Interne notater
STATISTISK SENTRALBYRÅ

86/23 5. mai 1986

EN LOGIT-MODELL FOR ANALYSE AV GIFTE KVINNERS ARBEIDSTILBUD

AV

KARI HOLST

FORORD

Ved å ta utgangspunkt i den tradisjonelle statiske teorien for tilbud av arbeid presenteres en økonometrisk modell for sannsynligheten for å arbeide. Denne utledningen er basert på en antatt klasse av nyttefunksjoner og på spesielle forutsetninger om de stokastiske variablene som inngår. Disse forutsetningene fører til at modellen blir på Logit-formen. Dette gir mulighet for sammenlikning av Logit og Probit-analyse basert på samme økonomiske modell.

Arbeidet som legges fram her, er en del av den forskning vedrørende tilbud av arbeid som pågår i Statistisk Sentralbyrå. Arbeidet er en spesialoppgave i det sosialøkonomiske embetsstudiet. Veiledere har vært John K. Dagsvik og Steinar Strøm. Vi takker for støtte fra Nordiska Skattevetenskapliga Forskningsråd og fra RFSP/NAVF.

INNHOLD

Kapittel 1. Innledning
Kapittel 2. Modell
Kapittel 3. Estimeringsmetode
Kapittel 4. Estimeringsresultater
Kapittel 5. Tilbudselastisiteter

Litteratur

1. INNLEDNING

Formålet med rapporten er å analysere gifte kvinners tilbud av arbeid på grunnlag av mikro-økonomisk teori og levekårsdata fra 1980/79. Vi ser bort fra eventuelle ledighetsproblemer ved å forutsette at observert sysselsetting er lik arbeidstilbudet. I oppgaven vil jeg drøfte teorien som ligger til grunn for en modell for sannsynligheten for å være sysselsatt. Denne modellen, som ofte benevnes Logit-modellen, blir sammenliknet med den tilsvarende Probit-modellen ved å beregne elastisiteter for begge modellene.

2. MODELL

Vi tar utgangspunkt i statisk konsument-teori, det betyr at vi ser på tilpasningen innenfor en gitt periode (her: et år) og uten at forhold av betydning for framtidige perioder trekkes inn. Vi ser bort fra sparing og lån, slik at kjøp av goder pr. tidsenhet ikke kan overskride nettoinntekten. Nettoinntekt defineres som inntekt etter skatt. I modellen vil beslutningen om tilbud av arbeid og etterspørsel etter varer og tjenester bli avledet fra nyttemaksimering under en gitt budsjettbeskrankning. Vi vil anta at de fire aksiomer fra konsument-teorien er oppfylt, og at individet (i vårt tilfelle den gifte kvinne) har en to ganger kontinuerlig deriverbar nytteindikator som er bestemt på en vilkårlig stigende transformasjon nær. Vi vil anta at kvinnens nytte avhenger av to desisjonsvariable:

- familiens samlede konsum
- kvinnens arbeidstid.

Vi innfører følgende symboler:

C = familiens samlede konsum pr. år, prisen på C (sammensatt gode) settes lik 1

U(C,h,S) = familiens nytteindikator

h = antallet av årets timer som kvinnen bruker til (lønnet) arbeid

S = en vektor med variable som karakteriserer kvinnens sosio-økonomiske bakgrunn

For å beskrive budsjettrestriksjonen, trenger vi følgende variable:

Årets totaltid er satt lik M

W = kvinnens timelønn

k = ektefellens bruttoinntekt, lik summen av lønns- og kapitalinntekter

t = familiens skatt, som avhenger av begges inntekter

Vi simulerer den gifte kvinnens tilpasning ved å anta at h bestemmes gjennom maksimering av familiens nytte gitt budsjettrestriksjon og gitt mannens bruttoinntekt. Vi ser altså bort fra den eventualitet at ektefellene foretar en simultan tilpasning. Dette er en framgangsmåte benyttet i et stort antall økonometriske arbeider av gifte kvinners arbeidstilbud, se Killingsworth (1983) og kan rettferdiggjøres gjennom den observasjon at menns arbeidstilbud er relativt uelastisk. Modellen er dermed

$\max_h U(C,h,S)$

gitt $C = wh + k - t(wh,k)$

der k = kjent og gitt

og $0 \leq h \leq 1$

Vi skal senere spesifisere formen på nytte- og skattefunksjon. Her skal vi nøye oss med å anta at nyttefunksjonen er ikke-avtakende i konsum, avtakende i arbeidstid og konkav. Skattefunksjonen forutsettes å være voksende i wh og k, samt konveks.

Første-ordens-betingelsene for nyttemaksimum gir at

$$(1) \qquad w(1-t'(w\bar{h},k)) \leq \frac{-U_h'(\bar{C},\bar{h},S)}{U_C'(\bar{C},\bar{h},S)}$$

hvor t' er den deriverte av t mhp. wh og hvor U_C' og U_h' er den deriverte av U mhp. C og h. (Vi har her sett bort fra eventuelle øvre skranker på arbeidstiden).

Uttrykket til høyre for ulikhetstegnet i (1) kalles den marginale substitusjonsbrøk, innsatt optimumsverdiene $\bar{C}$ og $\bar{h}$. Størrelsen på den marginale substitusjonsbrøk i punktet h=0 vil være individets reservasjonslønn. Med reservasjonslønn menes den laveste lønna individet vil akseptere for å være i arbeid.

Dette gir følgende desisjonsregel:

$$(2a) \qquad \text{Hvis } w(1-t'(0,k)) < \frac{-U_h'(C_0,0,S)}{U_C'(C_0,0,S)}$$

vil individet velge ikke å være i arbeid.

$$(2b) \qquad \text{Hvis } w(1-t'(0,k)) = \frac{-U_h'(C_0,0,S)}{U_c'(C_0,0,S)}$$

vil individet være indifferent til hvorvidt det er i arbeid eller ikke.

$$(2c) \qquad \text{Hvis } w(M-t'(0,k)) > \frac{-U_h'(C_0,0,S)}{U_c'(C_0,0,S)}$$

vil individet være i arbeid og tilpasse seg slik at 1.ordensbetingelsen er oppfylt med likhet.

C_0 er konsumet når h=0, det vil si $C_0 = k-t(0,k)$.

Individet vil bestemme simultant hvorvidt det ønsker å være sysselsatt eller ikke, og hvor mange timer det ønsker å arbeide. Som nevnt foran, ser vi bort fra eventuelle rasjoneringer i arbeidsmarkedet, noe som kan rettferdiggjøres ved den lave arbeidsledigheten i Norge i 1979/80. Dette er m.a.o. en modell med full sikkerhet og full informasjon fra beslutningstakers synspunkt, og "jobb/ikke jobb"-beslutningen samt "hvor mange timer skal jeg arbeide"-beslutningen er én og samme beslutning.

