12-14) gegeben ist.

12.3.5 Die Schätzfunktion C_a bei kleinem T : Multiplikative Überlagerung mit Faktorstruktur

Die Abschnitt 12.3.4 entsprechende Analyse für den Höhenfall hat zu beachten, daß die (Auto-)Kovarianzen der überlagerten Variablen **nicht** gleich den (Auto-)Kovarianzen der Originalvariablen sind. Siehe dazu auch die Ableitung des Erwartungswertes von S_a^2 in Abschnitt 11.4.3; auf die Ergebnisse jenes Abschnittes werden wir uns im Folgenden beziehen, insbesondere wird die in (11-32) definierte Autokovarianz

$$\gamma_{t-s}^H \equiv cov[X_t^a, X_s^a] = \gamma_0 \rho^{|t-s|} + \delta^2 \left\{ (1 + \rho + \dots + \rho^{t-s-1}) \frac{\rho_0^2}{1-\rho} + \rho^{(t-s)} \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right\}$$

eine Rolle spielen. Man beachte, daß diese "Pseudo"-Autokovarianz vom Überlagerungsparameter δ^2 abhängt.

Wir starten wieder von der Zerlegung der Schätzfunktion C_a in zwei Summanden

$$\begin{aligned} C_a &= \frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu_x) (X_{t-1} U_{t-1} - \mu_x) \\ &\quad - \frac{1}{(T-1)^2} \sum_{s=2}^T \sum_{s'=2}^T (X_s U_s - \mu_x) (X_{s'-1} U_{s'-1} - \mu_x) \end{aligned}$$

(siehe Abschnitt 12.3.4) und bestimmen nun die Erwartungswerte der einzelnen Summanden unter Beachtung der Höhe-Spezifikation $X_t^a = (1 + \delta D + \varepsilon_t) X_t$.

Für den ersten Summanden erhalten wir wegen (11-32)

$$E \left[\frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu_x) (X_{t-1} U_{t-1} - \mu_x) \right] = cov[X_t^a, X_{t-1}^a] = \gamma_1^H \quad .$$

Für den zweiten Summanden ergibt sich⁹⁸

$$\begin{aligned} &E \left[\frac{1}{(T-1)^2} \sum_{s=2}^T \sum_{s'=2}^T (X_s U_s - \mu_x) (X_{s'-1} U_{s'-1} - \mu_x) \right] \\ &= \frac{1}{(T-1)^2} \left(\sum_{s=2}^{T-1} var[X_s U_s] + \sum_{s=2}^T \sum_{s'=2, s \neq s'}^T cov[X_s U_s, X_{s'-1} U_{s'-1}] \right) \\ &= \frac{T-2}{(T-1)^2} \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) \\ &\quad + \frac{1}{(T-1)^2} (2(T-2)\gamma_1^H + 2(T-3)\gamma_2^H + \dots + 6\gamma_{T-4}^H + 4\gamma_{T-3}^H + 2\gamma_{T-2}^H + \gamma_{T-1}^H) \end{aligned}$$

⁹⁸Bei der Bestimmung der Kovarianzen gilt das Schema aus Fußnote 95.

Die Einzelergebnisse lassen sich wie folgt zusammenfassen: Für den Erwartungswert von C_a (asymmetrische Form) gilt bei multiplikativer Höhne-Überlagerung (Multiplikative Überlagerung mit Faktorstruktur)

$$E[C_a] = \gamma_1^H - \left\{ \frac{T-2}{(T-1)^2} \left((\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \zeta^H(\gamma^H, \delta^2, T) \right\} \quad (12-18)$$

wobei $\zeta^H(\gamma^H, \delta^2, T)$ durch

$$\begin{aligned} \zeta^H(\gamma^H, \delta^2, T) &= \frac{T-2}{(T-1)^2} \gamma_0 \\ &+ \left(\frac{1}{T-1} \right)^2 [2 \{ (T-2)\gamma_1^H + (T-3)\gamma_2^H + \dots + 3 \cdot \gamma_{T-4}^H + 2 \cdot \gamma_{T-3}^H + \gamma_{T-2}^H \} + \gamma_{T-1}^H] \end{aligned} \quad (12-19)$$

gegeben ist. Dabei deutet das Argument δ^2 in dieser Funktion darauf hin, daß die Biasfunktion über γ^H von diesem Parameter abhängt.

Zwar erscheint auf den ersten Blick die Struktur dieses Erwartungswertes große Ähnlichkeit mit dem Erwartungswert im Fall der allgemeinen multiplikativen Überlagerung zu haben, doch ergeben sich bei näherem Hinsehen deutliche Unterschiede.

- Da für $T \rightarrow \infty$ der zweite und dritte Term in (12-18) verschwinden, strebt der IV-Schätzer im Höhnefall gegen $\gamma_1^H \neq \gamma_1$ und ist deshalb nicht konsistenter Schätzer der Autokovarianz.
- Für den Fall $\rho = 0$ gilt

$$\gamma_k^H = \delta^2 \rho_0^2, \quad k = 1, 2, \dots,$$

d.h. die Pseudo-Auto-Kovarianzen γ_k^H verschwinden auch für $\rho = 0$ nicht.

- Für kleines T ist die Beschreibung der Verzerrung kompliziert. Siehe jedoch die Simulationsergebnisse für diese Anonymisierungsvariante in Abschnitt 12.5.3.

12.4 Relevanz für den IV-Schätzer

12.4.1 Allgemeine Ausführungen

Entsprechend den Ergebnissen für den naiven Panelschätzer $\hat{\beta}_W^a$ in Abschnitt 11.7 lassen sich die vorhergehenden Ergebnisse für den Erwartungswert der Schätzfunktionen C und C_a für die Ableitung der Wahrscheinlichkeitsgrenzwerte des IV-Schätzers nutzen, der bei Verwendung des verzögerten Regressors als Instrument gemäß (12-5) durch

$$\begin{aligned} \hat{\beta}^{a,IV} &= \frac{\sum_t (x_{t-1}^a - \bar{x}_1^a)(y_t - \bar{y}_2)}{\sum_t (x_{t-1}^a - \bar{x}_1^a)(x_t^a - \bar{x}_2^a)} \\ &= \frac{\sum_t (x_{t-1}^a - \bar{x}_1^a)(x_t - \bar{x}_2)}{\sum_t (x_{t-1}^a - \bar{x}_1^a)(x_t^a - \bar{x}_2^a)} \beta + \frac{\sum_t (x_{t-1}^a - \bar{x}_1^a)(\eta_t - \bar{\eta}_2)}{\sum_t (x_{t-1}^a - \bar{x}_1^a)(x_t^a - \bar{x}_2^a)} \end{aligned}$$

gegeben ist, wenn die "asymmetrische Form" des Schätzers verwendet wird.

Wir betrachten nun die Erweiterung des "asymmetrischen" Schätzers für den Panelfall. Dann lautet der obige Schätzer wie folgt (siehe (12-7) und (12-8)):

$$\hat{\beta}_P^a = \left(\sum_{i=1}^N \mathbf{Z}'\langle i \rangle \left(\mathbf{I}_{T-1} - \frac{1}{T-1} \boldsymbol{\nu}_{T-1} \boldsymbol{\nu}'_{T-1} \right) \mathbf{X}^a \langle i \rangle \right)^{-1} \sum_{i=1}^N \mathbf{Z}'\langle i \rangle \left(\mathbf{I}_{T-1} - \frac{1}{T-1} \boldsymbol{\nu}_{T-1} \boldsymbol{\nu}'_{T-1} \right) \mathbf{y} \langle i \rangle . \quad (12-20)$$

bzw. für den Fall $K = 2$ (nur ein Regressor sowie Absolutglied)

$$\begin{aligned} \hat{\beta}_P^{a,IV} &= \frac{\sum_{i=1}^N \sum_{t=2}^T \left(x_{i,t-1}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{i,s-1}^a \right) \right) \left(y_{it} - \left(\frac{1}{T-1} \sum_{s=2}^T y_{is} \right) \right)}{\sum_{i=1}^N \sum_{t=2}^T \left(x_{i,t-1}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{i,s-1}^a \right) \right) \left(x_{it}^a - \left(\frac{1}{T-1} \sum_{s=2}^T x_{i,s}^a \right) \right)} \\ &= \frac{\sum_{i=1}^N \sum_{t=2}^T \left(x_{i,t-1}^a - \bar{X}_{i\bullet 1}^a \right) \left(y_{it} - \bar{Y}_{i\bullet 2} \right)}{\sum_{i=1}^N \sum_{t=2}^T \left(x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a \right) \left(x_{it}^a - \bar{x}_{i\bullet 2} \right)} , \end{aligned} \quad (12-21)$$

mit

$$\bar{X}_{i\bullet 1}^a = \left(\frac{1}{T-1} \sum_{t=2}^T x_{i,t-1}^a \right) , \quad \bar{X}_{i\bullet 2}^a = \left(\frac{1}{T-1} \sum_{t=2}^T x_{i,t}^a \right) , \quad \bar{Y}_{i\bullet 2} = \left(\frac{1}{T-1} \sum_{t=2}^T y_{i,t-1} \right) .$$

Für den Wahrscheinlichkeitsgrenzwert dieser Schätzfunktion (mit $N \rightarrow \infty$ sowie T fest) ergibt sich

$$\begin{aligned} \text{plim } \hat{\beta}_P^{a,IV} &= \frac{\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a \right) \left(y_{it} - \bar{y}_{i\bullet 2} \right)}{\text{plim } \sum_t \sum_i \left(x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a \right) \left(x_{it}^a - \bar{x}_{i\bullet 2}^a \right)} \\ &= \frac{\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a \right) \left(x_{it} - \bar{x}_{i\bullet 2} \right)}{\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(x_{i,t-1}^a - \bar{X}_{i\bullet 1}^a \right) \left(x_{it}^a - \bar{x}_{i\bullet 2}^a \right)} \beta \\ &\quad + \frac{\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a \right) \left(\eta_{it} - \bar{\eta}_{i\bullet 2} \right)}{\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a \right) \left(x_{it}^a - \bar{x}_{i\bullet 2}^a \right)} \\ &= \frac{\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(x_{i,t-1} - \bar{x}_{i\bullet 1} \right) \left(x_{it} - \bar{x}_{i\bullet 2} \right)}{\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a \right) \left(x_{it}^a - \bar{x}_{i\bullet 2}^a \right)} \beta \quad (12-22) \\ &= \frac{\text{plim } \bar{C}}{\text{plim } \bar{C}_a} \beta \\ &= \frac{\text{plim } C}{\text{plim } C_a} \beta \\ &= \frac{E[C]}{E[C_a]} \beta \end{aligned}$$

mit C und C_a aus (12-10) und (12-11).

Bei dieser Ableitung werden zunächst die Unkorreliertheit von X_{it} und η_{it} sowie die bereits weiter oben verwendete Gleichheit der Kovarianz von X_{t-1}^a und X_t mit der Kovarianz von $X_{i,t-1}$ mit X_{it} , also

$$\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(X_{i,t-1}^a - \bar{X}_{i\bullet}^a \right) \left(X_{it} - \bar{X}_{i\bullet} \right) = \text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i \left(X_{i,t-1} - \bar{X}_{i\bullet} \right) \left(X_{it} - \bar{X}_{i\bullet} \right)$$

benutzt. Sodann betrachten wir das arithmetische Mittel der N individuellen empirischen Autokovarianzen:

$$\bar{C} = \frac{1}{N} \sum_i \left\{ \frac{1}{T-1} \sum_t (x_{i,t-1} - \bar{x}_{i\bullet 1})(x_{it} - \bar{x}_{i\bullet 2}) \right\} = \frac{1}{N} \sum_i C_i$$

und

$$\bar{C}_a = \frac{1}{N} \sum_i \left\{ \frac{1}{T-1} \sum_t (x_{i,t-1}^a - \bar{x}_{i\bullet 1}^a)(x_{it}^a - \bar{x}_{i\bullet 2}^a) \right\} = \frac{1}{N} \sum_i C_{a,i}$$

Wegen der angenommenen identischen Verteilung über alle N Unternehmen gilt jedoch $E[\bar{C}] = E[C]$ sowie $E[\bar{C}_a] = E[C_a]$ und damit selbstverständlich auch $\text{plim } \bar{C} = \text{plim } C$ bzw. $\text{plim } \bar{C}_a = \text{plim } C_a$. Schließlich konvergieren die beiden Schätzfunktionen C und C_a im quadratischen Mittel gegen die jeweiligen Erwartungswerte.

Wir benutzen nun die Ergebnisse aus den Abschnitten 12.3.2, 12.3.3, 12.3.4 sowie 12.3.5, um für additive und multiplikative Überlagerung die Wahrscheinlichkeitsgrenzwerte des "asymmetrischen" Panel-IV-Schätzers (12-28) zu bestimmen.

12.4.2 Additive Überlagerung

Da im Fall der additiven Überlagerung die beiden Erwartungswerte mit dem Bias $\zeta(\gamma, T)$ für kleines T behaftet sind, und $E[C_a]$ ferner eine weitere Verzerrung aufweist, die von T sowie der Überlagerungsvarianz σ_u^2 abhängt, ergibt sich in diesem Fall für den Panel-IV-Schätzer (12-21)

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_P^{a,IV} = \frac{\gamma_1 - \zeta(\gamma, T)}{\gamma_1 - \left\{ \frac{T-2}{(T-1)^2} \sigma_u^2 + \zeta(\gamma, T) \right\}} \beta \neq \beta \quad (12-23)$$

mit

$$\zeta(\gamma, T) = \frac{T-2}{(T-1)^2} \gamma_0 + \frac{1}{(T-1)^2} \left(2 \sum_{k=1}^{T-2} (T-k-1) \gamma_k + \gamma_{T-1} \right) .$$

Siehe (12-13) und (12-14). Dabei nimmt die Biasfunktion ζ für **positive Autokorrelation** stets nur negative Werte an und verschwindet für $T \rightarrow \infty$.

Demnach ist der IV-Panelschätzer bei additiver Fehlerüberlagerung nicht konsistent, sofern die verzögerte, ebenfalls fehlerüberlagerte Regressorvariable verwendet wird. Für großes T (d.h. $\zeta(\gamma, T) \rightarrow \infty$) ergibt sich eine **Überschätzung** des wahren Parameter-Wertes. Dagegen ist für kleines T keine eindeutige Aussage über die Richtung der Verzerrung möglich.

Die Konsistenz ist gegeben, sofern keine Fehlerüberlagerung besteht und damit $\sigma_u^2 = 0$ gilt. Unerheblich ist dabei, wie groß der Wert der Biasfunktion ζ ist! Allerdings ist die durch den Ausdruck

$$\frac{T-2}{(T-1)^2} \sigma_u^2$$

verursachte Verzerrung nicht besonders folgenreich: Selbst für einen so kleinen Wert wie $T = 4$ ergibt sich $2/9 \sigma_u^2 = 0.222 \cdot \sigma_u^2$ und für $T = 10$ erhalten wir $0.098 \cdot \sigma_u^2$. Falls die

Tabelle 12.8: Untersuchung der Bedingungen für $E[C_a] = 0$ bei additiver Überlagerung

T	Überlagerungsvarianz (σ_u^2)					
	0.01	0.09	0.25	1.00	2.25	4.00
3	-	-	-	-	-	-
4	0.5701	0.6321	0.7661	-	-	-
5	0.3691	0.4041	0.4751	0.8341	-	-
6	0.2681	0.2921	0.3391	0.5671	0.9331	-
7	0.2111	0.2281	0.2641	0.4321	0.6951	-
8	0.1741	0.1881	0.2161	0.3501	0.5601	0.8031
9	0.1481	0.1601	0.1841	0.2961	0.4711	0.6781
10	0.1291	0.1391	0.1601	0.2571	0.4091	0.5911
11	0.1141	0.1231	0.1411	0.2271	0.3621	0.5261
12	0.1021	0.1111	0.1271	0.2031	0.3251	0.4751
13	0.0931	0.1001	0.1151	0.1841	0.2951	0.4341
14	0.0851	0.0921	0.1051	0.1691	0.2701	0.3991
15	0.0791	0.0851	0.0971	0.1551	0.2501	0.3701
20	0.0571	0.0611	0.0701	0.1121	0.1801	0.2711
25	0.0441	0.0481	0.0551	0.0881	0.1411	0.2141
30	0.0371	0.0391	0.0451	0.0721	0.1161	0.1771
40	0.0271	0.0291	0.0331	0.0531	0.0861	0.1311
50	0.0211	0.0231	0.0261	0.0421	0.0681	0.1041

Bemerkung: Die Werte in der Tabelle geben den jeweiligen Wert von ρ an.

