

# Arbeidsnotater

T A T I S T I S K S E N T R A L B Y R Å

IO 70/18

Oslo, 3. desember 1970

## OM ETTERSPORSELEN ETTER FERJETURER

### I NORSKE HUSHOLDNINGER

Av

Nils Johan Korsvik

#### INNHOOLD

	Side
Innl�dning .....	2
KAPITTEL I: Ettersp�rsele etter ferjeturer. En teoretisk dr�fting av problemstillingen .....	2
1. Ettersp�rselefunksjoner .....	2
2. Ettersp�rselefunksjon for ferjeturer .....	3
KAPITTEL II: Opplegg til den �konometriske analysen .....	5
1. Presisering av de variable .....	5
2. Valg av analysemetode .....	7
KAPITTEL III: Datamaterialet .....	8
1. Frafall .....	8
2. M�lefeil .....	8
KAPITTEL IV: Regresjonsanalysen .....	9
1. Modellene .....	9
2. Den �konomiske mening med relasjonene .....	10
3. Diskusjon av relasjonenes statistiske egenskaper ....	10
KAPITTEL V: Resultater av regresjonsberegningene. Kommentarer .....	11
1. Resultater .....	11
2. Tester .....	15
3. Dr�fting av fortegnene p� koeffisientene .....	16
4. Virkning p� ferjeturutgiftene ved endring i husholdningsinntekten .....	18
5. Sluttord .....	19
Litteratur .....	20

Dette arbeid er opprinnelig skrevet som spesialoppgave ved det sosial konomiske studium. Forfatteren har st tt fritt i valg av opplegg og unders kelsesmetoder. Arbeidet gjengis her en del forkortet og med en del endringer som forfatteren har  nsket   foreta. Synspunkter og konklusjoner st r for forfatterens regning.

*Ikke for offentligg ring. Dette notat er et arbeidsdokument og kan siteres eller refereres bare etter spesiell tillatelse i hvert enkelt tilfelle. Synspunkter og konklusjoner kan ikke uten videre tas som uttrykk for Statistisk Sentralbyr s oppfatning.*

## INNLEDNING

Hensikten med dette arbeidet er å analysere etterspørselen etter det vi kan kalle godet ferieturer.

Med bakgrunn i økonomisk teori og ved hjelp av økonometriske metoder vil jeg forsøke å "forklare" denne etterspørselen.

Det konkrete datamateriale som ligger til grunn, er hentet fra Ferieundersøkelsen for ferieåret 1/5 1967 - 30/4 1968 som ble foretatt av Statistisk Sentralbyrå.

I kapittel I drøftes det hvilke variable som kan tenkes å inngå i en etterspørselsfunksjon for ferieturer.

Kapittel II gir et opplegg til en økonometrisk analyse.

Datamaterialet drøftes i kapittel II.

I kapittel IV presenteres og drøftes de økonometriske modeller som skal benyttes.

I kapittel V følger resultatene av analysen og kommentarer til disse.

## KAPITTEL I: ETTERSSPØRSEL ETTER FERIETURER. EN TEORETISK-ØKONOMISK DRØFTING AV PROBLEMSTILLINGEN

### 1. Etterspørselsfunksjoner

Etterspørselsfunksjoner blir utledet på bakgrunn av teorier for konsumentenes atferd.<sup>1)</sup> I det vi betrakter en enkelt konsument med en gitt preferansestruktur får vi at hans etterspørsel etter et gode i en periode kan skrives som en funksjon av prisene på alle goder og av inntekten:

$$(I.1) \quad q_i = f_i(p_1, p_2, \dots, p_n, r) \quad (i = 1, \dots, n)$$

hvor  $q_i$  betegner etterspurt mengde av gode  $i$ ,  $p_i$  er prisen på gode  $i$  og  $r$  konsumentens inntekt i perioden.

Som mål for virkningene på etterspørselen av en gitt partiell endring i ett av etterspørselsfunksjonens argumenter brukes vanligvis den partielle elastisitet.

Engel-elastisiteten er den partielle elastisitet av (I.1) m.h.p. inntekten

$$(I.2) \quad E_i = E_i(p_1, p_2, \dots, p_n, r) = \frac{\delta q_i}{\delta r} \frac{r}{q_i} \quad (p_1, \dots, p_n = \text{konstante})$$

Litt unøyaktig sier  $E_i$  med hvor mange prosent  $q_i$  endres om  $r$  endres partielt med 1 prosent.

1) Se f.eks. [4], [11], [12].

Cournot-elastisiteten for gode nr.  $i$  er:

$$(I.3) \quad \epsilon_{ij} = \epsilon_{ij}(p_1 \dots p_n, r) = \frac{\delta q_i}{\delta p_j} \cdot \frac{p_j}{q_i} \quad (\text{De \u00f8vrige priser og } r = \text{konst.})$$

$\epsilon_{ij}$  sier (litt upresist) med hvor mange prosent  $q_i$  endres om  $p_j$  endres partielt med 1 prosent.

Forutsetningen om gitt preferansestruktur regnes vanligvis ikke som realistisk i praktiske undersøkelser. Variable som antas \u00e5 p\u00e5virke denne trekkes derfor eksplisitt inn i ettersp\u00f8rselsfunksjonen.

I denne unders\u00f8kelsen er det husholdningen som forutsettes \u00e5 v\u00e8re desisjonsenheten. Det forutsettes alts\u00e5 at husholdningens atferd er karakterisert p\u00e5 tilsvarende m\u00e5te som atferden til en enkelt konsument. Dette er en ganske vanlig forutsetning i \u00f8konometriske undersøkelser, s\u00e5 jeg finner ingen grunn til noen n\u00e6rmere dr\u00f8fting her.