Nyttefunksjonen antas å ha følgende spesifiserte form:

$$(3) \qquad U(C,h) = \frac{A((1-h/M)^{\alpha}-1)}{\alpha} + \frac{B(C^{\beta}-1)}{\beta}$$

Denne funksjonen er to ganger kontinuerlig deriverbar, additiv separabel i konsum og fritid og strengt konkav når α og β begge er mindre enn 1. Funksjonen er avtagende i arbeidstid og voksende i konsum. Når α og β går mot 0, går U(C,h) mot A log(1-h) + B log C, som er en separabel translog nyttefunksjon i konsum og fritid.

Variabelen S, som inneholder informasjon om kvinnens sosio-økonomiske bakgrunn, er representert i størrelsene A og B. Dette vil bli mer utførlig kommentert senere.

Den spesifiserte skattefunksjonen er ment å representere en faktisk stykkevis lineær skattefunksjon, og den skal ivareta de gjeldende lovbestemte fradrag. Vi approksimerer den virkelig skattestruktur med en funksjon som er deriverbar og strengt konveks, se Offerdal og Strøm (1979). I henhold til norske skatteregler, liknes ektepar særskilt hvis begge ektefeller har en inntekt som overskrider 22 000.

Ved felles likning vil skattefunksjonen være

$$(4a) \qquad t(wh,k) = g(wh+k) = g(I)$$

der $I = wh + k$

Ved særskilt likning

(4b) $t(wh,k) = g(wh) + g(k)$

Som en tilnærming, lar vi funksjonsformen g være den samme enten det er felles eller særskilt likning. Vi merker oss at den deriverte av skattefunksjonene mhp. kvinnens lønnsinntekt er $g'(wh+k)$ ved felles likning og $g'(wh)$ ved særskilt likning. Den valgte g-formen er

(4c) $g(x) = 3{,}38 \cdot 10^{-4}(0{,}81x+6467)^{1{,}61} + 0{,}053x$

med justeringer for lave x slik at $g(0) = 0$.

Vi lar nå f^* stå for disponibel inntekt. Det vil si

$$f^* = f^*(wh,k) = wh + k - t(wh,k)$$

og fra budsjettbetingelsen har vi at

$$C = f^*(wh,k).$$

Når betingelsene for felles likning er oppfylt, har vi

(5a) $f(wh,k) = wh + k - g(wh+k) = f(wh+k)$
der $f(X) = 1 - g(X)$.

Ved særskilt likning har vi

(5b) $f^*(wh,k) = wh + k - g(wh) - g(k) = f(wh) + f(k)$.

De korresponderende deriverte av de disponible inntektsfunksjonene mhp. h er ved felles likning

(5c) $f_h^{*'} = w(1-g'(wh+k)) = wf'(wh+k)$

og ved særskilt likning

(5d) $f_h^{*'} = w(1-g'(wh)) = wf'(wh)$.

Når kvinnen ikke er i arbeid, dvs. når $h = 0$, liknes ektefellene sammen. I en slik situasjon vil den deriverte av den disponible inntektsrelasjonen mhp. h være gitt ved funksjonen for felles likning, dvs. lik

$$wf'(0+k) = wf'(k).$$

Denne størrelsen kan passende kalles den marginale nettoinntekten av å arbeide "en time" gitt at en i utgangspunktet ikke arbeider.

Av (2c) og av definisjonen $C_0 = k - t(k) = f(k)$ ser vi da at (2c) kan skrives

(6) $$wf'(k) > \frac{A}{B} f(k)^{1-\beta}.$$

Uttrykket til høyre er den marginale substitusjonsbrøken regnet ut i det punkt hvor h = 0 og kalles i litteraturen den marginale reservasjonslønnen. Ulikheten i (6) sier da at kvinnen ønsker å arbeide hvis den marginale nettoinntekten overstiger den marginale reservasjonslønnen for den "første" timen arbeidet.

Hvis (6) er oppfylt, så blir det optimale antall timer tilbudt bestemt ved likhet mellom marginal nettoinntekt og marginal substitusjonsbrøk. Vi merker oss at den marginale inntekten i en slik situasjon er avhengig av om kvinnen velger å arbeide et antall timer som gjør at ektefellene blir særskilt eller felles liknet. Vi går ikke nærmere inn på dette her siden denne oppgaven dreier seg om analysen av om kvinner er i arbeid eller ikke.

Vi ser av (6) at betingelsen om at kvinnen ønsker arbeid, kan skrives

(7) $$\log wf'(k) > (1-\beta)\log f(k) + \log \frac{A}{B}.$$

Ulikheten i (7) er utgangspunktet for den stokastiske modellen og det estimeringsopplegget som vil bli drøftet i neste kapittel.

3. ESTIMERINGSMETODEN

I dette kapitlet skal vi drøfte hvordan vi skal estimere de ulike parametrene i modellen. Estimeringen vil skje på grunnlag av et tverrsnittsmateriale, dvs. Levekårsundersøkelsen 1980 og inntektsundersøkelsen for 1979. Foran skrev jeg at størrelsen A/B inneholder informasjon om kvinnens sosio-økonomiske bakgrunn, hvilket vil si at A/B avhenger av S. Bare endel av de sosio-økonomiske bakgrunnsvariable vil være observerbare for økonometrikeren. Vi forutsetter derfor om A/B at

(8) $$\mathrm{Log}(A/B) = X_2'\kappa + \varepsilon_2$$

hvor ε_2 er et stokastisk restledd som skal fange opp ikke-observerbare faktorer som påvirker preferansene.

X_2 er en vektor som består av variablene:

- alder
- alder^2/100
- antall barn under 6 år
- antall barn over 6 år

- begrensning i arbeidsevne

κ er en ukjent koeffisientvektor.

Vi skal også innføre en lønnsrelasjon. Det gjør vi blant annet fordi vi mangler lønnsobservasjoner for gifte kvinner som ikke er i arbeid. For denne gruppen må vi derfor predikere timelønna. Lønnsfunksjonen antas å være

$$(9) \qquad \log w = X_1'\eta + \varepsilon_1$$

hvor ε_1 er et stokastisk restledd som fanger opp ikke observerte komponenter som påvirker lønna. Vi antar a priori at lønna kan være en klokkeformet funksjon av alder slik at i forklaringsvariabelvektoren X_1 inngår alder og alder^2/100. Dessuten inngår antall år med utdanning i X_1, og vi forventer at lønna stiger med utdanning.

Det kan tenkes at endel av de uobserverte variable som påvirker ε_2, slik som motivasjon og type jobb/yrke, også påvirker ε_1. Vi åpner derfor for korrelasjon mellom ε_1 og ε_2.

Logaritmeformen er valgt fordi flere analyser har vist at denne passer best med data, se Heckmann (1974) og Dagsvik (1984).

