

"Kaos" og ikke-lineariteter i de finansielle markeder i Danmark

Allan Timmermann

Department of Economics, Birkbeck College, University of London

SUMMARY: The paper shows that there is evidence of non-linear dynamic components in a number of daily returns series for Danish stocks and exchange rates, although nothing conclusively shows that the non-linearities are of a chaotic nature. Based on a non-parametric nearest neighbour algorithm which is applicable in the presence of stochastic as well as deterministic non-linearities we show that it is possible to predict a statistically significant proportion of the variation in daily returns. The predicted values can be used to form a portfolio which mean-variance dominates a buy-and-hold strategy in the Copenhagen Stock Index provided that transaction costs are very low.

1. Indledning

Udviklingen af nye metoder til at analysere ikke-lineære systemer hvor dynamikken er deterministisk ("kaos") har vakt megen opmærksomhed inden for en række discipliner såsom biologi, fysik, økonomi¹ og finansiering. Ikke overraskende har flere empiriske studier (Hsieh (1992)), Scheinkman og LeBaron (1989) søgt at anvende lange tidsserier til at identificere en kaos-komponent i finansielle markeder. Disse studier har primært fokuseret på det amerikanske aktiemarked. Formålet med denne artikel er at analysere den statistiske og økonomiske betydning af identificerbare ikke-lineære komponenter i en række danske tidsserier og diskutere om disse komponenter er kaotiske.

Ikke-lineære stokastiske modeller er vidt udbredte indenfor økonomi, og den mest kendte model er nok Engle's ARCH model (Engle (1982)). Kaos kan defineres som deterministisk ikke-lineær dynamik, og udgør blot en delmængde af ikke-lineære dynamiske systemer. Der er således et langt skridt fra at finde, at en tidsserie er ikke-lineær til at konkludere, at serien er kaotisk. "Kaos" er en uheldig betegnelse for deterministisk ikke-lineær dynamik, idet en "kaotisk" model kan være særdeles simpel og fuldt ud forstået af forskere. Derfor er det mere dækkende at bruge betegnelsen kompleksitet, idet dette begreb bedre beskriver det vigtige princip bag ikke-lineære deterministiske systemer, nemlig deres evne til at genere komplekse tidsserier. Fra et heuristisk synspunkt har kaos en speciel interesse, siden viden om deterministiske

Jeg er taknemmelig for kommentarer fra en anonym referee samt fra Erik Reimer Larsen.

1. For en oversigt omkring kaos-modellers anvendelse inden for økonomi se Baumol og Benhabib (1989).

"Kaos" og ikke-lineariteter i de finansielle markeder i Danmark

Allan Timmermann

Department of Economics, Birkbeck College, University of London

SUMMARY: The paper shows that there is evidence of non-linear dynamic components in a number of daily returns series for Danish stocks and exchange rates, although nothing conclusively shows that the non-linearities are of a chaotic nature. Based on a non-parametric nearest neighbour algorithm which is applicable in the presence of stochastic as well as deterministic non-linearities we show that it is possible to predict a statistically significant proportion of the variation in daily returns. The predicted values can be used to form a portfolio which mean-variance dominates a buy-and-hold strategy in the Copenhagen Stock Index provided that transaction costs are very low.

1. Indledning

Udviklingen af nye metoder til at analysere ikke-lineære systemer hvor dynamikken er deterministisk ("kaos") har vakt megen opmærksomhed inden for en række discipliner såsom biologi, fysik, økonomi¹ og finansiering. Ikke overraskende har flere empiriske studier (Hsieh (1992)), Scheinkman og LeBaron (1989) søgt at anvende lange tidsserier til at identificere en kaos-komponent i finansielle markeder. Disse studier har primært fokuseret på det amerikanske aktiemarked. Formålet med denne artikel er at analysere den statistiske og økonomiske betydning af identificerbare ikke-lineære komponenter i en række danske tidsserier og diskutere om disse komponenter er kaotiske.

Ikke-lineære stokastiske modeller er vidt udbredte indenfor økonomi, og den mest kendte model er nok Engle's ARCH model (Engle (1982)). Kaos kan defineres som deterministisk ikke-lineær dynamik, og udgør blot en delmængde af ikke-lineære dynamiske systemer. Der er således et langt skridt fra at finde, at en tidsserie er ikke-lineær til at konkludere, at serien er kaotisk. "Kaos" er en uheldig betegnelse for deterministisk ikke-lineær dynamik, idet en "kaotisk" model kan være særdeles simpel og fuldt ud forstået af forskere. Derfor er det mere dækkende at bruge betegnelsen kompleksitet, idet dette begreb bedre beskriver det vigtige princip bag ikke-lineære deterministiske systemer, nemlig deres evne til at genere komplekse tidsserier. Fra et heuristisk synspunkt har kaos en speciel interesse, siden viden om deterministiske

Jeg er taknemmelig for kommentarer fra en anonym referee samt fra Erik Reimer Larsen.

1. For en oversigt omkring kaos-modellers anvendelse inden for økonomi se Baumol og Benhabib (1989).

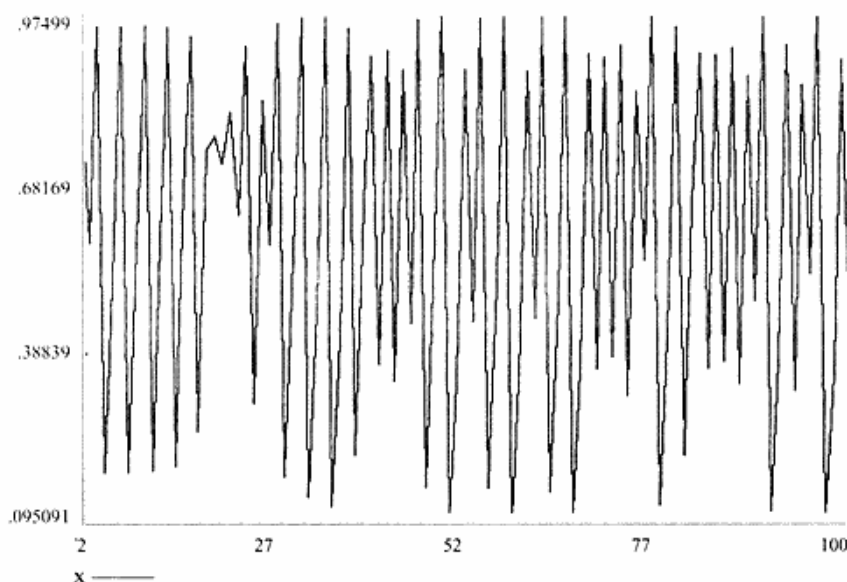
komponenter i en økonomisk serie kan lede til en bedre forståelse af det underliggende system og tillige kan anvendes til forudsigelse og kontrol af systemet.

Det er væsentligt at gøre sig den økonomiske betydning af ikke-lineariteter på de finansielle markeder klar. Standard teori for finansielle markeder definerer et efficient marked som et marked, hvor al offentlig tilgængelig information til enhver tid er reflekteret i kurserne. Hvis denne betingelse er opfyldt, vil nutidig information ikke kunne bruges til at forudsige fremtidige kursændringer. Det er imidlertid *ikke* en konsekvens af den efficiente markedshypotese, at aktiekurserne følger en lineær process, og den oftest anvendte model for prisdannelsen på de finansielle markeder, Lucas (1978), indebærer ikke at aktiekurserne er lineære. For det andet er mange anvendte modeller – specielt de såkaldte tekniske eller "chartist" – af natur ikke-lineære. Ikke-lineær dynamik kan også skyldes agents læreprocesser, idet agenterne ikke kan antages at kende den "sande" underliggende model for økonomien og gennem deres læreproces introducerer ikke-lineære mønstre i økonomiske variable (Timmermann (1993b)).

