politica, 39. årg. nr. 3 2007, 294-316

Simon Calmar Andersen

Multilevel-modeller: en introduktion og et eksempel

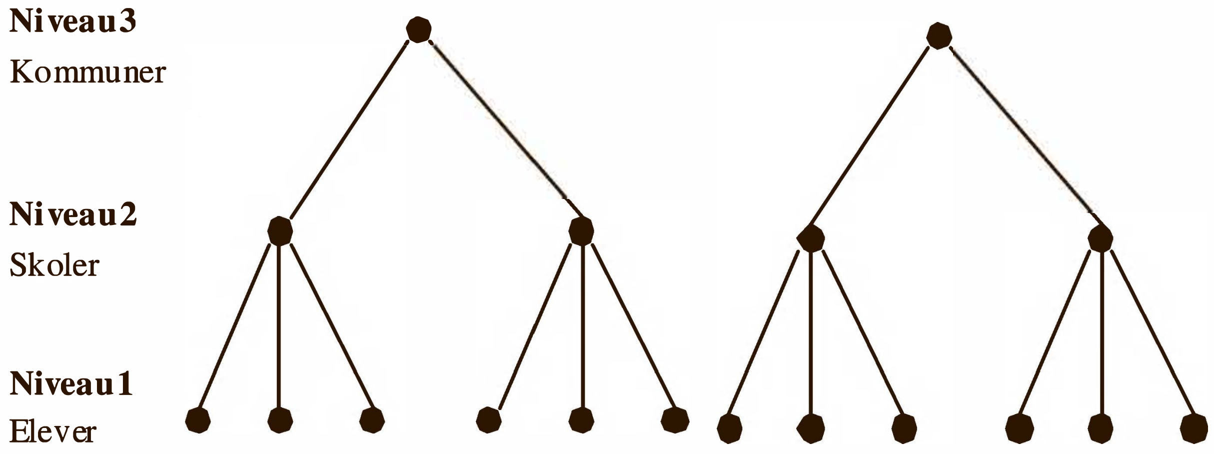
I samfundsvidenskaberne arbejdes der ofte med data, som er struktureret i hierarkier. Når det er tilfældet, giver såkaldte multilevel-modeller ikke bare mulighed for en mere korrekt beregning af standardfejl og signifikanstest, men også for at udnytte strukturen i datasættet til at undersøge forhold af substantiel interesse. Til at illustrere dette benyttes data om danske skoleelevers prøvekarakterer. Datasættet er struktureret i elever i skoler i kommuner, og multilevel-modellering gør det derfor muligt at undersøge, om der er signifikante forskelle mellem enhederne på det politiske niveau, kommunerne, om kommunerne gør en større forskel for elever med svag socioøkonomisk baggrund, og om man kan identificere enkelte kommuner, som klarer sig signifikant bedre end andre. Afslutningsvis henvises der til supplerende litteratur om de statistiske og de praktiske aspekter i metodens anvendelse.

Når data er hierarkisk opbygget

Inden for samfundsvidenskaberne er de data, som ligger til grund for statistiske analyser, ofte hierarkisk opbygget. Det gælder fx den rumlige struktur, når man har oplysninger om regioner eller individer i forskellige lande, om forskellige kommuner i samme land eller om vælgere i forskellige opstillingskredse og afstemningsdistrikter. Eller hvis man har oplysninger om brugere af forskellige institutioner såsom patienter på sygehuse eller elever i skoler. Det gælder også den tidslige struktur, når data rummer flere observationer af de samme personer eller samme institutioner på forskellige tidspunkter. Den sociale struktur i datasættet kan også være hierarkisk som i studier af familier med flere børn. Eller man kan have kombinationer af disse forskellige typer, så man får hierarkier med tre eller flere niveauer. Figur 1 illustrerer princippet med elever i skoler i kommuner som eksempel.

Generelt gælder det for sådanne hierarkisk opbyggede datasæt, at de ikke kan analyseres korrekt med den klassiske lineære regressionsmodels *ordinary least squares* (OLS)-estimater, blandt andet fordi den model bygger på en forudsætning om, at observationerne er uafhængige af hinanden. Den forudsætning holder sjældent stik i denne type datasæt, fordi enheder inden for samme gruppe ofte vil være stærkere korreleret med hinanden end med enheder i andre grupper. Heldigvis er der blevet udviklet en type statistiske modeller, som ikke blot kan estimere sammenhænge i sådanne datasæt korrekt, men som også kan udnytte strukturen til at undersøge forhold af substantiel interesse. For eksempel kan det undersøges, om grupper på overordnet niveau, såsom

Figur 1. Strukturen i et hierarkisk datasæt. Skoleelever som eksempel



kommuner, samlet set har en signifikant sammenhæng med en afhængig variabel på lavere niveau, som for eksempel elevers karakterer. Eller om en forklarende variabel har forskellige effekter på forskellige niveauer eller i forskellige enheder inden for samme niveau.

Disse modeller går under navne som multilevel regressionsmodeller, hierarkiske lineære modeller, *mixed effects* regressioner eller *random coefficients* modeller, og de har de senere år vundet i både udbredelse og tilgængelighed, blandt andet fordi metoden til at beregne dem med efterhånden er blevet integreret i de store statistiksoftwareprogrammer som SPSS (ver. 11), STATA (ver. 9) og SAS.

Næste afsnit beskriver først problemerne ved OLS-regressioner, når data er hierarkisk struktureret, og derefter præsenteres grundprincipperne i multilevel-modeller. Tredje afsnit illustrerer nogle af mulighederne i disse relativt nye modeller på et stort datasæt med oplysninger om danske folkeskoleelevers prøvekarakterer og sociale baggrund. Datasættet er struktureret i tre niveauer, elever i skoler i kommuner, og spørgsmålet er, om det politiske niveau, kommunerne, gør en forskel for elevernes karakterer, når der kontrolleres for betydningen af socioøkonomiske forhold på elevniveau. Artiklen her kan kun give en kort introduktion til og forsmag på multilevel-modeller, men sidste afsnit sammenfatter de tekniske og substantielle fordele ved metoden, og der henvises til supplerende litteratur om både den statistiske og den praktiske side af arbejdet med den.

Statistikken bag multilevel-modeller

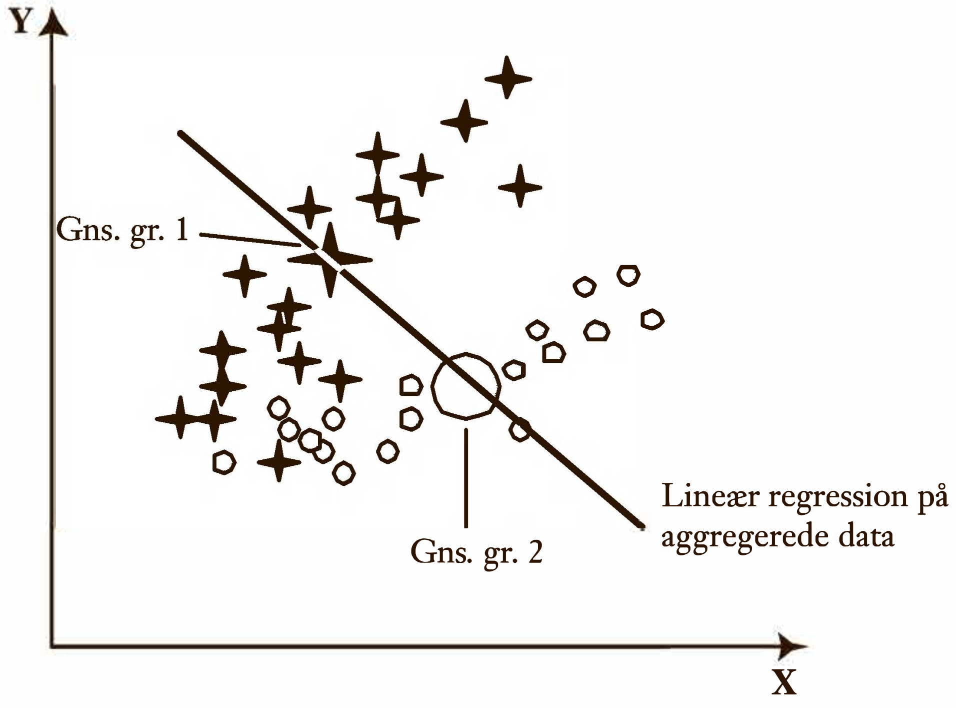
Man har længe været opmærksom på problemerne ved at drage økologiske fejlslutninger om individuel adfærd på baggrund af aggregerede data (i hvert fald siden Robinson, 1950). For at illustrere problematikken kan vi betragte figur 2 som en kortlægning af sammenhængen mellem to variable, individers indtægt (X) og deres almene sundhedstilstand (Y), for to grupper (to kommuner) markeret med henholdsvis cirkler og stjerner. De små cirkler og stjerner markerer så de enkelte individer, mens den store cirkel og den store stjerne markerer de to kommuners gennemsnit på de to variable. De konstruerede observationer i figuren viser på individniveau en tydeligt positiv sammenhæng mellem indtægt og sundhedstilstand. Men hvis vi alene betragter disse data på aggregeret niveau, ser vi en klar negativ sammenhæng mellem kommunernes gennemsnit på de to variable, markeret ved den grå regressionslinje. En sådan aggregeret betragtning kan formelt formuleres med den statistiske model,

$$\bar{y}_j = \beta_0 + \beta_1\bar{x}_j + e_j \qquad (1.)$$

hvor $\bar{y}_j$ er gennemsnittet på den afhængige variabel for hver gruppe på niveau 2, kaldet j. $\bar{x}_j$ er gennemsnittet af den forklarende variabel på niveau j, β_0 er en konstant, som i den grafiske fremstilling svarer til skæringspunktet med Y-aksen, mens β_1 er regressionskoefficienten, der svarer til hældningen. e_j er residualen og repræsenterer den del af Y, som ikke kan indfanges af den lineære model. På grafen kommer den til udtryk som den lodrette afstand mellem en observation og regressionslinjen. Det antages almindeligvis, at e_j er normalfordelt og har en fast varians.