Setter vi inn for log w og log(A/B) i (7), får vi følgende sysselsettingsbetingelse

$$X_1'\eta + \varepsilon_1 + \log f'(k) > (1-\beta)\log f(k) + X_2'\kappa + \varepsilon_2$$

Lar vi

$$v_1 = X_1'\eta + \log f'(k)$$

$$v_2 = (1-\beta)\log f(k) + X_2'\kappa$$

ser vi at ulikheten over kan skrives

$$(10) \qquad v_1 + \varepsilon_1 > v_2 + \varepsilon_2$$

Vi definerer P som sannsynligheten for at en gift kvinne ønsker å arbeide. Av (10) får vi da

$$(11) \qquad P = \Pr(h>0) = \Pr\left[v_1 - v_2 > \varepsilon_2 - \varepsilon_1\right]$$

Vi antar nå at $(\varepsilon_1, \varepsilon_2)$ er bivariat ekstremverdifordelt. For begrunnelse av valg av ekstremverdifordeling, se Dagsvik (1984).

Vi forutsetter at $E\varepsilon_1 = E\varepsilon_2 = \gamma$ = Eulers konstant $\simeq$ 0,5772 og at $\text{var}\varepsilon_1 = \text{var}\,\varepsilon_2 = \sigma^2$. Ekstremverdifordelingen impliserer at variansene er like.

Vi lar $a = \frac{\sigma\sqrt{6}}{\pi}$

En type bivariat ekstremverdi-fordeling ser ut som følger:

$$pr[\varepsilon_1 < x, \varepsilon_2 < y] = \exp(-(e^{\frac{-x}{a\varrho}} + e^{\frac{-y}{a\varrho}})^{\varrho})$$

Hvor ϱ har tolkningen $\varrho^2 = 1-\mathrm{corr}(\varepsilon_1, \varepsilon_2)$ og $0 \leq \varrho \leq 1$.

Vi skal benytte oss av følgende setning som er vist i Johnson og Kotz (1972) og er også demonstrert nedenfor:

Setning 1: Når to variable er ekstremverdifordelt (av type ekstremverdifordeling som angitt ovenfor) vil differansen mellom disse to variablene være logistisk fordelt.

Av (11) og med disse forutsetningene om fordelingen til $(\varepsilon_1, \varepsilon_2)$ får vi da:

$$(12) \qquad P = \Pr[h>0] = \Pr\left[\varepsilon_2-\varepsilon_1 < v_1-v_2\right] = \frac{1}{1+e^{\frac{-(v_1-v_2)}{a\varrho}}}$$

som er den logistiske fordelingen.

Før jeg går videre, vil jeg kommentere litt om ekstremverdifordelingen.

Hvis en variabel X er ekstremverdifordelt, har vi at

$$(13) \qquad \Pr(X<x) = e^{-e^{\frac{-(x-b)}{a}}} = F(x) \text{ (Den kummulative ekstremverdifordeling)}$$

$$EX = \gamma + b$$

hvor γ = Eulers konstant $\approx$ 0,5772 og hvor a og b er konstanter, $a > 0$.

$$\mathrm{var}\, X = \frac{\pi^2}{6}a^2.$$

Tetthetsfunksjonen er gitt ved

$$(14) \qquad f(x) = \frac{1}{a}e^{\frac{-(x-b)}{a}} \exp(-e^{\frac{-(x-b)}{a}})$$

Den standardiserte ekstremverdifordeling er:

$$\Pr(X<x) = e^{-e^{\frac{-(x+\gamma)}{\pi}\sqrt{6}}}$$

Som vi ser, har denne forventning 0 og varians 1.

Ekstremverdifordelingen har følgende egenskaper:

Anta Y_1, Y_2, ..., Y_n er identisk uavhengig fordelt og definer:

$$X_n = \max_{j \leq n} Y_j$$

Da finnes det konstanter α_n og β_n slik at fordelingen til

$$Z_n = \frac{X_n - \alpha_n}{\beta_n}$$

går mot ekstremverdifordelingen når n går mot uendelig, dvs.

$$\lim \Pr(Z_n \leq x) = EV(x) \qquad n \to \infty$$

hvor EV(x) angir den kumulative ekstremverdifordelingen i punktet x. Relevansen av dette for vår modell er at det i ε_1 og ε_2 kan være et stort antall variable som er ukjent for økonometrikeren, men ikke for individet. Siden vår modell er en mikroøkonomisk optimeringsmodell, er det rimelig å anta at individet tilpasser seg rasjonelt, dvs. optimerer over de variable vi ikke kjenner til. Hvis antall utelatte variable er "tilstrekkelig mange", er dermed ekstremverdifordelingen en rimelig, nærmest aksiomatisk begrunnet, fordeling.

Vi vender nå tilbake til relasjon (12), som definerer jobbsannsynligheten. Vi har da at sannsynligheten for <u>ikke</u> å være i arbeid blir

$$1-p = \frac{e^{\frac{-(v_1-v_2)}{a\varrho}}}{1+e^{\frac{-(v_1-v_2)}{a\varrho}}}$$

Vi får derfor

$$(15) \qquad L^* = \log\left(\frac{P}{1-P}\right) = (v_1 - v_2)$$

Når P går fra 0 til 1, går L^* fra $-\infty$ til $+\infty$. Med andre ord, (v_1-v_2) kan variere over alle grenser, men P vil holde seg innenfor intervallet [0,1].

Hadde vi i stedet antatt at P var lineær i parametrene, ville vi kunnet risikere å få prediksjoner utenfor [0,1] intervallet.

Funksjonen (15) sikrer oss at sannsynlighetsprediksjonene ligger innenfor [0,1] intervallet og er således mer egnet enn den lineære funksjonsformen.

Fra definisjonene (8), (9) og (10) ser vi at

$$(X_1\eta - X_2\kappa) = (\eta_2-\kappa_2)(\text{alder}) + (\eta_3-\kappa_3)(\text{alder}^2/100)$$
$$+ \eta_1(\text{utdannelse}) + \kappa_1(\text{antall egne barn, 0-6 år})$$
$$+ \kappa_4(\text{antall egne barn, 7-18 år})$$
$$+ \kappa_5 \cdot \text{begrensning i arbeidsevne}$$

som for korthets skyld skrives $X'\varrho = (X_1'\eta - X_2'\kappa)$

Vi kan derfor skrive

$$(16) \quad \frac{-(v_1-v_2)}{a\varrho} = -\frac{1}{a\varrho}\{(\beta-1)\log f(k)+\log f'(k)+X'\varrho\}$$

Definerer vi $\frac{(\beta-1)}{a\varrho} = b_1$

$$\frac{1}{a\varrho} = b_2$$

$\frac{\varrho}{a\varrho} = b_3$ vil

$$(17) \quad \frac{-(v_1-v_2)}{a\varrho} = -\{b_1\log f(k)+b_2\log f'(k)+Xb_3\}$$

Setter vi (17) inn i (12), får vi

$$(18) \quad P = pr[h>0] = \frac{1}{1+e^{-\left[b_1\log f(k)+b_2\log f'(k)+X'b_3\right]}}$$

I det følgende vil vi gå gjennom <u>3 estimeringstrinn</u>, hvorav estimering av parametrene i (18) utgjør første trinn. I første trinn får vi <u>ikke</u> tatt hensyn til en restriksjon vi påla i (7), nemlig at log w og log $f'(k)$ skal ha samme koeffisient.