## 2. Ettersp\u00f8rselsfunksjon for ferieturer

Med ferietur vil vi forst\u00e5: "Opphold utenfor hel\u00e5rsboligen i helse- eller rekreasjons\u00f8yemed som inkluderer 4 eller flere overnattinger. Forretnings- og studiereiser, sykehus- og rekonvalesensopphold skal alts\u00e5 ikke tas med. Det samme gjelder skoleungdoms og studenters opphold hjemme hos foreldre i skoleferien".

Denne definisjonen ble brukt i Ferieunders\u00f8kelsen for ferie\u00e5ret 1/5 1967-30/4 1968, foretatt av Statistisk Sentralbyr\u00e5.

Alle prisene er forutsatt \u00e5 ha v\u00e8rt de samme for alle husholdninger. Som uttrykk for den "mengde" av godet den enkelte husholdning ettersp\u00f8r brukes husholdningens utgifter til posten ferieturer. Det kan selvsagt diskuteres om dette er et godt valg. Svaret vil for en stor del avhenge av hva det er vi spesielt er ute etter. I en \u00f8konomisk analyse m\u00e5 en vel kunne si at valget kan forsvares som et summarisk m\u00e5l for "mengden av ferieturer".

Det f\u00f8lgende er en liste over - og en diskusjon omkring - de variable jeg mener har betydning n\u00e5r det gjelder \u00e5 undersøke husholdningenes ettersp\u00f8rsel etter ferieturer. Siden prisene oppfattes som faste vil disse ikke inng\u00e5 i rekken av forklaringsvariable.

### Husholdningsinntekten

Det er vel ganske rimelig \u00e5 oppfatte denne som en viktig variabel i det den for de fleste legger beskrekninger p\u00e5 hvor mye husholdningen i alt kan skaffe seg av forskjellige goder. Til \u00e5 karakterisere inntektsavhengigheten

for et gode brukes Engel-elastisiteten. A priori vil det være rimelig å anta at Engel-elastisiteten for ferieturer er ganske høy, iallfall for middels og lavt inntektsnivå. På grunn av den vide definisjonen for godet ferietur er det imidlertid tvilsomt om inntektsvirkningen vil være så sterk at vi får en Engel-elastisitet større enn 1, eller m.a.o. om virkningen vil være så sterk at ferieturers budsjettandel vil øke med økende inntekt.

#### Husholdningsstørrelsen

Det er sannsynlig at husholdningsstørrelsen og dens sammensetning virker inn på etterspørselen. For matvarer er det nokså innlysende at etterspørselen øker når husholdningsstørrelsen øker. Det er atskillig mer usikkert i hvilken retning virkningen går når det gjelder ferieturer.

Ser vi på en partiell økning i husholdningsstørrelsen, vil det bety at husholdningen må foreta en omfordeling av sine kjøp av de forskjellige goder. Hvordan resultatet blir for ferieturenes vedkommende er det vanskelig å si noe om a priori.

En ting som vi ha betydning er hvordan husholdningsstørrelsen måles. Det vil neppe være rimelig å tillegge alle husholdningsmedlemmene like stor vekt.

Jeg kan tenke med to alternative måter å behandle dette problemet på. Den første måten går ut på å dele husholdningen i grupper etter alder og la hvert husholdningsmedlem telle 1 i den gruppen det tilhører. For eksempel kan husholdningen deles i to grupper, de under 16 år og de 16 år og over.

Det andre opplegget går ut på å benytte seg av såkalte dummy-variable. Det vil si at den dummy-variable har verdien 1 dersom det er et bestemt antall personer i husholdningen og 0 ellers. Hvor mange slike variable som skal innføres blir da et sentralt spørsmål. Dette opplegget forutsetter at endringer i husholdningsstørrelsen bare gir seg uttrykk i etterspørselen som faste nivåforskjeller mellom størrelsesgrupper.

#### Kvantitative variable

Det er utvilsomt en rekke variable av kvalitativ art som virker inn på folks preferanser. I denne analysen vil jeg ta med: Bosted, hovedpersonens yrke og utdanningsnivå, om husholdningen eier eller disponerer feriested eller fremkomstmiddel.

Jeg vil undersøke om det kan sies å være nivåforskjeller som skyldes disse kjennetegn.

Når det gjelder utdanning er min a priori-antakelse at folk med høyere utdanning bruker mer til ferieturer. F.eks. vil slike ting som språkvanser i utlandet være en mindre hindring for dem med høyere utdannelse.

Jeg vil også teste om det er noen forskjell mellom de husholdninger som eier eller disponerer feriested og de som ikke gjør det. Det er rimelig å anta at de som har feriested vil reise oftere på ferie enn de andre. På den andre siden vil det trekke i motsatt retning at vårt mengdemål, som er utgifter, ikke tar hensyn til at ferieturene blir billigere for disse. De utgifter husholdningen ville hatt dersom de skulle leid et tilsvarende sted måtte eventuelt ha vært anslått. Hva total-virkningen vil bli er vanskelig å si.

For de med fremkomstmiddel kan kombinasjoner av disse kjennetegn sammen med inntekten, gi forskjellig verdi på Engel-elastisiteten.

Når det gjelder den nærmere gruppering av kjennetegnene henviser jeg til variabel-listen. Grupperingen følger stort sett den som ble gjort i Ferieundersøkelsen.

Hypotesen m.h.t. bosted er at husholdningenes ferieturutgifter vil øke med graden av boligtetthet. Dette er f.eks. ut fra en antakelse om at rekreasjonsbehovet er større for folk jo tettere de bor.

Det jeg har kalt for hovedpersonens yrke ville det muligens vært riktigere å kalle hovedpersonens økonomiske aktivitet. Jeg er spesielt interessert i å undersøke en hypotese om at selvstendige, og da særlig selvstendige i jordbruk, skogbruk, fiske; bruker mindre til ferie enn ansatte.