Tekniske metoder til analyse af kaos-komponenter i tidsserier er primært udviklede for systemer bestående af et meget stort antal observationer for en enkelt variabel. Vores analyse vil således fokusere på daglige aktiekurser og to kronekurser. Denne begrænsning skal ikke tolkes som en tro på, at andre makroøkonomiske og 'fundamentals' faktorer ikke spiller en rolle for de finansielle markeder. For længere tidshorisonter end den vi vil anvende i vores papir, har Fama & French (1989), Campbell (1987) og Timmermann (1993a) vist, at forskellige makroøkonomiske indikatorer såsom renter, risikopræmier, inflationsrater og udbytterater kan anvendes til at forudsige konjunkturelle komponenter i aktiekurserne. Derfor skal de ikke-lineære analysemetoder kun anvendes til høj-frekvens data og ikke til at teste og forudsige månedlige data eller data målt for endnu længere tidshorisonter.

Fokuseringen på individuelle tidsserier kan forsvares fra tre synspunkter. For det første fordi vi betragter daglige kurser, og variationen i 'fundamentals' for denne tidshorizont ikke er så væsentlig som for længere perioder. For det andet fordi data for BNP, inflationsraten etc kun er tilgængelige på måneds- eller kvartalsbasis. Endelig er det i sig selv interessant at se i hvilken udstrækning, vi kan forklare dynamikken i en række aktiekurser uden at referere til eksogene faktorer. En mulig tolkning af den interne prisdynamik i de finansielle markeder er således, at den repræsenterer investorernes psykologiske reaktionsmønstre. Vi understreger dog, at vores approach er ateoretisk, idet formålet med denne artikel er at undersøge den empiriske betydning af ikke-lineariteter i nogle finansielle markeder i Danmark. Det er op til efterfølgende studier at forklare ikke-lineariteter fra et teoretisk perspektiv.

I den resterende del af dette papir vil vi analysere tre punkter, der er vigtige i en analyse af den økonomiske betydning af ikke-lineariteter i de finansielle markeder. Først vil vi diskutere, hvordan vi kan teste for eksistensen af ikke-lineariteter i tidsse-



Figur 1. Den Logistiske ligning $x = 3.9x(-1)*(1-x(-1))$.

rier. Dette følges op af en analyse af hvorledes vi kan forudsige data gennem anvendelse af en algoritme, der kan opfange ikke-lineære mønstre. Endelig vil vi analysere den økonomiske størrelse af afkastet fra en investeringsstrategi baseret på disse forudsigelser.

2. Hvad er kaos?

Vi vil indlede vores analyse med en intuitiv definition af kaos, som vi vil anvende i vores undersøgelse af finansielle tidsserier. Populært sagt er kaos en deterministisk serie, der ser "tilfældig ud". Mere præcist siges en tidsserie at være kaotisk, hvis den udviser

1. Følsomhed over for begyndelsesbetingelser.
2. Stationaritet.
3. Ingen cyklisk gentagende perioder.

Den første betingelse indebærer, at små ændringer i kursen i dag kan have en utrolig stor betydning for fremtidige kurser. Dette er i modstrid med standard lineære modeller, der er betydeligt mere stabile, og hvor små ændringer i dag typisk ikke har nogen dramatisk effekt på kurserne i fremtiden. Den anden betingelse udelukker dataserier for perioder, hvor der har været institutionelle og/eller makroøkonomiske regime-skift af betydning for de finansielle markeder. Endelig udelukker den tredje betingelse sinus- og andre trigonometriske funktioner og den udelukker også historiske episoder, der gentages præcist i fremtiden.

I praksis vil vi ikke forvente, at tidsserier er "rene" kaos serier eller "rene" stokastiske serier. Men selv i en ideel situation hvor den studerede serie er deterministisk, kan det være svært at bestemme, om serien er kaotisk eller stokastisk. Man skal især gøre sig to punkter klart, når man søger at anvende principperne bag kaos i den praktiske analyse af data. For det første kan tidsserier genereret af kaos nogle gange udvise skarpe kvalitative spring som nemt kan forveksles med choks i en stokastisk serie. Derfor kan standard tests for deterministiske elementer ofte heller ikke skelne mellem kaotiske mønstre og mange ikke-lineære stokastiske processer.

For det andet kan kaotiske tidsserier være ekstremt følsomme over for mikroskopiske ændringer i parameterverdier for den underliggende proces. Dette punkt gør det svært at estimere modeller med kaos, og gør det også vanskeligt at forudsige fremtidige værdier for serier med kaos. Tænk bare på den såkaldte logistiske differensligning $x_t = ax_{t-1}(1-x_{t-1})$.

Figur 1 viser en tidsserie for den logistiske differensligning for $a = 3,9$, dvs. $x_t = 3,9x_{t-1}(1-x_{t-1})$. De kvalitative ændringer og store fluktuationer i denne deterministiske tidsserie fremgår tydeligt. Lineære stokastiske modeller kan kun generere fire typer af dynamiske mønstre: (i) stabile svingninger, (ii) ustabile svingninger, (iii) stabile uden svingninger og (iv) eksplosive uden svingninger. I modsætning til dette kan modeller med kaos generere langt mere komplekse tidsserier, der fuldt ud matcher den dynamiske kompleksitet, som vi finder i finansielle data.

3. En test for eksistensen af kaos i tidsserier

Det er generelt meget svært at påvise eksistensen af kaos i en tidsserie. En række betingelser skal være opfyldt for, at vi kan gøre os håb om at teste eksistensen af kaos for et økonomisk system. For det første, at det betragtede økonomiske system er af "lav kompleksitet" – dvs. kan indlejres i et system, der har lav dimension. Hvis det økonomiske system er meget komplekst, er det i praksis umuligt at "rekonstruere" den underliggende model, og systemet kan lige så godt betragtes som værende stokastisk.

For det andet skal måleunøjagtigheder i data samt stokastiske støjkluder være af relativ lille betydning i forhold til den deterministiske komponent i data. Det er i forvejen svært nok at skelne mellem kaotisk og stokastisk dynamik. Hvis derfor det underliggende system er en blanding af disse to kilder, vil det i praksis være umuligt at reddegøre for kaoskomponenten (Granger (1991)).