Resultatene fra beregningene i trinn 1 vil jeg bruke i <u>trinn 2</u>, hvor jeg skal estimere parametrene i timelønnsrelasjonen. Når parametrene i timelønnsrelasjonen er estimert, kan vi også lage et estimat for $E(\log wf'(k))$. Dette estimatet vil bli brukt som høyresidevariabel i trinn 3, slik at restriksjonen fra (7) blir respektert.

I trinn 1 finner vi estimatene $\hat{b}_1$, $\hat{b}_2$, $\hat{b}_3$ ved hjelp av sannsynlighetsmaksimeringsmetoden. Det vi si, vi maksimerer likelihoodfunksjonen

$$M = \prod_{\substack{i=\text{alle}\\ \text{sysselsatte}}} P_i \prod_{\substack{i=\text{alle ikke-}\\ \text{sysselsatte}}} (1-P_i)$$ med hensyn på b_1, b_2 og b_3.

Maksimum likelihood estimatorene er de verdiene av b_1, b_2 og b_3 som gir løsning av systemet

$$\left\{\frac{\delta M}{\delta b_1} = \frac{\delta M}{\delta b_2} = \frac{\delta M}{\delta b_3} = 0\right\}$$

Løsningsverdiene for b_1, b_2 og b_3 definerer vi som $\hat{b}_1$, $\hat{b}_2$ og $\hat{b}_3$. Disse blir funnet med et standard LOGIT program.

Estimatet for $a\varrho$ får vi ved:

$$\hat{a\varrho} = \frac{1}{\hat{b}_2}$$

Dermed får vi også avledet $\hat{\beta}$ og $\hat{\Omega}$. Vi får ikke estimert κ og η hver for seg i denne omgang.

Vi går så over til trinn 2. Et problem her er at lønna bare er observert for de som er i arbeid, så vi må derfor korrigere for hva Heckmann (1979) kaller "selectivity bias" ("utvalgsskjevhet"). Med utvalgsskjevhet menes at man ikke uten videre bør predikere lønna for alle kvinner på grunnlag av lønnsobservasjoner for de som jobber. Jeg vil illustrere dette med et eksempel: Vi har observasjoner for utdannelsens lengde for alle kvinner, men vi har ingen observasjoner for utdannelsens type. La oss nå si at alle sysselsatte kvinner har en utdannelse som kvalifiserer dem til høytlønnsyrker og at alle ikke-sysselsatte kvinner har en utdannelse av samme varighet som de sysselsatte, men denne kvalifiserer bare til lavtlønnsyrker. Skulle vi predikere timelønna for de ikke-sysselsatte ved å regne ut gjennomsnittslønna for de sysselsatte, ville vi åpenbart risikere å gjøre en dårlig prediksjon.

Seleksjonsskjevhet oppstår med andre ord fordi lønna kan være korrelert med jobb/ikke jobb status. Altså er den betingede forventningen til $\varepsilon 1$ (se (9)) gitt at kvinnen jobber, ikke nødvendigvis lik den ubetingede forventningen til ε_1.

Vi ønsker derfor å komme fram til et uttrykk for

(19) $$E[\log w | h>0]$$

hvor w som foran, er timelønn.

Vi trenger noe ny notasjon for den videre drøfting.

W = marginallønn (= $w(1-g')$)

W^* = marginal reservasjonslønn

og vi husker at betingelsen for $h > 0$ (på logaritmisk form) er $\log W > \log W^*$.

Vi starter med å utlede

(20) $\quad E(\log W|h>0)$

Vi har

$$E(\log W|h>0) = E[\log W|\log W>\log W^*] = E\left[\log W|v_1-v_2>\varepsilon_2-\varepsilon_1\right]$$

La fordelingsfunksjonen H være definert ved

(21) $$H(x,y) = \Pr[\log W<x,\log W^*<y] = \Pr\left[v_1+\varepsilon_1<x,v_2+\varepsilon_2<y\right] = F(x-v_1,y-v_2)$$

hvor F(x,y) er definert ved (15).

Videre har vi at

(22) $$H_1'(x,x)dx \approx \Pr(\log W\varepsilon(x,x+dx),\log W>\log W^*)$$

$$\approx \Pr(\log W\varepsilon(x,x+dx),\log W^*<x)$$

fordi

$$\frac{H(x+\Delta x,x)-H(x,x)}{\Delta x} \xrightarrow[\Delta x \to 0]{} H_1'(x,x)$$

Dermed blir

(23) $$\Pr[\log W>\log W^*] \quad = \quad \int H_1'(x,x)dx$$

Den betingede sannsynlighet for log(W) gitt at kvinnen jobber, finner vi ved

(23) $$\Pr[\log W\varepsilon(x,x+dx)|\log W>\log W^*]$$

$$\approx \frac{\Pr[\log W\varepsilon(x,x+dx),\log W^*<x]}{\Pr[\log W>\log W^*]} \approx \frac{H_1'(x,x)dx}{\int_{-\infty}^{\infty} H_1'(y,y)dy}$$

Det følger av dette at

(25) $$E[\log W|h>0] = \int_{-\infty}^{\infty} \frac{xH_1'(x,x)dx}{\int_{-\infty}^{\infty} H_1'(y,y)dy} = \frac{\int_{-\infty}^{\infty} xH_1'(x,x)dx}{\int_{-\infty}^{\infty} H_1'(y,y)dy}$$

La oss gå tilbake til (21). Ved å derivere (21) mhp. første argument, får vi at

$$H_1'(x,y) = \left\{\exp\left[-\left[e^{\frac{v_1-x}{a\varrho}}+e^{\frac{v_2-y}{a\varrho}}\right]^{\varrho}\right]\right\}\left\{\left(e^{\frac{v_1-x}{a\varrho}}+e^{\frac{v_2-y}{a\varrho}}\right)^{\varrho-1}\right\}\varrho e^{\frac{v_1-x}{a\varrho}}\frac{1}{a\varrho}\,.$$

som impliserer at

$$(26)\qquad H_1'(x,x) = \{\exp(-(e^{\frac{v_1-x}{a\varrho}}+e^{\frac{v_2-x}{a\varrho}})^{\varrho-1})\}((e^{\frac{v_1-x}{a\varrho}}+e^{\frac{v_2-x}{a\varrho}})^{\varrho}\frac{e^{\frac{v_1-x}{a\varrho}}}{a(e^{\frac{v_1-x}{a\varrho}}+e^{\frac{v_2-x}{a\varrho}})}$$

$$= \{\exp(-e^{-\frac{x}{a}}(e^{\frac{v_1}{a\varrho}}+e^{\frac{v_2}{a\varrho}})^{\varrho}\}e^{-\frac{x}{a}}(e^{\frac{v_1}{a\varrho}}+e^{\frac{v_2}{a\varrho}})^{\varrho}\frac{e^{\frac{v_1}{a\varrho}}}{a(e^{\frac{v_1}{a\varrho}}+e^{\frac{v_2}{a\varrho}})}$$

