

**SAMFUNNSØKONOMISKE STUDIER**

**42**



**ANALYSE AV SAMMENHENGEN  
MELLOM  
FORBRUK, INNTEKT OG FORMUE  
I NORSKE HUSHOLDNINGER**

**ANALYSING THE RELATIONSHIP BETWEEN  
CONSUMPTION, INCOME AND WEALTH  
IN NORWEGIAN HOUSEHOLDS**

Av/By  
**ERIK BIØRN**

**STATISTISK SENTRALBYRA  
CENTRAL BUREAU OF STATISTICS OF NORWAY  
OSLO 1979**



**SAMFUNNSØKONOMISKE STUDIER NR. 42**

**ANALYSE AV SAMMENHENG  
MELLOM  
FORBRUK, INNTEKT OG FORMUE  
I NORSKE HUSHOLDNINGER**

**ANALYSING THE RELATIONSHIP BETWEEN  
CONSUMPTION, INCOME AND WEALTH  
IN NORWEGIAN HOUSEHOLDS**

Av/By  
**ERIK BIØRN**

**STATISTISK SENTRALBYRÅ  
CENTRAL BUREAU OF STATISTICS OF NORWAY  
OSLO 1979**

ISBN 82-537-1012-7



**ANALYSE AV SAMMENHENGEN MELLOM  
FORBRUK, INNTEKT OG FORMUE  
I NORSKE HUSHOLDNINGER**



## FORORD

Statistisk Sentralbyrås 3-årige inntekts- og formuesundersøkelser - som bygger på individuelle skatteligningsoppgaver - og de forbruksundersøkelser - basert på regnskapsføring og intervju - som blir foretatt samtidig, dekker for en stor del de samme husholdninger. En sammenkobling av undersøkelsene for året 1973 gir således et komplett materiale av inntekts-, formues- og forbrukstall for over 3 000 norske husholdninger.

I denne studien presenteres hovedresultatene av en økonometrisk analyse av dette materialet. Et viktig formål med undersøkelsen har vært å forsøke å bringe klarhet i en del spørsmål som det erfaringsmessig har vist seg vanskelig å få tilfredsstillende svar på ved bare å studere aggregerte tidsserier (f.eks. nasjonalregnskapsdata). Dette er samtidig spørsmål som står sentralt i arbeidet med å spesifisere forbruksstrukturen i Byråets skattemodeller og makroøkonomiske analysemodeller.

Studien har også interesse fra et mer generelt metodesynspunkt. Det anvendes dels regresjonsanalyse kombinert med et opplegg for multiple sammenligninger (multipel testing), dels simultanestimering innenfor modeller med uobserverbare (latente) variable.

Statistisk Sentralbyrå, Oslo, 31. august 1979

Odd Aukrust

## *PREFACE*

The income and property statistics, prepared at 3 years intervals by the Central Bureau of Statistics on the basis of a sample of individual reports to the tax authorities, and the surveys of consumer expenditure, which are prepared simultaneously by using mixed interview and accounting, cover largely the same households. By combining the two data files from the year 1973, we get complete reports on income, wealth and consumption expenditures from more than 3 000 Norwegian households.

This study presents the main results of an econometric analysis based on this combined data file. The main attention is focused on questions to which aggregate time series have not yet given satisfactory answers. These questions are also essential in formalising household consumption behaviour in taxation models and models for macro-economic analysis and forecasting.

Moreover, the study is interesting from a methodological point of view. One part is based on multiple testing within a regression framework, another utilises estimation of simultaneous models with unobservable (latent) variables.

Central Bureau of Statistics, Oslo, 31 August 1979

Odd Aukrust



## INNHOLD

	Side
1. Innledning. Bakgrunn .....	9
2. Datagrunnlaget .....	12
2.1. Observasjonsenheten .....	12
2.2. Avgrensningen av de sentrale variable .....	13
2.3. Summarisk oversikt over datamaterialet. Deskriptive korrelasjons- og regresjonsberegninger .....	19
Appendiks: Om betydningen av at forbruksutgiftsperioden er forskjøvet i forhold til inntektsregistreringsperioden .....	26
3. Konsumutgiftens avhengighet av inntekt, formue, alder og husholdningstype. Regresjonsanalyse og testing av multiple hypoteser .....	29
3.1. Innledende bemerkninger .....	29
3.2. Nærmere om analyseopplegget .....	30
3.3. Konsumets avhengighet av inntektens og formuens artssammensetning: Analysekjema I .....	33
3.4. Konsumtilbøyelighetens avhengighet av inntektens artssammensetning og hovedinntektstagerens sosialgruppe: Analysekjema II ..	39
3.5. Avvikelser fra linearitet. Samspill mellom inntekt, formue, alder og husholdningsstørrelse: Analysekjema III .....	44
3.6. Konsumfunksjonens homogenitetsegenskaper analysert ved en modifisert "translog" - funksjon: Analysekjema IV .....	51
3.7. Oppsummering .....	55
4. Estimering av konsumfunksjoner i modeller med uobserverbare (latente) variable .....	57
4.1. Utgangspunktet .....	57
4.2. En variant av Friedman's konsumteori. Estimering ved hjelp av instrumentvariable .....	59
4.3. En generell lineær strukturmodell med latente variable .....	60
4.4. Estimeringsresultater for tre enkle modeller .....	64
Appendiks: Identifiserbarheten av modell A .....	82
5. Oversikt og hovedkonklusjoner .....	84
Litteraturhenvisninger .....	87
Sammendrag på engelsk .....	90
Utkommet i serien SØS .....	93

*CONTENTS*

	Page
1. Introduction. Background .....	9
2. The data base .....	12
2.1. The unit of observation .....	12
2.2. Definition of the main variables .....	13
2.3. Survey of the data. Descriptive correlation coefficients and regressions .....	19
Appendix: Some implications of the fact that the periods of reporting of consumption expenditure and income do not coincide .....	26
3. The dependence of consumption expenditure on income, wealth, age and type of household. Regression analysis and multiple testing	29
3.1. Introductory remarks .....	29
3.2. Outline of procedure .....	30
3.3. The dependence of consumption on the composition of income of wealth: Scheme I .....	33
3.4. The dependence of the marginal propensity to consume on type of income and social status of the head of household: Scheme II .....	39
3.5. Deviations from linearity. Interaction between income, wealth, age and number of household members: Scheme III ..	44
3.6. The homogeneity of the consumption function analysed by means of a modified "translog" function: Scheme IV .....	51
3.7. A summing-up .....	55
4. Estimation of consumption functions in models with unobservable (latent) variables .....	57
4.1. Point of departure .....	57
4.2. A variant of Friedman's consumption theory. Estimation by means of instrumental variables .....	59
4.3. A general linear structural model with latent variables ...	60
4.4. Estimation results for three simple models .....	64
Appendix: The identification of Model A .....	82
5. Summary of results .....	84
References .....	87
Summary in English .....	90
Issued in the series Social Economic Studies (SES) .....	93

## FIGURREGISTER

	Side
1. Hypotesestrukturen i analyseskjema I .....	36
2. Hypotesestrukturen i analyseskjema II .....	41
3. Hypotesestrukturen i analyseskjema III .....	46
4. Hypotesestrukturen i analyseskjema IV .....	53
5. Kausalitetsstrukturen i en generell modell med latente variable .	62
6. Kausalitetsstrukturen i modell A .....	66
7. Kausalitetsstrukturen i modell B .....	71
8. Kausalitetsstrukturen i modell C .....	77

## TABELLREGISTER

1. Gjennomsnitt og empirisk standardavvik for analysens hovedvariable	20
2. Inntektssammensetning for de enkelte sosialgrupper .....	23
3A. Korrelasjonsmatrise for inntektsvariable .....	24
3B. Korrelasjonsmatrise for formuesvariable .....	24
3C. Korrelasjonsmatrise for inntektsvariable med hensyn på formues- variable .....	24
4. Analyseskjema I. Testresultater .....	35
5. Analyseskjema I. F-observatorer for deltestene .....	37
6. Analyseskjema I. Koeffisientestimater svarende til testresultatet	38
7. Hypoteser i analyseskjema II .....	40
8. Analyseskjema II. Testresultater .....	42
9. Analyseskjema II. F-observatorer for deltestene .....	42
10. Analyseskjema II. Koeffisientestimater svarende til testresul- tatet .....	43
11. Hypoteser i analyseskjema III .....	45
12. Analyseskjema III. Testresultater .....	47
13. Analyseskjema III. F-observatorer for deltestene .....	48
14. Analyseskjema III. Koeffisientestimater svarende til test- resultatet .....	50
15. Analyseskjema IV. F-observatorer for deltestene .....	54
16. Analyseskjema IV. Koeffisientestimater svarende til test- resultatet .....	55
17. Estimater for konsumelastisiteter ved proporsjonal endring av inntekt og antall husholdningsmedlemmer .....	55
18. Maximum Likelihood (FIML)-estimater for modell A .....	68
19. Maximum Likelihood (FIML)-estimater for modell B .....	72
20. Maximum Likelihood (FIML)-estimater for modell C .....	78

## INDEX OF FIGURES

	Page
1. The structure of hypotheses in Scheme I .....	36
2. The structure of hypotheses in Scheme II .....	41
3. The structure of hypotheses in Scheme III .....	46
4. The structure of hypotheses in Scheme IV .....	53
5. The structure of causality in a general model with latent variables .....	62
6. The structure of causality in Model A .....	66
7. The structure of causality in Model B .....	71
8. The structure of causality in Model C .....	77

## INDEX OF TABLES

1. Mean and standard deviation of the main variables .....	20
2. Composition of income for the different social groups .....	23
3A. Correlation matrix for income variables .....	24
3B. Correlation matrix for wealth variables .....	24
3C. Correlation matrix for income variables with respect to wealth variables .....	24
4. Scheme I. Testing results .....	35
5. Scheme I. F statistics for the sub-tests .....	37
6. Scheme I. Coefficient estimates corresponding to the testing result .....	38
7. Hypotheses in Scheme II .....	40
8. Scheme II. Testing results .....	42
9. Scheme II. F statistics for the sub-tests .....	42
10. Scheme II. Coefficient estimates corresponding to the testing result .....	43
11. Hypotheses in Scheme III .....	45
12. Scheme III. Testing results .....	47
13. Scheme III. F statistics for the sub-tests .....	48
14. Scheme III. Coefficient estimates corresponding to the testing result .....	50
15. Scheme IV. F statistics for the sub-tests .....	54
16. Scheme IV. Coefficient estimates corresponding to the testing result .....	55
17. Estimates of consumption elasticities when income and the number of household members change proportionally .....	55
18. Maximum Likelihood (FIML) estimates of Model A .....	68
19. Maximum Likelihood (FIML) estimates of Model B .....	72
20. Maximum Likelihood (FIML) estimates of Model C .....	78

## 1. INNLEDNING. BAKGRUNN\*

Økonomer har lenge vært opptatt av hvordan inntekten, sammen med andre variable, påvirker husholdningenes totale etterspørsel etter konsumgoder. Gjennom årene har det teoretiske innhold i denne sammenhengen vært viet betydelig oppmerksomhet, og matematiske presiseringer i form av såkalte konsumfunksjoner er i dag et uunnværlig verktøy for økonomisk analyse.

På tross av at konsumfunksjoner har vært gjenstand for omfattende empirisk forskning i en rekke land, står fortsatt flere vesentlige spørsmål ubesvart. Riktignok er det noe nær samstemmighet om at inntekten er en viktig forklaringsvariabel for konsumutgiften - mange mener den viktigste - men hvilke andre variable som har betydning, og valget av matematisk funksjonsform har vært stadig tilbakevendende diskusjonstemaer i litteraturen. Det er betegnende at den kjente amerikanske økonom James Tobin i en nylig utgitt artikkelsamling bruker over 250 sider på å gi en - ganske visst utførlig, men ikke utpreget detaljert - oversikt over konsumfunksjonens utvikling og status pr. i dag.<sup>1)</sup>

Ved siden av inntektens størrelse er det følgende faktorer som oftest bringes inn for å forklare størrelsen av konsumet i en bestemt periode:

- (i) sammensetningen av inntekten etter art,
- (ii) formuens og gjeldens størrelse og sammensetning,
- (iii) likviditets- og kredittforholdene,
- (iv) rentenivået,
- (v) prisstigningstakten,
- (vi) konsumentenes alder, utdanning og andre demografiske kjennetegn,
- (vii) konsumentenes forventninger om fremtidig inntektsutvikling, rentenivå og prisstigning og deres oppfatning av usikkerheten i disse forventninger, og endelig
- (viii) psykologiske og sosiologiske faktorer, bl.a. konsumplaner og holdninger til det å konsumere.

Tidsstrukturen i konsumfunksjonen, dvs. hvorvidt de enkelte variable virker momentant eller med tidsforskyvninger ("lag"), er også ofret betydelig oppmerksomhet.

I økonometriske studier av konsumfunksjonen har både aggregerte tidsserier (blant annet nasjonalregnskapsdata) og tverrsnittsdata fra enkelthusholdninger vært benyttet. Men det er antagelig riktig å si at den første datatype har vært den langt dominerende. Tre forhold har nok medvirket til dette. For det første er det *makro*-konsumfunksjoner - dvs. funksjoner som bestemmer

---

\* Forfatteren vil gjerne takke Arne Amundsen, Olav Bjerkholt, Svein Longva, Håvard Røyne og Odd Skarstad for nyttige synspunkter og kommentarer til deler av et tidligere manuskriptutkast og Inger Holm for verdifull bistand med å tilrettelegge datamaterialet for analysen.

1) Tobin [44], pp. 61-318.

et lands totalkonsum ved totalverdier eller gjennomsnittsverdier for de enkelte forklaringsvariable - som modellanalytikernes hovedinteresse har vært knyttet til i de siste 30-40 år. Dette skyldes naturligvis den sentrale rolle slike aggregerte funksjoner spiller i modellresonnementer for totaløkonomien. For det annet har aggregerte tidsserier i praksis vært lettere tilgjengelig for analytiske formål enn tilsvarende individuelle tverrsnittsdata. Den tredje, og ikke minst viktige, årsak er at sammenhengende tidsserier er en nødvendighet for å kunne studere eventuelle tidsforskyvninger mellom konsumet og dets forklaringskomponenter.<sup>2)</sup>

I Statistisk Sentralbyrås modellarbeid har økonometriske studier av konsumfunksjoner spesiell interesse av minst to grunner. Den ene er at analysemodellen MODIS IV inneholder en makro-konsumfunksjon. Denne er et hovedelement i modellen, og det er karakteristisk at modellens viktigste bruker, Finansdepartementet, har omfattet denne relasjon med særlig stor interesse og årvåkenhet i den tid modellen har vært i virksomhet. Hittil har nasjonalregnskapsdata vært den dominerende datakilde for tallfesting av makro-konsumfunksjonen i MODIS.<sup>3)</sup> Men det har også vært gjort enkelte sporadiske forsøk på å utnytte inntekts- og forbruksdata for enkelthusholdninger. Generelt tyder erfaring på at mikrodata vil være et nyttig supplement til aggregerte tidsserier ved estimering av økonometriske strukturlikninger. Estimering på mikronivå kan ha verdi selv om det bare er i aggregert form en ønsker å bruke resultatet. I en kommentar til en engelsk undersøkelse av husholdningssparing i 1953 sier Trygve Haavelmo bl.a. følgende:<sup>4)</sup>

"There are at least three important reasons why cross-section data on savings are essential as alternatives to, or as supplements to, time series data. One reason is that the average disposable income of consumers seldom shows sufficiently large short-run variations to permit a reasonably accurate measurement of the propensity to consume. A second reason is that cross-section data can tell us a good deal about how to develop approximate aggregation formulae. A third reason is that cross-section data can give us information on the effects upon savings of certain variables (age, family size, etc.) that do not vary at all, or vary only slightly, in aggregate time series for the community as a whole. It is of course true that the use of survey data requires some rather bold hypotheses concerning individual 'likeness' under 'similar conditions'. But that is also to some extent true for a study based on time series, in the sense that then one usually has to build on some hypotheses of inter-temporal (instead of inter-personal) comparability."

Det annet, og minst like viktige, argument for å arbeide med konsumfunksjoner som ledd i modellprosjekter i Byrået er hensynet til skatteinsidensmodellene. Dette ble aktualisert da det for noen år siden ble satt i gang

2) Muligheten for å skaffe tidsserier av individualdata (kombinerte tidsserie-/tverrsnittsdata) har i praksis vært meget begrenset.

3) Se Biørn [4]. Noen av beregningene er senere revidert og ajourført; se Cappelen [9].

4) Haavelmo [21], p. 145.

arbeid med å utvikle en empirisk modell for analyse av fordelingsvirkninger av samtidige endringer i direkte og indirekte skatter og subsidier. Et kjernepunkt i modellen viste seg å bli beskrivelsen av sammenhengen mellom husholdningenes disponible inntekt og utgiften til konsumgoder. I mangel av pålitelige empiriske resultater ble det valgt - som en midlertidig løsning - å forutsette at alle husholdninger forbruker hele sin disponible inntekt, dvs. at spareraten er null.<sup>5)</sup> Det ble dermed ikke tatt vare på noen av faktorene nevnt under (i) - (viii) ovenfor. Som generell beskrivelse er dette selvsagt utilfredsstillende - ikke minst fordi det må være et hovedpoeng i skattefordelingsanalyser å ta hensyn til at husholdninger med forskjellige karakteristika vanligvis har forskjellig konsum- og spareatferd.

Formålet med denne studien er å undersøke - på en mer systematisk måte enn tidligere - hvilke holdepunkter norske husholdningsdata gir for å bestemme konsumfunksjoner empirisk. Datamaterialet er dannet ved å kombinere oppgaver fra Forbruksundersøkelsen for 1973 med opplysninger fra Inntekts- og formuesundersøkelsen for samme år. Det gir detaljert informasjon om konsumets, inntektens og formuens størrelse og sammensetning i et representativt utvalg av husholdninger og muliggjør en rekke interessante økonomiske beregninger. Vi vil konsentrere oppmerksomheten om et utvalg av de hypoteser som synes å være særlig interessante fra et modellutviklingssynspunkt. Det faktum at observasjonene bare dekker ett enkelt år har selvsagt fått innflytelse på valget av hypoteser; eventuelle tidsforskyvninger i konsumentenes reaksjoner og andre typer av dynamikk har vi bare i meget beskjeden grad kunnet ta hensyn til. Blant de 8 typer av konsumforklarende variable vi nevnte ovenfor, er det (i) og (ii) vi særlig vil feste oppmerksomheten ved. Men deler av analysen vil også berøre (iii) og (vi).<sup>6)</sup>

Fremstillingen er disponert på følgende måte: Kapittel 2 gir en oversikt over datagrnnlaget. Avgrensningen av de sentrale variable, inntekt, formue og konsumutgift og deres komponenter, samt svakheter ved de indikatorer som er valgt, diskuteres spesielt (avsnitt 2.2). Hovedtrekk ved datamaterialet karakteriseres ved enkle korrelasjons- og regresjonsberegninger (avsnitt 2.3). I kapittel 3 presenteres resultatene av en forholdsvis omfattende regresjonsanalyse. Her undersøkes konsumutgiftens avhengighet av inntektens og formuens sammensetning og hvorvidt de marginale konsumtilbøyeligheter varierer med husholdningens sosialgruppe. Vi forsøker dessuten å kartlegge i hvilken utstrekning konsumfunksjonene er lineære eller om effekter av høyere orden må trekkes inn. I dette kapittel støtter vi oss for en vesentlig del til et opplegg for multiple sammenligninger (multiple tester). Erfaringene fra denne del av analysen har

5) Modellen som den i dag foreligger, er beskrevet i Biørn og Garaas [7].

6) En videreføring av analysen som bringer (iv) og (vii) inn som sentrale forklaringsfaktorer ved siden av (i), (ii) og (vi), er presentert i Biørn [6].

også interesse fra et generelt metodesynspunkt. I kapittel 4 utvider vi perspektivet ved at vi bringer inn muligheten for at noen av de variable kan være observert med målefeil. Vi starter med en enkel utgave av Milton Friedman's konsummodell (avsnitt 4.2) og utvider den i forskjellige retninger til tre modeller som alle har form av lineære simultane ligningssystemer med latente (uobserverbare) variable. Modellene er simultane fordi de ved siden av konsumfunksjoner også inneholder ligninger til bestemmelse av inntekten. I avsnitt 4.3 behandles kort spesifisering og estimering av lineære strukturmodeller med latente variable på generelt grunnlag før vi i avsnitt 4.4 presenterer estimeringsresultater for de tre modellene basert på Maximum Likelihood-metoden (sannsynlighetsmaksimeringsmetoden). Hovedresultatene oppsummeres i kapittel 5.

## 2. DATAGRUNNLAGET

### 2.1. Observasjonsenheten

Datamaterialet består av oppgaver over forbruksutgifter, inntekter og formuesverdier fra 3 271 private husholdninger. Det er dannet ved å kombinere opplysninger om forbruksutgifter fra Statistisk Sentralbyrås forbruksundersøkelse for året 1973 med inntekts- og formuesoppgaver fra Byråets inntekts- og formuesundersøkelse for samme år. Begge undersøkelser er utvalgsundersøkelser. *Husholdningen*<sup>1)</sup> er observasjonsenhet i forbruksundersøkelsen, mens observasjonsenheten i inntekts- og formuesundersøkelsen er den enkelte *skattyter*.<sup>2)</sup> Utvalget av skattytere i inntekts- og formuesundersøkelsen inneholder - med visse unntagelser - samtlige skattytere i de husholdninger som deltok i forbruksundersøkelsen. Det er dermed lett å beregne inntekts- og formuestall for den enkelte husholdning ved å summere tallene for de skattytere som inngår i husholdningen. Grunnlaget for sammenkoblingen av de to materialene er skattyternes personnumre.

Observasjonsenheten i datamaterialet blir altså husholdningen slik den er definert i forbruksundersøkelsen.<sup>3)</sup>

---

1) Husholdningsbegrepet er den såkalte "kosthusholdning", dvs. et kollektiv av personer med et visst forbruksfellesskap. En mer presis definisjon er gitt i [39], som også inneholder en redegjørelse for opplegget av forbruksundersøkelsen for 1973.

2) Opplegget av inntekts- og formuesundersøkelsene er beskrevet i [41] og [42].

3) Merk at dette husholdningsbegrep avviker fra det som er brukt i tabellpublikasjonene fra inntekts- og formuesundersøkelsene. Der er det forsørgelsesbyrde, ikke forbruksfellesskap, som er kriteriet for aggregering av individene.



## 2.2. Avgrensningen av de sentrale variable

### 2.2.1. Generelt

Det eksisterer neppe alment aksepterte definisjoner av "inntekt", "formue" og "forbruk" - de tre hovedbegreper i vår undersøkelse. Men man ønsker som regel at sammenhengen

$$(2.1) \quad \text{Inntekt etter skatt} = \text{Verdi av forbruk} + \text{Verdi av formuesøkning (sparing)}$$

skal gjelde. Dette betyr i prinsippet at bare to av de tre størrelser kan defineres fritt. En naturlig fremgangsmåte når en skal avgrense inntekt, formue og forbruk som *teoretiske* begreper, kan dermed være først å bestemme hvilke objekter som skal regnes som forbruk og som formue, dernest finne et prinsipp å verdsette disse objekter etter og til slutt bestemme inntektsbegrepet ved hjelp av (2.1). I den grad en aksepterer det forbruks- og formuesbegrep og de vurderingsprinsipper som er valgt, gir denne fremgangsmåte et "teoretisk riktig" inntektsbegrep.<sup>4)</sup>

De indikatorer for inntekt, formue og forbruk som vi i praksis kan observere, vil imidlertid vanligvis ikke oppfylle (2.1). Dette skyldes dels at enkelte poster i den teoretiske "idealinntekt" ikke er statistisk registrerbare, dels at målefeilene i de formues- og forbruksoppgaver som foreligger, er så store at sammenhengen bare i beste fall vil gjelde som en grov tilnærming. Mer konkret: Hvis vi forsøker å beregne en husholdnings samlede (real- og finans-)sparing (formuesøkning) i et år ved å kombinere inntektsoppgaver og skatteligningsopplysninger med forbruksutgiftstall fra en forbruksundersøkelse, vil vi bare unntagelsesvis få samme resultat som ved å ta differensen mellom den skattepliktige nettoformue ved årets slutt og årets begynnelse. Ingen av beregningsmåtene gir korrekt uttrykk for husholdningens faktiske sparing; det finnes to ufullkomne måter å måle samme størrelse på.

Dette at tallmaterialet ikke "stemmer økosirkisk" begrenser selvsagt spillerommet for valg av analyseopplegg og hypoteser. Men materialet inneholder likevel så verdifull informasjon at det fortjener en nærmere økonomisk behandling. Vi vil nedenfor kort redegjøre for innholdet i de inntekts-, formues- og forbruksutgiftsvariable som er benyttet. Spesielt vil vi forklare hvordan primærmaterialet er forsøkt utnyttet til å danne en analytisk interessant inndeling av inntekten og formuen. Nyten av det kombinerte datamaterialet ligger blant annet nettopp i at det gjør det mulig å studere hvordan konsumet og dets komponenter varierer med inntektens og formuens sammensetning.

4) I Biørn [5], avsnitt 2, er denne tankegangen forsøkt utdypet.

### 2.2.2. Inntektsvariablene

Selvangivelsesskjemaene gir detaljert informasjon om postene i bruttoinntekten og fradragspostene ved inntektslikningen. I inntektsundersøkelsen er disse opplysninger supplert med oppgaver over barnetrygd, forsørgerstønad og økonomisk bidrag fra sosialhjelpen samt opplysninger om de samlede utlignede inntekts- og formuesskatter og medlemsavgifter til folketrygden. Derimot er en del skattefrie trygdeytelser ikke medregnet. På dette grunnlag er det beregnet et uttrykk for husholdningenes *disponible inntekt* etter at det er foretatt en del *korreksjoner* i det "inntektsregnskap" som fremkommer ved skattelikningen. Hensikten er å oppnå et analytisk sett mer tjenlig inntektsbegrep. Derfor er flere poster som kommer til fradrag ved inntektslikningen, lagt til den skattepliktige nettoinntekt igjen. Dette gjelder i første rekke skatteteknisk og finanspolitisk begrunnede fradrag som minstefradrag, skattefritt fradrag (fribehold) for renter og utbytte av finanskapital, skattefrie banksparing, premie for frivillig livsforsikring, fradrag for tidligere års underskudd i næring samt skattefrie fondsavsetninger. Tilsvarende korreksjoner er gjort for "utgifter til inntekts ervervelse", bl.a. utgifter til arbeidsreiser, merutgifter ved å bo utenfor hjemmet etc. De to siste poster inngår i forbruksutgiften slik den registreres ved forbruksundersøkelsen, og for å få et uttrykk for "inntekt disponibel for forbruk" bør de være med også i inntekten (selv om de fleste nok vil protestere mot at disse utgiftene bidrar nevneverdig til deres økonomiske velferd).

Behandlingen av *gjeldsrentefradraget* - som for mange husholdninger representerer et betydelig beløp - reiser et spesielt problem. Vi har her valgt ikke å gjøre korreksjoner for dette fradraget, men regne alle renteutgifter som fradragspost. Begrunnelsen er at dette er beløp husholdningene må dekke på grunnlag av sine løpende innbetalinger, og de vil være fradragspost i alle (?) rimelige teoretiske definisjoner av inntekt. Forutsetningen er at en eventuell avkastning av formuesobjekter som anskaffes ved lånte midler, regnes med på plussiden i inntektsregnskapet. For de fleste husholdninger, bortsett fra dem hvor hovedpersonen er selvstendig næringsdrivende, vil renter på boliglån utgjøre hoveddelen av gjeldsrentefradraget. Nå er bruksverdien av egen (selveid) bolig registrert som forbruksutgift i forbruksundersøkelsen - riktignok lavt verdsatt. Men dette er i seg selv intet argument for å fravike prinsippet om å regne alle renteutgifter som fradrag i inntekten; regulært vil nemlig skattytere som disponerer egen bolig, bli beskattet for denne fordel ved at en viss andel (i 1973 2½ prosent) av boligens skattetakst inkluderes i den skattepliktige inntekt. Leieverdien av egen bolig er altså *i prinsippet* med i den inntekt som registreres ved inntektsundersøkelsen - selv om det er en utbredt oppfatning at "2½-prosentregelen" gir en urimelig lav verdi sammenlignet med markedsverdien av

boligtjenestene. Etter dette må vi regne med at inntekten og verdien av boligkonsumet i noen grad undervurderes for husholdninger med selveid bolig. Dette gjelder ca. 60 prosent av husholdningene i datamaterialet.

Formålet med oppdelingen av inntekten er å danne indikatorer for de viktigste *inntektsarter*, nemlig arbeidsinntekt, trygdeinntekt, inntekt av finans- og realkapital samt kapitalgevinster. Ut fra det som er sagt ovenfor, er det uten videre klart at skatteligningsmaterialet, selv etter at det er bearbeidet i inntektsundersøkelsen, bare vil kunne gi ufullkomne mål for disse variable. Særlig mangelfullt registrert er nok inntekt av realkapital og kapitalgevinster, hovedsakelig fordi slike inntekter bare i begrenset grad er skattepliktige for personlige skattytere.<sup>5)</sup>

Vi har først delt bruttoinntekten i følgende komponenter (hvor tilføyselsen "brutto" markerer at det ikke er gjort fradrag for skatter):

$R_1^*$  : Lønnsinntekt, tjenestepensjon mv., brutto.

$R_2^*$  : Ytelser fra folketrygden (ekskl. barnetrygd), brutto.

$R_3^*$  : Næringsinntekt, brutto.

$R_4^*$  : Renter og utbytte av bankinnskudd, aksjer, obligasjoner og andre utestående fordringer, brutto.

$S^*$  : Gjeldsrenter, brutto.

$R_5^*$  : Inntekt av bolig og annen fast eiendom, brutto.

$R_6^*$  : Gevinster ved tomte- og aksjesalg, brutto.

Her representerer  $R_1^*$  arbeidsinntekt,  $R_2^*$  trygdeinntekt,  $R_4^*$  inntekt av finanskapital,  $R_5^*$  inntekt av realkapital og  $R_6^*$  kapitalgevinster. Næringsinntekten  $R_3^*$  er fra vårt synspunkt inhomogen, idet den inneholder både arbeidsinntekt - verdien av den næringsdrivendes arbeidsinnsats - og realkapitalinntekt - avkastningen av den produksjonskapital han disponerer i næringsvirksomheten. Med det foreliggende datamateriale er det ikke mulig å splitte denne komponenten. Det ville også være betydelige prinsipielle problemer forbundet med en slik oppdeling.

Totalverdien av bruttoinntekten er lik

$$(2.2) \quad R^* = \sum_{i=1}^6 R_i^* - S^*.$$

5) For en utdypning, se f.eks. Biørn [5], avsnitt 3.

Det bearbejdede inntektsmateriale, etter at det er foretatt korreksjoner som nevnt ovenfor, gir størrelsen

$R$  : Disponibel inntekt.