Antag at vores økonomiske system afhænger af n faktorer (observerbare såvel som ikke-observerbare) sammenfattet i en n -dimensional vektor x_t . Den deterministiske lov ("law of motion") for systemet vil vi repræsentere i form af en funktion, f :

$$f: R^n \rightarrow R^n: x_{t+1} = f(x_t), \quad (1)$$

hvor t er det betragtede tidspunkt, og $t+1$ er punktet for en periode frem i tiden. Det sædvanlige problem, som vi står overfor, er, at vi ikke kender det "sande" økonomiske system og derfor hverken kender n – antallet af relevante komponenter i x_t – eller hvilke variable x_t består af. Den bedste alternativ i denne situation er at betragte en vektor bestående af N observerbare økonomiske variable, m_t , som afhænger af det økonomiske system gennem en funktion, som vi vil kalde g :

$$g: R^n \rightarrow R^N: m_t = g(x_t), \quad (2)$$

N er den såkaldte indlejningsdimension. Princippet er, at vi kan lære om det underliggende økonomiske system, x_t , gennem de observerbare variable i m_t . For enhver værdi af indlejningsdimensionen N (antallet af elementer i m_t) kan vi definere en historie af systemet af længde N :

$$\Phi_N(x) = \{g(x), g[f(x)], \dots, g[f^{N-1}(x)]\}. \quad (3)$$

Φ_N er en funktion fra $R^n \rightarrow R^N$. Undersøgelsen af kaos bygger på et resultat, der er kendt som Takens Embedding Theorem (Takens 1981)). Dette teorem siger løseligt, at under nærmere definerede tekniske betingelser for funktionen g , kan al relevant information om x_t udledes fra m_t såfremt $N \geq 2n+1$. Med andre ord: prisen for ikke at kende identiteten af de underliggende økonomiske faktorer er, at vi skal inkludere lidt over dobbelt så mange variable i vores system ($2n+1$), som det havde været nødvendigt, hvis vi kendte identiteten af det "sande" økonomiske system (n). Dette er en lav pris at betale, såfremt n er lille. Vi betaler for vores uvidenhed om det sande økonomiske system gennem anvendelsen af flere datapunkter.

3.1. Ikke-lineariteter i data

Mange statistiske tests for nulhypotesen om, at data er uafhængigt og identisk stokastisk fordelte, opfanger blot skift i den datagenererende mekanisme (såkaldt ikke-stationaritet). For at håndtere dette problem anvender vi specielle tests, der er designet til at opfange ikke-lineariteter. Lad $X_1, X_2, \dots, X_T$ være en tidsserie af længde T . For hver værdi af indlejningsdimensionen N kan vi konstruere N -dimensionale vektorer $Y_i = (X_i, \dots, X_{i+N-1})$ ($i = 1, \dots, T-N+1$). Korrelationsintegralet $C_T(\gamma, N)$ er defineret som andelen af par (Y_i, Y_j) , som ligger inden for en afstand γ af hinanden. Vi vil anvende følgende definition

$$X_{ij} = \begin{cases} 1, & \text{hvis } \|Y_i - Y_j\| < \gamma \\ 0, & \text{hvis } \|Y_i - Y_j\| \geq \gamma \end{cases} \quad (4)$$

hvor $\|\cdot\|$ er det absolutte afstandsmål. Korrelationsintegralet er nu defineret som følger

$$C_T = \frac{2}{(T-N)(T-N+1)} \sum_{i=1}^{T-N} \sum_{j=i+1}^{T-N+1} X_{ij} \quad (5)$$

Lad korrelationskoefficienten, $C(\gamma)$, være lig grænseværdien $(\lim) \lim_T \rightarrow \infty C_T(\gamma, N)$. I denne definition antages det, at uendeligt mange observationer (T) er tilgængelige. γ er et filter i følgende forstand: Alle afstande mindre end γ tolkes som naboer og bidrager til korrelationsintegralet, mens afstande større end γ ignoreres. Korrelationsintegralet er således baseret på et mål for "tæthed" mellem datapunkterne. I vores test anvender vi de observerbare punkter m_i . For at få en ide om hvor mange beregninger der skal foretages, kan det nævnes, at hvis vi blot har 1000 data punkter, skal der beregnes ca. 500.000 afstande, idet alle mulige afstande mellem disse punkter skal beregnes og sammenlignes med γ .

Ideen bag vores test har en simpel grafisk forklaring. Hvis eksempelvis x består af to variable (x_1, x_2) , der er relaterede gennem en funktionel regel $x_2 = f(x_1)$, vil alle (x_1, x_2) punkter ligge omkring en graf i et todimensionalt diagram, der plottes x_2 mod x_1 . Når indlejningsdimensionen vokser, vil antallet af naboer således ikke vokse proportionalt med N , idet hele det N -dimensionale rum ikke er fyldt ud. Dette skal ses i kontrast til en situation hvor x_1 og x_2 er uafhængigt fordelte, og hvor (x_1, x_2) punkterne fylder hele diagrammet ud. I denne situation har x_1 naboer til alle fire sider (og ikke blot langs en linie), således at korrelationsintegralet bliver fordoblet, når vi går fra et endimensionalt til et todimensionalt billede. Denne logik holder også for større værdier af n . Hvis data er genereret af stokastisk støj, vil korrelationsintegralet være lig indlejningsdimensionen, N , idet et plot af $(x_1, x_2, \dots, x_N)$ fylder hele det N -dimensionale rum ud. Hvis data er genereret af kaos, vil korrelationsintegralet derimod ikke vokse som funktion af indlejningsdimensionen for $N \geq 2n+1$. Opgaven er således at finde en tilstrækkelig stor værdi for indlejningsdimensionen, således at datapunkterne ikke fylder hele det N -dimensionale rum ud. Bemærk at korrelationsdimensionen ikke behøver at være et naturligt tal, og at den ikke repræsenterer antallet af økonomiske faktorer, der driver systemet (n). Korrelationsdimensionen er defineret som den værdi af N , for hvilke korrelationskoefficienten når sit maksimum.

Der er to metoder til at beregne korrelationsintegralet, Grassberger-Procaccias metode og Takens metode. Begge metoder er baseret på de observerbare punkter m_i frem for de ukendte faktorer, x_i . Vi anvendte Taken's metode i vores beregninger og har beskrevet denne beregningsteknik i det tekniske appendiks. For begge metoder går testen for kaos ud på at plote en graf for det beregnede korrelationsintegral (y -aksen) som funktion af indlejningsdimensionen (x -aksen). Vi har følgende tilfælde:

Tabel 1. Takens' mål for korrelations-koefficienten.

Indlejr. dimension	København Fondsbørs	Carlsberg B	DS1912A	Danske Bank	DFDS	Jyske Bank	Dollar/ DKK	ECU/DKK
1	0.92	1.01	0.95	0.96	0.97	1.04	0.88	0.51
2	1.04	1.14	1.07	1.12	1.20	1.25	0.99	0.63
3	1.15	1.30	1.18	1.31	1.41	1.48	1.04	0.73
4	1.26	1.47	1.29	1.50	1.64	1.78	1.11	0.79
5	1.39	1.71	1.39	1.75	1.96	2.11	1.18	0.83
6	1.50	1.86	1.53	1.93	2.15	2.32	1.26	0.84
7	1.65	2.08	1.68	2.21	2.41	2.66	1.31	0.85
8	1.83	2.25	1.80	2.39	2.48	2.88	1.43	0.87
9	1.98	2.55	1.90	2.68	2.82	3.28	1.50	0.88
10	2.11	2.98	1.94	2.95	2.85	3.66	1.65	0.91
15	2.96	3.90	2.69	4.14	3.96	7.09	2.16	1.02
20	3.71	4.05	3.37	7.58	3.25	21.74	2.64	1.10