Vi definerer nå

$$(27)\qquad d = (e^{\frac{v_1}{a\varrho}}+e^{\frac{v_2}{a\varrho}})^{\varrho}$$

Innsatt i (26) gir dette:

$$H_1'(x,x) = \exp\left[-e^{-\frac{x}{a}}d\right] e^{-\frac{x}{a}}\frac{d}{a}P$$

der P fremdeles betegner jobbsannsynligheten. Jobbsannsynligheten finner vi ved å integrere over alle mulige verdier av x, det vil si:

$$(29)\qquad \int_{-\infty}^{\infty} H_1'(x,x)dx = P\int_{-\infty}^{\infty}\exp\left[-de^{-\frac{x}{a}}\right]\frac{d}{a}e^{-\frac{x}{a}}dx$$

$$= P\Big|_{-\infty}^{\infty}\exp\left[-de^{-\frac{x}{a}}\right]dx = P.$$

Vi har dermed gitt et bevis for påstanden på side 14 om at P er en Logit funksjon. Merk at en annen skrivemåte for $H_1'(x,x)$ er:

$$(30)\qquad H_1'(x,x) = \exp\left[-e^{\frac{-(x-a\log d)}{a}}\right]e^{\frac{-(x-a\log d)}{a}}\frac{P}{a}$$

Vi skal nå gå tilbake til vårt utgangspunkt, nemlig (20)

$$E[\log W|\log W>\log W^*]$$

Med utgangspunkt i resultatene fra (25), (29) og (30) kan vi skrive

$$(31)\qquad E[\log W|h>0] = \frac{\int_{-\infty}^{\infty} xH_1'(x,x)dx}{P}$$

$$= \int_{-\infty}^{\infty} x\left[\exp\left(-e^{\frac{-(x-a\log d)}{a}}\right)e^{\frac{-(x-a\log d)}{a}}\frac{P}{a}\right]\frac{dx}{P}$$

$$= \int_{-\infty}^{\infty} x\left[\exp\left(-e^{\frac{-x-a\log d}{a}}\right)e^{\frac{-x-a\log d}{a}}\frac{1}{a}\right]dx$$

hvor

$$(32)\qquad b = a\log d.$$

som et mellomresultat, har vi at

$$\frac{d}{dx}\{\exp(-e^{\frac{-x-b}{a}})\} = \exp(-e^{\frac{-x-b}{a}})e^{\frac{-(x-b)}{a}}\frac{1}{a}$$

slik at

$$\int_{-\infty}^{\infty} xH_1'(x,x)dx = P\int_{-\infty}^{\infty}\left[\exp(-e^{\frac{-(x-b)}{a}})e^{\frac{-(x-b)}{a}}\frac{1}{a}\right]dx = P\int_{-\infty}^{\infty} xf(x),$$

hvor $f(x)$ er definert ved (14).
Vi ser da at

$$\int_{-\infty}^{\infty} xH_1'(x,x) = P(\gamma+b)$$

slik at (31) kan skrives

$$(33)\qquad E[\log W|h>0] = \frac{P(\gamma+b)}{P} = (\gamma+b)$$

Vi definerte b ved (32) og d ved (27). Dette gir

$$(34)\qquad E(\log W|h>0) = \gamma + a\log(e^{\frac{v_1}{a\varrho}} + e^{\frac{v_2}{a\varrho}})^{\varrho} = \gamma + v_1 - a\varrho\log(P)$$

$$= X_1'\eta + \log f'(k) + \gamma - a\varrho\log P$$

Vi ser at differansen mellom den ubetingete forventningen v_1 og den betingete, betinget av at $h>0$, er leddet $\gamma - a\varrho \log P$. Når vi senere skal bruke observerte lønnsdata for å estimere parametrene i lønnsrelasjonen, er det nettopp den betingete forventningen vi må ta utgangspunkt i.

(35) $E(\log w|h>0) = X_1'\eta + \gamma - a\varrho \log(P) = X_1'\varphi + \psi\lambda$

hvor $\varphi = \begin{bmatrix} \gamma+\eta_0 \\ \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix}$ og $\begin{cases} \psi = -a\varrho \\ \lambda = \log(P) \end{cases}$

Parametrene φ og ψ kan nå estimeres på grunnlag av de som er i arbeid ved bruk av minste kvadraters metode ved å bruke log W som avhengig variabel, X_1 og $\hat{\lambda}$ som forklaringsvariable, der $\hat{\lambda} = \log(\hat{P})$ og $\hat{P}$ er den predikerte jobbsannsynlighet fra estimeringstrinn 1.

Vi kjører regresjonen

(36) $\log w = X_1'\eta + \psi\hat{\lambda} + u_1$

hvor u_1 er et restledd med forventning 0.

Vi får dermed estimatene $\hat{\eta}$ og $\hat{\psi}$. Når dette er gjort, er vi ferdig med trinn 2 i prosedyrene.

Estimat $\hat{\varphi}$ vil bli brukt i trinn 3 der jobbsannsynligheten re-estimeres. I trinn 3 skal vi ta henyn til at log w og log $f'(k)$ skal ha samme koeffisient.

Vi bruker estimatene fra (36) til å beregne

(37) $\hat{v}_1 = X_1'\hat{\eta} + \log f'(k)$.

Vi får nå

(38) $P = \Pr[\hat{v}_1+\varepsilon_1>v_2+\varepsilon_2] = \Pr[(\beta-1)\log f'(k)+\hat{v}_1-X_2'\kappa>\varepsilon_2-\varepsilon_1]$

Tidligere definerte vi

(39a) $\frac{(\beta-1)}{a\varrho} = b_1$

(39b) $\frac{1}{a\varrho} = b_2$

I tillegg definerer vi nå

(39c) $\frac{\kappa}{a\varrho} = b_4$

Når vi definerer med $(a\varrho)$ på begge sider av ulikhetstegnet, kan jobbsannsynligheten skrives

$$(40) \qquad P = Pr[h>0] = Pr\left[b_1 \log f(k) + b_2\hat{v}_1 + X_2'b_4 > \frac{\varepsilon_2-\varepsilon_1}{a\varrho}\right]$$

$$= \frac{1}{1+e^{-[b_1 \log f(k)+b_2\hat{v}_1+X_2'b_4]}}$$

Parametrene b_1, b_2 og vektoren b_4 estimeres nå på grunnlag av alle kvinner i undersøkelsen (1 205 personer), og som i trinn 1, bruker vi et standard LOGIT-program som beregner ML estimatorer $\hat{b}_1$, $\hat{b}_2$ og $\hat{b}_4$ for b_1, b_2 og b_4.

