Arbeidsnotater

STATISTISK SENTRALBYRÅ

Dronningensgt. 16, Oslo-Dep., Oslo 1. Tlf. 41 38 20

IO 77/4 21. januar 1977

HVORDAN PÅVIRKES FORDELING AV DISPONIBEL INNTEKT FOR PERSONLIGE SKATTYTERE AV ENDRINGER I INNTEKTS- OG FORMUESBESKATNINGEN?x

Av

Ole K. Hovland

INNHOLD

---

x) Dette arbeid er opprinnelig skrevet som spesialoppgave ved det sosialøkonomiske studium. Forfatteren har stått fritt i valg av opplegg og analysemetoder. Arbeidet gjengis her i noe bearbeidet form med endringer som forfatteren har ønsket å foreta. Synspunkter og konklusjoner står for forfatterens regning.

1. Innledning

Jeg skal ta for meg noen konkrete skatteendringer, og se nærmere på fordelingen av disponibel inntekt for personlige skatteytere ved hvert skattealternativ. Først og fremst er jeg interessert i å studere virkningen av endringer i formuesbeskatningen. Kan vi få en "jevnere" fordeling av disponibel inntekt ved å gå over til å beskatte formuen hardere, og for eksempel redusere inntektsskatten til staten tilsvarende?

Jeg skal bruke data fra inntekts- og formuesundersøkelsen i 1973, og de endringene i skattesystemet jeg kommer til å foreta, blir endringer ut fra gjeldende skatteregler i 1973. Alle endringer er foretatt slik at de gir provenymessig balanse med det faktiske skattesystem i 1973.

For å analysere endringer i inntektsfordelingen skal jeg bruke programmet LOTTE. LOTTE er et regnemaskinprogram som er utviklet ved skatteforskningsgruppen i SSB. Det brukes til beregning av inntekts- og formuesskatt under ulike definisjoner av inntekt og under ulike skatteregler. Programmet er spesielt egnet til å belyse virkninger på privat disponibel inntekt og direkte skatter av endringer i skattesystemet.

For å sammenligne inntektsfordelingene under de ulike skattealternativer vil jeg bruke Lorenz-kurver og en enkel Gini-koeffisient. Jeg vil også bruke en χ^2-homogenitetstest for å teste om de utslag jeg får virkelig er signifikante.

2. Presisering av problemstilling

Utgangspunktet er fordeling av disponibel inntekt for personlige skatteytere i 1973 under dagjeldende skattesystem, situasjon 0. Følgende skatter er med:

- inntektsskatt til kommune
- fellesskatt til Skattefordelingsfond
- inntektsskatt til stat
- folketrygdpremie, pensjonsdel og sykedel
- formuesskatt til kommune
- formuesskatt til stat
- barnetrygd (negativ skatt)
- forsørgerstønad (negativ skatt)

Disponibel inntekt definerer vi slik:

Disponibel inntekt = nettoinntekt ved statsskattelikning
+ særfradrag - sum av skatter.

I situasjon 1 tenker vi oss at formuesskatten til stat og kommune faller bort. Vi skal se på to alternative måter å erstatte denne på:

Situasjon 1: Formuesskatten til stat og kommune faller bort, og erstattes av en like stor prosentvis økning (a%) i alle prosentsatsene i inntektsskatten til staten.

Situasjon 2: Formuesskatten til stat og kommune faller bort, og erstattes av at alle prosentsatsene i inntektsskatten til staten får et like stort tillegg (+b%).

Vi forutsetter at skatteendringene i situasjon 1 og 2 gir provenymessig balanse, dvs. at a og b er av en slik størrelsesorden at det totale innbetalte skattebeløp i situasjon 1 og 2 er lik det totale innbetalte skattebeløp i situasjon 0.

I situasjon 3 tenker vi oss at inntektsskatten til staten faller delvis bort og erstattes av økt formuesskatt til staten:

Situasjon 3: Inntektsskatten til staten gjøres "flat", dvs. skattefritaking på 18 000 kr i klasse 1 og 27 000 kr i klasse 2 beholdes, overskytende beløp beskattes med laveste sats, 10%. Prosentsatsene i formuesskatten til staten justeres opp, og skattefritt fradrag i formuen justeres ned slik at vi får provenymessig balanse.

På forhånd er det vanskelig å si noe mer eksplisitt om utformingen av formuesskatten i situasjon 3. Jeg vil prøve å ha omtrent samme progresjon (men med høyere satser) som gjeldende regler i 1973, men vil forsøke å unngå at høyeste sats overstiger 5%.