(i) Data er stokastisk; Grafen vil være en ret linje

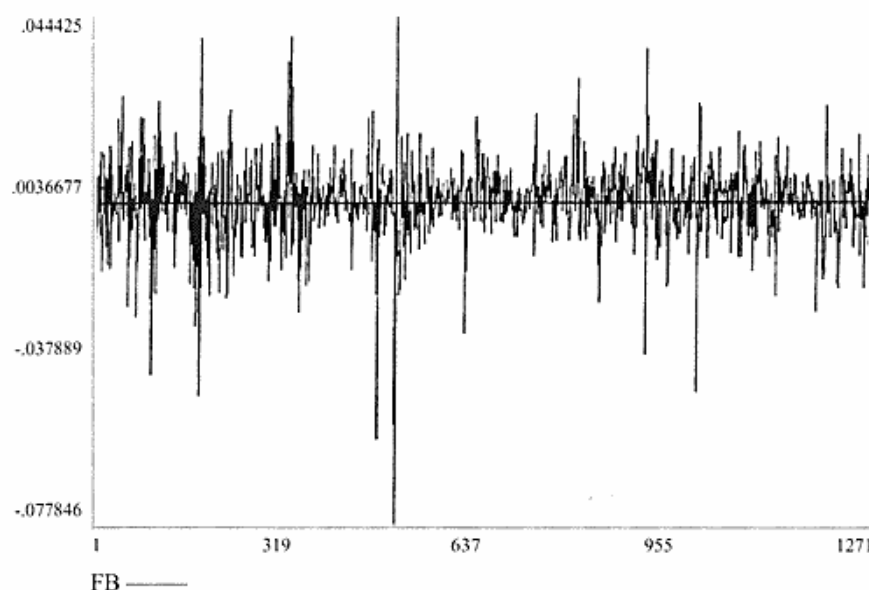
(II) Det økonomiske system indeholder uendeligt mange faktorer (dvs. N er meget stor); Grafen vil være en ret linje.

(III) Kaos med en lav indlejringsdimension; Grafen vil stige i begyndelsen, men herefter flade ud og tilnærmelsesvist være konstant.

3.2. Empiriske resultater for danske data

Vi analyserede en række finansielle tidsserier for Danmark over perioden 1980-1990. DataStream blev anvendt som kilde. Den præcise måleperiode afviger fra aktiv til aktiv afhængigt af, hvornår DataStream begyndte at have daglige observationer. Data blev rensset for punkter, hvor kurserne var præcis den samme to dage i træk for at udelukke episoder, hvor en aktie ikke blev handlet, eller markedet var lukket. De analyserede tidsserier og det anvendte antal observationer var som følger: Københavns Fondsbørsindex (1777), Carlsberg B (1540), DS1912 A (1507), Danske Bank (1664), DFDS (1622), Jyske Bank (1522), Dollar/DKK (1187) og ECU/DKK (2853). Figur 2 viser de daglige afkast for Københavns Fondsbørsindeks for de sidste 1271 dage i undersøgelsen og Figur 3 viser autokorrellogrammet for denne tidsserie. Tidsserien ligner tilnærmelsesvist "hvid støj".

Tabel 1 viser de beregnede værdier af korrelationsintegralet for ovenstående data baseret på Takens beregningsmetode. For de fleste serier er der tegn på ikke-linearitet, idet de estimerede korrelationsintegraler ikke stiger proportionalt med værdien af indlejringsdimensionen. Det kan ikke konkluderes, at der er kaos i serierne, idet visse former for stokastisk ikke-linearitet (f.x. ARCH) kan skabe lignende mønstre i korrelationsintegralerne. Det skal understreges, at den statistiske usikkerhed ved bestem-



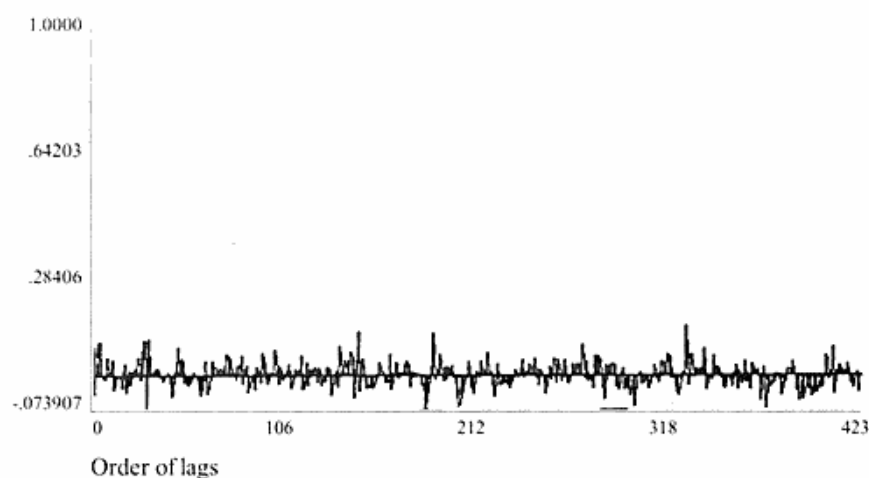
Figur 2. Daglige afkast for Københavns Fondsbørs Indeks.

melsen af korrelationsintegralet for en indlejningsdimension højere end 6-7 er meget stor for vores forholdsvis små datasæt.

Bemærk at løsningen på problemet med for små datasæt ikke nødvendigvis er at indsamle data meget ofte. Grunden er, at data i så fald kan indeholde for megen støj fra bid-ask spreads og lignende. Kunsten er at indsamle data tilstrækkeligt ofte til, at data er indsamlet i et og samme økonomiske regime, uden at data er sløret af for megen idiosynkratisk støj.

Som nævnt i indledningen kan vi imidlertid ikke konkludere fra indikationerne af ikke-lineære komponenter i de finansielle tidsserier, at der er kaos i disse serier. Der er flere faktorer der peger på en naturlig skepsis over for kaos i økonomiske tidsserier. For det første er der ikke nogen naturlig tidsskala i de fleste økonomiske systemer. Selv hvis der er mekanismer, der genererer kaos i markederne, vil vi med al sandsynlighed ikke kunne finde disse kaosmønstre, hvis vi anvender kalendertid i stedet for en tidsskala relateret til begivenhederne der introducerer kaos. Eksempelvis kunne vi lade salgsvolumen være vores tidsmåler, hvis vi tror at aktivitet i markederne er relateret til transaktionsniveauet.

For det andet er de fleste økonomiske tidsserier aggregerede (BNP etc.), og aggregering (i sig selv en lineær transformation) kan fjerne ikke-lineariteter i en tidsserie. Selvom der er kaos på det mikroøkonomiske plan, er det langt fra sikkert, at dette vil kunne påvises gennem anvendelse af makroøkonomiske tidsserier (Granger (1991)). Endelig er der problemet med måleunøjagtigheder og stokastisk støj i mange tidsserier, der kan hindre opdagelsen af underliggende ikke-lineære mønstre, hvis systemets signal-til-støj forhold er lavt.



Figur 3. Autokorrelation funktion for Københavns fondsbørsindeks.