4. ESTIMERINGSRESULTATER

I dette kapitlet vi jeg forklare nærmere inneholdet i de variable som ble brukt ved data-kjøringene, samt kommentere estimeringsresultatene fra trinn 1, 2 og 3. Resultatene fra trinn3 vil også bli sammenliknet med en tilsvarende kjøring hvor PROBIT-modellen er benyttet, altså fra en modell hvor ε_1 og ε_2 er bivariat normalt fordelt.

Forklaring av variable og resultater fra trinn 1

Avhengig variabel: denne antar verdien 1 for personer som er observert som "sysselsatt lønnstaker". Jeg presiserer lønnet arbeid fordi selvstendig næringsdrivende får, i likhet med de som oppgav at de ikke var sysselsatt, verdien 0 på denne variabelen. (Av 1 205 personer hadde 608 verdien 1).

Som uavhengige variable har vi brukt: log f(k), log f'(k) og vektoren X, der log f(k) og log f'(k) er forklart tidligere. Variablene i X-vektoren er: alder

alder^2/100
antall egne barn mellom 0 og 6 år
antall egne barn mellom 7 og 18 år
begrensning i arbeidsevne
utdannelse

De 4 første variablene trenger ingen nærmere forklaring. Begrensning i arbeidsevne er en dummy med verdi 0 for de som ikke har begrensning i arbeidsevne, verdien 1 for de som har begrenset arbeidsevne.

Utdannelse angir antall år med fullført utdannelse multiplisert med 10. Laveste verdi på denne variabelen er 90, høyeste verdi er 180.

Data er fra levekårsundersøkelsen for 1980 og fra koplete data til Inntektsundersøkelsen fra 1979 (basert på et utvalg av selvangivelser).

Fra kjøringene på trinn 1 fikk jeg følgende resultater:

Tabell 1. Tabell over estimater i trinn 1. Foreløpig estimering av jobbsannsynlighet

Variabelnavn	estimat-symbol	estimat-størrelse	t-verdi
Konstantleddet	$\hat{b}_{30}$	-7,790	7,173
Begrensning i arbeidsevne	$\hat{b}_{31}$	-0,621	4,18
Antall små barn	$\hat{b}_{32}$	-0,87	7,28
Antall store barn	$\hat{b}_{33}$	-0,181	2,57
Utdanning	$\hat{b}_{34}$	0,03	6,9
Alder	$\hat{b}_{35}$	0,286	6,09
$\text{Alder}^2/100$	$\hat{b}_{36}$	-0,361	6,55
$\log f'(k)$	$\hat{b}_2$	1,909	7,19
$\log f(k)$	$\hat{b}_1$	1,157	2,91

t-verdiene viser at koeffisientene er signifikant forskjellig fra null.

Som vi ser av tabell 1, virker begrensning i arbeidsevne, små barn, store barn og $(\text{alder}^2/100)$ negativt inn på jobbsannsynligheten. Utdannelse, alder og log f(k) virker positivt på jobbsannsynligheten. Estimatet foran $\log f'(k)$ er også positivt. Ut fra denne kjøringen, isolert betraktet, kan vi derfor ikke si om k virker positivt eller negativt på jobbsannsynligheten når vi tar hensyn til virkninger av k gjennom både marginalskatt og inntekt.

Siden trinn 1 bare utgjør et foreløpig estimeringstrinn, vil jeg ikke kommentere resultatene ytterligere.

Forklaring av variable og resultater fra trinn 2 (timelønnsrelasjonen):

Avhengig variabel: logaritmen til observert timelønn

Uavhengig variable: alder

$\text{alder}^2/100$

utdannelse (samme som trinn 1)

$\hat{\lambda} = \log(\hat{P})$

$\hat{P}$ er den predikerte jobbsannsynlighet fra trinn 1. Siden $0 < \hat{P} < 1$, vil $\hat{\lambda}$ være en negativ variabel.

Resultatene fra trinn 2 er gitt i tabell 2.

Tabell 2. Tabell over estimater i trinn 2. Timelønnsrelasjoner

Variabelnavn	estimatsymbol	estimatstørrelse	t-verdi
Konstantleddet	$\hat{\eta}_1$	2,082	8,02
Utdannelse	$\hat{\eta}_2$	0,0067	7,56
Alder	$\hat{\eta}_3$	0,031	2,89
Alder^2/100	$\hat{\eta}$	-0,034	2,63
$\hat{\lambda}$	$\hat{\psi}$	0,069	1,36

Vi har fått signifikante estimater foran alder, alder^2/100 og utdannelse. Utdannelse virker positivt på jobbsannsynligheten. Seleksjonsvariabelen $\hat{\lambda}$ har ikke signifikant innflytelse på lønna.

Deriverer vi

$$\log w = 0{,}031 \cdot \text{alder} - 0{,}034 \cdot \frac{\text{alder}^2}{100} + \text{konstantledd}$$

mhp. alder, finner vi at lønnsfunksjonen når sitt høydepunkt med hensyn på alder når alder antar verdien ca. 50 år. Modellforutsetningen tilsier at vi hadde ventet et negativt estimat for ψ, da $(a\rho)$ skal være positiv. Da $\hat{\psi}$ viste seg å være negativ, kan dette tyde på feilspesifikasjon av modellen. På den annen side er ikke ψ signifikant forskjellig fra null.

Trinn 3: Reestimering av jobbsannsynlighet: Resultattolkning og forklaring av variable

I dette trinnet reestimeres jobbsannsynligheten fra trinn 1. Vi bruker følgende høyreside-variable: log f(k), $\hat{v}_1$ og X_2 (vektor). Størrelsen log f(k) er forklart tidligere, og $\hat{v}_1$ er logaritmen til predikert marginallønn beregnet på grunnlag av resultatene fra trinn 2. X_2 er en vektor som består av variablene

begrensning i arbeidsevne
egne barn mellom 0 og 6 år
egne barn mellom 7 og 18 år
alder
alder^2/100

Vi fikk følgende estimater fra reestimering av jobbsannsynligheten:

Tabell 3. Tabell over estimater fra trinn 3. Reestimater av jobbsannsynlighet.

Variabelnavn	Symbol	Estimat	t-verdi
Konstantledd	$\hat{b}_{41}$	-8,749	6,404
Begrensning i arbeidsevne ..	$\hat{b}_{42}$	-0,657	4,55
Antall små barn	$\hat{b}_{43}$	-0,503	4,61
Alder	$\hat{b}_{44}$	0,244	5,36
$\text{Alder}^2/100$	$\hat{b}_{45}$	-0,319	5,97
log f(k)	$\hat{b}_1$	0,0749	1,32
Antall store barn	$\hat{b}_{46}$	0,1296	1,85
$\hat{v}_1$= log (pred. marginallønn)	$\hat{b}_2$	1,760	7,01

Vi legger merke til at koeffisienten foran $\hat{v}_1$ er positivt og signifikant. Det vil si at en økning i marginallønna vil virke positivt på jobbsannsynligheten (gitt at ektefellens inntekt og alle andre variable holdes konstant).