Den atskiller seg fra bruttoinntekten  $R^*$  først og fremst ved at direkte skatter (inntekts- og formuesskatter) er trukket fra og barnetrygd og forsørgerstønad lagt til. Men differensen inneholder også noen, mindre, til dels uspesifiserte, poster som vanskelig kan henføres til én av de seks inntektskomponenter. Dermed gir  $(R^* - R)/R^*$  tilnærmet uttrykk for gjennomsnittsskattesatsen for direkte skatter, relatert til bruttoinntekten  $R^*$ .

Som annet trinn er nettoverdien av inntektskomponentene bestemt ved å fordele skattebeløpet proporsjonalt på inntektsposter. Denne metode er selvsagt ikke ideell, men med det datamateriale som foreligger, antagelig den beste vi kunne velge.<sup>6)</sup> Nettoinntektene og nettoverdien av gjeldsrentene blir altså

$$(2.3) \quad \begin{cases} R_i = R_i^* \frac{R}{R^*} \\ S = S^* \frac{R}{R^*} \end{cases} \quad (i=1, \dots, 6),$$

De oppfyller betingelsen

$$(2.4) \quad \sum_{i=1}^6 R_i - S = R.$$

### 2.2.3. Formuesvariablene

For formuestallene er det ikke gjort andre korreksjoner i forhold til skatteligningens registreringer enn at det ikke er tatt hensyn til det skattefrie fradrag (fribeløpet) for finansiell formue. Bruttoformuen er splittet i seks grupper - tre av disse inneholder finansobjekter og tre realobjekter.

6) Et alternativ kunne være å "belaste" bare kapitalinntekt ( $R_4^* + R_5^* - S^*$ ) med formuesskatt og anvende forskjellige skattesatser for forskjellige inntektskomponenter, f.eks. bruke marginalsattesatsen på den del av inntekten som ut fra et eller annet kriterium er "marginalinntekt" for den enkelte husholdning (f.eks. renteinntekt for en lønnstagerhusholdning og lønns- og trygdeinntekt for en husholdning hvor hovedpersonen er selvstendig næringsdrivende). Slike opplegg ville imidlertid lett bli nokså spekulative og ville dessuten gjøre det nødvendig å trekke alle progresjonstabeller og en rekke detaljer i skattereglene inn i analysen.

Formuesvariablene er

$F_1$ : Bank- og postgiroinnskudd.

$G$ : Gjeld.

$F_2$ : Ihendehaverobligasjoner, aksjer, utestående fordringer mv.

$F_3$ : Faste eiendommer, produksjonsmidler etc.

$F_4$ : Innbo og løsøre.

$F_5$ : Privatbiler og lystfartøyer.

$F_6$ : Leieboerinnskudd, forsikringspoliser etc.

Disse seks gruppene er forholdsvis homogene både når det gjelder likviditetsgrad og prinsippene for den skattemessige taksering. Av den totale nettoformue

$$(2.5) \quad F = \sum_{i=1}^6 F_i - G$$

utgjør *realformuen*  $F_3 + F_4 + F_5$ , mens *finansformuen*, regnet netto, er  $F_1 + F_2 + F_6 - G$ .

Alle formuesregistreringer refererer seg til *utgangen* av året 1973; sparing i løpet av året er følgelig *i prinsippet* med i formuestallene. Informasjon om begynnelsesformuen hadde utvilsomt også vært ønskelig, men ut fra de opplysninger som forelå, var det ikke mulig å beregne verdier som er definisjonsmessig konsistente med anslagene for sluttformuen.<sup>7)</sup>

#### 2.2.4. Forbruksutgiftsvariablene

Oppgavene over forbruksutgifter bygger direkte på forbruksundersøkelsens registreringer. I økonometriske analyser av sammenhengen mellom konsum, inntekt og formue på husholdningsnivå har oppmerksomheten stort sett vært begrenset til den totale konsumutgift.<sup>8)</sup> Dette skyldes nok i de fleste tilfelle at informasjon om utgiftens sammensetning ikke foreligger - f.eks. fordi oppgavene bygger på et summarisk intervju. Vårt datamateriale gir slike opplysninger, og det ville i prinsippet være mulig å gå ned på et temmelig detaljert spesifikasjonsnivå. For å begrense omfanget av beregningene har vi imidlertid valgt en aggregert inndeling med fem grupper - tre varegrupper og to tjenestegrupper. *Vare*gruppene er inndelt etter

7) Vi kommer tilbake til dette problemet på slutten av avsnitt 3.3 og i avsnitt 4.4.2.

8) Se f.eks. de oversikter som er gitt i Ferber [15] og Mayer [29].

kriteriet "varighet" - de tre gruppene betegnes som henholdsvis "varige", "halv-varige" og "ikke-varige" goder. De varige goder omfatter utgifter til møbler, elektriske husholdningsapparater, egne transportmidler (hovedsakelig kjøp av biler) samt TV- og radioapparater. Halv-varige goder er bl.a. klær og skotøy, sportsutstyr, en del drifts- og vedlikeholdsutgifter for egne transportmidler, bøker etc., mens matvarer, drikkevarer, elektrisitet og brensel, bensin og olje samt toalettartikler, kosmetiske preparater mv. betraktes som ikke-varige goder. *Tjenestegruppene* er dannet ved at utgifter til bruk og vedlikehold av boliger er skilt ut fra de øvrige tjenesteutgifter.

Denne inndeling er interessant for det første fordi inntekt, formue og alder står i en annen stilling som kjøpsmotiverende variable for varige og halv-varige goder enn for ikke-varige goder og tjenester, og for det annet fordi boligutgiften, som til dels er en beregnet størrelse, har en annen karakter enn tjenesteutgiftene for øvrig.<sup>9)</sup> Vi finner da også at inndelinger av denne type er i bruk i flere makro-økonometriske analysemodeller, men har foreløpig vært lite anvendt i modellbygging i Norge.

En kvalitetsvurdering av utgiftstallene må ta utgangspunkt i at registreringene som hovedregel bygger på regnskapsføring over en *14 dagersperiode*. Det gjelder de ikke-varige goder, de fleste tjenester og en god del av de halv-varige goder. Anslagene er "oppblåst" til årsbasistall ved multiplikasjon med 26. For ikke-varige goder og tjenester vil kjøpsutgiften i prinsippet være en god indikator for konsumet, men når registreringsperioden er så kort som den er i vårt tilfelle, kan lagervariasjoner representere en ikke ubetydelig feilkilde. Det er velkjent at mange husholdninger kjøper f.eks. matvarer for lagring, slik at forbruket vil vise et annet tidsforløp enn innkjøpet.

Utgiftene til goder som erfaringsmessig anskaffes forholdsvis sjelden - i første rekke varige goder, men også en del halv-varige goder og tjenester - er rapportert ved intervju. Registreringene gjelder da 12-måneders-perioden umiddelbart forut for regnskapsperioden. Varige (og til en viss grad halv-varige) goder er kjennetegnet ved at kjøpsutgiften begrepsmessig atskiller seg fra verdien av konsumet. Fordi slike goder yter konsumtjenester over flere år, vil utgiftstall for et kalenderår - uansett hvor nøyaktig de er registrert - ikke kunne gi fullgodt uttrykk for faktisk konsum. Differensen vil være avhengig av godenes "levetid".

---

9) Et tilleggsmoment er at utgiftsregistreringene for varige og halv-varige goder sannsynligvis er beheftet med vesentlig større målefeil enn registreringene for ikke-varige goder og tjenester.

Forbrukstallene må følgelig antas å inneholde ikke ubetydelige målefeil. Sannsynligvis er feilene størst for varige goder med lang levetid, goder med utpreget sesongmønster i forbruket (dersom registreringsperioden er 14 dager) og ikke-varige goder hvor innkjøp for lagring hyppig forekommer.

I det følgende benyttes symbolene:

$C_{IV}$  : Utgift til ikke-varige goder.

$C_{HV}$  : Utgift til halv-varige goder.

$C_V$  : Utgift til varige goder.

$C_T$  : Utgift til tjenester eksklusive boligjenester.

$C_{BT}$  : Utgift til boligjenester.

Den totale konsumutgift er altså lik

$$(2.6) \quad C = C_{IV} + C_{HV} + C_V + C_T + C_{BT}.$$

### 2.3. Summarisk oversikt over datamaterialet. Deskriptive korrelasjons- og regresjonsberegninger

Tabell 1 gir gjennomsnitt og standardavvik for inntekts-, formues- og forbruksutgiftsvariablene, hovedpersonens alder (A) og antall husholdningsmedlemmer (N).<sup>10)</sup> Blant *inntektsartene* dominerer lønnsinntekten ( $R_1$ ) med ca. tre fjerdedeler av den totale disponible inntekt. Det er samtidig den komponent som viser relativt sett minst variasjon og har en variasjonskoeffisient på under 1.<sup>11)</sup> Gevinster ved tomte- og aksjesalg ( $R_6$ ) er den minste av inntektskomponentene med et gjennomsnitt på bare ca. 100 kroner, men viser store individuelle variasjoner. Posten gjeldsrenter (S) er i gjennomsnitt nesten tre ganger så stor som renter og utbytte av finanskapital ( $R_4$ ); saldoen av renter og utbytte er med andre ord negativ for gjennomsnittshusholdningen.

Blant *formueskomponentene* er faste eiendommer ( $F_3$ ) den dominerende, med et gjennomsnitt på nesten 60 000 kroner. Minst er formuesverdien av innbo og løsøre ( $F_4$ ), hvor observasjonsgjennomsnittet bare er 130 kroner. Dette avspeiler selvfølgelig de liberale vurderingskriterier som legges til grunn ved fastsettelse av ligningsverdien av denne del av formuen.<sup>12)</sup>

10) Lesere som sammenligner tallene med gjennomsnittstallene i Skarstad [37], som hovedsakelig bygger på samme primærdata, vil finne enkelte avvikelser. Disse skyldes dels definisjonsforskjeller, dels at Skarstad forsøker å korrigere for forskjeller i frafallsprosent mellom de enkelte hovedgrupper av husholdninger.

11) Variasjonskoeffisienten er definert som forholdet mellom (det empiriske) standardavviket og gjennomsnittet.

12) I 1973 ble således innbo og løsøre regnet som skattepliktig formue bare hvis brannforsikringsverdien oversteg 167 000 kr. Av det overskytende ble regulært medregnet bare 30-40 prosent av forsikringsverdien.

Tabell 1. Gjennomsnitt og empirisk standardavvik for analysens hovedvariable. Kroner. Data fra Forbruks-, Inntekts- og Formuesundersøkelsene 1973 (3 271 husholdninger)

Variabel	Observasjons- gjennomsnitt	Empirisk standard- avvik
R <sub>1</sub> : Lønnsinntekt mv. <sup>a)</sup> .....	28 445.43	23 806.98
R <sub>2</sub> : Folketrygdytelser <sup>a)</sup> .....	4 811.00	7 776.89
R <sub>3</sub> : Næringsinntekt <sup>a)</sup> .....	4 906.24	16 396.94
R <sub>4</sub> : Renteinntekter, utbytte mv. <sup>a)</sup> ....	603.86	1 974.51
S : Gjeldsrenter .....	1 735.49	8 809.13
R <sub>4</sub> -S : Renteinntekter og utbytte minus gjeldsrenter <sup>a)</sup> .....	-1 131.63	8 653.36
R <sub>5</sub> : Inntekt av fast eiendom <sup>a)</sup> .....	665.33	2 356.34
R <sub>6</sub> : Salgsgevinster, tomter og aksjer <sup>a)</sup>	101.49	1 448.23
R : Total disponibel inntekt .....	37 797.86	19 983.35
F <sub>1</sub> : Bankinnskudd mv. ....	16 023.77	35 084.75
G : Gjeld .....	44 352.15	159 007.16
F <sub>1</sub> -G : Bankinnskudd mv. minus gjeld ....	-28 328.38	158 019.28
F <sub>2</sub> : Ihendehaverobligasjoner, aksjer ..	10 274.57	70 763.32
F <sub>3</sub> : Faste eiendommer .....	59 175.56	166 328.07
F <sub>4</sub> : Innbo, løsøre .....	130.11	2 731.67
F <sub>5</sub> : Privatbiler, lystfartøyer .....	5 398.29	8 701.92
F <sub>6</sub> : Leieboerinnskudd, forsikrings- poliser mv. ....	3 042.18	12 520.80
F : Total nettoformue .....	49 692.32	111 280.03
C <sub>IV</sub> : Utgift til ikke-varige goder ....	15 523.77	9 063.53
C <sub>HV</sub> : Utgift til halv-varige goder ....	8 186.28	9 740.36
C <sub>V</sub> : Utgift til varige goder .....	5 496.14	8 650.48
C <sub>T</sub> : Utgift til tjenester ekskl. bolig- tjenester .....	5 394.98	8 041.85
C <sub>BT</sub> : Utgift til bolig tjenester .....	3 262.52	6 541.66
C : Total forbruksutgift .....	37 863.75	25 707.18
A : Hovedpersonens alder .....	51.01	16.07
N : Antall husholdningsmedlemmer ....	3.10	1.58

a) Etter proporsjonal fordeling av skattebeløpet.

Relativt sett minst variasjon viser bankinnskudd mv. (F<sub>1</sub>) og privatbiler og lystfartøyer (F<sub>5</sub>) med variasjonskoeffisienter på henholdsvis ca. 2.2 og ca. 1.6. Gjennomsnittshusholdningen i materialet har en total nettoformue på nesten 50 000 kroner - etter fradrag av en gjeld på vel 44 000 kroner. Det er verdt å bemerke at gjennomsnittlig formuesverdi av faste eiendommer er større enn gjennomsnittsverdien av den totale nettoformue - gjelden overstiger altså verdien av de øvrige formueskomponenter.



Den største av *konsumutgiftsgruppene* er utgifter til ikke-varige goder (ca. 40 prosent av totalutgiften), og den har den klart minste variasjonskoeffisient, ca. 0.6. Ikke uventet er det varige goder som viser størst relativ utgiftsvariasjon, men variasjonskoeffisienten er ikke større enn ca. 1.6. (Her må vi imidlertid ta i betraktning at utgiftsregistreringene for de varige goder stort sett er basert på lengre rapporteringsperiode enn for resten av konsumet; se avsnitt 2.2.4. Hadde rapporteringsperioden vært 14 dager for samtlige goder, er det sannsynlig at utgiften til varige goder ville ha vist vesentlig større spredning relativt sett.)

Den totale forbruksutgift er ubetydelig større enn den totale disponible inntekt for gjennomsnittshusholdningen; tallene er henholdsvis 37 864 kroner og 37 798 kroner. Hvis hele forbruksutgiften representerte umiddelbart forbruk, ville differansen, - 66 kroner, være et mål for gjennomsnittshusholdningens *sparing*. Dette tallet dekker imidlertid over store individuelle variasjoner; standardavviket er hele 22 079 kroner. I lys av de svakheter som hefter ved R som inntektsmål og ved C som uttrykk for verdien av konsumet, er dette ikke overraskende.<sup>13)</sup> Vi må her ta i betraktning at sparing også kan ta form av investering i konsumkapital, som inngår i utgiften til varige og til dels ikke-varige goder. Gjennomsnittstallet ovenfor er altså nærmest en indikator for den *finansielle* sparing. Direkte investering i fast eiendom, bl.a. egen bolig, kommer eksempelvis ikke med. Selv om det er vanskelig å akseptere at "gjennomsnittshusholdningen i Norge" ikke skulle ha en positiv sparing totalt sett, er det ikke urimelig at finanssparingen er negativ eller null for en stor del av husholdningene. Det hadde vært ønskelig å kunne konfrontere anslagene ovenfor med oppgaver over formuesbevegelser, men det er vi avskåret fra da det, som nevnt, kun finnes opplysninger om formuen ved slutten av 1973. Dette er dataproblemer vi må ta ad notam, men som vi i liten grad kan gjøre noe med. Ved å innhente supplerende informasjon vil de kanskje kunne elimineres - helt eller delvis - ved fremtidige undersøkelser.

En spesiell vanskelighet når det gjelder muligheten for å finne pålitelige anslag for sparingen, oppstår fordi forbruksutgiftstallene ikke refererer seg til samme tidsrom som inntektstallene og fordi forbruksutgiftsperioden har forskjellig "lokalisering" i forhold til inntektsregistreringsperioden for de enkelte husholdninger. De som observeres i desember, vil således gjennomgående ha høyere konsummotiverende inntekt ("normalinntekt") *i registreringsperioden* enn de som observeres tidligere i året. Dette betyr at de første gjennomgående vil figurere med en høyere forbruksutgift enn de

13) Hvis tallene justeres for frafall og det skjønnsmessig korrigeres for de skattefrie trygdeytelser mv. som ikke registreres i inntektsundersøkelsen, blir gjennomsnittsanslaget for sparingen positivt. (Skarstad [37], pp. 9,13.)

siste, selv om årsinntekten er den samme.<sup>14)</sup> Utgiftstallene viser da også signifikant positiv korrelasjon med regnskapsperiodens nummer for alle goder unntatt de varige (signifikansnivå 5 prosent). At de varige goder skiller seg ut, kan i noen grad tilskrives at utgiften til disse stort sett er rapportert for en 12-måneders-periode under ett. (Se avsnitt 2.2.4.) Derimot kan verken inntektsvariablene, formuesvariablene, hovedpersonens alder eller antall husholdningsmedlemmer påvises å ha noen signifikant korrelasjon med regnskapsperiodens nummer. Den rekkefølge husholdningene fører regnskap i, kan altså sies å være tilfeldig i forhold til disse variable. Betydningen av dette vil vi komme tilbake til i kapittel 3. Noen virkninger av tidsforskjvninger mellom forbruksutgiftsperioden og inntektsregistreringsperioden er diskutert mer inngående i appendiks til dette kapitlet.

Inntektssammensetningen viser betydelige forskjeller mellom sosialgruppene i datamaterialet. I følge tabell 2 mottar således den gjennomsnittlige lønnsstagerhusholdning over 95 prosent av inntekten som lønnsinntekt. I husholdninger hvor hovedpersonen er klassifisert som selvstendig næringsdrivende, utgjør næringsinntekt og realkapitalinntekt ca. 70 prosent av inntekten, mens husholdninger i gruppen ikke yrkesaktive figurerer med ca. 45 prosent av inntekten som folketrygdytelser og et omtrent like stort beløp som lønnsinntekt. Nettoverdien av renteinntekter og utbytte er negativ for samtlige sosialgrupper, men varierer betydelig, fra -0.7 prosent for ikke yrkesaktive til -9.0 prosent for selvstendige utenom primærnæringene. De forskjeller som her er nevnt, må naturligvis for en stor del tilskrives det faktum at hovedinntektstagerens "viktigste kilde til livsopphold" er et hovedkriterium for sosialgruppeinndelingen.

Tabell 3, som inneholder korrelasjonskoeffisienter mellom samtlige inntekts- og formuesvariable, gir et interessant supplement til tabell 1. Ikke overraskende finner vi at lønnsinntekt er negativt korrelert både med folketrygdytelser og næringsinntekt. (Korrelasjonskoeffisientene er henholdsvis -0.45 og -0.23.) Trygdeinntekten er til og med negativt korrelert med summen av inntektspostene. Dette gjenspeiler at husholdninger hvor trygdytelser utgjør en vesentlig del av inntekten, ofte er lavinntekts-husholdninger. Både gjelden og gjeldsrentebeløpet er positivt korrelert med totalinntekten. (Korrelasjonskoeffisientene er henholdsvis 0.24 og 0.25.) For gjeldsrentebeløpet er samvariasjonen så sterk at saldoen av renteinntekter/utbytte og gjeldsrenter er negativt korrelert med totalinntekten (korrelasjonskoeffisient -0.22). Blant inntektsartene er det næringsinntekten som viser sterkest samvariasjon med gjeldsrentene (korrelasjonskoeffisient 0.64).

14) De husholdninger som observeres på slutten av året, vil også ha bedre oversikt over årsinntekten og formuen ved årsskiftet, som de i sin tid vil rapportere på selvangivelsen. Økonometrisk sett er altså årsinntekten og formuen tilnærmet predeterminerte variable for disse husholdninger.

Tabell 2. Inntektssammensetning for de enkelte sosialgrupper

Hoved- personens sosial- gruppe	I n n t e k t s a r t					Antall hus- hold- nings- regn- skaper
	Lønns- inntekt  $R_1$	Folke- trygd- ytel- ser  $R_2$	Nærings- inntekt og inntekt av real- kapital $R_3+R_5+R_6$	Renter og utbytte, netto  $R_4-S$	Total inntekt  $R$	
<i>Kroner</i>						
Lønnstager ...	41 673.90	1 323.34	1 791.24	-1 159.99	43 628.49	1 829
Selvstendig, primærnæring .	7 330.20	4 160.03	25 189.78	-2 213.39	34 466.62	234
Selvstendig ellers .....	14 261.68	2 347.61	29 205.11	-3 768.46	42 045.94	232
Ikke yrkes- aktiv .....	12 089.62	12 088.41	2 674.59	-192.35	26 660.27	976
Alle .....	28 445.43	4 811.00	5 673.06	-1 131.63	37 797.86	3 271
<i>Prosent av totalinntekt</i>						
Lønnstager ...	95.5	3.0	4.1	-2.6	100	1 829
Selvstendig, primærnæring .	21.2	12.1	73.1	-6.4	100	234
Selvstendig ellers .....	33.9	5.6	69.5	-9.0	100	232
Ikke yrkes- aktiv .....	45.4	45.3	10.0	-0.7	100	976
Alle .....	75.3	12.7	15.0	-3.0	100	3 271

Alle de seks komponenter i bruttoformuen ( $F_1, \dots, F_6$ ) viser positiv samvariasjon innbyrdes, men den er svak i de fleste tilfelle. Høyest er korrelasjonskoeffisienten mellom ihendehaverobligasjoner/aksjer og faste eiendommer (0.26). Positiv samvariasjon er det også mellom gjelden og hver enkelt av bruttoformueskomponentene; særlig sterk er den for faste eiendommer, hele 0.92. Dette gjenspeiler at kjøp av slike formuesobjekter i stor utstrekning finansieres ved lån. Ved å beregne regresjonen av gjelden mhp. den samlede bruttoformue finner vi (standardavvikestimater er gitt i parentes)

$$G = -16\,188.98 + 0.64375 \sum_{i=1}^6 F_i, \\ (0.00675)$$

dvs. over observasjonsmaterialet løper en formuesøkning på 1 000 kroner gjennomgående parallelt med en gjeldsøkning på 644 kroner. Det skulle være unødvendig å påpeke at dette ikke er ment å skulle avspeile et årsaks-virkningsforhold, men bare er en høyst foreløpig sammenfatning av et sentralt trekk ved observasjonsmaterialet.

Tabell 3A. Korrelasjonsmatrise for inntektsvariable<sup>a)</sup>

	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	R <sub>6</sub>	S	R <sub>4</sub> -S	R
R <sub>1</sub>	1								
R <sub>2</sub>	-0.45052	1							
R <sub>3</sub>	-0.23165	-0.09290	1						
R <sub>4</sub>	-0.04154	0.05942	0.18347	1					
R <sub>5</sub>	0.14321	-0.03862	0.04016	0.13127	1				
R <sub>6</sub>	-0.02604	0.04025	-0.00202	0.04204	-0.00381	1			
S	0.11399	-0.07715	0.63562	0.19026	0.36327	0.00040	1		
R <sub>4</sub> -S	-0.12552	0.09210	-0.60520	0.03449	-0.33975	0.00919	-0.97459	1	
R	0.78659	-0.18553	0.25093	0.15770	0.15905	0.05899	0.24816	-0.21664	1

a) Symbolforklaring er gitt i tabell 1.

Tabell 3B. Korrelasjonsmatrise for formuesvariable<sup>a)</sup>

	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	F <sub>6</sub>	G	F <sub>1</sub> -G	F
F <sub>1</sub>	1								
F <sub>2</sub>	0.17897	1							
F <sub>3</sub>	0.18410	0.26030	1						
F <sub>4</sub>	0.01721	0.01803	0.00454	1					
F <sub>5</sub>	0.06073	0.04079	0.12751	0.03767	1				
F <sub>6</sub>	0.14741	0.14750	0.16296	0.00117	0.08826	1			
G	0.13839	0.28947	0.91647	0.03869	0.14022	0.20429	1		
F <sub>1</sub> -G	0.08277	-0.25154	-0.88132	-0.03512	-0.12761	-0.17283	-0.97552	1	
F	0.52828	0.68800	0.43713	-0.00398	0.12437	0.21139	0.20354	-0.08752	1

a) Symbolforklaring er gitt i tabell 1.

Tabell 3C. Korrelasjonsmatrise for inntektsvariable med hensyn på formuesvariable<sup>a)</sup>

	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	F <sub>6</sub>	G	F <sub>1</sub> -G	F
R <sub>1</sub>	-0.03094	0.00454	0.02088	0.09684	0.33804	0.03329	0.07916	-0.08652	-0.05621
R <sub>2</sub>	0.06543	-0.01275	-0.07168	-0.02713	-0.17916	-0.06099	-0.10225	0.11742	0.02995
R <sub>3</sub>	0.24856	0.17420	0.56851	-0.01017	0.08401	0.17333	0.54478	-0.49300	0.28628
R <sub>4</sub>	0.56202	0.45461	0.26068	0.00388	0.04166	0.29779	0.22835	-0.10500	0.56649
R <sub>5</sub>	0.04468	0.07511	0.27183	0.00398	0.06668	0.02413	0.20597	-0.19734	0.18187
R <sub>6</sub>	0.06466	0.00849	0.01165	-0.00245	0.02658	0.02323	0.00299	0.01134	0.04356
S	0.08060	0.16054	0.65851	0.01650	0.06786	0.11656	0.69505	-0.68150	0.13743
R <sub>4</sub> -S	0.04619	-0.05970	-0.61088	-0.01591	-0.05958	-0.05071	-0.65546	0.66981	-0.01065
R	0.22251	0.12700	0.23182	0.08987	0.38593	0.14071	0.24219	-0.19430	0.19958

a) Symbolforklaring er gitt i tabell 1.

Korrelasjonen mellom inntekt og formue er gjennomgående forholdsvis svak; korrelasjonskoeffisienten mellom disponibel inntekt og nettoformue er så lav som 0.20. Men det er enkelte interessante unntak. Renter og utbytte er, som rimelig kan være, positivt korrelert med finansformuen. (Korrelasjonskoeffisienten mhp. bankinnskudd er 0.56 og mhp. obligasjoner og aksjer 0.45.) Ved regresjon finner vi

$$R_4 = 51.21 + 0.02795 F_1 + 0.01021 F_2,$$

(0.00075)                      (0.00037)

med en multippel korrelasjonskoeffisient på 0.667. Dette indikerer at husholdningene i observasjonsmaterialet gjennomgående hadde en marginalavkastning etter skatt på 2.8 prosent på bankinnskudd og 1.0 prosent på aksjer og obligasjoner. En tilsvarende beregning basert på avkastning før skatt gir

$$R_4^* = 31.45 + 0.03717 F_1 + 0.02570 F_2.$$

(0.00125)                      (0.00062)

Dette stemmer godt med vanlige antagelser om finanskapitalavkastning i Norge.<sup>15)</sup> Også næringsinntekten og verdien av faste eiendommer er forholdsvis sterkt korrelert; korrelasjonskoeffisienten er 0.57. Dette avspeiler at faste eiendommer, bl.a. driftskapital, er betydelige formuesobjekter for selvstendig næringsdrivende. Som ventet er også gjeldsrentebeløpet og gjelden sterkt korrelert; korrelasjonskoeffisienten er 0.70. Ved regresjon finner vi

$$S = 27.64 + 0.03851 G$$

(0.00070)

og

$$S^* = 105.25 + 0.04617 G,$$

(0.00036)

som tilsvarende en marginal gjeldsrentesats på 4.6 prosent før skatt og 3.9 prosent etter skatt. Heller ikke dette er urimelig ut fra de opplysninger om gjennomsnittlig utlånsrente som finnes i den norske kredittmarkedstatistikk.<sup>16)</sup>

15) Se f.eks. [32]. Supplerende beregninger hvor også annengradsledd i  $F_1$  og  $F_2$  inngår, indikerer at avkastningsraten før skatt er avtagende med formuens størrelse for aksjer og obligasjoner, men ikke for bankinnskudd. Etter skatt er det signifikant avtagende utbytte både for aksjer/obligasjoner og for bankinnskudd. Det siste må naturligvis i stor grad tilskrives progressiviteten i inntekts- og formuesbeskatningen.

16) Se [40], tabellene 49 og 62.

OM BETYDNINGEN AV AT FORBRUKSUTGIFTSPERIODEN ER FORSKJØVET I FORHOLD TIL INNTEKTSREGISTRERINGSPERIODEN

Som påpekt i avsnitt 2.3 faller den periode da forbruksutgiftene er registrert, ikke sammen med registreringsperioden for inntekt. Formålet med dette appendiks er å se nærmere på hvilken betydning denne tidsforskyvningen kan ha for estimeringen av konsumfunksjonen.

Anta at konsumutgiften for et gode i en vilkårlig 14-dagers-periode bestemmes ved funksjonen

$$(1) \quad C_t = a + bR_t + cz_t + u_t \quad (t=1,2,\dots,26),$$

hvor  $C_t$  betegner konsumutgiften i 14-dagers-periode  $t$ ,  $R_t$  normalverdien (trendverdien) av inntekten i denne perioden,  $z_t$  verdien av en annen konsummotiverende variabel, f.eks. antall husholdningsmedlemmer eller hovedpersonens alder, og  $u_t$  er et stokastisk restledd med forventning lik null. Vi velger å la normalverdien av inntekten basert på lengre tids erfaringer være konsummotiverende fordi observert inntekt over et så kort tidsrom som 14 dager erfaringsmessig viser liten samvariasjon med konsumutgiften i denne perioden. I det helt korte løp er dessuten meningsinnholdet i begrepet inntekt diffust. For enkelhets skyld ser vi bort fra prisendringer i løpet av de 26 observasjonsperiodene, slik at vi kan forsvare å formulere modellen i nominelle verdier.

Betrakt en husholdning som fører regnskap i periode nr.  $\theta$  og intervjues ved utgangen av denne perioden. For de goder hvor forbruksutgiftstallene baseres på *regnskap*, er det altså  $C_\theta$  som blir registrert. Den tilsvarende årsutgift beregnes som  $26 C_\theta$ . For de resterende goder, hvor utgiftene rapporteres ved *intervju*, er det i prinsippet utgiftssummen fra intervjutidspunktet og ett år (=26 perioder) bakover, dvs.

$$(2) \quad C^* = \sum_{t=\theta-25}^{\theta} C_t$$

som registreres. Observasjonen av  $z$ -variabelen refererer seg vanligvis til intervjutidspunktet, dvs. den verdi som registreres, er  $z_\theta$ . For samtlige husholdninger er inntekten registrert på kalenderårsbasis. Den observerte inntekt er altså

$$(3) \quad R = \sum_{t=1}^{26} R_t.$$

Hvilken feil begår vi om vi estimerer  $a$ ,  $b$  og  $c$  ved å ta regresjonen av  $26C_{\theta}$ , resp.  $C^*$ , mhp.  $R$  og  $z_{\theta}$ ?