4. Ikke-parametrisk forudsigelse af daglige kursændringer

Hvis en tidsserie indeholder deterministiske komponenter, er det logisk at spørge, om disse komponenter kan anvendes til at forudsige fremtidige værdier gennem anvendelsen af historisk information. Det er helt klart ikke nogen nem opgave, vi her bliver stillet, siden den deterministiske komponent er kaotisk og typisk er svær at identificere. Forsøg på at anvende specifikke former for ikke-lineær dynamik vil ofte ikke have succes, idet usikkerhed omkring estimerne kan påvirke de forudsagte værdier ganske dramatisk (tænk på den logistiske funktion). For at undgå denne skæbne vil vi bruge en ikke-parametrisk metode, der ikke bygger på en speciel antagelse om den underliggende model. En sådan metode er klart ateoretisk, idet vi ikke søger at estimere en model baseret på teoretiske antagelser. Til gengæld behøver vi ikke at lægge os fast på en bestemt model, som genererer ikke-lineariteterne i tidsserierne. Dette er specielt en fordel i lyset af resultaterne i den foregående sektion, hvor vi udtrykte skepsis over for eksistensen af kaos i de undersøgte tidsserier. Vores ikke-parametriske metode kan anvendes uafhængigt af, om dynamikken er stokastisk og lineær, stokastisk og ikke-lineær eller kaotisk.

Vi vil anvende følgende såkaldte "nearest neighbour" procedure foreslået af Sugihara og May (1990) og Linden, Satchell og Yoon (1992):

- (1) konstruer historiske tidsserier ($m_t, m_{t-1}, \dots, m_{t-N+1}$) af observerede data for indlejningsdimensionen $N = 2, \dots, 15$.
- (2) lad de foregående 500 observationer udgøre et historisk bibliotek af data.
- (3) for alle tidligere observationer i biblioteket find de P nærmeste punkter ($P = 10$ i vores undersøgelse) og for disse punkter beregn vægte således, at jo længere væk en historisk observation er fra punktet i dag, jo lavere vægt får denne observation i forudsigelsen.

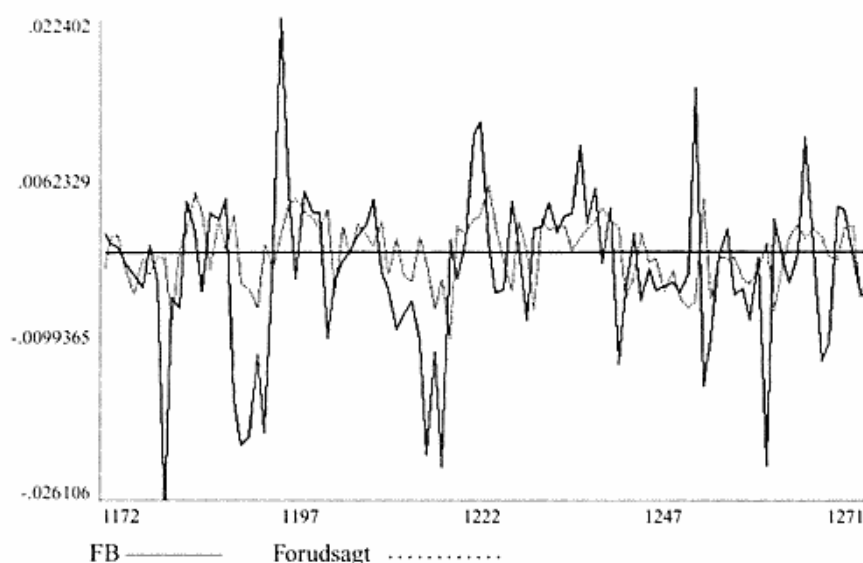
Tabel 2. Korrelation Mellem Realiserede og Forudsagte Værdier for Daglige Afkast.

Indlejr. Dimension	Københavns Fondsbørs	Carlsberg B	DS1912A	Danske Bank	DFDS	Jyske Bank	Dollar/ DKK	ECU/DKK
1	.120 (4.87)	-.007 (-.23)	.058 (1.93)	.115 (4.43)	.131 (5.04)	-.010 (-.32)	.006 (.15)	0.40 (2.05)
2	.164 (6.99)	.016 (.53)	.045 (1.50)	.072 (2.63)	.082 (2.99)	.028 (.91)	.030 (.81)	.018 (.89)
3	.117 (4.73)	.053 (1.80)	-.013 (-.41)	.095 (3.55)	.122 (4.66)	.070 2.40	-.044 (-1.09)	-.001 (-.05)
4	.117 (4.74)	.053 (1.79)	-.023 (-.70)	.080 (2.97)	.097 (3.60)	.091 3.17	.029 (.77)	.000 (.02)
5	.148 (6.20)	.034 (1.14)	-.004 (-.14)	.055 (1.97)	.120 (4.54)	.078 2.71	.015 (.39)	0.25 (1.24)
6	.120 (4.86)	.052 (1.76)	.014 (.45)	.078 (2.88)	.116 (4.40)	.062 2.12	.047 (1.29)	.049 (2.47)
7	.108 (4.29)	.070 (2.41)	.005 (.16)	.049 (1.75)	.099 (3.68)	.067 2.11	.057 (1.57)	.067 (3.47)
8	.100 (3.96)	.083 (2.90)	.005 (.16)	.045 (1.62)	.104 (3.89)	.080 2.77	.043 (1.17)	.053 (2.67)
9	.122 (4.45)	.022 (.75)	-.016 (-.51)	.032 (1.15)	.100 (3.73)	.144 5.38	.047 (1.29)	.032 (1.57)
10	.106 (4.22)	.050 (1.69)	.014 (.43)	.041 (1.46)	.090 (3.29)	.099 3.51	.000 (.011)	.037 (1.86)
11	.095 (3.74)	.012 (.40)	.009 (.28)	.046 (1.64)	.098 (3.61)	.099 3.51	.007 (0.19)	.045 (2.27)
12	.079 (3.06)	.006 (.21)	-.003 (-.10)	.042 (1.48)	.084 (3.05)	.062 2.11	.029 (.77)	0.068 (3.50)
13	.097 (3.83)	-.005 (-.16)	.003 (1.12)	.049 (1.75)	.060 (2.11)	.114 4.11	.034 (.93)	.062 (3.18)
14	.079 (3.07)	-.005 (-.17)	.002 (.06)	.064 (2.31)	.090 (3.31)	.074 2.56	.023 (.63)	.062 (3.17)
15	.058 (2.20)	-.027 (-.83)	.011 (.35)	.085 (3.15)	.081 (2.95)	.010 3.56	.030 (.81)	.042 (2.10)

Anm.: Forudsigelserne er baseret på en »nearest neighbour« algoritme der anvendte de 10 nærmeste nabopunkter til at forudsige afkastet for den næste dag. Den anvendte biblioteksstørrelse var 500 observationer.

(4) Anvend vægtene fra punkt (3) til at forudsige værdien af kursen et trin frem i tiden.

Vi anvendte geometriske vægte under punkt (3), men andre papirer har anvendt "tri-cubic" vægte med lignende resultater (Diebold og Nason (1989)). Vores algoritme er baseret på den samme antagelse om "tæthed" mellem observationerne, som korrelationsintegralet bygger på. Bemærk at vi holder størrelsen af biblioteket konstant (500 observationer, svarende til cirka to år for daglige data) for at gardere os mod gradvise skift i den datagenererende proces.



Figur 4. Faktiske (FB) og forudsagte værdier for 100 daglige ændringer i Københavns fondsbørs index (slutningen af 1990).

Hvis den anvendte indlejningsdimension (N) er større end den sande værdi, introduceres støj i systemet og forudsigeligheden af daglige kursændringer kan blive negativt påvirket: støjen fra tidligere punkter, der ligger langt fra hinanden, kan forstyrre vores vægtprocedure under punkt (3). Imidlertid er det bedre at anvende en større indlejningsdimension end den sande korrelationsdimension frem for at anvende en for lav indlejningsdimension: med en for lav indlejningsdimension vil systematiske mønstre i en tidsserie ikke kunne forklares.