Vi ventet at $\hat{b}_1$ skulle bli negativ all den stund β forutsettes mellom 0 og 1. Estimatet er imidlertid ikke signifikant forskjellig fra 0, dvs. at estimatet på β ikke er signifikant forskjellig fra 1. Vi kan tolke resultatet slik at en økning i ektefelleinntekten vil være uten betydning for jobbsannsynligheten (gitt at marginallønna og alle andre variable holdes konstant). $\hat{b}_{46}$ ble heller ikke signifikant på 5 prosent nivå, og antallet store barn ser derfor ikke ut til å ha noen betydning for jobbsannsynligheten. $\hat{b}_{41}$, $\hat{b}_{42}$, $\hat{b}_{43}$, $\hat{b}_{44}$ og $\hat{b}_{45}$ ble alle signifikante. Vi kan derfor konkludere med at jobbsannsynligheten avtar med økende begrensing i arbeidsevne, og den avtar også med antall små barn.

Deriverer vi P med hensyn på alder, finner vi at jobbsannsynligheten når sitt høydepunkt med hensyn på alder når alder antar verdien 38 år.

Vi skal nå sammenlikne resultatene fra trinn 3 (LOGIT) med en tilsvarende reestimering hvor Probit-modellen er benyttet.

Tabell 4. Tabell over estimater fra PROBIT's reestimering av jobbsannsynligheten

Variabelnavn	Estimatstørrelse	t-verdi
Konstantledd	-5,352	6,53
Begrensing i arbeidsevne	-0,402	4,55
Antall små barn	-0,300	4,54
Alder	0,1524	5,56
Alder^2/100	-0,1987	6,18
log f(k)	0,0425	1,26
Antall store barn	-0,0791	1,878
$\hat{v}_1$= log (pred. marginallønn)	1,054	7,02

Vi ser at estimatene fra PROBIT-kjøringene er lavere enn tilsvarende estimater i LOGIT. Koeffisientene i Probit er i henhold til Amemyia (1981) omtrent lik koeffisienten i Logit multiplisert med 1,7, hvilket stort sett er tilfelle også i det vi har gjort.

Tolkningen av resultatene blir imidlertid den samme i Logit og Probit da også Probit-kjøringene ikke gir signifikante estimater foran log f(k) og store barn. Vi får samme fortegn og signifikante estimater både i Logit og Probit for log marginallønn, små barn, alder, alder^2 og begrensing i arbeidsevne.

Deriverer vi P med hensyn på alder, får vi at i Probit blir jobbsannsynligheten med hensyn på alder størst for 38-åringer.

5. ELASTISITETER

I dette kapitlet skal jeg drøfte nærmere hvordan jobbsannsynligheten vil endre seg når det skjer en økning i marginallønna (gitt at ektefellens inntekt og andre variable er konstante).

Jeg skal presenterer elastisiteten av jobbsannsynligheten med hensyn på marginallønna for noen av kvinnene i utvalget.

Marginallønna er definert som følger:

$$M = e^{v_1} \quad \text{hvor } v_1 = X_1\eta + \log f'(k).$$

Jeg skal elastisitere både Logits og Probits jobbsannsynligheter. Deretter vil jeg sammenlikne resultatene innsatt de estimater vi fikk av beregningene.

Utledning av elastisitetene:

Vi tar utgangspunkt i trinn 3's reestimerte jobbsannsynlighet. Vi har tidligere definert

(41) $$P_{Logit} = \frac{1}{1+e^{-\chi}} = L(\chi) \quad \text{hvor} \quad \chi = (b_1 \log f(k) + b_2 \hat{v}_1 + X_2' b_4)$$

(42) $$P_{Probit} = \Phi(Z)$$

hvor Φ er den kumulative normalfordelingen og L er den kummulative logistiske fordelingsfunksjonen.

$$Z = (\xi_1 \log f(k) + \xi_2 \hat{v}_1 + x_2 \xi_4)$$

Elastisiteten av P med hensyn på M uttrykker den prosentvise endring i P når M endres med én prosent.

Merk at v_1 = log M.

Logit elastisiteten blir som følger

(43) $$El_M P_{Logit} = \frac{M}{L(\chi)} L'(\chi) \frac{\delta\chi}{\delta M} = b_2 \left[\frac{e^{-\chi}}{1+e^{-\chi}}\right] = b_2[1-P]$$

og Probit elastisiteten blir

$$El_M P_{Probit} = \frac{M}{\Phi(Z)} \Phi'(Z) \frac{\delta Z}{\delta M} = \xi_2 \lambda$$

hvor $\lambda = \frac{\Phi'(Z)}{\Phi(Z)}$

Innsatt estimatene fra den reestimerte jobbsannsynlighet, vil disse elastisitetene kunne skrives:

(45) $$El_M \hat{P}_{Logit} = \hat{b}_2[1-\hat{P}]$$

(46) $$El_M \hat{P}_{Probit} = \hat{\xi}_2 \hat{\lambda}$$

Jeg vil presentere de 10 laveste og de 10 høyeste elastisitene i tabell 5.

Tabell 5. Tabell over de 10 laveste og de 10 høyeste elastisiteter etter stigende rekkefølge, samt de tilhørende verdier på andre variable

Elastisiteter av jobbsannsynligheter mhp. marginallønnen Logit	Probit	Alder	Utdan-nelses-lengde	Ant. små barn	Ant. store barn	begr. i ar-beids-evne	$\hat{v}_1$Logit ≈ $\hat{v}_1$Probit	$\hat{P}$Logit	$\hat{\lambda}$Probit	M
0,2	0,18	40	135	0	1	0	3,5	0,93	0,12	33
0,12	0,19	40	155	0	1	0	3,4	0,93	0,12	30
0,14	0,22	36	155	0	0	0	3,3	0,92	0,15	27
0,15	0,22	47	135	0	1	0	3,5	0,92	0,15	33
0,17	0,26	34	155	0	0	0	3,2	0,90	0,17	24
0,17	0,26	41	155	0	2	0	3,3	0,90	0,17	27
0,19	0,29	27	155	0	0	0	3,3	0,89	0,20	27
0,20	0,30	34	135	0	1	0	3,2	0,89	0,20	24
0,20	0,30	32	135	0	0	0	3,1	0,89	0,20	22
0,20	0,31	33	135	0	0	0	3,1	0,12	0,21	22
1,54	2,53	66	90	0	0	0	2,5	0,12	1,7	12
1,55	2,55	66	100	0	0	1	2,5	0,12	1,7	12
1,55	2,55	66	100	0	0	0	2,1	0,12	1,7	8
1,56	2,58	64	90	0	0	1	2,3	0,12	1,7	10
1,56	2,59	22	90	1	0	1	1,8	0,11	1,7	6
1,57	2,64	65	90	0	0	0	1,9	0,11	1,8	7
1,60	2,80	63	90	0	0	1	2,0	0,09	1,9	7
1,66	3,15	63	100	0	0	1	1,7	0,06	2,1	5
1,66	3,15	63	100	0	0	1	1,7	0,06	2,1	5
1,66	3,17	64	90	0	0	1	1,8	0,06	2,1	6
1,67	3,44	66	100	0	0	1	1,8	0,04	2,3	6

Sortering av elastisitene ga samme rankering i Logit og Probit. Kvinnen med lavest Logit-elastisitet hadde lavest Probit elastisitet. Kvinnen med høyest Logit-elastisitet hadde også høyest Probit-elastisitet.