Anta at inntektsutviklingen tilnærmet følger en lineær trend:

$$R_t = \frac{R}{26} (kt+h).$$

Vi bestemmer  $h$  slik at (3) blir oppfylt og får da

$$(4) \quad R_t = \frac{R}{26} \{k(t-13.5)+1\}.$$

La oss først se på tilfellet da *utgiftsregistreringene baseres på regnskap*. Av (1), (3) og (4) følger

$$(5) \quad 26C_{\theta} = 26a + bR + 26cz_{\theta} + v,$$

hvor

$$(6) \quad v = bk(\theta-13.5)R + 26u_{\theta} = bk\theta' R + u',$$

idet vi har satt  $\theta' = \theta-13.5$  og  $u' = 26u_{\theta}$ . Her representerer  $\theta'$  registreringsperiodens nummer målt som avvik fra sitt gjennomsnitt.<sup>1)</sup> Den betingede forventning og varians til det sammensatte restledd  $v$  blir dermed

$$(7) \quad E(v|R, z_{\theta}) = bkRE(\theta'|R, z_{\theta}),$$

$$(8) \quad \text{var}(v|R, z_{\theta}) = b^2 k^2 R^2 \sigma_{\theta'}^2 + (26\sigma_u)^2,$$

idet vi antar at  $u'$  er ukorrelert med  $R$  og  $z_{\theta}$  og lar  $\sigma_{\theta'}^2$  og  $\sigma_u^2$  betegne variansene til henholdsvis  $\theta'$  og  $u_t$ . Restleddet i den konsumfunksjon vi estimerer, (5), er altså ukorrelert med de to høyresidevariable dersom  $E(\theta'|R, z_{\theta}) = 0$ , dvs. dersom  $\theta'$  er ukorrelert med disse variable. Denne forutsetning synes å være oppfylt i vårt tilfelle. Som vi påpekte i avsnitt 2.3, viser verken inntektsvariablene, formuesvariablene, hovedpersonens alder eller antall husholdningsmedlemmer signifikant samvariasjon med regnskapsperiodens nummer. Følgelig vil minste kvadratets metode anvendt på (5) gi konsistente estimater for strukturcoeffisientene  $a$ ,  $b$  og  $c$ . Estimaten vil derimot ikke være effisiente, fordi variansen til  $v$ , gitt ved (8), ikke er konstant men stiger med inntekten - forutsatt at inntektens vekstrate  $k$  er forskjellig fra null.

La oss dernest betrakte tilfellet da *utgiftene rapporteres ved intervju*. Vi bruker da  $C^*$  som "proxyvariabel" for konsumutgiften. Av (1)-(4) følger

1) Strengt tatt gjelder dette bare tilnærmet fordi observasjonene ikke fordeles seg helt jevnt over året.

$$(9) \quad C^x = 26a + b(1-12.5k)R + 26cz_\theta + w,$$

hvor

$$(10) \quad w = bk(\theta-13.5)R + 26c \left( \sum_{t=\theta-25}^{\theta} z_t / (26z_\theta) - 1 \right) z_\theta + \sum_{t=\theta-25}^{\theta} u_t.$$

$$= bk\theta'R + 26c\mu z_\theta + u''.$$

Vi har her innført symbolene

$$\mu = \sum_{t=\theta-25}^{\theta} z_t / (26z_\theta) - 1,$$

$$u'' = \sum_{t=\theta-25}^{\theta} u_t.$$

$\mu$  betegner den relative forskjell mellom gjennomsnittsverdien for  $z$  i perioden fra  $\theta-25$  til  $\theta$  og verdien på tidspunkt  $\theta$ , intervjutidspunktet. Følgelig er

$$(11) \quad E(w|R, z_\theta) = bkRE(\theta'|R, z_\theta) + 26cz_\theta E(\mu|R, z_\theta),$$

$$(12) \quad \text{var}(w|R, z_\theta) = b^2 k^2 R^2 \sigma_\theta^2 + (26c)^2 z_\theta^2 \sigma_\mu^2 + 26\sigma_u^2$$

forutsatt at  $\theta'$  og  $\mu$  er ukorrelerte og  $\mu$  har varians lik  $\sigma_\mu^2$ . Restleddet i den konsumfunksjon vi estimerer, (9), vil altså være ukorrelert med  $R$  og  $z_\theta$  dersom både  $\theta'$  og  $\mu$  er ukorrelert med disse variable ( $E(\theta'|R, z_\theta) = E(\mu|R, z_\theta) = 0$ ).

Den essensielle forskjell mellom dette tilfelle og tilfellet da utgiftsregistreringene baseres på regnskap, er at inntekten i gjennomsnitt vil være datert senere enn konsumutgiften. Dette kommer til uttrykk ved at  $R$  i ligning (9) opptrer med en koeffisient lik  $b(1-12.5k)$ , som betyr at minste kvadraters metode anvendt på (9) vil underestimere  $b$  med en andel omtrent lik halvparten av inntektens vekstrate på årsbasis.

Siden inntekten gjennomgående vokser gjennom året, må vi derfor alt i alt regne med at forskyvningen av forbruksutgiftsperioden i forhold til inntektsregistreringsperioden fører til en viss underestimering av inntektens marginale konsumtilbøyelighet. Men dette momentet alene vil neppe bidra til en feil på mer enn høyst et par prosent.



### 3. KONSUMUTGIFTENS AVHENGIGHET AV INNTEKT, FORMUE, ALDER OG HUSHOLDNINGSTYPE. REGRESJONSANALYSE OG TESTING AV MULTIPLE HYPOTESER

#### 3.1. Innledende bemerkninger

Hvilken rolle spiller inntektens størrelse og sammensetning for husholdningenes konsumterspørsmål? Hva betyr formuen? Er alder og sosialgruppe viktige forklaringsfaktorer? Hvordan bør vi ta hensyn til antall personer i husholdningen? Viser konsummønsteret spesielle trekk for visse goder, eller kan vi betrakte hele konsumutgiften under ett? Disse generelle spørsmål danner utgangspunktet for dette kapitlet. Bortsett fra at vi begrenser interessen til de variable som er nevnt ovenfor, vil vi være forholdsvis lite restriktive når det gjelder a priori forutsetninger om konsumfunksjonens form. Vår analyse kan derfor sies å ha et noe videre perspektiv enn de fleste undersøkelser på dette området.<sup>1)</sup> Spesielt gir litteraturen relativt få eksempler på forsøk på å kartlegge betydningen av inntektens og formuens artssammensetning. I teoretiske analyser av vekst- og fordelingsproblemer spiller forskjeller i konsum- og sparetilbøyelighetene for lønnsinntekt på den ene siden og kapitalinntekt og bedriftsoverskudd (profitt) på den annen ofte en sentral rolle (se f.eks. Kaldor [25]). Slike analyser bygger gjerne stilltiende på den antagelse at lønnsinntekt har en markert høyere konsumtilbøyelighet enn kapitalinntekt og bedriftsoverskudd. Vårt datamateriale kan kaste lys over dette spørsmålet.

Med et så generelt utgangspunkt blir valget av statistisk metode et kjerneproblem. De spørsmål vi stilte innledningsvis, kan presiseres i forskjellige retninger. Mange konkurrerende hypoteser om forbruksstrukturen vil være i fokus samtidig. Skal vi ha håp om å oppnå velfunderte og interessante resultater, kan vi ikke betrakte de enkelte hypoteser isolert - vi må utforme en analysemetode som eksplisitt tar hensyn til måten de enkelte hypoteser er knyttet sammen på.

1) Undersøkelser med til dels noe mer avgrensede problemstillinger er Crockett [12] og Crockett og Friend [13], som retter søkelyset mot formuen som konsummotiverende variabel ved siden av inntekten; Modigliani og Brumberg [31], Fisher [17] og Thurow [43], som tar for seg alderens betydning; Watts [45], som både studerer formuen og alderen, foruten diverse sosioøkonomiske variable; Bodkin [8] og Reid [33], som spesielt undersøker konsumeffekten av tilfeldige, uventede inntekter ("windfall gains") og Fisher [17] og Liviatan [26], som blant annet analyserer strukturforskjeller mellom lønnsstagerhusholdninger og husholdninger hvor hovedpersonen er selvstendig næringsdrivende. Oversikter over litteraturen er gitt i Mayer [29] og Ferber [15], [16].

Ideelt sett burde vi begynne med å formulere en modell som omfatter all den forhåndsinformasjon som vi mener å ha om konsumfunksjonen, og som er så generell at den som spesialtilfelle inkluderer alle de hypoteser vi kan ønske å analysere. Dernest burde vi, med utgangspunkt i modellen, utforme en *strategi* for å skjære oss gjennom "nettets av konkurrerende hypoteser" for å nå fram til den beste ut fra et på forhånd valgt optimalitetskriterium. Når hele strategien er fastlagt - dvs. når vi har en klar oppskrift på hva vi skal gjøre i alle tenkelige situasjoner - bringes datamaterialet inn, og vi trekker vår konklusjon på basis av den strategi som er valgt.

### 3.2. Nærmere om analyseopplegget

Det er sjelden praktisk mulig å følge et slikt "idealopplegg" fullt ut, men vi vil her anvende det et stykke på vei. I stedet for å formulere én modell som alle aktuelle hypoteser er spesialtilfelle av, stiller vi opp *fire modeller* og fire tilhørende *analyse skjemaer* (dvs. skjemaer av sammenknyttede hypoteser). Hvert skjema konsentrerer oppmerksomheten om visse sider av konsumets avhengighet av sine potensielle forklaringsvariable og behandler de øvrige summarisk. Innenfor hvert skjema følger vi en strategi av den type som er angitt ovenfor, men når vi går over til et nytt skjema, starter vi "på bar bakke". Vi legger altså liten vekt på å se de fire delanalysene i sammenheng. Grunnen til at vi har måttet modifisere "idealoppskriften" på denne måten er at antall variable (originale og transformerte) i basismodellen ellers ville bli svært stort og tallet på hypoteser nærmest prohibitivt.<sup>2)</sup>

Analyseskjemaene kan stikkordmessig beskrives på følgende måte: Skjema I er utformet med sikte på å kartlegge konsumets avhengighet av inntektens og formuens sammensetning. I skjema II er hovedformålet å undersøke om konsumtilbøyelighetene for de enkelte inntektsarter er de samme for alle sosialgrupper. Begge disse skjemaene bygger på at konsumfunksjonen er lineær i inntekts- og formuesvariablene. Denne forutsetning fravikes i skjemaene III og IV, men her sløyfer vi til gjengjeld inntekts- og formues-sammensetningen og hovedpersonens sosialgruppe fra listen av forklaringsvariable. Basismodellen i skjema III er en konsumfunksjon uttrykt som et fullstendig annengradspolynom i inntekt, formue, hovedpersonens alder og antall husholdningsmedlemmer. Hovedformålet er her å undersøke om funksjonsformen avviker fra den lineære. Endelig tar skjema IV sikte på å undersøke hvorvidt konsumfunksjonen er homogen i matematisk forstand.

---

2) Et trivielt, men ikke desto mindre reelt, praktisk problem var at det datamaskinprogram for regresjonsberegninger som stod til rådighet, har begrenset variabelkapasitet; maksimalt 44 regressorvariable (høyresidevariable) kan behandles samtidig.

Vi vil kort beskrive analysemetoden. For hvert av de fire analyse-skjemaene er utgangspunktet at det foreligger en relativt vid modell - *basismodellen* - og at vi ønsker å undersøke hvilke av de spesifiserte forklaringsvariable (og transformasjoner av slike variable) i denne modellen som er essensielle og hvilke som er irrelevante.<sup>3)</sup>

Vi danner et *hierarki av hypoteser* (konkurrerende modellspesifikasjoner)  $H_A$ ,  $H_{B_i}$  ( $i=1,2,\dots$ ),  $H_{C_{ij}}$  ( $i=1,2,\dots; j=1,2,\dots$ ),  $H_{D_{ijk}}$  ( $i=1,2,\dots; j=1,2,\dots; k=1,2,\dots$ ) etc., hvor  $H_A \supset H_{B_i} \supset H_{C_{ij}} \supset H_{D_{ijk}} \supset \dots$  for alle verdier av  $i$ ,  $j$ ,  $k$  etc.<sup>4)</sup> Fotskriftene A, B, C, D, ... betegner nivåer i hypotesehierarkiet.  $H_A$  er basismodellen. Hypotesene på B-nivået er spesialtilfelle av denne, hypotesene på C-nivået spesialtilfelle av én eller flere av B-hypotesene etc. Regulært vil det nederste nivå i hierarkiet inneholde kun én hypotese, som følgelig er et spesialtilfelle av alle de øvrige.

Vi vil teste

$H_{B_i}$  mot  $H_A$  for alle  $i$ ,

$H_{C_{ij}}$  mot  $H_{B_i}$  for alle  $i$  og  $j$ ,

$H_{D_{ijk}}$  mot  $H_{C_{ij}}$  for alle  $i$ ,  $j$  og  $k$ ,

etc.

og ønsker å sikre oss kontroll med *sannsynligheten for å begå forkastningsfeil minst én gang*, dvs. påstå minst én gang at den restriktive hypotese bør forkastes til fordel for den mer generelle når det faktisk ikke er grunnlag

3) I engelskspråklig litteratur betegnes basismodellen med "*maintained hypothesis*". Et problem beslektet med det vi har beskrevet ovenfor, kan bestå i at vi har utført en enkelt signifikanstest innenfor basismodellen og er blitt ledet til å forkaste en restriktiv hypotese (f.eks. en konsumfunksjon med bare et par forklaringsvariable) mot en vesentlig videre (f.eks. med 10 tilleggsvariable). Vi finner imidlertid dette svaret utilfredsstillende fordi det er for upresist. Vi ønsker å vite *på hvilken måte* den restriktive hypotese blir forkastet: hvilke av de potensielle tilleggsvariable som "burde ha vært med", hvorvidt to koeffisienter som er satt like "burde ha vært forskjellige" etc.

4) Symbolet  $\supset$  er å tolke på følgende måte: La A og B være to påstander (hypoteser). Utsagnet  $A \supset B$  betyr da at hvis B gjelder, så gjelder også med nødvendighet A, dvs. at alle mulige "virkeligheter" som er forenlige med B, også er forenlige med A.

for en slik konklusjon. For hver av deltestene anvender vi en F-test basert på observatoren

$$F = \frac{Q_0 - Q_1}{Q_1} \frac{n - K_1 - 1}{K_1 - K_0} = \frac{R_1^2 - R_0^2}{1 - R_1^2} \frac{n - K_1 - 1}{K_1 - K_0},$$

hvor  $Q_0$  og  $Q_1$  er residual kvadratsum og  $R_0$  og  $R_1$  den multiple korrelasjonskoeffisient under henholdsvis nullhypotese og alternativ,  $K_0$  og  $K_1$  det tilsvarende antall forklaringsvariable i regresjonsligningene og  $n$  antall observasjonssett (= 3271). Under de respektive nullhypoteser vil denne observatoren være F-fordelt med  $K_1 - K_0$  og  $n - K_1 - 1$  frihetsgrader, forutsatt at restleddene i regresjonsligningene er uavhengige og identisk normalt fordelte. Hvis  $N$  er antall deltester som er spesifisert a priori, og hver har testnivå  $\alpha$ , vil  $N\alpha$  være en øvre grense for sannsynligheten for å begå forkastningsfeil minst én gang.<sup>5)</sup> Vi "tester oss nedover" i hierarkiet, fra basismodellen til de stadig mer spesialiserte hypoteser, inntil en hypotese forkastes med nivå  $\alpha$ . Den alternative hypotese ved den siste testen - forutsatt at bare én hypotese har denne egenskap - velges som "løsning" på testproblemet. Den er kjennetegnet ved at den ikke inneholder "unødvendige" regressorvariable samtidig som den gir "effektivt bedre forklaringskraft" enn de mer restriktive hypoteser. Men det kan tenkes at metoden gir flere løsninger, og da har vi intet objektivt kriterium for hvilken som er best. Som vi skal se, oppstår dette problem i forbindelse med analyseskjema III. En annen svakhet ved metoden er at det "sanne" nivå for totaltesten kan være vesentlig lavere enn maksimalnivået  $N\alpha$ . Det later til at begge disse komplikasjoner gjennomgående oppstår desto hyppigere jo større antall tester er - uten at vi kan påstå at det gjelder en klar sammenheng.

Denne inferensmetode (dvs. metode for å trekke statistisk velfunderte slutninger) er selvfølgelig bare én blant flere mulige. Andre kontrollkriterier enn ønsket om å begrense sannsynligheten for å begå forkastningsfeil minst én gang kan være aktuelle. Metoden ovenfor har det fortrinn at den er relativt enkel å anvende, selv når hypotesestrukturen er komplisert - som

5) Beviset er antydningvis følgende: La  $b$  betegne den begivenhet at vi begår forkastningsfeil ved deltest nr.  $i$  ( $i \in \{1, \dots, N\}$ ). Vi har da:

$$P(\text{begå forkastningsfeil minst én gang}) \\ = P(b_1 \cup b_2 \cup \dots \cup b_N) \leq \sum_{i=1}^N P(b_i) = N\alpha,$$

hvor ulikhetstegnet følger av grunnregler i sannsynlighetsregning (likhetstegnet gjelder det tilfelle da  $b_1, \dots, b_N$  er uavhengige.) (Jfr. det tilsvarende resonnement i forbindelse med "Bonferroni t-tester" for variansanalysesituasjoner i Spjøtvoll [38], p. 102.)

den er i vårt tilfelle.<sup>6)</sup> I mer oversiktlige situasjoner, f.eks. situasjoner hvor det eksisterer en entydig rangordning mellom hypotesene (alle hypoteser er "nested", ifølge engelsk terminologi), finnes andre og mer interessante måter å formulere inferensproblemet på.<sup>7)</sup>

På den annen side er vår metode klart å foretrekke fremfor standard-metodene for "trinnsvis regresjon".<sup>8)</sup> En hovedsvakheter ved disse - i hvert fall slik de finnes implementert i programpakker for datamaskiner - er at de ikke krever at man formulerer alle aktuelle hypoteser før man ser på data-materialet. Med slike metoder blir det derfor ikke mulig å holde den statistiske usikkerhet i konklusjonene under kontroll.

### 3.3. Konsumets avhengighet av inntektens og formuens artssammensetning: Analyse-skjema I

En husholdnings inntekt kommer vanligvis fra flere kilder, og vi beskrev i avsnitt 2.2.2 hvordan våre primærdata kan utnyttes til å splitte totalinntekten i forholdsvise grove artsgrupper. A priori er det lite som tilsier at konsumtilbøyeligheten er den samme for alle inntektsarter, f.eks. at en lønnsøkning på 1 000 kroner skulle anvendes på samme måte som en økning i trygdeytelser eller aksjeutbytte på 1 000 kroner. Noe tilsvarende kan gjelde for formue. I den utstrekning formuen er direkte konsummotiverende - utover den indirekte virkning den har ved at formuesavkastning inngår i inntekten - vil trolig dens *likviditetsgrad* og fysiske form ha betydning for konsumtilbøyeligheten, fordi dette er egenskaper som bestemmer hvor lett formuen kan transformeres til konsumgoder. Høy likviditet har de fleste former for bankinnskudd, noe mindre likvide er ihendehaverobligasjoner og aksjer, mens likviditeten vanligvis er lav for faste eiendommer og for de fleste typer av konsumkapital.

Husholdningenes gjeld og gjeldsrentebetaling må også inn i dette skjemaet. En rimelig hypotese er at disse to variable isolert sett virker negativt på trangen til å konsumere fordi de begrenser spillerommet for husholdningenes økonomiske disposisjoner. Spesielt kunne vi tenke oss at gjeld og gjeldsrenter inngår symmetrisk med positive formues- og inntektskomponenter i konsumfunksjonen. Mest nærliggende ville det i så fall være å se gjelden som en motpost til bankinnskudd og gjeldsrentene som en motpost til renteinntekter.<sup>9)</sup>

6) I et par nyere arbeider av Dale Jorgenson mfl. er det benyttet i hovedtrekk samme inferensmetode. Se Christensen, Jorgenson og Lau [10], pp. 379-381, og Conrad og Jorgenson [11].

7) Se f.eks. Anderson [2], avsnitt 6.4, og Mizon [30], pp. 1225-1229.

8) En oversikt over de viktigste av disse metodene finnes i Draper og Smith [14], kapittel 6.

9) For langsiktig gjeld ville imidlertid ikke dette uten videre være noen god løsning. Datamaterialet inneholder dessverre ingen opplysning om hvor langsiktige de enkelte gjeldsforhold er, så vi har ikke kunnet bringe denne variabel inn i modellen.

Dette motiverer til å la funksjonen

$$(3.1) \quad C = \alpha_0 + \alpha_A A + \alpha_N N + \sum_{i=1}^6 \beta_i R_i + \beta_4^* S + \sum_{i=1}^6 \gamma_i F_i + \gamma_1^* G + u_I$$

være basismodell i analyseskjema I. Her betegner C konsumutgiften - enten totalt eller til én av de fem utgiftsgruppene - A hovedpersonens alder, N antall husholdningsmedlemmer,  $R_i$  inntekt av type i etter skatt,  $F_i$  formue av type i, S gjeldsrenter, G samlet gjeld<sup>10)</sup> og  $u_I$  et stokastisk restledd. Vi forutsetter at verdien av  $u_I$  for de enkelte husholdninger er uavhengige og har identisk samme normale fordeling. Variablene A og N har karakter av "bakgrunnsvariable" og gir til sammen en summarisk beskrivelse av "husholdningstype". Normalt vil det nemlig være slik at når hovedpersonens alder er gitt, vil alderen til de øvrige husholdningsmedlemmer befinne seg innenfor relativt snevre grenser. En mer raffinert representasjon av husholdningstype vil neppe være mulig uten en betydelig økning i koeffisientantallet. Men selv en så enkel parametrisering som (3.1) fanger formodentlig opp hovedtendensene. Også i analyseskjemaene II-IV benytter vi A og N til å karakterisere husholdningstype.

Hypotesehierarkiet er spesifisert i figur 1.  $H_A$ , basismodellen, er ligning (3.1) uten restriksjoner,  $H_{B1}$  innebærer at renteinntekter og -utgifter påvirker konsumet symmetrisk, slik at bare differensen har betydning,  $H_{B2}$  at bankinnskudd og gjeld inngår symmetrisk, mens  $H_C$  kombinerer disse to hypoteser. Dette er den eneste hypotese på det tredje nivå i skjemaet. Også D-nivået inneholder én hypotese. Den innfører som tilleggsrestriksjoner at  $R_5$  (nettoinntekt av bolig og annen fast eiendom) og  $R_6$  (gevinster ved tomte- og aksjesalg) har samme marginale konsumtilbøyelighet som  $R_3$  (næringsinntekt), og at det tilsvarende gjelder de tre finansielle formueskomponenter  $F_1$  (bank- og postgiroinnskudd),  $F_2$  (ihendehaverobligasjoner og aksjer) og  $F_6$  (leieboerinnskudd og forsikringspoliser) på den ene side og de to konsumkapitalkomponentene  $F_4$  (innbo og løsøre) og  $F_5$  (privatbiler, lystfartøyer) på den annen. Spillerommet innskrenkes ytterligere på E- og F-nivået: i  $H_{F1}$  inngår f.eks. hele nettoformuen som én samlevariabel, mens inntekten deles i to, lønnsinntekt pluss folketrygdytelser ( $R_1+R_2$ ) og resten. I  $H_{F2}$  slås hele nettoinntekten sammen, mens formuen deles i to: netto finansformue ( $F_1+F_2+F_6-G$ ) og realformue ( $F_3+F_4+F_5$ ). Trådene samles i  $H_G$ , hvor bare totalverdiene av inntekten og formuen har betydning. Pilene indikerer bevegelse fra det generelle til det stadig mer spesialiserte; den eller de tilleggsrestriksjoner som pålegges når en går fra en hypotese til den neste, er angitt i de ovale figurer på hver enkelt pil.

10) Nærmere detaljer om variabeldefinisjoner og beregningsprinsipper er gitt i avsnitt 2.2.

Verdiene av observatoren  $F$  og de tilhørende 99 prosent- og 99,9 prosent-fraktiler for de enkelte deltester er gitt i tabell 5. Ved å forkaste de respektive nullhypoteser - angitt i kolonnen under  $H_0$  i tabell 5 - når disse fraktiler overskrides, oppnår vi at hver deltest får et nivå  $\alpha = 1$  prosent, resp. 0.1 prosent. Siden det totale antall tester (antall piler i figur 1) er 14, gir dette et nivå for totaltesten på maksimalt 14, resp. 1.4 prosent.

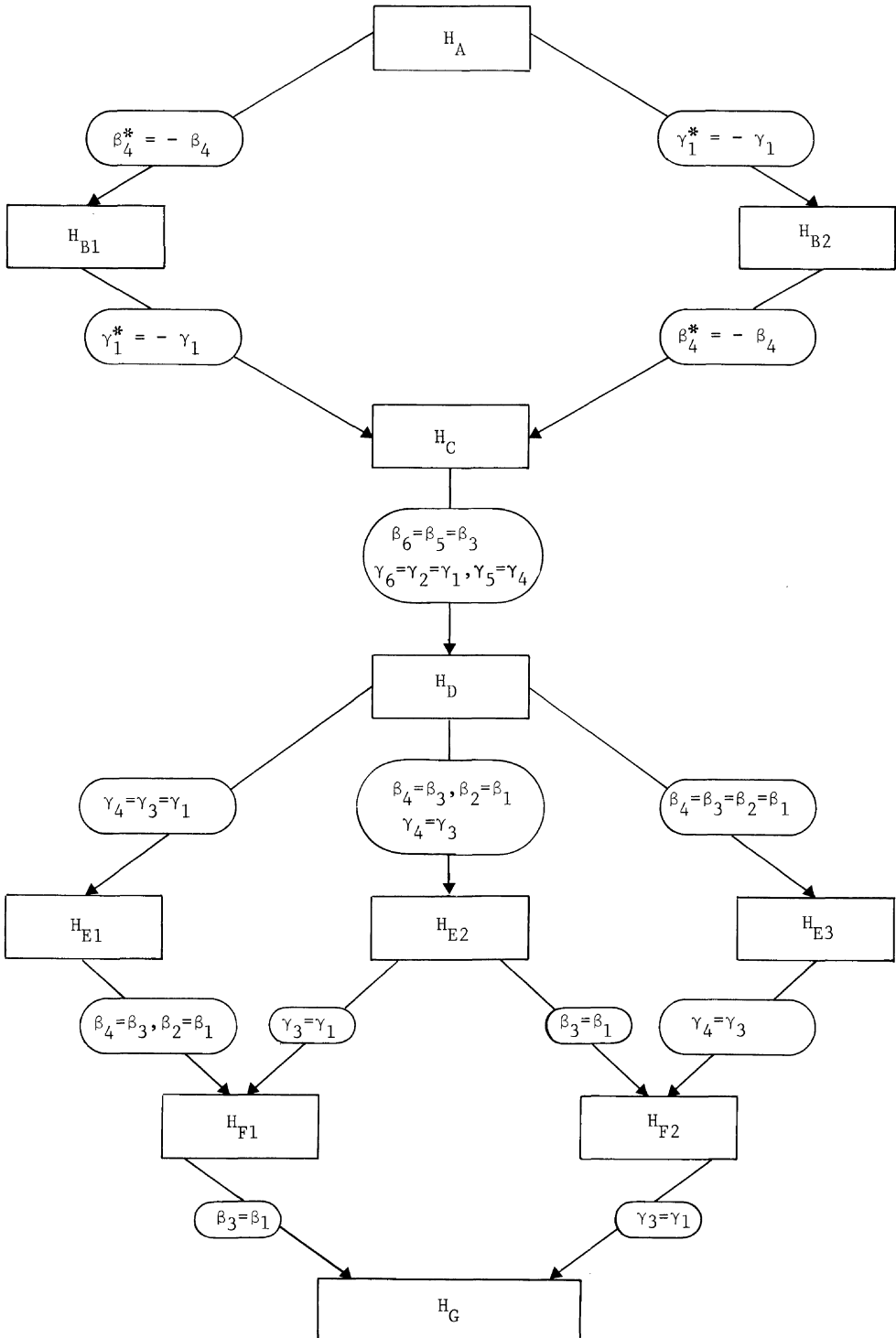
Testresultatet er oppsummert i tabell 4. For ikke-varige goder ( $C_{IV}$ ) og for begge tjenestegrupper ( $C_T$  og  $C_{BT}$ ) blir utfallet det samme enten vi velger  $\alpha = 1$  prosent eller 0.1 prosent. Med begge valg av testnivå finner vi at renteinntekter og gjeldsrenter virker symmetrisk for samtlige utgiftsgrupper ( $\beta_4^* = -\beta_4$ ;  $H_{B1}$  aksepteres), mens gjeld inngår symmetrisk med bankinnskudd for alle goder unntatt de ikke-varige ( $\gamma_1^* = -\gamma_1$ ;  $H_{B2}$  aksepteres). De ikke-varige goder synes altså å ha den egenskap at en økning i bankinnskuddet ikke påvirker konsumutgiften på samme måte som en tilsvarende nedgang i gjelden. Ikke i noe tilfelle aksepteres hypotesene  $H_{F1}$ ,  $H_{F2}$  eller  $H_G$  - dvs. både inntekts- og formuessammensetningen later til å ha betydning for konsumetterspørselen.

Tabell 4. Analysekjema I. Testresultater

Avhengig variabel	Testresultat med et nivå for deltestene lik	
	$\alpha = 0.01$	$\alpha = 0.001$
C (alle goder)	$H_{B1}$	$H_C$
$C_{IV}$ (ikke-varige goder)	$H_{B1}$	$H_{B1}$
$C_{HV}$ (halv-varige goder)	$H_D$	$H_{E3}$
$C_V$ (varige goder)	$H_C$	$H_{E3}$
$C_T$ (tjenester ekskl. bolig)	$H_{E1}$	$H_{E1}$
$C_{BT}$ (boligtjenester)	$H_C$	$H_C$

De koeffisientestimer som motsvarer testresultatet (med 1 prosentnivå for deltestene), er gitt i tabell 6. Lønns- og trygdeinntekt synes ikke å ha klart høyere konsumtilbøyeligheter enn inntekten for øvrig. For tre av utgiftsgruppene - ikke-varige goder, halv-varige goder og tjenester utenom boligtjenester - er endog trygdeinntekt den inntektsart som har lavest konsumtilbøyelighet. Dette er bemerkelsesverdig fordi det i makroøkonomiske resonnementer ofte stilltiende forutsettes at lønns- og trygdeøkninger har en sterkere konsumeffekt enn en tilsvarende økning i annen husholdningsinntekt. Det er renteinntekt som kommer ut med høyest konsumtilbøyelighet for både ikke-varige og halv-varige goder. Lønnsinntektens konsumtilbøyelighet er høyest bare for tjenester utenom boligtjenester.