Fordelen ved den anvendte metode til forudsigelse af daglige kursændringer er, at mens lineære modeller kun kan opfange simple mønstre i aktiekurserne (og i praksis ikke er succesrige i denne henseende), så kan vores metode opfange tilbagevendende ikke-lineære mønstre såsom "heads and shoulders", der er velkendte fra anvendelsen af charts. I denne henseende kan metoden ses som en raffineret udgave af ateoretiske teknikker. Ydermere er vores resultater robuste over for kritikken af teknisk analyse: hvis ikke der er noget om troen på tilbagevendende mønstre, vil vores metode heller ikke finde sådanne tegn, og i praksis give resultater meget lig resultaterne fra en lineær model. Den ikke-parametriske procedure har vist sig velegnet til at forudsige en række deterministiske og stokastiske ikke-lineære processer (Linden, Satchell og Yoon (1992)).

4.1. Resultater for de danske finansielle markeder

Ovenstående metode anvendtes til at forudsige de daglige kursændringer for de undersøgte finansielle tidsserier. Resultatet fremgår af Tabel 2, og Figur 4 viser de faktiske og forudsagte værdier for 100 daglige ændringer i Københavns Fondsbørs Indeks

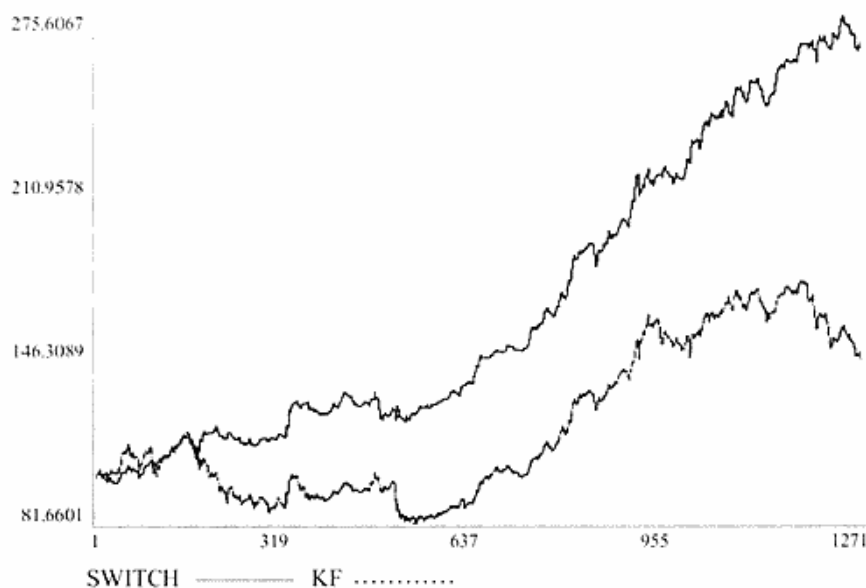
i slutningen af 1990. Tabel 2 viser korrelationen mellem de forudsagte procentvise ændringer en dag frem i tiden og den faktiske procentvise ændring. desuden giver tabellen t -værdier for den statistiske signifikans af sammenhængen mellem de forudsagte og de faktiske daglige kursændringer. T -værdier større end 2 viser, at der med 95 procent sandsynlighed er en signifikant positiv sammenhæng mellem forudsagte og faktiske værdier. For sådanne t -værdier tilbageviser vi nul-hypotesen om, at sammenhængen mellem de to tidsserier er ren tilfældighed.

Det fremgår af Tabel 2, at vores forudsigelser af daglige kursændringer er statistisk signifikante for Københavns Fondsbørindeks, Carlsberg B, Den Danske Bank, DFDS, Jyske Bank og Ecu'en, hvorimod der ikke er tegn på forudsigelige mønstre i kursændringerne for DS1912 A og Dollarkursen. Det er interessant at bemærke, hvor succesfulde vi er i at forudsige Københavns Fondsbørsindeks, hvor vi for en relativ lav indlejringsdimension på 2 opnår en meget høj værdi af korrelationen mellem faktiske og forudsagte værdier (0.16), hvilket giver sig udslag i en t -værdi på 6.99. Ligeledes skal det bemærkes, at der ikke synes at være noget systematisk mønster i for hvilke værdier af indlejringsdimensionen den højeste korrelation mellem forudsagte og faktiske kursændringer opnås. For nogle tidsserier synes den bedste indlejringsdimension at være ganske lav (Københavns Fondsbørindeks, Danske Bank, DFDS), mens det for andre tidsserier synes at være en høj værdi af indlejringsdimensionen (Jyske Bank, Carlsberg B, ECU/DKK).

Opmærksomme læsere kunne indvende, at daglige data i praksis repræsenterer et for kort tidsinterval, og at investorer snarere er interesseret i at ændre deres portefølje på månedsbasis eller sågar en gang hver kvartal. Imidlertid er der mange investorer på de finansielle markeder, der handler på basis af høj-frekvens (daglig) data, og som vil finde de daglige forudsigelser af kurserne nyttige. Fristelsen til at forudsige kurser en måned eller længere frem i tiden på basis af kaos skal man ikke falde for. Pointen ved ikke-lineær dynamik er netop, at forudsigelser langt fremme i tiden er umådeligt følsomme over for begyndelsesbetingelser og den estimerede model. Derfor er forudsigelser en uge eller længere frem i tiden baseret på en model for daglige kurser ikke bedre end et tilfældigt gæt.

5. Kan de ikke-parametriske forudsigelser udnyttes profitabelt i en investeringsstrategi?

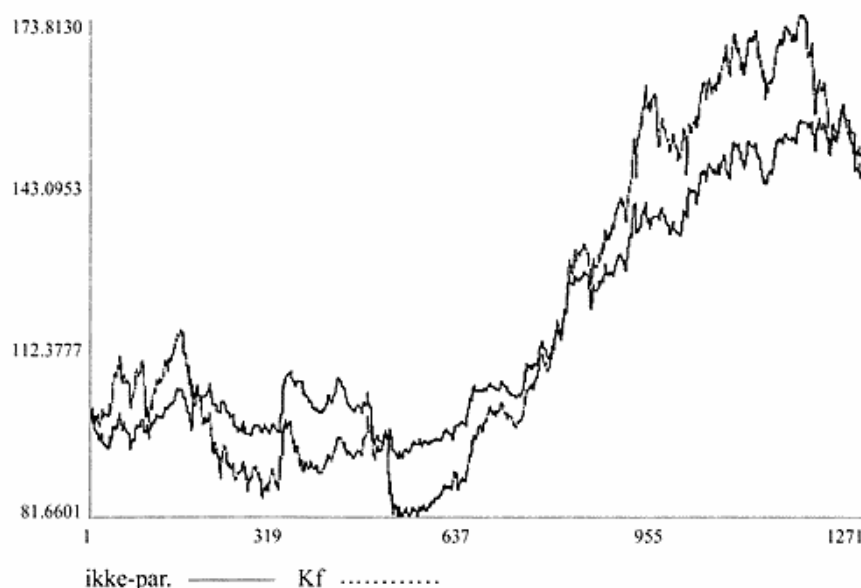
Som det afsluttende punkt i vores analyse anvendte vi de forudsagte kursændringer for Københavns Fondsbørindeks i en aktiv investeringsstrategi byggende på følgende simple princip: Hvis kursen forudsiges at gå op næste dag, hold en portefølje svarende til Fondsbørindekset; hvis kursen forudsiges at gå ned næste dag, sælg Fondsbørindekset og hold kontanter.