En økologisk fejlslutning ville bestå i at slutte fra det aggregerede resultat, en negativ sammenhæng, og til at der for det enkelte individ er en negativ sammenhæng mellem indtægt og sundhedstilstand. Resultatet viser alene en

Figur 2. Konstrueret sammenhæng mellem to variable analyseret på aggregeret niveau



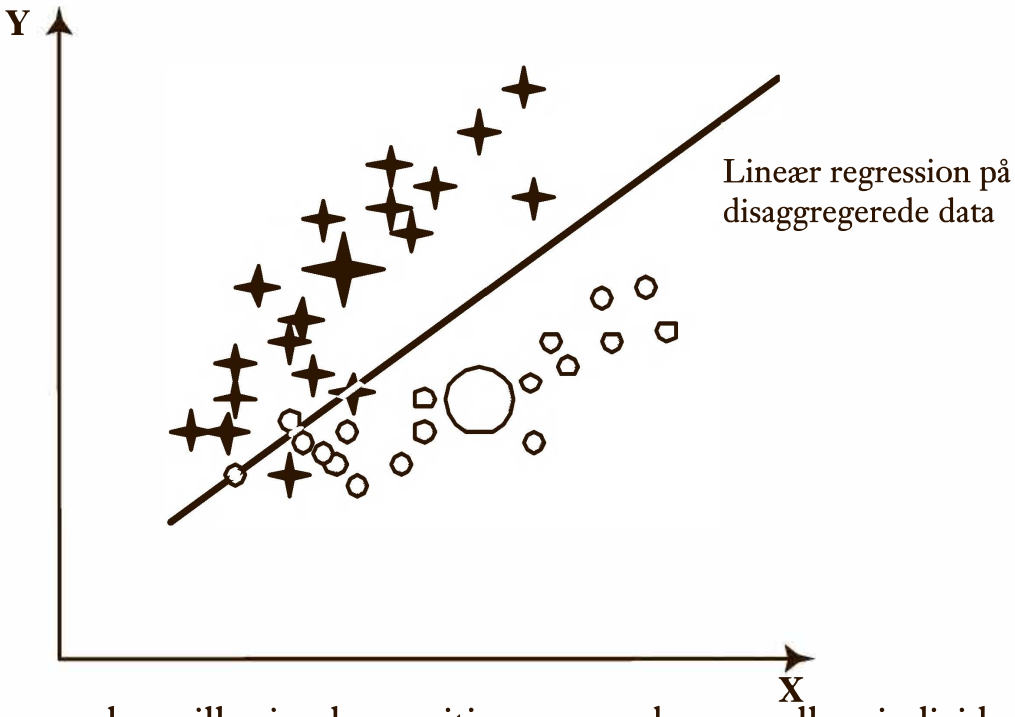
sammenhæng *mellem* (ikke inden for) kommuner. På kommuneniveau gælder det i eksemplet, at kommunernes gennemsnitsindtægt er negativt relateret til den gennemsnitlige sundhedstilstand – måske fordi den rigeste kommune har en større tæthed af forurenende fabrikker.

Et alternativ til den aggregerende model er at nøjes med at betragte observationerne på individniveau uden at skelne mellem grupperne. Det kan formelt modelleres således:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + e_{ij} \qquad (2.)$$

Her markerer fodtegnet *i*, at variablerne *Y* og *X* varierer mellem enhederne på niveau 1, *i*, og på niveau 2, *j*, men det gør β-koefficienterne ikke. Det forudsættes med andre ord, at en enkelt regressionslinje kan give en tilfredsstillende beskrivelse af data. En OLS estimering af β-koefficienterne i denne model ville give en linje nogenlunde svarende til den, der ses i figur 3.

Figur 3. Sammenhæng mellem to variable analyseret på disaggregeret niveau



Denne analyse ville vise den positive sammenhæng mellem individers indtægt og sundhedstilstand, men den ville for det første ikke vise, at den modsatte sammenhæng optræder på aggregeret niveau. Den ville for det andet heller ikke vise den rumlige autokorrelation i datasættet, som udspringer af, at enheder på niveau 1 ikke er udtrukket uafhængigt af hinanden, og som kommer til udtryk ved, at observationerne fra den ene kommune, stjernerne, har en systematisk tendens til at ligge over regressionslinjen og dermed have positive residualer, mens observationerne i den anden kommune, cirklerne, generelt har negative residualer. For det tredje ville den ikke tage højde for den heteroskedasticitet, den uensartehed i residualerne, som kommer af, at observationer med høje X-værdier generelt ligger længere fra regressionslinjen end observationer med lave X-værdier.

Dette er helt generelle problemer ved OLS-estimering af den klassiske lineære regressionsmodel, når data er hierarkisk strukturerede:

(1) OLS-estimaterne korrigerer ikke for den rumlige autokorrelation, som opstår, hvis observationer på niveau 1 ikke er uafhængige af hinanden, fordi de er grupperet i enheder på niveau 2. I det tilfælde vil OLS-estimater af regressionskoefficienter stadig være konsistente, men de estimerede standardfejl vil ofte være for små, og der vil derfor være en tendens til, at man for ofte konkluderer, at en sammenhæng er signifikant (type I-fejl) (Rabe-Hesketh og Skrondal, 2005: 34).[1]

(2) Derudover baseres OLS-estimaterne på den forudsætning, at residualerne er homoskedastiske (har ensartet varians). Men hvis sammenhængen mellem to variable varierer mellem grupperne på niveau 2, som i det konstruerede eksempel i figur 3, så opstår der heteroskedasticitet, og estimeringen af standardfejlene bliver unøjagtig.

(3) Den klassiske lineære regressionsmodel tager heller ikke højde for, at samme variabel kan have forskellige effekter på forskellige niveauer, og den rummer dermed en risiko enten for de ovenfor nævnte økologiske fejlslutninger eller for individualistiske fejlslutninger, hvor det konkluderes, at en sammenhæng på individniveau også gælder på et højere hierarkisk niveau.

(4) Ud over disse statiske problemer indebærer brugen af den klassiske lineære regressionsmodel også, at der er forhold på de enkelte niveauer, som kan være af substantiel interesse, men som ikke kan analyseres. Den side af sagen vender vi tilbage til i det empiriske eksempel i næste afsnit.

Men først skal vi se, hvordan en multilevel-model kan opbygges. Det grundlæggende princip ligger i at tillade den mulighed, at der er variation ikke kun mellem individer på niveau 1, i, men også mellem grupper på niveau 2, j, (eller højere). Første trin består i at lade skæringspunktet, β_0, variere mellem grupperne på niveau 2, j:

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij} \qquad (3.)$$

Når skæringspunktet varierer mellem grupperne, er det muligt at opdele det i et gennemsnit og et residualled, som repræsenterer variation omkring gennemsnittet:

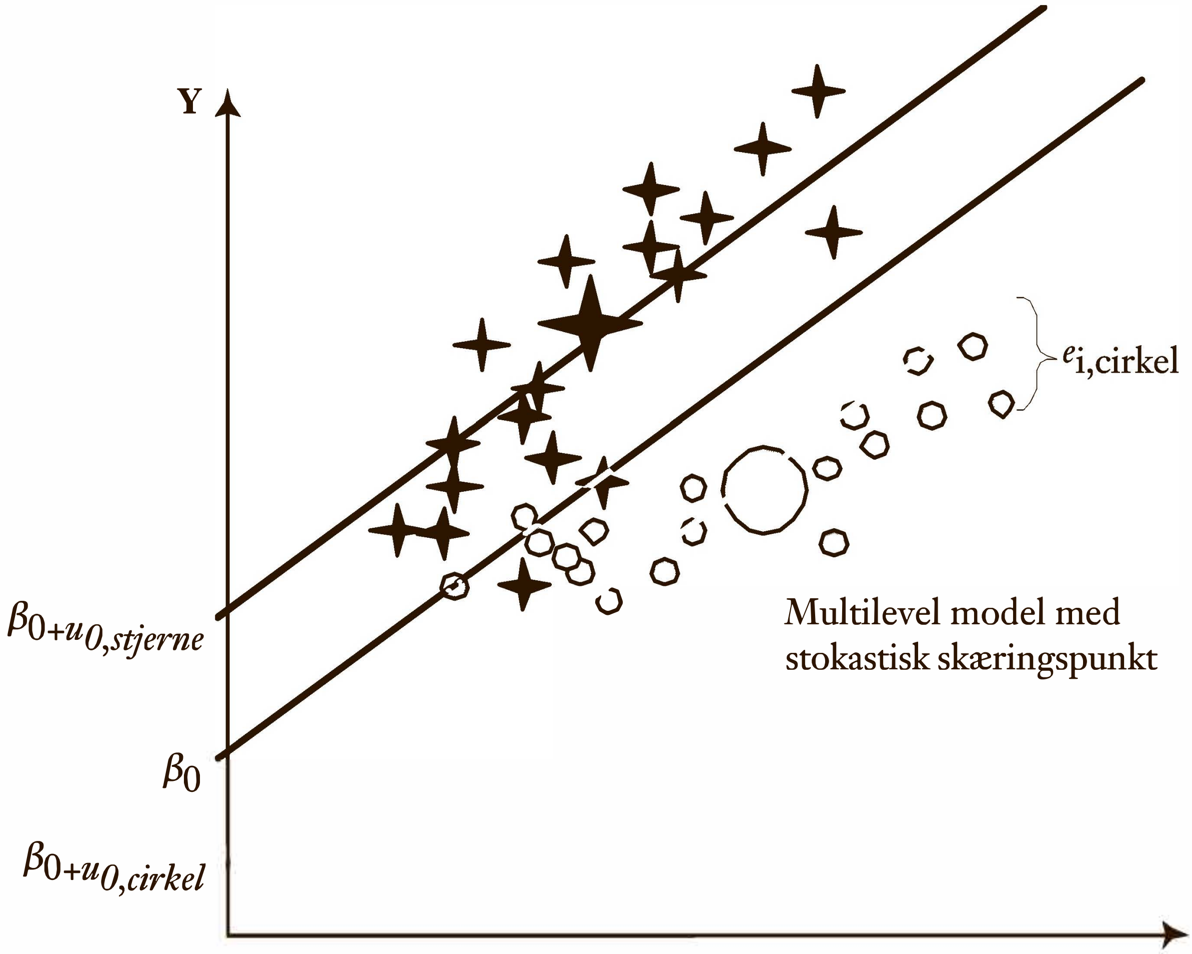
$$\beta_{0j} = \beta_0 + u_{0j} \qquad (4.)$$

Sættes denne ligning (4.) ind i den ovenfor (3.), fås følgende model:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + e_{ij} \qquad (5.)$$

Dermed fås en model, som angiver et fælles skæringspunkt, en fælles hældning, en variation omkring skæringspunktet og en residual variation. Grafisk kan det illustreres som i figur 4.

Figur 4. Multilevel-model med stokastisk skæringspunkt



Her genfindes den fælles regressionslinje fra figur 3, men nu suppleret med parallelle linjer for de enkelte grupper på niveau 2. u_{0j} angiver afstanden fra det fælles skæringspunkt og til de enkelte gruppers linje. Den kan fortolkes som hovedeffekten af den pågældende gruppe. e_{ij} angiver afstanden fra den enkelte observation til gruppens linje. På den måde tages der højde for punkt 1 ovenfor, den rumlige autokorrelation mellem observationer inden for en gruppe, fordi u_{0j} estimeres, så summen af residualerne, e_{ij}, for den enkelte gruppe bliver 0.

Der er to måder at estimere denne model på. u_{0j} kan opfattes som *faste* (*fixed*) parametre i en model, der estimerer et parameter for hver gruppe. Men hvis man har mange grupper, for eksempel mange kommuner, vil det hurtigt give uoverskueligt mange parameterestimater. Og hvis ikke man er specielt interesseret i de enkelte grupper (kommuner), men i stedet opfatter grupperne som en tilfældig stikprøve af en større population (eller opfatter populations-data som en stikprøve fra en stokastisk verden[2]), som man ofte gør med enheder på niveau 1 (individer), så vil det være mere relevant at estimere u_{0j} som stokastiske variable, der udtrykker den uforklarede gruppeeffekt (for en diskussion af de to tilgange se Snijders og Bosker, 1999: 41-45; se også Goldstein, 2003: 15; Rabe-Hesketh og Skrondal, 2005: 61).

Andet trin i modelbygningen består i at tillade ikke bare skæringspunktet, men også hældningen på regressionslinjerne at variere mellem grupperne. Det markeres ved at sætte fodtegnet j på hældnings- eller regressionskoefficienten, β_1:

$$y_{ij} = \beta_0 + \beta_{1j}x_{ij} + u_{0j} + e_{ij} \qquad (6.)$$

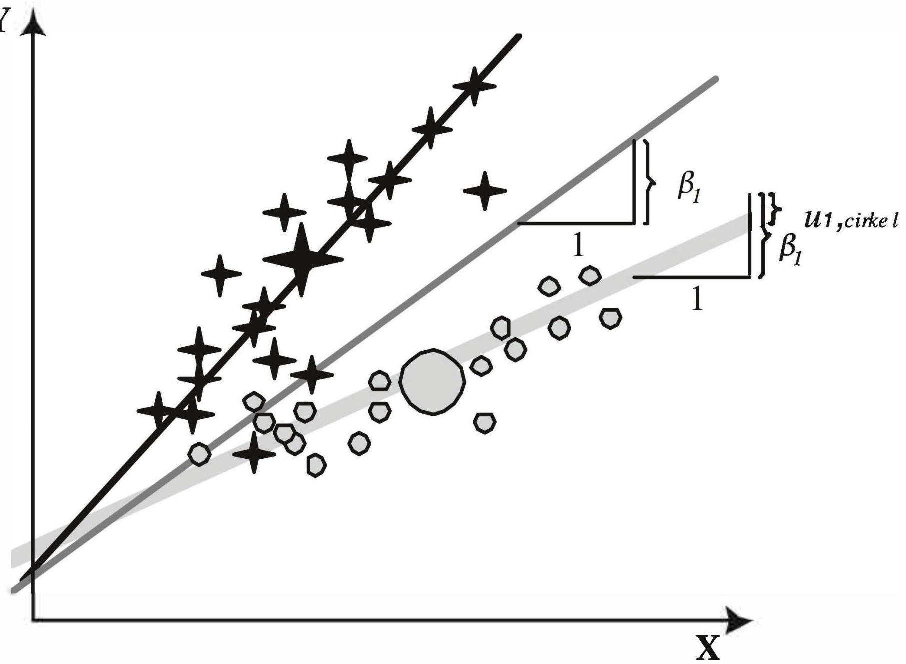
Dermed kan regressionskoefficienten ligesom skæringspunktet opdeles i en fast og en stokastisk del:

$$\beta_{1j} = \beta_1 + u_{1j} \qquad (7.)$$

Sættes denne ligning (7.) ind i multilevel-modellen ovenfor (6.), får vi:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + u_{1j}x_{ij} + e_{ij} \qquad (8.)$$

Figur 5. Multilevel-model med stokastiske skæringspunkter og stokastiske hældninger



Her kan β_1 fortolkes som den gennemsnitlige hældningskoefficient, mens u_{1j} kan betragtes som en stokastisk effekt af X, som varierer mellem grupperne. Grafisk kan det illustreres som i figur 5. Nu har hver gruppe ikke bare sit eget skæringspunkt men også sin egen hældning. u_{1j} er afstanden mellem den gennemsnitlige hældning og den enkelte gruppes hældning. Som det fremgår, modellerer denne model eksplicit en potentiel heteroskedasticitet (jf. punkt 2 ovenfor), som opstår, hvis variationen omkring regressionslinjen afhænger af værdien af *X*. Det gør modellen ved at lade en del af residualerne, u_{1j}, være afhængige af *X*.