## KAPITTEL II: OPPLÉGG TIL DEN ØKONOMETRISKE ANALYSEN

### 1. Presisering av de variable

Det følgende er en liste over de variable som skal inngå i etter-spørselsfunksjonen. Ved siden av det symbol som vil bli brukt følger en kort forklaring om den enkelte variable. Hvordan de forskjellige variable konkret skal inngå i funksjonen kommer jeg tilbake til i et senere avsnitt. Alle variable refererer seg til én husholdning.

Avhengige variable:

$q_F$  : Husholdningens totale ferieturutgifter i ferieåret 1/5-67 - 30/4-68, målt i kr. Dvs. de totale utgifter mens husholdningens medlemmer har vært på ferietur. Utgifter til utstyr og annet som er kjøpt før ferieturen tok til, er ikke tatt med.

$$Y_1 = \begin{cases} 1: & \text{Dersom et eller flere husholdningsmedlemmer har vært på ferietur} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Y_2 = \ln q_F \quad (\text{naturlig logaritme})$$

#### Forklaringsvariable

$$X_1 : \text{ Antall personer i husholdningen 16 år og over pr. 1/5-67.}$$

$$X_2 : \text{ Antall personer i husholdningen under 16 år pr. 1/5-67.}$$

$$X_3 : \text{ Husholdningens samlede antatte inntekt i skatteåret 1966, målt i 100 kr.}$$

$$Z_1 = \begin{cases} 1 & \text{Hvis bosatt i "tettbygd ellers"} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_2 = \begin{cases} 1 & \text{Hvis bosatt i "spredtbygd strøk"} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_3 = \begin{cases} 1 & \text{Hvis hovedpersonen (Hp) selvstendig i jordbruk, skogbruk, fiske} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_4 = \begin{cases} 1 & \text{Hvis Hp selvstendig i andre næringer} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_5 = \begin{cases} 1 & \text{Hvis Hp student eller skoleelev} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_6 = \begin{cases} 1 & \text{Hvis Hp "Annet" (bl.a. pensjonist, husmor)} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_7 = \begin{cases} 1 & \text{Hvis Hp realskole eller tilsvarende} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_8 = \begin{cases} 1 & \text{Hvis Hp artium eller tilsvarende} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_9 = \begin{cases} 1 & \text{Hvis Hp høgskole eller universitetseksamen} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{10} = \begin{cases} 1 & \text{Hvis husholdningen eier eller disponerer feriested} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{11} = \begin{cases} 1 & \text{Hvis husholdningen eier eller disponerer bil, campingvogn,} \\ & \text{motorsykkkel, motorbåt, seilbåt} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{12} = \begin{cases} 1 & \text{Hvis 2 personer i husholdningen} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{13} = \begin{cases} 1 & \text{Hvis 3 personer i husholdningen} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{14} = \begin{cases} 1 & \text{Hvis 4 personer i husholdningen} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{15} = \begin{cases} 1 & \text{Hvis 5 personer i husholdningen} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{16} = \begin{cases} 1 & \text{Hvis 6 eller flere personer i husholdningen} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{17} = \begin{cases} 1 & \text{Hvis husholdningsinntekten kr. 12 000 - 24 900} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{18} = \begin{cases} 1 & \text{Hvis husholdningsinntekten kr. 25 000 - 39 900} \\ 0 & \text{Hvis ikke} \end{cases}$$

$$Z_{19} = \begin{cases} 1 & \text{Hvis husholdningsinntekten kr. 40 000 eller over} \\ 0 & \text{Hvis ikke} \end{cases}$$

Det er verd å merke seg at inntektsoppgavene refererer seg til året 1966, mens utgiftene til feieturer refererer seg til tidsrommet 1/5-67 til 30/4-68. Det er en alvorlig begrensing at inntektene og utgiftene ikke refererer seg til samme periode. Vi våger likevel å benytte oss av inntekten fra 1966, idet vi betrakter den som en grov indikator for nivået av inntekten i det etterfølgende tidsrom.

Andre svakheter ved å benytte tall fra skattelikningen er blitt så ofte diskutert at jeg ikke finner det nødvendig med noen diskusjon her.

## 2. Valg av analysemetode

I hovedsaken vil jeg benytte meg av multippel regresjonsanalyse. (Se f.eks. [14] bind III.) Ved at jeg vil innføre såkalte dummy-variable får analysen også et variansanalyttisk aspekt.

Jeg har tenkt å benytte meg av to forskjellige opplegg. Det første er en vanlig regresjonsmodell der det forutsettes en lineær sammenheng mellom den absolutte verdi av ferieturutgiftene og et sett av forklaringsvariable. Modellen er lineær i den forstand at den er lineær i koeffisientene.

Det andre opplegget følger en metode som er utarbeidd av G. Orcutt m.fl. (Se [9].) Metoden er forøvrig også omtalt i [5].

I korthet går den ut på følgende:

Materialet deles i to grupper.

1. De husholdninger der et eller flere medlemmer har vært på ferietur.
2. De husholdninger der det ikke har vært noen på ferietur.

En er så interessert i å si noe om i hvilken grad forskjellige kjennetegn er bestemmende for om folk drar på ferie eller ikke. Dette analyseres ved at en benytter bare dummy-variable. Med andre ord at den avhengige variable har verdien 1 dersom et eller flere husholdningsmedlemmer har vært på ferietur og 0 ellers. På denne måten får en estimert en slags betinget sannsynlighet.

For den del av samplet som har vært på ferietur analyserer en så sammenhengen mellom verdien av den avhengige variable og et sett av forklaringsvariable.

En mer konkret drøfting av metodene vil bli foretatt i et senere kapittel.

### KAPITTEL III: DATAMATERIALET

Bakgrunns materialet for analysen av en landsomfattende intervjuundersøkelse som ble foretatt av Statistisk Sentralbyrå i tiden 1.- 31. mai 1968. Det skulle i alt intervjues et utvalg på ca. 3 000 husholdninger. Tidsrommet det ble samlet inn opplysninger for var ferieåret 1/5 1967 - 30/4 1968.