Figur 1. Hypotesestrukturen i analyseeskjema I.





Estimatene for formueskoeffisientene gir ikke et éntydig bilde. En forklaring på de urimelig høye estimater for innbo og løsøre ( $\gamma_4$ ) og for privatbiler og lystfartøyer ( $\gamma_5$ ) må være at ligningsverdien gjennomgående er vesentlig lavere enn markedsverdien (og gjenanskaffelsesverdien)<sup>11)</sup>, slik at konsumtilbøyeligheten overestimeres. Det er i denne forbindelse interessant at estimatene for  $\gamma_4$  og  $\gamma_5$  er desidert høyest for varige konsumgoder. Siden det er sluttformuen i 1973 som er registrert, slik at formuestallene vil inneholde ligningsverdien av konsumkapital anskaffet i løpet av året, er dette et rimelig resultat. Det kan også gi noe av forklaringen på at  $\gamma_1$  kommer ut med negativt og  $\gamma_1^*$  med positivt estimat for fire av de fem utgiftsgruppene. Noen husholdninger kan ha økt sitt konsum ved å stifte gjeld eller tære på bankinnskudd. Isolert sett trekker dette i retning av at C er positivt korrelert med G og negativt korrelert med  $F_1$ .<sup>12)</sup>

Tabell 5. Analysekjema I. F-observatorer for deltestene

$H_0$	$H_1$	$K_0$	$K_1$	Avhengig variabel						$F_{0.99}$	$F_{0.999}$
				C	$C_{IV}$	$C_{HV}$	$C_V$	$C_T$	$C_{BT}$		
$H_{B1}$	$H_A$	15	16	0.21	0.53	0.19	0.49	1.02	0.66	6.63	10.83
$H_{B2}$	$H_A$	15	16	6.76	30.80	0.27	0.08	0.32	0.13	6.63	10.83
$H_C$	$H_{B1}$	14	15	7.02	33.52	0.61	0.00	0.03	0.61	6.63	10.83
$H_C$	$H_{B2}$	14	15	0.50	3.24	0.53	0.41	0.71	1.12	6.63	10.83
$H_D$	$H_C$	9	14	5.12	5.32	2.60	3.72	0.38	5.50	3.02	4.10
$H_{E1}$	$H_D$	7	9	38.09	7.99	7.79	149.57	1.80	0.04	4.61	6.91
$H_{E2}$	$H_D$	6	9	35.15	11.13	10.78	99.95	6.00	0.73	3.78	5.42
$H_{E3}$	$H_D$	6	9	9.74	9.17	5.29	3.90	9.26	8.69	3.78	5.42
$H_{F1}$	$H_{E1}$	5	7	14.30	9.22	8.35	0.32	8.18	1.11	4.61	6.91
$H_{F1}$	$H_{E2}$	5	6	0.00	1.11	0.00	0.00	1.96	0.13	6.63	10.83
$H_{F2}$	$H_{E2}$	5	6	8.71	3.45	0.81	4.85	11.45	24.16	6.63	10.83
$H_{F2}$	$H_{E3}$	5	6	84.46	9.24	17.14	292.75	1.72	0.27	6.63	10.83
$H_G$	$H_{F1}$	4	5	9.06	2.79	0.83	5.02	13.99	24.40	6.63	10.83
$H_G$	$H_{F2}$	4	5	0.40	0.50	0.02	0.18	4.45	0.38	6.63	10.83

11) Se fotnote 12 i kapittel 2.

12) Det ligger imidlertid her an til et "simultanitetsproblem": Sluttverdien av finansformuen er begrepsmessig lik begynnelsesformuen pluss inntekten minus verdien av den samlede konsumutgift, mens sluttverdien av konsumkapitalen er lik begynnelsesverdien pluss utgiften til kjøp av varige konsumgoder minus verdien av depresieringen av disse godene. Dette peker i retning av en modell hvor konsumutgift og sluttformue bestemmes simultant og hvor inntekten og begynnelsesformuen opptrer som predeterminerte variable. Hvis dette er en korrekt beskrivelse, vil minste kvadraters metode føre til skjeve (inkonsistente) estimater. Vi kommer tilbake til dette problemet i avsnitt 4.4.2.

Tabell 6. Analysekjema I. Koeffisientestimater svarende til testresultatet.  $\alpha = 0,01$ 

Koeffisient	Avhengig variabel					
	C	C <sub>IV</sub>	C <sub>HV</sub>	C <sub>V</sub>	C <sub>T</sub>	C <sub>BT</sub>
$\alpha_0$	11633.15	2942.00	1271.11	4147.90	-792.60	4304.00
$\alpha_A$	-76.45	8.90	-20.87	-56.43	28.69	-41.55
$\alpha_N$	3345.74	2120.57	1028.23	159.70	232.31	-191.81
$\beta_1$	0.5200	0.1613	0.1307	0.0537	0.1163	0.0518
$\beta_2$	0.2958	0.0990	0.0503	0.0628	0.0483	0.0312
$\beta_3$	0.4571	0.1571	0.1196	0.0954	0.0758	0.0049
$\beta_4$	0.6061	0.2403	0.1664	0.0732	0.1131	0.0048
$\beta_4^*$	-0.6061	-0.2403	-0.1664	-0.0732	-0.1131	-0.0048
$\beta_5$	0.4438	0.1398	0.1196	0.1039	0.0758	0.0676
$\beta_6$	0.4775	0.1946	0.1196	0.0406	0.0758	-0.0441
$\gamma_1$	-0.0454	-0.0307	-0.0024	-0.0023	0.0033	-0.0032
$\gamma_1^*$	0.0109	0.0046	0.0024	0.0023	-0.0033	0.0032
$\gamma_2$	0.0130	0.0032	-0.0024	0.0038	0.0033	0.0019
$\gamma_3$	-0.0078	-0.0013	-0.0010	-0.0043	0.0033	-0.0032
$\gamma_4$	0.2039	0.0199	0.0666	0.1154	0.0033	0.0818
$\gamma_5$	0.3722	0.0338	0.0666	0.3031	0.0033	-0.0084
$\gamma_6$	0.0640	0.0089	-0.0024	-0.0183	0.0033	0.0345
$\hat{\sigma}_{uI}$	20011.53	6929.09	8787.60	7830.22	7640.11	6376.35

Selv om vi ser bort fra det siste momentet, er det i og for seg ikke urimelig at konsum og formue skulle vise negativ korrelasjon over et tverrsnittsmateriale. Det kan meget vel tenkes at husholdningenes preferanser for konsum kontra sparing ikke er konstante over observasjonsmaterialet, men at "styrken" i preferansene i retning av sparing er positivt korrelert med formuens størrelse. Jean Crockett formulerer denne hypotesen slik:

"To a considerable degree, persons with high assets (particularly those with high assets in relation to income) are persons who *want* to have high assets because their taste for accumulating assets is relatively strong. Thus, by selecting a high-asset group, in effect we select "high savers". The downward pressure on consumption exerted by this trait tends to conceal the upward pressure which large asset holdings should exert for given savings preferences." (Crockett [12], p. 112.)

### 3.4. Konsumtilbøyelighetens avhengighet av inntektens artssammensetning og hovedinntektstagerens sosialgruppe: Analyseeskjema II

Som vi påpekte i avsnitt 2.3 (tabell 2), varierer inntektssammensetningen betydelig med hovedinntektstagerens sosialgruppe. Dette reflekterer blant annet at sosialgruppeinndelingen i inntekts- og forbruksundersøkelsen i betydelig grad bygger på inntektens artssammensetning. Imidlertid er det en kjennsgjerning at mange modellresonnementer regner med sosialgruppespesifikke forskjeller i forbruksstrukturen - med til dels forskjellig begrunnelse. Dette leder naturlig til spørsmålet om det empirisk kan påvises genuine forskjeller i konsumatferden mellom sosialgrupper, slik det for eksempel forutsettes i MODIS IV, eller om de forskjeller som tilsynelatende eksisterer, i sin helhet kan tilskrives forskjellene i inntektssammensetning. Vårt opplegg for multippel testing ligger godt til rette for å kaste lys over dette spørsmål.

I analyseeskjema II vil vi spesielt skille mellom følgende fire hypoteser: Inntektens marginale konsumtilbøyelighet varierer (A) både med inntektsart og sosialgruppe, (B1) bare med inntektsart, (B2) bare med sosialgruppe, (C) verken med inntektsart eller sosialgruppe. Hvis den siste hypotese aksepteres, kan det være spørsmål om konsumfunksjonens form overhodet ikke avhenger av sosialgruppe eller om det kan påvises nivåforskjeller i konstantleddet.

Vi velger følgende basismodell:

$$\begin{aligned}
 (3.2) \quad C &= \alpha_L Z_L + \alpha_{SP} Z_{SP} + \alpha_{SA} Z_{SA} + \alpha_{IY} Z_{IY} + \alpha_A A + \alpha_{AA} A^2 + \alpha_N N + \alpha_{NN} N^2 \\
 &+ \{ \beta_{L1} Z_L + \beta_{SP1} Z_{SP} + \beta_{SA1} Z_{SA} + \beta_{IY1} Z_{IY} \} R_1 \\
 &+ \{ \beta_{L2} Z_L + \beta_{SP2} Z_{SP} + \beta_{SA2} Z_{SA} + \beta_{IY2} Z_{IY} \} R_2 \\
 &+ \{ \beta_{L3} Z_L + \beta_{SP3} Z_{SP} + \beta_{SA3} Z_{SA} + \beta_{IY3} Z_{IY} \} (R_3 + R_5 + R_6) \\
 &+ \{ \beta_{L4} Z_L + \beta_{SP4} Z_{SP} + \beta_{SA4} Z_{SA} + \beta_{IY4} Z_{IY} \} (R_4 - S) \\
 &+ \delta t + \delta_1 Z_1 + \delta_2 Z_2 + \delta_3 Z_3 + u_{II},
 \end{aligned}$$

hvor  $Z_L$ ,  $Z_{SP}$ ,  $Z_{SA}$ ,  $Z_{IY}$  er binærvariable med verdi lik 1 når hovedpersonen er henholdsvis lønnstager, selvstendig næringsdrivende i primærnærings, selvstendig næringsdrivende i andre næringer, og ikke yrkesaktiv - og verdi 0 ellers. Her betegner f.eks.  $\beta_{L1}$  lønnsinntektens marginale konsumtilbøyelighet

for en lønnstagerhusholdning,  $\beta_{SP2}$  folketrygdtelsenes konsumtilbøyelighet for en husholdning hvor hovedpersonen er selvstendig i primærnæring etc. Inntektens artsinndeling er her noe mindre detaljert enn i analyseskjema I; næringsinntekt, boliginntekt og tomte- og aksjegevinster - hvor forskjellene i konsumtilbøyelighetene gjennomgående synes å være lite utsagnskraftige (jfr. tabell 6) - er slått sammen.

Som bakgrunnsvariable i tillegg til alder (A) og antall husholdningsmedlemmer (N) har vi i dette skjema spesielt valgt å inkludere t, nummeret på den 14-dagers-periode da husholdningen rapporterte sine forbruksutgifter ( $t=1,2,\dots,26$ ), og tre binærvariable  $Z_i$  ( $i=1,2,3$ ) med verdi lik 1 i i-te kvartal, 0 ellers. Binærvariablene er med for å oppfange eventuelle sesongsvingninger i forbruksutgiften, mens t ivaretar veksten i konsumet gjennom året som følge av inntektsendringer over året. Fordi husholdninger som ble observert i begynnelsen av året, gjennomgående har hatt lavere inntekt i registreringsperioden enn de som førte regnskap i slutten, vil de inntekter som registreres på kalenderårsbasis, kunne overvurdere den konsummotiverende "normalinntekt" for de første og undervurdere den for de siste. (Jfr. avsnitt 2.3 og appendiks til kapittel 2.)<sup>13)</sup>

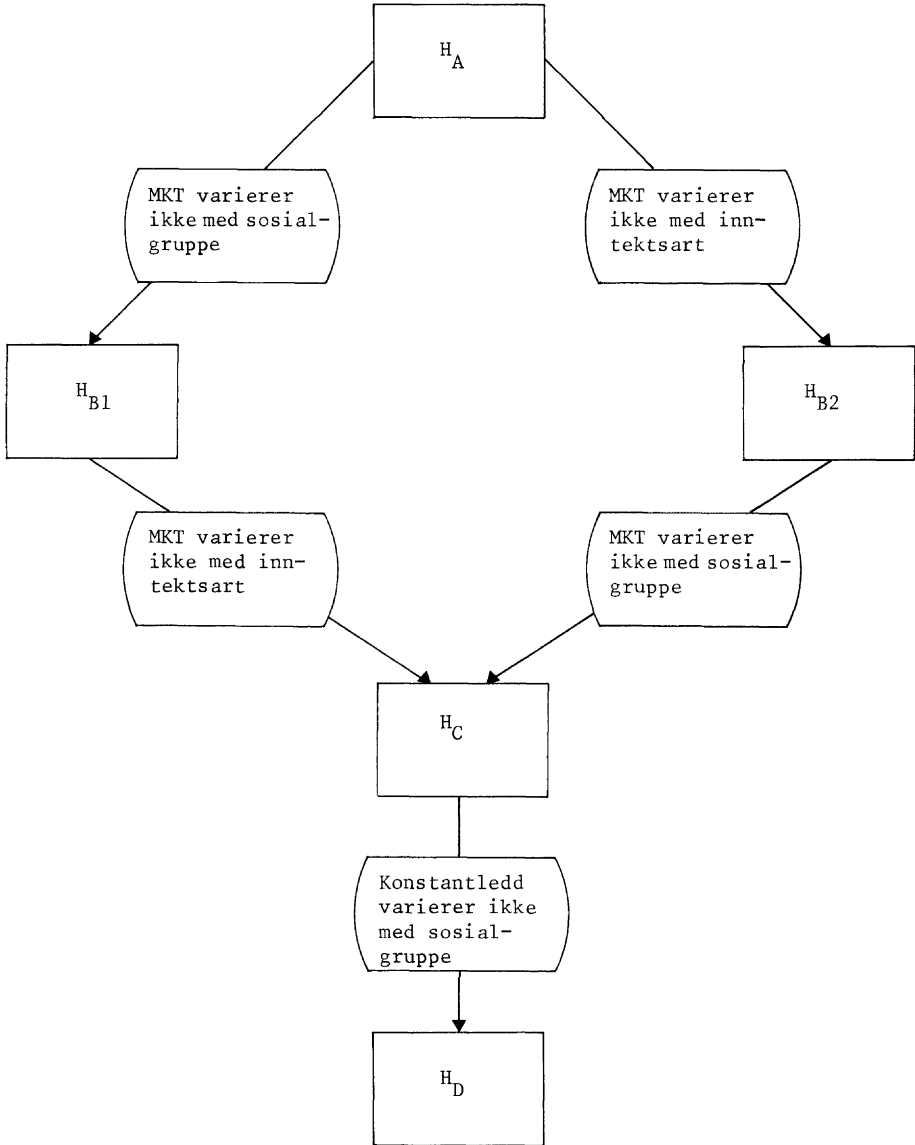
Tabell 7. Hypoteser i analyseskjema II<sup>a)</sup>

Symbol	Restriksjoner
$H_A$	Ingen
$H_{B1}$	$\beta_{L1} = \beta_{SP1} = \beta_{SA1} = \beta_{IY1}$ , $\beta_{L2} = \beta_{SP2} = \beta_{SA2} = \beta_{IY2}$ , $\beta_{L3} = \beta_{SP3} = \beta_{SA3} = \beta_{IY3}$ , $\beta_{L4} = \beta_{SP4} = \beta_{SA4} = \beta_{IY4}$
$H_{B2}$	$\beta_{L1} = \beta_{L2} = \beta_{L3} = \beta_{L4}$ , $\beta_{SP1} = \beta_{SP2} = \beta_{SP3} = \beta_{SP4}$ , $\beta_{SA1} = \beta_{SA2} = \beta_{SA3} = \beta_{SA4}$ , $\beta_{IY1} = \beta_{IY2} = \beta_{IY3} = \beta_{IY4}$
$H_C$	Alle $\beta$ -koeffisienter like
$H_D$	Alle $\beta$ -koeffisienter like, og $\alpha_L = \alpha_{SP} = \alpha_{SA} = \alpha_{IY}$

a) Basismodell: (3.2).

13) Det kan her selvsagt anføres at trend- og sesongkorreksjoner burde ha vært foretatt også i de tre øvrige analyseskjemaer. Dette har imidlertid neppe vesentlig betydning for tolkningen og sammenlignbarheten av resultatene. Som påpekt i avsnitt 2.3, er verken årsinntekten, formuen, alderen eller antall husholdningsmedlemmer nevneverdig korrelert med regnskapsperiodens nummer - fordelingen av husholdningene på regnskapsperiode er så å si "randomisert" med hensyn på disse variable. Det å utelate t,  $Z_1$ ,  $Z_2$  og  $Z_3$  fra regresjonsligningen vil derfor ikke i seg selv bringe skjevheter i estimatene. Se f.eks. Lovell [27], teorem 4.1.

Figur 2. Hypotesestrukturen i analyseskjema II.  
(MKT = marginal konsumtilbøyelighet.)



Hypotesehierarkiet er spesifisert i tabell 7. Sammenhengen mellom enkelthypotesene er vist i figur 2. Her uttrykker  $H_{B1}$  at det er artsspesifikke, men ikke sosialgruppespesifikke forskjeller i de marginale konsumtilbøyeligheter,  $H_{B2}$  kjennetegner den motsatte situasjon, mens  $H_C$  representerer tilfellet da konsumtilbøyeligheten er den samme uansett inntektsart og sosialgruppe. Slutthypotesen,  $H_D$ , pålegger den tilleggsrestriksjon at også konstantleddet har samme verdi for alle husholdninger.

Testresultatet - med nivå  $\alpha=1$  prosent for hver av deltestene, dvs. høyst 5 prosent nivå for totaltesten - er gitt i tabell 8. Tabell 9 inneholder de tilhørende verdier av F-observatorene. For både ikke-varige og varige goder er "løsningen"  $H_A$ , dvs. at konsumtilbøyeligheten både har en artsspesifikk og en sosialgruppespesifikk komponent. For begge tjenestegrupper viser konsumtilbøyeligheten signifikant variasjon med inntektsart, men ikke med sosialgruppe ( $H_{B1}$ ). Konsumfunksjonen for halv-varige goder derimot "slipper gjennom" hele hypotesenettet uten å bli forkastet, dvs. en spesifisering uten artsspesifikke og sosialgruppespesifikke elementer ( $H_D$ ) er statistisk sett "like god" som de mer generelle hypoteser.

Tabell 8. Analyseskjema II. Testresultater.  $\alpha = 0.01$

Avhengig variabel	Testresultat
C (alle goder)	$H_A$
$C_{IV}$ (ikke-varige goder)	$H_A$
$C_{HV}$ (halv-varige goder)	$H_D$
$C_V$ (varige goder)	$H_A$
$C_T$ (tjenester ekskl. bolig)	$H_{B1}$
$C_{BT}$ (boligtjenester)	$H_{B1}$

Tabell 9. Analyseskjema II. F-observatorer for deltestene

$H_0$	$H_1$	$K_0$ b)	$K_1$ b)	Avhengig variabel						$F_{0.99^c)}$
				C	$C_{IV}$	$C_{HV}$	$C_V$	$C_T$	$C_{BT}$	
$H_{B1}$	$H_A$	15	27	2.53	3.19	1.42	2.86	1.91	2.09	2.18
$H_{B2}$	$H_A$	15	27	3.91	3.84	1.79	3.15	3.53	2.40	2.18
$H_C$	$H_{B1}$	12	15	a)	a)	3.35	a)	7.76	6.95	3.78
$H_C$	$H_{B2}$	12	15	a)	a)	1.86	a)	a)	a)	3.78
$H_D$	$H_C$	9	12	a)	a)	2.85	a)	a)	a)	3.78

a) Ikke beregnet da den er uten interesse for gjennomføringen av testen.

b) Etter eliminasjon av  $Z_L$  ved hjelp av identiteten  $Z_L + Z_{SP} + Z_{SA} + Z_{IY} = 1$ .

c) 99 prosent-fraktilen i F-fordelingen med  $K_1 - K_0$  og  $n - K_1 - 1$  frihetsgrader.

Tabell 10. Analyteskjema II. Koeffisientestimater svarende til testresultatet

Koeffisient	Avhengig variabel					
	C	C <sub>IV</sub>	C <sub>HV</sub>	C <sub>V</sub>	C <sub>T</sub>	C <sub>BT</sub>
$\beta_{L1}$	0.5860	0.1498	0.1249	0.1107	0.1206	0.0565
$\beta_{SP1}$	0.5005	0.1978	0.1249	0.0543	0.1206	0.0565
$\beta_{SA1}$	0.5170	0.1150	0.1249	0.0569	0.1206	0.0565
$\beta_{IY1}$	0.5134	0.1318	0.1249	0.0959	0.1206	0.0565
$\beta_{L2}$	0.3917	0.1872	0.1249	0.0684	0.0444	0.0252
$\beta_{SP2}$	-0.0705	0.0038	0.1249	-0.0239	0.0444	0.0252
$\beta_{SA2}$	0.0548	0.0270	0.1249	0.0363	0.0444	0.0252
$\beta_{IY2}$	0.3454	0.0875	0.1249	0.0937	0.0444	0.0252
$\beta_{L3}$	0.4387	0.0662	0.1249	0.2044	0.0824	0.0060
$\beta_{SP3}$	0.3079	0.0380	0.1249	0.0936	0.0824	0.0060
$\beta_{SA3}$	0.4762	0.1239	0.1249	0.1318	0.0824	0.0060
$\beta_{IY3}$	0.2953	0.1354	0.1249	0.0528	0.0824	0.0060
$\beta_{L4}$	0.7251	0.2058	0.1249	0.1396	0.1227	-0.0027
$\beta_{SP4}$	-0.2070	-0.0980	0.1249	-0.1546	0.1227	-0.0027
$\beta_{SA4}$	0.6516	0.1446	0.1249	0.1827	0.1227	-0.0027
$\beta_{IY4}$	0.3528	0.1675	0.1249	0.0645	0.1227	-0.0027
$\alpha_L$	-15321.06	-11008.80	-8938.14	1465.58	-3543.32	6964.66
$\alpha_{SP}$	-11058.10	-9951.80	-8938.14	3310.74	-3226.16	6839.81
$\alpha_{SA}$	-9533.78	-9025.25	-8938.14	2790.74	-2789.31	7362.99
$\alpha_{IY}$	-12199.49	-11221.78	-8938.14	2472.74	-3247.51	7293.50
$\alpha_A$	260.52	301.28	158.32	-100.69	31.35	-103.17
$\alpha_{AA}$	-3.79	-3.09	-2.03	0.30	0.01	0.58
$\alpha_N$	7235.98	3454.46	1510.40	1388.71	510.71	284.33
$\alpha_{NN}$	-521.14	-176.31	-75.53	-166.96	-38.33	-62.37
$\delta$	824.22	340.18	328.21	117.34	76.76	-36.55
$\delta_1$	5427.21	2668.53	3242.15	324.75	1660.31	-2500.16
$\delta_2$	4684.64	2487.68	3523.03	-44.97	1356.53	-2549.30
$\delta_3$	685.17	959.60	295.12	-384.50	1172.36	-1363.58
$\hat{\sigma}_{uII}$	19711.27	6705.18	8152.45	8120.30	7648.47	6343.16

De koeffisientestimater som motsvarer testresultatet, er gitt i tabell 10. I de tilfelle vi kan etterspore sosialgruppespesifikke forskjeller i konsumtilbøyelighetene, er det lønnsinntektens konsumtilbøyelighet som viser minst variasjon med sosialgruppe. Størst er forskjellene for netto renteinntekter. Det er i første rekke selvstendige i primærnæringene som viser at atypisk mønster; i ligningen for totalkonsumet opptrer både trygdeinntekt ( $\beta_{SP2}$ ) og netto renteinntekt ( $\beta_{SP4}$ ) med negativ koeffisient. Det er ikke uventet at denne sosialgruppe skiller seg ut, da naturalhusholdning - med betydelig forbruk av varer fra egen produksjon (i datamaterialet vurdert til produsentpris) - er særlig utpreget i primærnæringene.<sup>14)</sup> Resultatene bekrefter ellers at konsumutgiften inneholder markante trend- og sesongkomponenter. For alle utgiftsgrupper bortsett fra tjenester eksklusive bolig tjenester forkastes hypotesen  $\delta = \delta_1 = \delta_2 = \delta_3 = 0$  ifølge en F-test med nivå 1 prosent.

### 3.5. Avvikelser fra linearitet. Samspill mellom inntekt, formue, alder og husholdningsstørrelse: Analyseeskjema III

Hittil er det forutsatt at de marginale konsumtilbøyeligheter for inntekt og formue ikke varierer med inntektens og formuens størrelse og heller ikke med alder og husholdningsstørrelse. Ved mindre endringer av de variable omkring observasjonsmaterialets "tyngdepunkt" kan nok dette være en brukbar hypotese, men ønsker vi konsumfunksjoner som dekker en større del av variasjonsområdet, kan streng linearitet være for restriktivt. Siden variablene viser til dels betydelig spredning i observasjonsmaterialet (tabell 1), bør vi derfor rette søkelyset mot mulige ikke-lineariteter i forbruksmønsteret.

For å holde analysen innenfor en rimelig ramme vil vi unnlate å splitte inntekten og formuen, dvs. vi lar bare totalverdiene  $R = \Sigma R_i - S$  og  $F = \Sigma F_i - G$  og transformasjoner av disse inngå som argumenter. Basismodellen er et fullstendig kvadratisk uttrykk i inntekt, formue, hovedpersonens alder og antall husholdningsmedlemmer:

14) For selvstendige bærer inntektsregnskapene ofte preg av at det er vanskelig å trekke en klar grense mellom næringsvirksomheten og eierens private husholdning. Av denne grunn burde vi kanskje ha utelatt selvstendige fra observasjonsmaterialet, noe som ville ha redusert antall observasjonsheter med ca. 15 prosent. Det som i første rekke har avholdt oss fra å gjøre dette, er at vi da ville svekke muligheten for å analysere konsumvirkningen av næringsinntekten vesentlig.



$$\begin{aligned}
 (3.3) \quad C = & a_o + a_{R^2} + a_{F^2} + a_{A^2} + a_{N^2} \\
 & + a_{RR}R^2 + a_{FF}F^2 + a_{AA}A^2 + a_{NN}N^2 \\
 & + a_{RF}RF + a_{RA}RA + a_{RN}RN \\
 & + a_{FA}FA + a_{FN}FN + a_{AN}AN + u_{III},
 \end{aligned}$$

som setter oss i stand til å studere både hvordan konsumfunksjonen er krummet i hver av de fire variable og hvilke typer av *samspill* (dvs. hvilke produkt-ledd) det er viktig å ta hensyn til. Spesielt kan den bidra til å kartlegge hvorvidt inntektens marginale konsumtilbøyelighet varierer med formue, alder og husholdningsstørrelse.

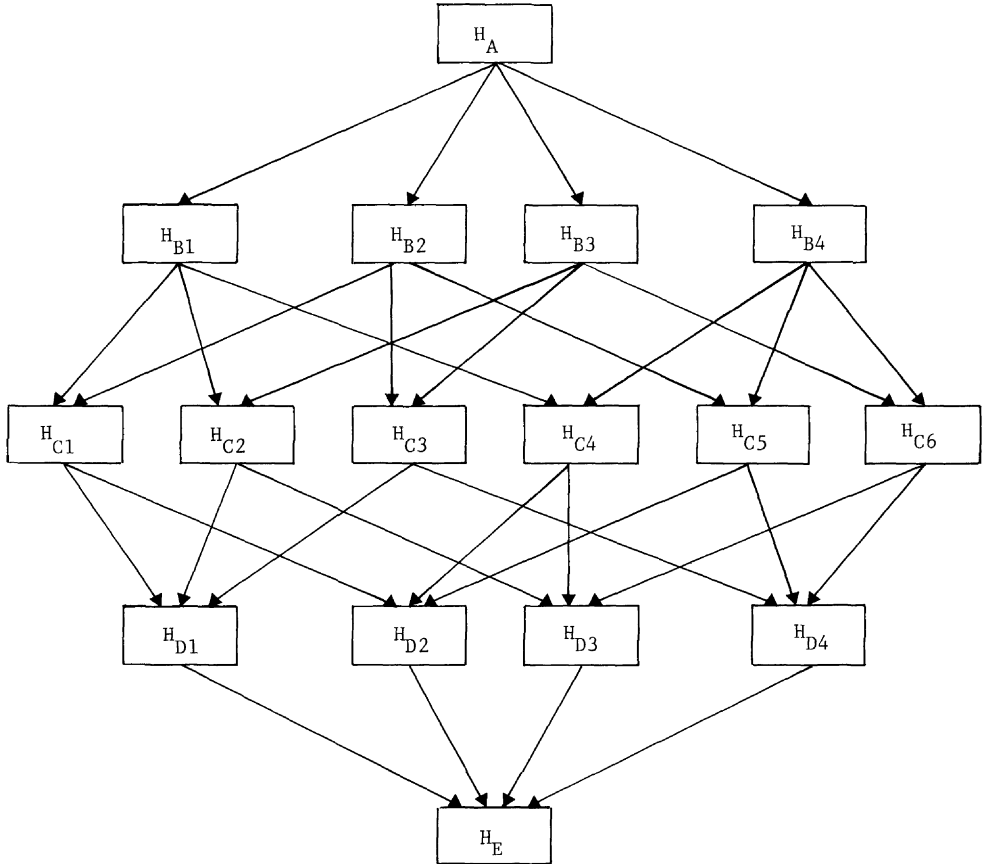
Hypotesestrukturen er spesifisert i tabell 11 og skissert i figur 3. Hierarkiet har her fem nivåer med i alt seksten hypoteser. Hypotesene på

Tabell 11. Hypoteser i analyseskjema III<sup>a)</sup>

Symbol	Restriksjoner
H <sub>A</sub>	Ingen
H <sub>B1</sub>	$a_{NN} = a_{RN} = a_{FN} = a_{AN} = 0$
H <sub>B2</sub>	$a_{AA} = a_{RA} = a_{FA} = a_{AN} = 0$
H <sub>B3</sub>	$a_{FF} = a_{RF} = a_{FA} = a_{FN} = 0$
H <sub>B4</sub>	$a_{RR} = a_{RF} = a_{RA} = a_{RN} = 0$
H <sub>C1</sub>	$a_{NN} = a_{AA} = a_{RA} = a_{RN} = a_{FA} = a_{FN} = a_{AN} = 0$
H <sub>C2</sub>	$a_{NN} = a_{FF} = a_{RF} = a_{RN} = a_{FA} = a_{FN} = a_{AN} = 0$
H <sub>C3</sub>	$a_{AA} = a_{FF} = a_{RF} = a_{RA} = a_{FA} = a_{FN} = a_{AN} = 0$
H <sub>C4</sub>	$a_{RR} = a_{NN} = a_{RF} = a_{RA} = a_{RN} = a_{FN} = a_{AN} = 0$
H <sub>C5</sub>	$a_{RR} = a_{AA} = a_{RF} = a_{RA} = a_{RN} = a_{FA} = a_{AN} = 0$
H <sub>C6</sub>	$a_{RR} = a_{FF} = a_{RF} = a_{RA} = a_{RN} = a_{FA} = a_{FN} = 0$
H <sub>D1</sub>	$a_{RR} \neq 0$ ; øvrige a-er med to fotskrifter lik null
H <sub>D2</sub>	$a_{FF} \neq 0$ ; øvrige a-er med to fotskrifter lik null
H <sub>D3</sub>	$a_{AA} \neq 0$ ; øvrige a-er med to fotskrifter lik null
H <sub>D4</sub>	$a_{NN} \neq 0$ ; øvrige a-er med to fotskrifter lik null
H <sub>E</sub>	Alle a-er med to fotskrifter lik null

a) Basismodell: (3.3)

Figur 3. Hypotesestrukturen i analyseskjema III.