Hvis man vil undersøge, om X har en anden effekt på gruppeniveau, kan man i et tredje trin opstille en model, hvor man skelner mellem variation inden for og mellem grupper. Det kan fx gøres ved at operere med to forklarende variable, henholdsvis $\bar{x}_j$, som er gruppens gennemsnit på X, og $(x_{ij} - \bar{x}_j)$, som er den enkelte enheds afvigelse fra gruppens gennemsnit. Det giver denne model:

$$y_{ij} = \beta_0 + \beta_1(x_{ij} - \bar{x}_j) + \beta_2 \bar{x}_j + u_{0j} + u_{1j}x_{ij} + e_{ij} \qquad (9.)$$

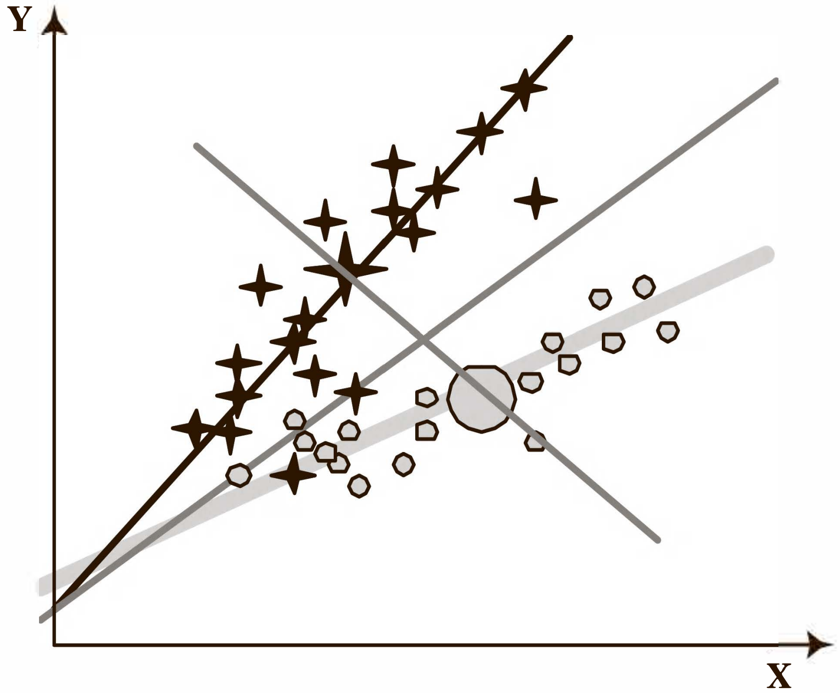
I denne model kan β_1 fortolkes som den gennemsnitlige effekt af X inden for grupperne (effekten af indtægt på sundhedstilstanden for borgere inden for samme kommune), mens β_2 kan tolkes som effekten af X mellem grupper (effekten af kommunernes gennemsnitlige indtægt på kommunernes gennemsnitlige sundhedstilstand). Grafisk kan modellen illustreres som i figur 6, hvor den faldende regressionslinje ligesom i figur 2 markerer mellem-grupper-effekten, mens de stigende regressionslinjer markerer inden-for-grupper-effekten.

For alle disse multilevel-modeller antages det som udgangspunkt,

- at residualerne, e_{ij}, u_{0j} og u_{1j}, er normalfordelte med et gennemsnit på 0, $E(e_{ij}) = E(u_{0j}) = E(u_{1j}) = 0$,
- at de er homoskedastiske med varianserne $\text{var}(e_{ij}) = \sigma_e^2$, $\text{var}(u_{0j}) = \sigma_{u0}^2$ og $\text{var}(u_{1j}) = \sigma_{u1}^2$,
- at kovariansen mellem u_{0j} og u_{1j} er $\text{cov}(u_{0j}, u_{1j}) = \sigma_{u01}$, og
- at e_{ij} er ukorreleret med u_{0j} og u_{1j} (Snijders og Bosker. 1999: 45-46; se også Rabe-Hesketh og Skrondal, 2005: 35-36, 63-64; Goldstein, 2003: 15-16).[3]

Under disse forudsætninger kan maximum likelihood- eller restricted maximum likelihood-estimatorerne anvendes (for detaljer, se Goldstein, 2003: 19-21; jf. Rabe-Hesketh og Skrondal, 2005: 16).

Figur 6. Multilevel-model med stokastiske skæringspunkter, hældningskoefficienter og separat mellem-grupper-effekt



For den simpleste multilevel-model med stokastisk skæringspunkt (se ligning 5.) kan variansen på *Y* dekomponeres som summen af varianserne inden for og mellem grupperne, fordi det forudsættes, at residualerne u_{0j} og e_{ij} er uafhængige. Det vil sige, at var(*Y*) = var(e_{ij}) + var(u_{0j}) = $\sigma_e^2 + \sigma_{u0}^2$. Den andel af variansen på *Y*, som er mellem grupperne (kommunerne), kan derfor bestemmes som

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0}^2} \qquad (10.)$$

ρ (rho) kan med andre ord fortolkes som niveau 2's andel af variansen, også kaldet varians-del-koefficienten (*variance partition coefficient*). ρ har imidlertid også en række andre meget nyttige fortolkninger. ρ kan således fortolkes analogt til R^2 som den del af den samlede variation i *Y*, som skyldes eller kan forklares af grupperne på niveau 2.

Hvis datasættet består af gentagne målinger på samme individer, hvor niveau 1 er målinger grupperet inden for individer på niveau 2, da er σ_{u0}^2 variansen mellem individernes „sande“ score, og σ_e^2 variansen på målingsfejlene i de enkelte observationer. I det tilfælde kan ρ fortolkes som *test-retest* reliabiliteten, idet høj reliabilitet, forstået som lille variation mellem de enkelte målinger af samme individ, vil få σ_e^2 til at gå mod 0 og ρ til at gå mod 1 (Rabe-Hesketh og Skrondal 2005: 7-8).

Endelig kan ρ også fortolkes som intra gruppe-korrelationen (intra kommune-korrelationen), forstået som et mål for, hvor meget enheder (borgere) inden for samme gruppe (kommune) korrelerer. Det ses ved, at kovariansen mellem to enheder, *i* og *i´*, som hører til samme gruppe, *j*, er

$$\text{cov}(y_{ij}, y_{i'j}) = \text{cov}(u_{0j} + e_{0ij}, u_{0j} + e_{0i'j}) = \text{cov}(u_{0j}, u_{0j}) = \sigma^2_{u0}, \qquad (11.)$$

fordi niveau 1 residualerne forudsættes at være uafhængige. Korrelationen mellem disse to enheder inden for samme gruppe (intra-gruppe-korrelationen) udgøres af kovariansen delt med produktet af deres standardafvigelser:

$$\text{cor}(y_{ij}, y_{i'j}) = \frac{\text{cov}(y_{ij}, y_{i'j})}{\sqrt{\text{var}(y_{ij})}\sqrt{\text{var}(y_{i'j})}} = \frac{\sigma^2_{u0}}{\sqrt{\sigma^2_{u0} + \sigma^2_{e0}}\sqrt{\sigma^2_{u0} + \sigma^2_{e0}}} = \frac{\sigma^2_{u0}}{\sigma^2_{u0} + \sigma^2_{e0}} = \rho \qquad (12.)$$

(Rabe-Hesketh og Skrondal, 2005: 8; Snijders og Bosker, 1999: 46).

ρ kan derfor også tages som et mål for, hvor megen rumlig autokorrelation, der er i datasættet, og kun hvis ρ er lille, er OLS-estimaterne korrekte.[4]

Hvis multilevel-modellen ikke kun har stokastiske skæringspunkter, men også stokastiske hældningskoefficienter (som i ligning 8.) er intra gruppe-korrelationen ikke den samme som varians-del-koefficienten (for detaljer, se Goldstein, 2003: 16-17, 68-69), og intra gruppe-korrelationen afhænger nu af *X*. Den har derfor ikke de samme fortolkningsmuligheder (Rabe-Hesketh og Skrondal, 2005: 65-66).

Endelig skal det nævnes, at de ovenstående modeller kan udvides til at indbefatte tre eller flere niveauer. I så fald opdeles residualerne i tre dele markeret med et ekstra fodtegn, *k*. I en model med stokastiske skæringspunkter og hældninger, svarende til toniveau-modellen i ligning 8 ovenfor, ser det således ud:

$$y_{ijk} = \beta_0 + \beta_1 x_{ijk} + v_{0k} + v_{1k}x_{ijk} + u_{0jk} + u_{1jk}x_{ijk} + e_{ijk} \qquad (13.)$$

v_{0k} og v_{1k} er den stokastiske variation omkring henholdsvis skæringspunktet og hældningen på niveau 3 med varianserne var(v_{0k}) = σ^2_{v0} og var(v_{1k}) = σ^2_{v1}. Bemærk at den stokastiske variation på niveau 2, *u*, nu varierer på både niveau 2 og 3 og tilsvarende for variationen på niveau 1.

I næste afsnit skal vi se en empirisk anvendelse af sådan en treniveau-model.