Formuesskatten til staten hadde i 1973 følgende utforming[1)]:

Klassefradrag i formuen:

kr 75 000 i skatteklasse 1
" 100 000 " " 2.

Netto formue blir redusert med klassefradraget, og av det gjenstående, skattbar formue, er skatten:

0,4 % av de første 250 000
0,8 % " " neste 250 000
1,2 % " overskytende beløp.

Opprinnelig hadde jeg tenkt å fjerne all inntektsskatt til staten i situasjon 3 og erstatte dette med økt formuesskatt, men ved hjelp av LOTTE fant jeg at satsene ovenfor måtte økes til ca. 10%, 20% og 30% for at vi skulle få provenymessig balanse.

I 1973 ble "80%-regelen" opphevet. Denne regelen innebar (noe forenklet) at de samlede inntekts- og formuesskatter ikke skulle overskride 80% av nettoinntekten. I 1974 ble en lignende "90%-regel" innført igjen. Regelen kommer særlig til anvendelse i tilfeller med lave inntekter og høye formuer. I analysen har vi ikke med noen "80%-regel". Dette medførte for situasjon 3's vedkommende at det ble forholdsvis mange med negativ disponibel inntekt, noe som skaper problemer ved tegning av Lorenz-kurver.

Vi innfører derfor i situasjon 3 en "100%-regel", dvs. at de samlede inntekts- og formuesskatter ikke må overskride nettoinntekten ved statsskattelikningen + særfradrag. For de andre alternativene gir en slik regel neglisjerbart utslag.

Før vi går videre kan det være nyttig å presisere problemstillingen i symboler. For $(i=1,\dots, N)$ og $(j=0,1,2,3)$ har vi at[2)]:

N - antall inntektstakere
r_i - nettoinntekt for individ i
f_i - netto formue for individ i
T_i^j - samlet skatt av inntekt og formue for individ i i situasjon j
$\phi^j = \phi^j(r_i)$ - inntektsskatt til stat for individ i i situasjon j
$\lambda^j = \lambda^j(r_i)$ - andre inntektsskatter for individ i i situasjon j
$\beta^j = \beta^j(f_i)$ - formuesskatt til stat for individ i i situasjon j
$\gamma^j = \gamma^j(f_i)$ - formuesskatt til kommune for individ i i situasjon j

"100%-regel" for situasjon 3 impliserer at[3)]:

$$T_i^3 = \begin{cases} (\phi^3(r_i)+\lambda^3(r_i)+\beta^3(f_i)+\gamma^3(f_i)) & \text{hvis } (\phi^3(r_i)+\lambda^3(r_i)+\beta^3(f_i)+\gamma^3(f_i)) < r_i \\ (r_i & \text{hvis } (\qquad " \qquad) \geq r_i \end{cases}$$

For situasjon 1 og 2 vet vi at:

$$\beta^1(f_i) = \beta^2(f_i) = \gamma^1(f_i) = \gamma^2(f_i) = 0$$

Videre vil:

$$\lambda^0(r_i) = \lambda^1(r_i) = \lambda^2(r_i) = \lambda^3(r_i)$$

og

$$\gamma^0(f_i) = \gamma^3(f_i)$$

1) Se "Historisk oversikt over skattesatser m.v." Arbeidsnotat fra SSB IO 76/17.
2) Noe forenklet da formen på skattefunksjonene er differensiert etter (bl.a.) forsørgelsesbyrde.
3) Se Engebretsen (1974) s. 13.

Da får vi:

$$T_i^0 = \phi^0(r_i) + \lambda^0(r_i) + \beta^0(f_i) + \gamma^0(f_i)$$

$$T_i^1 = \phi^1(r_i) + \lambda^0(r_i)$$

$$T_i^2 = \phi^2(r_i) + \lambda^0(r_i)$$

$$T_i^3 = \begin{cases} (\phi^3(r_i)+\lambda^0(r_i)+\beta^3(f_i)+\gamma^0(f_i)) & \text{hvis } (\phi^3(r_i)+\lambda^0(r_i)+\beta^3(f_i)+\gamma^0(f_i)) < r_i \\ (r_i & \text{hvis } (\qquad \text{"} \qquad) \geq r_i \end{cases}$$

Vi forlanger at alle skattealternativene skal gi samme proveny som situasjon 0. Det betyr at vi må ha:

(1) $\sum_{i=1}^{N} T_i^0 = \sum_{i=1}^{N} T_i^1 \iff \sum_{i=1}^{N} (\phi^0(r_i)+\beta^0(f_i)+\gamma^0(f_i)) = \sum_{i=1}^{N} \phi^1(r_i)$

(2) $\sum_{i=1}^{N} T_i^0 = \sum_{i=1}^{N} T_i^2 \iff \sum_{i=1}^{N} (\phi^0(r_i)+\beta^0(f_i)+\gamma^0(f_i)) = \sum_{i=1}^{N} \phi^2(r_i)$

(3) $\sum_{i=1}^{N} T_i^0 = \sum_{i=1}^{N} T_i^3$

Vi kjenner alle funksjonene som inngår i T_i^0, og kjenner følgelig også $\sum_{i=1}^{N} T_i^0$ som er summen av alle skattene nevnt i begynnelsen av avsnittet i 1973.

Ligning (2) og (3) ovenfor er det vanskelig å komme videre med. Ligning (2) beskriver at vi skal ha samme proveny i situasjon 2 som i situasjon 0. $\phi^2(r_i)$ er en ukjent funksjon. Vi har forutsatt at vi skal oppnå samme proveny ved å gi alle prosentsatser i inntektsskatten til staten, $\phi^2(r_i)$, et like stort tillegg. Vi må da gjette på et tillegg, bruke LOTTE og se hvilken proveny dette gir, gjette på nytt til vi får et tilfredsstillende resultat.

Ligning (3) beskriver at vi skal ha samme proveny i situasjon 3 som i situasjon 0. $\phi^3(r_i)$ kjenner vi i og med at vi tidligere har spesifisert hvordan vi vil ha inntektsskatten til staten i situasjon 3. $\beta^3(f_i)$ er den ukjente funksjonen. Her må vi bruke samme framgangsmåte som for ligning (2)'s vedkommende, bruke LOTTE og gjette til vi får omtrent proveny.

Ligning (1) beskriver at vi skal ha samme proveny i situasjon 1 som i situasjon 0. $\phi^1(r_i)$ er ukjent for oss. For å oppnå samme proveny, når vi sløyfer all formuesskatt, har vi forutsatt at vi skal gi alle prosentsatsene i inntektsskatten til staten en like stor prosentvis økning, a%. Dette er imidlertid det samme som å øke selve inntektsskatten med a%. Da kan ligning (1) omskrives til:

(1)^x $\sum_{i=1}^{N} (\phi^0(r_i)+\beta^0(f_i)+\gamma^0(f_i)) = \sum_{i=1}^{N} (1+a/100)\cdot\phi^0(r_i)$

$$\iff \qquad a = \frac{\sum_{i=1}^{N} (\beta^0(f_i)+\gamma^0(f_i))}{\sum_{i=1}^{N} \phi^0(r_i)} \cdot 100$$

a blir altså lik forholdet mellom totale formuesskatter og total inntektsskatt til staten i 1973, multiplisert med 100 for å få svare i %.

Når vi ved hjelp av de metoder jeg har skissert ovenfor, har kommet fram til skatteregler som gir lik proveny med situasjon 0 for hvert enkelt skattealternativ, skal vi videre bruke programmet LOTTE for å se nærmere på hvordan fordeling av disponibel inntekt for personlige skattytere er under de ulike alternativ. Først skal vi imidlertid se på noen av de problemer som er knyttet til å måle ulikheter mellom fordelinger.

3. Fordelinger og mål for ulikhet

Å måle ulikheter mellom inntektsfordelinger er et omstridt problem. Diskusjonen går særlig på om en kan bruke rent statistiske mål for å måle forskjeller mellom inntektsfordelinger, eller om en må ta utgangspunkt i velferdsfunksjoner eller velferdsbetraktninger for å bygge opp slike mål. En kan skille mellom to typer av mål for ulikhet mellom inntektsfordelinger (Sen (1973) s. 2 og s. 24 og Bigsten (1976) s. 282):

- positive mål som ikke eksplisitt bygger på velferdsbetraktninger eller velferdsfunksjoner
- normative mål som er avledet av eksplisitte velferdsbetraktninger eller velferdsfunksjoner

Ikke alle vil være enige i et slikt skille. Atkinson (1970) hevder at det bak et hvert sammenfattende mål for ulikhet må ligge et velferdsbegrep. Derfor har til og med rent statistiske mål en normativ karakter, idet de implisitt inneholder vurderinger om hvilken vekt en skal tillegge inntektsforskjeller i ulike deler av inntektsskalaen. Ifølge Atkinson bør en derfor egentlig spesifisere en velferdsfunksjon for samfunnet og utifra den avgjøre hvordan en skal beskrive ulikheter mellom inntektsfordelinger.