B-nivået innebærer at alle annengrads- og samspillsledd for én av de variable bortfaller (for  $H_{B1}$  gjelder dette N, for  $H_{B2}$  gjelder det A osv.). C-hypotesene utelater annengrads- og samspillsleddene for to av de variable, mens D-hypotesene utelater slike ledd for tre av de variable. Slutthypotesen,  $H_E$ , innebærer at konsumutgiften er lineær i samtlige fire variable. Skjemaet inneholder i alt  $4+12+12+4=32$  deltester. Vi velger nivået  $\alpha=0.001$  for hver av dem, slik at totaltesten får et nivå på maksimalt 3.2 prosent.

Tabell 12 viser testresultatet og tabell 13 de tilhørende verdier av F-observatorene. For fire av de fem utgiftsgrupper og for totalkonsumet er testresultatet ikke entydig, dvs. vår strategi setter oss ikke i stand til å plukke ut én hypotese som "den beste". Vi møter her, som nevnt, en underliggende svakhet ved den statistiske metode vi har valgt. Resultatet kan imidlertid også sees som en indikasjon på at antall hypoteser og deltester er for stort. Vi har så å si stilt "for mange spørsmål" på én gang; datamaterialet kan ikke greie å skille dem klart fra hverandre.<sup>15)</sup>

Tabell 12. Analyseskjema III. Testresultater.  $\alpha = 0.001$

Avhengig variabel	Resultat
C (alle goder)	$H_{D1}$ eller $H_{C3}$
$C_{IV}$ (ikke-varige goder)	$H_{C2}$ eller $H_{B3}$
$C_{HV}$ (halv-varige goder)	$H_E$
$C_V$ (varige goder)	$H_E$ eller $H_{D4}$
$C_T$ (tjenester ekskl. bolig)	$H_{D1}$ eller $H_{C2}$
$C_{BT}$ (boligtjenester)	$H_{C2}$ eller $H_{C6}$

Det entydige testresultatet får vi for halv-varige goder, hvor konsumfunksjonen slipper gjennom hele hypotesenettet uten å bli forkastet, og vi ender med en funksjon som er lineær i alle variable. For de øvrige utgiftsgrupper får vi to "løsninger".

De tilhørende koeffisientestimater er gitt i tabell 14. Det er annengradsleddene i inntekt og husholdningsstørrelse som ser ut til å spille størst rolle. Der hvor slike ledd inngår, opptrer både kvadrert inntekt og kvadrert husholdningsstørrelse med negativ koeffisient og produktleddet med positiv. Tallverdien av koeffisienten til den kvadrerte inntekt er imidlertid liten, så konsumfunksjonens krumning "i inntektsretningen" er

15) Denne tendens ville selvfølgelig ha blitt forsterket om vi hadde operert med flere inntekts- og formuesvariable.

Tabell 13. Analyseeskjema III. F-observatorer for deltestene

$H_0$	$H_1$	$K_0$	$K_1$	Avhengig variabel						$F_{0.999}^{b)}$
				C	$C_{IV}$	$C_{HV}$	$C_V$	$C_T$	$C_{BT}$	
$H_{B1}$	$H_A$	10	14	4.26	4.50	1.78	3.47	0.34	3.85	4.62
$H_{B2}$	$H_A$	10	14	3.80	12.88	3.59	0.74	4.49	6.70	4.62
$H_{B3}$	$H_A$	10	14	2.24	1.88	0.84	2.06	0.47	2.09	4.62
$H_{B4}$	$H_A$	10	14	6.39	8.84	2.55	0.69	7.88	3.42	4.62
$H_{C1}$	$H_{B1}$	7	10	6.58	17.09	5.08	1.82	6.75	5.49	5.42
$H_{C2}$	$H_{B1}$	7	10	2.87	2.52	1.15	2.47	0.49	2.37	5.42
$H_{C4}$	$H_{B1}$	7	10	11.86	15.64	3.47	0.88	13.37	6.38	5.42
$H_{C1}$	$H_{B2}$	7	10	7.18	a)	2.65	5.47	1.21	a)	5.42
$H_{C3}$	$H_{B2}$	7	10	1.64	a)	0.86	1.85	0.97	a)	5.42
$H_{C5}$	$H_{B2}$	7	10	9.36	a)	3.67	1.14	7.91	a)	5.42
$H_{C2}$	$H_{B3}$	7	10	5.57	6.03	2.41	4.35	0.31	4.73	5.42
$H_{C3}$	$H_{B3}$	7	10	3.73	15.61	4.53	0.09	6.33	9.63	5.42
$H_{C6}$	$H_{B3}$	7	10	7.06	11.63	3.31	0.17	10.21	3.58	5.42
$H_{C4}$	$H_{B4}$	7	10	a)	a)	2.44	4.59	a)	6.95	5.42
$H_{C5}$	$H_{B4}$	7	10	a)	a)	5.07	1.21	a)	6.45	5.42
$H_{C6}$	$H_{B4}$	7	10	a)	a)	1.02	1.99	a)	1.81	5.42
$H_{D1}$	$H_{C1}$	5	7	a)	a)	0.05	1.01	1.38	a)	6.91
$H_{D2}$	$H_{C1}$	5	7	a)	a)	4.78	0.60	10.78	a)	6.91
$H_{D1}$	$H_{C2}$	5	7	6.09	23.32	5.94	0.04	10.78	9.49	6.91
$H_{D3}$	$H_{C2}$	5	7	15.09	21.80	4.21	0.07	20.06	8.77	6.91
$H_{D1}$	$H_{C3}$	5	7	8.86	a)	2.74	6.44	1.73	a)	6.91
$H_{D4}$	$H_{C3}$	5	7	12.11	a)	5.60	0.20	10.47	a)	6.91
$H_{D2}$	$H_{C4}$	5	7	a)	a)	7.21	2.02	a)	a)	6.91
$H_{D3}$	$H_{C4}$	5	7	a)	a)	0.73	2.46	a)	a)	6.91
$H_{D2}$	$H_{C5}$	5	7	a)	a)	3.26	7.09	a)	a)	6.91
$H_{D4}$	$H_{C5}$	5	7	a)	a)	1.38	1.26	a)	a)	6.91
$H_{D3}$	$H_{C6}$	5	7	a)	a)	2.86	6.36	a)	10.49	6.91
$H_{D4}$	$H_{C6}$	5	7	a)	a)	7.44	0.08	a)	9.44	6.91

a) Ikke beregnet da den er uten interesse for gjennomføringen av testen.

b) 99.9 prosent-fraktilen i F-fordelingen med  $K_1 - K_0$  og  $n - K_1 - 1$  frihetsgrader.

Tabell 13 (forts.). Analysekjema III. F-observatorer for deltestene

H <sub>0</sub>	H <sub>1</sub>	K <sub>0</sub>	K <sub>1</sub>	Avhengig variabel						F <sub>0.999</sub> <sup>b)</sup>
				C	C <sub>IV</sub>	C <sub>HV</sub>	C <sub>V</sub>	C <sub>T</sub>	C <sub>BT</sub>	
H <sub>E</sub>	H <sub>D1</sub>	4	5	32.49	a)	9.42	0.14	18.84	a)	10.83
H <sub>E</sub>	H <sub>D2</sub>	4	5	a)	a)	0.00	0.96	a)	a)	10.83
H <sub>E</sub>	H <sub>D3</sub>	4	5	a)	a)	12.96	0.07	a)	a)	10.83
H <sub>E</sub>	H <sub>D4</sub>	4	5	a)	a)	3.76	12.60	a)	a)	10.83

a) Ikke beregnet da den er uten interesse for gjennomføringen av testen.

b) 99.9 prosent-fraktilen i F-fordelingen med K<sub>1</sub>-K<sub>0</sub> og n-K<sub>1</sub>-1 frihetsgrader.

svak.<sup>16)</sup> Alder opptrer med negativt kvadratledd for ikke-varige goder og med et (svakt) positivt kvadratledd for tjenester. Vi kan derimot ikke påvise at annengradsledd i formue har betydning for noen av utgiftsgruppene. Alt i alt må vi kunne si at det er konsumfunksjonen for ikke-varige goder som viser sterkest tegn til ikke-linearitet.

Som en slags sammenfatning av resultatene i tabell 14 gjengir vi anslag for inntektens marginale konsumtilbøyelighet for hver av de fem utgiftsgruppene. For de grupper hvor den multiple testen ikke gir et éntydig resultat, baserer vi oss på den "løsningen" som er mest generell, dvs. inneholder flest utsagnskraftige høyresidevariable. Vi finner da

$$\frac{\partial C_{IV}}{\partial R} = 0.1745 - 0.00000128R + 0.00020A + 0.00825N,$$

$$\frac{\partial C_{HV}}{\partial R} = 0.1439,$$

$$\frac{\partial C_V}{\partial R} = 0.0950,$$

$$\frac{\partial C_T}{\partial R} = 0.0307 - 0.00000072R + 0.00248A,$$

$$\frac{\partial C_{BT}}{\partial R} = 0.1668 - 0.00000028R - 0.00186A.$$

16) Den konklusjon at konsumfunksjonens annenderiverte mhp. inntekten er negativ er i tråd med resultatet av en analyse foretatt av Ralph Husby [22] på basis av mikrodata fra USA.

Tabell 14. Analysekjema III. Koeffisientestimater svarende til testresultatet.  $\alpha = 0.001$

Koeffisient	Avhengig variabel										
	C		C <sub>IV</sub>		C <sub>HV</sub>	C <sub>V</sub>		C <sub>T</sub>		C <sub>BT</sub>	
	H <sub>D1</sub>	H <sub>C3</sub>	H <sub>C2</sub>	H <sub>B3</sub>		H <sub>E</sub>	H <sub>D4</sub>	H <sub>D1</sub>	H <sub>C2</sub>	H <sub>C2</sub>	H <sub>C6</sub>
$a_o$	11348.70	6455.57	-3442.81	-4078.18	1985.41	4870.10	2863.70	-1329.62	3542.77	1990.26	131.23
$10^3 \cdot a_R$	769.68	701.27	202.51	174.52	143.89	102.47	95.00	170.73	30.66	166.78	59.20
$10^3 \cdot a_F$	-0.74	-0.54	-2.37	-2.17	-1.27	2.41	2.45	1.94	1.22	-1.33	-1.76
$a_A$	-166.90	-148.11	292.09	283.62	-44.19	-75.87	-68.50	15.43	-82.68	-27.32	13.45
$a_N$	2889.61	6545.21	1920.98	2782.52	987.72	157.73	1397.21	80.00	94.04	-272.73	1683.11
$10^6 \cdot a_{RR}$	-1.68	-1.52	-0.64	-0.64	0	0	0	-0.48	-0.36	-0.14	0
$10^6 \cdot a_{FF}$	0	0	0	0	0	0	0	0	0	0	0
$a_{AA}$	0	0	-3.16	-3.09	0	0	0	0	0.20	0.35	0.06
$a_{NN}$	0	-504.03	0	-180.39	0	0	-159.30	0	0	0	-95.26
$10^6 \cdot a_{RF}$	0	0	0	0	0	0	0	0	0	0	0
$10^3 \cdot a_{RA}$	0	0	0.41	0.20	0	0	0	0	2.48	-1.86	0
$10^3 \cdot a_{RN}$	0	7.52	0	8.25	0	0	0	0	0	0	0
$10^3 \cdot a_{FA}$	0	0	0	0	0	0	0	0	0	0	0
$10^3 \cdot a_{FN}$	0	0	0	0	0	0	0	0	0	0	0
$a_{AN}$	0	0	0	4.93	0	0	0	0	0	0	26.75
$\hat{\sigma}_{uIII}$	20321.08	20272.46	6913.81	6897.87	8822.02	8203.86	8189.28	7650.14	7627.32	6402.83	6399.48

Inntektens konsumtilbøyelighet synes altså å være uavhengig av alderen for halv-varige og varige goder, stigende med alderen for ikke-varige goder og tjenester utenom boligjenester og synkende for boligjenester. En økning av antall husholdningsmedlemmer fører til en (svak) heving av konsumtilbøyeligheten for ikke-varige goder, men ser ikke ut til å ha betydning for resten av konsumet.

### 3.6. Konsumfunksjonenes homogenitetsegenskaper analysert ved en modifisert "translog"-funksjon: Analysekjema IV

Konsumfunksjoner utformes undertiden slik at konsumutgift pr. person avhenger av inntekt (og evt. formue) pr. person. Matematisk sett er da funksjonene homogene av grad én i antall personer og inntekt (og evt. formue). Det påstås imidlertid ofte at det er "stordriftsfordeler" (eller "-ulemper") knyttet til konsumaktiviteten og at konsumfunksjonen derfor ikke vil ha en så enkel form. Hensikten med analysekjema IV er å undersøke dette spørsmålet med vår sterkt aggregerte utgiftsgruppeinndeling.

Det er i dette tilfelle hensiktsmessig å la de variable opptre på logaritmisk form. Vi har valgt en basismodell hvor (den naturlige) logaritmen til konsumutgiften er lineær i logaritmene til formue og alder og kvadratisk i logaritmene til inntekt og antall husholdningsmedlemmer:

$$(3.4) \quad \ln C = b_o + b_R \ln R + b_N \ln N + b_F \ln F + b_A \ln A \\ + b_{RR} (\ln R)^2 + b_{RN} \ln R \ln N + b_{NN} (\ln N)^2 + u_{IV}.$$

Funksjonen er altså en såkalt "translog"-funksjon i R og N. Av hensiktsmessighetsgrunner omformer vi den til<sup>17)</sup>

$$(3.4^*) \quad \ln C - \ln N = b_o + b_R (\ln R - \ln N) + (b_R + b_N - 1) \ln N + b_F \ln F \\ + b_A \ln A + b_{RR} (\ln R - \ln N)^2 \\ + (b_{NN} - b_{RR}) (\ln N)^2 + (b_{RN} + 2b_{RR}) \ln R \ln N + u_{IV}.$$

Vi har valgt å spesifisere annengradsledd bare for  $\ln R$  og  $\ln N$  fordi det er disse variable som er sentrale for spørsmålet om konsumfunksjonens homogenitetsegenskap; F og A har her karakter av bakgrunnsvariable.

Med (3.4) som basismodell er det spesifisert i alt ti hypoteser, fordelt på fire nivåer. Hypotesehierarkiet er illustrert i figur 4. Innholdet i de ti hypotesene er stikkordmessig følgende (med "homogen" menes "homogen av grad én"):

17) Det er velkjent at lineærtransformasjoner av denne type ikke påvirker minste kvadratets estimatorene.

- $H_A$  : Ingen restriksjoner.  
 $H_{B1}$  : Formue uten betydning.  
 $H_{B2}$  : Annengradsledd (i R og N) uten betydning.  
 $H_{B3}$  : Homogen i R, N og F.  
 $H_{B4}$  : Homogen i R og N.  
 $H_{C1}$  : Verken annengradsledd (i R og N) eller formue har betydning.  
 $H_{C2}$  : Homogen i R og N. Formue uten betydning.  
 $H_{C3}$  : Homogen i R, N og F. Annengradsledd uten betydning.  
 $H_{C4}$  : Homogen i R og N. Annengradsledd uten betydning.  
 $H_D$  : Homogen i R og N. Verken annengradsledd eller formue har betydning.

Testskjemaet inneholder i dette tilfelle 17 deltester. Som nivå for hver deltest velger vi alternativt  $\alpha = 0.01$  og  $\alpha = 0.001$  og oppnår da et totalt testnivå på (maksimalt) henholdsvis 17 og 1.7 prosent.

F-observatorene for deltestene er gitt i tabell 15.<sup>18)</sup> Utfallet av totaltesten blir det samme for samtlige utgiftsgrupper, nemlig at hypotese  $H_{B1}$  aksepteres. Dette er den eneste hypotese på B-nivået som ikke forkastes mot basishypotesen  $H_A$ , mens den forkastes mot både  $H_{C1}$  og  $H_{C2}$ . Resultatet er det samme enten vi velger 1 prosent eller 0.1 prosent som nivå for deltestene. Konklusjonen blir dermed at vi for samtlige utgiftsgrupper må avvise både at konsumfunksjonen er homogen av grad én i inntekt og antall personer og at homogenitetsgraden er konstant (men forskjellig fra én). Ikke for noen av gruppene gir formuen signifikant bidrag.

De tilhørende koeffisientestimater er gitt i tabell 16. De innebærer blant annet at elastisiteten av konsumutgiften med hensyn på inntekten,

$$\frac{\partial \ln C}{\partial \ln R} = b_R + 2b_{RR} \ln R + b_{RN} \ln N,$$

er en stigende funksjon av inntekten ( $b_{RR} > 0$ ) og en avtagende funksjon av antall husholdningsmedlemmer ( $b_{RN} < 0$ ) for samtlige utgiftsgrupper. Den tilsvarende elastisitet ved en proporsjonal endring av inntekt og antall husholdningsmedlemmer er gitt ved:<sup>19)</sup>

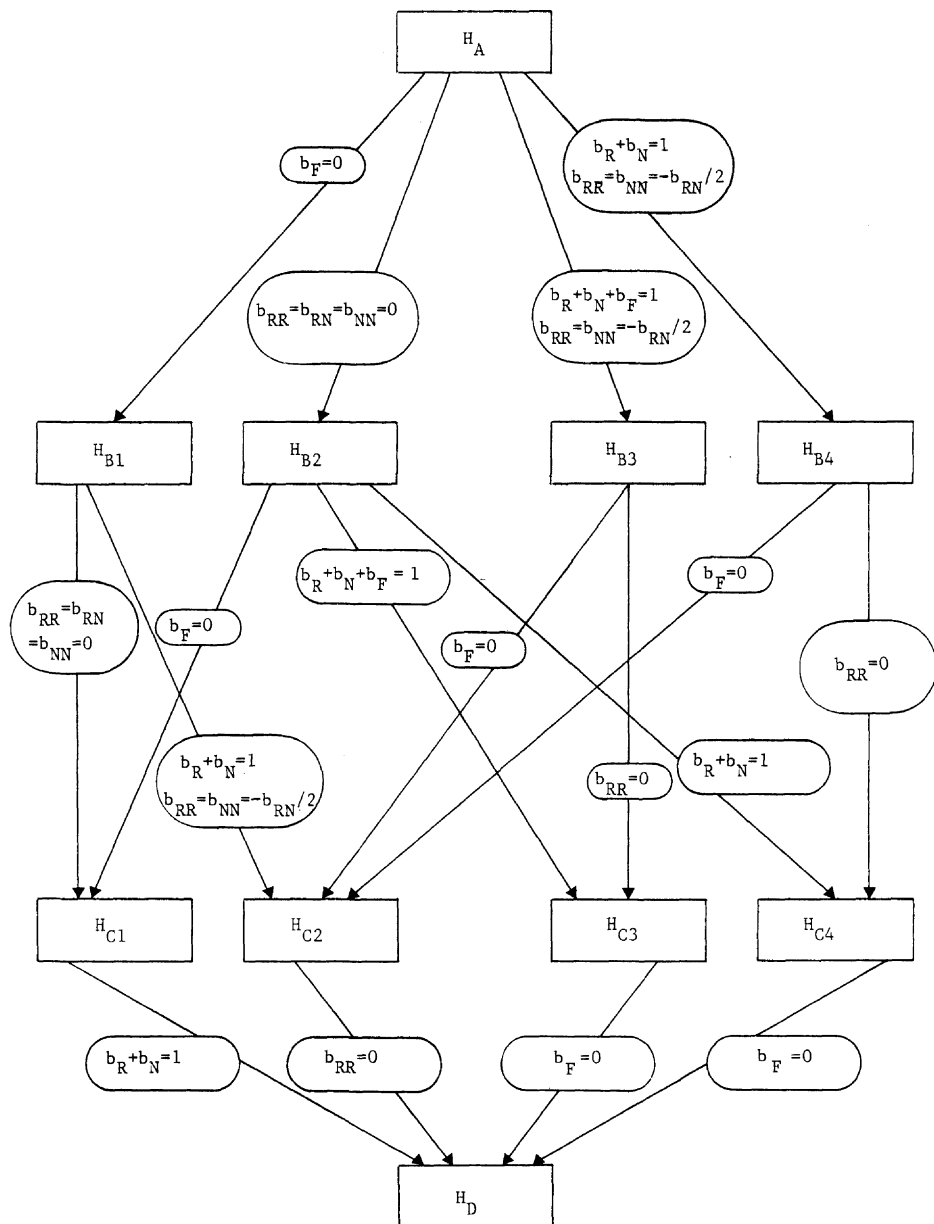
$$\frac{\partial \ln C}{\partial \ln R} + \frac{\partial \ln C}{\partial \ln N} = b_R + b_N + (2b_{RR} + b_{RN}) \ln R + (2b_{NN} + b_{RN}) \ln N.$$

18) Noen få av husholdningene i materialet figurerer med null i inntekt eller formue. For å kunne transformere til logaritmer valgte vi i slike tilfelle å sette inntekten, resp. formuen, lik 1 krone, dvs. erstatte  $\ln R$  og  $\ln F$  med  $\ln [\max(R, 1)]$  og  $\ln [\max(F, 1)]$ .

19) For enkelhets skyld regner vi som om N kan variere kontinuerlig.



Figur 4. Hypotesestrukturen i analyseskjema IV.



Estimater for denne elastisiteten for utvalgte inntektsnivåer og med N satt lik sitt observasjonsgjennomsnitt er gitt i tabell 17. Anslagene ligger under 1 for ikke-varige goder og boligjenester og over 1 for de tre øvrige godene. Variasjonen med inntektsnivået er forholdsvis moderat.

Tabell 15. Analyseeskjema IV. F-observatorer for deltestene

H <sub>0</sub>	H <sub>1</sub>	K <sub>0</sub>	K <sub>1</sub>	C	C <sub>IV</sub>	C <sub>HV</sub>	C <sub>V</sub>	C <sub>T</sub>	C <sub>BT</sub>	F <sub>0.99</sub>	F <sub>0.999</sub>
H <sub>B1</sub>	H <sub>A</sub>	6	7	0.34	0.84	0.11	2.63	0.05	0.70	6.63	10.83
H <sub>B2</sub>	H <sub>A</sub>	4	7	175.32	93.35	45.38	20.12	50.17	36.19	3.78	5.42
H <sub>B3</sub>	H <sub>A</sub>	4	7	26.64	31.90	55.90	32.54	15.92	6.42	3.78	5.42
H <sub>B4</sub>	H <sub>A</sub>	4	7	26.64	30.68	53.36	30.32	15.38	5.99	3.78	5.42
H <sub>C1</sub>	H <sub>B1</sub>	3	6	163.56	93.38	46.29	21.28	50.58	36.00	3.78	5.42
H <sub>C2</sub>	H <sub>B1</sub>	3	6	11.78	32.75	57.48	34.68	16.09	6.80	3.78	5.42
H <sub>C1</sub>	H <sub>B2</sub>	3	4	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>C3</sub>	H <sub>B2</sub>	3	4	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>C4</sub>	H <sub>B2</sub>	3	4	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>C2</sub>	H <sub>B3</sub>	3	4	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>C3</sub>	H <sub>B3</sub>	3	4	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>C2</sub>	H <sub>B4</sub>	3	4	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>C4</sub>	H <sub>B4</sub>	3	4	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>D</sub>	H <sub>C1</sub>	2	3	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>D</sub>	H <sub>C2</sub>	2	3	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>D</sub>	H <sub>C3</sub>	2	3	a)	a)	a)	a)	a)	a)	6.63	10.83
H <sub>D</sub>	H <sub>C4</sub>	2	3	a)	a)	a)	a)	a)	a)	6.63	10.83

a) Observatoren er ikke beregnet, da den er uten interesse for gjennomføringen av testen.

Tabell 16. Analysekjema IV. Koeffisientestimater svarende til testresultatet

Koeffisient	Avhengig variabel					
	C	C <sub>IV</sub>	C <sub>HV</sub>	C <sub>V</sub>	C <sub>T</sub>	C <sub>BT</sub>
b <sub>o</sub>	10.4869	8.4343	8.6580	9.7013	8.3331	8.4059
b <sub>R</sub>	-0.5207	-0.2719	-0.8953	-0.8569	-1.2997	-0.6766
b <sub>R</sub> +b <sub>N</sub> <sup>-1</sup>	-0.2051	0.0522	1.4489	2.7260	-0.1523	0.7853
b <sub>F</sub>	0	0	0	0	0	0
b <sub>A</sub>	-0.2973	-0.0539	-0.8577	-1.4540	-0.4597	-1.1306
b <sub>RR</sub>	0.0554	0.0326	0.1046	0.0963	0.1319	0.0976
b <sub>NN</sub> <sup>-b<sub>RR</sub></sup>	-0.1385	-0.1351	-0.2043	-0.5835	-0.1152	-0.1004
b <sub>RN</sub> <sup>+2b<sub>RR</sub></sup>	0.0340	0.0053	-0.0245	-0.0683	0.0784	-0.0802
$\hat{\sigma}_{uIV}$	0.4944	0.4101	1.7295	2.8252	2.0352	1.7904

Tabell 17. Estimater for konsumelastisiteter ved proporsjonal endring av inntekt og antall husholdningsmedlemmer

Inntektsnivå, kr	Utgiftsgruppe				
	C <sub>IV</sub>	C <sub>HV</sub>	C <sub>V</sub>	C <sub>T</sub>	C <sub>BT</sub>
20 000	0.805	1.716	1.652	1.452	0.673
40 000	0.809	1.699	1.604	1.507	0.617
60 000	0.811	1.689	1.577	1.538	0.585
80 000	0.812	1.682	1.557	1.561	0.562

### 3.7. Oppsummering

Estimerings- og testresultatene ovenfor gir på ingen måte noe enhetlig bilde. Vi har stilt mange spørsmål om konsumfunksjonens form og fått nesten like mange svar. Det er åpenbart at en stor del av konsumvariasjonene i datamaterialet ikke oppfanges av de modeller vi har betraktet - estimatet for restleddsstandardavviket er betydelig både sett i forhold til standardavviket og til gjennomsnittet for konsumet. Den individuelle komponent i konsum- og spareatferden er med andre ord betydelig (jfr. Haavelmo [21], pp. 146-147). Skal vi - med fare for å overforenkle - forsøke å oppsummere de sentrale konklusjoner i noen få punkter, kan vi kanskje gjøre det på følgende måte:

1. Renteinntekter og -utgifter virker symmetrisk på konsumentterspørselen. Derimot ser det ikke ut til å være full symmetri mellom virkningen av bankinnskudd og gjeld. Her er det gruppen ikke-varige goder som skiller seg ut.
2. Testresultatene bekrefter *ikke* den antagelse som ofte gjøres at lønns- og trygdeinntekt har høyere marginal konsumtilbøyelighet enn resten av inntekten. Dette gjelder både når vi ser på totalutgiften samlet og betrakter hver av de fem utgiftsgruppene for seg.
3. Resultatene viser klarere tegn på at inntektens marginale konsumtilbøyelighet varierer med inntektsart enn at den varierer med sosialgruppe. Det er derfor grunn til å anta at de forskjeller i konsumtilbøyeligheter mellom sosialgrupper som ofte registreres, for en stor del er en refleks av forskjellene i inntektssammensetning. For de fem utgiftsgruppene finner vi følgende interessante mønster: Konsumtilbøyeligheten for halv-varige goder kan ikke påvises å variere verken med inntektsart eller med sosialgruppe. For tjenestegruppene finner vi signifikant variasjon med inntektsart, men ikke med sosialgruppe. Bare for ikke-varige og varige goder synes konsumtilbøyeligheten å ha både en artsspesifikk og en sosialgruppespesifikk komponent. I de tilfelle da en sosialgruppespesifikk komponent kan etterspores, er det særlig selvstendige i primærnæringene som skiller seg ut.
4. Vår analyse gir ikke grunnlag for entydige svar på om konsumfunksjonens form avviker fra den lineære. Dette skyldes bl.a. at spørsmålet ikke kan analyseres uten å la mange konkurrerende hypoteser være i bildet på én gang, og vår metode vil da ha en tendens til å gi flertydige svar. For halv-varige goder finner vi ingen tegn til avvikelse fra linearitet i noen retning. For ikke-varige goder og for tjenester peker resultatene i retning av at inntektens marginale konsumtilbøyelighet avtar med inntektsnivået og stiger med alderen.
5. Vi finner ingen støtte for å anta at konsumfunksjonen er homogen av grad 1 i inntekt og antall husholdningsmedlemmer. Resultatene tyder heller ikke på at homogenitetsgraden er konstant, men forskjellig fra 1.
6. Formuen, slik den registreres ved skatteliggingen, gir som helhet lite selvstendig bidrag til å forklare variasjonene i konsumutgiften. Gjennomgående synes den å ha mindre "forklaringskraft" enn både inntekt, alder og antall personer. Riktignok kan vi i denne henseende spore en viss forskjell mellom de enkelte formuesarter, men siden verdien av realformue trolig er systematisk underregistrert i datamaterialet, er resultatene vanskelige å tolke. En grundigere analyse, hvor skatteliggingens formuesoppgaver suppleres med den informasjon om beholdningen av konsumkapital som ble innhentet i forbindelse med Forbruksundersøkelsen 1973, men som det ikke er blitt anledning til å innarbeide i vårt datamateriale, vil kunne kaste et

klarere lys over spørsmålet.<sup>20)</sup>

7. Skal vi peke på ett trekk som fremfor noe preger resultatene, må det bli at punkttestimatene for inntektens marginale konsumtilbøyelighet gjennomgående er uventet lave. For totalkonsumet ligger de stort sett mellom 0.3 og 0.7, som er markert lavere enn dem man vanligvis oppnår ved estimering på grunnlag av aggregerte tidsseriedata (se f.eks. Biørn [4]). Noen klare årsaker til dette er det vanskelig å peke på. En del av svaret kan ligge i at våre variable er annerledes definert enn dem det vanligvis er aktuelt å bruke i makroanalyser, men dette kan neppe forklare en betydelig del av nivåforskjellen. Simultanitetsproblemet påpekt i fotnote 12 ovenfor er et annet moment, likeledes det forhold at registreringen av forbruksutgifter og inntekter ikke faller sammen i tid. Muligheten for at vi har feilspesifisert modellstrukturen og dermed valgt en inadekvat estimeringsmetode er altså i høy grad til stede. I kapittel 4 vil vi vende oppmerksomheten mot en type av feilspesifikasjon som kan ha gjort seg gjeldende.