Kommuners betydning for skoleelevers karakter

Formålet med den følgende analyse er at illustrere mulighederne i multilevel-modeller gennem en analyse af danske kommuners betydning for deres elevers faglige færdigheder – målt ved elevernes gennemsnitskarakter i skriftlig matematik og dansk ved folkeskolens afgangsprøve i 2002. Der hersker blandt danske politikere en stor interesse for at øge folkeskoleelevers faglige udbytte af deres undervisning. Men det er et spørgsmål, hvor stor indflydelse politikere overhovedet har på det forhold (Andersen, 2006). I Danmark er en stor del af styringen af folkeskolerne decentraliseret til kommunerne. Hvis de forskelle, som eksisterer mellem de enkelte kommuners skolepolitik, har en væsentlig betydning for elevernes faglige præstationer, så skulle man forvente signifikante og væsentlige forskelle mellem karakterniveauerne kommunerne imellem. Med en almindelig OLS-model eller med en faste effekter (*fixed effects*) model, der estimerer et parameter for hver kommune, vil det ikke være muligt at få et samlet mål for, om kommunerne gør en signifikant forskel på elevernes karakterer. Det er derimod muligt med en multilevel-model.

Til formålet benyttes et datasæt, som består af oplysninger om alle elevers karakterer ved folkeskolens afgangsprøve i 2002, og som via CPR-numre kobler disse elever med oplysninger om deres forældres sociale og økonomiske baggrund. Formålet med Folkeskolen er væsentligt bredere, end hvad der kan måles med karakterer. Men i forhold til alternative måder at måle skoleelevers udbytte af deres skolegang, må karaktererne betragtes som relativt valide og reliable mål for en central del af målet med undervisningen, de faglige færdigheder. Eftersom elevernes sociale og økonomiske baggrund må forventes at have stor betydning for deres præstationer ved afgangsprøverne, er det væsentligt at kontrollere for sådanne forhold. I dette tilfælde inddrages detaljerede oplysninger om forældres indkomst, uddannelse, bolig- og familieforhold. Der inddrages også en række kontrolvariable for forskelle mellem kommuner, som politikere ikke har nogen direkte indflydelse på (for nærmere oplysninger om og diskussion af variablene i datasættet, se Andersen og Serritzlew, 2007). Den resterende, uforklarede variation mellem kommuner, som kan estimeres med en multilevel-model, kan derfor tolkes som et øvre mål for, hvor megen betydning forskelle i de kommunale skolepolitikker har.

Vi begynder med en simpel multilevel-model med stokastiske skæringspunkter (som ligning 13, men uden stokastiske hældningskoefficienter). Resultatet ses i tabel 1.

Som det fremgår, har variablene på elevniveau alle sammen en signifikant sammenhæng med elevernes gennemsnitskarakter i skriftlig dansk og matematik. Specielt ses det, at piger scorer højere end drenge og indvandrere, og efterkommere af indvandrere lavere end øvrige danskere. Børn, som bor hos begge forældre, scorer også højere end alle andre typer. Jo længere uddannelse forældrene har, desto højere karakterer opnår eleverne, og en større (ejer-)bolig, højere indtægt og formue for forældrenes vedkommende har også en positiv sammenhæng med elevernes karakterer. Til gengæld har kommune-variablerne indbyggertal, beskatningsgrundlag og urbaniseringstype ikke nogen signifikant sammenhæng med karaktererne (med undtagelse af kommuner i Køben-

Tabel 1. Gennemsnitkarakter i skriftlig dansk og matematik forklaret ud fra variable på elev- og kommuneniveau. Multilevel-model med stokastiske skæringspunkter

	Skr. dansk og matematik		
	Koeffi-cient		**Standard fejl**
ELEV-VARIABLE			
Pige	0,261	**	(0,012)
Indvandrer/efterkommer	-0,611	**	(0,028)
Forældres uddannelse			
(reference: Højest grundskoleuddannelse)			
Højest ungdomsuddannelse	0,397	**	(0,019)
Højest kort videregående uddannelse	0,799	**	(0,030)
En mellemlang eller lang videregående uddannelse	0,921	**	(0,023)
To mellemlange eller lange videregående uddannelser	1,380	**	(0,027)
Husstand			
(reference: Bor hos begge forældre)			
Bor hos mor i et nyt par	-0,183	**	(0,022)
Bor hos enlig mor	-0,071	**	(0,021)
Bor hos far i et nyt par	-0,132	**	(0,050)
Bor hos enlig far	-0,155	**	(0,040)
Bor ikke hos forældrene	-0,470	**	(0,070)
Boligtype (reference: ikke-ejerbolig)	0,232	**	(0,018)
Boligstørrelse (100 m²)	0,105	**	(0,014)
Forældres bruttoindkomst (100 000 kr.)	0,014	**	(0,002)
Forældres formue (100 000 kr.)	0,007	**	(0,001)

- fortsætter -

	Skr. dansk og matematik		
	Koeffi-cient		**Standard fejl**
KOMMUNE-VARIABLE			
Indbyggertal (10 000 kr.)	-0,001		(0,003)
Beskatningsgrundlag (1000 kr.)	0,001		(0,002)
Urbaniseringstype			
(reference: landkommune)			
Hovedstaden	0,071		(0,119)
Hovedstadens forstæder	0,130	**	(0,051)
Provinsbyer med >100 000 indbyggere	0,029		(0,086)
Provinsbyer med >10 000 indbyggere	0,0460		(0,031)
Konstant	6,726	**	(0,262)
STOKASTISKE EFFEKTER			
σ^2_{v0} (varians mellem kommuner)	0,007	*	(0,003)
σ^2_{u0} (varians mellem skoler)	0,066	**	(0,005)
σ^2_{e0} (varians mellem elever)	1,411	**	(0,010)
-2*loglikelihood	120087		
N	37408		

Note: Modellen er estimeret med maximum-likelihood ved hjælp af Iterative Generalized Least Squares (se Goldstein, 2003: 19-21, 49-51). * Signifikant på 0,05 niveauet, to-sidet test; ** signifikant på 0,01 niveauet.

havns forstæder, der ser ud til at opnå signifikant højere karaktergennemsnit end landkommuner, hvor største by har færre end 10.000 indbyggere).

Nederst i tabellen angives estimeringen af den stokastiske del af modellen. Konstanten (skæringspunktet), som elever, skoler og kommuner varierer omkring, når der kontrolleres for de andre variable i modellen, er estimeret til 6,73. Kommunerne fordeler sig om dette gennemsnit med en estimeret varians på 0,007, hvilket kun er signifikant på et 5 pct. signifikansniveau.[5] Der er med andre ord en vis variation mellem kommunerne, når man kontroller for denne række af variable, som politikere ikke umiddelbart har nogen indflydelse på. Variansen mellem skoler er dog ni gange større og mellem elever 200 gange større end variansen på kommuneniveau. Den klart største variation i karaktererne i skriftlig dansk og matematik ligger altså mellem elever inden for samme skole, og ikke mellem kommuner. Varians-dels-koefficienten ρ, giver et mål for kommunernes andel af den samlede varians. I dette tilfælde:

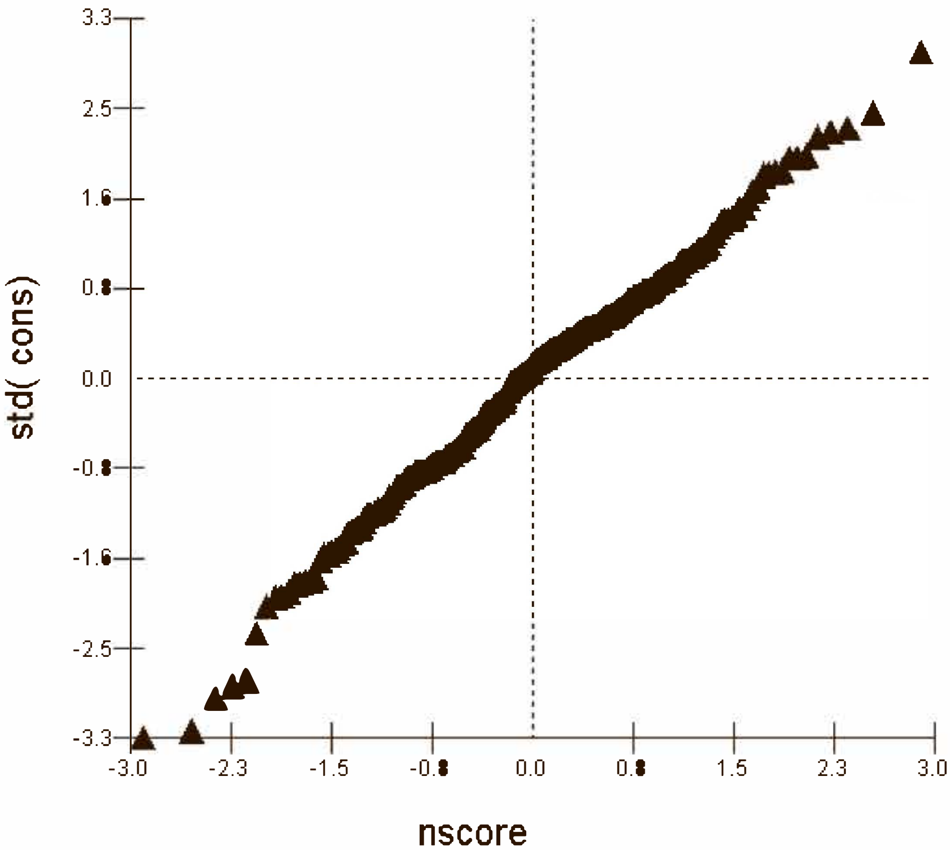
$$\rho = \frac{\sigma_{v0}^2}{\sigma_{v0}^2 + \sigma_{u0}^2 + \sigma_{e0}^2} = \frac{0{,}007}{0{,}007 + 0{,}066 + 1{,}411} = 0{,}0046 \qquad (14.)$$

Det er altså kun omkring 0,5 pct. af variansen, som ligger mellem kommuner. Det indikerer, at variationen i kommunernes styring af folkeskolerne ikke har stor betydning for elevernes faglige præstationer.