Andre og deriblant Sen mener at hvis en gjør dette, så vil ulikhetsmålet bli helt avhengig av velferdsfunksjonens utseende, og begrepet ulikhet vil miste den beskrivende betydning det har i normalt språkbruk. De skiller derfor normative mål fra positive mål.

Jeg skal i det følgende se på noen enkle mål for ulikhet som har vært brukt, og si litt om deres svakheter, og hvorfor de ikke egner seg for vårt formål. Så vil jeg gå mer grundig inn på de mål jeg skal bruke, Lorenzkurver og Gini-koeffisienten. Alle disse mål faller inn under gruppen positive mål.

3.1 Noen enkle ulikhetsmål

Vi tenker oss en fordeling av inntekt over n personer, $i=1,\dots,n$.

r_i = inntekt for person nr. i

μ = gjennomsnittsinntekten, $\mu = \frac{1}{n}\sum_{i=1}^{N} r_i$

x_i = den relative andel av totalinntekten som tilfaller person i, $x_i = \frac{r_i}{n\cdot\mu}$

Variasjonsbredden er kan hende et av de aller enkleste mål for ulikhet. Det baserer seg på å sammenligne de ekstreme verdier i fordelingen, dvs. høyeste og laveste inntekt. Vi definerer variasjonsbredden slik:

$$E = \frac{1}{\mu}\cdot(\text{Maks}_i\ r_i - \text{Min}_i\ r_i) \qquad E \in (0,n)$$

Dette er et meget ufullstendig mål, idet det helt ignorerer fordelingen av inntektene mellom de to ekstremverdiene.

Tallverdiavviket tar utgangspunkt i inntektsnivået for hver enkelt person og sammenligner med gjennomsnittsinntekten, tallverdien av disse differansene summeres opp og sammenlignes så med totalinntekten:

$$M = \frac{1}{n\cdot\mu}\sum_{i=1}^{n}|\mu - r_i|$$

Min. verdi for M får vi når alle inntektene er lik gjennomsnittsinntekten. M blir da lik 0. Maks.verdi for M får vi når en person har hele inntekten:

$$\text{Maks } M = \frac{1}{n\mu}\left((n-1)\mu + |\mu - n\mu|\right)$$

$$= \frac{1}{n\mu}\left((n-1)\mu + \mu(n-1)\right) = \frac{2(n-1)}{n}$$

Altså vil $M \in (0, \frac{2(n-1)}{n})$

Svakheten ved det relative standardavviket er at det ikke er følsomt for transaksjoner mellom to personer så lenge inntektene deres befinner seg på samme side av gjennomsnittsinntekten. En slik overføring vil i det hele tatt ikke endre M.

Variansen summerer kvadratet av avvikene fra gjennomsnittsinntekten og veier dette i forhold til antallet:

$$V = \frac{1}{n} \sum_{i=1}^{n} (\mu - r_i)^2$$

En reduksjon i r_i og en tilsvarende økning i r_j vil alltid øke V dersom $r_i < r_j$ (men begge kan befinne seg på samme side av μ). Dette ser ut til å være en attraktiv egenskap for ulikhetsmål. Dalton påpekte alt i 1920 at dette var det minste en måtte kunne forlange av mål for ulikhet mellom inntektsfordelinger. Betingelsen er senere blitt kalt Pigou-Dalton betingelsen.

Svakheten ved variansen er at den avhenger av nivået på gjennomsnittsinntekten. En fordeling kan vise mye større relativ variasjon enn en annen og likevel ha lavere varians hvis gjennomsnittsinntekten er lavere enn for den fordeling en sammenligner med.

Varianskoeffisienten har ikke samme svakhet som variansen. Den er definert som kvadratroten av variansen delt på gjennomsnittsinntekten:

$$C = \frac{1}{\mu}\sqrt{V}$$

C oppfyller Pigou-Dalton betingelsen, og er uavhengig av nivået på gjennomsnittsinntekten. Varianskoeffisienten er like følsom overfor inntektsoverføringer på alle inntektsnivåer. Det vil si at overføringer mellom en person med 1 000 kr inntekt og en med 900 kr i inntekt teller likt med tilsvarende overføringer mellom en person med 1 000 100 kr i inntekt og en med 1 000 000 kr i inntekt. Er en slik nøytralitet en ønsket egenskap? Begynner vi å svare på slike spørsmål, kommer vi lett over på Atkinsons tanker; bak et hvert mål for ulikhet ligger det velferdsbetraktninger.