#### 4. ESTIMERING AV KONSUMFUNKSJONER I MODELLER MED UOBSERVERBARE (LATENTE) VARIABLE

##### 4.1. Utgangspunktet

At individualdata for inntekt, formue og forbruk inneholder *målefeil*, er et faktum som erkjennes av alle som arbeider med å fremskaffe slike data. Diskusjonen i kapittel 2 viser at det til dels er betydelig avstand mellom det vi kunne ønske å måle og det vi er i stand til å registrere. Nå er det velkjent at målefeil i de variable skaper spesielle problemer for økonometriske analyse av samvariasjon. Blant annet vil de estimerer minste kvadraters regresjon gir, vanligvis inneholde systematiske skjevheter. De er *inkonsistente* i statistisk forstand. Vi har tidligere konstatert at regresjonsberegningene i kapittel 3 gjennomgående gav uventet lave estimerer for inntektens marginale konsumtilbøyelighet.

Det har derfor interesse å analysere konsumfunksjoner innenfor modeller som er mer generelle enn den enkle regresjonsmodellen ved at de tar hensyn til at noen av de variable kan være feilaktig observert. I utgangspunktet kunne vi tenke oss (minst) tre forklaringer på at punkttestimatene for inntektens marginale konsumtilbøyelighet i kapittel 3 stort sett var påfallende lave: (i) konsumet er systematisk underregistrert, (ii) inntekten er

20) Anskaffelsesverdien av beholdningen av konsumkapital (utenom boliger) er anslått til ca. 18 000 kroner for gjennomsnittshusholdningen i 1973. (Skarstad [37], p. 11.)

systematisk overregistrert, eller (iii) inntekten inneholder tilfeldige målefeil.<sup>1)</sup> De to første hypotesene er i vårt tilfelle lite rimelige. Vi må snarere regne med at inntekten systematisk undervurderes i skattemålingsmaterialet og at forbruksundersøkelsens utgiftstall trolig har en tendens til å overvurdere verdien av konsumet totalt sett.<sup>2)</sup> Den tredje hypotese kan derimot ikke avvises. Noen husholdninger kan "tilfeldigvis" ha hatt store inntektsposter eller små fradragposter i observasjonsåret, for andre kan det motsatte ha vært tilfelle - i gjennomsnitt jevner disse forskjeller seg mer eller mindre ut.

1) Formelt kan dette vises på følgende måte: Betrakt konsumfunksjonen

$$C^* = a_0 + aR^*,$$

hvor  $C^*$  og  $R^*$  er de "sanne" verdier av konsum og inntekt. Disse er imidlertid uobserverbare; det vi observerer, er

$$Y = \alpha_0 + \alpha C^* + u,$$

$$X = \beta_0 + \beta R^* + v,$$

hvor  $\alpha_0$ ,  $\alpha$ ,  $\beta_0$ ,  $\beta$  er ukjente og hvor vi antar at  $u$ ,  $v$  og  $R^*$  er innbyrdes ukorrelerte. Hva skjer hvis vi forsøker å estimere  $a$  ved å ta regresjonen av  $Y$  mhp.  $X$ ? Vi bruker da som "estimator" ( $M_{PQ}$  betegner generelt den empiriske kovarians mellom  $P$  og  $Q$ ):

$$\hat{a} = \frac{M_{YX}}{M_{XX}} = \frac{\alpha\beta M_{R^*R^*} + \alpha M_{C^*v} + \beta M_{R^*u} + M_{uv}}{\beta^2 M_{R^*R^*} + 2\beta M_{R^*v} + M_{vv}} = \frac{\alpha\beta M_{R^*R^*} + \alpha M_{R^*v} + \beta M_{R^*u} + M_{uv}}{\beta^2 M_{R^*R^*} + 2\beta M_{R^*v} + M_{vv}}$$

Denne konvergerer stokastisk mot

$$\text{plim } \hat{a} = \frac{\alpha\beta\sigma_{R^*R^*}}{\beta^2\sigma_{R^*R^*} + \sigma_{vv}} = \frac{\frac{\alpha}{\beta} a}{1 + \sigma_{vv}/(\beta^2\sigma_{R^*R^*})},$$

hvor  $\sigma_{R^*R^*}$  og  $\sigma_{vv}$  er sannsynlighetsgrensene for henholdsvis  $M_{R^*R^*}$  og  $M_{vv}$ . Følgelig er  $\text{plim } \hat{a} < a$  hvis  $\alpha/\beta < 1$  og/eller  $\sigma_{vv} > 0$ . Uten ytterligere restriksjoner på modellen er  $a$  ikke identifiserbar, så noen konsistent estimator eksisterer ikke.

2) Denne tendens motvirkes noe ved at inntekten i observasjonsmaterialet i gjennomsnitt er datert senere enn forbruksutgiften. Se appendiks til kapittel 2.

#### 4.2. En variant av Friedman's konsumteori. Estimering ved hjelp av instrumentvariable

La oss starte med å betrakte følgende variant av Milton Friedman's konsummodell:<sup>3)</sup>

$$(4.1) \begin{cases} C = a_0 + aR^{**} + \epsilon, \\ R = R^{**} + \delta, \end{cases}$$

hvor C og R er de observerte verdier av henholdsvis konsum og inntekt, mens  $R^{**}$ , som er uobserverbar, betegner den konsummotiverende normalinntekt (den permanente inntekt). Restleddsvariablene  $\epsilon$  og  $\delta$  har forventning null, er innbyrdes ukorrelerte og ukorrelert med  $R^{**}$ , og alle tre har endelige varianser.  $\epsilon$  og  $\delta$  betegnes gjerne som henholdsvis transitorisk (tilfeldig) konsum og transitorisk inntekt.

Konsistent estimering av  $a_0$  og  $a$  vil i dette tilfellet kunne oppnås ved bruk av *instrumentvariable*, dvs. variable som er ukorrelert med feilleddene  $\epsilon$  og  $\delta$ , men som samtidig er korrelert med normalinntekten  $R^{**}$  og følgelig også med den observerte inntekt R.<sup>4)</sup> I lys av resultatene i kapittel 3 fremtrer alder (A), antall husholdningsmedlemmer (N) og formue (F) som potensielle instrumenter da de alle er korrelert med inntekten og dermed tilfredsstillende det andre av de to kravene. Instrumentvariableestimatene for  $a$  svarende til disse tre valg av instrumenter er henholdsvis

$$a_A^{**} = \frac{M_{CA}}{M_{RA}} = 1.82,$$

$$a_N^{**} = \frac{M_{CN}}{M_{RN}} = 1.22,$$

$$a_F^{**} = \frac{M_{CF}}{M_{RF}} = 0.41,$$

hvor generelt  $M_{PQ}$  betegner den empiriske kovarians mellom P og Q. Som anslag for inntektens langsiktige marginale konsumtilbøyelighet er ingen av disse akseptable. Siden verken A, N eller F er sterkt korrelert med

3) Friedman [18], kapittel 3. Vanligvis settes konstantleddet  $a_0$  lik null.

4) Hvis dette er egenskaper ved populasjonen og ikke bare gjelder som sampelegenskaper, vil "instrumentvariabelmetoden" gi konsistente estimater. I modell (4.1), med de angitte restleddsforutsetninger og normal fordeling av de variable, er  $a$  ikke identifiserbar (og kan følgelig heller ikke estimeres konsistent) med mindre det foreligger ekstern tilleggsinformasjon. Men den opplysning at det "utenfor" den formaliserte modell eksisterer instrumenter med de nevnte korrelasjonsegenskaper, er tilstrekkelig til å gjøre  $a$  identifiserbar.

inntekten, vil variansene til estimatene være store.<sup>5)</sup> Konsistensegenskapen har derfor liten interesse i dette tilfelle, og det er ingen grunn til å foretrekke disse estimatene fremfor minste kvadraters estimatet  $\hat{a} = M_{CR}/M_{RR} = 0.73$ , selv om det er inkonsistent (jfr. fotnote 1).

Spørsmålet er da om vi ved å velge en passende lineær kombinasjon av A, N og F og eventuelt andre variable som instrument kan forbedre estimatene. Fra dette er ikke skrittet langt til å utvide modellen ved å supplere konsumfunksjonen med en ligning som beskriver hvordan  $R^*$  bestemmes ved disse tre variable.<sup>6)</sup> Modellen får da form av et simultant lignings-system med latente variable.

#### 4.3. En generell lineær strukturmodell med latente variable<sup>7)</sup>

Et lineært system av strukturligninger med latente (uobserverbare) variable kan generelt skrives på formen

$$(4.2) \quad BY^* = \Gamma X^* + u,$$

hvor B og  $\Gamma$  er matriser av strukturkoeffisienter av dimensjon henholdsvis  $(m \times m)$  og  $(m \times n)$ , og  $Y^*$ ,  $X^*$  og  $u$  er stokastiske vektorer av dimensjon henholdsvis  $(m \times 1)$ ,  $(n \times 1)$  og  $(m \times 1)$ .  $Y^*$  er modellens endogene variable,  $X^*$  er de eksogene variable, mens  $u$  er feil i ligningene (disturbances, shocks). Noen av eller alle elementene i  $X^*$  og  $Y^*$  er uobserverbare. Dette er de latente, strukturelle variable. Det vi observerer, er

$$(4.3) \quad Y = Y^* + \varepsilon,$$

$$(4.4) \quad X = X^* + \delta,$$

5) Korrelasjonskoeffisientene i samplet er henholdsvis  $R_{AR} = -0.20$ ,  $R_{NR} = 0.49$ ,  $R_{FR} = 0.20$ . Under ellers like forhold er estimatets varians omvendt proporsjonal med kvadratet av korrelasjonskoeffisienten. Se f.eks. Malinvaud [28], p. 350.

6) Jfr. Zellner [47], hvor "instrumentvariabeltankegangen" anvendt på en énrelasjonsmodell er generalisert til estimering av et simultant lignings-system. En oversikt over de senere års litteratur om simultane modeller med uobserverbare variable finnes i Griliches [20].

7) Dette avsnitt bygger i betydelig grad på Jöreskog [23], avsnittene 1-3.



hvor  $\epsilon$  og  $\delta$  er vektorer av målefeil (errors).<sup>8)</sup> Her har  $Y$  og  $\epsilon$  dimensjon  $(m \times 1)$  og  $X$  og  $\delta$  dimensjon  $(n \times 1)$ .

Vi forutsetter at alle feilledd har forventning null:

$$(4.5) \quad E(u) = E(\epsilon) = E(\delta) = 0,$$

og at  $\epsilon$ ,  $\delta$  og  $u$  er innbyrdes ukorrelerte og ukorrelerte med  $X^*$ . Av dette følger at  $\epsilon$  og  $\delta$  også vil være ukorrelert med  $Y^*$  - alle målefeil er altså ukorrelert med de tilsvarende sanne verdier. I det følgende forutsetter vi at  $X$  og  $Y$  er målt fra sine respektive gjennomsnitt og at alle verdier av  $X^*$  er generert ved samme sannsynlighetsfordeling. Dermed blir

$$(4.6) \quad \begin{cases} E(Y) = E(Y^*) = 0, \\ E(X) = E(X^*) = 0. \end{cases}$$

La dessuten

$$(4.7) \quad E(X^* X^{*'}) = \Phi, \quad (\text{dimensjon } (n \times n))$$

$$(4.8) \quad E(uu') = \psi, \quad (\text{dimensjon } (m \times m))$$

$$(4.9) \quad E(\epsilon\epsilon') = \theta_\epsilon^2, \quad (\text{dimensjon } (m \times m))$$

$$(4.10) \quad E(\delta\delta') = \theta_\delta^2, \quad (\text{dimensjon } (n \times n))$$

hvor  $\theta_\epsilon$  og  $\theta_\delta$  er diagonalmatriser. Vi forutsetter altså at det ikke er korrelasjon mellom målefeilene for de enkelte variable. Men vi tillater korrelasjon mellom feilene i de enkelte strukturlikninger i modellen - dvs.  $\psi$  er ikke nødvendigvis diagonal.

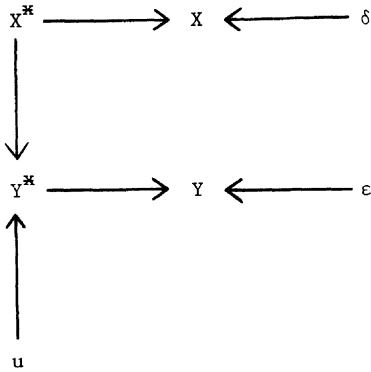
Lar vi piler symbolisere årsaks-virkningsforhold, kan strukturen i modellen beskrives som i figur 5. Pilene peker fra de variable som bestemmes utenfor modellen, mot de variable som bestemmes i modellen.

Formelt sett er ikke bare  $Y^*$  og  $Y$ , men også  $X$  endogene.<sup>9)</sup>

8) I tilfelle noen av elementene i  $Y^*$  og  $X^*$  er observerbare uten feil, er naturligvis de tilhørende elementer i  $\epsilon$  og  $\delta$  lik null. Modellen er av den type som i den økonometriske litteratur går under betegnelsen "strukturell modell" (structural model) - til forskjell fra en "funksjonell modell" (functional model) hvor den latente verdi av den eksogene variable  $X^*$  er ikke-stokastisk. Se f.eks. Malinvaud [28], p. 330.

9) Om kausalitetsbegrepet i økonometriske modeller, se f.eks. Simon [36].

Figur 5. Kausalitetsstrukturen i en generell modell med latente variable



La  $Z$  betegne vektoren av de observerte variable i modellen:

$$(4.11) \quad Z = \begin{pmatrix} Y \\ X \end{pmatrix} \quad (\text{dimensjon } ((m+n) \times 1)).$$

Av forutsetningene ovenfor følger - etter noe regning - at varians-kovariansmatrisen til  $Z$  vil være gitt ved<sup>10)</sup>

$$(4.12) \quad E(Z Z') = \Sigma = \begin{bmatrix} B^{-1} \Gamma \Phi \Gamma' B'^{-1} + B^{-1} \Psi B'^{-1} + \theta_{\epsilon}^2 & B^{-1} \Gamma \Phi \\ \Phi \Gamma' B'^{-1} & \Phi + \theta_{\delta}^2 \end{bmatrix}.$$

Elementene i  $\Sigma$  er altså funksjoner av elementene i  $B$ ,  $\Gamma$ ,  $\Phi$ ,  $\Psi$ ,  $\theta_{\epsilon}$  og  $\theta_{\delta}$ , som inneholder alle strukturkoeffisientene i modellen. I praksis vil en del av disse koeffisientene ha kjente verdier, f.eks. 0 eller 1, de resterende er ukjente og skal estimeres ved hjelp av det datamateriale som står til disposisjon.

10) Jfr. Jöreskog [23], ligning (4). Vi forutsetter at  $B$  er ikke-singulær.

For å avgjøre om konsistent estimering overhodet er mulig bør vi i prinsippet undersøke modellens *identifikasjonsegenskaper*. I denne type av modeller er det vanlig å forutsette at alle eksogene variable (inklusive restleddene og feilleddene) er *multinormalt fordelt*, selv om det ofte kan være urealistisk for de strukturelle variable  $X^*$ 's vedkommende.<sup>11)</sup> Da vil også  $X$  og  $Y$  - og dermed  $Z$  - være multinormalt fordelt. Normalfordelinger karakteriseres éntydig ved sine første- og annenordensmomenter. Siden  $X$ ,  $Y$  og  $Z$  er transformert slik at de har forventningsverdier lik null, betyr det at all informasjon vi vil kunne skaffe oss om fordelingen til  $Z$  fra et observasjonsmateriale, vil være i form av informasjon om elementene i matrisen  $\Sigma$ . Vi kan følgelig ikke få vite noe mer om de ukjente parametre i  $B$ ,  $\Gamma$ ,  $\Phi$ ,  $\Psi$ ,  $\theta_\epsilon$  og  $\theta_\delta$  enn det vi kan avlede av ligning (4.12) når  $\Sigma$  er kjent. Dette gjelder uansett hvor mange observasjoner vi har. Spørsmålet om identifikasjon reduserer seg derfor til spørsmålet om vi ut fra kjennskap til  $\Sigma$  kan "regne oss tilbake" til de ukjente parametre i disse 6 matrisene på en éntydig måte. Er det slik at det til hver gitt  $\Sigma$  svarer én og bare én "struktur", eller finnes det flere "strukturer" som gir samme verdi av  $\Sigma$ ?

Å besvare dette spørsmålet generelt er vanskelig. Bare visse spesialtilfelle er behandlet i litteraturen.<sup>12)</sup> En *nødvendig* betingelse for at alle parametre i modellen er identifiserbare er imidlertid lett å utlede. Vi har at antall parametre i  $\Sigma$ , når vi tar hensyn til at den er symmetrisk, er  $(m + n)(m + n + 1)/2$ . Dette blir følgelig også antall ligninger i systemet (4.12) og gir dermed en øvre grense for antall ukjente parametre i matrisene  $B$ ,  $\Gamma$ ,  $\Phi$ ,  $\Psi$ ,  $\theta_\epsilon$ ,  $\theta_\delta$  som kan bestemmes når  $\Sigma$  er fullstendig kjent. La  $t$  betegne antall elementer i disse 6 matriser som er ukjente a priori (når vi tar hensyn til at  $\Phi$  og  $\Psi$  er symmetriske). En nødvendig betingelse for at hele modellen er identifiserbar blir dermed

$$(4.13) \quad t \leq (m + n)(m + n + 1)/2.$$

Hvis modellen i sin helhet er identifiserbar, kan den estimeres ved *Full Information Maximum Likelihood-metoden* (sannsynlighetsmaksimeringsmetoden). Estimatenes er konsistente og (regulært) asymptotisk effisiente. Metoden kan i korthet beskrives på følgende måte:

11) Som kjent viser fordelingene av formue, inntekt og alder oftest en utpreget skjevhet (asymmetri). De kan derfor neppe approksimeres særlig godt ved en symmetrisk fordeling som den normale. Dette forhindrer selvsagt ikke at den *betingede* fordeling av  $Y^*$  (og  $Y$  og  $X$ ) gitt  $X^*$  med god tilnærming kan antas å være normal. Jfr. skillet mellom det "funksjonelle" og det "strukturelle" tilfelle.

12) Se f.eks. Reiersøl [34] og Geraci [19].

La  $N$  betegne antall observasjonssett og  $Z_i$  verdien av vektoren  $Z$  for  $i$ -te observasjonssett ( $i=1,2,\dots,N$ ). Vi forutsetter at observasjonene er uavhengige. Logaritmen til den simultane tetthetsfunksjon for  $Z_1, Z_2, \dots, Z_N$  - også kalt log-likelihood-funksjonen - er da gitt ved

$$(4.14) \quad \ln L = -\frac{1}{2}(m+n)N \ln(2\pi) - \frac{1}{2}N \ln |\Sigma| - \frac{1}{2} \sum_{i=1}^N Z_i' \Sigma^{-1} Z_i,$$

idet vi tar hensyn til at alle  $Z_i$  har forventning null. Den kvadratiske form i dette uttrykket kan skrives som<sup>13)</sup>

$$\sum_{i=1}^N Z_i' \Sigma^{-1} Z_i = \text{tr}(\Sigma^{-1}Q) = N \text{tr}(\Sigma^{-1}S),$$

hvor  $Q = \sum_{i=1}^N Z_i Z_i'$  og  $S = Q/N$ . Maksimering av log-likelihood-funksjonen

(4.14) er følgelig ekvivalent med å minimere funksjonen

$$(4.15) \quad F = \ln |\Sigma| + \text{tr}(\Sigma^{-1}S)$$

med hensyn på  $\Sigma$  under hensyntagen til (4.12) og de restriksjoner som er pålagt strukturkoeffisientene i modellen. Problemet blir altså å minimere  $F$  med hensyn på de ukjente elementer i matrisene  $B, \Gamma, \Phi, \psi, \theta_\epsilon$  og  $\theta_\delta$ . Den svenske statistiker Karl G. Jöreskog har utviklet et datamaskinprogram, LISREL, som utfører slike beregninger numerisk. Alle beregninger i dette kapittel er gjennomført ved hjelp av dette programmet.<sup>14)</sup>

#### 4.4. Estimeringsresultater for tre enkle modeller

Vi har dannet tre enkle modeller som alle har form av et simultant ligningssystem med latente variable og hvor én av ligningene er en konsumfunksjon. De kan oppfattes som generaliseringer av den enkle Friedman-modellen i avsnitt 4.2.

13) Her symboliserer "tr" trasen, dvs. summen av elementene langs matrisens hoveddiagonal. Se f.eks. Anderson [1], pp. 44-46.

14) LISREL-programmet kan behandle modeller som er noe mer generelle enn (4.2)-(4.10). Den nyeste versjon av programmet er beskrevet i Jöreskog og Sörbom [24]. Ved estimeringen ble det benyttet en tidligere programversjon, som ikke skiller seg vesentlig fra denne.

#### 4.4.1. Modell A

Den første modellen har formen

$$(4.16) \quad C^{\mathbf{x}} = aR^{\mathbf{x}} + bF^{\mathbf{x}},$$

$$(4.17) \quad R^{\mathbf{x}} = eF^{\mathbf{x}} + fA + gN + u_2,$$

$$(4.18) \quad C = C^{\mathbf{x}} + \varepsilon_C,$$

$$(4.19) \quad R = R^{\mathbf{x}} + \varepsilon_R,$$

$$(4.20) \quad F = F^{\mathbf{x}} + \delta_F,$$

hvor  $C$ ,  $R$ ,  $F$ ,  $N$  og  $A$  er de observerte verdier av henholdsvis total konsumtgift, inntekt, formue, antall husholdningsmedlemmer og hovedpersonens alder, målt som avvik fra sine respektive gjennomsnitt. Variable forsynt med stjerne er de tilsvarende latente variable. Vi gjør den realistiske antagelse at  $A$  og  $N$  er observert uten feil. Konsum og inntekt er modellens endogene variable; deres målefeil er henholdsvis  $\varepsilon_C$  og  $\varepsilon_R$ . Målefeilen i formuen er  $\delta_F$ .<sup>15)</sup>

Inntektslikningen (4.17) kan begrunnes på flere måter. Enklest er det vel å tenke seg at formuesleddet oppfanger den del av inntekten som er avkastning av real- og finansformue, mens aldersvariabelen på en summarisk måte representerer det kompleks av variable som påvirker inntektsopptjeningsvevnen og arbeidskapasiteten til husholdningens hovedperson, f.eks. hans utdanning, yrkeserfaring, utholdenhet mv. (Hadde vi hatt observasjoner av de sistnevnte variable, burde vi selvsagt ha trukket dem inn i inntektslikningen ved siden av alderen.) Antall husholdningsmedlemmer er selvsagt også med på å bestemme husholdningens inntektsopptjeningspotensial. Et annet forhold som kan bringe korrelasjon mellom  $R^{\mathbf{x}}$  og  $N$ , er at inntektsopptjeningsvevnen kan være bestemmende for hvor mange personer husholdningen vil ha anledning til å forsørge. Er dette momentet det dominerende, vil kausaliteten i modellen gå fra  $R^{\mathbf{x}}$  til  $N$  istedenfor omvendt.

---

15) Restleddet i konsumfunksjonen er satt lik null. Siden vi har åpnet muligheten for målefeil i konsumet, er dette ingen effektiv restriksjon. Vi kan ikke observasjonsmessig skjelve en modell med målefeil i konsumet fra en modell med restledd i konsumfunksjonen i dette tilfelle.

Formelt tilsvarende modellen følgende spesialtilfelle av den generelle modellen i avsnitt 4.3:

$$B = \begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} b & 0 & 0 \\ e & f & g \end{bmatrix}$$

$$Y^* = \begin{pmatrix} C^* \\ R^* \end{pmatrix}, \quad Y = \begin{pmatrix} C \\ R \end{pmatrix}, \quad X^* = \begin{pmatrix} F^* \\ A \\ N \end{pmatrix}, \quad X = \begin{pmatrix} F \\ A \\ N \end{pmatrix}$$

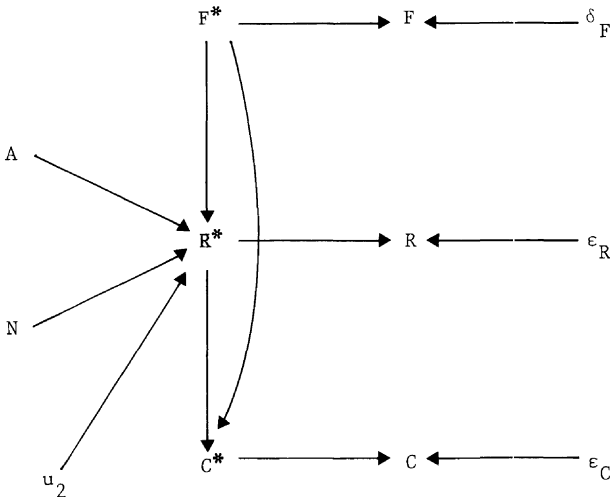
$$u = \begin{pmatrix} 0 \\ u_2 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_C \\ \varepsilon_R \end{pmatrix}, \quad \delta = \begin{pmatrix} \delta_F \\ 0 \\ 0 \end{pmatrix}$$

$$m=2, n=3.$$

Den har 5 ukjente strukturcoeffisienter (matrisene B og  $\Gamma$ ), 6 varianser og kovarianser til (latente) eksogene variable (matrisen  $\Phi$ ), én restleddsvarians (matrisen  $\psi$ ), 2 varianser til målefeil i endogene variable (matrisen  $\theta_\varepsilon^2$ ) og én varians til målefeil i eksogene variable (matrisen  $\theta_\delta^2$ ) - til sammen 15 ukjente strukturparametre. Siden  $(m+n)(m+n+1)/2 = 15$ , er altså den nødvendige betingelse for identifiserbarhet, (4.13), oppfylt med likhet.<sup>16)</sup>

Kausalitetsstrukturen i modellen kan anskueliggjøres som i figur 6.

Figur 6. Kausalitetsstrukturen i modell A.



16) Et resonnement som sannsynliggjør at også de tilstrekkelige betingelser for identifikasjon er tilfredsstillet selv om vi ikke pålegger tilleggsrestriksjoner, er gjengitt i appendiks til dette kapitlet.

I det spesialtilfellet da både inntekten og formuen er observert uten målefeil ( $\epsilon_R = \delta_F = 0$ ), er modellen *rekursiv*. Kausaliteten er da, noe forenklet uttrykt, rettet fra F, A og N, via R til C.<sup>17)</sup>

Maximum Likelihood-estimer for fem forskjellige varianter av denne modellen er gitt i tabell 18. I de to første, A1 og A2, forutsettes det at verken inntekten eller formuen er observert med målefeil, A3 og A4 åpner for målefeil i inntekten, men ikke i formuen, mens A5 tillater målefeil i begge.

Resultatet er slående. Når vi ikke regner med målefeil i inntekten og formuen, får vi et estimat for inntektens marginale konsumtilbøyelighet  $a$  på 0.73 hvis inntekten er eneste argument i konsumfunksjonen (variant A1) og 0.74 hvis også formuen inkluderes (variant A2). Begge estimater er signifikant mindre enn 1 - selv med et signifikansnivå på 0.1 prosent. Regner vi med målefeil i inntekten, øker estimatene til henholdsvis 1.17 (variant A3) og 1.27 (variant A4), som begge er signifikant større enn 1. For  $e$ , som kan tolkes som formuens marginale avkastningsrate (renten), får vi et lignende resultat. Estimatet er 3.8 prosent når vi ser bort fra målefeil i formuen (variant A4) og stiger til hele 17.8 prosent når vi åpner for målefeil (variant A5).

---

17) En modell som ligner modell A på vesentlige punkter, er estimert av Attfield [3] på grunnlag av data fra en engelsk spareundersøkelse fra 1953. Attfield har formue, alder og husholdningsstørrelse med i inntektslikningen, men formuen inngår ikke i konsumfunksjonen. Dessuten ser han bort fra målefeil i formuen. Til gjengjeld lar han den transitoriske inntekt  $\epsilon_R$  opptre som egen konsummotiverende variabel ved siden av den permanente inntekt.

Tabell 18. Maximum Likelihood (FIML) - estimater for modell A: a) b) c)

$$C^{\#} = aR^{\#} + bF^{\#}$$

$$R^{\#} = eF^{\#} + fA + gN + u_2$$

$$C = C^{\#} + \epsilon_C$$

$$R = R^{\#} + \epsilon_R$$

$$F = F^{\#} + \delta_F$$

Variant	$\hat{a}$	$\hat{b}$	$\hat{e}$	$\hat{f}$	$\hat{g}$
A1 <sup>d)</sup>	0.7258 (0.0187)	-	0.029 (0.002)	-0.089 (0.019)	5.806 (0.203)
A2 <sup>d)</sup>	0.7384 (0.0191)	-0.011 (0.003)	0.029 (0.002)	-0.089 (0.019)	5.806 (0.203)
A3	1.168 (0.039)	-	0.029 (0.002)	-0.089 (0.019)	5.806 (0.203)
A4	1.266 (0.043)	-0.031 (0.004)	0.038 (0.003)	-0.092 (0.018)	5.587 (0.202)
A5	1.237 (0.047)	-0.125 (0.020)	0.178 (0.038)	-0.258 (0.053)	4.905 (0.311)

Variant	$\hat{\sigma}_{u_2}^2$	$\hat{\sigma}_{\epsilon_C}$	$\hat{\sigma}_{\epsilon_R}$	$\hat{\sigma}_{\delta_F}$	Antall ukjente struktur-koef.	Antall ukjente restleddsvarianser	Antall varianser til feilledd	Antall 2. ordensmomenter	Totalt antall ukjente parametre
A1	139.081	21.287	-	-	4	1	1	6	12
A2	139.081	21.253	-	-	5	1	1	6	13
A3	139.081 (7.860)	18.069 (0.348)	12.413 (0.323)	-	4	1	2	6	13
A4	129.931 (7.324)	17.436 (0.371)	12.737 (0.303)	-	5	1	2	6	14
A5	107.244 (12.096)	18.139 (0.411)	11.096 (0.528)	97.054 (3.055)	5	1	3	6	15

a) C, R og F er målt i 1 000 kr.

b) Parametre hvis verdi a priori er satt lik null, er markert med tankestrek. Estimater for asymptotiske standardavvik er gjengitt i parentes.

c)  $\sigma_{u_2}^2 = \text{var } u_2$ .