Modellen baseres som nævnt på den forudsætning, at residualerne på de enkelte niveauer er normalfordelte. Det kan testes ved at plotte rangordnede residualer mod tilsvarende punkter på en normalfordelingskurve. Hvis normalitetsforudsætningen er overholdt, skal resultatet blive en tilnærmelsesvist ret linje (Rasbash et al., 2003: 37-38).

Figur 7 viser resultatet af den analyse for residualerne på kommuneniveau.

Figur 7. Normalitetsplot af residualerne på kommuneniveau



Tilsvarende analyser er lavet for residualerne på elev- og skoleniveau. I alle tre tilfælde ser forudsætningen ud til at være rimelig.

Selvom der ikke er megen variation i karaktergennemsnittene mellem kommuner, kan der jo godt tænkes at være enkelte kommuner, som klarer sig signifikant bedre end gennemsnittet. Kommuneniveau-residualet, v_{0k}, udtrykker, hvor meget den enkelte kommune ligger over eller under kommunernes gennemsnitlige skæringspunkt, og det kan derfor betragtes som et mål for, hvor meget bedre eller dårligere, den enkelte kommune klarer sig i forhold til modellens forudsigelse.

Når den enkelte kommunes residual skal estimeres, er der to muligheder: „rå“ eller „krympede“ residualer. De rå residualer, $\tilde{y}_{ijk}$, beregnes simpelthen som forskellen mellem den observerede værdi, y_{ijk}, og den forudsagte værdi, $\hat{y}_{ijk}$:

$$\tilde{y}_{ijk} = y_{ijk} - \hat{y}_{ijk} \qquad (15.)$$

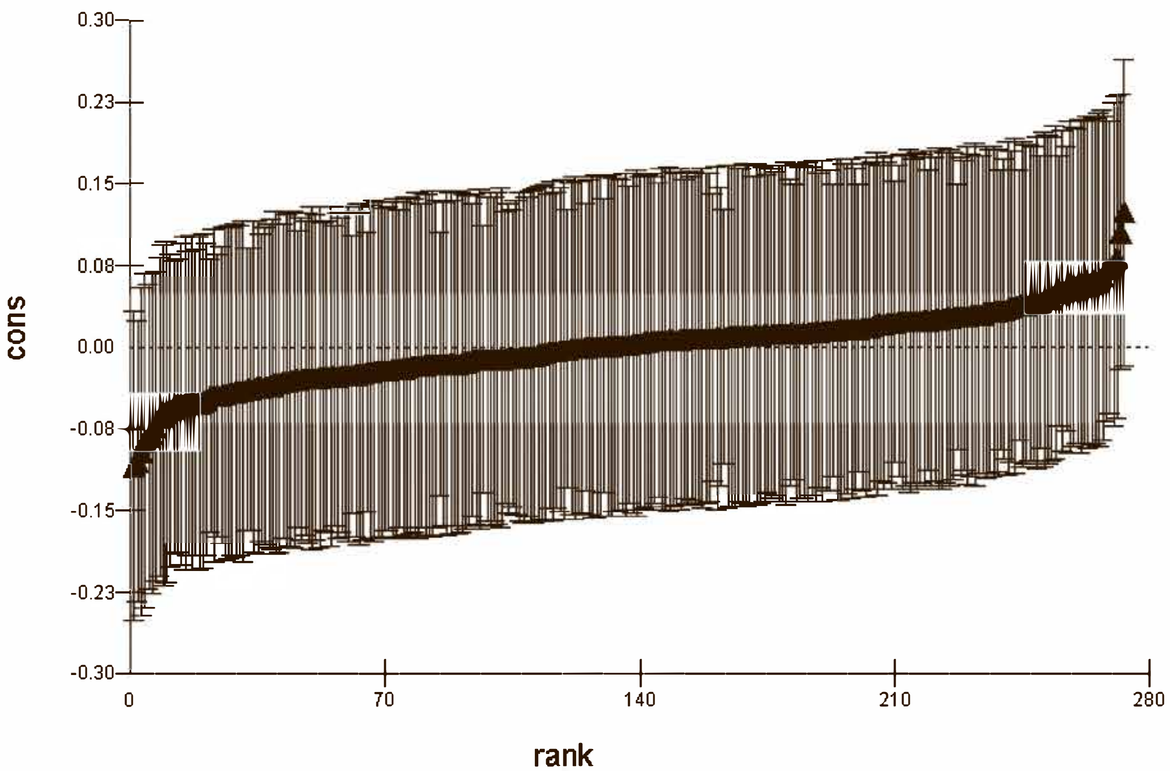
Men hvis man ikke blot er interesseret i at beregne den enkelte kommunes residual i datasættet, men betragter grupperne, i dette tilfælde kommunerne, som en stikprøve af en større population (eller som en stikprøve fra en stokastisk verden, jf. note 2), så kan der være grund til at benytte de krympede residualer. Her vægtes de rå residualer i forhold til, hvor mange enheder der er i den enkelte gruppe (elever i kommunen), og i forhold til hvor stor variationen på gruppeniveau er i forhold til variationen på individniveau, som det sker i denne formel:

$$\hat{v}_{0k} = \frac{\sigma^2_{v0}}{\sigma^2_{v0} + (\sigma^2_{u0} + \sigma^2_{e0})/n_k}\tilde{y}_k \qquad (16.)$$

hvor n_k er antallet af enheder i den k'te niveau 3-gruppe, og gennemsnittet af de rå residualer i samme gruppe. Som det fremgår, er krympefaktoren altid mindre end 1 og større end 0, hvilket betyder, at den altid gør den krympede residual mindre end den rå residual. En begrundelse for at benytte de krympede residualer er, at hvis man havde en kommune, som ikke var med i datasættet, ville bedste gæt på dens residual være den gennemsnitlige residual. Hvis man tilsvarende har en kommune med meget få observationer, så har man et meget upræcist estimat af dens parametre. Derfor får man et bedre gæt ved at trække den type kommuner i retning af gennemsnittet. Det er, hvad den krympede residual gør. Krympefaktoren er den bedste lineære unbiased prædiktor (BLUP) for v_{0j} – i det her tilfælde, om en kommune virkelig klarer sig bedre eller dårligere, end man skulle forvente ud fra dens elev- og skolesammensætning (for nærmere diskussioner, se Rabe-Hesketh og Skrondal, 2005: 22; Goldstein, 2003: 22; Rasbash et al., 2003: 33).

For at undersøge, om enkelte kommuner klarer sig signifikant bedre end gennemsnittet, er disse krympede residualer og deres standardfejl estimeret. På den baggrund er kommunerne i figur 8 rangordnet i forhold til, hvor godt de klarer sig, og der er lagt et 95 pct. konfidensinterval omkring hver kommune. Hvis en kommunes konfidensinterval ligger over (under) gennemsnittet, der er 0, så viser det, at der er mindre end 5 pct. sandsynlighed for, at den kommune ikke rent faktisk klarer sig bedre (værre) end dens forudsagte gennemsnitskarakter (som er et gennemsnit justeret for kontrolvariablene). Som det fremgår, er der ikke nogen kommune, som ud fra den beregning klarer sig signifikant bedre eller værre end gennemsnittet.

Figur 8. Klarer nogen kommuner sig signifikant bedre eller værre end gennemsnittet?



Note: Kommunernes krympede residualer i forhold til deres forudsagte score. 95 pct. konfidensintervaller.