En kan også stille seg andre spørsmål. Hvorfor kvadrerer vi differensen mellom gjennomsnittsinntekten og person i's inntekt? Hvorfor ikke velge andre former som også ville oppfylle Pigou-Dalton betingelsen? Hvorfor sammenligne med gjennomsnittsinntekten?. Det er ikke sikkert at det er noen som har denne inntekten i det hele tatt. Ville det ikke vært mer naturlig å sammenligne hvert par av inntekter?

Logaritmisk standardavvik tillegger overføringer mellom "fattige" større vekt enn overføringer mellom "rike". Vi definerer denne størrelsen slik:

$$H = \left(\frac{1}{n} \sum_{i=1}^{n} (\log\mu - \log r_i)^2\right)^{\frac{1}{2}}$$

H bygger altså på et geometrisk gjennomsnitt i stedet for aritmetiske gjennomsnitt som vi har omtalt tidligere. Det logaritmiske standardavviket er lite følsomt overfor overføringer mellom svært rike personer. For ekstremt høye inntekter kan det skje at målet ikke oppfyller Pigou-Dalton betingelsen (Sen (1973) s. 32).

3.2 Gini-koeffisienten og Lorenz-kurver

Gini-koeffisienten er et mål for ulikhet som kan knyttes sammen med Lorenz-kurvene. Det er derfor naturlig å kommentere dem først[1].

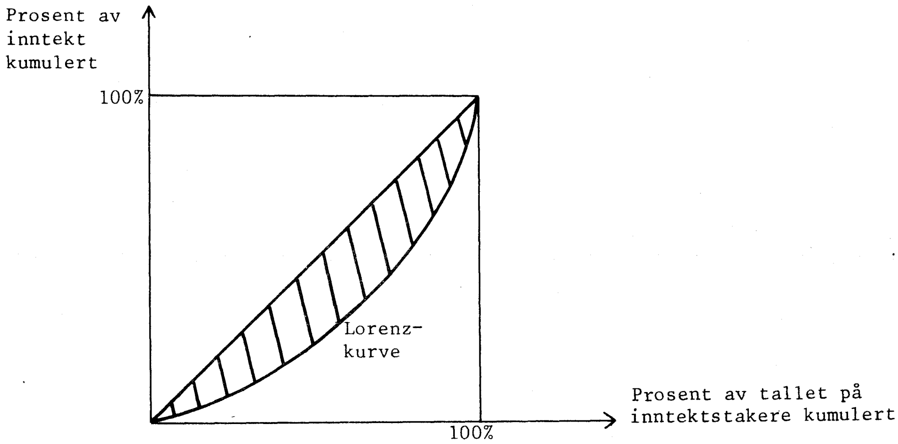
En Lorenz-kurve gir uttrykk for hvor stor prosent av de totale inntekter som tilfaller de p prosent av inntektstakerne med lavest inntekt. Denne prosenten må nødvendigvis være mindre eller lik p. Hvis den er lik p må alle inntektstakerne ha samme inntekt, og en kunne i så fall ikke rangere personene etter inntektens størrelse.

---

1) Se "Den personlige inntektsfordeling 1958, 1962 og 1967" s. 16 og Serck-Hanssen (1967) s. 22.

Grafisk kan en Lorenz-kurve tegnes slik:

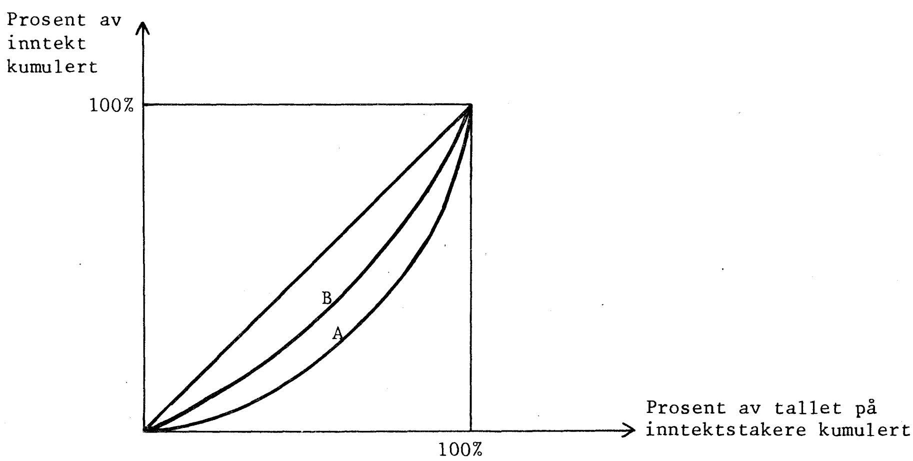
FIGUR 1.



