

Skovenes pris

Handelsprisens variation med skovens lokalitet og areal

Lars J. Ravn-Jonsen

Juni 5:2005

akf working paper indeholder foreløbige resultater af undersøgelser eller forarbejder til artikler eller rapporter. Læseren bør derfor være opmærksom på, at resultater og fortolkninger i den færdige rapport eller artikel vil kunne afvige fra et working paper. Working paper er ikke omfattet af de procedurer for kvalitetssikring og redigering, som gælder for akf-rapporter. akf working paper udgives kun på www.akf.dk og ikke i trykt form.

Skovens pris

Handelsprisens variation med skovens lokalitet og areal

working paper

af Lars J. Ravn-Jensen

Resumé

Værdien af skovenes vedproduktion falder. Alligevel stiger handelspriserne på skov-ejendomme. Dette paradoks kan forklares med, at stigningen i skovenes rekreative eller herlighedsværdi mere end opvejer faldet i værdien af traditionelle skovprodukter. For at belyse denne tese analyseres data for samtlige handler af »særskilt vurderet skov og plantage« i Danmark fra perioden 1999-2004. Da disse data har en fordeling med ekstreme værdier, er der benyttet robuste lineære estimationsmetoder.

I den benyttede model er hektarprisen proportional med en tidsfaktor en stedfaktor og en arealfaktor. Tidsfaktoren udvikler sig eksponentielt over tid med en årlig stigning på 8,7%. Stedfaktoren er knyttet til det pågældende områdes befolkningstæthed og udviser forholdsvis højere priser i områder med højere befolkningstæthed. Arealfaktoren er en hyperbelfunktion af arealet, og selv om en anden type funktion kunne være tilpasset data, er tendensen med faldende hektarpris med størrelsen klar.

Tesen, der forklarer den paradoksale stigning i skovenes handelspris med en delvis rekreativ motiveret efterspørgsel, modsiges ikke af resultaterne. Da den rekreative værdi ikke kan forventes at stige proportionalt med arealet, har de mindre ejendomme en forholdsvis højere hektarpris, og da den rekreative benyttelse befordres af en nærhed af ejendomme, vil ejendomme i forholdsvis tætbefolkede områder efterspørges højere. Denne reaktive motiverede efterspørgsel af skovejendomme får prisen til at stige trods faldende dækningsbidrag fra skovens traditionelle produktion.

Nøgleord: Skov, Produktion, Rekreativ, Robuste estimater, Priser, Danmark.

Abstract

Even though the value of the wood production of the forest declines the market value of forest estates increases. If the increase in recreation and amenity value of the forest more than outweighs the decline in the value of traditional forest products, this paradox can be explained. In order to examine this hypothesis, data of all sales of »separate assessed forest and plantation« in Denmark in the period 1999-2004 are analysed. Since these data possess a distribution with many extreme observations, robust linear estimation methods are applied.

In the utilized model the price per hectare is proportional with a time factor, a location factor and an area factor. The time factor develops exponential by time with an annual increase of 8.7%. The location factor is tied to the population density of the place in question and prices increase with increase in population density. The area factor is a hyperbolic function of the acreage, and even though this function is not the only possibly function that might be fitted to the data, the trend of declining prices per hectare with increasing prices is evident.

The hypothesis that explains the paradoxical increase in the market value of forest estates with a recreational demand is not contradicted by the results: As the recreational value of the forest cannot be expected to rise proportional with the acreage, the smaller estates are sold to a relatively higher price per hectare, and as the recreational use of the forest is promoted by the proximity of the forest, estates in relatively high population density areas are in higher demand. These recreationally motivated demands of forest estates induce an increase in the prices of forest despite the declining contribution margin from traditional forest products.

Keywords: Forest, Production, Recreation, Robust estimation, Price, Denmark.

Forord

Der redegøres i dette papir for nogle statistiske analyser af skoves salgspriser. Formålet med dette arbejde er at skabe et grundlag for en artikel, Nils Groes og jeg vil forfatte. Den kommende artikel skal diskutere og perspektivere resultaterne af dette papirs undersøgelse med særlig henblik på at diskutere, i hvor høj grad skovens produktion består af rekreative værdier, en idé Nils Groes har advokeret for længe. Meget af ideen til og indholdet af dette papir vil derfor være sammenfaldende med artiklen, og Nils Groes har været inspirator og kritiker ved udformningen af dette papir.

Lars J. Ravn-Jonsen
Maj 2005

Indholdsfortegnelse

1. Indledning.....	6
2. Metode	9
3. Resultater	13
4. Diskussion	13
5. Konklusion.....	16
Bilag	
1. Danmark inddelt i områder efter befolkningstæthed.....	17
2. Supplerende resultater vedrørende model I og II	18
3. Robust estimation af lineærmodeller.....	21
Litteratur	24

1. Indledning

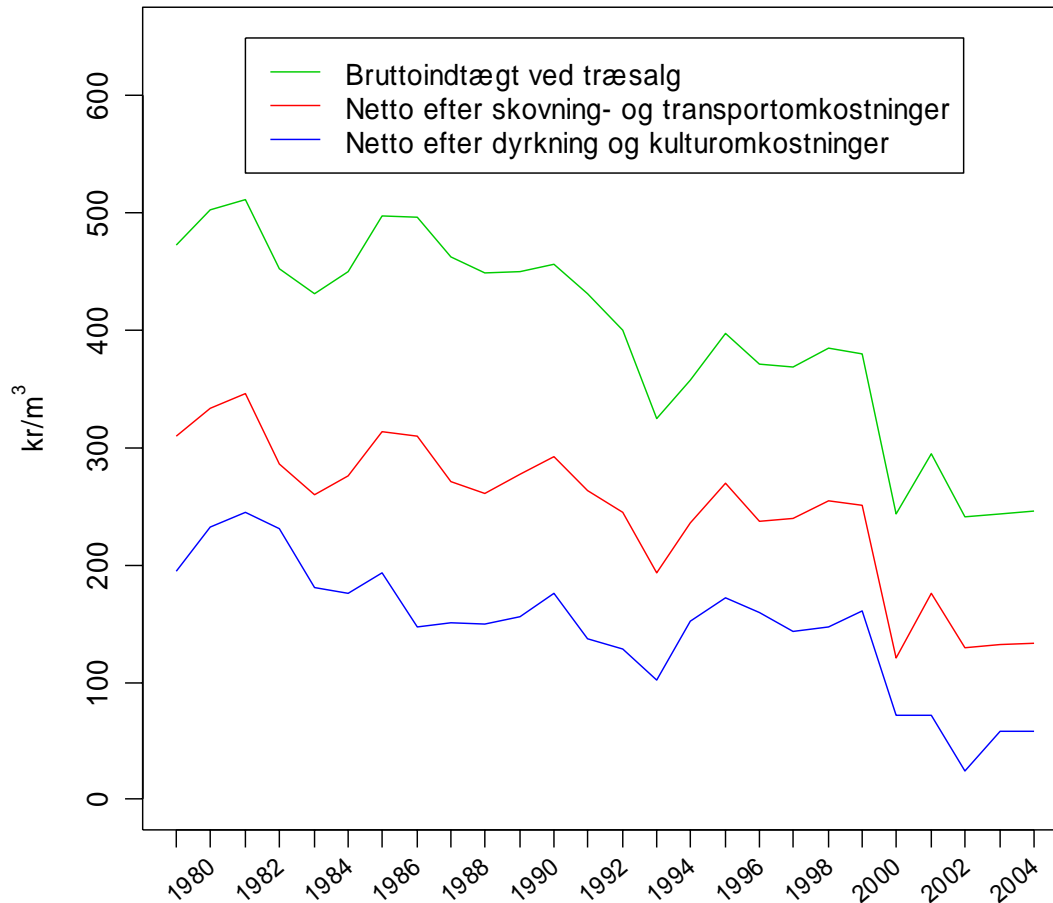
Prisen på jord afspejler den rente (dvs. afkast), der produceres af det givne stykke jord – i princippet således, at renten er lig prisen på jorden multipliceret med ejerens diskonteringsrate. Den rente, jorden producerer, kan i nogle tilfælde knyttes sammen med produktionen, der foregår på jorden, fx korndyrkning, eller knyttes sammen med jordens betydning for en anden produktion, fx er jorden tæt knyttet til svineproduktion i form af et medium for gyllespredning. Her vil svineproduktionens skyggepris for mere jord til gyllespredning svare til jordrenten. Der er dog også jorder, hvor jordens rente ikke er knyttet til en egentlig produktion, fx jord til boliger og sommerhuse. Her producerer jorden et medium til at bygge på, men prisen, og dermed jordrenten, afspejler andet end det. Her afspejles den pris, folk er villige til at betale for at opholde sig lige der, dvs. en beliggenheds-, en rekreativ eller en herlighedsværdi.