Det ble benyttet to typer skjemaer:

- a. Skjema for husholdningsintervju
- b. Skjema for ferietur

Se forøvrig vedlegg.

#### 1. Frafall

Av de 2 999 husholdninger som ble oppsøkt var det 306, eller 10,20 prosent som det ikke ble oppnådd intervju med. Det som har interesse er hvordan frafallet har innvirket på materialets representativitet.

Det viste seg at frafallet var større i tettbygde enn i tettbygde strøk. Spesielt var frafallet stort i Oslo. Videre hadde én-personshusholdninger større frafall enn fler-personshusholdninger.

Dette betyr at en kan regne med å få skjeve resultater på grunn av frafallet.

#### 2. Målefeil

I en intervjuundersøkelse kan en si at måleinstrumentet er spørsmål. Hvor gode mål vi skal få vil da i en viss grad avhenge av hvor gode spørsmålene er. De spørsmål i Ferieundersøkelsen som har relevans til denne oppgaven anser jeg som tilfredsstillende.

Grunnlaget for å vurdere målefeilene er de variable en er interessert i for et bestemt formål. Av de variable jeg skal benytte er det spesielt to som kan tenkes å være beheftet med målefeil. Det er den avhengige variable, husholdningens totale ferieturutgifter og husholdningsinntekten.



Graden av målefeil for ferieturutgiftene er vanskelig å vurdere. Som regel var det lenge siden turen ble foretatt og de færreste hadde noe nøyaktig regnskap for utgiftene. For å peile inn beløpet mest mulig nøyaktig ble det stilt spørsmål av følgende type: "Var utgiftene over x kr.?" "Under y kr.?"

At det er muligheter for målefeil ved at en benytter inntektsdata fra skattestatistikken har jeg nevnt tidligere.

#### KAPITTEL IV: REGRESJONSANALYSEN

##### 1. Modellene

Det vil bli benyttet to modeller. Modellene må ses på som alternative. Noen konkret test for å se hvilken som er best kan ikke gjennomføres. For eksempel kan en ikke sammenligne den multiple regresjonskoeffisient fordi modellene har forskjellig venstreside variabel. Modellene kan uttrykkes ved følgende relasjoner: (Kfr. variabellisten.)

##### Modell I.

$$q_F = \alpha_0 + \alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 z_3 + \alpha_4 z_4 + \alpha_5 z_5 + \alpha_6 z_6 + \alpha_7 z_7 + \alpha_8 z_8 + \alpha_9 z_9 + \alpha_{10} z_{10} + \\ \alpha_{11} z_{11} + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 z_4 x_3 + \beta_5 z_2 x_3 + \beta_6 z_3 x_3 + \beta_7 z_4 x_3 + \beta_8 z_5 x_3 + \\ \beta_9 z_6 x_3 + \beta_{10} z_7 x_3 + \beta_{11} z_8 x_3 + \beta_{12} z_9 x_3 + \beta_{13} z_{10} x_3 + \beta_{14} z_{11} x_3 + \beta_{15} x_3^2 + u$$

Forutsetninger:

- (1)  $E(u_i) = 0$ ,  $i = 1, 2, \dots, n$  ( $n =$  antall husholdninger)
- (2)  $E(u_i u_j) = \begin{cases} 0 & \text{for } i \neq j \\ \delta^2 & \text{for } i = j \end{cases}$   $i, j = 1, 2, \dots, n$

##### Modell II.

$$(a) Y_1 = \gamma_0 + \gamma_1 z_1 + \gamma_2 z_2 + \gamma_3 z_3 + \gamma_4 z_4 + \gamma_5 z_5 + \gamma_6 z_6 + \gamma_7 z_7 + \gamma_8 z_8 + \gamma_9 z_9 + \\ \gamma_{10} z_{10} + \gamma_{11} z_{11} + \gamma_{12} z_{12} + \gamma_{13} z_{13} + \gamma_{14} z_{14} + \gamma_{15} z_{15} + \gamma_{16} z_{16} + \\ \gamma_{17} z_{17} + \gamma_{18} z_{18} + \gamma_{19} z_{19} + w$$

Forutsetning:  $E w_i = 0$  for alle  $i$ .

$$(b) Y_2 = a_0 + a_1 z_1 + a_2 z_2 + a_3 z_3 + a_4 z_4 + a_5 z_5 + a_6 z_6 + a_7 z_7 + a_8 z_8 + a_9 z_9 + a_{10} z_{10} + \\ a_{11} z_{11} + a_{12} z_{12} + a_{13} z_{13} + a_{14} z_{14} + a_{15} z_{15} + a_{16} z_{16} + b_1 \ln x_3 + b_2 (\ln x_3)^2 + v$$

Forutsetninger:

- (1)  $E(v_i) = 0$  for alle  $i$
- (2)  $E(v_i v_j) = \begin{cases} 0 & \text{for } i \neq j \\ \delta^2 & \text{for } i = j \end{cases}$   $i, j = 1, 2, \dots, n$

## 2. Den økonomiske mening med relasjonene

Analysen bygger på et cross-section materiale slik at prisene oppfattes som faste. Modell I sier da at etterspørselen etter ferieturer, representert ved ferieturutgiftene er en funksjon av inntekten, husholdningsstørrelsen og en rekke kvalitative variable. Funksjonen er lineær i  $\alpha$ -ene og  $\beta$ -ene. Den er imidlertid ikke lineær i inntekten ( $x_3$ ). Det tas også hensyn til at det kan være nivåforskjeller i kosnumtilbøyeligheten (etterspørselstilbøyeligheten). Dette er gjort ved at relasjonen har 11 samspillvariable.  $z_1x_3, z_2x_3, \dots, z_{11}x_3$ .