$\sigma_{\epsilon_i}$  = standardavvik til  $\epsilon_i$  ( $i = C, R, F$ ).  $\sigma_{\delta_F}$  = standardavvik til  $\delta_F$ .

d) Estimeringen er gjennomført ved å bruke minste kvadraters metode på hver av strukturligningene. Siden modellen er rekursiv, faller disse estimater sammen med FIML-estimaten.



Formuens marginale konsumtilbøyelighet,  $b$ , har et negativt estimat uansett hva vi forutsetter om målefeil i inntekten og formuen. Nå representerer  $b$  bare den *partielle* virkning av formuen. Tar vi hensyn til at formuesendringer også påvirker konsumet via inntekten, viser det seg at *totaleffekten* er positiv. Av (4.16) og (4.17) finner vi at ligningen for konsumutgiften i modellens reduserte form er gitt ved

$$C^* = (b+ae)F^* + afA + agN + au_2.$$

De tilhørende koeffisientestimater er følgende:

<u>Variant</u>	<u><math>\hat{b} + \hat{ae}</math></u>	<u><math>\hat{af}</math></u>	<u><math>\hat{ag}</math></u>
A1	0.021	-0.065	4.214
A2	0.010	-0.066	4.287
A3	0.034	-0.104	6.779
A4	0.017	-0.117	7.070
A5	0.095	-0.320	6.067

Det er her interessant at det er i tilfellet med målefeil i formuen, altså A5, at vi oppnår det klart høyeste estimat på formueskoeffisienten  $b + ae$  (9.5 prosent) og samtidig det laveste estimat på alderskoeffisienten  $af$  (-0.320, som innebærer at konsumutgiften reduseres med 320 kroner for hvert år alderen øker). Dette har nok sammenheng med at formue og alder er positivt korrelerte variable (korrelasjonskoeffisienten i samplet er 0.18), slik at en "oppjustering" av formueskoeffisienten vil ha en tendens til å ledsages av en "nedjustering" av alderskoeffisienten.

Variansen til målefeilen utgjør en betydelig større del av totalvariansen for formue enn for inntekt. I variant A5 er de estimerte andeler henholdsvis 31 prosent og 76 prosent. At formuen skulle være mer unøyaktig målt enn inntekten virker plausibelt, men en "målestøy" på 76 prosent er vel i overkant av det en med rimelighet kan akseptere.

#### 4.4.2. Modell B

Formuestallene i datamaterialet er, som nevnt i avsnitt 2.2.3, basert på registreringer ved utgangen av 1973. De vil dermed være påvirket av sparing i løpet av året. I prinsippet har vi at

$$\text{Formue ved slutten av 1973} = \text{Inntekt i 1973} - \text{Konsum i 1973} + \text{Formue ved begynnelsen av 1973}.$$

I modell A tok vi ikke hensyn til dette og betraktet formuen som eksogen.<sup>18)</sup>

Vi vil nå revidere modellen slik at den latente verdi av formuen opptrer som en lineær funksjon av de latente verdier av inntekten og konsumutgiften i håp om å få tatt hensyn til "økosirkeffekter" av denne typen. Samtidig utelater vi formuen som argument i konsumfunksjonen. Etter disse endringene får modellen formen

$$(4.21) \quad C^* = aR^*,$$

$$(4.22) \quad R^* = eF^* + fA + gN + u_2,$$

$$(4.23) \quad F^* = hR^* + iC^* + u_3,$$

$$(4.24) \quad C = C^* + \epsilon_C,$$

$$(4.25) \quad R = R^* + \epsilon_R,$$

$$(4.26) \quad F = F^* + \epsilon_F.$$

Vi antar at  $u_2$  og  $u_3$  er ukorrelerte. Restleddet i formuesligningen skal i prinsippet oppfange effekten av initialformuen. Vi må derfor a priori regne med at det har stor varians.

Modellen fremtrer som følgende spesialtilfelle av den generelle modellen i avsnitt 4.3:

$$B = \begin{bmatrix} 1 & -a & 0 \\ 0 & 1 & -e \\ -i & -h & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 & 0 \\ f & g \\ 0 & 0 \end{bmatrix}$$

$$Y^* = \begin{pmatrix} C^* \\ R^* \\ F^* \end{pmatrix}, \quad Y = \begin{pmatrix} C \\ R \\ F \end{pmatrix}, \quad X^* = X = \begin{pmatrix} A \\ N \end{pmatrix}$$

$$u = \begin{pmatrix} 0 \\ u_2 \\ u_3 \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_C \\ \epsilon_R \\ \epsilon_F \end{pmatrix}, \quad \delta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

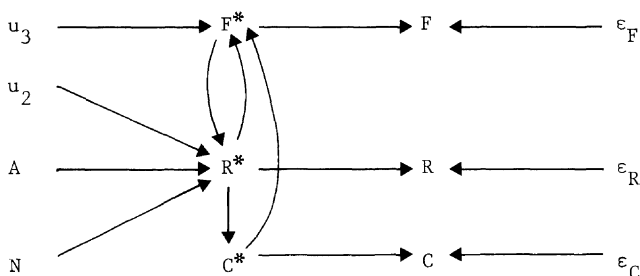
$$m = 3, \quad n = 2.$$

18) Det samme gjorde vi i regresjonsmodellene i kapittel 3. Det kan gis et visst forsvar for å neglisjere denne sammenhengen i vårt tilfelle. For det første gjør målefeilene i konsumet pga. den korte registreringsperioden at null kan være et bedre anslag for formuesøkningen i 1973 enn differensen mellom registrert inntekt og konsumutgift. For det annet vil den nettoformue som registreres ved skatteligningen, vanligvis bare oppfange en brøkdel av husholdningenes sparing (i real- og finansobjekter) i løpet av året. Jfr. også fotnote 12 i kapittel 3.

Den har 6 ukjente strukturcoeffisienter, 3 varianser og kovarianser til eksogene variable, 2 restleddsvarianser og 3 varianser til målefeil i endogene variable - til sammen 14 ukjente strukturparametre. Siden dette antallet er mindre enn  $(m+n)(m+n+1)/2 = 15$ , er den nødvendige betingelse for identifiserbarhet oppfylt.

Kausalitetsstrukturen i denne modellen kan anskueliggjøres som i figur 7.

Figur 7. Kausalitetsstrukturen i modell B.



I dette tilfelle er det en to-veis sammenheng mellom alle de endogene variable. Modellen er *interdependent*. Dette gjelder uansett om de variable er observert med feil eller ikke.

Estimeringsresultater er gitt i tabell 19. I det første tilfelle, B1, er det ikke pålagt restriksjoner utover de som er nevnt ovenfor. I det andre, B2, forlanger vi at koeffisientene foran inntekt og konsumutgift i formuleslikningen (4.23) skal være lik henholdsvis 1 og -1. Dette ville være et opplagt krav dersom ligningen var ment som en streng definisjonslikning hvor  $u_3$  er verdien av initialformuen, men hvis vi må regne med betydelige "nivåfeil" i minst én av de tre variable, er det ikke uten videre rimelig.

I begge tilfelle er estimatet for  $a$  bare utbetydelig høyere enn det tilsvarende estimat i modell A, som er 1.168.<sup>19)</sup> "Endogeniseringen" av formuen har altså liten innvirkning på anslaget for inntektens marginale konsumtilbøyelighet. Derimot påvirkes estimatene for koeffisientene i inntektsgenereringslikningen (4.22) (= (4.17)) vesentlig. Spesielt gjelder dette i tilfellet hvor det ikke er lagt restriksjoner på koeffisientene i formuleslikningen. Estimater for  $e$ , formuens marginalavkastningsrate, er da nesten 20 prosent.

19) Dessverre var det ikke mulig å beregne standardavvikestimater med LISREL-programmet i dette tilfelle.

Tabell 19. Maximum Likelihood (FIML)- estimator for modell B: a) b) c)

$$C^* = aR^*$$

$$R^* = eF^* + fA + gN + u_2$$

$$F^* = hR^* + iC^* + u_3$$

$$C = C^* + \epsilon_C$$

$$R = R^* + \epsilon_R$$

$$F = F^* + \epsilon_F$$

Variant	$\hat{a}$	$\hat{e}$	$\hat{f}$	$\hat{g}$	$\hat{h}$	$\hat{i}$
B1	1.176	0.198	-0.030	4.933	0.927	-0.024
B2	1.174	0.078	-0.095	5.814	1 <sup>d)</sup>	-1 <sup>d)</sup>

Vari- ant	$\hat{\sigma}_{u2}^2$	$\hat{\sigma}_{u3}^2$	$\hat{\sigma}_{\epsilon C}$	$\hat{\sigma}_{\epsilon R}$	$\hat{\sigma}_{\epsilon F}$	Antall ukjen- te struk- tur- koef.	Antall ukjen- te rest- ledds- vari- anser	Antall vari- anser til feil- ledd	Antall 2. or- dens momen- ter, para- eks. var.	Totalt antall ukjen- te para- metre t
B1	99.746	100.147	17.863	12.412	109.45	6	2	3	3	14
B2	122.564	4999.791	18.034	12.450	86.633	4	2	3	3	12

a) C, R og F er målt i 1 000 kr.

b) Asymptotiske standardavvikestimater ikke beregnet.

c)  $\hat{\sigma}_{u_i}^2 = \text{var } u_i$  ( $i=2,3$ ).

$\hat{\sigma}_{\epsilon_i}$  = standardavvik til  $\epsilon_i$  ( $i=C, R, F$ ).

d) Fiksert a priori.

Det er ellers verdt å bemerke at estimatet for h når formuleslikningen estimeres uten restriksjoner, er nær 1, mens estimatet for i er nær null. Dette tyder på at inntekten i løpet av året er en betydelig forklaringsfaktor for formuen ved årets slutt, mens konsumutgiften ikke spiller nevneverdig rolle.

Variansestimaterne til målefeilene er på omtrent samme nivå som de tilsvarende estimater i modell A. For formuen utgjør målefeilvariansen hele 97 prosent (!) av totalvariansen når formuesligningen estimeres uten restriksjoner, mens andelen faller til 61 prosent når vi pålegger "økosirk-betingelsen"  $h=1, i=-1$ . Denne nedgangen ledsages av en kraftig økning av variansen til restleddet i formuesligningen,  $\hat{\sigma}_{u_3}^2$ .<sup>20)</sup>

#### 4.4.3. Modell C

Den tredje modellen går et skritt videre, idet den tar hensyn til inntekts og formuens sammensetning. Vi deler her inntekten i fire grupper<sup>21)</sup>

RL: Lønnsinntekt,

RT: Trygdeinntekt,

RF: Nettoinntekt av finanskapital (= Renter og utbytte minus gjeldsrenter),

RR: Inntekt av realkapital, næringsinntekt mv.,

og formuen i to<sup>22)</sup>

FF: Nettofinansformue,

FR: Realformue.

Dessuten skiller vi ut 3 av konsumutgiftsgruppene<sup>23)</sup>

CI: Utgift til ikke-varige goder,

CV: Utgift til varige goder,

CT: Utgift til tjenester eksklusive boligjenester

som egne variable på linje med totalutgiften C.

20) Siden  $F^*$  og  $\epsilon_F$  pr. forutsetning er ukorrelerte, er  $\text{var } F^* = \text{var } F - \text{var } \epsilon_F$ . Benytter vi den empiriske varians til  $F$ , som er 12383, som estimat for  $\text{var } F$ , får vi i tilfelle B2 estimatet  $\text{var } \hat{F}^* = 12383 - 86.6^2 = 4883$ . Dette er ikke så langt unna FIML-estimatet for  $\text{var } u_3$  i dette tilfelle. Tolker vi (4.23) med  $h=1$  og  $i=-1$  som en definisjonsligning, hvor  $u_3$  representerer initialformuen, er dette et rimelig resultat, fordi formuen ved årets begynnelse og årets slutt etter all sannsynlighet har omtrent like stor varians.

21) Med de symboler vi brukte i kapitlene 2 og 3, har vi  $RL = R_1, RT = R_2, RF = R_4 - G, RR = R_3 + R_5 + R_6$ .

22) Med de symboler vi brukte i kapitlene 2 og 3, har vi  $FF = F_1 + F_2 + F_6 - G, FR = F_3 + F_4 + F_5$ .

23) Uttrykt med symbolene i kapitlene 2 og 3 er altså  $CI = C_{IV}, CV = C_V, CT = C_T$ .

Vi forutsetter at lønnsinntekten og trygdeinntekten har lik marginal konsumtilbøyelighet. Det samme gjelder inntekten av finanskapital, realkapitalinntekten og næringsinntekten. Vi setter altså

$$(4.27) \quad C^{\mathbf{x}} = a_1 (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_2 (RF^{\mathbf{x}} + RR^{\mathbf{x}}),$$

$$(4.27a) \quad CI^{\mathbf{x}} = a_{1I} (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_{2I} (RF^{\mathbf{x}} + RR^{\mathbf{x}}),$$

$$(4.27b) \quad CV^{\mathbf{x}} = a_{1V} (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_{2V} (RF^{\mathbf{x}} + RR^{\mathbf{x}}),$$

$$(4.27c) \quad CT^{\mathbf{x}} = a_{1T} (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_{2T} (RF^{\mathbf{x}} + RR^{\mathbf{x}}),$$

hvor stjerner, som før, markerer latente verdier. Lønns- og trygdeinntekten antas å bli bestemt ved hovedpersonens alder og antall husholdningsmedlemmer,

$$(4.28) \quad RL^{\mathbf{x}} = f_L A + g_L N + u_1,$$

$$(4.29) \quad RT^{\mathbf{x}} = f_T A + g_T N + u_2,$$

mens inntekten av finanskapital bestemmes ved størrelsen av finanskapitalen,

$$(4.30) \quad RF^{\mathbf{x}} = e_F FF^{\mathbf{x}} + u_3,$$

og realkapitalinntekten og næringsinntekten ved realkapitalens størrelse,

$$(4.31) \quad RR^{\mathbf{x}} = e_R FR^{\mathbf{x}} + u_4.$$

Fordi næringsinntekten inneholder et betydelig element av arbeidsinntekt, kan den siste hypotesen være noe tvilsom, men er valgt for å forenkle modellstrukturen.

Vi forutsetter at restleddene i inntektsgenereringsligningene (4.28)-(4.31) er ukorrelerte.<sup>24)</sup> Disse ligningene kan betraktes som en disaggregering av ligning (4.17) i modell A. Sammenhengen mellom de observerte og de latente variable er gitt ved

24) Siden LISREL-modellen er basert på at alle målefeil er ukorrelerte (jfr. ligning (4.9)), innebærer denne spesifikasjonen dessuten at målefeilene i CI, CV og CT forutsettes å være ukorrelert med målefeilen i totalkonsumet. Dette er utvilsomt en dristig hypotese, men skulle vi løsne på den, ville vi måtte utvide modellen med atskillige parametre.

$$(4.32) \quad C = C^{\mathbf{x}} + \varepsilon_C,$$

$$(4.33) \quad CI = CI^{\mathbf{x}} + \varepsilon_{CI},$$

$$(4.34) \quad CV = CV^{\mathbf{x}} + \varepsilon_{CV},$$

$$(4.35) \quad CT = CT^{\mathbf{x}} + \varepsilon_{CT},$$

$$(4.36) \quad RL = RL^{\mathbf{x}} + \varepsilon_{RL},$$

$$(4.37) \quad RT = RT^{\mathbf{x}} + \varepsilon_{RT},$$

$$(4.38) \quad RF = RF^{\mathbf{x}} + \varepsilon_{RF},$$

$$(4.39) \quad RR = RR^{\mathbf{x}} + \varepsilon_{RR},$$

$$(4.40) \quad FF = FF^{\mathbf{x}} + \delta_{FF},$$

$$(4.41) \quad FR = FR^{\mathbf{x}} + \delta_{FR}.$$

Modellen er følgende spesialtilfelle av den generelle modellen i avsnitt 4.3:<sup>25)</sup>

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & -a_1 & -a_1 & -a_2 & -a_2 \\ 0 & 1 & 0 & 0 & -a_{1I} & -a_{1I} & -a_{2I} & -a_{2I} \\ 0 & 0 & 1 & 0 & -a_{1V} & -a_{1V} & -a_{2V} & -a_{2V} \\ 0 & 0 & 0 & 1 & -a_{1T} & -a_{1T} & -a_{2T} & -a_{2T} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

25) Den kan eventuelt oppfattes som en generalisering av den "dis-aggregerte Friedman-modell" som er studert av Willassen [46]. Hans modell er estimert på grunnlag av aggregerte tidsseriedata for Norge.

$$\Gamma = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & f_L & g_L \\ 0 & 0 & f_T & g_T \\ e_F & 0 & 0 & 0 \\ 0 & e_R & 0 & 0 \end{bmatrix}$$

$$Y^{**} = \begin{pmatrix} C^{**} \\ CI^{**} \\ CV^{**} \\ CT^{**} \\ RL^{**} \\ RT^{**} \\ RF^{**} \\ RR^{**} \end{pmatrix}, \quad Y = \begin{pmatrix} C \\ CI \\ CV \\ CT \\ RL \\ RT \\ RF \\ RR \end{pmatrix}, \quad X^{**} = \begin{pmatrix} FF^{**} \\ FR^{**} \\ A \\ N \end{pmatrix}, \quad X = \begin{pmatrix} FF \\ FR \\ A \\ N \end{pmatrix}$$

$$u = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_C \\ \epsilon_{CI} \\ \epsilon_{CV} \\ \epsilon_{CT} \\ \epsilon_{RL} \\ \epsilon_{RT} \\ \epsilon_{RF} \\ \epsilon_{RR} \end{pmatrix}, \quad \delta = \begin{pmatrix} \delta_{FF} \\ \delta_{FR} \\ 0 \\ 0 \end{pmatrix}$$

$$m=8, \quad n=4.$$





Tabell 20. Maximum Likelihood (FIML)-estimator for modell C: a) b) c)

$$\begin{aligned}
 C^{\mathbf{x}} &= a_1 (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_2 (RF^{\mathbf{x}} + RR^{\mathbf{x}}) & C &= C^{\mathbf{x}} + \epsilon_C \\
 CI^{\mathbf{x}} &= a_{1I} (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_{2I} (RF^{\mathbf{x}} + RR^{\mathbf{x}}) & CI &= CI^{\mathbf{x}} + \epsilon_{CI} \\
 CV^{\mathbf{x}} &= a_{1V} (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_{2V} (RF^{\mathbf{x}} + RR^{\mathbf{x}}) & CV &= CV^{\mathbf{x}} + \epsilon_{CV} \\
 CT^{\mathbf{x}} &= a_{1T} (RL^{\mathbf{x}} + RT^{\mathbf{x}}) + a_{2T} (RF^{\mathbf{x}} + RR^{\mathbf{x}}) & CT &= CT^{\mathbf{x}} + \epsilon_{CT} \\
 RL^{\mathbf{x}} &= f_L A + g_L N + u_1 & RL &= RL^{\mathbf{x}} + \epsilon_{RL} \\
 RT^{\mathbf{x}} &= f_T A + g_T N + u_2 & RT &= RT^{\mathbf{x}} + \epsilon_{RT} \\
 RF^{\mathbf{x}} &= e_F FF^{\mathbf{x}} + u_3 & RF &= RF^{\mathbf{x}} + \epsilon_{RF} \\
 RR^{\mathbf{x}} &= e_R FR^{\mathbf{x}} + u_4 & RR &= RR^{\mathbf{x}} + \epsilon_{RR} \\
 & & FF &= FF^{\mathbf{x}} + \delta_{FF} \\
 & & FR &= FR^{\mathbf{x}} + \delta_{FR}
 \end{aligned}$$

Variant	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_{1I}$	$\hat{a}_{2I}$	$\hat{a}_{1V}$	$\hat{a}_{2V}$	$\hat{a}_{1T}$	$\hat{a}_{2T}$
C1	0.727 (0.016)	0.656 (0.023)	0.240 (0.006)	0.233 (0.008)	0.119 (0.006)	0.146 (0.009)	0.124 (0.006)	0.096 (0.008)
C2	0.805	2.529	0.240	0.610	0.134	0.531	0.141	0.424
C3	1.915 (0.053)	0.545 (0.050)	0.494 (0.015)	0.167 (0.017)	0.371 (0.014)	0.144 (0.014)	0.328 (0.012)	0.074 (0.011)
C4	1.914	0.550	0.494	0.170	0.371	0.145	0.328	0.074

Variant	$\hat{f}_L$	$\hat{g}_L$	$\hat{f}_T$	$\hat{g}_T$	$\hat{e}_F$	$\hat{e}_R$
C1	-0.310 (0.026)	4.611 (0.262)	0.250 (0.008)	-0.106 (0.081)	0.036 (0.001)	0.059 (0.001)
C2	-0.310	4.611	0.250	-0.106	0.035	0.038
C3	-0.356 (0.016)	3.347 (0.179)	0.245 (0.008)	-0.243 (0.078)	0.036 (0.001)	0.059 (0.001)
C4	-0.356	3.336	0.245	-0.245	0.047	0.068

a) C, CI, CV, CT, RL, RT, RF, RR, FF og FR er målt i 1 000 kr.

b) Parametre hvis verdi a priori er satt lik null, er markert med tankestrek. Estimator for asymptotiske standardavvik er gjengitt i parentes. I varianstene C2 og C4 lot standardavvikestimater seg ikke beregne.

c)  $\sigma_{ui}^2 = \text{var } u_i$  ( $i=1, \dots, 4$ ).

$\sigma_{\epsilon_i}$  = standardavvik til  $\epsilon_i$  ( $i = C, CI, CV, CT, RL, RT, RF, RR$ ).

$\sigma_{\delta_i}$  = standardavvik til  $\delta_i$  ( $i = FF, FR$ ).

Tabell 20 (forts.). Maximum Likelihood (FIML)-estimer for modell C:<sup>a)b)c)</sup>

Vari- ri- ant	$\hat{\sigma}_{\epsilon C}$	$\hat{\sigma}_{\epsilon CI}$	$\hat{\sigma}_{\epsilon CV}$	$\hat{\sigma}_{\epsilon CT}$	$\hat{\sigma}_{\epsilon RL}$	$\hat{\sigma}_{\epsilon RT}$	$\hat{\sigma}_{\epsilon RF}$	$\hat{\sigma}_{\epsilon RR}$	$\hat{\sigma}_{\delta FF}$	$\hat{\sigma}_{\delta FR}$
C1	21.324 (0.264)	7.700 (0.095)	8.287 (0.102)	7.666 (0.095)	-	-	-	-	-	-
C2	0.000	6.076	6.997	6.751	-	-	4.624	12.675	-	-
C3	0.000 (21143.9)	6.130 (0.084)	7.013 (0.089)	6.747 (0.085)	18.255 (0.240)	6.007 (0.128)	0.000 (13168.4)	5.404 (1.486)	-	-
C4	0.000	6.129	7.013	6.747	18.256	6.008	0.000	5.424	74.674	60.879

Vari- ant	$\hat{\sigma}_{u1}^2$	$\hat{\sigma}_{u2}^2$	$\hat{\sigma}_{u3}^2$	$\hat{\sigma}_{u4}^2$	Antall ukjen- te struk- tur- koef.	Antall ukjen- te rest- ledds- vari- anser	Antall vari- anser til feil- ledd	Antall 2. or- dens- momen- ter, eks. var.	Totalt antall ukjen- te para- metre t
C1	457.855 (11.323)	43.682 (1.080)	43.177 (1.068)	179.958 (4.451)	14	4	4	10	32
C2	457.854	43.682	23.737	82.135	14	4	6	10	34
C3	143.905 (8.017)	7.728 (1.268)	43.177 (5.753)	150.719 (16.478)	14	4	8	10	36
C4	144.023	7.717	33.626	135.396	14	4	10	10	38

a) Se note a, side 78.

b) Se note b, side 78.

c) Se note c, side 78.

Estimatene for lønns- og trygdeinntektens marginale konsumtilbøyelighet er stort sett vesentlig forskjellig fra de tilsvarende estimater for formuesinntekten. Når vi åpner for målefeil i formuesinntekten, firedobles estimatet for konsumtilbøyeligheten til denne inntektskomponenten. Tillater vi målefeil også i lønns- og trygdeinntekten, synker estimatene for  $a_2$ ,  $a_{2I}$ ,  $a_{2V}$  og  $a_{2T}$  litt, mens estimatene for  $a_1$ ,  $a_{1I}$ ,  $a_{1V}$  og  $a_{1T}$  mer enn fordobles (sammenlign variantene C3 og C1). Estimater for  $a_1$  (dvs. konsumtilbøyeligheten for lønns- og trygdeinntekt) er da så høyt som 1.91, mens estimatet for  $a_2$  (dvs. konsumtilbøyeligheten for formuesinntekt og næringsinntekt) bare er 0.55.

Forholdet mellom estimatene er gitt i følgende tabell:

Variant	$\hat{a}_1/\hat{a}_2$	$\hat{a}_{1I}/\hat{a}_{2I}$	$\hat{a}_{1V}/\hat{a}_{2V}$	$\hat{a}_{1T}/\hat{a}_{2T}$
C1	1.11	1.03	0.82	1.29
C2	0.32	0.39	0.25	0.33
C3	3.51	2.96	2.58	4.43
C4	3.48	2.91	2.56	4.43

Vi finner at  $\hat{a}_{1V}/\hat{a}_{2V}$  er mindre enn både  $\hat{a}_{1I}/\hat{a}_{2I}$  og  $\hat{a}_{1T}/\hat{a}_{2T}$  i samtlige varianter. Dette innebærer at nærings- og kapitalinntekten slår *relativt sett* svakere ut i kjøp av ikke-varige goder og tjenester enn i kjøp av varige goder og at denne konklusjonen er "robust" overfor valget av forutsetninger om målefeil.

Punkttestimatene for  $e_F$  og  $e_R$ , som kan tolkes som de marginale avkastningsrater for henholdsvis finansformue og realformue<sup>26)</sup>, har plausibel størrelse - i varianten uten målefeil henholdsvis 3.6 prosent og 5.9 prosent. De øker når det åpnes for målefeil, men utslaget er ikke på langt nær så markert som i modell A.

"Målestøyen" i formuesregistreringene synes å være betraktelig mindre i denne modellen enn i de to foregående. Vi finner anslag for målefeilvariansens andel av totalvariansen på 23 prosent for finansformue og 13 prosent for realformue.

#### 4.4.4. Konklusjon

Vi stilte innledningsvis spørsmålet om en neglisjering av målefeil i de variable kunne medføre en vesentlig underestimering av de marginale konsumtilbøyeligheter for inntekt og formue. Denne hypotesen har vi - i en viss forstand - fått bekreftet til overmål for inntektens vedkommende. Maximum Likelihood-estimatene for konsumtilbøyeligheten til inntektens latente del er så å si uten unntagelse utsagnskraftig større enn 1, mens de tilsvarende regresjonsestimater når målefeil neglisjeres, sjelden overstiger 0.7 og ofte er betydelig lavere.

26) Vi må her erindre at næringsinntekten i sin helhet er oppfattet som inntekt av realkapital.

Målefeilvariansen til inntekten estimeres til mellom trdjedelen av fjerdedelen av totalvariansen. For formuen synes den tilsvarende andel å være betydelig større. Dette beviser selvsagt ikke at inntekts- og formuesundersøkelsenes registreringer inneholder målefeil av slike dimensjoner. Men sett på bakgrunn av det faktum at både skattyterne og ligningsmyndighetene ofte viser betydelig grad av skjønn i sine vurderinger peker det nokså klart i retning av en modell med latente variable.

Det er ikke umiddelbart lett å godta de numeriske verdier av konsumtilbøyelighetene som modellvariantene med latente variable kommer ut med. Økonomer har nå engang vanskelig for å akseptere verdier av inntekts marginale konsumtilbøyelighet på over 1 - i hvert fall som gjennomsnittsanslag. Men tar vi samtidig med i vurderingen at inntektstall basert på skatteligningsmateriale inneholder ikke bare en tilfeldig, men høyst sannsynlig også en systematisk negativ feilkomponent,<sup>27)</sup> og at det er visse indikasjoner på at registreringsmetodene i forbruksundersøkelsene tenderer til å overvurdere verdien av konsumet,<sup>28)</sup> er våre estimater på ca. 1.15-1.25 lettere å forsone seg med.

Modelleksperimentene i dette avsnitt tyder ellers på at videre fremstøt med estimering av konsumfunksjoner innenfor mikromodeller hvor det tas hensyn til samspillet mellom konsum, inntekt og formue, kan være vel verdt. Men full avkastning vil slike forsøk neppe gi før definisjons-sammenhengene mellom de variable blir datamessig bedre dekket enn de er i dag. Vi bør antagelig også ta flere eksogene variable med i modellenes inntektsligninger enn det er blitt anledning til her.

---

27) Se avsnittene 2.2.2 og 4.1 ovenfor, Skarstad [37], avsnittene II.2 og II.6, og Biørn [5], avsnitt 5.

28) Dette gjelder spesielt i år med høy omsetning av varige konsumgoder.

## IDENTIFISERBARHETEN AV MODELL A

I dette appendiks vil vi se nærmere på muligheten for å identifisere strukturparametrene i modell A i avsnitt 4.4.1 ovenfor.<sup>1)</sup> Vi vil begrense oss til den minst restriktive varianten, A5, og skissere et resonnement som sannsynliggjør at de 15 ukjente parametre som opptrer der, er identifiserbare. Da vil selvsagt parametrene i de mer restriktive variantene, A1-A4, også være identifiserbare - med mulig unntagelse for visse ekstreme spesialtilfelle.

La  $E(Y Y') = \Sigma_{YY}$ ,  $E(Y X') = \Sigma_{YX}$ ,  $E(X X') = \Sigma_{XX}$ , slik at varians-kovariansmatrisen til modellens observerbare variable kan skrives på partisjonert form som

$$\Sigma = \begin{bmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{YX}' & \Sigma_{XX} \end{bmatrix}.$$

Matriseligningen (4.12) kan dermed spaltes i tre sub-systemer:

- (1)  $\Sigma_{XX} = \Phi + \theta_{\delta}^2$ ,
- (2)  $\Sigma_{YX} = B^{-1} \Gamma \Phi$ ,
- (3)  $\Sigma_{YY} = B^{-1} \Gamma \Phi \Gamma' B'^{-1} + B^{-1} \Psi B'^{-1} + \theta_{\epsilon}^2$ .

I modell A er

$$B = \begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} b & 0 & 0 \\ e & f & g \end{bmatrix}$$

$$\Phi = \begin{bmatrix} \sigma_{F^*}^2 & \sigma_{F^*A} & \sigma_{F^*N} \\ \sigma_{F^*A} & \sigma_A^2 & \sigma_{AN} \\ \sigma_{F^*N} & \sigma_{AN} & \sigma_N^2 \end{bmatrix}, \quad \Psi = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}$$

$$\theta_{\epsilon}^2 = \begin{bmatrix} \sigma_{\epsilon C}^2 & 0 \\ 0 & \sigma_{\epsilon R}^2 \end{bmatrix}, \quad \theta_{\delta}^2 = \begin{bmatrix} \sigma_{\delta F}^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

1) Modellene B og C kan behandles på tilsvarende måte, men algebraen blir en del mer komplisert.