Værdien af skovenes vedproduktion falder. I figur 1 er den gennemsnitlige værdi af 1 m³ råtræ solgt fra danske skove, som den har varieret 1979-2004, vist med grønt. Priserne er gennemsnitlige priser for alle effekter korrigeret med forbrugerprisindekset til 2000-priser. Der er i skovbruget sket en kraftig produktionsudvikling, men selv når der tages højde for dette, er der stadig et kraftigt fald. I figur 1 er med blå vist et gennemsnit for »netto på rod«-værdien af træ, skovet i de danske skove. Det er udregnet ved, at alle skovningsomkostninger er trukket fra salgsprisen og repræsenterer derfor værdien af træet, lige før det fældes, når alle tidligere afholdte omkostninger anses som begravede (sunk cost). Hvis alle dyrknings- og kulturomkostninger for de enkelte år fordeles på årets skovning, fås den røde kurve. Kurven repræsenterer det, som i skovbruget kaldes dækningsbidrag II, det dækningsbidrag produktion giver, dels til administration, ejendomsskatter og afskrivninger på produktionsudstyr, dels som skovejeren kan hjemtage som rente og driftsherregæst.

Figur 1

Udvikling for prisen på skovens vedproduktion. Alle priser er i kr./m³ justeret med forbrugerprisindekset til 2000-priser

Realprisudvikling



Kilde: Dansk Skovforening.

Bem.: Med grønt er den gennemsnitlige pris for råtræ vist, med rødt er vist nettoprisen, når skovnings- og transportomkostninger er fratrukket, og med blå er vist dækningsbidraget efter årets omkostninger til dyrkning, kultur, veje og vand er afholdt. Værdier for 2004 er skønnet.

Med denne kraftigt faldende tendens for værdien af vedproduktionen burde, hvis det eneste produkt skoven producerer er ved, og ejerens diskonteringsrate er konstant, priserne på skovejendomme falde. Ejendomsværdien for skovene er dog paradoksalt nok stigende, også i realpriser (Thorsen 2003a). Skovejendomsprisers manglende sammenhæng med råtræprisen blev diskuteret i Groes (1979). Her foreslås, at skovejendomme tillige har en herligheds- og rekreativ værdi, der kom til udtryk i skovejendommens værdisætning. Thorsen (2003a) viser, at skovens samlede afkast – på trods af sin dårlige driftsøkonomi som investeringsobjekt – ikke fejler noget, når kapitalafkastet indregnes. Thorsen (2003b) tilskriver stigning i ejendomspriserne på trods af skovenes ringe driftsøkonomi til en værdisætning af skovenes herlighedsværdi.

Herlighedsværdien af et gode er den nydelse, som godet giver mulighed for at opnå uden yderligere omkostninger. Eksempler på begreber, der normalt henregnes under herlighedsværdier er: udsigt, stilhed, naturskønhed og dyreliv. En del af den herlig-

hedsværdi skoven giver er tilgængelig for offentligheden i almindelighed, idet offentlighedens adgang er sikret via naturbeskyttelsesloven. Adgangen er ikke uden begrænsninger, men det konkrete privilegium, ud over retten til drift af skoven, som ejeren har, er retten til jagten. Jagt anses som et biprodukt af skovdriften. Produktet er ikke kødet, men retten til at gå på jagt. Dette biprodukt, som er en rekreativ måde at udnytte herlighedsværdien på, forbruges i nogle tilfælde af ejerne, i andre tilfælde sælges det. Til underbygning af, at efterspørgslen efter rekreativ nydelse af herlighedsværdien i skoven er stigende kan nævnes, at jagtindtægt pr. hektar i perioden 1998-2003 i realpriser, dvs. rettet med forbrugerindeks, er steget med ca. 15% for skove på øerne og hedeområdet i Jylland og ca. 5% i øvrige Jylland (Dansk Skovforening 2004).

Drift af skov er reguleret gennem skovloven, hvor begrebet fredsskov defineres. Langt hovedparten af skove og plantager i Danmark er underlagt fredsskovspligt. Det medfører, at skovene skal holdes bevokset med træer og ikke kan anvendes til andet formål som fx afgræsningsareal for dyr eller ager med etårige afgrøder. Samtidig må sammenhængende fredsskovspligtige arealer ikke udstykkes eller formindskes ved arealoverførsel. Skovarealet er således ved lov sikret mod at blive mindre, hvorimod der ikke er noget i vejen for, at arealet øges gennem etablering af nye skovområder. Skovarealet er da også ifølge Danmarks Statistik øget fra 445.390 ha i 1990 til 486.234 ha i år 2000. En del af denne øgning skyldes formodentlig amternes mere præcise registrering af skove, specielt offentligt ejede fredsskove, og ikke nødvendigvis nyplantning.

Nyetablering af skov er således en mulighed, hvorimod nedlæggelse af skov ikke er muligt, udbudskurven for skov er derfor i praksis nærmest lodret. Skovens stigende (faldende) pris skyldes derfor en øget (mindsket) efterspørgsel og ikke en afspejling af en alternativ anvendelse af jorden. Efterspørgsel efter skove fra personer, der ønsker at eje skov af delvis rekreative motiver, vil derfor afspejle sig i priserne:"

1. Hvis rekreativt forbrug er et normalt gode, vil rekreativ motiveret efterspørgsel stige, når indkomsten stiger. Hvis værdien af skovenes rekreative afkast således stiger på grund af øget velstand, og denne stigning dominerer over skovenes produktive rente, vil det komme til udtryk i salgspriserne på skovene: salgsprisen vil stige på trods af faldende rente fra produktionen.
2. Da den rekreative værdi af skoven afhænger af ejerens mulighed for at komme i skoven, kan det forventes, at efterspørgslen efter bynære skove vil være forholdsvis højere med en deraf følgende højere prissætning.
3. Umiddelbart er skovbrug et område, hvor der eksisterer stordriftsfordele, dvs. at der op til en meget stor skovstørrelse vil være tiltagende marginalt økonomisk udbytte pr. hektar, men ud fra en formodning om, at den rekreative værdi ikke er ligefrem proportional med størrelse, men derimod marginalt faldende, vil mindre skovejendomme blive handlet til forholdsvis højere pris pr. hektar.