Überlagerung also nicht zu groß ist, wird diese Verzerrung nicht ins Gewicht fallen. Man vergleiche dazu die Simulationsergebnisse in Abschnitt 12.5.1.

Man beachte, daß der Ausdruck

$$\gamma_1 - \left\{ \frac{T-2}{(T-1)^2} \sigma_u^2 + \zeta(\gamma, T) \right\}$$

durchaus negative Werte annehmen kann. Wenn der Ausdruck allerdings gleich Null ist, ist der Wahrscheinlichkeitsgrenzwert nicht definiert. Außerdem kann es vorkommen, daß $E[C] > 0$ und $E[C_a] < 0$ (oder umgekehrt) gilt, wodurch sich sogar ein Vorzeichenwechsel im Wahrscheinlichkeitsgrenzwert ergeben kann!

Schwerwiegender ist, daß der Nenner-Ausdruck

$$\sum_t (x_{t-1}^a - \bar{x}_1^a)(x_t^a - \bar{x}_2^a)$$

des Panel-IV-Schätzers bei bestimmten Parameter-Konstellationen beliebig nahe bei Null sein kann und damit - wegen der dann bestehenden Ausreißer - zu einer sehr unzuverlässigen Schätzung führen kann.⁹⁹ Siehe dazu die Simulationsergebnisse in Abschnitt 12.5.1. Tabelle 12.8 gibt einen Eindruck von denjenigen Parameter-Konstellationen, bei denen der Nennerausdruck nahe bei Null liegt. Dazu bestimmen wir für ausgewählte Werte von T und verschiedene Werte von σ_u^2 diejenigen Werte von ρ , für die bei additiver Überlagerung $E[C_a] = 0$ gilt.¹⁰⁰

⁹⁹Dieses Problem des IV-Schätzers ist natürlich in der Literatur bekannt. Siehe beispielsweise Carroll et al (2006), S. 133 sowie Fuller(1987) S 54.

¹⁰⁰Die Parameterwerte für σ_ε und ρ_0 sind wie folgt: ****(nachtragen!)

Tabelle 12.9: Die Bedingung $E[C_a] = 0$ bei allgemeiner multiplikativer Überlagerung

T	Überlagerungsvarianz (σ_u^2)					
	0.01	0.09	0.25	1.00	2.25	4.00
3						
4	0.5729	0.7043, 0.8669				
5	0.3697	0.4116, 0.9637	0.5162, 0.8777			
6	0.2683	0.2940, 0.9814	0.3493, 0.9439			
7	0.2104	0.2290, 0.9885	0.2678, 0.9665	0.5075, 0.8041		
8	0.1732	0.1880	0.2183, 0.9775	0.3795, 0.8884		
9	0.1474	0.1597	0.1847, 0.9838	0.3113, 0.9243		
10	0.1283	0.1389	0.1603, 0.9877	0.2659, 0.9447	0.4855, 0.8275	
11	0.1137	0.1230	0.1417	0.2328, 0.9576	0.4064, 0.8799	
12	0.1021	0.1104	0.1271	0.2075, 0.9663	0.3544, 0.9094	
13	0.0926	0.1001	0.1152	0.1873, 0.9726	0.3159, 0.9285	0.5452, 0.8159
14	0.0848	0.0916	0.1054	0.1708, 0.9773	0.2857, 0.9419	0.4730, 0.8648
15	0.0782	0.0845	0.0971	0.1571, 0.9808	0.2612, 0.9517	0.4237, 0.8935
20	0.0563	0.0608	0.0698	0.1123	0.1843, 0.9765	0.2890, 0.9533
25	0.0440	0.0475	0.0546	0.0875	0.1431, 0.9861	0.2222, 0.9733
30	0.0362	0.0390	0.0448	0.0718	0.1170, 0.9908	0.1811, 0.9826
40	0.0266	0.0288	0.0330	0.0528	0.0859, 0.9951	0.1326, 0.9910
50	0.0211	0.0228	0.0261	0.0418	0.0679, 0.9970	0.1047, 0.9945

Bemerkung: $\sigma_\varepsilon = 1, \varrho_0 = 0$
 Die Werte in der Tabelle geben den (die) jeweiligen Wert(e) von ϱ an.
 In den leeren Feldern existiert kein $\varrho \in \{-1, +1\}$.

12.4.3 Allgemeine multiplikative Überlagerung

Im Fall der "allgemeinen" multiplikativen Überlagerung erhalten wir als Wahrscheinlichkeitsgrenzwert für den Panel-IV-Schätzer (12-21)

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^{a,IV} = \frac{\gamma_1 - \zeta(\gamma, T)}{\gamma_1 - \zeta(\gamma, T) - \frac{T-2}{(T-1)^2} \left(\sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right)} \beta \quad (12-24)$$

mit $\zeta(\gamma, T)$ aus (12-14). Siehe (12-17). Dabei nimmt die Biasfunktion ζ für positive Autokorrelation stets nur positive Werte an und verschwindet für $T \rightarrow \infty$.

Die für den Fall der additiven Überlagerung geschilderten Probleme, die sich für den Fall, daß der Nenner von (12-24) (annähernd) gleich Null ist, gelten natürlich auch hier. In Tabelle 12.9 werden für ausgewählte Parameterwerte die kritischen Werte von ϱ angegeben. Man beachte, daß für größere Überlagerungsvarianz auch mehr als ein Wert "kritisch" sein kann, wobei alle Werte im positiven Bereich liegen, aber sowohl "geringe" als auch "starke" Autokorrelation beschreiben.

12.4.4 Multiplikative Höhne-Überlagerung

Im Fall der multiplikativen Höhne-Überlagerung erhalten wir als Wahrscheinlichkeitsgrenz-

wert für den Panel-IV-Schätzer (12-21)

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^{a,IV} = \frac{\gamma_1^H - \zeta^H(\gamma^H, T)}{\gamma_1^H - \zeta^H(\gamma^H, T) - \frac{T-2}{(T-1)^2} \left((\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right)} \beta \quad (12-25)$$

mit $\zeta^H(\gamma^H, T)$ aus (12-19). Siehe (12-18).

Die für den Fall der additiven Überlagerung geschilderten Probleme, die sich für den Fall, daß der Nenner von (12-25) (annähernd) gleich Null ist, gelten natürlich auch hier. Dabei ist zu beachten, daß in diesem Fall für den Erwartungswert des Nenners, d.h. für den Erwartungswert von C_a

$$\begin{aligned} E[C_a] &= \gamma_0 \rho + \delta^2 \left\{ \frac{\rho_0^2}{1-\rho} + \rho \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right\} \\ &\quad - \zeta^H(\gamma^H, T) - \frac{T-2}{(T-1)^2} \left((\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) \end{aligned}$$

gilt. Wegen $\gamma_1 = \rho\gamma_0$ ist sowohl der zweite Summand in der ersten Zeile als auch beide Summanden in der zweiten Zeile als Bias von γ_1 zu interpretieren. Auch hier ist zu untersuchen, unter welchen Bedingungen dieser Erwartungswert gleich Null ist. Darauf gehen wir nicht mehr im Einzelnen ein!

12.4.5 Korrektur-IV-Schätzer bei multiplikativer Überlagerung

Der Wahrscheinlichkeitsgrenzwert des (naiven) Panel-IV-Schätzers bei allgemeiner multiplikativer Überlagerung weist dieselbe Struktur auf wie der naive Panelschätzer $\hat{\beta}_W^a$. Man vergleiche dazu (C-25) mit Formel (11-45), die hier der Einfachheit halber noch einmal wiederholt wird:

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^a = \frac{\gamma_0 - \psi(\gamma, T)}{\gamma_0 - \psi(\gamma, T) + (1 - \frac{1}{T})\sigma_u^2} \beta$$

. Der entsprechende Korrektorschätzer lautete (siehe (11-48))

$$\widehat{\beta}_W^{a,korr} = \frac{S_a^2}{S_a^2 - (1 - \frac{1}{T})\sigma_u^2} \hat{\beta}_W^a$$

Entsprechend konstruieren wir nun einen Korrektorschätzer zu dem Panel-IV-Schätzer:

$$\widehat{\beta}_W^{a,IV,korr} = \frac{S_a^2}{S_a^2 + \frac{1}{T} \left(\sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right)} \hat{\beta}_W^{IV,a} \quad (12-26)$$

Das Unangenehme an diesem Vorschlag ist, daß im Gegensatz zum oben präsentierten Korrektorschätzer in diesem Fall nicht nur die Überlagerungsvarianz σ_u^2 , sondern auch die Parameter (ρ_0 , ρ und σ_ω^2) des autoregressiven Prozesses bekannt sein oder geschätzt werden müssen. In der Realität, auch bei anonymisierten Daten, wird damit dasselbe Problem virulent, das wir bereits bei Konstruktion von Korrektorschätzern zum naiven Panelschätzer bei multiplikativer Überlagerung kennen gelernt haben:¹⁰¹ Im empirischen Kontext kann

¹⁰¹Siehe Abschnitt 11.6.

Tabelle 12.10: Simulations-Design für IV-Schätzung bei additiver Überlagerung

ÜBERLAGERUNG mit konstantem γ_0	
Anzahl Wiederholungen	2000
Anzahl Wellen (T)	4, 10, 30, 50
Anzahl Beobachtungseinheiten (N)	1, 200
AR(1) - Parameter ρ_0	4.35
AR(1) - Parameter ρ_1	-0.9, -0.5, -0.1, 0.0, +0.1, +0.5, +0.9
AR(1) - Parameter $\gamma_0 = \sigma_\omega^2 / (1 - \rho_1^2)$	1.00
Überlagerungsvarianz σ_u^2	0.114 ²
Regressionsparameter α	0.00
Regressionsparameter β	-2.50
Regressionsparameter σ_η^2	1.00

nicht ein identisches ρ bzw., allgemeiner, ein identischer AR(1)-Prozeß, für alle N Unternehmen unterstellt werden. Wir haben es also nicht nur mit einem Schätzproblem, sondern auch mit einem Spezifikationsproblem zu tun.

12.5 Simulationsergebnisse für den IV-Schätzer

12.5.1 IV-Schätzung bei additiver Überlagerung

In diesem Abschnitt untersuchen wir den Panel-IV-Schätzung im Zusammenhang mit der **additiven** Überlagerung, um einen Eindruck von der in Abschnitt 12.4.2 hergeleiteten Verzerrung zu erhalten. Außerdem wurde dort bereits auf die Problematik hingewiesen, die sich durch die Verzerrung für endliches T ergibt: Falls diese Verzerrung zu einem Wert von annähernd Null für

$$E[C_a] = \gamma_1 - \zeta(\gamma, T) - \frac{T-2}{(T-1)^2} \sigma_u^2]$$

führt, ergibt sich - so zumindest unsere Behauptung - eine numerisch bedingte (unendlich) große Varianz für den IV-Schätzer.

In Tabelle 12.10 sind die Details des Designs der Simulations-Studie zusammengestellt. Im Gegensatz zum Design beim KQ-Schätzer in Abschnitt 11.5.2 wurde mehr Wert auf sehr kleines T gelegt, außerdem wurden mehr unterschiedliche Werte für ρ untersucht. Um die dadurch erzeugte Variation in der Varianz der Regressor-Variablen zu vermeiden, wurde die Varianz des AR(1)-Prozesses so gewählt, daß γ_0 für alle ρ konstant bleibt. In den Simulationen, über die in der Tabelle 12.11 berichtet wird, wurde $\gamma_0 = 1.0$ gewählt. Die Ergebnisse der Simulationen für den in (12-7) bzw. (12-8) gegebenen Panel-IV-Schätzer sind in den Tabellen 12.11 zusammengefaßt.

Die Tabelle 12.11 berichtet über die Panel-IV-Schätzung für "großes" N . In der ersten Spalte wird jeweils T angegeben, das für unterschiedliche Werte von ρ die Ergebnisse ausweist. Die folgenden fünf Spalten geben einen Hinweis auf die Art der (theoretischen) Verzerrung: Der wahre Parameter γ_1 wird gemeinsam mit den beiden Verzerrungskomponenten ζ und $(T-2)\sigma_u^2/(T-1)^2$ und dem Resultat $E[C_a] = \gamma_1 - \zeta - (T-2)\sigma_u^2/(T-1)^2$ sowie der resultierende Wahrscheinlichkeitsgrenzwert $\text{plim} \hat{\beta}^{IV}$ ausgewiesen. Es folgt der Schätzwert für β . Dieser wird berechnet als arithmetisches Mittel der Schätzwerte aus allen 2000 Simulationsläufen und ist zu vergleichen mit dem Median, der das 50%-Quantil aus diesen Werten angibt. Die Standardabweichung wurde aus den 2000 Schätzwerten der einzelnen

Tabelle 12.11: Panel-IV-Schätzung bei additiver Überlagerung

T	ϱ	γ_1	$\zeta(\gamma, T)$	$\frac{T-2}{(T-1)^2} \sigma_u^2$	$E[C_a]$	$\text{plim} \hat{\beta}^{IV}$	estimate	std. dev.	min.	median	max.
4	-0.9000	-0.9000	-0.0788	0.0029	-0.8241	-2.4912	-2.4917	0.0494	-2.6986	-2.4921	-2.2973
4	-0.5000	-0.5000	0.0417	0.0029	-0.5446	-2.4867	-2.4853	0.0733	-2.8090	-2.4858	-2.2295
4	-0.1000	-0.1000	0.1799	0.0029	-0.2828	-2.4745	-2.4723	0.1244	-2.9570	-2.4692	-2.1020
4	0.0000	0.0000	0.2222	0.0029	-0.2251	-2.4679	-2.4711	0.1627	-3.2483	-2.4680	-1.9334
4	0.1000	0.1000	0.2690	0.0029	-0.1719	-2.4580	-2.4662	0.2027	-3.1961	-2.4695	-1.8158
4	0.5000	0.5000	0.5139	0.0029	-0.0168	-2.0696	0.3200	176.2937	-2107.1278	-2.1942	7565.4026
4	0.9000	0.9000	0.8832	0.0029	0.0139	-3.0198	-3.2052	1.9881	-43.0973	-2.9972	2.6374
10	-0.9000	-0.9000	-0.0027	0.0013	-0.8985	-2.4964	-2.4961	0.0275	-2.5933	-2.4958	-2.4097
10	-0.5000	-0.5000	0.0302	0.0013	-0.5314	-2.4940	-2.4947	0.0450	-2.6538	-2.4952	-2.3557
10	-0.1000	-0.1000	0.0806	0.0013	-0.1819	-2.4824	-2.4780	0.1323	-2.9228	-2.4791	-1.9322
10	0.0000	0.0000	0.0988	0.0013	-0.1000	-2.4679	-2.4736	0.2547	-3.6498	-2.4765	-1.2845
10	0.1000	0.1000	0.1204	0.0013	-0.0217	-2.3521	-2.6684	16.6783	-549.4932	-2.4488	163.9280
10	0.5000	0.5000	0.2717	0.0013	0.2270	-2.5141	-2.5123	0.0934	-2.8223	-2.5105	-2.2016
10	0.9000	0.9000	0.7423	0.0013	0.1565	-2.5205	-2.5206	0.0832	-2.7825	-2.5230	-2.2393
30	-0.9000	-0.9000	0.0012	0.0004	-0.9016	-2.4988	-2.4992	0.0156	-2.5503	-2.4989	-2.4378
30	-0.5000	-0.5000	0.0108	0.0004	-0.5113	-2.4979	-2.4988	0.0264	-2.5935	-2.4979	-2.4082
30	-0.1000	-0.1000	0.0272	0.0004	-0.1277	-2.4915	-2.4926	0.1099	-2.9073	-2.4901	-2.1308
30	0.0000	0.0000	0.0333	0.0004	-0.0337	-2.4679	-3.1742	31.5420	-1412.0212	-2.4641	25.4017
30	0.1000	0.1000	0.0407	0.0004	0.0589	-2.5184	-2.5249	0.2427	-3.5637	-2.5301	-1.7208
30	0.5000	0.5000	0.0975	0.0004	0.4021	-2.5027	-2.5015	0.0322	-2.6158	-2.5023	-2.3904
30	0.9000	0.9000	0.4501	0.0004	0.4495	-2.5024	-2.5020	0.0227	-2.5688	-2.5017	-2.4321
50	-0.9000	-0.9000	0.0009	0.0003	-0.9011	-2.4993	-2.4992	0.0116	-2.5389	-2.4994	-2.4625
50	-0.5000	-0.5000	0.0066	0.0003	-0.5068	-2.4987	-2.4976	0.0204	-2.5703	-2.4976	-2.4255
50	-0.1000	-0.1000	0.0163	0.0003	-0.1166	-2.4944	-2.4936	0.0932	-2.7899	-2.4954	-2.1605
50	0.0000	0.0000	0.0200	0.0003	-0.0203	-2.4679	-2.2448	9.1938	-125.8658	-2.4349	361.1547
50	0.1000	0.1000	0.0244	0.0003	0.0753	-2.5086	-2.5090	0.1387	-3.0593	-2.5099	-2.0578
50	0.5000	0.5000	0.0591	0.0003	0.4406	-2.5015	-2.5013	0.0231	-2.5851	-2.5012	-2.4241
50	0.9000	0.9000	0.3128	0.0003	0.5869	-2.5011	-2.5012	0.0148	-2.5480	-2.5016	-2.4558

Bemerkung: Alle Ergebnisse beziehen sich auf den wahren Parameter $\beta = -2.50$.