Ved å eliminere  $\phi$  - som vi forutsetter er ikke-singulær - fra ligningene (1) og (2) får vi

$$(4) \quad B^{-1}\Gamma = \begin{bmatrix} b + ae & af & ag \\ e & f & g \end{bmatrix} = \Sigma_{YX} (\Sigma_{XX} - \theta_{\delta}^2)^{-1}.$$

Hvis dette ligningssystemet har slike egenskaper at vi entydig kan bestemme  $a, b, e, f, g$  og  $\sigma_{\delta F}^2$  når  $\Sigma_{YX}$  og  $\Sigma_{XX}$  er kjent, vil disse 6 parametre være identifiserbare. Når  $\theta_{\delta}^2$  er identifiserbar, følger direkte av (1) at det samme vil gjelde parametrene i  $\phi$ .<sup>2)</sup>

De gjenstående 3 parametre, dvs. den ukjente restleddsvariansen i  $\psi$  og målefeilvariansene i  $\theta_{\epsilon}^2$ , vil da også være identifiserbare. Det innsees på følgende måte: Siden

$$B^{-1} = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix},$$

følger av (3) at

$$(5) \quad B^{-1}\psi B'^{-1} + \theta_{\epsilon}^2 = \sigma_{u2}^2 \begin{bmatrix} a^2 & a \\ a & 1 \end{bmatrix} + \begin{bmatrix} \sigma_{\epsilon C}^2 & 0 \\ 0 & \sigma_{\epsilon R}^2 \end{bmatrix} = \Sigma_{YY} - B^{-1}\Gamma\Phi\Gamma'B'^{-1}.$$

Når uttrykket på høyre side av likhetstegnet i (5) er kjent, er det da lett å bestemme  $\sigma_{u2}^2$ ,  $\sigma_{\epsilon C}^2$  og  $\sigma_{\epsilon R}^2$ . Følgelig er de identifiserbare.

---

2) En formell diskusjon av identifikasjonsegenskapene for denne type av modeller er gitt i Rothenberg [35], avsnitt 6.

## 5. OVERSIKT OG HOVEDKONKLUSJONER

Erfaringene fra denne undersøkelsen kan sammenfattes i fire hovedpunkter.

### Estimerings- og testresultater

Denne studien har klarlagt at individuelle husholdningsoppgaver over forbruksutgifter koblet med inntekts- og formuesoppgaver basert på skatteligningsmateriale inneholder betydelig informasjon om forbruksstrukturen. Materialet bidrar til å belyse ikke bare inntektens og formuens stilling som konsummotiverende variable, men til en viss grad også virkningen av alder, husholdningsstørrelse og andre demografiske og sosioøkonomiske variable. Selvsagt er ikke alle spørsmål tilfredsstillende besvart. Noen av de svar vi har oppnådd, er i tråd med det vi ventet på forhånd; i andre tilfelle har vi fått overraskelser eller skuffelser.

Den klareste positive konklusjon er kanskje at både størrelsen av inntekten og dens sammensetning på artsgrupper har betydning for størrelsen av konsumet. Inntektssammensetningen slår sterkere igjennom enn hovedpersonens sosialgruppetilknytning. I den grad en sosialgruppespesifikk forklaringskomponent kan etterspores, er det i første rekke fordi selvstendig næringsdrivende i primærnæringene skiller seg ut. Resultatene bekrefter ikke at lønns- og trygdeinntekten har høyere marginal konsumtilbøyelighet enn andre inntektskomponenter, men gir heller ikke grunnlag for den motsatte konklusjon.

Verdien av den enkelte husholdnings formue, slik den registreres ved skatteligningen, har overraskende liten betydning for konsumutgiften. Hovedpersonens alder er derimot en klart utsagnskraftig variabel. Som hovedregel ser det ut til at konsumutgiften avtar med alderen, under ellers like forhold. På den annen side er det visse indikasjoner på at inntektens *marginale* konsumtilbøyelighet stiger med alderen.<sup>1)</sup>

Deler av undersøkelsen bygger på en inndeling av konsumgodene i fem grupper. Analysen avdekker forskjeller mellom disse gruppene på flere punkter, blant annet når det gjelder betydningen av alders- og formuesvariablene. Vi mister altså vesentlige nyanser i totalbildet ved å aggregere konsumutgiften til én samlestørrelse, slik det har vært vanlig å gjøre i konsum- og spareanalyser på grunnlag av mikrodata.

1) Resultater basert på det samme datamateriale, men på et annet modellopplegg (livsinntektshypotesen), peker i samme retning. (Biørn [6], avsnitt 5.)



Vi får bekreftet den erfaring andre tidligere har gjort, at regresjonsestimering av konsumfunksjoner ved hjelp av tverrsnittsdata gir markert lavere konsumtilbøyeligheter for inntekten enn tilsvarende estimering på grunnlag av aggregerte tidsserier. I vårt tilfelle er imidlertid forskjellene mer utpreget enn i de fleste andre undersøkelser. Målefeil i inntektsdataene ser ut til å kunne gi noe av forklaringen på de lave estimatene. Det later til at målefeilvariansen utgjør en ikke ubetydelig del av totalvariansen både for inntekt og formue.

Resultatene støtter ellers, i noen grad, en hypotese om at konsumutgift, inntekt og formue er simultant bestemte variable - altså at vi begår en feilspesifikasjon hvis vi estimerer som om inntekt og formue er eksogene forklaringsfaktorer for konsumet. Men datakvaliteten er ikke så god at den tillater en klar konklusjon på dette punkt.

### Metode

Med denne studien har vi vist at en strategi for multiple sammenligninger innenfor et regresjonsopplegg kan være et slagkraftig verktøy for analyse av samvariasjonsproblemer hvor det er viktig å holde mange konkurrerende hypoteser under oppmerksomhet på én gang. Vårt valg av kriterium for å holde den statistiske usikkerhet under kontroll kan selvfølgelig diskuteres, men gir et oversiktlig beregningsopplegg selv når hypotesestrukturen er komplisert. På den annen side kan vi risikere å få flertydige "løsninger" på testproblemet - spesielt når antall deltester er stort.

Videre har vi godtgjort at simultan-estimering av flerligningsmodeller med latente strukturvariable ligger innenfor mulighetenes grense rent beregningsteknisk. Men kravet om at hele modellen må være identifiserbar på den ene side og ønsket om noenlunde rask konvergens ved iterasjonsberegningene på den annen setter forholdsvis snevre grenser for hva slags modeller som kan behandles i praksis.

### Data

Erfaringene fra en empirisk studie som denne bør også nedfelle seg i synspunkter på datamaterialet. Det er ingen tvil om at vårt sammenkoblede materiale av konsum-, inntekts- og formuestall gir et bedre grunnlag for å splitte de variable i meningsfylte undergrupper enn det materiale som vanligvis står til rådighet for konsumanalyser. Men mangelen på fullstendig avstemte husholdningsregnskaper er følbar. Spesielt har det vist seg å være et problem at bare formuen ved utgangen av året og ikke initialformuen er kjent. Dessuten er sammenhengen mellom verdien av de varige konsumgoder slik de registreres i forbruksundersøkelsen, og de skattemessige anslag for verdien av innbo og løsøre uklar. Endelig medfører den

korte registreringsperioden for forbruksutgifter at det blir vanskelig å finne gode indikatorer for sparingen på årsbasis i de enkelte husholdninger.

Ved en mer omfattende databearbeidelse enn det er blitt anledning til her, kunne nok noen av problemene ovenfor elimineres. Men et virkelig gjennombrudd ville det først kunne bli om inntekts- og formuesundersøkelsene ble utvidet til årlige undersøkelser med en viss "rotering" av utvalgene - slik det i dag gjøres i forbruksundersøkelsene. Da ville vi kunne oppnå sammenhengende tidsserier av forbruks-, inntekts- og formuestall for enkelthusholdninger og ville dermed kunne utvide perspektivet til dynamiske modeller. Det er bare de aller enkleste teorier for konsum og sparing som kan analyseres fullt ut tilfredsstillende ved tverrsnittsdata alene.

### Anvendelser

Et viktig formål med undersøkelser av denne type er at de skal gi impulser til utvikling og forbedring av økonometriske analysemodeller. Vi er ikke nådd så langt at vi har oppnådd koeffisientestimater som direkte kan settes inn i eksisterende modeller. Men det er ingen tvil om at våre resultater når det gjelder betydningen av inntektssammensetning, husholdningstype og alder kan gi holdepunkter for å forbedre beskrivelsen av det private konsum i makroøkonomiske modeller som MODIS IV. De bør også kunne gi impulser til en mer raffinert behandling av sammenhengen mellom konsumutgift og inntekt i modeller for skatteinsidensanalyse.

Alt i alt gir resultatene en påminnelse om at forbindelseslinjene mellom konsum, inntekt og formue er vesentlig mer kompliserte enn en vanligvis forestiller seg i stiliserte teori-resonnementer.

## LITTERATURHENVISNINGER

- [ 1 ] Anderson, T.W.: *An Introduction to Multivariate Statistical Analysis*. (New York: John Wiley & Sons, Inc., 1958.)
- [ 2 ] Anderson, T.W.: *The Statistical Analysis of Time Series*. (New York: John Wiley & Sons, Inc., 1971.)
- [ 3 ] Attfield, C.L.F.: Estimation of the Structural Parameters in a Permanent Income Model. *Economica*, vol. 43 (1976), pp. 247-254.
- [ 4 ] Biørn, E.: *Estimering av makro-konsumfunksjoner for etterkrigs-tiden: Metodespørsmål og empiriske resultater*. Artikler fra Statistisk Sentralbyrå, nr. 63. (Oslo: Statistisk Sentralbyrå, 1974.)
- [ 5 ] Biørn, E.: Sammenhengen mellom forbruksutgift og beskattet inntekt - noen prinsipielle betraktninger supplert med norske data. Finnes i: *Personbeskattningen i högskattesamhället*. Nordiska skattevetenskapliga forskningsrådets skriftserie nr. 4 (Stockholm: Liber Förlag, 1977), pp. 71-85.
- [ 6 ] Biørn, E.: The Consumption Function and the Life-Cycle Hypothesis: An Analysis of Norwegian Household Data. Upublisert notat, oktober 1979.
- [ 7 ] Biørn, E. og Garaas, E.: *Inntekts- og forbruksbeskatning fra et fordelingssynspunkt - en modell for empirisk analyse*. Samfunnsøkonomiske studier nr. 30. (Oslo: Statistisk Sentralbyrå, 1976.)
- [ 8 ] Bodkin, R.: Windfall Income and Consumption. *American Economic Review*, vol. 49 (1959), pp. 602-614.
- [ 9 ] Cappelen, Å.: Makrokonsumfunksjonen i MODIS IV. Arbeidsnotater fra Statistisk Sentralbyrå, IO 78/6 (1978).
- [10] Christensen, L.R., Jorgenson, D.W. og Lau, L.J.: Transcendental Logarithmic Utility Functions. *American Economic Review*, vol. 65 (1975), pp. 367-383.
- [11] Conrad, K. og Jorgenson, D.W.: The Structure of Technology: Nonjointness and Commodity Augmentation, Federal Republic of Germany, 1950-1973. *Empirical Economics*, vol. 3 (1978), pp. 91-113.
- [12] Crockett, J.: Income and Asset Effects on Consumption: Aggregate and Cross Section. Finnes i: *Models of Income Determination*. Studies in Income and Wealth, vol. 28 (Princeton: Princeton University Press, 1964), pp. 97-132.
- [13] Crockett, J. og Friend, I.: Consumer Investment Behavior. Finnes i R. Ferber (ed.): *Determinants of Investment Behavior*. (New York: National Bureau of Economic Research, 1967), pp. 15-127.
- [14] Draper, N.R. og Smith, D.: *Applied Regression Analysis*. (New York: John Wiley & Sons, Inc., 1966.)
- [15] Ferber, R.: Research on Household Behaviour. *American Economic Review*, vol. 52 (1962), pp. 19-63.

- [16] Ferber, R.: Consumer Economics, a Survey. *Journal of Economic Literature*, vol. 11 (1973), pp. 1303-1342.
- [17] Fisher, M.R.: Exploration in Savings Behaviour. *Bulletin of the Oxford University Institute of Statistics*, vol. 18 (1956), pp. 201-277.
- [18] Friedman, M.: *A Theory of the Consumption Function*. (Princeton: Princeton University Press, 1957.)
- [19] Geraci, V.: Identification of Simultaneous Equation Models with Measurement Error. *Journal of Econometrics*, vol. 4 (1976), pp. 263-283.
- [20] Griliches, Z.: Errors in Variables and Other Unobservables. *Econometrica*, vol. 42 (1974), pp. 971-998.
- [21] Haavelmo, T.: Econometric Analysis of the Savings Survey Data. *Bulletin of the Oxford University Institute of Statistics*, vol. 19 (1957), pp. 145-149.
- [22] Husby, R.: A Nonlinear Consumption Function Estimated from Time-Series and Cross-Section Data. *Review of Economics and Statistics*, vol. 53 (1971), pp. 76-79.
- [23] Jöreskog, K.G.: A General Method for Estimating a Linear Structural Equation System. Kap. 5 i: A.S. Goldberger og O.D. Duncan (eds.): *Structural Equation Models in the Social Sciences*. (New York: Seminar Press, 1973.)
- [24] Jöreskog, K.G. og Sörbom, D.: LISREL IV: *Analysis of Linear Structural Relationships by the Method of Maximum Likelihood*. (Chicago: International Educational Services, 1978.)
- [25] Kaldor, N.: Alternative Theories of Distribution. *Review of Economic Studies*, vol. 23 (1955-56), pp. 94-100.
- [26] Liviatan, N.: Tests of the Permanent-Income Hypothesis Based on a Reinterview Saving Survey. Finnes i C. Christ (ed.): *Measurement in Economics*. (Stanford: Stanford University Press, 1963.)
- [27] Lovell, M.C.: Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis. *Journal of the American Statistical Association*, vol. 58 (1963), pp. 993-1010.
- [28] Malinvaud, E.: *Statistical Methods of Econometrics* (Amsterdam: North-Holland Publishing Company, 1966.)
- [29] Mayer, T.: *Permanent Income, Wealth and Consumption*. (Berkeley: University of California Press, 1972.)
- [30] Mizon, G.H.: Inferential Procedures in Nonlinear Models: An Application in a U.K. Industrial Cross Section Study of Factor Substitution and Returns to Scale. *Econometrica*, vol. 45 (1977), pp. 1221-1242.

- [31] Modigliani, F. og Brumberg, R.: Utility Analysis and the Consumption Function: An Interpretation of Cross Section Data. Finnes i: K.K. Kurihara (ed.): *Post-Keynesian Economics*. (London: George Allen and Unwin Ltd., 1955), pp. 388-436.
- [32] Norges Bank: *En vurdering av renteutviklingen og rentestrukturen i Norge*. Innstilling fra et utvalg nedsatt av Det kredittpolitiske samarbeidsutvalg. (Oslo: Norges Bank, 1974.)
- [33] Reid, M.G.: Consumption, Savings and Windfall Gains. *American Economic Review*, vol. 52 (1962), pp. 728-737.
- [34] Reiersøl, O.: Identifiability of a Linear Relation Between Variables which are Subject to Error. *Econometrica*, vol. 18 (1950), pp. 375-389.
- [35] Rothenberg, T.J.: Identification in Parametric Models. *Econometrica*, vol. 39 (1971), pp. 577-591.
- [36] Simon, H.A.: Causal Ordering and Identifiability. Finnes i: Hood, W.C. og Koopmans, T.C. (eds.): *Studies in Econometric Method*. (New York: John Wiley & Sons, Inc., 1953), pp. 49-74.
- [37] Skarstad, O.: *Levestandard for private husholdninger*. Artikler fra Statistisk Sentralbyrå, nr. 102. (Oslo: Statistisk Sentralbyrå, 1977.)
- [38] Spjøtvoll, E.: Multiple Testing in the Analysis of Variance. *Scandinavian Journal of Statistics*, vol. 1 (1974), pp. 97-114.
- [39] Statistisk Sentralbyrå: *Forbruksundersøkelse 1973*. Norges offisielle statistikk A 705. (Oslo: Statistisk Sentralbyrå, 1975.)
- [40] Statistisk Sentralbyrå: *Kredittmarkedstatistikk 1972-1974*. Norges offisielle statistikk A 748. (Oslo: Statistisk Sentralbyrå, 1975.)
- [41] Statistisk Sentralbyrå: *Inntektsstatistikk 1973*. Norges offisielle statistikk A 892. (Oslo: Statistisk Sentralbyrå, 1977.)
- [42] Statistisk Sentralbyrå: *Formuesstatistikk 1973*. Norges offisielle statistikk A 922. (Oslo: Statistisk Sentralbyrå, 1977.)
- [43] Thurow, L.: The Optimum Lifetime Distribution of Consumption Expenditures. *American Economic Review*, vol. 59 (1969), pp. 324-330.
- [44] Tobin, J.: *Essays in Economics. Volume 2: Consumption and Econometrics*. (Amsterdam: North-Holland Publishing Company, 1975.)
- [45] Watts, H.W.: Long-Run Income Expectations and Consumer Saving. Finnes i: Dernburg, T.F., Rosett, R.N. og Watts, H.W. (eds.): *Studies in Household Economic Behavior*. (New Haven: Yale University Press, 1958), pp. 101-144.
- [46] Willassen, Y.: An Analysis of the Effect of Transitory Income on Consumption. *Scandinavian Journal of Economics*, vol. 80 (1978), pp. 299-310.
- [47] Zellner, A.: Estimation of Regression Relationships Containing Unobservable Independent Variables. *International Economic Review*, vol. 11 (1970), pp. 441-454.

*SUMMARY IN ENGLISH*

This is an econometric study of the relationship between consumption, income, and wealth in Norwegian households. Its basic aim is to obtain cross-sectional information on the form of the consumption function for use in tax incidence models and models for macro-economic analysis and forecasting.

A survey of the data is given in Chapter 2. The data file is constructed by combining data from two sources. The consumption data are from the Survey of Consumer Expenditure of 1973, which is a sampling survey based on mixed interview and accounting of expenditures during a period of two weeks. The data on income and wealth are extracted from the individual reports to the tax authorities for the income year 1973, processed by the Central Bureau of Statistics in connection with the Income and Property Statistics of 1973. The sample contains 3 271 households. From the basic income data file the following aggregates are constructed: wage income; income from social security schemes; entrepreneurial income; income from financial assets net of debt interests; income from dwellings and other property income; and capital gains. For each household, the total value of net direct taxes is allocated proportionally on the different income components. The main categories of wealth are: bank deposits net of liabilities; bearer bonds, shares etc.; real property; and private cars and other consumer capital. The consumption goods are divided into five groups: non-durables; semi-durables; durables; housing and maintenance; and other services. Simple coefficients of correlation and empirical regressions are calculated to illustrate main features of the data.

In Chapter 3, we explore several hypotheses concerning the parametric form of the consumption function. For this purpose, we construct four schemes of mutually interrelated hypotheses, each of which is analysed by means of a multiple testing strategy. The criterion employed in selecting the "best" hypothesis in each different scheme is to control the probability of making at least one false rejection among the various sub-tests involved. In the first scheme, we investigate the dependence of the consumption expenditure on the composition of income and wealth. The second scheme deals with the interaction of type of income and social status of the head of household, whereas the third is concerned with the possible interactions between income, wealth, the age of the head of household, and the number of household members. The purpose of the fourth scheme is to analyse the homogeneity of the consumption function by using a slightly modified trans-log function.

The main testing results are the following:

- The effect on the consumption expenditure of an increase in income from financial assets is not significantly different from the effect of a corresponding decrease in debt interests. As regards the effect of the wealth variables, the hypothesis of complete symmetry between bank deposits and liabilities is rejected for non-durable goods, but cannot be rejected for the four other commodity groups.

- Our tests do not confirm the assumption often made that consumption expenditure is more responsive to changes in wage income and income from social security schemes than it is to changes in entrepreneurial income and income from capital (profits). For only one of the commodity groups, services excl. housing and maintenance, wage income comes out with the highest estimate of the marginal propensity to consume. For this group, however, income from social security schemes has the lowest estimate. Income from financial assets has the highest estimate of the marginal propensity to consume for two of the commodity groups, viz. non-durable and semi-durable goods.

- For non-durable and durable goods, our results indicate that the marginal propensity to consume depends not only on the type of income, but also on the social status of the head of household. Households headed by self-employed persons in agriculture, forestry and fishing are the most atypical in this respect. For the other commodity groups, there is no evidence that the marginal propensity to consume depends on the social status of the household when differences in the composition of income are accounted for.

- Our multiple testing strategy only partly allows firm conclusions to be drawn concerning the degree of non-linearity of the consumption function. For semi-durable goods, we find no evidence of departure from linearity. For non-durables and services, the marginal propensity to consume of income seems to be a (slowly) decreasing function of income and an increasing function of age.

- We get no support of the hypothesis that the consumption function is homogeneous of degree one in income and the number of persons in the household. The hypothesis that the degree of homogeneity is constant (but different from unity) is rejected too.

- The ordinary least squares estimates of the marginal propensity to consume of income are generally rather low - between 0.3 and 0.7 for the total expenditure - and considerably lower than those usually obtained from aggregate time series data.

Chapter 4 is concerned with the estimation of consumption functions within the framework of linear simultaneous equation models with latent (unobservable) structural variables. The computer program LISREL, developed by Karl G. Jöreskog, is utilized to calculate Full Information Maximum Likelihood (FIML) estimates of three models, each containing a consumption function and an income generating function, i.e. a relation giving the latent (permanent) component of income as a linear function of wealth, age of the head of household, and the number of household members.

A main conclusion is that when errors of measurement in income are allowed for, the FIML estimates of the marginal propensity to consume of (latent) income are about 1.2. This holds regardless of the assumptions made concerning errors of measurement in wealth. The corresponding estimates in the absence of errors of measurement are about 0.7 or lower. Correspondingly, the estimate of the coefficient of the wealth variable in the income generating function - which may be interpreted as an implicit estimate of the rate of return of investment - increases considerably (from about 3 per cent to about 18 per cent) when errors of measurement in wealth are allowed for.

The estimate of the variance of the error (transitory) component of income is about 30 per cent of the total variance. For wealth, the corresponding fraction is considerably higher. Of course, this does not prove that our data are affected by errors of this order of magnitude, but it suggests that models with latent variables are strong competitors to simple regression models in analysing consumption functions from micro data.

Finally, our experiences reported in Chapter 4 also suggest that micro models determining consumption, income, and wealth simultaneously may well deserve a closer examination by practising econometricians. However, owing to inadequacies of our data, e.g. the fact that the tax reports only give wealth at the end of the year and not initial wealth, several interesting questions concerning this simultaneity could not be properly answered in this study.



## Utkommet i serien SØS

*Issued in the series Social Economic Studies (SES)*

- Nr. 1 Det norske skattesystems virkninger på den personlige inntektsfordeling *The Effects of the Norwegian Tax System on the Personal Income Distribution* 1954 Sidetall 103 Pris kr 3,00
- 2 Skatt på personleg inntekt og midel *Tax on Personal Income and Capital* 1954 Sidetall 120 Pris kr 3,00
- 3 Økonomisk utsyn 1900-1950 *Economic Survey* 1955 Sidetall 217 Pris kr 4,00
- 4 Nasjonalregnskap. Teoretiske prinsipper *National Accounts. Theoretical Principles* 1955 Sidetall 123 Pris kr 3,00
- 5 Avskrivning og skattlegging *Depreciation and Taxation* 1956 Sidetall 85 Pris kr 3,00
- 6 Bedriftsskatter i Danmark, Norge og Sverige *Corporate Taxes in Denmark, Norway and Sweden* 1958 Sidetall 101 Pris kr 4,00
- 7 Det norske skattesystemet 1958 *The Norwegian System of Taxation* 1958 Sidetall 159 Pris kr 6,50
- 8 Produksjonsstruktur, import og sysselsetting *Structure of Production, Imports and Employment* 1959 Sidetall 129 Pris kr 5,50
- 9 Kryssløpsanalyse av produksjon og innsats i norske næringer 1954 *Input-Output Analysis of Norwegian Industries* 1960 Sidetall 614 Pris kr 10,00
- 10 Dødeligheten og dens årsaker i Norge 1856-1955 *Trend of Mortality and Causes of Death in Norway* 1962 Sidetall 246 Pris kr 8,50
- 11 Kriminalitet og sosial bakgrunn *Crimes and Social Background* 1962 Sidetall 194 Pris kr 7,00
- 12 Norges økonomi etter krigen *The Norwegian Post-War Economy* 1965 Sidetall 437 Pris kr 15,00
- 13 Ekteskap, fødsler og vandringer i Norge 1856-1960 *Marriages, Births and Migrations in Norway* 1965 Sidetall 221 Pris kr 9,00
- 14 Foreign Ownership in Norwegian Enterprises *Utenlandske eierinteresser i norske bedrifter* 1965 Sidetall 213 Pris kr 12,00
- 15 Progressiviteten i skattesystemet 1960 *Statistical Tax Incidence Investigation* 1966 Sidetall 95 Pris kr 7,00
- 16 Langtidslinjer i norsk økonomi 1955-1960 *Trends in Norwegian Economy* 1966 Sidetall 150 Pris kr 8,00
- 17 Dødelighet blant spedbarn i Norge 1901-1963 *Infant Mortality in Norway* 1966 Sidetall 74 Pris kr 7,00
- 18 Storbyutvikling og arbeidsreiser En undersøkelse av pendling, befolkningsutvikling, næringsliv og urbanisering i Oslo-området *Metropolitan Growth, Commuting and Urbanization in the Oslo Area* 1966 Sidetall 298 Pris kr 12,00

- Nr. 19 Det norske kredittmarked siden 1900 *The Norwegian Credit Market since 1900* Sidetall 395 Pris kr 11,00
- 20 Det norske skattesystemet 1967 *The Norwegian System of Taxation 1968* Sidetall 146 Pris kr 9,00
- 21 Estimating Production Functions and Technical Change from Micro Data. An Exploratory Study of Individual Establishment Time-Series from Norwegian Mining and Manufacturing 1959-1967 *Estimering av produktfunksjoner og tekniske endringer fra mikro data. Analyser på grunnlag av tidsrekker for individuelle bedrifter fra norsk bergverk og industri* 1971 Sidetall 226 Pris kr 9,00 ISBN 82-537-0014-8
- 22 Forsvarets virkninger på norsk økonomi *The Impact of the Defence on the Norwegian Economy* 1972 Sidetall 141 Pris kr 9,00 ISBN 82-537-0149-7
- 23 Prisutvikling og prisatferd i 1960-årene En presentasjon og analyse av nasjonalregnskapets prisdata 1961-1969 *The Development and Behaviour of Prices in the 1960's Presentation and Analysis of the Price-Data of the Norwegian National Accounts* 1974 Sidetall 478 Pris kr 15,00 ISBN 82-537-0279-5
- 24 Det norske skattesystemet I Direkte skatter 1974 *The Norwegian System of Taxation I Direct Taxes* 1974 Sidetall 139 Pris kr 9,00 ISBN 82-537-0399-6
- 25 Friluftsliv, idrett og mosjon *Outdoor Recreation, Sport and Exercise* 1975 Sidetall 114 Pris kr 8,00 ISBN 82-537-0469-0
- 26 Nasjonalregnskap, modeller og analyse En artikkelsamling til Odd Aukrusts 60-årsdag *National Accounts, Models and Analysis To Odd Aukrust in Honour of his Sixtieth Birthday* 1975 Sidetall 320 Pris kr 13,00 ISBN 82-537-0530-1
- 27 Den representative undersøgelsesmetode *The Representative Method of Statistical Surveys* 1976 Sidetall 64 Pris kr 8,00 ISBN 82-537-0538-7
- 28 Statistisk Sentralbyrå 100 år 1876-1976 *Central Bureau of Statistics 100 years* 1976 Sidetall 128 Pris kr 9,00 ISBN 82-537-0557-3
- 29 Statistisk Sentralbyrås 100-årsjubileum Prolog og taler ved festmøtet i Universitetets aula 11. juni 1976 *Central Bureau of Statistics Prologue and Addresses at the Centenary Celebration, University Hall* 1976 Sidetall 32 Pris kr 7,00 ISBN 82-537-0637-5
- 30 Inntekts- og forbruksbeskatning fra et fordelingssynspunkt - En modell for empirisk analyse *Taxation of Income and Consumption from a Distributional Point of View - A Model for Empirical Analysis* 1976 Sidetall 148 Pris kr 9,00 ISBN 82-537-0647-2
- 31 Det norske skattesystemet II Indirekte skatter og offentlige trygdeordninger 1976 *The Norwegian System of Taxation II Indirect Taxes and Social Security Schemes* 1977 Sidetall 124 Pris kr 13,00 ISBN 82-537-0713-4
- 32 Inntekt og forbruk for funksjonshemmede *Income and Consumer Expenditure of Disabled Persons* 1977 Sidetall 166 Pris kr 13,00 ISBN 82-537-0732-0

- Nr. 33 Prinsipper og metoder for Statistisk Sentralbyrås utvalgsundersøkelser *Sampling Methods Applied by the Central Bureau of Statistics of Norway* 1977 Sidetall 105 Pris kr 11,00 ISBN 82-537-0771-1
- 35 Flyttemotivundersøkelsen 1972 *Survey of Migration Motives* 1978 Sidetall 233 Pris kr 15,00 ISBN 82-537-0783-5
- 36 Konjunkturbølger fra utlandet i norsk økonomi *International Cycles in Norwegian Economy* 1979 Sidetall 141 Pris kr 13,00 ISBN 82-537-0910-2
- 37 Norske lytter- og seervaner *Radio Listening and Television Viewing in Norway* 1979 Sidetall 216 Pris kr 13,00 ISBN 82-537-0931-5
- 38 Analyse av investeringsatferd: Problemer, metoder og resultater *Analysing Investment Behaviour: Problems, Methods, and Results* 1979 Sidetall 91 Pris kr 13,00 ISBN 82-537-0952-8
- 39 Kvinners yrkesdeltaking i Norge *Female Labour Activity in Norway* 1979 Sidetall 162 Pris kr 13,00 ISBN 82-537-0961-7
- 40 Framskrivning av befolkningens utdanning til år 2000 *Projections of the Education Characteristics of the Population to the Year 2000* 1979 Sidetall 112 Pris kr 13,00 ISBN 82-537-0998-6
- 42 Analyse av sammenhengen mellom forbruk, inntekt og formue i norske husholdninger *Analysing the Relationship between Consumption, Income and Wealth in Norwegian Households* 1979 Sidetall 95 Pris kr 13,00 ISBN 82-537-1012-7

Publikasjonen utgis i kommisjon hos  
H. Aschehoug & Co. og Universitetsforlaget, Oslo,  
og er til salgs hos alle bokhandlere.

Pris kr. 13,00

ISBN 82-537-1012-7

Engers Boktrykkeri A/S - Otta