Punkt 3 er til dels en følge af punkt 1, idet de begge bygger på en antagelse om, at rekreativ erhvervelse af skoven er et normalt forbrugsgode. Punkt 3 kan være svært at

skelne fra punkt 2, hvis de store skove også er dem, der ligger langt fra byerne. Undersøgelsen af punkterne 2 og 3 kan også vanskeliggøres, hvis boniteten af bynære skove er forholdsvis bedre.

2. Metode

Til analysen af skovpriser benyttes data for handler af »særskilt vurderet skov og plantage« i tiden 1999-2004, indsamlet af Told og Skat. Datasættet indeholder 1.583 observationer, men indeholder observationer, der ingen pris har eller har arealet 0. Når disse er fjernet, er der 1.350 tilbage. Ud over salgsdato, salgspris, areal og antal bygninger er der også en ejendomscode, hvis første tre cifre er kommunens nummer.

Modelmæssigt ønskes at forklare hektarprisen for skov ud fra året, arealet og området samt overveje bygningers indflydelse på prisen. Simple gennemsnit og medianer af data, grupperet efter årstal, størrelse og amt, er gengivet i tabel 1, 2 og 3. De forventede tendenser: stigende pris over tid, faldende pris med størrelse og højere pris i visse områder af landet, er til en vis grad til stede i simple gennemsnit af grupperede data, men tendenserne er langt tydeligere til stede, når medianer benyttes som midtpunktsmål. Priserne i datamaterialet er yderst varierende med priser fra 8 kr. til 550 mio.kr. pr. ha. De ekstreme værdier har stor indflydelse på gennemsnit, men lille indflydelse på medianer. Forsøg med at danne en fornuftig normalfordelingsmodel til forklaring af prisen på skov må af samme grund opgives: de ekstreme prisværdier giver en fordeling af residualen, der ikke kan opfattes som normalfordelt, idet der optræder mange ekstreme værdier, som har stor indflydelse på estimatet.

Som det ses af den store forskel på gennemsnit og medianer, er fordeling af hektarpriser skæv. Dette kan til dels afhjælpes med en logaritmisk transformation af priserne. Herved kommer store værdier til at have mindre indflydelse på estimatet, men residualernes fordeling er stadig med mange ekstreme værdier, der har stor indflydelse på estimatet. Der er derfor benyttet en metode, der antager, at residualen har en central symmetrisk fordeling, der er »forurennet« med elementer med stor spredning, en såkaldt robust metode (robust over for ekstreme værdier). Metoden er foreslået af Yhoai et al. (1991) og benytter en maksimum likelihood-lignende estimeringsrutine, der vægter residualerne, ikke med kvadratfunktionen som i normalfordelingsmodeller, men med en funktion, der sætter ekstreme værdier til nul. Hvor store residualer, der skal vægtes med nul, afgøres ved en skalering af residualerne med et estimat for spredningen. Rutinen foregår trinvis. Først udføres en spredningsestimation, hvorefter kvotienterne i den lineære model estimeres. Derefter udføres en kontrol for, om metoden giver et centralt estimat, og endelig udregnes kovariansmatrixen for kvotientestimatene. I bilag 3 er givet en mere detaljeret gennemgang af metoden. Rutinen er gennemført i R (R Development Core Team (2004) og biblioteket MASS ved hjælp af funktionen `rlm(..., method = 'MM')` (se Venables og Ripley (2002) og Ripley (2005) for flere detaljer vedrørende `rlm`).

Tabel 1**Gennemsnit og medianer for handlet skov efter år i kr./ha**

År	Fuldt datasæt			Kun skove uden bygninger		
	Gennemsnit	Median	Antal	Gennemsnit	Median	Antal
1999	142.963	42.102	212	84.674	39.695	136
2000	3.934.936	45.259	266	5.428.597	40.845	190
2001	170.026	47.588	220	144.504	44.450	163
2002	385.261	50.992	242	512.001	48.344	173
2003	366.285	57.000	221	200.809	52.318	151
2004	189.070	63.093	189	172.421	55.382	135
I alt			1350			948

Bem.: »Antal« angiver antallet af observationer. Beregningerne er fortaget for både fuldt datasæt og observationer med skove uden bygninger.

Tabel 2**Gennemsnit og medianer for handlet skov efter størrelse i kr./ha**

Størrelse Ha	Fuldt datasæt			Kun skove uden bygninger		
	Gennemsnit	Median	Antal	Gennemsnit	Median	Antal
0-1	11.888.794	158.466	104	14.144.841	77.662	82
1-5	94.097	51.493	360	72.014	49.782	316
5-10	80.003	48.033	228	55.663	42.210	178
10-50	56.366	46.701	466	42.549	41.179	317
50-100	58.372	54.411	104	48.627	46.235	48
100+	39.043	23.266	88	44.970	57.617	7
I alt			1350			948

Bem.: »Antal« angiver antallet af observationer. Beregningerne er fortaget for både fuldt datasæt og observationer med skove uden bygninger.

Tabel 3**Gennemsnit og medianer for handlet skov efter amt i kr./ha**

Amt	Fuldt datasæt			Kun skove uden bygninger		
	Gennemsnit	Median	Antal	Gennemsnit	Median	Antal
Nordsjællands	2.817.992	357.971	10	2.612.472	75.499	6
Vestsjællands	268.702	65.795	32	94.547	74.538	9
Storstrøms	10.242.114	14.605	55	25.462.297	48.155	22
Bornholms	109.022	39.683	17	109.022	39.683	17
Fyns	189.889	73.807	70	187.939	62.287	44
Sønderjyllands	63.116	47.948	159	49.833	46.611	141
Ribe	66.869	44.698	93	66.856	40.685	72
Vejle	194.517	54.278	283	204.807	48.921	217
Ringkøbing	3.679.481	41.837	126	5.020.198	40.455	92
Århus	460.969	55.382	253	411.990	51.136	159
Viborg	151.182	49.463	122	145.983	42.182	90
Nordjyllands	300.743	48.421	130	317.382	41.408	79
I alt			1350			948

Bem.: Nordsjælland omfatter Roskilde, Frederiksborg og Københavns Amter. Der er ingen skove handlet i Københavns og Frederiksberg Kommuner. »Antal« angiver antallet af observationer. Beregningerne er foretaget for både fuldt datasæt og observationer med skove uden bygninger.