Det er vanlig å bruke det skraverte arealet i figur 1 som et mål for inntektsfordelingens skjevhet. Jo mindre dette arealet er, desto jamnere vil en si at inntektsfordelingen er. Ettersom vi nærmer oss grensetilfellet, lik inntekt for alle inntektstakere, nærmer Lorenz-kurven seg 45^{o}-linjen i figur 1.

Hvis vi skal sammenligne to fordelinger ved hjelp av Lorenz-kurver og har følgende tilfelle:

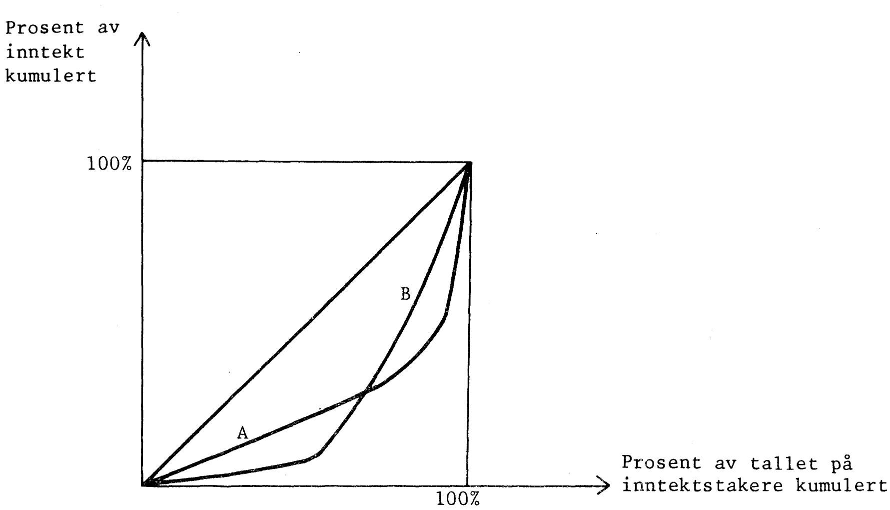
FIGUR 2.



Da vil vi entydig, når vi bruker arealet mellom 45^{o}-linjen og Lorenz-kurvene som mål for inntektsfordelingens skjevhet, slutte at inntektsfordelingen som Lorenz-kurve B er tegnet på grunnlag av er mer jamn enn inntektsfordelingen som Lorenz-kurve A er tegnet på grunnlag av. For Lorenz-kurver som ikke skjærer hverandre gjelder det altså at den Lorenz-kurve som ligger nærmest 45^{o}-linjen, er tegnet på grunnlag av den jamneste fordelingen.

En annen mulighet er at Lorenz-kurvene skjærer hverandre:

FIGUR 3.



Når Lorenz-kurvene krysser hverandre, kan vi ikke uten videre slutte noe om hvilken inntektsfordeling som er jamnest. For å komme videre med dette problemet, kan vi gå over til å se på Gini-koeffisienten.

Gini-koeffisienten er definert som forholdet mellom arealet mellom Lorenz-kurven og 45^{o}-linjen (det skraverte arealet i figur 1) og hele arealet under 45^{o}-linjen. Koeffisienten varierer da mellom 0 og 1. Den er lik 0 når alle inntektene er like og lik 1 når en person har hele inntekten. Det finnes flere måter å skrive Gini-koeffisienten på. En kan vise at den er lik akkurat halvparten av det aritmetiske gjennomsnittet av absoluttverdiene av differensene mellom alle inntektspar (Sen (1973) s. 31):

$$(4) \qquad G = \frac{1}{2n^2\mu} \sum_{i=1}^{n} \sum_{j=1}^{n} |r_i - r_j|$$

$$(5) \quad \Longleftrightarrow \quad G = 1 - \frac{1}{n^2\mu} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Min}\,(r_i, r_j)$$

$$(6) \quad \Longleftrightarrow \quad G = 1 + \frac{1}{n} - \frac{2}{n^2\mu}(r_1 + 2r_2 + \ldots + nr_n) \quad \text{for } r_1 \geqq r_2 \geqq \ldots \geqq r_n$$