Simulationsläufe berechnet. Dazu korrespondieren Minimum und Maximum, die ebenfalls das Ausmaß der Streuung der Schätzung charakterisieren.

Die Ergebnisse zeigen zum einen, daß die durch die additive Überlagerung erzeugte Verzerrung für die meisten untersuchten Parameterkonstellationen praktisch ohne Bedeutung ist. Man vergleiche dazu die Spalte $\text{plim} \hat{\beta}^{IV}$. In den meisten Fällen stimmt der aus den Simulationen berechnete Schätzwert auch gut mit dem theoretischen Wert gemäß Wahrscheinlichkeitsgrenzwert überein.

Es gibt jedoch auch eine beträchtliche Anzahl von Fällen, in denen der Schätzwert mit einer sehr großen Varianz behaftet ist und demnach als Wert keine Relevanz hat, selbst wenn der betreffende Wert in der Nähe des wahren Wertes - 2.50 liegt. Für $T = 4$ beispielsweise ergibt sich diese Anomalie bei $\varrho = 0.50$ und $\varrho = 0.90$ und für $T = 10$ bei $\varrho = 0.10$. Ein Blick auf die jeweiligen Werte von $E[C_a]$ offenbart, daß in diesen Fällen die betreffenden Fälle besonders nahe bei Null liegen. Ferner bestätigt ein Blick auf die Tabelle 12.8, daß für $T = 4$ und Überlagerungsvarianz von $\sigma_u^2 = 0.01$ der Erwartungswert ungefähr bei $\varrho = 0.57$ den Wert Null aufweist. Der entsprechende Wert für $T = 10$ lautet $\varrho = 0.129$. Dabei ist zu beachten, daß in den Simulationen $\sigma_u^2 = 0.114^2 = 0.013$ gewählt wurde. Demnach ist offensichtlich das bereits in Abschnitt 12.4.2 angesprochene Problem existent: Für bestimmte Bereiche von ϱ , für die $E[C_a]$ annähernd Null ist, sind die Schätzungen sehr unzuverlässig wenn nicht gar unbrauchbar.

Um dieses Problem näher zu untersuchen, wird in den Tabellen 12.12 für $T = 4$ und 12.13 für $T = 10$ für unterschiedlich großes γ_0 sowie für Werte von ϱ im kritischen Bereich über zusätzliche Simulationsergebnisse berichtet.¹⁰² Für $T = 4$ ergibt sich im Bereich "positives

¹⁰²In den im Folgenden berichteten Simulationen wurde γ_0 wie folgt variiert: Es wurde der Parameter κ_ω aus der Menge $\{0.5, 0.7, 1.0, 2.0\}$ vorgegeben und daraus für bestimmtes ϱ die Varianz $\sigma_\omega^2 = \kappa_\omega^2(1 - \varrho^2)$ des AR(1)-Prozesses bestimmt. Damit gilt für jedes ϱ die Beziehung $\gamma_0 = \kappa_\omega^2$.

Tabelle 12.12: Panel-IV-Schätzung bei additiver Überlagerung - Ergebnisse für $T = 4$

$\kappa_\omega = \sqrt{\gamma_0}$	ρ	γ_1	$\zeta(\gamma, T)$	$\frac{T-2}{(T-1)^2} \sigma_u^2$	$E[C_a]$	$\text{plim}\beta^{IV}$	estimate	std. dev.	min.	median	max.
0.5000	0.2000	0.0500	0.0802	0.0029	-0.0331	-2.2819	-2.2726	0.5425	-4.6623	-2.2721	-0.0820
0.7000	0.2000	0.0980	0.1572	0.0029	-0.0621	-2.3838	-2.3861	0.3957	-4.2301	-2.3803	-0.6711
1.0000	0.2000	0.2000	0.3209	0.0029	-0.1238	-2.4417	-2.4303	0.2811	-3.5431	-2.4389	-1.0510
2.0000	0.2000	0.8000	1.2836	0.0029	-0.4864	-2.4852	-2.4867	0.1450	-3.0734	-2.4821	-1.7783
0.5000	0.3000	0.0750	0.0946	0.0029	-0.0225	-2.1795	-2.1642	0.8129	-6.2679	-2.1869	2.3736
0.7000	0.3000	0.1470	0.1855	0.0029	-0.0414	-2.3255	-2.3427	0.6399	-7.8699	-2.3114	0.8970
1.0000	0.3000	0.3000	0.3786	0.0029	-0.0814	-2.4113	-2.3973	0.4690	-4.3828	-2.4076	2.7603
2.0000	0.3000	1.2000	1.5142	0.0029	-0.3171	-2.4772	-2.4710	0.3141	-9.1409	-2.4756	3.4850
0.5000	0.4000	0.1000	0.1107	0.0029	-0.0136	-1.9673	-1.6737	28.5578	-274.0506	-1.9846	1201.3271
0.7000	0.4000	0.1960	0.2169	0.0029	-0.0238	-2.1966	-2.2326	4.1672	-83.6942	-2.2091	77.6507
1.0000	0.4000	0.4000	0.4427	0.0029	-0.0456	-2.3415	-2.5084	9.3533	-411.4039	-2.3472	24.5702
2.0000	0.4000	1.6000	1.7707	0.0029	-0.1736	-2.4584	-2.3617	3.1780	-24.2265	-2.4675	121.5913
0.5000	0.5500	0.1375	0.1381	0.0029	-0.0035	-0.4263	-2.3790	75.4800	-2043.1472	-1.5279	1103.8679
0.7000	0.5500	0.2695	0.2707	0.0029	-0.0041	-0.7181	-0.9902	73.8092	-665.8099	-2.0528	2819.5611
1.0000	0.5500	0.5500	0.5524	0.0029	-0.0053	-1.1282	-6.1135	226.7231	-9987.0822	-2.2409	1311.3230
2.0000	0.5500	2.2000	2.2095	0.0029	-0.0124	-1.9172	-0.9170	78.0645	-255.9409	-2.4828	3436.5548
0.5000	0.5750	0.1437	0.1431	0.0029	-0.0022	0.7358	-0.3164	129.4680	-1513.7416	-1.6562	3662.3847
0.7000	0.5750	0.2818	0.2805	0.0029	-0.0016	2.0100	0.5079	104.0616	-1223.2615	-2.2823	4148.8886
1.0000	0.5750	0.5750	0.5724	0.0029	-0.0003	25.1349	-1.5438	112.3559	-2111.2373	-2.5073	4344.0875
2.0000	0.5750	2.3000	2.2895	0.0029	0.0076	-3.4476	-1.2762	38.7139	-207.1781	-2.5034	1524.0109
0.5000	0.6000	0.1500	0.1482	0.0029	-0.0011	4.0032	-0.0416	73.8415	-572.9250	-1.7969	2699.0352
0.7000	0.6000	0.2940	0.2905	0.0029	0.0006	-14.6051	-1.9783	52.5163	-910.4642	-2.4601	1271.2841
1.0000	0.6000	0.6000	0.5929	0.0029	0.0042	-4.2096	2.8960	204.1129	-1040.4897	-2.5619	8825.6519
2.0000	0.6000	2.4000	2.3716	0.0029	0.0256	-2.7825	-2.4575	13.5400	-248.4783	-2.4763	291.7354
0.5000	0.6500	0.1625	0.1589	0.0029	0.0007	-12.3428	3.7981	604.3900	-12290.5900	-2.6320	23879.0225
0.7000	0.6500	0.3185	0.3114	0.0029	0.0042	-4.2149	-0.2181	107.8508	-1480.7130	-2.8729	4255.0709
1.0000	0.6500	0.6500	0.6355	0.0029	0.0116	-3.1225	-2.3476	97.1376	-2279.6583	-2.7482	3245.3277
2.0000	0.6500	2.6000	2.5421	0.0029	0.0551	-2.6311	-5.6873	100.1862	-4317.3386	-2.5753	109.2411
0.5000	0.7000	0.1750	0.1701	0.0029	0.0020	-6.0590	-9.4249	197.2981	-7212.4671	-3.2678	2718.0842
0.7000	0.7000	0.3430	0.3334	0.0029	0.0067	-3.5698	-10.5190	220.0383	-7062.0817	-3.0612	1137.6661
1.0000	0.7000	0.7000	0.6803	0.0029	0.0168	-2.9303	-3.3503	36.3913	-961.3121	-2.8823	382.6526
2.0000	0.7000	2.8000	2.7213	0.0029	0.0758	-2.5953	-3.5239	27.2669	-927.3822	-2.5848	90.8097
0.5000	0.7500	0.1875	0.1819	0.0029	0.0028	-5.1213	-2.5920	72.6750	-786.3934	-3.8336	2198.9430
0.7000	0.7500	0.3675	0.3564	0.0029	0.0082	-3.3836	-3.6552	38.3704	-1326.4063	-3.2086	862.0765
1.0000	0.7500	0.7500	0.7274	0.0029	0.0197	-2.8668	-3.2207	14.3612	-510.2465	-2.8188	128.9106
2.0000	0.7500	3.0000	2.9097	0.0029	0.0874	-2.5826	-2.8258	12.7786	-542.4809	-2.5879	85.6688
0.5000	0.8000	0.2000	0.1942	0.0029	0.0029	-4.9985	-2.1775	105.7082	-1338.5880	-4.0546	3792.2764
0.7000	0.8000	0.3920	0.3807	0.0029	0.0084	-3.3558	-4.2729	19.8844	-599.8657	-3.2939	222.6945
1.0000	0.8000	0.8000	0.7769	0.0029	0.0202	-2.8570	-4.4921	66.1130	-2934.9510	-2.8775	234.0335
2.0000	0.8000	3.2000	3.1076	0.0029	0.0896	-2.5806	-2.5982	1.3063	-22.9965	-2.5743	32.1554
0.5000	0.8500	0.2125	0.2072	0.0029	0.0024	-5.4908	21.9885	1059.0498	-1297.5238	-4.4914	47183.7096
0.7000	0.8500	0.4165	0.4061	0.0029	0.0075	-3.4621	-2.4583	37.4402	-91.2334	-3.3991	1520.3016
1.0000	0.8500	0.8500	0.8288	0.0029	0.0183	-2.8941	-2.8434	12.8523	-199.9107	-2.8872	514.7525
2.0000	0.8500	3.4000	3.3152	0.0029	0.0819	-2.5881	-2.5923	0.9132	-11.3731	-2.5881	25.2449

Bemerkung: Alle Ergebnisse beziehen sich auf den wahren Parameter $\beta = 2.50$.

großes" ρ eine äußerst unzuverlässige Schätzung, während aus Tabelle 12.11 klar ist, daß für "kleines positives" sowie "negatives" ρ diese Probleme nicht bestehen. Man beachte, daß teilweise der Wahrscheinlichkeitsgrenzwert einen **positiven** Wert ausweist und damit das Vorzeichen umkehrt! In allen Fällen ist dieses Ergebnis auf die Tatsache zurückzuführen, daß der Ausdruck im Nenner des Panel-IV-Schätzers in der Nähe von Null liegt. Die Ergebnisse für $T = 10$ in Tabelle 12.13 zeigen, daß für größeres T der kritische Bereich von ρ näher bei Null liegt und nicht ein so großes Intervall einnimmt: Für $\rho = 0.05$ im unteren Bereich und $\rho = 0.20$ im oberen Bereich sind die Ergebnisse in der Tabelle als durchaus zufriedenstellend zu bezeichnen, mit Ausnahme des Falls $\rho = 0.05, \gamma_0 = 0.01$. Wegen der äußerst geringen Varianz der Regressorvariable wirkt sich hier die additive Überlagerung besonders stark aus und führt zu einer größeren Variation des Schätzers.

12.5.2 IV-Schätzung bei allgemeiner multiplikativer Überlagerung

Die Simulationsergebnisse für den Panel-IV-Schätzer bei allgemeiner multiplikativer Überlagerung faßt die Tabelle 12.14 zusammen. Dabei gelten für das Simulationsdesign dieselben Daten wie für die additive Überlagerung. Siehe Tabelle 12.10. Insbesondere wurde die Varianz der Überlagerung mit $\sigma_u^2 = 0.114^2$ genau wie bei der additiven Überlagerung

Tabelle 12.13: Panel-IV-Schätzung bei additiver Überlagerung - Ergebnisse für $T = 10$

$\kappa_\omega = \sqrt{\gamma_0}$	ρ	γ_1	$\zeta(\gamma, T)$	$\frac{T-2}{(T-1)^2} \sigma_u^2$	$E[C_a]$	$\text{plim}\beta^{IV}$	estimate	std. dev.	min.	median	max.
0.5000	0.0500	0.0125	0.0273	0.0013	-0.0161	-2.3002	-1.9917	12.7212	-20.0457	-2.3267	559.1388
0.7000	0.0500	0.0245	0.0535	0.0013	-0.0302	-2.3939	-2.4722	1.6780	-41.1877	-2.4080	13.5916
1.0000	0.0500	0.0500	0.1091	0.0013	-0.0604	-2.4469	-2.4324	1.6954	-26.8193	-2.4389	58.6863
2.0000	0.0500	0.2000	0.4364	0.0013	-0.2377	-2.4865	-2.4564	2.0787	-8.1946	-2.4966	89.0441
0.5000	0.1000	0.0250	0.0301	0.0013	-0.0064	-1.9975	2.7029	185.7333	-266.4821	-2.0940	8243.7900
0.7000	0.1000	0.0490	0.0590	0.0013	-0.0113	-2.2156	3.5783	376.9342	-2777.0969	-2.3270	16440.1672
1.0000	0.1000	0.1000	0.1204	0.0013	-0.0217	-2.3521	-14.9792	638.7854	-28432.7572	-2.4185	2177.7277
2.0000	0.1000	0.4000	0.4816	0.0013	-0.0829	-2.4613	-2.1623	16.3090	-100.9811	-2.4858	674.5581
0.5000	0.1250	0.0313	0.0316	0.0013	-0.0017	-0.5595	-2.1789	89.9022	-1475.9449	-2.1948	2416.8242
0.7000	0.1250	0.0612	0.0620	0.0013	-0.0020	-0.9026	-18.6413	1152.4166	-50397.4876	-2.2890	7981.9249
1.0000	0.1250	0.1250	0.1265	0.0013	-0.0028	-1.3389	-2.5858	28.5848	-708.2962	-2.4573	498.4158
2.0000	0.1250	0.5000	0.5059	0.0013	-0.0072	-2.0546	-2.2224	19.9697	-294.9419	-2.4775	723.4143
0.5000	0.1500	0.0375	0.0332	0.0013	0.0030	-3.5687	-3.0439	69.7711	-1000.8711	-2.6845	2124.8894
0.7000	0.1500	0.0735	0.0651	0.0013	0.0071	-2.9508	-7.8494	127.0002	-4135.6457	-2.6681	657.2797
1.0000	0.1500	0.1500	0.1329	0.0013	0.0159	-2.7023	-3.1888	54.6031	-2009.2168	-2.5417	1070.3468
2.0000	0.1500	0.6000	0.5314	0.0013	0.0673	-2.5477	-2.5829	17.2853	-444.4950	-2.4916	503.6825
0.5000	0.2000	0.0500	0.0367	0.0013	0.0121	-2.7660	-3.2310	24.0328	-1044.9392	-2.7141	171.8453
0.7000	0.2000	0.0980	0.0718	0.0013	0.0249	-2.6290	-2.7269	1.8887	-30.3274	-2.6544	20.1048
1.0000	0.2000	0.2000	0.1466	0.0013	0.0521	-2.5616	-2.6050	1.2370	-32.3131	-2.5598	8.3742
2.0000	0.2000	0.8000	0.5864	0.0013	0.2123	-2.5151	-2.5798	3.7868	-166.8395	-2.5092	21.1795
0.5000	0.2500	0.0625	0.0405	0.0013	0.0208	-2.6546	-2.6684	1.2320	-16.1142	-2.6585	42.6204
0.7000	0.2500	0.1225	0.0793	0.0013	0.0419	-2.5766	-2.5979	0.4475	-6.1807	-2.5718	0.6716
1.0000	0.2500	0.2500	0.1619	0.0013	0.0869	-2.5369	-2.5470	0.2791	-3.7644	-2.5415	-0.8561
2.0000	0.2500	1.0000	0.6475	0.0013	0.3513	-2.5091	-2.5115	0.1447	-3.3356	-2.5089	-1.7142

Bemerkung: Alle Ergebnisse beziehen sich auf den wahren Parameter $\beta = 2.50$.

gewählt.