I tabel 1 ses, at medianerne af data, fordelt efter år, viser en jævn stigning på omkring 8% pr. år. Derfor modelleres logaritmen til hektarprisen som funktion af datoen. Det ses af tabel 2, at prisen falder med størrelsen, stejlt i starten og fladt for større værdier, en hyperbelfunktionen synes rimelig, hvorfor logaritmen til hektarprisen modelleres som funktion af logaritmen til arealet. Videre viser tabel 3, at der i amter med høj befolkningstæthed er en tendens til højere skovpriser. Det er dog ikke muligt at estimere en robust lineær regression med amterne som forklarende variabler. Der er for få observationer i de enkelte amter, og ved gruppering af indbyrdes naboamter opnås ikke meningsfulde resultater. I stedet er valgt at inddele kommunerne efter befolkningstæthedskvartiler, dvs. i fire lige store dele efter stigende befolkningstæthed. Disse inddelinger kaldes områder, hvor »område I« har lavest befolkningstæthed og »område IV« har højest befolkningstæthed. Andelen af observationer er dog ikke lige fordelt, der er langt flere skovhandler i tyndbefolkede områder. Et kort med de benyttede områder er vist i bilag 1. Bemærk, at områderne ikke er geografisk sammenhængende, men alene er et udtryk for befolkningstæthed.

For at bedømme værdien af bygningernes indflydelse på resultaterne, er modellerne estimeret for både det fulde datasæt og for et datasæt med udelukkende de ejendomme, der ingen bygninger har.

Den benyttede model er:

Model I:

$$\log(\text{pris}_i / \text{ha}_i) = \alpha(\text{område}_i) + \beta_1 \text{dato}_i + \beta_2 \log(\text{ha}_i) + \varepsilon_i$$

$i = 1, 2 \dots 1350$ for fuldt datasæt og $i = 1, 2 \dots 948$ for datasæt med ejendomme uden bygninger. pris_i er ejendommens handelspris, ha_i er ejendommens areal i hektar, dato_i er handelsdato målt i dage siden 2.12.2001. Denne dato er gennemsnitlig handelsdato. Valget er uden indflydelse på resultaterne, men er valgt som udgangspunkt for at minimere kovariansen mellem estimationen af β_1 og de øvrige kvotienter. $\text{område}_i \in \{\text{område I, område II, område III, område IV}\}$, der er landets kommuner inddelt i kvartiler efter befolkningstæthed. Der er henholdsvis 582, 361, 244 og 163 observationer i hvert af områderne. Residualen ε_i antages at være indbyrdes uafhængig og identisk fordelte (dvs. der er ingen forudsætning om normalfordeling).

Model I giver signifikante resultater (se Resultater) og viser som forventet, at prisen stiger i de tætbefolkede områder og falder med arealet. Dette fald i prisen med arealet kan dog skyldes, at store skove har en tendens til at ligge i tyndtbefolkede områder. Selv om der er korrigeret ved inddeling af landet i de fire områder, vil tendensen med store arealer i tyndere befolkede områder også være gældende inden for det enkelte område. Derfor bruges i model II logaritmen til indbygger pr. ha som kontinuert variabel.

Inden for skovbruget grupperes skovejendomme groft efter bonitet i tre grupper: Heden, øvrige Jylland, og øerne. Heden, dvs. vest for den jyske højderyg, har ringere bonitet end det øvrige Danmark. Idet store ejendomme er mere almindelige på heden, og heden er forholdsvis tyndt befolket, kan prisforskelle, som har sin oprindelse i bonitet, komme til udslag som pristendenserne med hensyn til areal og befolkningstæthed. Derfor er denne grove bonitetsinddeling inddraget i model II for at sikre, at det ikke er bonitetsforskelle, der forårsager pristendenserne med hensyn til areal og befolkningstæthed.

Model II:

$$\log(\text{pris}_i / \text{ha}_i) = \alpha + \beta_1 \text{dato}_i + \beta_2 \log(\text{ha}_i) + \beta_3 \text{ind}_i + \beta_4 \text{hede}_i + \beta_5 \text{øer}_i + \varepsilon_i$$

$i = 1, 2 \dots 1350$ for fuldt datasæt og $i = 1, 2 \dots 948$ for datasæt med ejendomme uden bygninger. pris_i er ejendommens handelspris, ha_i er ejendommens areal i hektar, dato_i er handelsdato målt i dage siden 2.12.2001. Heden¹ og øerne er inddraget som dummyvariabler, dvs. $\text{hede}_i = 1$, hvis ejendommen er i hedeområdet ellers nul, og $\text{øer}_i = 1$, hvis ejendommen er på øerne, ellers 0. α vil så give den systematiske effekt for øvrige Jylland, og β_4 og β_5 vil angive forskellen mellem den systematiske effekt fra henholdsvis heden og øerne fra øvrige Jylland. Residualen ε_i antages at være indbyrdes uafhængig og identisk fordelte.

1 Hedeområdet i skovbonitets betydning er området vest og syd for Weichselistidens hovedopholdslinje. I denne undersøgelse er observationerne knyttet til kommuner. Kommuner, der overvejende ligger i hedeområdet, er regnet som hede.

3. Resultater

I tabel 4 og 5 er gengivet parameterestimerne. Bemærk, at tiden måles i dage: model I for fuldt datasæt siger, at priserne efter et år er $\exp(365,25 * \beta) = 1,09$ højere, dvs. 9% årlig stigning. I bilag 2 er gengivet spredningsestimater, kovariansmatrixer og teststørrelser og -resultater. Desuden er et diagram med residuaer og den benyttede vægtningsfunktion.

Tabel 4
Parameterestimer for model I

Parameter	Fuldt datasæt		Uden bygninger	
	Estimat	Spredning	Estimat	Spredning
$\alpha(\text{område I})$	10,870	0,0621	10,865	0,0625
$\alpha(\text{område II})$	10,932	0,0678	10,933	0,0677
$\alpha(\text{område III})$	11,027	0,0689	11,047	0,0681
$\alpha(\text{område IV})$	11,130	0,0746	11,121	0,0742
β_1	2,286e-4	0,330e-4	1,823e-4	0,317e-4
β_2	-0,0333	0,0118	-0,0542	0,0129

Tabel 5
Parameterestimer for model II

Parameter	Fuldt datasæt		Uden bygninger	
	Estimat	Spredning	Estimat	Spredning
α	10,610	0,1435	10,5466	0,1440
β_1	2,31e-4	3,20e-05	1,91e-4	3,28e-05
β_2	-0,0318	0,0119	-0,0494	0,0135
β_3	0,0953	0,0293	0,1031	0,0290
β_4	-0,1559	0,0428	-0,1062	0,0437
β_5	0,1350	0,0615	0,1018	0,0688

4. Diskussion

I de benyttede modeller er hektarprisen proportional med en tidsfaktor, en arealfaktor og en stedfaktor. Tidsfaktoren udvikler sig eksponentielt over tid, arealfaktoren er proportional med en funktion af arealet og stedfaktoren er knyttet til det pågældende områdes befolkningstæthed. Endelig er der i model II en faktor, der korrigerer for, om skoven ligger på heden eller på øerne. De enkelte faktorer diskuteres hver for sig.

Tabel 6
Årlig stigning af skovenes hektarpris i procent

	Fuldt datasæt		Uden bygninger	
	Model I	Model II	Model I	Model II
Årlig stigning	8,71	8,79	6,89	7,23

Der er i alle de benyttede modeller fundet en tendens til stigende hektarpriser på skov- ejendomme over tid. I tabel 6 er de estimerede parametre for prisstigninger omregnet til årlig stigning. I figur 1 er gengivet gennemsnitlig dækningsbidrag fra almindelig skovdrift. Diagrammet opererer med forbrugerindekskorrigerede tal, hvorimod tabel 6 er i nominelle priser. I tidsrummet 1999-2004 er forbrugerindekset steget med gennemsnitligt 2,1% pr. år. Skovenes handelspriser overgår langt denne stigning i modsætning til dækningsbidraget, der er faldende. Stigningen i pris afspejler en øget efterspørgsel efter skov, en efterspørgsel, der ikke kan være begrundet i skovenes almindelige driftsindtægter. En del af stigningen kan tilskrives stigning i bygningernes værdi, men selv uden bygninger er stigningen betydelig.