Figur 5. Værdien af at investere 100 dkr. i Københavns Fondsbørsindeks (KF) eller i den ikke-parametriske portefølje (switch).

En sådan strategi er kendt som en "switching strategy", fordi der skiftes mellem investeringsobjekter (aktier eller kontanter). Resultatet af at anvende denne strategi for Københavns Fondsbørsindeks på basis af 1270 daglige ændringer fremgår af Figur 5. Figuren viser værdien af at investere 100 kr. i henholdsvis Københavns Fondsbørs Aktieindeks eller i en portefølje baseret på forudsigelserne fra kaos-modellen. Værdien af vores portefølje ved afslutningen af de 1270 dage (cirka 5 år) er cirka den dobbelte af værdien af at holde fondsbørsindekset. Gennemsnitsafkastet per dag af de to porteføljer var .03229 procent (Københavns Fondsbørs Indeks) og .07792 procent (porteføljen baseret på de ikke-lineære forudsigelser), så det daglige afkast på vores portefølje er mere end dobbelt så stor som afkastet fra en passiv investering i Københavns Fondsbørs Indeks. Eftersom ikke alle aktier i Københavns Fondsbørsindeks handles dagligt, skal disse resultater ikke tolkes som et præcist udtryk for hvad en investor kunne have tjent ved at følge vores forudsigelser. Resultaterne tjener snarere som en illustration af størrelsesordenen af den økonomiske værdi af de ikke-parametriske forudsigelser.

Ovenstående analyse antog, at der ikke var nogen transaktionsomkostninger forbundet med at skifte mellem aktier og kontanter. Dette er selvfølgelig ikke tilfældet i praksis, men er en god approksimation for handel i et futures-indeks hvor transaktionsomkostningerne er yderst små. For ovenstående analyse fremgår det af Figur 6, at porteføljen baseret på ikke-parametriske forudsigelser giver et højere afkast end fondsbørsindekset, så længe transaktionsomkostningerne er mindre end 0.1 procent.



Figur 6. Værdien af at investere 100 kr. i KF indekset (KF) eller i den ikke-parametriske portefølje. Incl. transaktionsomkostninger.

Ydermere er volatiliteten af det daglige afkast fra den ikke-parametriske portefølje betydeligt lavere end volatiliteten af de daglige kursændringer for Københavns Fondsbørs Aktieindeks. Standardafvigelsen på de daglige afkast var .0086 og .0061, henholdsvis for KF indekset og porteføljen baseret på ikke-parametriske forudsigelser. Dette giver i sig selv porteføljen baseret på de ikke-lineære forudsigelser en økonomisk værdi, idet investorer typisk er villige til at betale for at undgå risiko (de er risiko averse).

Vi konkluderer, at det er muligt at basere en simpel investeringsstrategi på de ikke-parametriske forudsigelser af daglige kursændringer, der vil dominere fondsbørsindekset for meget små transaktionsomkostninger. Høje transaktionsomkostninger er imidlertid en barriere, som det er svært at overvinde, når vi anvender en investeringsstrategi baseret på ikke-lineære forudsigelser. For at finde ikke-lineære mønstre er det nødvendigt at have lange tidsserier, hvilket i praksis betyder høj-frekvens data for de finansielle markeder (daglige observationer). Derfor vil de ikke-parametriske forudsigelser generere signaler, der let kan indebære en meget stor volumen for køb og salg af finansielle aktiver. Fra dette følger problemet, at selv små transaktionsomkostninger kan fjerne den økonomiske fordel af at anvende ikke-lineære metoder.

6. Konklusion

Meget tyder på, at de daglige afkast for en række finansielle tidsserier i Danmark indeholder ikke-lineære komponenter, skønt der ikke er overbevisende tegn på, at der

er kaos i serierne. Sandsynligvis kan ikke-lineariteterne bedre modelleres og forstås som stokastiske ikke-lineariteter, hvilket det vil være en udfordring for fremtidige studier at undersøge. I vores analyse anvendte vi disse indikationer til at konstruere ikke-parametriske forudsigelser af daglige afkast. Analysen viste, at de forudsagte værdier var signifikant korrelerede med de faktiske værdier af daglige afkast. Endvidere demonstrerede vi, at disse forudsigelser kunne anvendes i en investeringsstrategi, der gav et større afkast end det gennemsnitlige afkast på Københavns Fondsbørsindeks, når transaktionsomkostningerne er meget lave. Når realistiske transaktionsomkostninger inkluderes, er profitten fra investeringsstrategien baseret på de ikke-parametriske forudsigelser ikke længere økonomisk signifikant.

Teknisk Appendiks

Dette appendiks beskriver hvorledes korrelationsintegralet kan beregnes ved hjælp af Takens (1984) maximum likelihood metode. For en bestemt værdi af "cutoff" punktet γ beskrevet i hovedteksten (f.x. 1-2 standard-afvigelser af data) antager vi at den følgende relation holder eksakt: $C(\gamma) = \gamma^\alpha$, således at korrelationsintegralet følger en "power law" for alle afstande mindre end eller lig γ . Alle afstande større end γ ignoreres og de tilbageværende afstande divideres med γ for at få en fordeling af afstande, d_i , defineret på intervallet $(0;1]$. Vi vil nu betragte vores afstande $d_1, \dots, d_m \in (0,1)$ som en tilfældig stikprøve fra en sandsynlighedsfordeling. Hvis vi antager, at disse afstande er uafhængigt fordelte vil, for en given værdi af α , sandsynligheden for at finde en stikprøve $(d_1, d_1 + \Delta d_1), \dots, (d_m, d_m + \Delta d_m)$ være lig $\prod_{i=1}^m d_i^\alpha$.

Derfor kan vi estimere α gennem anvendelse af maximum likelihood:

$$\hat{\alpha} = \frac{-m}{\sum \ln(d_i)},$$

hvor $\hat{\alpha}$ er den estimerede korrelationskoefficient.

Litteratur

- Baumol, W. og Benhabib, J. 1989. Chaos: significance, Mechanism, and Economic Applications. *Journal of Economic Perspectives*, 3, 77-105.
- Campbell, J.Y. 1987. Stock Returns and the Term Structure. *Journal of Financial Economics*, 18, 373-99.
- Diebold, F.X. og Nason, J.A. 1990. Nonparametric Exchange Rate Prediction? *Journal of International Economics*, 28, 315-32.
- Engle, R.F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- Fama, E.F. og French, K.R. 1989. Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics*, 25, 23-49.
- Granger, C.W.J. 1991. Developments in the Nonlinear Analysis of Economic Series. *Scandinavian Journal of Economics*, 93 (2), 263-76.
- Hsieh, D.A. 1991. Chaos and non-linear Dynamics: Applications to Financial Markets. *Journal of Finance*, 46, 1839-77.
- Linden, N., Satchell, S.E. og Yoon, Y. 1991.

er kaos i serierne. Sandsynligvis kan ikke-lineariteterne bedre modelleres og forstås som stokastiske ikke-lineariteter, hvilket det vil være en udfordring for fremtidige studier at undersøge. I vores analyse anvendte vi disse indikationer til at konstruere ikke-parametriske forudsigelser af daglige afkast. Analysen viste, at de forudsagte værdier var signifikant korrelerede med de faktiske værdier af daglige afkast. Endvidere demonstrerede vi, at disse forudsigelser kunne anvendes i en investeringsstrategi, der gav et større afkast end det gennemsnitlige afkast på Københavns Fondsbørsindeks, når transaktionsomkostningerne er meget lave. Når realistiske transaktionsomkostninger inkluderes, er profitten fra investeringsstrategien baseret på de ikke-parametriske forudsigelser ikke længere økonomisk signifikant.