Für $T = 4$ sind die ergebnisse in Ordnung, obwohl Tabelle 12.9 für Werte von ρ in der Nähe von 0.56 Probleme signalisiert. Auch für $T = 10$ ergeben sich weitgehend befriedigende Schätzergebnisse, nur für $\rho = 0.10$ sind die Ergebnisse schlechter. Dies rührt offensichtlich von der Nähe zum "kritischen" Wert 0.1281 her. Für größeres T ergeben sich mehrfach unbefriedigende Schätzergebnisse für "großes" positives ρ . Dies scheint - wie im additiven Fall - durch die große Überlagerungsvarianz verursacht zu sein, die selbst bei einem $E[C_a]$ deutlich von Null entfernt dafür sorgt, daß bestimmte Realisationen des Nenner-Ausdrucks (= Schätzwerte von $E[C_a]$) in der Nähe von Null liegen und damit die Ausreißerproblematik anstoßen. Andererseits gibt es ja auch Fälle in der Tabelle, in denen trotz einem nahe bei Null liegenden $E[C_a]$ die Schätzergebnisse durchaus akzeptabel sind.

12.5.3 IV-Schätzung bei multiplikativer Überlagerung mit Faktorstruktur

Simulationen zu diesem Fall wurden nicht durchgeführt. Allerdings zeigt die Analyse in Abschnitt 12.4.4, daß sich hier ebenfalls für bestimmte Parameterkonstellationen Probleme bei der Schätzung ergeben, weil der Nenner des Schätzers einen Wert nahe bei Null annimmt.

12.6 Ein (fast) optimaler IV-Schätzer

Bei der Überprüfung der Eigenschaften E1 bis E3 der Instrumentvariablen in Abschnitt ?? wird darauf hingewiesen, daß im Rahmen der Anonymisierung, nicht aber im "normalen" Ökonometriebetrieb die Möglichkeit besteht, eine "zweite" fehlerbehaftete Variable

$$Z_t = X_t^b = \begin{cases} X + W & \text{(additive Überlagerung)} \\ X \cdot W & \text{(multiplikative Überlagerung)} \end{cases}$$

Tabelle 12.14: Panel-IV-Schätzung bei multiplikativer Überlagerung

T	ρ	γ_1	$\zeta(\gamma, T)$	$\frac{T-2}{(T-1)^2} \sigma_u^2$	$E[C_a]$	$\text{plim} \hat{\beta}^{IV}$	estimate	std. dev.	min.	median	max.
4	-0.9000	-0.9000	-0.0788	0.0180	-0.8392	-2.4463	-2.4490	0.0608	-2.7567	-2.4474	-2.2496
4	-0.5000	-0.5000	0.0417	0.0272	-0.5688	-2.3806	-2.3809	0.0910	-2.6671	-2.3789	-2.0873
4	-0.1000	-0.1000	0.1799	0.0481	-0.3279	-2.1337	-2.1278	0.1640	-2.6954	-2.1262	-1.4209
4	0.0000	0.0000	0.2222	0.0575	-0.2798	-1.9858	-1.9834	0.2088	-2.7621	-1.9814	-1.1378
4	0.1000	0.1000	0.2690	0.0704	-0.2394	-1.7652	-1.7659	0.2519	-2.6495	-1.7654	-0.5878
4	0.5000	0.5000	0.5139	0.2215	-0.2354	-0.1475	-0.1262	0.3988	-1.2314	-0.1669	2.1384
4	0.9000	0.9000	0.8832	5.4677	-5.4509	0.0077	0.0076	0.0422	-0.1578	0.0080	0.1726
10	-0.9000	-0.9000	-0.0027	0.0080	-0.9053	-2.4779	-2.4779	0.0339	-2.5932	-2.4782	-2.3704
10	-0.5000	-0.5000	0.0302	0.0121	-0.5422	-2.4443	-2.4462	0.0598	-2.6641	-2.4442	-2.2705
10	-0.1000	-0.1000	0.0806	0.0214	-0.2020	-2.2356	-2.2410	0.1812	-2.9056	-2.2491	-1.6459
10	0.0000	0.0000	0.0988	0.0256	-0.1243	-1.9858	-1.9894	0.3316	-4.9401	-1.9936	-0.3355
10	0.1000	0.1000	0.1204	0.0313	-0.0517	-0.9873	-1.7931	46.7308	-1453.2898	-1.0454	494.3697
10	0.5000	0.5000	0.2717	0.0984	0.1298	-4.3954	-4.8263	2.3126	-44.9374	-4.3853	32.3270
10	0.9000	0.9000	0.7423	2.4301	-2.2724	0.1735	0.1897	0.1149	-0.1338	0.1750	1.1658
30	-0.9000	-0.9000	0.0012	0.0027	-0.9039	-2.4925	-2.4924	0.0198	-2.5551	-2.4922	-2.4304
30	-0.5000	-0.5000	0.0108	0.0041	-0.5149	-2.4802	-2.4794	0.0350	-2.5868	-2.4791	-2.3665
30	-0.1000	-0.1000	0.0272	0.0072	-0.1344	-2.3661	-2.3692	0.1625	-3.2011	-2.3658	-1.7710
30	0.0000	0.0000	0.0333	0.0086	-0.0419	-1.9858	-2.0684	8.6398	-335.0462	-1.9796	182.2935
30	0.1000	0.1000	0.0407	0.0105	0.0488	-3.0400	-3.0937	3.8929	-14.6369	-3.0224	140.4798
30	0.5000	0.5000	0.0975	0.0332	0.3693	-2.7246	-2.7343	0.1466	-3.3194	-2.7234	-2.2936
30	0.9000	0.9000	0.4501	0.8192	-0.3693	3.0459	5.0260	101.7905	-528.0309	2.3035	3102.4501
50	-0.9000	-0.9000	0.0009	0.0016	-0.9025	-2.4955	-2.4956	0.0148	-2.5633	-2.4956	-2.4404
50	-0.5000	-0.5000	0.0066	0.0024	-0.5090	-2.4880	-2.4882	0.0289	-2.5829	-2.4884	-2.3897
50	-0.1000	-0.1000	0.0163	0.0043	-0.1207	-2.4104	-2.4153	0.1367	-2.9812	-2.4146	-1.9109
50	0.0000	0.0000	0.0200	0.0052	-0.0252	-1.9858	-1.9123	3.4645	-44.0738	-1.9847	102.2631
50	0.1000	0.1000	0.0244	0.0063	0.0692	-2.7285	-2.7705	0.3316	-4.3304	-2.7436	-1.8029
50	0.5000	0.5000	0.0591	0.0199	0.4209	-2.6183	-2.6241	0.0944	-2.9452	-2.6211	-2.3591
50	0.9000	0.9000	0.3128	0.4919	0.0953	-15.4024	-233.9785	10292.5065	-460231.1005	-3.9894	2588.4932
Bemerkung: Alle Ergebnisse beziehen sich auf den wahren Parameter $\beta = -2.50$.											

mit additiver über multiplikativer Überlagerung zu erzeugen, die hoch mit der zuvor erzeugten Variablen X^a korreliert ist, wobei andererseits der Fehler W nicht mit dem Fehler U korreliert sein darf. In diesem Abschnitt sollen die Eigenschaften des betreffenden IV-Panelschätzers kurz untersucht werden.

Der IV-Schätzer lautet in diesem Fall

$$\begin{aligned}
\hat{\beta}_P^{ab,IV} &= \frac{\sum_{i=1}^N \sum_{t=2}^T \left(x_{i,t}^b - \left(\frac{1}{T} \sum_{s=1}^T x_{i,s}^b \right) \right) \left(y_{it} - \left(\frac{1}{T} \sum_{s=1}^T y_{is} \right) \right)}{\sum_{i=1}^N \sum_{t=1}^T \left(x_{i,t-1}^b - \left(\frac{1}{T} \sum_{s=1}^T x_{i,s}^b \right) \right) \left(x_{it}^a - \left(\frac{1}{T} \sum_{s=1}^T x_{i,s}^a \right) \right)} \\
&= \frac{\sum_{i=1}^N \sum_{t=1}^T \left(x_{i,t}^b - \bar{x}_{i\bullet}^b \right) \left(y_{it} - \bar{y}_{i\bullet} \right)}{\sum_{i=1}^N \sum_{t=1}^T \left(x_{i,t}^b - \bar{x}_{i\bullet}^b \right) \left(x_{it}^a - \bar{x}_{i\bullet}^a \right)}, \tag{12-27}
\end{aligned}$$

Für den Wahrscheinlichkeitsgrenzwert dieser Schätzfunktion (mit $N \rightarrow \infty$ sowie T fest)

ergibt sich

$$\begin{aligned}
\text{plim } \hat{\beta}_P^{ab,IV} &= \frac{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(y_{it} - \bar{y}_{i\bullet})}{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(x_{it} - \bar{x}_{i\bullet}^a)} \\
&= \frac{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(x_{it} - \bar{x}_{i\bullet})}{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(x_{it} - \bar{x}_{i\bullet}^a)} \beta \\
&\quad + \frac{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(\eta_{it} - \bar{\eta}_{i\bullet})}{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(x_{it} - \bar{x}_{i\bullet}^a)} \\
&= \frac{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(x_{it} - \bar{x}_{i\bullet})}{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t}^b - \bar{x}_{i\bullet}^b)(x_{it} - \bar{x}_{i\bullet}^a)} \beta \\
&= \frac{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t} - \bar{x}_{i\bullet})(x_{it} - \bar{x}_{i\bullet})}{\text{plim } \frac{1}{T} \sum_t \frac{1}{N} \sum_i (x_{i,t} - \bar{x}_{i\bullet})(x_{it} - \bar{x}_{i\bullet})} \beta \\
&= \frac{E[S^2]}{E[S^2]} \beta \\
&= \beta
\end{aligned} \tag{12-28}$$

mit S^2 aus (11.2.2). Dabei wurde, genau wie in Abschnitt ??, ausgenutzt, daß

$$\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i (X_{i,t-1}^b - \bar{X}_{i\bullet}^b)(X_{it} - \bar{X}_{i\bullet}) = \text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i (X_{i,t-1} - \bar{X}_{i\bullet})(X_{it} - \bar{X}_{i\bullet})$$

gilt. Entsprechend kann man verifizieren, daß bei Unkorreliertheit von U und W

$$\text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i (X_{i,t-1}^b - \bar{X}_{i\bullet}^b)(X_{it}^a - \bar{X}_{i\bullet}^a) = \text{plim } \frac{1}{T-1} \sum_t \frac{1}{N} \sum_i (X_{i,t-1} - \bar{X}_{i\bullet})(X_{it} - \bar{X}_{i\bullet})$$

Beide Resultate gemeinsam wurden beim Übergang von der viertletzten zur drittletzten Zeile verwendet, was zur Konsistenz des Schätzers führt.

Eigene Simulationsrechnungen haben allerdings gezeigt, daß die Konsistenz oftmals nicht gegeben ist. Auch hier spielt offensichtlich die Tatsache eine Rolle, daß der Nenner des Schätzers 12-27 für kleines T und bei Autokorrelation des Regressors nahe an den Wert Null herankommt. Zur Erinnerung sei nochmals darauf hingewiesen, daß in diesem Fall der Erwartungswert des Nenners durch

$$E[S^2] = \gamma_0 - \psi(\gamma, T)$$

(siehe (11-7)) gegeben ist, wobei im Fall des AR(1) Prozesses die Biasfunktion durch

$$\psi(\gamma, T) = \frac{\gamma_0}{T^2} \{T + 2(T-1)\rho^1 + 2(T-2)\rho^2 + \dots + 2 \cdot 2\rho^{T-2} + 2 \cdot 1\rho^{T-1}\} .$$

gegeben ist. Siehe (11-9).

Es soll nun untersucht werden, ob dieser Erwartungswert negativ bzw. gleich Null sein kann. Dazu ist zunächst zu bemerken, daß wegen der monotonen Beziehung des Erwartungswertes bezüglich ρ dieser Ausdruck für $\rho \rightarrow +1$ gegen sein Minimum strebt. Wir untersuchen deshalb $E[S^2]$ an der Stelle $\rho = 1$. Dann erhalten wir

$$E[S^2] \geq \gamma_0 \left\{ 1 - \frac{T+2(T-1)+2(T-2)+\dots+2(T-(T-2))+2(T-(T-1))}{T^2} \right\}$$

Für den Ausdruck im Zähler des Bruchs in den geschweiften Klammern ergibt sich

$$\begin{aligned} T + (T-1)T - \sum_{i=1}^{T-1} &= \begin{cases} T + (T-1)T - \frac{T-1}{2} [1 + (T-1)] & \text{falls } T-1 \text{ gerade} \\ T + (T-1)T - \left(\frac{T-2}{2} [1 + (T-1)] + \frac{T}{2}\right) & \text{falls } T-1 \text{ ungerade} \end{cases} \\ &= T^2 \end{aligned}$$

Demnach erhalten wir

$$E[S^2] \geq \gamma_0 \left\{ 1 - \frac{T^2}{T^2} \right\} = 0 \quad ,$$

womit gezeigt ist, daß dieser Erwartungswert (zumindest im Fall des AR(1)-Prozesses) niemals negativ sein wird. Gleichwohl kann die Nenner-Zufallsvariable

$$\sum_{i=1}^N \sum_{t=1}^T \left(X_{i,t-1}^b - \left(\frac{1}{T} \sum_{s=1}^T X_{i,s}^b \right) \right) \left(X_{it}^a - \left(\frac{1}{T} \sum_{s=1}^T X_{i,s}^a \right) \right)$$

des Schätzers $\hat{\beta}_P^{ab,IV}$ in (12-27) natürlich deutlich von diesem Erwartungswert abweichen, so daß vor allem bei starker Überlagerung und großem positivem ϱ mit Problemen bei diesem Schätzer zu rechnen ist. Dabei ist auch noch zu unterscheiden, ob die Überlagerung in additiver oder in multiplikativer Form stattfindet.¹⁰³

¹⁰³Dazu sei auch auf die Untersuchungen von Elena Biewen im Rahmen des FAWE-Projektes verwiesen.
Noch genauer zitieren

13 Literatur

- Anderson, T.W: (1971). The Statistical Analysis of Time Series. Wiley: New York.
- Biørn, Erik(1996) "Panel data with measurement errors". in: Matyas, L. and P. Sevestre: The Econometrics of Panel Data: A Handbook of the Theory with Applications, Kluwer, Boston. Zweite revidierte Auflage, 236-279.
- Biørn, Erik, und Krishnakumar, J. (2008)."Measurement Errors and Simultaneity". in: Matyas, L. and P. Sevestre: The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice. Springer, New York. Dritte Auflage, 323-367.
- Bound, J., D.A. Jaeger und R.M. Baker (1995). Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak. Journal of the American Statistical Association 90, 443-450.
- Box, G.E.P, G.M. Jenkins und G.C. Reinsel (1994). Time Series Analysis. Forecasting and Control. Prentice Hall, Englewood Cliffs (NJ).Dritte Auflage.
- Cheang, W.-K., und G.C. Reinsel (2000). Bias Reduction of Autoregressive Estimates in Time Series Regression Model Through Restricted Maximum Likelihood. Journal of the American Statistical Association 96, 1173-1184.
- Dunson, D.B., und C. R. Weinberg (2000). Modeling Human Fertility in the Presence of Measurement Error. Biometrics 56, 288-292.
- Evans,T., L. Zayatz und J. Slanta (1998). Using Noise for disclosure limitation of establishment tabular data. Journal of Official Statistics 14, 537-551.
- Hamilton, J.D. (1994). Time series analysis . Princeton, NJ : Princeton Univ. Press.
- Hwang, J. (1986). Multiplicative errors-in-variables models with applications to recent data released by the U.S. Department of Energy. Journal of the American Statistical Association 81, 680-688.
- Kim, J.J., und W. E. Winkler : "Masking Microdata Files." *American Statistical Association. Proceedings of the Section on Survey Research Methods* 1995 , 114-119.
- Lechner, S., und W Pohlmeier (2003). Schätzung ökonomischer Modelle auf der Grundlage anonymisierter Daten, in: Gnoss, R. and G. Ronning (Herausgeber:) Anonymisierung wirtschaftsstatistischer Einzelndaten, Forum der Bundesstatistik, Band 42, Wiesbaden.
- Lin, A.(1989). Estimation of multiplicative Measurement error models and some simulation results. Economics letters 31, 13-20
- Marshall,A.W., und I. Olkin(1979). Inequalities.Theory of Majorization and Its Application. Academic Press: New York.
- Massell,P., L. Zayatz und J. Funk (2006). Protecting the confidentiality of survey tabular data by adding noise to the underlying micro data: Application to the Commodity Flow Survey. In: J. Domingo und L. Franconi (Herausgeber): Privacy in Statistical Data Bases. CENEX-SDC Project International Conference, PSD 2006, Rome, Italy, December 2006, Proceedings, Springer, Berlin, 304-317.