At priserne på skov stiger over tid på trods af forringelse af driftsindtægter fra skoven kan skyldes spekulation, dvs. en forventning om, at priserne stiger i fremtiden. Men hvor kommer den forventning fra, hvorfor forventes priserne at stige? Den spekulative forventning må bygge på en opfattelse af, at efterspørgslen efter skov i fremtiden vil stige. Der kan næppe være en rationel forventning til, at driftsresultaterne i skoven skal trække prisen op, hvorfor den spekulative motiverede erhvervelse må bygge på en opfattelse af, at noget andet end driftsresultaterne fra traditionel skovdrift vil trække priserne yderligere op.

Arealfaktorens indflydelse er i tabel 7 udregnet for nogle skovstørrelser med 100 ha som indeks. Faldet i indekset med størrelsen er påvirket af, om skove med bygninger indgår i datasættet eller ej og er betydeligt større, når bygninger ikke indgår. Arealfaktoren er ikke nævneværdig forskellig mellem model I og model II. Den fundene sammenhæng er derfor næppe et udtryk for ubalance i datasættet, dvs. flere stor ejendomme i tyndbefolkede områder eller hedeområder og må tages som udtryk for, at efterspørgslen efter mindre ejendomme realt er forholdsvis større. Modellens hyperbelbeskrivelse er ikke den eneste funktion, der vil kunne beskrive data, og bør tages med forbehold; men tendensen med klart højere priser for mindre skove er klar.

Tabel 7**Arealfaktorens indflydelse på skoves hektarpris forklaret med indekstal (100 ha=100)**

Størrelse	Fuldt datasæt		Uden bygninger	
	Model I	Model II	Model I	Model II
0,1 ha	125	125	145	141
1 ha	117	116	128	126
5 ha	111	110	118	116
20 ha	106	105	109	108
50 ha	102	102	104	103
100 ha	100	100	100	100
300 ha	96	97	94	95
1000 ha	93	93	88	89

Stedfaktoren, der er knyttet til stedets befolkningstæthed, er som indeks gengivet i tabel 8. I tabellen er under model I med ens bogstaver markeret, hvilke værdier der ikke kan anses for forskellige på 5%-niveau, jf. tabel 13 og 14 i bilag 2. Under model II er udregnet, hvad gennemsnitlig forventet værdi i de fire områder er efter model II, givet de handler, der er i datasættet. Der er ikke nævneværdi forskel på estimerne efter om skove med bygninger indgår eller ej, hvorimod opdelingen af landet i de traditionelle bonitetsområder (hede, øvrige Jylland og øerne) i model II reducerer indflydelsen en anelse. På trods af dette er tendensen klar: skove prissættes højere i tætbefolkede områder.

Tabel 8**Områdefaktorens indvirkning på skoves hektarpris forklaret med indekstal (område I =100)**

	Fuldt datasæt		Uden bygninger	
	Model I	Model II	Model I	Model II
Område I	100 a	100	100 a	100
Område II	106 a b	105	107 a	105
Område III	117 c b	110	120 b	111
Område IV	129 c	122	129 b	125

Bem.: Tal, der under model I ikke er mærket med samme bogstav, er signifikant forskellige på 5%-niveau ved ensidige test, se bilag 2. Tal under model II er den forventede værdi, fundet under model II, givet de handler der er i datasættet.

Tabel 9**Områderne heden, øvrige Jylland og øernes indvirkning, som de er fundet i model II, på hektarprisen forklaret med indekstal (øvrige Jylland=100)**

	Fuldt datasæt	Uden Ejendomme
Heden	86	90
Øvrige Jylland	100	100a
Øerne	114	111a

Bem.: Tal, der ikke er mærket med samme bogstav, er signifikant forskellige på 5%-niveau ved tosidig test, se bilag 2.

I tabel 9 er gengivet model II's forventninger til priser i den traditionelle inddeling af skove efter heden, øvrige Jylland og øerne. Den traditionelle forventede prisforskel eksisterer, dog er forskellen mellem øvrige Jylland og øerne ikke signifikant, når kun skove uden bygninger betragtes. Hvis resultaterne fra model I sammenholdes med model II kan konkluderes, at boniteten nok kan forklare noget af den fundne forskel imellem områderne i model I, men langt fra hele forskellen.