Tabell 5 viser at kvinner som allerede har en høy predikert jobbsannsynlighet, er lite følsomme overfor endringer i marginallønna. Disse kvinnene har også en høy predikert marginallønn i utgangspunktet. Andre felles kjennetegn er at de alle er under 50 år, de har ikke små barn og alle har utdannelse utover gymnasnivå.

De høye elastisitetene fant vi hos kvinner med lav predikert jobbsannsynlighet. Med unntak av én, er alle disse kvinnene rundt pensjonsalderen , og de fleste har en begrenset arbeidsevne. Den ene unge kvinnen har et mindreårig barn samt begrenset arbeidsevne. Alle 10 har lav utdannelse og lav predikert marginallønn.

Vi ser at Logit-elastisiteten er langt lavere enn Probit-elastisitetene selv om $\hat{b}_2$ er litt større enn $\hat{\xi}_2$ ($\hat{b}_2$=1,76, $\hat{\xi}_2$=1,054).

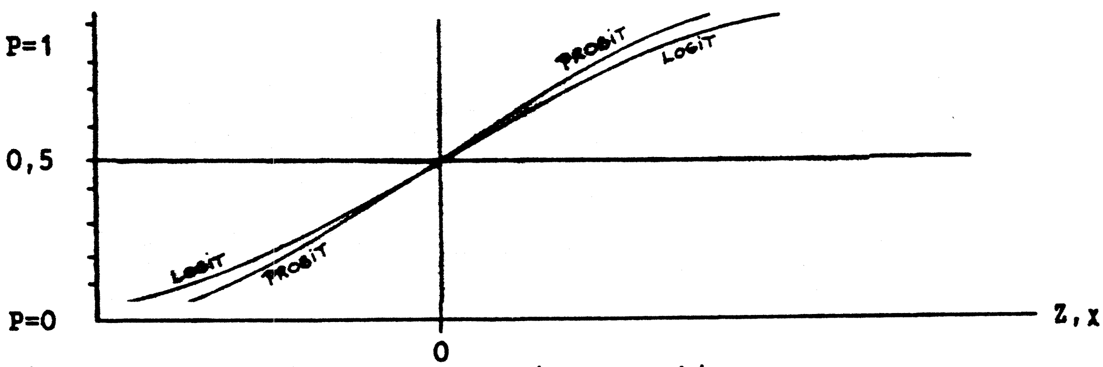
Dette kan forklares ut fra funksjonsformene i henholdsvis Logit og Probit. Probit-elastisiteten kan skrives som følger:

$$El_M P_{Probit} = \xi_2 \frac{\Phi'(Z)M}{\Phi(Z)} \quad \text{hvor } \Phi(Z) = P_{Probit}$$

og Logit-elastisitetene kan skrives

$$El_M P_{Logit} = b_2 \frac{L'(\chi)M}{L(\chi)}, \quad L(\chi) = P_{Logit}.$$

Tegnet inn i et koordinatsystem, ser funksjonene slik ut:



Figur 1. Sammenlikningen av Logit og Probit.

For lave verdier av P vil $El_M P_{Probit} > El_M P_{Logit}$ fordi helningen er er omtrent like $[(\Phi'(Z) \approx L'(\chi)]$, men nivået er størst i Logit: $L(\chi) > \Phi(Z)$. Koeffisient-estimatene er såpass like at forskjellene i disse ikke oppveier den effekten nevneren har på elastisitetens størrelse. Logit-elastisiteten har størst nevner, derfor blir Logit-elastisiteten minst. For høye verdier av P, vil nevneren i elastisitetsbrøkene være tilnærmet lik 1. Da vil det være tellerne i $\Phi'(Z)$ og $L'(\chi)$ samt koeffisient-estimatene, som har betydning for elastisitetens størrelse. Siden Probit-funksjonen stiger raskere enn Logit, vil derfor Probit-elastisitene være størst.

Som påpekt av Amemyia (1981), er det bånd mellom Probit og Logit som gjør at estimatene i Logit er en faktor på 1,7 høyere enn Probit-estimatene. Ved en passende normalisering av Probit, kan en dermed få praktisk talt like estimat. Av den grunn kan en bli ledet til å tro at det er likegyldig hvilken modell en anvender. Elastisitetsberegningene her viser imidlertid at dette ikke er tilfelle. Ved bruken av modellen kan det ha betydning om en har en Probit eller Logit. Vi ser av figur 1 at Logit er noe "tjukkerere" i halene enn Probit. Dette kan være en viktig egenskap i analyse av et materiale hvor "mye" sannsynlihetsmasse er samlet rundt verdier som avviker noe fra gjennomsnittsverdiene i sampelet. Dette kan nettopp være tilfelle ved et tverrsnittsmateriale over gifte kvinners tilpasning i arbeidsmarkedet. På grunn av kohorteffekter, kan det være stor spredning i materialet. Dette gir den forsiktige konklusjon at Logit bør foretrekkes framfor Probit.

LITTERATUR

Amemyia, T. (1981): "Qualitative Response Models: A Survey". Journal of Economic Literature, December.

Dagsvik, J. (1984): Stochastic Properties of the Market Wage and the Labour supply Functions. A theoretical Justification Manuscript, Statistisk Sentralbyrå.

Dagsvik, J, O. Ljones, S. Strøm og R. Aaberge (1986): "Gifte kvinners arbeidstilbud, skatter og fordelingsvirkninger". Kommer i Rapportserien fra Statistisk Sentralbyrå, 1986.

Heckmann, J. (1974): "Shadow Prices, Market Wages and Labour Supply". Econometrica 42.

Johnson, N. and S. Kotz (1972): "Distribution i Statistics: Continuous Multivariate Distributions". John Wiley, New York.

Killingsworth, M. (1983): "Labour Supply". Cambrigde University Press.

Offerdal, E. og S. Strøm (1979): "Arbeidsløshet, betalingsbalanse og inflasjon". Memorandum fra Sosialøkonomisk Institutt.