Modell II består av to relasjoner. Det grunnleggende ved denne modellen er at den foretar en oppdeling av materialet av de husholdninger som har vært og de som ikke har vært på ferietur. Ved den første relasjonen postulerer en en slags betinget sannsynlighet for at en husholdning har vært på ferietur. Den andre relasjonen er en mer "vanlig" etterspørselsrelasjon. Det er ingen prinsipiell forskjell på denne relasjonen og relasjonen i modell I. Konkret er det den forskjell at i relasjon II.b. brukes det logaritmiske verdier for den avhengige variable og for inntekten. Videre er det ingen samspillvariable med og husholdningsstørrelsen er representert ved dummy-variable.

Det kan selvsagt diskuteres hvor relevante disse relasjonene er. De må uansett bare ses på som approksimasjoner og ikke som uttrykk for sammenhenger som er sanne i en eller annen absolutt forstand. Problemer i forbindelse med spesifikasjonsfeil i økonomiske relasjoner er bl.a. behandlet av H. Theil [13].

Det at vi kan ha spesifikasjonsfeil i modellene får konsekvens for den vekt vi kan tillegge de konklusjoner estimeringer på modellene gir.

## 3. Diskusjon av relasjonenes statistiske egenskaper

Dersom det er målefeil i de forklaringsvariable i tillegg til spesifikasjonsfeil i modellen vil en få forventningsskjevne estimater på koeffisientene ved minste kvadraters metode. Estimaten vil heller ikke være konsistente. Se [5] Ch. 6 eller [8] Ch. 10.

Det er nevnt tidligere at spesielt for husholdningsinntekten må en regne med målefeil. I [8] s. 258-260 blir det vist at den asymptotiske feil en får på estimatene blir mindre jo lavere korrelasjonen mellom restleddet og den eksogene variable er og jo mindre forholdet er mellom variansen for restleddet og variansen for den eksogene variable. Ved positiv korrelasjon mellom restleddet og den eksogene variable vil en få et underestimat på koeffisienten.

I relasjonene I og II.b. er det forutsatt konstant varians på restleddet. Dette må sies å være en ganske hard forutsetning. En variabel som

husholdningsinntekten har et stort variasjonsområde. Vanligvis vil restleddsvariansen øke med økende inntekt. Se f.eks. [10] s. 55. Ved vanlig minste kvadraters estimering på en lineær funksjon vil en få forventningsriktige estimater for relasjonens koeffisienter selv om det ikke er konstant restledd varians. Estimaten vil imidlertid ikke være effesiente, det vil si at de har ikke minimum varians.

Dersom en antar at den relasjonen som gjelder er lineær i logaritmene, f.eks. II.b. er en forutsetning om konstant varians mer plausibel. Se [5] s. 207-211 eller [10] s. 48-50 og s. 55.

Relasjon II.a. består bare av dummy-variable. Minste kvadraters estimatene i regresjonen av  $Y_1$  m.h.p. de dummy-variable vil da være den gjennomsnittlige ferieturhyppighet i hver gruppe.

En svakhet ved metoden er at det er umulig å forutsette konstant varians på restleddet. Dette er vist i [3] s. 249. Det blir her foreslått en to-trinns prosedyre for å oppnå de beste forventningsrette estimater. Jeg skal ikke benytte denne her. De resultatene jeg får må derfor tolkes med en viss forsiktighet.

Null-observasjoner skaper problemer når det brukes logaritmiske variable. I denne analysen er problemet aktuelt bare når det gjelder husholdningsinntekten. Jeg har gjort det slik at null-inntekt er erstattet med 100 kr. Siden inntekten måles i 100 kr. betyr dette at logaritmen settes lik null. (Gjelder bare relasjon II.b.)

## KAPITTEL V: RESULTATER AV REGRESJONSBEREGNINGENE. KOMMENTARER

### 1. Resultater

Det følgende er resultatene av regresjonsberegningene oppstilt i tabellform.

Tabell V.1. Resultater Modell I.

Variabel	Regresjons- koeffisient	Standardavvik regresjonskoeffisient	t-verdi
X <sub>1</sub> .....	76,7	25,8	2,98
X <sub>2</sub> .....	4,1	19,4	0,21
X <sub>3</sub> .....	1,5	0,6	2,37
X <sub>3</sub> <sup>2</sup> .....	0,0	0,0	3,75
Z <sub>1</sub> .....	-200,5	108,0	-1,86
Z <sub>2</sub> .....	-160,5	106,3	-1,51
Z <sub>3</sub> .....	-329,5	138,9	-2,37
Z <sub>4</sub> .....	-49,1	153,2	-0,32
Z <sub>5</sub> .....	-821,0	427,0	-1,92
Z <sub>6</sub> .....	-222,3	107,8	-2,06
Z <sub>7</sub> .....	225,6	122,6	1,84
Z <sub>8</sub> .....	1 001,3	204,9	4,89
Z <sub>9</sub> .....	1 060,4	268,0	3,96
Z <sub>10</sub> .....	297,5	108,5	2,74
Z <sub>11</sub> .....	-41,4	96,7	-0,43
Z <sub>1</sub> X <sub>3</sub> .....	-1,0	0,4	-2,60
Z <sub>2</sub> X <sub>3</sub> .....	-2,8	0,4	-6,94
Z <sub>3</sub> X <sub>3</sub> .....	0,1	0,6	0,13
Z <sub>4</sub> X <sub>3</sub> .....	0,0	0,5	0,07
Z <sub>5</sub> X <sub>3</sub> .....	3,9	3,7	1,06
Z <sub>6</sub> X <sub>3</sub> .....	-0,2	0,5	-0,44
Z <sub>7</sub> X <sub>3</sub> .....	0,5	0,4	1,24
Z <sub>8</sub> X <sub>3</sub> .....	-1,4	0,6	-2,34
Z <sub>9</sub> X <sub>3</sub> .....	-0,7	0,6	-1,16
Z <sub>10</sub> X <sub>3</sub> .....	-0,4	0,4	-0,97
Z <sub>11</sub> X <sub>3</sub> .....	0,5	0,4	1,13

Konstantledd: 481,0

Multippel korrelasjonskoeffisient: 0,54144

F-verdi: 42,5      X<sup>2</sup>-observator: 1 849,2

Tabell V.2. Resultater Modell II. Regresjonsligning a.