Det er imidlertid vigtigt at være opmærksom på, at med disse konfidensintervaller, må vi forvente, at vi for hver 20. kommune (5 pct.) identificerer en signifikant residual alene af tilfældige årsager. Konfidensintervallet kan altså kun bruges til at teste, om en bestemt kommune, som vi i forvejen forventer klarer sig bedre end gennemsnittet, faktisk også gør det. Hvis vi ville bruge plottet til med 95 pct. sikkerhed at identificere signifikant gode kommuner, skulle konfidensintervallerne justeres for, at vi på én gang sammenligner 275 kommuner. Der eksisterer en hel litteratur om, hvordan sådanne multiple sammenligninger skal foretages, men generelt gælder det, at de giver meget bredere konfidensintervaller, og i dette tilfælde ville alle kommuner være langt fra at gøre en signifikant forskel (se fx Benjamini og Hochberg, 1995).[6]

Et andet spørgsmål er, om elevernes sociale baggrund har forskellig betydning for deres faglige præstationer i forskellige kommuner. For at undersøge det spørgsmål, er de forskellige variable for den sociale baggrund samlet i et formativt indeks (for detaljer om indekset, se appendiks). Indekset er anvendt i en multilevel-model først med stokastiske skæringspunkter og derefter med både stokastiske skæringspunkter og hældningskoefficienter. Den sidste model svarer til ligning 10, men suppleret med kontrol for elevens køn og de kommunale faktorer anvendt i tabel 1. Resultatet af at estimere de to modeller ses i tabel 2 i henholdsvis første og anden søjle.

Søjle 1 svarer til tabel 1, blot er de sociale baggrundsvariable samlet i et indeks. Modellen bekræfter en stærkt signifikant sammenhæng mellem social baggrund og karakterer og tilsvarende mellem køn og karakterer. Kommunevariablerne har ikke nogen stærk sammenhæng med karaktererne, og der er en signifikant, men lille variation omkring kommunernes gennemsnitlige skæringspunkt. (Kommunernes andel af den samlede variation, varians-dels-koefficient, er blot 0,7 pct.). I søjle 2 tillades effekten af elevernes sociale baggrund at variere både mellem skoler og mellem kommuner. Det påvirker ikke de faste estimater i nævneværdig grad, men i den nederste del af tabellen kan vi se, at der er en signifikant variation ikke blot i kommunernes og skolernes skæringspunkter, men også i deres hældningskoefficienter for indekset for social baggrund.

Det kan testes, om tilføjelsen af stokastiske hældningskoefficienter samlet set er signifikant. Når man sammenligner to multilevel-modeller er det i følge Goldstein (2003: 25) mest korrekt at benytte en *likelihood ratio test*. Det gøres ved at tage differencen mellem de to modellers -2*loglikelihood estimater (i dette tilfælde 55), som har en chi-i-anden-fordeling med antal frihedsgrader svarende til antallet af yderligere estimater (i dette tilfælde 4). Det giver en sandsynlighed <0,000 og bekræfter, at der er signifikant forskel mellem de to modeller og dermed signifikant variation i effekten af den sociale baggrund.

Det fremgår også, at der er en signifikant negativ kovarians mellem skæringspunktet og hældningskoefficienten på både skole- og kommuneniveau. Kovariansen har også en substantiel fortolkning: Hvis man forestiller sig, at man tegnede en regressionslinje for hver kommunes sammenhæng mellem social baggrund og karakterer (som i figur 5), så ville kommuner med et højt skæringspunkt (karaktergennemsnit for værdien 0 på indekset for social baggrund) have en lav hældning (og omvendt). Kommunernes regressionslinjer

Tabel 2. Har elevers sociale baggrund forskellig betydning for deres karakterer i forskellige kommuner? Multilevel-modeller med stokastiske skæringspunkter (søjle 1) og stokastiske hældningskoefficienter (søjle 2).

	Skr. dansk og matematik			
	(1)		(2)	
	Koef.	**Std.fejl**	**Koef.**	**Std.fejl**
ELEV-VARIABLE				
Elevens sociale baggrund (index)	0,374 **	(0,005)	0,364 **	(0,007)
Køn (reference: dreng)	0,260 **	(0,013)	0,260 **	(0,013)
KOMMUNE-VARIABLE				
Indbyggertal (10 000 kr.)	-0,004	(0,003)	-0,002	(0,004)
Beskatningsgrundlag (1000 kr.)	0,003	(0,002)	0,003	(0,002)
Urbaniseringstype				
(reference: landkommune)				
Hovedstaden	0,203	(0,133)	0,243	(0,132)
Hovedstadens forstæder	0,122 *	(0,055)	0,135 **	(0,055)
Provinsbyer med >100 000 indbyggere	0,131	(0,101)	0,136	(0,101)
Provinsbyer med >10 000 indbyggere	0,053	(0,034)	0,053	(0,034)
Konstant	6,034 **	(0,283)	6,119 **	(0,279)

- fortsætter -

ville altså samle sig mere og mere, jo højere social baggrund eleven har (modsat figur 5). Det vil sige, at kommunerne gør en større forskel for elever med svag social baggrund end for elever med høj social baggrund (og det samme gælder skolerne). Korrelationen mellem skæringspunktet og interceptet på kommuneniveau er så høj som:

$$\mathrm{cor}(\mathit{skæring}, \mathit{hældning}) = \frac{\mathrm{cov}(v_{0k}, v_{1k})}{\sqrt{\sigma_{v0}^2 * \sigma_{v1}^2}} = \frac{-0{,}0099}{\sqrt{0{,}052 * 0{,}0024}} = -0{,}89$$

	Skr. dansk og matematik			
	(1)		(2)	
	Koef.	**Std.fejl**	**Koef.**	**Std.fejl**
STOKASTISKE EFFEKTER				
Kommuneniveau				
σ^2_{v0} (skæringspunkt)	0,011 **	(0,004)	0,0520 **	(0,0161)
σ_{v01} (kovarians)			-0,0099 **	(0,0036)
σ^2_{v1} (hældningskoeff., soc. baggrund)			0,0024 **	(0,0009)
Skoleniveau				
σ^2_{u0} (skæringspunkt)	0,072 **	(0,006)	0,1299 **	(0,0203)
σ_{u01} (kovarians)			-0,0152 **	(0,0049)
σ^2_{u1} (hældningskoeff., soc. baggrund)			0,0035 **	(0,0014)
Elevniveau				
σ^2_{e0} (varians mellem elever)	1,480 **	(0,011)	1,4719 **	(0,0011)
-2*loglikelihood	124517		124462	
N	38208		38208	

Note: Gennemsnitkarakter i skriftlig dansk og matematik forklaret ud fra variable på elev- og kommuneniveau. Modellen er estimeret med maximum-likelihood ved hjælp af Iterative Generalized Least Squares (se Goldstein, 2003: 19-21, 49-51). * Signifikant på 0,05 niveauet, to-sidet test; ** signifikant på 0,01 niveauet.

Konklusionen fra den første del af multilevel-analysen modificeres altså, når de stokastiske hældningskoefficienter introduceres i modellen. Godt nok er der ikke megen variation i karakterniveauet mellem kommunerne, og der er ikke nogen kommune, som klarer sig signifikant bedre eller dårligere end gennemsnittet. Men forskellen mellem kommunerne er større for elever med svag social baggrund end for elever med stærk social baggrund – efter at der er kontrolleret for kommunevariable som urbaniseringstype og beskatningsgrundlag (som i øvrigt med en enkelt undtagelse ikke korrelerer signifikant med karaktergennemsnittet).[7] En forklaring på det kunne være, at hvis elever har en stærk social baggrund, så skal deres forældre under alle omstændigheder nok sørge for, at de får lært de ting, der skal læres, mens det omvendt gælder, at hvis eleverne har en svag social baggrund, så har kommunernes og skolernes indsats en stor betydning for elevernes faglige udbytte af deres skolegang.

Sammenfatning og supplerende litteratur

Sammenfattende kan man sige, at ved at benytte multilevel-modeller til at analysere karakterdataene var det muligt at undersøge nogle substantielle og politologisk set højst relevante spørgsmål, som det ikke ville have været muligt at besvare med almindelig OLS-modellering. Multilevel-modellerne gjorde det muligt at undersøge, om kommuner med deres forskellige skolepolitikker som sådan gør en forskel på elevernes faglige udbytte af deres skolegang – og også om de gør en større forskel for bestemte grupper af elever.

Mere generelt gælder det, at multilevel-modeller tilbyder både nogle statistisk-tekniske og nogle substantielle fordele, når ens data er struktureret hierarkisk (jf. Jones og Duncan, 1998). Statistisk-teknisk tilbyder de for det første en måde at håndtere rumlig autokorrelation, som opstår, hvis enheder inden for en gruppe er korreleret indbyrdes. Det sker ved at opdele residualledet i niveauer og dermed estimere variationen omkring gruppernes skæringspunkt. For det andet kan de håndtere nogle former for heteroskedasticitet, som opstår, hvis variansen på residualerne afhænger af de forklarende variable. Det gør de ved at tillade, at de forklarende variable har forskellig sammenhæng med den afhængige variabel i de enkelte grupper. Endelig kan de håndtere variable på flere niveauer, og de undgår dermed den risiko for økologiske eller individualistiske fejlslutninger, som opstår, hvis man er tvunget til enten at aggregere eller disaggregere hierarkisk ordnede data.