Gini-koeffisienten er nok det mest direkte mål for ulikhet mellom inntektsfordelinger av dem vi har omtalt. Fordeling om gjennomsnittsinntekten inngår ikke, men koeffisienten tar hensyn til differansen mellom hvert par av inntekter. Kvadreringen som vi har omtalt tidligere, inngår heller ikke. Gini-koeffisienten oppfyller Pigou-Dalton betingelsen, men har ikke samme skjevhet som det logaritmiske standardavviket, noe som ofte er omtalt som en negativ faktor for Gini-koeffisienten. Når det foregår inntektsoverføring mellom to inntekter, er ikke utslaget i Gini-koeffisienten avhengig av disse inntektenes størrelse, men av hvor mange andre inntekter som ligger mellom disse to (se (6)).

3.3 45^{o}-linjen i Lorenz-diagrammet

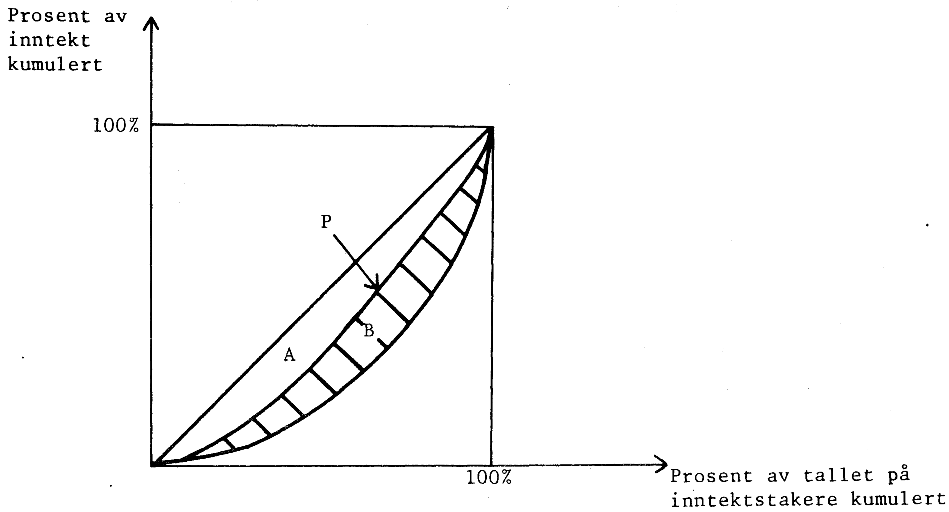
Det er ofte blitt kritisert å bruke 45^{o}-linjen i Lorenz-diagrammet uten videre som referanselinje når vi beregner Gini-koeffisienter. Vi ser på fordelinger av disponibel inntekt for personlige inntektstakere, og forutsetter at vi har perfekt likhet når Gini-koeffisienten er lik 0, dvs. når alle inntektstakere har samme inntekt. Dette er nok en situasjon som de fleste ikke vil kalle perfekt likhet. Faktorer som ulik forsørgelsesbyrde og alder kommer inn i bildet. En kan tenke seg problemet løst på flere måter:

Atkinson har skissert en metode som går ut på å ikke sammenligne inntekter for personlige inntektstakere direkte, men korrigere for forsørgelsesbyrde. Vi kan da tenke oss å sammenligne for eksempel inntekter for husholdninger pr. husholdningsenhet. Barn i lav alder vil da kunne gis mindre vekt enn voksne. På et slikt grunnlag vil det være mer akseptabelt å bruke 45^{o}-linjen i Lorenz-diagrammet som referanselinje[1].

Paglin (1975) har tatt et annet utgangspunkt. Han kritiserer 45^{o}-linjen ut fra en livssykelteori: "Når en gjør beregninger på tverrsnittsdata, kommer alle aldersgrupper med i undersøkelsen. Det er videre åpenbart at inntekt og behovet for inntekt varierer over livet for alle mennesker. En bedre måte å definere likhet på, er da å forutsette at vi har likhet når livsinntekten er den samme for alle mennesker. Så kan vi heller akseptere at inntekten varierer fra år til år, for eksempel at den er stor når en har stor forsørgelsesbyrde og avtar etter hvert som en blir eldre". En måte å konstruere en slik referanselinje på er å gruppere ulike aldersgrupper etter deres gjennomsnittsinntekt. De unge og de gamle vil da havne i bunnen av skalaen.