- Press, S.J. (1972). *Applied Multivariate Analysis*. Holt, Rhinehart & Winston: New York.
- G. McLachlan und D. Peel: *Finite Mixture Models*. Wiley, New York, 2000.
- Ronning, G.(1977). Pseudo-Bayessche Schätzer in der Ökonometrie. *Operations Research Verfahren 26* , 861-871.
- Ronning, G.(2005) *Statistische Methoden in der empirischen Wirtschaftsforschung*. Münster : Lit-Verlag.
- Ronning, G. et al (2005). Handbuch zur Anonymisierung wirtschaftsstatistischer Mikrodaten. Statistisches Bundesamt, Wiesbaden , Reihe "Statistik und Wissenschaft", Band 4, 2005, (gemeinsam mit Roland Sturm, Jörg Höhne, Rainer Lenz, Martin Rosemann, Michael Scheffler und Daniel Vorgrimler)
- Rosemann, M.(2006). Auswirkungen datenverändernder Anonymisierungsverfahren auf die Analyse von Mikrodaten. IAW Forschungsbericht Nr. 66, Tübingen.
- Gina Marie Roque. Masking Microdata Files with Mixtures of Multivariate Normal Distributions. Dissertation, June 2000, University of California, Riverside.
- Springer, M.D. (1979). *The Algebra of Random Variables*. Wiley: New York.
- Staiger, D., und J.H. Stock (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica* 65, 557-586.
- Stata Corporation (2003). *Stata Statistical Software*. Release 8.0. "Cross-sectional Time-Series".
- Staudenmayer, J., und John Buonaccorsi. (2005). Measurement Error in Linear Autoregressive Models. *Journal of the American Statistical Association* . 100, 841-852.
- Stock, J.H., J.H. Wright und M. Yogo (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics* 20, 518-529.
- Wooldridge, J.M. 2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press: Cambridge (MA).
- . Yancey,E., W. E. Winkler und R.H. Creecy: "Disclosure Risk Assessment in Perturbative Micro Data Protection." in: J. Domingo-Ferrer (Herausgeber): *Inference Control in Statistical Databases*. Springer: Berlin, 2002, 135-152.

A Einige nützliche Matrizen-Resultate

Bei der Analyse der multiplikativen Überlagerung spielt das Hadamard-Produkt eine wichtige Rolle. Ich liste deshalb im folgenden wichtige Eigenschaften auf, die auch Rosemann (2006) bereits verwendet hat.

Für "Hadamard-Produkt von zwei $(m \times n)$ -Matrizen \mathbf{A} und \mathbf{B} gilt

$$\mathbf{A} \odot \mathbf{B} = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1,n-1}b_{1,n-1} & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \dots & a_{2,n-1}b_{2,n-1} & a_{2n}b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m-1,1}b_{m-1,1} & a_{m-1,2}b_{m-1,2} & \dots & a_{m-1,n-1}b_{m-1,n-1} & a_{m-1,n}b_{m-1,n} \\ a_{m,1}b_{m,1} & a_{m,2}b_{m,2} & \dots & a_{m,n-1}b_{m,n-1} & a_{m,n}b_{m,n} \end{pmatrix}, \quad (\text{A-1})$$

d.h. die resultierende Matrix hat ebenfalls m Zeilen und n Spalten. Mit \div wird die entsprechende elementweise **Division** bezeichnet. Insbesondere gilt für beliebige Rechtecksmatrix

$$\mathbf{A} \div \mathbf{A} = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} = \boldsymbol{\iota} \boldsymbol{\iota}' \quad \text{und} \quad \mathbf{A} \odot \boldsymbol{\iota} \boldsymbol{\iota}' = \mathbf{A}, \quad (\text{A-2})$$

und für beliebigen Vektor \mathbf{a} gilt

$$\mathbf{a} \div \mathbf{a} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} = \boldsymbol{\iota} \quad \text{und} \quad \mathbf{a} \odot \boldsymbol{\iota} = \mathbf{a}. \quad (\text{A-3})$$

Unmittelbar aus (A-1) folgt¹⁰⁴

$$\text{vec}(\mathbf{A} \odot \mathbf{B}) = \text{vec}(\mathbf{A}) \odot \text{vec}(\mathbf{B}) \quad (\text{A-4})$$

Für m -dimensionale Vektoren \mathbf{a} , \mathbf{c} und n -dimensionale Vektoren \mathbf{b} , \mathbf{d} sind manchmal folgende Ergebnisse nützlich:

$$\mathbf{a} \odot \mathbf{b} = \mathbf{D}_a \mathbf{b} = \mathbf{D}_b \mathbf{a} \quad (\text{A-5})$$

Dabei sind \mathbf{D}_a und \mathbf{D}_b Diagonalmatrizen mit den Elementen des jeweiligen Vektors auf der Diagonalen. Ferner gilt

$$(\mathbf{a} \odot \mathbf{b})(\mathbf{c} \odot \mathbf{d})' = \mathbf{a} \mathbf{c}' \odot \mathbf{b} \mathbf{d}' \quad (\text{A-6})$$

In der Arbeit von Lin(1989) ist folgendes Ergebnis von besonderer Bedeutung, das nur indirekt mit dem Hadamard-Produkt zu tun hat: Für eine beliebige $(m \times n)$ -Matrix \mathbf{A} gilt

$$\mathbf{A}' \mathbf{A} = \sum_{j=1}^n [(\mathbf{I}_n \otimes \mathbf{e}_j) \text{vec}(\mathbf{A}')] [(\mathbf{I}_n \otimes \mathbf{e}_j) \text{vec}(\mathbf{A})]' \quad (\text{A-7})$$

¹⁰⁴Der vec -Operator stapelt die Spalten einer Matrix übereinander. $\text{vec}(\mathbf{A})$ ergibt einen mn -dimensionalen Vektor.

Dabei ist \mathbf{e}_j ein n -dimensionaler Vektor und bezeichnet die j -te Spalte der Einheitsmatrix

Beweis: Zunächst schreiben wir die Matrix \mathbf{A} wie folgt:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \vdots \\ \mathbf{a}'_{m-1} \\ \mathbf{a}'_m \end{pmatrix} \quad \text{mit} \quad \mathbf{a}'_j = (a_{j1}, \dots, a_{jn})$$

und damit

$$\mathbf{A}' \mathbf{A} = \sum_{j=1}^m \mathbf{a}_j \mathbf{a}'_j$$

Ferner gilt

$$\mathbf{A}' = (\mathbf{a}_1, \mathbf{a}_2 \dots \mathbf{a}_{m-1}, \mathbf{a}_m) \quad \text{und} \quad \text{vec}(\mathbf{A}') = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \vdots \\ \mathbf{a}_m \end{pmatrix}$$

sowie

$$(\mathbf{I}_n \otimes \mathbf{e}'_j) \text{vec}(\mathbf{A}') = (\mathbf{0}, \dots, \mathbf{I}_n \dots, \mathbf{0}) \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_m \end{pmatrix} = \mathbf{a}_j \quad .$$

Daraus folgt (A-7).

Für die lineare Form $\mathbf{A} \mathbf{x}$ gilt folgendes Ergebnis: Es sei \mathbf{A} eine $(m \times n)$ -Matrix und \mathbf{x} ein n -dimensionaler Vektor. Dann läßt sich die Linearform wie folgt darstellen:

$$\mathbf{A} \mathbf{x} = (\mathbf{I}_m \otimes \mathbf{x}') \text{vec}(\mathbf{A}') \quad (\text{A-8})$$

Beweis: Zunächst gilt

$$\mathbf{A} \mathbf{x} = \begin{pmatrix} \mathbf{a}'_1 \mathbf{x} \\ \vdots \\ \mathbf{a}'_m \mathbf{x} \end{pmatrix}$$

$$\mathbf{I}_m \otimes \mathbf{x}' = \begin{pmatrix} \mathbf{x}' & & & & \\ & \mathbf{x}' & & & \\ & & \ddots & & \\ & & & \mathbf{x}' & \\ & & & & \mathbf{x}' \end{pmatrix}$$

und damit

$$(\mathbf{I}_m \otimes \mathbf{x}') \text{vec}(\mathbf{A}') = \begin{pmatrix} \mathbf{x}' \mathbf{a}_1 \\ \vdots \\ \mathbf{x}' \mathbf{a}_m \end{pmatrix} ,$$

womit (A-8) bewiesen ist.

B Beweis von Resultaten

Die folgenden Resultate werden in Abschnitt 10.11.2 verwendet.

B.1 Drei Resultate

Wir beweisen im Zusammenhang mit (10-91) und (10-92) in Abschnitt 10.11.2 folgendes Resultat:

Es seien X_i, Y_i, U_i, V_i , $i = 1, \dots, n$, vier Stichprobenvariablen mit $E[X_i] = \mu_x$, $\text{var}[X_i] = \sigma_x^2$, $E[Y_i] = \mu_y$, $E[U_i] = \mu_u$, $E[V_i] = \mu_v$, wobei (X_i, Y_j) für alle i, j, k, ℓ stochastisch unabhängig von (U_j, V_ℓ) ist und außerdem bezüglich der Kovarianzen

$$\text{cov}[X_i, Y_j] = \begin{cases} \sigma_{xy} & i = j \\ 0 & i \neq j \end{cases}, \quad \text{cov}[U_i, V_j] = \begin{cases} \sigma_{uv} & i = j \\ 0 & i \neq j \end{cases}$$

gilt. Ferner gelte $\bar{V} = \frac{1}{n} \sum_{j=1}^n V_j$ sowie $\overline{XU} = \frac{1}{n} \sum_{j=1}^n X_j U_j$. Dann erhalten wir

$$E[(Y_i - \mu_y)(X_i U_i - \overline{XU})(V_i - \bar{V})] = \left(1 - \frac{1}{n}\right)^2 \sigma_{xy} \sigma_{uv} \quad (\text{B-1})$$

$$E[(X_i U_i - \overline{XU})(V_i - \bar{V})] = \left(1 - \frac{1}{n}\right) \mu_x \sigma_{uv} \quad (\text{B-2})$$

$$E[(X_i U_i - \overline{XU})(X_i V_i - \overline{XV})] = \left(1 - \frac{1}{n}\right) (\sigma_x^2 \mu_u \mu_v + \sigma_{uv}(\sigma_x^2 + \mu_x^2)) \quad (\text{B-3})$$

Um die Ergebnisse (10-91) bis (10-93) in Abschnitt 10.11.2 zu beweisen, ist bei obigen Ergebnissen der Index i durch den Index t sowie n durch T zu ersetzen. Ferner tritt die Zufallsvariable τ mit Erwartungswert 0 an die Stelle der Zufallsvariablen Y mit Erwartungswert μ_y . Schließlich ist zu beachten, daß die Zufallsvariablen U und V jeweils Erwartungswert 1 besitzen, also gilt $\mu_u = \mu_v = 1$.

B.2 Hilfsresultate

$$E[X_i Y_i U_i V_i] = E[X_i Y_i] E[U_i V_i] \quad (\text{B-4})$$

Beweis: Dies folgt unmittelbar aus der angenommenen stochastischen Unabhängigkeit. ■

$$E[U_i \bar{V}] = \frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \quad (\text{B-5})$$

Beweis:

$$\begin{aligned} E[U_i \bar{V}] &= \frac{1}{n} \sum_j E[U_i V_j] \\ &= \frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \quad \blacksquare \\ E[\bar{U} \bar{V}] &= \frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \quad (\text{B-6}) \end{aligned}$$

Beweis:

$$\begin{aligned} E[\bar{U} \bar{V}] &= \frac{1}{n^2} \sum_i \sum_j E[U_i V_j] \\ &= \frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \quad \blacksquare \end{aligned}$$

$$E[U_i X_i \bar{V}] = \frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_x \mu_u \mu_v \quad (\text{B-7})$$

Beweis:

$$\begin{aligned} E[U_i X_i \bar{V}] &= \frac{1}{n} \sum_j E[U_i X_i V_j] \\ &= \frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_x \mu_u \mu_v \blacksquare \end{aligned}$$

$$E[\overline{UX} \bar{V}] = \frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_x \mu_u \mu_v \quad (\text{B-8})$$

Beweis:

$$\begin{aligned} E[\overline{UX} \bar{V}] &= \frac{1}{n^2} \sum_j \sum_k E[U_j X_j V_k] \\ &= \frac{1}{n^2} E[U_i V_i] + \frac{(n-1)n}{n^2} \mu_x \mu_u \mu_v \blacksquare \end{aligned}$$

$$E[V_i \overline{UX}] = \frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_x \mu_u \mu_v \quad (\text{B-9})$$

Beweis:

$$\begin{aligned} E[V_i \overline{UX}] &= \frac{1}{n} \sum_j E[V_i U_j X_j] \\ &= \frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_x \mu_u \mu_v \blacksquare \end{aligned}$$

$$E[V_i X_i \overline{UX}] = \frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x^2 \quad (\text{B-10})$$

Beweis:

$$\begin{aligned} E[V_i X_i \overline{UX}] &= \frac{1}{n} \sum_j E[V_i X_i U_j X_j] \\ &= \frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x^2 \blacksquare \end{aligned}$$

$$E[U_i X_i \overline{VX}] = \frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x \quad (\text{B-11})$$

Beweis:

$$\begin{aligned} E[U_i X_i \overline{VX}] &= \frac{1}{n} \sum_j E[U_i X_i V_j X_j] \\ &= \frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x^2 \blacksquare \end{aligned}$$

$$E[\overline{UX} \overline{VX}] = \frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x \quad (\text{B-12})$$

Beweis:

$$\begin{aligned} E[\overline{UX} \overline{VX}] &= \frac{1}{n^2} \sum_i \sum_j E[U_i X_i V_j X_j] \\ &= \frac{1}{n^2} E[V_i X_i] E[X_i^2] + \frac{(n-1)n^2}{n^2} \mu_u \mu_v \mu_x^2 \blacksquare \end{aligned}$$

$$E[Y_i V_i \overline{UX}] = \frac{1}{n} E[X_i Y_i] E[U_i V_i] + \frac{n-1}{n} \mu_x \mu_y \mu_u \mu_v \quad (\text{B-13})$$

Beweis:

$$\begin{aligned} E[Y_i V_i \overline{UX}] &= \frac{1}{n} \sum_j E[Y_i V_i U_j X_j] \\ &= \frac{1}{n} \sum_j E[Y_i X_j] E[V_i U_j] \\ &= \frac{1}{n} E[X_i Y_i] E[U_i V_i] + \frac{n-1}{n} \mu_x \mu_y \mu_u \mu_v \blacksquare \end{aligned}$$

$$\begin{aligned} E[Y_i \overline{UX} \bar{V}] &= \frac{1}{n^2} E[X_i Y_i] E[U_i V_i] + \frac{n-1}{n^2} E[U_i V_i] \mu_x \mu_y \\ &\quad + \frac{n-1}{n^2} E[X_i Y_i] \mu_u \mu_v + \left(\frac{n-1}{n}\right)^2 \mu_x \mu_y \mu_u \mu_v \quad (\text{B-14}) \end{aligned}$$

Beweis: Im Folgenden sind **für festes** i folgende Fälle zu unterscheiden, zu denen jeweils die Häufigkeit hinzugefügt wird:

Fall	Häufigkeit
$i = j = k$	1
$j = k \neq i$	$n - 1$
$i = j \neq k$	$n - 1$
$i \neq j \neq k$	$(n - 1)^2 - (n - 1) = (n - 1)(n - 2)$
Summe:	n^2

$$\begin{aligned}
 E[Y_i \overline{UX} \overline{V}] &= \frac{1}{n^2} \sum_j \sum_k E[Y_i U_j X_j V_k] \\
 &= \frac{1}{n^2} E[X_i Y_i] E[U_i V_i] \\
 &\quad + \frac{n-1}{n^2} E[U_i V_i] \mu_x \mu_y \\
 &\quad + \frac{n-1}{n^2} E[X_i Y_i] \mu_u \mu_v \\
 &\quad + \frac{n-1}{n^2} \mu_x \mu_y \mu_u \mu_v \\
 &\quad + \frac{(n-1)(n-2)}{n^2} \mu_x \mu_y \mu_u \mu_v \blacksquare
 \end{aligned}$$

B.3 Beweise

Unter Verwendung der Hilfsresultate (B-5) bis (B-13) aus Abschnitt B.2 beweisen wir nun die Resultate aus Abschnitt B.1.