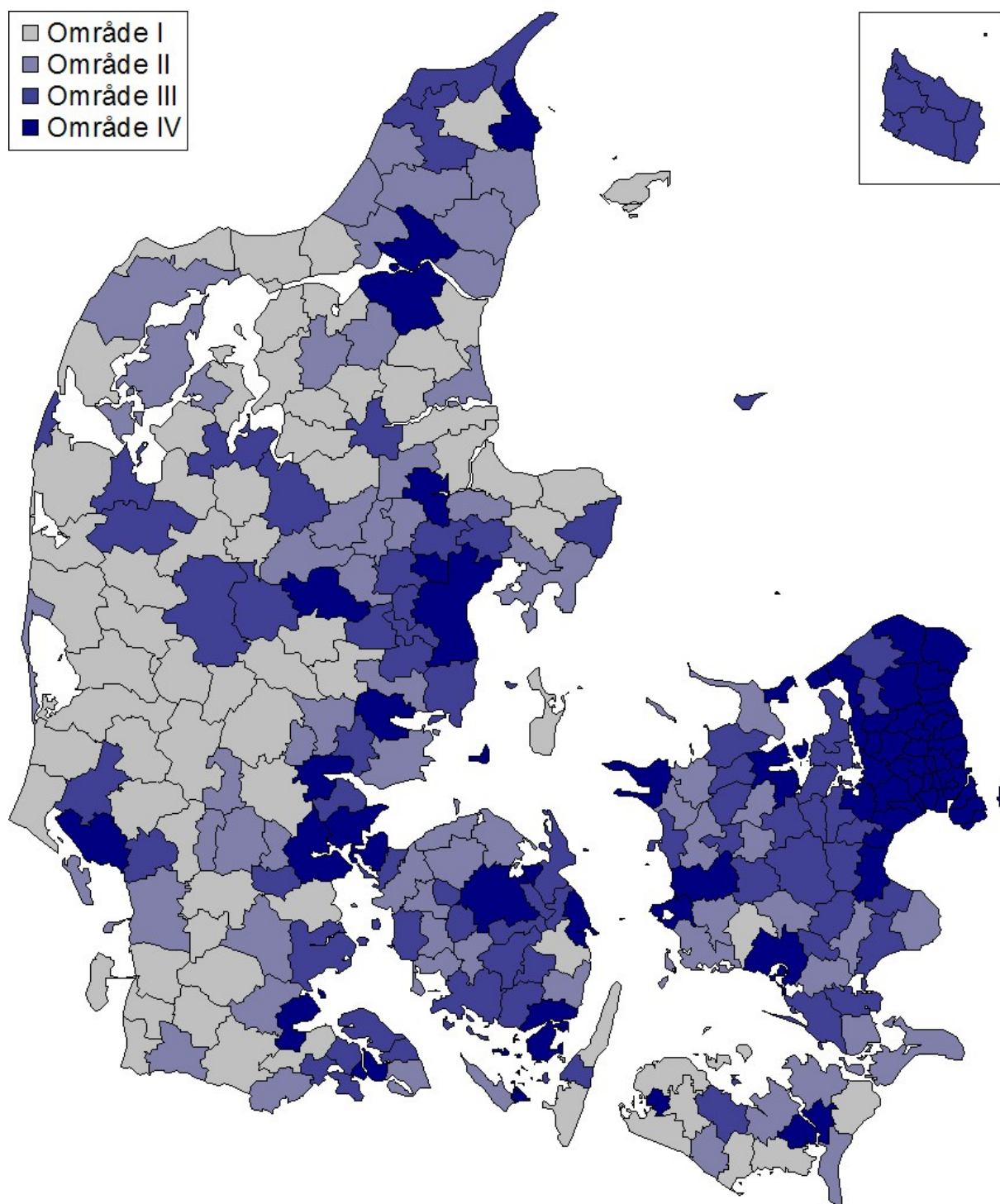
5. Konklusion

Analysen af skovejendommens handelspriser i perioden 1999-2004 viser, at ejendomspriserne handles til højere priser, jo mindre skovkene er, og jo tættere befolkede området er, skoven ligger i. Skovens handelsværdi er steget med omkring 8,7% årligt. En del af denne stigning kan dog forklares ved, at bygningerne er blevet mere værd, men selv for skove uden bygninger er stigningen omkring 7% årligt. Ejendommens nærhed til byer og ejendommenes størrelse har indflydelse på hektarprisen, således at mindre skove og skove i tæt befolkede områder prissættes forholdsvis højere; begge tendenser er kraftigst for skove uden bygninger. For disse vil en skov i tætbeholdt område som median prissættes omkring 25% højere end i et tyndbefolket område, og en 1 ha stor skov vil som median prissættes 28% højere pr. hektar end en skov på 100 ha. Skovenes prisforskel kan også forklares ud fra den traditionelle inddeling af skove efter hede, øvrige Jylland og øerne, men inddragelsen af denne inddeling ændrer ikke ved konklusionen: Skovens pris stiger på trods af forringede driftsresultater, mindre skove og skove i tættere befolkede områder efterspørges forholdsvis mere.

Den i indledningen fremsatte tese, der forklarer den paradoksale stigning i skovens handelspris med en delvis rekreativ motiveret efterspørgsel, modsiges ikke af resultaterne. Da den rekreative værdi ikke kan forventes at stige proportionalt med arealet, har de mindre ejendomme en forholdsvis højere hektarpris, og da den rekreative benyttelse befordres af en nærhed af ejendomme, vil ejendomme i forholdsvis tætbeholdt områder har en større efterspørgsel. Denne rekreative motiverede efterspørgsel af skovejendomme får prisen til at stige trods faldende dækningsbidrag fra skovens traditionelle produktion.

Bilag 1

Danmark inddelt i områder efter befolkningstæthed



Bilag 2

Supplerende resultater vedrørende model I og II

I dette bilag gives detaljeret resultater fra den statistiske analyse. I tabel 10 er angivet skaleringsparameteren $\widehat{\sigma}_0$ og den spredning kovariansmatrixen vægtes med $s = \sqrt{\widehat{\sigma}_0^2 \widehat{\tau} n^{-1}}$ (se bilag 3 for forklaring) samt antallet af frihedsgrader i modellerne. I tabel 11 og tabel 12 er gengivet korrelationsmatrixer for model I, da de indgår ved beregningen af teststørrelser. I tabel 13, 14, 15 og 16 er givet relevante teststørrelser, og i figur 2 er vist de standardiserede residuaer for de fire modelkørsler. Endvidere er i figuren angivet den funktion, der bruges i rutine under iterationen, summen af de standardiserede residualer ganget med denne funktion skal give nul. Af figuren fås et indtryk af de lange haler, der er til begge sider, og hvordan den brugte metode undlader at vægte udliggere.

Tabel 10

Spredningsmål og frihedsgrader for alle modeller

	Fuldt datasæt		Uden bygninger	
	Model I	Model II	Model I	Model II
$\widehat{\sigma}_0$	0,6489	0,6345	0,5572	0,5544
s	0,7295	0,7139	0,6105	0,6063
df	1344	1344	942	942

Tabel 11

Correlationsmatrix for estimation for helt datasæt, model I

	β_1	β_2	$\alpha(\text{område I})$	$\alpha(\text{område II})$	$\alpha(\text{område III})$
β_2	0,0150				
$\alpha(\text{område I})$	0,0033	-0,8770			
$\alpha(\text{område II})$	-0,0341	-0,8259	0,7240		
$\alpha(\text{område III})$	-0,0062	-0,7407	0,6497	0,6117	
$\alpha(\text{område IV})$	-0,0220	-0,6548	0,5741	0,5411	0,4850

Tabel 12**Correlationsmatrix for estimation for datasæt uden bygninger, model II**

	β_1	β_2	$\alpha(\text{område I})$	$\alpha(\text{område II})$	$\alpha(\text{område III})$
β_2	0,0068				
$\alpha(\text{område I})$	0,0029	-0,8849			
$\alpha(\text{område II})$	-0,0176	-0,8193	0,7249		
$\alpha(\text{område III})$	-0,0133	-0,7605	0,6728	0,6232	
$\alpha(\text{område IV})$	-0,0208	-0,6772	0,5991	0,5550	0,5151

Tabel 13**Relevante teststørrelser for model I, fuldt datasæt. H_0 er den hypotese, der testes, og H_A er den hypotese, der gælder, hvis H_0 afvises**

H_A	t	$p(H_0)$
$\beta_1 \neq 0$	7,000	4,0e-12
$\beta_2 \neq 0$	-2,774	0,0056
$\alpha(\text{område II}) - \alpha(\text{område I}) > 0$	1,264	0,1032
$\alpha(\text{område III}) - \alpha(\text{område I}) > 0$	2,819	0,0024
$\alpha(\text{område IV}) - \alpha(\text{område I}) > 0$	3,999	3,4e-05
$\alpha(\text{område III}) - \alpha(\text{område II}) > 0$	1,571	0,0581
$\alpha(\text{område IV}) - \alpha(\text{område II}) > 0$	2,859	0,0022
$\alpha(\text{område IV}) - \alpha(\text{område III}) > 0$	1,387	0,0828

Tabel 14**Relevante teststørrelser for model I med datasæt uden bygninger. H_0 er den hypotese, der testes, og H_A er den hypotese, der gælder, hvis H_0 afvises**

H_A	t	$p(H_0)$
$\beta_1 \neq 0$	5,532	4,1e-08
$\beta_2 \neq 0$	-4,034	5,9e-05
$\alpha(\text{område II}) - \alpha(\text{område I}) > 0$	1,366	0,0862
$\alpha(\text{område III}) - \alpha(\text{område I}) > 0$	3,323	4,6e-04
$\alpha(\text{område IV}) - \alpha(\text{område I}) > 0$	4,017	3,2e-05
$\alpha(\text{område III}) - \alpha(\text{område II}) > 0$	1,911	0,0281
$\alpha(\text{område IV}) - \alpha(\text{område II}) > 0$	2,764	0,0029
$\alpha(\text{område IV}) - \alpha(\text{område III}) > 0$	1,025	0,1527