Substantielt tilbyder multilevel-modeller en mulighed for at analysere effekten af den kontekst, som enheder på niveau 1 agerer inden for. Det blev illustreret med en undersøgelse af kommuners betydning for skoleelevers karakterer ved afgangsprøverne i skriftlig dansk og matematik. Her var det muligt at vise, at der var signifikante forskelle mellem kommunerne, men også at det kun var en meget lille del af den samlede variation, som skyldtes kommunerne. Endvidere kunne man se, at effekten af elevernes sociale baggrund var forskellig i forskellige kommuner, og endda at der var en generel tendens til, at kommuner gør en større forskel for elever med svag social baggrund. Endelig kan multilevel-modeller også bruges til at estimere, hvordan de enkelte kommuner klarer sig i forhold til deres forudsagte performance. Her var det ikke muligt at identificere nogen kommune, som klarede sig signifikant bedre eller værre end det forudsagte gennemsnit.

Og dette er endda bare nogle af mulighederne i multilevel-modeller. De kan også anvendes til logistiske regressioner af en binær afhængig variabel, til regressioner med flere afhængige variable, til longitudinale studier med mange manglende observationer, til faktoranalyse og til datasæt, hvor enhederne er medlemmer af forskellige grupper på kryds og tværs. Der findes efterhånden en hel del litteratur, som beskriver både de statistiske og de praktiske sider af metoden. En god, kort, ikke-matematisk introduktion til modellerne findes i Jones og Duncan (1998). Snijders og Bosker (1999) har skrevet en hel bog, som introducerer både basale og mere avancerede multilevel-modeller med en begrænset brug af matematik. En af de mest omfattende introduktioner, som samtidig redegør detaljeret for multilevel-modellernes statistiske grundlag, er skrevet af en af ophavsmændene, Harvey Goldstein (2003). En diskussion af de nyeste statiske udviklinger inden for metoden findes i et særnummer af *Political Analysis* (Special Issue on Multilevel Modeling for Large Clusters: Autumn 2005; Vol. 13, No. 4).

Når det gælder det praktiske arbejde med modellerne, eksisterer der også en række gode tekster, som ofte samtidig introducerer til modellerne. Analyserne til denne artikel er blevet foretaget i det lille specialprogram til multilevel-modeller, MLwiN (ver.2.0).[8] Fordelen ved dette program er, at det er designet til at foretage mange af de analyser, som er særligt relevante for multilevel-modeller. Ulemperne er, at det tager lidt tid at sætte sig ind i et nyt statistikprogram, at det ikke er så let at arbejde med syntaksfiler, og at outputtet skal kopieres som grafik fra skærmen. En meget god introduktion til programmet findes i Rasbash et al. (2003).

Som nævnt i indledningen er multilevel-modeller også blevet integreret i de store statistiksoftwarepakker. Lolle (2003) giver en kort dansk introduktion til multilevel analyse i SPSS (ver. 11.5). Singer (1998) viser, hvordan man kan benytte SAS-kommandoen PROC MIXED, som har eksisteret længe, men efterhånden er blevet forbedret, så den nu også kan bruges til at estimere multilevel-modeller. Endelig er multilevel-modellering også blevet relativt brugervenligt i STATA (ver. 9). Rabe-Hesketh og Skrondal (2005) giver en god og omfattende introduktion, som samtidig giver mange nyttige tips til datahåndtering i STATA.

Hvis ens data er struktureret i hierarkier, er det derfor både relativt let at anvende multilevel-modellering, det giver mere korrekte estimeringer, og det åbner for besvarelse af nye forskningsspørgsmål.

Noter

1. Mere præcist gælder det, at hvis der kun er en enhed (et individ) på niveau 1 i hver niveau 2-gruppe (kommune), er OLS-estimaterne af standardfejlene korrekte. Men jo flere enheder der er i hver gruppe, desto mere underestimerer OLS (Goldstein, 2003: 23; se også Snijders og Bosker, 1999: 15-16, 40-41).

2. Om statistisk inferens på populationsdata, se Thomsen (1997) og King et al. (1994, specielt pp. 55-63).
3. Det er dog også muligt at modellere en heteroskedastisk kompleks variation på niveau 1 (se Goldstein 2003: 63-76).
4. Jf. note 1: Hvis der kun er en enhed i hver gruppe, er der ingen korrelation inden for grupperne, hvorved $\rho = 0$, og så kan OLS-estimatoren godt anvendes.
5. Variansestimaterne er kun tilnærmelsesvist normalfordelte, og derfor kan signifikanstest kun foretages ved store stikprøver (mange observationer), hvilket må siges at være tilfældet her med 37.408 elever (Goldstein, 2003: 24).
6. Det er i den forbindelse værd at bemærke, at politikere, forskere og tænketanke har vist stor interesse for at sammenligne og rangordne skoler ud fra deres karaktergennemsnit. Problemet med disse rangordninger er – ud over at de ikke alle sammen kontrollerer adækvat for elevernes sociale baggrund – at de ikke tager højde for den usikkerhed, der er forbundet med de estimerede placeringer på ranglisten. I England, hvor man er gået langt med at rangordne skoler, har nogle forskere gjort meget for at diskutere de statiske vanskeligheder, der knytter sig til sådanne ranglister (jf. diskussionen i Goldstein og Spiegelhalter, 1996).
7. Som et sidste led i analysen kunne man undersøge, om den sociale baggrund har forskellige effekter på elev- og skoleniveau. Den samlede model præsenteres ikke her, men resultatet viser, at effekten af indekset for social baggrund faktisk er noget stærkere mellem skoler (en partiel korrelation på 0,56 (standardfejl 0,02)) end inden for skoler (partiel korrelation på 0,36 (standardfejl 0,01)).
8. Programmet udvikles og sælges af forskningsenheden The Centre for Multilevel Modelling. Det kan bestilles på hjemmesiden www.cmm.bristol.ac.uk (12.8.2007).

Litteratur

Andersen, Simon Calmar (2006). „Styring gennem autonomi? Et forskningsprojekt om politisk styring af de danske skoler“, *Politica*, 36. årgang, nr. 4, pp. 426-445.

Andersen, Simon Calmar and Søren Serritzlew (2007). „The unintended effects of private school competition“, *Journal of Public Administration Research and Theory*, Vol. 17, No. 2, pp. 335-356.

Benjamini, Yoav and Yosef Hochberg (1995). „Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing“, *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 57, No. 1, pp. 289-300.

Goldstein, Harvey (2003). *Multilevel Statistical Models*, London: Arnold.

Goldstein, Harvey and David J. Spiegelhalter (1996). „League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance“, *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, Vol. 159, No. 3, pp. 385-443.

Jones, Kelvyn and Craig Duncan (1998). „Modelling Context and Heterogeneity: Applying Multilevel Models“, pp. 95-123 in Elinor Scharbrough and Eric Tanenbaum (eds.), *Research Strategies in the Social Sciences*, Oxford: Oxford University Press.

King, Gary, Robert O. Keohane and Sidney Verba (1994). *Designing Social Inquiry*, Princeton, NJ: Princeton University Press.

Lolle, Henrik (2003). *Multilevel analyse i SPSS 11.5*, Aalborg: Institut for Økonomi, Politik og Forvaltning, Aalborg Universitet.

Rabe-Hesketh, Sophia and Anders Skrondal (2005). *Multilevel and Longitudianl Modeling Using Stata*, College Station, TX: Stata Press.

Rasbash, Jon, Fiona Steele and William Browne (2003). *A user's guide to MLwiN*, London: Centre for Multilevel Modelling, Institute of Education, University of London.
Robinson, W.S. (1950). „Ecological correlations and the behavior of individuals", *American Sociological Review*, Vol. 15, [**mangler nummer?**] pp. 351-357.
Singer, Judith D. (1998). „Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models", *Journal of Educational and Behavioral Statistics*, Vol. 23, No. 4, pp. 323-355.
Snijders, Tom A.B. and Roel J. Bosker (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*, London, Thousand Oaks (CA), New Delhi: Sage Publications.
Thomsen, Søren Risbjerg (1997). „Om anvendelse af signifikanstests i ikke-stikprøve situationer", upubliceret manuskript, Institut for Statskundskab, Aarhus Universitet.

Appendiks

Indekset over elevernes sociale baggrund baseres på fem variable, der er dikotomiseret så tæt på medianen som muligt, jf. tabel A.1, og givet værdierne 0 og 1. Indekset ligger dermed i intervallet [0;5].

Tabel A.1. Operationalisering af formativt indeks for den enkelte elevs socioøkonomiske status

	Værdi	
Variabel	0	1
Etnicitet	Indvandrer/efterkommer	Etnisk dansker
Forældres uddannelse	Højest ungdoms-uddannelse	Resten
Logaritmen til forældres bruttoindkomst	[0-13,1]	[13,1-17,9]
Husstand	Bor hos ingen af eller den ene af forældre (evt. i et nyt par)	Bor hos begge forældre
Boligtype	Lejebolig	Ejerbolig

Skolernes gennemsnitlige sociale baggrund er beregnet som et simpelt gennemsnit af elevernes sociale baggrund.