Beweis von (B-1)

$$\begin{aligned}
E[(Y_i - \mu_y)(U_i X_i - \overline{UX})(V_i - \overline{V})] &= E[U_i V_i X_i Y_i] - E[U_i X_i Y_i \overline{V}] \\
&\quad - E[V_i Y_i \overline{UX}] + E[Y_i \overline{UX} \overline{V}] \\
&\quad - E[U_i V_i X_i] \mu_y + E[U_i X_i \overline{V}] \mu_y \\
&\quad + E[V_i \overline{UX}] \mu_y - E[\overline{UX} \overline{V}] \mu_y \\
&= E[U_i V_i] E[X_i Y_i] - E[U_i \overline{V}] E[X_i Y_i] \\
&\quad - E[V_i Y_i \overline{UX}] + E[Y_i \overline{UX} \overline{V}] \\
&\quad - E[U_i V_i] \mu_x \mu_y + E[U_i \overline{V}] \mu_x \mu_y \\
&\quad + E[V_i \overline{UX}] \mu_y - E[\overline{UX} \overline{V}] \mu_y \\
&= E[U_i V_i] E[X_i Y_i] \\
&\quad - \left(\frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \right) E[X_i Y_i] \\
&\quad - \left(\frac{1}{n} E[X_i Y_i] E[U_i V_i] + \frac{n-1}{n} \mu_x \mu_y \mu_u \mu_v \right) \\
&\quad + \left(\frac{1}{n^2} E[X_i Y_i] E[U_i V_i] + \frac{n-1}{n^2} E[U_i V_i] \mu_x \mu_y \right. \\
&\quad \quad \left. + \frac{n-1}{n^2} E[X_i Y_i] \mu_u \mu_v + \left(\frac{n-1}{n} \right)^2 \mu_x \mu_y \mu_u \mu_v \right) \\
&\quad - E[U_i V_i] \mu_x \mu_y \\
&\quad + \left(\frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \right) \mu_x \mu_y \\
&\quad + \left(\frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_u \mu_v \mu_x \right) \mu_y \\
&\quad - \left(\frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_u \mu_v \mu_x \right) \mu_y \\
&= E[U_i V_i] E[X_i Y_i] \\
&\quad - \left(\frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \right) E[X_i Y_i] \\
&\quad - \left(\frac{1}{n} E[X_i Y_i] E[U_i V_i] + \frac{n-1}{n} \mu_x \mu_y \mu_u \mu_v \right) \\
&\quad + \left(\frac{1}{n^2} E[X_i Y_i] E[U_i V_i] + \frac{n-1}{n^2} E[U_i V_i] \mu_x \mu_y \right. \\
&\quad \quad \left. + \frac{n-1}{n^2} E[X_i Y_i] \mu_u \mu_v + \left(\frac{n-1}{n} \right)^2 \mu_x \mu_y \mu_u \mu_v \right) \\
&\quad - E[U_i V_i] \mu_x \mu_y \\
&\quad + \left(\frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \right) \mu_x \mu_y \\
&= \left(1 - \frac{2}{n} + \frac{1}{n^2} \right) E[U_i V_i] E[X_i Y_i] \\
&\quad + \left(\frac{n-1}{n^2} - \frac{n-1}{n} \right) E[X_i Y_i] \mu_u \mu_v \\
&\quad + \left(\frac{1}{n} + \frac{n-1}{n^2} - 1 \right) E[U_i V_i] \mu_x \mu_y \\
&\quad + \left(\frac{(n-1)^2}{n^2} \right) \mu_u \mu_v \mu_x \mu_y \\
&= \left(1 - \frac{1}{n} \right)^2 E[U_i V_i] E[X_i Y_i] \\
&\quad - \left(\frac{(n-1)^2}{n^2} \right) E[X_i Y_i] \mu_u \mu_v \\
&\quad - \left(\frac{(n-1)^2}{n^2} \right) E[U_i V_i] \mu_x \mu_y \\
&\quad + \left(\frac{(n-1)^2}{n^2} \right) \mu_u \mu_v \mu_x \mu_y \\
&= \left(\frac{(n-1)^2}{n^2} \right) (E[U_i V_i] - \mu_u \mu_v) (E[X_i Y_i] - \mu_x \mu_y) \\
&= \left(\frac{(n-1)^2}{n^2} \right) \sigma_{uv} \sigma_{xy} \blacksquare
\end{aligned}$$

Beweis von (B-2)

$$\begin{aligned}
E[(U_i X_i - \overline{UX})(V_i - \overline{V})] &= E[U_i V_i X_i] - E[V_i \overline{UX}] \\
&\quad - E[U_i X_i \overline{V}] + E[\overline{UX} \overline{V}] \\
&= E[U_i V_i] \mu_x - E[V_i \overline{UX}] \\
&\quad - E[U_i X_i \overline{V}] + E[\overline{UX} \overline{V}] \\
&= E[U_i V_i] \mu_x - \left(\frac{1}{n} E[U_i V_i] + \frac{n-1}{n} \mu_u \mu_v \right) \\
&\quad - \left(\frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_x \mu_u \mu_v \right) \\
&\quad + \left(\frac{1}{n} E[U_i V_i] \mu_x + \frac{n-1}{n} \mu_x \mu_u \mu_v \right) \\
&= \left(1 - \frac{1}{n} \right) \mu_x \sigma_{uv} \blacksquare
\end{aligned}$$

Beweis von (B-3)

$$\begin{aligned}
E[(U_i X_i, -\overline{UX})(V_i X_i - \overline{VX})] &= E[U_i V_i X_i^2] - E[U_i X_i \overline{VX}] \\
&\quad - E[V_i X_i \overline{UX}] + E[\overline{UX} \overline{VX}] \\
&= E[V_i X_i] E[X_i^2] \\
&\quad - \left(\frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x^2 \right) \\
&\quad - \left(\frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x^2 \right) \\
&\quad + \left(\frac{1}{n} E[U_i V_i] E[X_i^2] + \frac{n-1}{n} \mu_u \mu_v \mu_x^2 \right) \\
&= \left(1 - \frac{1}{n} \right) \{ E[U_i V_i] E[X_i^2] - \mu_u \mu_v \mu_x^2 \} \blacksquare
\end{aligned}$$

C Ergebnisse für für symmetrischen IV-Schätzer**C.1 Einleitung**

In Abschnitt 12.2 wird darauf hingewiesen, daß bei Verwendung der verzögerten Regressor-Variablen als Instrument zwei verschiedene Formen des Schätzers existieren: die symmetrische und die asymmetrische Form. Siehe insbesondere (12-3) und (12-5). Da die asymmetrische Form der Logik des KQ-Schätzers entspricht, wird im Abschnitt 11 nur dieser betrachtet. In diesem Anhang werden nun die Ergebnisse für den "symmetrischen" Schätzer (12-3), d.h.

$$\begin{aligned}
\hat{\beta}^{a,IV} &= \frac{\sum_{t=2}^T (x_{t-1}^a - \overline{x^a})(y_t - \overline{y})}{\sum_{t=2}^T (x_{t-1}^a - \overline{x^a})(x_t^a - \overline{x^a})} \quad \text{"symmetrischer IV-Schätzer"} \\
&= \frac{\sum_{t=2}^T (x_{t-1}^a - \overline{x^a})(x_t - \overline{x})}{\sum_{t=2}^T (x_{t-1}^a - \overline{x^a})(x_t^a - \overline{x^a})} \beta + \frac{\sum_{t=2}^T (x_{t-1}^a - \overline{x^a})(\eta_t - \overline{\eta})}{\sum_{t=2}^T (x_{t-1}^a - \overline{x^a})(x_t^a - \overline{x^a})}
\end{aligned}$$

nachgeliefert.

C.2 Die Schätzfunktion C bei kleinem T

Ähnlich dem Vorgehen in Abschnitt 11.2.2 schreiben wir die Schätzfunktion als¹⁰⁵

$$\begin{aligned}
 C &= \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu + \mu - \bar{X})(X_{t-1} - \mu + \mu - \bar{X}) \\
 &= \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(X_{t-1} - \mu) + \frac{1}{T-1} \sum_{t=2}^T (\bar{X} - \mu)^2 \\
 &\quad - \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(\bar{X} - \mu) - \frac{1}{T-1} \sum_{t=2}^T (X_{t-1} - \mu)(\bar{X} - \mu) \\
 &= \frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(X_{t-1} - \mu) + \frac{1}{T-1} \left(\frac{1}{T}\right)^2 \sum_{t=2}^T \left(\sum_{s=1}^T \sum_{s'=1}^T (X_s - \mu)(X_{s'} - \mu)\right) \\
 &\quad - \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_t - \mu)(X_s - \mu) - \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_{t-1} - \mu)(X_s - \mu)
 \end{aligned}$$

Erwartungswertbildung der vier Summanden ergibt für den ersten Summanden

$$E\left[\frac{1}{T-1} \sum_{t=2}^T (X_t - \mu)(X_{t-1} - \mu)\right] = \gamma_1$$

für den zweiten Summanden (siehe dazu auch Abschnitt 11.2.2)

$$\begin{aligned}
 &E\left[\frac{1}{T-1} \left(\frac{1}{T}\right)^2 \sum_{t=2}^T \left(\sum_{s=1}^T \sum_{s'=1}^T (X_s - \mu)(X_{s'} - \mu)\right)\right] \\
 &= \left(\frac{1}{T}\right)^2 [T\gamma_0 + 2\{(T-1)\gamma_1 + (T-2)\gamma_2 + \dots + \cdot 3\gamma_{T-3} + \cdot 2\gamma_{T-2} + \gamma_{T-1}\}]
 \end{aligned}$$

und für die beiden restlichen Summanden¹⁰⁶

$$\begin{aligned}
 &E\left[\frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_t - \mu)(X_s - \mu)\right] \\
 &= \frac{1}{T-1} \frac{1}{T} [(T-1)\gamma_0 + (2T-3)\gamma_1 + (2T-5)\gamma_2 + \dots + 3\gamma_{T-2} + 1\gamma_{T-1}]
 \end{aligned}$$

und identisch dazu

$$\begin{aligned}
 &E\left[\frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_{t-1} - \mu)(X_s - \mu)\right] \\
 &= \frac{1}{T-1} \frac{1}{T} [(T-1)\gamma_0 + (2T-3)\gamma_1 + (2T-5)\gamma_2 + \dots + 3\gamma_{T-2} + 1\gamma_{T-1}]
 \end{aligned}$$

¹⁰⁵Man beachte, daß dasselbe Symbol C wie in der asymmetrischen Form benutzt wird.

¹⁰⁶Die Ergebnisse lassen sich am leichtesten verifizieren, wenn man (hier beispielsweise für den dritten Summanden) die folgende Tabelle bildet:

		t							
		2	3	4	5	...	$T-2$	$T-1$	T
s	1	γ_1	γ_2	γ_3	γ_4	...	γ_{T-3}	γ_{T-2}	γ_{T-1}
	2	γ_0	γ_1	γ_2	γ_3	...	γ_{T-4}	γ_{T-3}	γ_{T-2}
	3	γ_1	γ_0	γ_1	γ_2	...	γ_{T-5}	γ_{T-4}	γ_{T-3}
	⋮					⋱			
	$T-3$	γ_{T-5}	γ_{T-6}	γ_{T-7}	γ_{T-8}	...	γ_1	γ_2	γ_3
	$T-2$	γ_{T-4}	γ_{T-5}	γ_{T-6}	γ_{T-7}	...	γ_0	γ_1	γ_2
	$T-1$	γ_{T-3}	γ_{T-4}	γ_{T-5}	γ_{T-6}	...	γ_1	γ_0	γ_1
T	γ_{T-2}	γ_{T-3}	γ_{T-4}	γ_{T-5}	...	γ_2	γ_1	γ_0	

Die Tabelle zeigt, daß beispielsweise γ_0 insgesamt $(T-1)$ -mal auftaucht, daß γ_1 $(T-2)$ -mal unterhalb der Diagonalen und $(T-1)$ -mal oberhalb der Diagonalen auftaucht, schließlich daß γ_{T-3} zwei und drei mal, insgesamt also fünf mal in der Tabelle vorhanden ist. Entsprechende Tabellen lassen sich bezüglich der anderen Summanden bilden.

Damit erhalten wir für den Erwartungswert der Schätzfunktion C folgenden Ausdruck:

$$\begin{aligned} E[C] &= \gamma_1 + \left(\frac{1}{T}\right)^2 [T\gamma_0 + 2\{(T-1)\gamma_1 + (T-2)\gamma_2 + \dots + \cdot 3\gamma_{T-3} + \cdot 2\gamma_{T-2} + \gamma_{T-1}\}] \\ &\quad - \frac{2}{T(T-1)} [(T-1)\gamma_0 + (2T-3)\gamma_1 + (2T-5)\gamma_2 + \dots + 5\gamma_{T-3} + 3\gamma_{T-2} + 1\gamma_{T-1}] \\ &= \gamma_1 + \xi(\gamma, T) \end{aligned} \quad (\text{C-15})$$

mit

$$\begin{aligned} \xi(\gamma, T) &= \left(\frac{1}{T}\right)^2 [T\gamma_0 + 2\{(T-1)\gamma_1 + (T-2)\gamma_2 + \dots + \cdot 3\gamma_{T-3} + \cdot 2\gamma_{T-2} + \gamma_{T-1}\}] \\ &\quad - \frac{2}{T(T-1)} [(T-1)\gamma_0 + (2T-3)\gamma_1 + (2T-5)\gamma_2 + \dots + 5\gamma_{T-3} + 3\gamma_{T-2} + 1\gamma_{T-1}] \\ &= \frac{2}{T} \sum_{k=1}^{T-1} \left(\frac{T-k}{T} - \frac{2T-(2k+1)}{T-1} \right) \gamma_k - \frac{\gamma_0}{T} . \end{aligned} \quad (\text{C-16})$$

Eine Aussage über die Richtung der Verzerrung für kleines T läßt sich mit Hilfe des folgenden Ergebnisses treffen: Für jedes $k = 1, 2, \dots, T-1$, ergibt sich für den Koeffizienten von γ_k

$$\frac{2}{T} \left(\frac{T-k}{T} - \frac{2T-(2k+1)}{T-1} \right) = \frac{2}{T} \frac{k(T+1) - T^2}{T(T-1)} < 0 .$$

Dabei folgt die Ungleichung aus der Tatsache, daß der Zähler des zweiten Bruchs monoton in k wächst, jedoch für den maximalen Wert $k = T-1$ negativ ist. Daraus folgt:

$$E[C] < \gamma_1 \quad \text{falls } \gamma_k > 0 \text{ für } k = 1, 2, \dots, T-1, \quad (\text{C-17})$$

d.h. die Autokovarianz γ_1 wird durch C unterschätzt, falls alle Autokorrelationen positiv sind.

Speziell für den Fall, daß keine Autokorrelation vorliegt ($\gamma_\tau = 0$ für $\tau > 0$) ergibt sich demnach

$$E[C] = -\frac{\gamma_0}{T} < 0$$

und demnach wird der wahre Autokorrelationskoeffizient γ_1 ($= 0$) für endliches (und insbesondere für "kleines") T unterschätzt.

Die Tabelle C.1 zeigt ausgewählte Werte für den Bias, der durch die Funktion $\xi(\gamma, T)$ erzeugt wird. Dabei zeigt diese Tabelle die mit -1 multiplizierten Werte. Man erkennt, daß nur für extrem großes negatives ρ die Verzerrung positiv ist. Überdies sind die Verzerrungen für negatives ρ deutlich geringer.

Ferner wird aus der Tabelle deutlich, daß die durch C zu schätzende Autokovarianz γ_1 vor allem für kleines T sehr große Verzerrungen aufweist: Beispielsweise lautet für $T = 5$ und $\rho = 0.5$ der wahre Wert $\gamma_1 = \gamma_0\rho = 0.50$. Die ausgewiesene Verzerrung in der Tabelle zeigt dann, daß sich für den Erwartungswert

$$E[C] = 0.50 - 0.47375 = 0.02625$$

ergibt!