Tabel 15

Relevante teststørrelser for model II, fuldt datasæt. H_0 er den hypotese, der testes og H_A er den hypotese, der gælder, hvis H_0 afvises

H_A	t	$p(H_0)$
$\alpha \neq 0$	73,921	<2e-16
$\beta_1 \neq 0$	7,217	8,8e-13
$\beta_2 \neq 0$	-2,681	0,0074
$\beta_3 \neq 0$	3,247	0,0011
$\beta_4 \neq 0$	-3,642	2,8e-04
$\beta_5 \neq 0$	2,196	0,0282

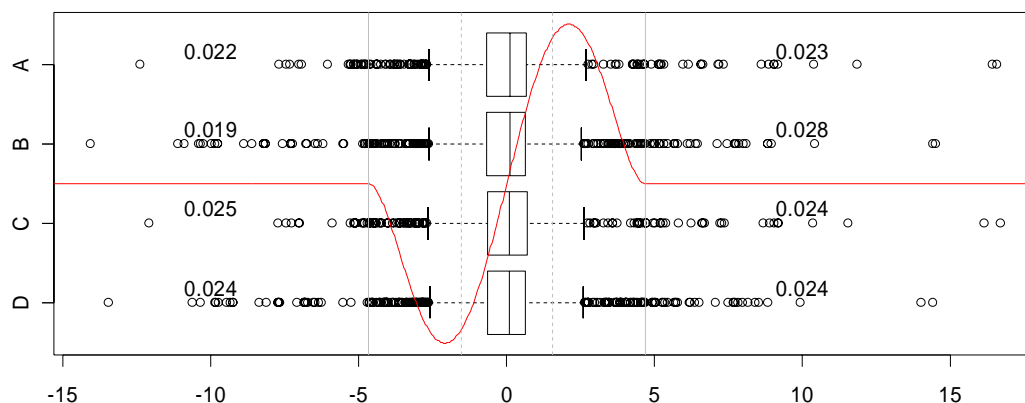
Tabel 16

Relevante teststørrelser for model II, med datasæt uden bygninger. H_0 er den hypotese, der testes, og H_A er den hypotese, der gælder, hvis H_0 afvises

H_A	t	$p(H_0)$
$\alpha \neq 0$	73,250	0
$\beta_1 \neq 0$	5,826	7,8e-09
$\beta_2 \neq 0$	-3,671	2,6e-04
$\beta_3 \neq 0$	3,548	4,1e-04
$\beta_4 \neq 0$	-2,428	0,0153
$\beta_5 \neq 0$	1,480	0,1392

Figur 2

Residualer og ψ_c



Bem.: Residualerne standardiseret med skaleringsparameteren $\hat{\sigma}_0$ vist som boksdiagram. A er model I, fuldt datasæt, B er model I, datasæt uden bygninger, C er model II, fuldt datasæt, og D er model II, datasæt uden bygninger. De to lodrette grå linjer viser $\pm c_1 = \pm 4,685$. Den røde kurve er vægtningsfunktionen ψ_{c_1} , observationer uden for de grå streger vægtes med 0. Andelen af observationer udenfor de grå streger er angivet med tal til hver side. De prikkede grå linjer viser $\pm c_0 = \pm 1,548$, der benyttes i det første trin i rutinen til estimation af skaleringsparameteren $\hat{\sigma}_0$.

Bilag 3

Robust estimation af lineærmodeller

Først gives kort baggrunden for den benyttede metode, derefter gennemgås metoden i detaljer.

For en lineær model $\mathbf{y} = \mathbf{x}\boldsymbol{\theta} + \varepsilon$ er maksimum likelihood-estimat (MLE):

$$\arg \min_{\boldsymbol{\theta}, s} \sum_{i=1}^n \rho \left(\frac{y_i - \boldsymbol{\theta}^T \mathbf{x}_i}{s} \right) + n \log s \quad (1)$$

Hvor ρ defineres som $\rho = -\log f$, hvor f er tætheds- eller sandsynlighedsfunktion for residualerne, og s er en skaleringsparameter svarende til spredningen i normalfordelingen. Den benyttede metode er en MLE-lignende estimator, hvilket betyder, at ρ ikke vælges ud fra en formodet fordeling af residualerne, men typisk vælges, så den giver et resultat, der ikke er påvirket af få ekstreme værdier.

Hvis s er kendt, simplificeres løsningen af ligning 1 til førsteordensbetingelsen:

$$\sum_{i=1}^n \mathbf{x}_i \psi \left(\frac{y_i - \boldsymbol{\theta}^T \mathbf{x}_i}{s} \right) = \mathbf{0} \quad (2)$$

hvor $\psi = \rho'$ (hvis den eksisterer). Dette inspirerer til en rutine, der først finder s og der efter $\boldsymbol{\theta}$. Da yderligere den benyttede ρ funktion ikke er konveks, er der behov for at udpege startpunktet for iterationen, således at der sikres konvergens mod det ønskede lokale minimum. Derfor foregår den benyttede metode i trin:

1. a. Find et godt startpunkt kaldet $\widetilde{\boldsymbol{\theta}}_0$
b. Find et godt koefficientestimat kaldet $\widehat{\boldsymbol{\theta}}_0$, der minimerer spredningsestimatet.
c. Find ud fra $\widehat{\boldsymbol{\theta}}_0$ spredningsestimatet $\widehat{\sigma}_0$.
2. Givet $\widehat{\sigma}_0$, estimer det endelige koefficientestimat $\widehat{\boldsymbol{\theta}}_1$.
3. Kontroller, om estimatet kan anses for centralt. Dvs. om $E\widehat{\boldsymbol{\theta}}_1 = \boldsymbol{\theta}_1$, og altså om kvotientestimatet kan anses for at være et veldefineret midtpunkt.
4. Udregn kovariansmatrixen for kvotientestimatet.

Den i opgaven benyttede estimatorfunktion kaldes »Tukey's bisquare« efter den afledte:

$$\rho_c(x) = \begin{cases} \left(\frac{x}{c}\right)^6 - 3\left(\frac{x}{c}\right)^4 + 3\left(\frac{x}{c}\right)^2 & |x| < c \\ 1 & |x| \geq c \end{cases}$$

og tilhørende afledte (ganget med $c/6$):

$$\psi_c(x) = \begin{cases} \frac{x}{c} \left[\left(\frac{x}{c} \right)^2 - 1 \right]^2 & |x| < c \\ 0 & |x| \geq c \end{cases}$$

Valget af c afgør, hvor store residualer der skal medtages i beregningerne. Når residualerne er delt med s , kan de betragtes som »standardiserede«, forstået sådan, at spredningen vil være tæt på 1. Den i rutinen fundne $\widehat{\sigma}_0$ er dog ikke den rigtige spredning. Den spredning, der bruges i vægtningen af kovariansmatrixen, er, som det ses af tabel 2, lidt større end $\widehat{\sigma}_0$. I den benyttede rutine indgår to valg af c . Først bruges $c = 1,548$, derefter bruges $c = 4,685$. I figur 1 er vist de »standardiserede« residualer, $\frac{y_i - \boldsymbol{\theta}^T \mathbf{x}_i}{s}$, fra de 4 estimerede modeller og funktionen $\psi_c(x)$ med $c = 4,685$.