Variabel	Regresjons- koeffisient	Standardavvik regresjonskoeffisient	t-verdi
Z <sub>1</sub> .....	-0,076	0,025	-3,09
Z <sub>2</sub> .....	-0,272	0,026	-10,63
Z <sub>3</sub> .....	-0,281	0,029	-9,84
Z <sub>4</sub> .....	-0,134	0,030	-4,46
Z <sub>5</sub> .....	0,134	0,115	1,16
Z <sub>6</sub> .....	-0,200	0,027	-7,47
Z <sub>7</sub> .....	0,085	0,024	3,60
Z <sub>8</sub> .....	0,170	0,037	4,60
Z <sub>9</sub> .....	0,188	0,049	3,84
Z <sub>10</sub> .....	0,156	0,021	7,30
Z <sub>11</sub> .....	0,112	0,020	5,66
Z <sub>12</sub> .....	-0,068	0,025	-2,71
Z <sub>13</sub> .....	-0,108	0,029	-3,75
Z <sub>14</sub> .....	-0,075	0,030	-2,48
Z <sub>15</sub> .....	-0,117	0,035	-3,33
Z <sub>16</sub> .....	-0,130	0,040	-3,23
Z <sub>17</sub> .....	-0,026	0,026	1,01
Z <sub>18</sub> .....	0,104	0,030	3,46
Z <sub>19</sub> .....	0,066	0,038	1,75

Konstantledd: 0,658

Multippel korrelasjonskoeffisient: 0,52842

F-verdi: 54,5       $X^2$ -observator: 212,7

Tabell V.3. Resultater Modell II. Regresjonsligning b.

Variabel	Regresjons- koeffisient	Standradavvik regresjonskoeffisient	t-verdi
$\ln X_3$ .....	-0,409	0,064	-6,39
$(\ln X_3)^2$ .....	0,083	0,010	7,88
$Z_1$ .....	-0,347	0,061	-5,69
$Z_2$ .....	-0,553	0,071	-7,79
$Z_3$ .....	-0,255	0,133	-1,91
$Z_4$ .....	0,041	0,084	0,49
$Z_5$ .....	-1,008	0,250	-4,02
$Z_6$ .....	-0,296	0,095	-3,11
$Z_7$ .....	0,235	0,061	3,86
$Z_8$ .....	0,332	0,085	3,90
$Z_9$ .....	0,443	0,111	3,98
$Z_{10}$ .....	0,073	0,053	1,36
$Z_{11}$ .....	0,022	0,061	0,36
$Z_{12}$ .....	0,125	0,075	1,67
$Z_{13}$ .....	0,265	0,085	3,10
$Z_{14}$ .....	0,230	0,086	2,68
$Z_{15}$ .....	0,306	0,102	3,00
$Z_{16}$ .....	0,207	0,124	1,67

Konstantledd: 6,636

Multiplere korrelasjonskoeffisient: 0,51973

F-verdi: 27,0  $\chi^2$ -observator: 58,7

## 2. Tester

Det var i alt 2 693 observasjoner. (Husholdninger.) Av disse hadde 1 334 hatt utgifter til ferietur.

Av en tabell over t-fordelingen (se [7]) finner en at en forkaster en hypotese om at en regresjonskoeffisient er null dersom t-observatoren er større enn 1,96 i tallverdi når signifikansnivået skal være 0,05. Med et signifikansnivå på 0,10 er den kritiske verdi 1,64.

Av en tabell over F-fordelingen (se [7]) finner en at en forkaster hypotesen om at alle koeffisientene er null dersom, i Modell I, F-verdien er større enn 1,50, i Modell II, regresjon a, F-verdien større enn 1,62 og i Modell II, regresjon b, F-verdien større enn 1,63.

Av en tabell over  $X^2$ -fordelingen, (se [7]) finner en at hypotesen om at restleddene er normalfordelte, forkastes dersom  $X^2$  observatoren er større enn 22,36.

Det fremgår av tabellene at for alle regresjonsligningene må hypotesen over normalfordelte restledd forkastes. Spørsmålet er så hvilken konsekvens dette får for t- og F-testene. Disse er utledet under den forutsetning at restleddene er normalfordelte. Problemet er behandlet i [8] Ch. 8 § 4. Konklusjonen er at t- og F-testene er lite følsomme overfor avvik fra normalitetsforutsetningen. Det skulle derfor ikke være noe i veien for å benytte seg av dem. Testene er muligens litt tvilsomme for Modell II.a. hvor det er varierende residualspredning.

Ser vi først på F-testen finner vi at for alle regresjonsligningene forkaster vi hypotesen  $H_0: \beta_j = 0$  for alle j, mot  $H_1: \beta_j \neq 0$  for unntatt én j verdi. Dette er a priori ganske opplagt.