Tabelle C.1: Werte der (negativen) Biasfunktion $(-1) \cdot \xi(\gamma, T)$ für AR(1)-Prozeß.

ρ	Anzahl Wellen (T)							
	3	4	5	10	30	50	100	500
-0.90000	-0.07667	0.00962	-0.02034	0.00278	0.00128	0.00087	0.00048	0.00010
-0.50000	0.08333	0.07813	0.05875	0.03137	0.01087	0.00658	0.00331	0.00067
-0.10000	0.27889	0.20663	0.16446	0.08182	0.02726	0.01636	0.00818	0.00164
0.00000	0.33333	0.25000	0.20000	0.10000	0.03333	0.02000	0.01000	0.00200
0.10000	0.39000	0.29838	0.24074	0.12167	0.04070	0.02443	0.01222	0.00244
0.50000	0.63889	0.54687	0.47375	0.27342	0.09755	0.05915	0.02979	0.00599
0.90000	0.92333	0.89538	0.86778	0.74458	0.45027	0.31284	0.17346	0.03735
<u>Hinweis:</u>	$\sigma_\varepsilon^2 = 1 - \rho^2$ und damit $\gamma_0 = 1$ sowie $\gamma_1 = \rho$ für jedes ρ .							

C.3 Die Schätzfunktion C_a bei kleinem T

C.3.1 Additive Überlagerung

Wir schreiben die Schätzfunktion C_a als¹⁰⁷

$$C_a = \frac{1}{T-1} \sum_{t=2}^T (X_t + U_t - \bar{X} - \bar{U})(X_{t-1} + U_{t-1} - \bar{X} - \bar{U})$$

wobei unterstellt wird, daß X_t^a durch additive Überlagerung generiert wird.

Erwartungswertbildung ergibt wegen der Unabhängigkeit von X_t und U_s für alle t und s :

$$\begin{aligned} cclE[C_a] &= E\left[\frac{1}{T-1} \sum_{t=2}^T (X_t - \bar{X})(X_{t-1} - \bar{X})\right] + E\left[\frac{1}{T-1} \sum_{t=2}^T (U_t - \bar{U})(U_{t-1} - \bar{U})\right] \\ &= E[C] - \frac{T-2}{(T-1)^2} \sigma_u^2 \quad (\text{siehe (C-15)}) \\ &= \gamma_1 + \xi(\gamma) - \frac{1}{(T)} \sigma_u^2 \quad , \end{aligned} \tag{C-18}$$

wobei in der zweiten Zeile ausgenutzt wurde, daß¹⁰⁸

$$\begin{aligned} &E\left[\frac{1}{T-1} \sum_{t=2}^T (U_t - \bar{U})(U_{t-1} - \bar{U})\right] \\ &= E\left[\frac{1}{T-1} \sum_{t=2}^T \left(U_t - \left(\frac{1}{T} \sum_{s=1}^T U_s\right)\right) \left(U_{t-1} - \left(\frac{1}{T} \sum_{s'=1}^T U_{s'}\right)\right)\right] \\ &= E\left[\frac{1}{T-1} \sum_{t=2}^T U_t U_{t-1}\right] + E\left[\frac{T-1}{T^2(T-1)} \sum_{s=1}^T \sum_{s'=1}^T U_s U_{s'}\right] \\ &\quad - E\left[\frac{1}{(T(T-1))} \sum_{t=2}^T \sum_{s=1}^T U_{t-1} U_s\right] - E\left[\frac{1}{(T(T-1))} \sum_{t=2}^T \sum_{s'=1}^T U_t U_{s'}\right] \\ &= 0 + \frac{1}{T} \sigma_u^2 - \frac{2}{T} \sigma_u^2 \\ &= -\frac{1}{T} \sigma_u^2 \end{aligned}$$

¹⁰⁷Man beachte, daß dasselbe Symbol C_a wie bei der asymmetrischen Form benutzt wird.

¹⁰⁸Sowohl Hans Schneeweiß als auch Andy Tremayne haben - zeitgleich - auf das bisher verwendete fehlerhafte Ergebnis hingewiesen, wonach der folgende Erwartungswert gleich Null sein soll!

gilt. Demnach ist der Erwartungswert von C_a bei additiver Überlagerung und $T \geq 3$ stets kleiner als der Erwartungswert von C . Damit ist klar, daß der "naive" IV-Schätzer bei additiver Überlagerung nicht konsistent ist. Siehe dazu im Einzelnen Abschnitt C.4.

Speziell für den Fall, daß keine Autokorrelation vorliegt, gilt

$$E[C_a] = -\frac{\gamma_0 + \sigma_u^2}{T} < 0$$

und demnach wird der wahre Autokorrelationskoeffizient $\gamma_1 (= 0)$ stärker als im Fall der Originaldaten (siehe Abschnitt C.2) unterschätzt.

C.3.2 Allgemeine multiplikative Überlagerung

Im Fall der "allgemeinen" multiplikativen Überlagerung schreiben wir die Schätzfunktion als

$$\begin{aligned} C_a &= \frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \frac{1}{T} \sum_{s=1}^T X_s U_s) (X_{t-1} U_{t-1} - \frac{1}{T} \sum_{s=1}^T X_s U_s) \\ &= \frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu + \mu - \frac{1}{T} \sum_{s=1}^T X_s U_s) (X_{t-1} U_{t-1} - \mu + \mu - \frac{1}{T} \sum_{s=1}^T X_s U_s) \\ &= \frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu_x) (X_{t-1} U_{t-1} - \mu_x) \\ &\quad + \frac{1}{T-1} \sum_{t=2}^T \left(\frac{1}{T} \sum_{s=1}^T (X_s U_s - \mu_x) \right) \left(\frac{1}{T} \sum_{s'=1}^T (X_{s'} U_{s'} - \mu_x) \right) \\ &\quad - \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_t U_t - \mu_x) (X_s U_s - \mu_x) \\ &\quad - \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_{t-1} U_{t-1} - \mu_x) (X_s U_s - \mu_x) \end{aligned}$$

Erwartungswertbildung für die vier Summanden ergibt folgendes: Für den ersten Summanden erhalten wir wegen (11-16)

$$E \left[\frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu_x) (X_{t-1} U_{t-1} - \mu_x) \right] = cov[X_t^a, X_{t-1}^a] = cov[X_t, X_{t-1}] = \gamma_1 \quad .$$

Für den zweiten Summanden ergibt sich (wie bei der Analyse von $E[S_a^2]$)

$$\begin{aligned} &E \left[\frac{1}{T-1} \sum_{t=2}^T \left(\frac{1}{T} \sum_{s=1}^T (X_s U_s - \mu_x) \right) \left(\frac{1}{T} \sum_{s'=1}^T (X_{s'} U_{s'} - \mu_x) \right) \right] \\ &= \frac{1}{T^2} \left(\sum_{s=1}^T var[X_s U_s] + \sum_{s=1}^T \sum_{s'=1, s \neq s'}^T cov[X_s U_s, X_{s'} U_{s'}] \right) \\ &= \frac{1}{T^2} \left(\sum_{s=1}^T var[X_s U_s] + \sum_{s=1}^T \sum_{s'=1, s \neq s'}^T cov[X_s, X_{s'}] \right) \end{aligned}$$

wobei das Ergebnis in der dritten Zeile wieder ausnutzt, daß bei multiplikativer Überlagerung

$$cov[X_s U_s, X_{s'} U_{s'}] = cov[X_s, X_{s'}]$$

gilt. Siehe (11-16). Damit erhalten wir

$$\begin{aligned} &E \left[\frac{1}{T-1} \sum_{t=2}^T \left(\frac{1}{T} \sum_{s=1}^T (X_s U_s - \mu_x) \right) \left(\frac{1}{T} \sum_{s'=1}^T (X_{s'} U_{s'} - \mu_x) \right) \right] \\ &= \frac{1}{T} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{T^2} (2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \dots + 2 \cdot 2\gamma_{T-2} + 2 \cdot 1\gamma_{T-1}) \end{aligned}$$

Für den dritten Summanden ergibt sich (unter Vernachlässigung des Minuszeichens)¹⁰⁹

$$\begin{aligned} & E \left[\frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_t U_t - \mu_x)(X_s U_s - \mu_x) \right] \\ &= \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T V[X_t U_t] + \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1, s \neq t}^T \text{cov}[X_t, X_s] \\ &= \frac{1}{T} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{T-1} \frac{1}{T} \{ (2T-3)\gamma_1 + (2T-5)\gamma_2 + \dots + 3\gamma_{T-2} + 1\gamma_{T-1} \} \end{aligned}$$

Für den vierten Summanden erhalten wir (genau wie bei der Analyse von $E[C]$) das identische Resultat:

$$\begin{aligned} & E \left[\frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_{t-1} U_{t-1} - \mu_x)(X_s U_s - \mu_x) \right] \\ &= \frac{1}{T} \left(\gamma_0 + \sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{T-1} \frac{1}{T} \{ (2T-3)\gamma_1 + (2T-5)\gamma_2 + \dots + 3\gamma_{T-2} + 1\gamma_{T-1} \} \end{aligned}$$

Diese Einzelergebnisse lassen sich wie folgt zusammenfassen: Für den Erwartungswert von C_a gilt bei allgemeiner multiplikativer Überlagerung

$$E[C_a] = \gamma_1 - \frac{1}{T} \left(\sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \xi(\gamma, T) \quad (\text{C-19})$$

oder auch

$$E[C_a] = E[C] - \frac{\sigma_u^2}{T} \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) < E[C] \quad , \quad (\text{C-20})$$

wobei $\xi(\gamma, T)$ durch (C-16) gegeben ist.

C.3.3 Multiplikative Überlagerung mit Faktorstruktur

Die Abschnitt C.3.2 entsprechende Analyse für den Höhenfall hat zu beachten, daß die (Auto-)Kovarianzen der überlagerten Variablen **nicht** gleich den (Auto-)Kovarianzen der

¹⁰⁹Die Ergebnisse der letzten Zeile der folgenden Gleichung lassen sich durch die folgende Tabelle verifizieren:

		t							
		2	3	4	5	...	$T-2$	$T-1$	T
s	1	γ_1	γ_2	γ_3	γ_4	...	γ_{T-3}	γ_{T-2}	γ_{T-1}
	2	$V[X_t U_t]$	γ_1	γ_2	γ_3	...	γ_{T-4}	γ_{T-3}	γ_{T-2}
	3	γ_1	$V[X_t U_t]$	γ_1	γ_2	...	γ_{T-5}	γ_{T-4}	γ_{T-3}
	⋮					⋮			
	$T-3$	γ_{T-5}	γ_{T-6}	γ_{T-7}	γ_{T-8}	...	γ_1	γ_2	γ_3
	$T-2$	γ_{T-4}	γ_{T-5}	γ_{T-6}	γ_{T-7}	...	$V[X_t U_t]$	γ_1	γ_2
	$T-1$	γ_{T-3}	γ_{T-4}	γ_{T-5}	γ_{T-6}	...	γ_1	$V[X_t U_t]$	γ_1
T	γ_{T-2}	γ_{T-3}	γ_{T-4}	γ_{T-5}	...	γ_2	γ_1	$V[X_t U_t]$	

Originalvariablen sind. Siehe dazu auch die Ableitung des Erwartungswertes von S_a^2 in Abschnitt 11.4.3; auf die Ergebnisse jenes Abschnittes werden wir uns im Folgenden beziehen, insbesondere wird die in (11-32) definierte Autokovarianz

$$\gamma_{t-s}^H \equiv \text{cov}[X_t^a, X_s^a] = \gamma_0 \rho^{|t-s|} + \delta^2 \left\{ (1 + \rho + \dots + \rho^{t-s-1}) \frac{\rho_0^2}{1-\rho} + \rho^{(t-s)} \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right\}$$

eine Rolle spielen.

Wir starten wieder von der Zerlegung der Schätzfunktion C_a in vier Summanden (siehe Abschnitt C.3.2) und bestimmen nun die Erwartungswert der einzelnen Summanden unter Beachtung der Höhne-Spezifikation $X_t^a = (1 + \delta D + \varepsilon_t)X_t$.

Für den ersten Summanden erhalten wir wegen (11-32)

$$E \left[\frac{1}{T-1} \sum_{t=2}^T (X_t U_t - \mu_x) (X_{t-1} U_{t-1} - \mu_x) \right] = \text{cov}[X_t^a, X_{t-1}^a] = \gamma_1^H \quad .$$

Für den zweiten Summanden ergibt sich

$$\begin{aligned} & E \left[\frac{1}{T-1} \sum_{t=2}^T \left(\frac{1}{T} \sum_{s=1}^T (X_s U_s - \mu_x) \right) \left(\frac{1}{T} \sum_{s'=1}^T (X_{s'} U_{s'} - \mu_x) \right) \right] \\ &= \frac{1}{T^2} \left(\sum_{s=1}^T \text{var}[X_s U_s] + \sum_{s=1}^T \sum_{s'=1, s \neq s'}^T \text{cov}[X_s U_s, X_{s'} U_{s'}] \right) \\ &= \frac{1}{T} \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{T^2} (2(T-1)\gamma_1^H + 2(T-2)\gamma_2^H + \dots + 2 \cdot 2\gamma_{T-2}^H + 2 \cdot 1\gamma_{T-1}^H) \end{aligned}$$

Für den dritten Summanden ergibt sich (unter Vernachlässigung des Minuszeichens)

$$\begin{aligned} & E \left[\frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1}^T (X_t U_t - \mu_x) (X_s U_s - \mu_x) \right] \\ &= \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T V[X_t U_t] + \frac{1}{T-1} \frac{1}{T} \sum_{t=2}^T \sum_{s=1, s \neq t}^T \text{cov}[X_t U_t, X_s U_s] \\ &= \frac{1}{T} \left(\gamma_0 + (\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \frac{1}{T-1} \frac{1}{T} \{ (2T-3)\gamma_1^H + (2T-5)\gamma_2^H + \dots + 3\gamma_{T-2}^H + 1\gamma_{T-1}^H \} \end{aligned}$$

Für den vierten Summanden erhalten wir dasselbe Resultat wie für den dritten Summanden (siehe auch Abschnitt C.3.2), so daß eine Darstellung hier überflüssig ist.

Die Einzelergebnisse lassen sich wie folgt zusammenfassen: Für den Erwartungswert von C_a gilt bei multiplikativer Höhne-Überlagerung

$$E[C_a] = \gamma_1^H - \frac{1}{T} \left((\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right) + \xi^H(\gamma^H, T) \quad (\text{C-21})$$

wobei $\xi^H(\gamma^H, T)$ durch

$$\begin{aligned} \xi^H(\gamma^H, T) &= \left(\frac{1}{T}\right)^2 [2\{(T-1)\gamma_1^H + (T-2)\gamma_2^H + \dots + 3\gamma_{T-3}^H + 2\gamma_{T-2}^H + \gamma_{T-1}^H\}] \\ &\quad - \frac{2}{T(T-1)} [(2T-3)\gamma_1^H + (2T-5)\gamma_2^H + \dots + 5\gamma_{T-3}^H + 3\gamma_{T-2}^H + 1\gamma_{T-1}^H] \\ &\quad - \frac{\gamma_0}{T} \\ &= \frac{2}{T} \sum_{k=1}^{T-1} \left(\frac{T-k}{T} - \frac{2T-(2k+1)}{T-1} \right) \gamma_k^H - \frac{\gamma_0}{T} . \end{aligned} \quad (\text{C-22})$$

gegeben ist.

Zwar erscheint auf den ersten Blick die Struktur dieses Erwartungswertes große Ähnlichkeit mit dem Erwartungswert im Fall der allgemeinen multiplikativen Überlagerung zu haben, doch ergeben sich bei näherem Hinsehen deutliche Unterschiede.

- Da für $T \rightarrow \infty$ der zweite und dritte Term in (C-21) verschwinden, strebt der IV-Schätzer im Höhnfall gegen $\gamma_1^H \neq \gamma_1$ und ist deshalb nicht konsistenter Schätzer der Autokovarianz.¹¹⁰
- Für den Fall $\varrho = 0$ gilt

$$\gamma_k^H = \delta^2 \varrho_0^2, \quad k = 1, 2, \dots,$$

d.h. die Auto-Kovarianzen γ_k^H verschwinden auch für $\varrho = 0$ nicht.

- Es fällt auf, daß sich nun $E[C]$ und $E[C_a]$ nicht nur durch den Term

$$\frac{1}{T} \left((\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1-\rho} \right)^2 \right) \right)$$

unterscheiden., was Konsequenzen für die Konstruktion eines Korrektorschätzers hat.