$\psi_c(x)$ angiver den vægt, de »standardiserede« residualer vægtes med ved løsningen af ligning 2. De grå linjer angiver, hvornår denne vægt er nul, dvs. ikke indgår i beregningen. Med tal er angivet, hvor stor andelen af observationer uden for de grå linjer er. Når $\widehat{\sigma}_0$ estimeres, er $c = 1,548$. Her vil vægten være nul et kort stykke uden for den centrale boks. For nærmere begrundelse af netop disse størrelser for c , henvises til Yohai et al. (1991).

MM-rutine efter Yohai, Stahel, og Zamar (1991)

Nedenfor er givet en detaljeret gennemgang af MM-rutinen foreslået af til Yohai et al. (1991) og benyttet i denne analyse. Trinnene refererer til rutinen foreslået ovenfor. Som »passende algoritme« til iteration benytter R-funktionen `rlm()` den almindelige Fisher-scoring eller »iterative re-weighted least squares«.

Indledningsvis defineres $s_c(\boldsymbol{\theta})$ som løsningen til

$$\frac{1}{n} \sum_{i=1}^n \rho_c \left(\frac{y_i - \boldsymbol{\theta}^T \mathbf{x}_i}{s} \right) = 0,5$$

Trin 1a:

$$c = c_0 = 1,548$$

Længden af $\boldsymbol{\theta}$ kaldes p . \mathbf{J} er et sæt af n_J vektorer af længden p , hver vektor bestående af p tilfældig heltal mellem 1 og antallet af observationer:

$$\mathbf{j} \in \mathbf{J}, \mathbf{j} = (j_1, j_2, \dots, j_p), j_i \in \{1, 2, \dots, n\}$$

Antallet af udtrukne vektorer udregnes som $n_J = 6.9 \cdot 2^p$

$\boldsymbol{\theta}_j$ defineres som løsningen til $y_i - \boldsymbol{\theta}_j^T \mathbf{x}_i = 0$

$$\widetilde{\boldsymbol{\theta}}_0 = \arg \min_{\boldsymbol{\theta} \in \{\boldsymbol{\theta}_j | j \in \mathbf{J}\}} s_c(\boldsymbol{\theta})$$

Trin 1b

Med en passende algoritme findes det lokale minimum med udgangspunkt i $\widetilde{\boldsymbol{\theta}}_0$ for $s_c(\boldsymbol{\theta})$:

$$\widehat{\boldsymbol{\theta}}_0 = \text{local arg min}_{\boldsymbol{\theta}} s_c(\boldsymbol{\theta})$$

Trin 1c:

$$\widehat{\sigma}_0 = s_c(\widehat{\boldsymbol{\theta}}_0)$$

Trin 2

$$c = c_1 = 4,685$$

Med en passende algoritme findes det endelige koefficientestimat som det lokale minimum med udgangspunkt i $\widehat{\boldsymbol{\theta}}_0$

$$\widehat{\boldsymbol{\theta}}_1 = \text{local arg min}_{\boldsymbol{\theta}} \sum_{i=1}^n \rho_c \left(\frac{y_i - \boldsymbol{\theta}^T \mathbf{x}_i}{\widehat{\sigma}_0} \right)$$

Trin 3

Bestem $\tilde{\sigma} = s_{c_0}(\widehat{\boldsymbol{\theta}}_1)$. Der udregnes en teststørrelse T_b^2 , der er en standardiseret udgave af forskellen mellem residualernes spredning opnået ved $\widehat{\boldsymbol{\theta}}_1$, og dem opnået ved det indledende estimat $\widehat{\boldsymbol{\theta}}_0$.

Udregn:

$$\begin{aligned} \widetilde{r}_i &= \frac{y_i - \boldsymbol{\theta}^T \mathbf{x}_i}{\widehat{\sigma}_0} \\ \widehat{v}_0 &= \frac{\text{ave}_i \psi'_{c_0}(\widetilde{r}_i)}{\text{ave}_i(\psi_{c_0}(\widetilde{r}_i)\widetilde{r}_i)} \\ \widetilde{\psi}_i^{(k)} &= \frac{\psi_{c_k}(\widetilde{r}_i)}{\text{ave}_i \psi'_{c_k}(\widetilde{r}_i)}, \quad k = 0, 1 \\ \widehat{\delta}^2 &= \text{ave}_i \left(\left(\widetilde{\psi}_i^{(1)} - \widetilde{\psi}_i^{(0)} \right)^2 \right) \\ T_b^2 &= \frac{2(\tilde{\sigma} - \widehat{\sigma}_0)}{\widehat{v}_0 \widehat{\delta}^2 \widehat{\sigma}_0} \end{aligned}$$

Hvis nT_b^2 er større end 0.95-fraktilen af χ_p^2 fordelingen, afgives en advarsel om, at estimationen ikke kan forventes at være central, og slutninger draget fra endelig estimation kan være farlige.

Trin 4

Kovariansmatrixen for de estimerede kvotienter:

$$w_i = \frac{\psi_{c_1}(\tilde{r}_i)}{\tilde{r}_i}$$
$$\hat{\mathbf{C}} = \frac{\text{ave}_i(\mathbf{x}_i \mathbf{x}_i^T w_i)}{\text{ave}_i w_i}$$
$$\hat{\tau} = \frac{\text{ave}_i(\psi_{c_1}(\tilde{r}_i))^2}{(\text{ave}_i \psi'_{c_1}(\tilde{r}_i))^2}$$
$$\hat{\mathbf{V}} = \hat{\mathbf{C}}^{-1} \hat{\sigma}_0^2 \hat{\tau} n^{-1}$$

Kvotient estimererne $\hat{\theta}_1$ og kovarians matrixen $\hat{\mathbf{V}}$ bruges til test og konfidensintervaller, hvis ikke der er af givet advarsel under trin 3.

Litteratur

Dansk Skovforening (2004): *Regnskabsoversigter for dansk privatskovbrug. Beretning nr. 58, 2003, Genel del.* Dansk Skovforening. København.

Groes, N. (1979): Skoven er ikke bare træer. *Dansk Skovforenings Tidsskrift*. LXIV.

R Development Core Team (2004): *R: A language and environment for statistical computing.* R foundation for Statistical Computing, Vienna. <http://www.R-project.org>

Ripley, B. (2005): *The VR Package.* The Comprehensive R Archive Network. Tilgængeligt på internettet: <http://mirrors.dotsrc.org/cran/doc/packages/VR.pdf>

Thorsen, B. J. 2003a. Skovbruget i 1900-tallet – erhverv og investering, *Skoven*, nr. 11 pp. 517-519.

Thorsen, B.J. (2003b): Skovbruget i 2000-tallet – investering eller filantropi, *Skoven*, nr. 11 pp. 520-523.

Venables, W.N. og B.D. Ripley (2002). *Modern Applied Statistics with S.* 4th edn. Springer-Verlag, Berlin.

Yohai, V.J.; W. Stahel og R.H. Zamar (1991): A procedure for robust estimation and inference in linear regression. In: *Directions in Robust Statistics and Diagnostics.* Eds. W. Stahel og S. Weisberg. Springer-Verlag. Berlin. Part II, pp. 365-374.