Av t-testen finner en at i modell I er bare 11 av de 26 variable signifikant forskjellig fra null med et nivå på 5 prosent. Med et nivå på 10 prosent er 14 variable signifikante. De 11 variable er:

$X_1$  : Antall personer i husholdningen 16 år og over

$X_3$  : Husholdningsinntekten

$X_3^2$  : Husholdningsinntekten kvadrert

$Z_3$  : Hp selvstendig i jordbruk, skogbruk, fiske

$Z_6$  : Hp annet (bl.a. husmor, pensjonist)

$Z_8$  : Hp artium

$Z_9$  : Hp universitets- eller høyskoleeksamen

$Z_{10}$ : Husholdningen eier eller disponerer feriested

$Z_1 X_3$ : Bosatt i tettbygd ellers/inntekt

$Z_2 X_3$ : Bosatt i spredtbygd strøk/inntekt

$Z_8 X_3$ : Artium/inntekt

De tre variable som kommer i tillegg ved et 10 prosent nivå er:

$Z_1$ : Bosatt i tettbygd strøk ellers

$Z_5$ : Hp student eller skoleelev

$Z_7$ : Hp realskole eller tilsvarende

Modell II, regresjonsligning a. Med et signifikansnivå på 5 prosent er koeffisientene foran 3 av de 19 variable ikke signifikante.

De tre variable er:

$Z_5$ : Hp student eller skoleelev

$Z_{17}$ : Husholdningsinntekten i klassen 12 000 - 24 900

$Z_{19}$ : Husholdningsinntekter i klassen 40 000 og over

$Z_{19}$  ville vært signifikant med et nivå på 10 prosent.

Modell II, regresjonsligning b. Av de 18 variable er det 6 som har ikke-signifikante koeffisienter ved et 5 prosent nivå. Disse er:

$Z_3$ : Selvstendige i jordbruk, skogbruk, fiske

$Z_4$ : Andre selvstendige

$Z_{10}$ : Husholdningen eier eller disponerer feriested

$Z_{11}$ : Husholdningen eier eller disponerer fremkomstmiddel

$Z_{12}$ : 2 personer i husholdningen

$Z_{16}$ : 6 eller flere personer i husholdningen

Med et nivå på 10 prosent er det bare  $Z_4$ ,  $Z_{10}$  og  $Z_{11}$  som ikke er signifikante.

### 3. Drøfting av fortegnene på koeffisientene

I kapittel I satte jeg fram en del a priori hypoteser om fortegnene på koeffisientene. Jeg skal nå jevnføre disse med resultatene fra regresjonsberegningene.

Når det gjelder regresjonen i modell I, så vi at svært mange av koeffisientene ikke var signifikant forskjellige fra null. I det hele tatt



må en si at de variable jeg har tatt med ikke har så veldig stor forklaringskraft når det gjelder variasjonen i ferieturutgiftene. Den multiple korrelasjonskoeffisient var 0,54.

Ser vi på de enkelte variable finner en at for husholdningsstørrelsen er det bare koeffisienten foran den variable som representerer antall personer over 16 år som er signifikant. Koeffisienten er som ventet positiv. Koeffisientene foran inntekten og inntekten kvadrert er begge positive. Det betyr at ferieturutgiftene øker med inntekten og at økningen er større jo større inntekten er.

De variable som representerer bostedstype,  $Z_1$  og  $Z_2$  har begge negative koeffisienter; altså en nivåsenkning av utgiftene i forhold til basisgruppen "bosatt i storby". Ingen av koeffisientene er imidlertid signifikante med et nivå på 5 prosent.

For variablene for yrke,  $Z_3$ -  $Z_6$ , hvor basisgruppen er "ansatt", ser en at alle koeffisientene er negative. Bare for "selvstendig i jordbruk osv." og "annet" er koeffisientene signifikante. Resultatene stemmer ellers noenlunde bra med den a priori antakelse.

Når det gjelder utdanning, variablene  $Z_7$ -  $Z_9$ , basis folkeskole, ser en at det er en markert heving av utgiftsnivået når utdanningen øker. Dette var for så vidt ventet, selv om jeg ikke hadde ventet en så markert økning.

Ellers ser det ut til at det å ha feriested øker ferieturutgiftene. Dette er vel rimelig siden hele materialet benyttes i denne regresjonen. Det å ha bil e.l. gir imidlertid ikke noe signifikant utslag på ferieturutgiftene.

Av de 11 kryssleddene er bare 3 signifikante. Stort sett ser det altså ut til at virkningen av en inntektsendring er nokså lik for alle grupper. De tre variable med signifikante koeffisienter er de to for bosted/inntekt, begge negative, og artium/inntekt, også negativ.

Vi går nå over til å se på modell II, regresjonsligning a. Begge bostedsvariablene har negative koeffisienter. Spesielt er nedgangen i sannsynligheten for at en husholdning har vært på ferietur stor dersom den er bosatt i spredtbygd strøk. Estimaten er en nedgang i forhold til basisgruppen på ca. 27,2 prosent.

For yrkesvariablene finner vi et mønster omtrent som vi kunne vente. Nedgangen er særlig stor dersom Hp er selvstendig i jordbruk, skogbruk, fiske, ca. 28,1 prosent.

Når det gjelder utdanning øker sannsynligheten med økt utdanningsnivå. Det er også en økning dersom husholdningen eier eller disponerer feriested og fremkomstmidler.

Den konklusjon en kan trekke ved å se på koeffisientene foran husholdningsstørrelsesvariablene,  $Z_{12}$  -  $Z_{16}$  ser ikke ut til å være annet enn at det er en nedgang for flerpersongshusholdninger i forhold til enpersons. Det ser ikke ut til å ha vært noen vits i å foreta en så detaljert oppdeling.

Vi så tidligere at bare en av variablene for inntektsklasse hadde signifikant koeffisient ved et 5 prosent nivå. Med en viss usikkerhet kan vi si at sannsynligheten for ferietur først øker når husholdningen er kommet opp i inntektsklassen 25 000 - 39 900. Tjener husholdningen mer enn 40 000 ser det ut for at sannsynligheten for ferietur går ned igjen. Alt i alt ser det ut til at det er andre ting enn inntekten som forklarer om en husholdning reiser på ferietur eller ikke.