- Für kleines T ist die Beschreibung der Verzerrung kompliziert.

¹¹⁰Es gilt

$$\begin{aligned} \gamma_1^H &= \gamma_0 \varrho + \delta^2 \left(\frac{\varrho^2}{1-\varrho} + \varrho \left(\gamma_0 + \left(\frac{\varrho_0}{1-\varrho} \right)^2 \right) \right) \\ &= (1 + \delta^2) \gamma_0 \varrho + \varrho_0^2 \left(\frac{\delta^2}{1-\varrho} + \frac{\delta^2 \varrho_0 \varrho}{(1-\varrho)^2} \right) \end{aligned}$$

und

$$\gamma_1 = \varrho \gamma_0 .$$

C.4 Relevanz für den IV-Schätzer

C.4.1 Ergebnisse für den Panel-IV-Schätzer

Wir betrachten nun die Erweiterung für den Panelfall. Dann lautet der obige Schätzer (symmetrische Form) wie folgt:

$$\begin{aligned}\hat{\beta}_W^{a,IV} &= \frac{\sum_t \sum_i (X_{i,t-1}^a - \bar{X}_{i\bullet}^a)(Y_{it} - \bar{Y}_{i\bullet})}{\sum_t \sum_i (X_{i,t-1}^a - \bar{X}_{i\bullet}^a)(X_{it}^a - \bar{X}_{i\bullet}^a)} \\ &= \frac{\sum_t \sum_i (X_{i,t-1}^a - \bar{X}_{i\bullet}^a)(X_{it}^a - \bar{X}_{i\bullet}^a)}{\sum_t \sum_i (X_{i,t-1}^a - \bar{X}_{i\bullet}^a)(X_{it}^a - \bar{X}_{i\bullet}^a)} \beta + \frac{\sum_t \sum_i (X_{i,t-1}^a - \bar{X}_{i\bullet}^a)(\eta_{it} - \bar{\eta}_{i\bullet})}{\sum_t \sum_i (X_{i,t-1}^a - \bar{X}_{i\bullet}^a)(X_{it}^a - \bar{X}_{i\bullet}^a)}\end{aligned}\quad (C-23)$$

Wir benutzen nun die Ergebnisse aus den Abschnitten C.2, C.3.1, C.3.2 sowie C.3.3, um für additive und multiplikative Überlagerung die Wahrscheinlichkeitsgrenzwerte des Panel-IV-Schätzers (C-23) zu bestimmen.

Additive Überlagerung

Im Fall der additiven Überlagerung erhalten wir

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^{a,IV} = \frac{\gamma_1 + \xi(\gamma, T)}{\gamma_1 + \xi(\gamma, T) - \frac{\sigma_u^2}{T}} \beta \neq \beta \quad (C-24)$$

mit

$$\xi(\gamma, T) = \frac{2}{T} \sum_{k=1}^{T-1} \left(\frac{T-k}{T} - \frac{2T-(2k+1)}{T-1} \right) \gamma_k - \frac{\gamma_0}{T} .$$

Siehe (C-16) und (C-18). Dabei nimmt die Biasfunktion ξ für positive Autokorrelation stets nur negative Werte an und verschwindet für $T \rightarrow \infty$. Man beachte, daß also auch bei additiver Fehlerüberlagerung der Schätzer nicht konsistent ist, selbst wenn T groß ist !

Man beachte, daß der Ausdruck

$$\gamma_1 + \xi(\gamma, T) ,$$

durchaus negative Werte annehmen kann, ohne daß die Konsistenz dadurch beeinträchtigt wird. Wenn der Ausdruck allerdings (annähernd) gleich Null ist, kann die Verzerrung durch die Funktion ξ zu einer äußerst unzuverlässigen Schätzung führen, die durch numerischen Instabilität hervorgerufen wird: Da der Nenner (annähernd) gleich Null sein kann, ergibt sich dadurch eine eher numerisch bedingte hohe Varianz der IV-Schätzung. So läßt sich beispielsweise zeigen¹¹¹, daß für

$$T = 4 \quad \text{und} \quad \varrho = 0.77200187265876$$

dieser Ausdruck exakt den Wert Null annimmt. (Der entsprechende Wert für $T = 10$ lautet $\varrho = 0.12867394027791$.)¹¹² Da er im Zähler als auch im Nenner auftaucht, ist dann der

¹¹¹Für $T = 4$ ergibt sich

$$\gamma_1 + \xi(\gamma, T) = \frac{\gamma_0}{48} (26\varrho - 12\varrho^2 - 2\varrho^3 - 12) .$$

Für $\varrho = 0.77200187265876$ hat die kubische Funktion in der Klammer eine Nullstelle.

¹¹²Auf diese Ergebnisse kommen wir bei der Interpretation der Simulationsergebnisse in Abschnitt 12.5.1 zurück.

Schätzer bzw. korrekter sein Wahrscheinlichkeitsgrenzwert nicht definiert. Allerdings wird sich ja bei der Anwendung der IV-Schätzung nie (bzw. nur mit Wahrscheinlichkeit Null) genau der Wert Null ergeben, sondern ein Wert nahe bei Null, der dann allerdings zu einer sehr unzuverlässigen Schätzung führt. Man vergleiche dazu die Simulationsergebnisse in Abschnitt 12.5.1.

Allgemeine multiplikative Überlagerung .

Im Fall der "allgemeinen" multiplikativen Überlagerung erhalten wir als Wahrscheinlichkeitsgrenzwert für den Panel-IV-Schätzer (C-23)

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^{a,IV} = \frac{\gamma_1 + \xi(\gamma, T)}{\gamma_1 + \xi(\gamma, T) - \frac{1}{T} \left(\sigma_u^2 \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right) \right)} \beta \quad (\text{C-25})$$

mit

$$\xi(\gamma, T) = \frac{2}{T} \sum_{k=1}^{T-1} \left(\frac{T-k}{T} - \frac{2T-(2k+1)}{T-1} \right) \gamma_k - \frac{\gamma_0}{T} .$$

Siehe (C-16) und (C-19). Dabei nimmt die Biasfunktion ξ **für positive Autokorrelation** stets nur negative Werte an und verschwindet für $T \rightarrow \infty$.

Die für den Fall der additiven Überlagerung geschilderten Probleme, die sich für den Fall, daß der Nenner von (C-25) (annähernd) gleich Null ist, gelten natürlich auch hier. Die explizite Analyse ist allerdings aufwendiger.

Multiplikative Höhne-Überlagerung .

Im Fall der multiplikativen Höhne-Überlagerung erhalten wir als Wahrscheinlichkeitsgrenzwert für den Panel-IV-Schätzer (C-23)

$$\text{plim}_{N \rightarrow \infty} \hat{\beta}_W^{a,IV} = \frac{\gamma_1 + \xi(\gamma, T)}{\gamma_1^H + \xi^H(\gamma^H, T) - \frac{1}{T} \left((\delta^2 + \sigma_\varepsilon^2) \left(\gamma_0 + \left(\frac{\rho_0}{1 - \rho} \right)^2 \right) \right)} \beta \quad (\text{C-26})$$

mit

$$\begin{aligned} \xi^H(\gamma^H, T) &= \left(\frac{1}{T} \right)^2 [2 \{ (T-1)\gamma_1^H + (T-2)\gamma_2^H + \dots + 3\gamma_{T-3}^H + 2\gamma_{T-2}^H + \gamma_{T-1}^H \}] \\ &\quad - \frac{2}{T(T-1)} [(2T-3)\gamma_1^H + (2T-5)\gamma_2^H + \dots + 5\gamma_{T-3}^H + 3\gamma_{T-2}^H + 1\gamma_{T-1}^H] - \frac{\gamma_0}{T} \\ &= \frac{2}{T} \sum_{k=1}^{T-1} \left(\frac{T-k}{T} - \frac{2T-(2k+1)}{T-1} \right) \gamma_k^H - \frac{\gamma_0}{T} . \end{aligned}$$

Siehe (C-21) und (C-22).

Die für den Fall der additiven Überlagerung geschilderten Probleme, die sich für den Fall, daß der Nenner von (C-26) (annähernd) gleich Null ist, gelten natürlich auch hier. Die explizite Analyse ist allerdings deutlich (!) aufwendiger.

IAW-Diskussionspapiere

Bisher erschienen:

- Nr. 1 (September 2001)
Das Einstiegsgeld – eine zielgruppenorientierte negative Einkommensteuer: Konzeption, Umsetzung und eine erste Zwischenbilanz nach 15 Monaten in Baden-Württemberg
Sabine Dann / Andrea Kirchmann / Alexander Spermann / Jürgen Volkert
- Nr. 2 (Dezember 2001)
Die Einkommensteuerreform 1990 als natürliches Experiment. Methodische und konzeptionelle Aspekte zur Schätzung der Elastizität des zu versteuernden Einkommens
Peter Gottfried / Hannes Schellhorn
- Nr. 3 (Januar 2001)
Gut betreut in den Arbeitsmarkt? Eine mikroökonomische Evaluation der Mannheimer Arbeitsvermittlungsagentur
Jürgen Jerger / Christian Pohnke / Alexander Spermann
- Nr. 4 (Dezember 2001)
Das IAW-Einkommenspanel und das Mikrosimulationsmodell SIMST
Peter Gottfried / Hannes Schellhorn
- Nr. 5 (April 2002)
A Microeconomic Characterisation of Household Consumption Using Quantile Regression
Niels Schulze / Gerd Ronning
- Nr. 6 (April 2002)
Determinanten des Überlebens von Neugründungen in der baden-württembergischen Industrie – eine empirische Survivalanalyse mit amtlichen Betriebsdaten
Harald Strotmann
- Nr. 7 (November 2002)
Die Baulandausweisungsumlage als ökonomisches Steuerungsinstrument einer nachhaltigkeitsorientierten Flächenpolitik
Raimund Krumm
- Nr. 8 (März 2003)
Making Work Pay: U.S. American Models for a German Context?
Laura Chadwick, Jürgen Volkert
- Nr. 9 (Juni 2003)
Erste Ergebnisse von vergleichenden Untersuchungen mit anonymisierten und nicht anonymisierten Einzeldaten am Beispiel der Kostenstrukturerhebung und der Umsatzsteuerstatistik
Martin Rosemann

IAW-Diskussionspapiere

- Nr. 10 (August 2003)
Randomized Response and the Binary Probit Model
Gerd Ronning
- Nr. 11 (August 2003)
Creating Firms for a New Century: Determinants of Firm Creation
around 1900
Joerg Baten
- Nr. 12 (September 2003)
Das fiskalische BLAU-Konzept zur Begrenzung des Siedlungsflächenwachstums
Raimund Krumm
- Nr. 13 (Dezember 2003)
Generelle Nichtdiskontierung als Bedingung für eine nachhaltige Entwicklung?
Stefan Bayer
- Nr. 14 (Februar 2003)
Die Elastizität des zu versteuernden Einkommens. Messung und erste Ergebnisse
zur empirischen Evidenz für die Bundesrepublik Deutschland.
Peter Gottfried / Hannes Schellhorn
- Nr. 15 (Februar 2004)
Empirical Evidence on the Effects of Marginal Tax Rates on Income –
The German Case
Peter Gottfried / Hannes Schellhorn
- Nr. 16 (Juli 2004)
Shadow Economies around the World: What do we really know?
Friedrich Schneider
- Nr. 17 (August 2004)
Firm Foundations in the Knowledge Intensive Business Service Sector.
Results from a Comparative Empirical Study in Three German Regions
Andreas Koch / Thomas Stahlecker
- Nr. 18 (Januar 2005)
The impact of functional integration and spatial proximity on the post-entry
performance of knowledge intensive business service firms
Andreas Koch / Harald Strotmann
- Nr. 19 (März 2005)
Legislative Malapportionment and the Politicization of Germany's
Intergovernmental Transfer System
Hans Pitlik / Friedrich Schneider / Harald Strotmann
- Nr. 20 (April 2005)
Implementation ökonomischer Steuerungsansätze in die Raumplanung
Raimund Krumm

IAW-Diskussionspapiere

- Nr. 21 (Juli 2005)
Determinants of Innovative Activity in Newly Founded Knowledge
Intensive Business Service Firms
Andreas Koch / Harald Strotmann
- Nr. 22 (Dezember 2005)
Impact of Opening Clauses on Bargained Wages
Wolf Dieter Heinbach
- Nr. 23 (Januar 2006)
Hat die Einführung von Gewinnbeteiligungsmodellen kurzfristige positive
Produktivitätswirkungen? – Ergebnisse eines Propensity-Score-Matching-Ansatzes
Harald Strotmann
- Nr. 24 (März 2006)
Who Goes East? The Impact of Enlargement on the Pattern of German FDI
Claudia M. Buch / Jörn Kleinert
- Nr. 25 (Mai 2006)
Estimation of the Probit Model from Anonymized Micro Data
Gerd Ronning / Martin Rosemann
- Nr. 26 (Oktober 2006)
Bargained Wages in Decentralized Wage-Setting Regimes
Wolf Dieter Heinbach
- Nr. 27 (Januar 2007)
A Capability Approach for Official German Poverty and Wealth Reports:
Conceptual Background and First Empirical Results
Christian Arndt / Jürgen Volkert
- Nr. 28 (Februar 2007)
Typisierung der Tarifvertragslandschaft – Eine Clusteranalyse der tarif-
vertraglichen Öffnungsklauseln
Wolf Dieter Heinbach / Stefanie Schröpfer
- Nr. 29 (März 2007)
International Bank Portfolios: Short- and Long-Run Responses to
the Business Cycles
Sven Blank / Claudia M. Buch
- Nr. 30 (April 2007)
Stochastische Überlagerungen mit Hilfe der Mischungsverteilung
Gerd Ronning
- Nr. 31 (Mai 2007)
Openness and Growth: The Long Shadow of the Berlin Wall
Claudia M. Buch / Farid Toubal

IAW-Diskussionspapiere

- Nr. 32 (Mai 2007)
International Banking and the Allocation of Risk
Claudia M. Buch / Gayle DeLong / Katja Neugebauer
- Nr. 33 (Juli 2007)
Multinational Firms and New Protectionisms
Claudia M. Buch / Jörn Kleinert
- Nr. 34 (November 2007)
Within-Schätzung bei anonymisierten Paneldaten
Elena Biewen
- Nr. 35 (Dezember 2007)
What a Difference Trade Makes – Export Activity and the Flexibility of Collective Bargaining Agreements
Wolf Dieter Heinbach / Stefanie Schröpfer
- Nr. 36 (Dezember 2007)
To Bind or Not to Bind Collectively? Decomposition of Bargained Wage Differences Using Counterfactual Distributions
Wolf Dieter Heinbach / Markus Spindler
- Nr. 37 (Dezember 2007)
Neue Ansätze zur flächenschutzpolitischen Reform des Kommunalen Finanzausgleichs
Raimund Krumm
- Nr. 38 (Januar 2008)
Banking Globalization: International Consolidation and Mergers in Banking
Claudia M. Buch / Gayle L. DeLong
- Nr. 39 (Januar 2008)
Multiplicative Measurement Error and the Simulation Extrapolation Method
Elena Biewen / Sandra Nolte / Martin Rosemann
- Nr. 40 (Juni 2008)
Das Konzept des „Regionalen Gewerbeflächenpools“ aus ökonomischer Sicht
Raimund Krumm
- Nr. 41 (Juli 2008)
Openness and Income Disparities: Does Trade Explain the 'Mezzogiorno' Effect?
Claudia M. Buch / Paola Monti
- Nr. 42 (August 2008)
Flächenschutzpolitische Implikationen eines Regionalen Gewerbeflächenpools
Raimund Krumm

IAW-Diskussionspapiere

- Nr. 43 (September 2008)
Mikroökonomische Determinanten und Effekte von FDI am Beispiel
Baden-Württemberg
Christian Arndt / Anselm Mattes
- Nr. 44 (September 2008)
The Impact of Macroeconomic Factors on Risks in the Banking Sector:
A Cross-Country Empirical Assessment
Olga Bohachova
- Nr. 45 (Oktober 2008)
Effects of Dismissal Protection Legislation on Individual Employment
Stability in Germany
Bernhard Boockmann / Daniel Gutknecht / Susanne Steffes
- Nr. 46 (November 2008)
Trade's Impact on the Labor Share: Evidence from German and Italian Regions
Claudia M. Buch / Paola Monti / Farid Toubal
- Nr. 47 (März 2009)
Network and Border Effects: Where Do Foreign Multinationals Locate
in Germany?
Julia Spies
- Nr. 48 (März 2009)
Stochastische Überlagerung mit Hilfe der Mischungsverteilung
(Stand: 18. März 2009 – Version 49)
Gerd Ronning