Teknisk Appendiks

Dette appendiks beskriver hvorledes korrelationsintegralet kan beregnes ved hjælp af Takens (1984) maximum likelihood metode. For en bestemt værdi af "cutoff" punktet γ beskrevet i hovedteksten (f.x. 1-2 standard-afvigelser af data) antager vi at den følgende relation holder eksakt: $C(\gamma) = \gamma^\alpha$, således at korrelationsintegralet følger en "power law" for alle afstande mindre end eller lig γ . Alle afstande større end γ ignoreres og de tilbageværende afstande divideres med γ for at få en fordeling af afstande, d_i , defineret på intervallet $(0;1]$. Vi vil nu betragte vores afstande $d_1, \dots, d_m \in (0,1)$ som en tilfældig stikprøve fra en sandsynlighedsfordeling. Hvis vi antager, at disse afstande er uafhængigt fordelte vil, for en given værdi af α , sandsynligheden for at finde en stikprøve $(d_1, d_1 + \Delta d_1), \dots, (d_m, d_m + \Delta d_m)$ være lig $\prod_{i=1}^m d_i^\alpha$.

Derfor kan vi estimere α gennem anvendelse af maximum likelihood:

$$\hat{\alpha} = \frac{-m}{\sum \ln(d_i)},$$

hvor $\hat{\alpha}$ er den estimerede korrelationskoefficient.

Litteratur

- Baumol, W. og Benhabib, J. 1989. Chaos: significance, Mechanism, and Economic Applications. *Journal of Economic Perspectives*, 3, 77-105.
- Campbell, J.Y. 1987. Stock Returns and the Term Structure. *Journal of Financial Economics*, 18, 373-99.
- Diebold, F.X. og Nason, J.A. 1990. Nonparametric Exchange Rate Prediction? *Journal of International Economics*, 28, 315-32.
- Engle, R.F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- Fama, E.F. og French, K.R. 1989. Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics*, 25, 23-49.
- Granger, C.W.J. 1991. Developments in the Nonlinear Analysis of Economic Series. *Scandinavian Journal of Economics*, 93 (2), 263-76.
- Hsieh, D.A. 1991. Chaos and non-linear Dynamics: Applications to Financial Markets. *Journal of Finance*, 46, 1839-77.
- Linden, N., Satchell, S.E. og Yoon, Y. 1991.

- Predicting British Financial Indices. An Approach Based on Chaotic Theory. Mimeo, University of Cambridge.
- Lucas, R.E. 1978. Asset Prices in an Exchange Economy. *Econometrica*, 46, 1429-45.
- Scheinkman, J.A. og LeBaron, B. 1989. Non-linear Dynamics and Stock Returns. *Journal of Business*, 62, 311-37.
- Sugihara, G. og May, R.M. 1990. Non-linear Forecasting as a Way of Distinguishing Chaos from Measurement Error in Time Series. *Nature*, 344, 734-741.
- Takens, F. 1981. Detecting Strange attractors in Turbulence. I D. Rand og L. Young (red.) *Dynamical Systems and Turbulence*. Springer-Verlag, Berlin.
- Takens, F. 1984. On the Numerical Determination of the Dimension of an Attractor. I *Dynamical Systems and Bifurcations*, vol. 1125 af Lecture Notes in Mathematics, Springer Verlag.
- Timmermann, A.G. 1993a. Learning, Specification Search and Market Efficiency. With an Application to the Danish Stock Market. Kommende i *Scandinavian Journal of Economics*, 95(2), 157-73.
- Timmermann, A.G. 1993b. How Learning in Financial Markets Generates Excess Volatility and Predictability of Stock Returns. Kommende i *Quarterly Journal of Economics*.

er kaos i serierne. Sandsynligvis kan ikke-lineariteterne bedre modelleres og forstås som stokastiske ikke-lineariteter, hvilket det vil være en udfordring for fremtidige studier at undersøge. I vores analyse anvendte vi disse indikationer til at konstruere ikke-parametriske forudsigelser af daglige afkast. Analysen viste, at de forudsagte værdier var signifikant korrelerede med de faktiske værdier af daglige afkast. Endvidere demonstrerede vi, at disse forudsigelser kunne anvendes i en investeringsstrategi, der gav et større afkast end det gennemsnitlige afkast på Københavns Fondsbørsindeks, når transaktionsomkostningerne er meget lave. Når realistiske transaktionsomkostninger inkluderes, er profitten fra investeringsstrategien baseret på de ikke-parametriske forudsigelser ikke længere økonomisk signifikant.

Teknisk Appendiks

Dette appendiks beskriver hvorledes korrelationsintegralet kan beregnes ved hjælp af Takens (1984) maximum likelihood metode. For en bestemt værdi af "cutoff" punktet γ beskrevet i hovedteksten (f.x. 1-2 standard-afvigelser af data) antager vi at den følgende relation holder eksakt: $C(\gamma) = \gamma^\alpha$, således at korrelationsintegralet følger en "power law" for alle afstande mindre end eller lig γ . Alle afstande større end γ ignoreres og de tilbageværende afstande divideres med γ for at få en fordeling af afstande, d_i , defineret på intervallet $(0;1]$. Vi vil nu betragte vores afstande $d_1, \dots, d_m \in (0,1)$ som en tilfældig stikprøve fra en sandsynlighedsfordeling. Hvis vi antager, at disse afstande er uafhængigt fordelte vil, for en given værdi af α , sandsynligheden for at finde en stikprøve $(d_1, d_1 + \Delta d_1), \dots, (d_m, d_m + \Delta d_m)$ være lig $\prod_{i=1}^m d_i^\alpha$.

Derfor kan vi estimere α gennem anvendelse af maximum likelihood:

$$\hat{\alpha} = \frac{-m}{\sum \ln(d_i)},$$

hvor $\hat{\alpha}$ er den estimerede korrelationskoefficient.

Litteratur

- Baumol, W. og Benhabib, J. 1989. Chaos: significance, Mechanism, and Economic Applications. *Journal of Economic Perspectives*, 3, 77-105.
- Campbell, J.Y. 1987. Stock Returns and the Term Structure. *Journal of Financial Economics*, 18, 373-99.
- Diebold, F.X. og Nason, J.A. 1990. Nonparametric Exchange Rate Prediction? *Journal of International Economics*, 28, 315-32.
- Engle, R.F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- Fama, E.F. og French, K.R. 1989. Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics*, 25, 23-49.
- Granger, C.W.J. 1991. Developments in the Nonlinear Analysis of Economic Series. *Scandinavian Journal of Economics*, 93 (2), 263-76.
- Hsieh, D.A. 1991. Chaos and non-linear Dynamics: Applications to Financial Markets. *Journal of Finance*, 46, 1839-77.
- Linden, N., Satchell, S.E. og Yoon, Y. 1991.