Regresjonsligning II.b. Ferieturutgiftene når det er gitt at husholdningen har vært på ferietur. Det første en legger merke til er at koeffisienten foran førstegradsleddet av logaritmen til inntekten er negativ. Imidlertid er annengradsleddet ganske sterkt positivt slik at det bare vil være for helt lave inntekter at virkningen av en inntektsendring vil være negativ. Denne inntekt er så lav at det vil være tvilsomt om det er riktig å trekke funksjonens gyldighetsområde så langt. Se forøvrig avsnittet om Engel-elastisitetene nedenfor.

En ser at bostedsstrukturen virker inn på utgiftsnivået på den måten at nivået blir lavere jo mindre befolkningstettheten er.

For yrkesvariablene er resultatet mer usikkert. Det er en markert nedgang dersom  $H_p$  er "student eller skoleelev". Nedgang også om  $H_p$  tilhører gruppen "annet".

Når det gjelder utdanningsnivå er det en signifikant økning i utgiftsnivået med økt utdanningsnivå.

Om husholdningen eier eller disponerer feriested eller fremkomstmiddel har ikke noe signifikant utslag på utgiftsnivået.

Bortsett fra at det er positivt fortegn for alle koeffisientene foran husholdningsstørrelsesvariablene er det vanskelig å peke på noe bestemt mønster. Det kan også her se ut som om oppdelingen har vært for detaljert.

#### 4. Virkning på ferieturutgiftene ved endring i husholdningsinntekten

I begge regresjonsligningene hvor utgiften var avhengig variabel fant vi at koeffisientene foran inntektsvariablene var forskjellig fra null. For å studere inntektsvirkningen bruker vi Engel-elastisiteten. Se ligning (I.2.)

Av modell I får vi at estimatet for ferieturutgiftene blir ca. kr. 875 når husholdningen tilhører basisgruppen (alle Z-ene lik null), har en inntekt lik gjennomsnittet (kr. 21 250) og består av to personer over 16 år og en under 16 år. Estimaten på Engel-elastisiteten blir da:

$$(V.1.) E_{qF} = (1,496 + 2 \cdot 0,003 \cdot 212,5) \frac{212,5}{875} \approx \underline{\underline{0,67}}$$

Engel-elastisiteten for modell II.b. blir:

$$(V.2.) E_{y2} = -0,409 + 2 \cdot 0,083 \ln x_3$$

Ved en inntekt lik gjennomsnittsinntekten i samplet (1 334 husholdninger), kr. 25 700 finner vi en Engel-elastisitet tilnærmet lik 0,51. Dersom inntekten er lik gjennomsnittet i hele samplet, dvs. kr. 21 250 blir Engel-elastisiteten tilnærmet lik 0,48.

Det kan være av interesse å se hvor stor inntekten må være for at Engel-elastisiteten skal være større enn null. Ved å sette inn for  $E_{y2} = 0$  finner vi at  $\log x_3$  må være 2,464. Dette tilsvarer en inntekt på knapt 1 200 kr. Det er neppe riktig å anta at funksjonens gyldighetsområde dekker så lave inntekter.

Ellers kan en vel si at Engel-elastisiteten er forbausende lav. Imidlertid må en som tidligere nevnt ta i betraktning den vide definisjonen begrepet ferietur har fått. Bl.a. er besøk hos slektninger med. Det er rimelig å anta at slike turer vil falle billigere enn mer "vanlige" ferieturer.

## 5. Sluttord

Det må vel innrømmes at de relasjonene jeg har brukt ikke har noen overvettets stor forklaringskraft når det gjelder variasjonen i ferieturutgiftene. Begge relasjonene hvor den avhengige variabel er ferieturutgiften har en multippel korrelasjonskoeffisient som bare er litt over 0,5. Det må forøvrig innskytes at relasjonene ikke er sammenlignbare med basis i denne koeffisienten. For det første er datagrunnlaget forskjellig og for det andre er den avhengige variable forskjellig i det det i modell II er brukt logaritmer. Tross alt mener jeg likevel det har kommet frem en del ting som har krav på interesse.

## L i t t e r a t u r

- [ 1 ] Aitchison, J. & Brown, J.A.C.: A Synthesis of Engel Curve Theory. The Review of Economic Studies; Vol. XXII 1954-55.
- [ 2 ] Amundsen, A.: Metoder i analysen av forbruksdata. Statistisk Sentralbyrå, Artikler nr. 6, 2. opplag 1968
- [ 3 ] Goldberger, A.S.: Econometric Theory, John Wiley & Sons Inc. 1964
- [ 4 ] Henderson, J.M. & Quandt, R.E.: Micro-economic Theory. McGraw-Hill 1958
- [ 5 ] Johnston, J.: Econometric Methods McGraw-Hill 1963
- [ 6 ] Klein, L.R.: An Introduction to Econometrics. Prentice Hall Inc. 1962
- [ 7 ] Lindley, D.V. & Miller, J.C.P.: Cambridge Elementary Statistical Tables. Cambridge University Press 1964
- [ 8 ] Malinvaud, E.: Statistical Methods of Econometrics. North Holland 1968
- [ 9 ] Orcutt, G.H., Greenberger, M., Korbel, J., Rivlin, A.M.: Microanalysis of Socio-economic System. Harper & Brothers 1961
- [10] Prais, S.J. & Houthakker, H.S.: The Analysis of Family Budgets. Cambridge 1955
- [11] Samuelson, P.A.: Foundations of Economic Analysis. Harvard University Press 1947
- [12] Serck-Hanssen, J.: Teorier for konsumentenes atferd. Universitetsforlaget 1964
- [13] Theil, H.: Specification Errors and the Estimation of Economic Relationships. Revue de l'institute internationale de statistique. Vol. 25 1957
- [14] Thorén Amundsen, H.: Innføring i teoretisk statistikk. 3. bind. Universitetsforlaget 1962
- [15] Wold, H.: Demand Analysis. Uppsala 1952