1) Se Atkinson (1974).

FIGUR 4.



Kurven P i figur 4 tenkes konstruert på en slik måte. For å få fram den nye Paglin-Gini-koeffisienten trekker vi arealet A fra arealet mellom Lorenz-kurven og 45^{o}-linjen, før vi dividerer på hele arealet av trekanten. Gini-koeffisienten blir da redusert med den del som skyldes at inntektstakerne befinner seg i ulike aldersgrupper. Når vi bruker en framgangsmåte som dette, må vi beregne en ny P for hver periode vi betrakter.

Vi skal i det følgende fortsette med å bruke 45^{o}-linjen som referenselinje. Det begrunnes med at vi ikke skal vurdere hvordan vi i situasjon 0 befinner oss i forhold til en idealsituasjon. Vi er bare interessert i å studere hvordan de forskjellige skattealternativene ligger i forhold til utgangssituasjonen, fortegnet på endringene så og si. Det er også tvilsomt om det er mulig å skissere en idealsituasjon de fleste vil være enige i.

3.4 Estimering av Gini-koeffisienten

Ved hjelp av programmet LOTTE finner vi hvor mange personer som har disponibel inntekt i det og det inntektsintervall. Vi tenker oss at vi har S inntektsintervall. Det siste intervallet er halvåpent, dvs. at det inneholder inntekter over en gitt grense.

Det er utviklet flere avanserte metoder for å estimere Gini-koeffisienten. Gastwirth (1972) beskriver hvordan en kan finne en øvre og nedre grense for Gini-koeffisienten. Forutsetningen er da at vi kjenner gjennomsnittsinntekten i hvert intervall. Den nedre grensen finner vi ved å sette alle inntektene i hvert intervall lik gjennomsnittsinntekten i intervallet (påvist av Morgan i 1962). En øvre grense finner vi ved å ta utgangspunkt i spredningen av inntekten i hvert intervall. For å komme fram på denne måten behøver en ikke tilnærme inntektsfordelingen til en bestemt fordeling.

Andre metoder tar utgangspunkt i nettopp en slik tilnærming. Den fordeling som har vært mest brukt er Paretofordelingen, men en har også foretatt en oppsplitting og tilnærmet nederste del for eksempel til en lognormal fordeling og øverste del til en paretofordeling[1].

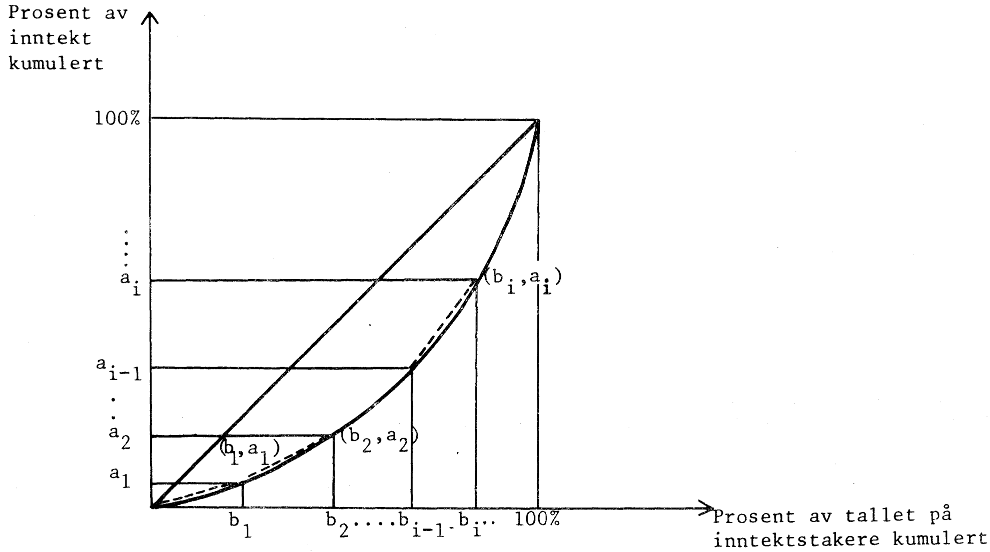
Vi skal beregne Gini-koeffisienten på en enklere måte. Først skal vi finne punktestimater slik at vi kan tegne Lorenz-kurven. Så beregner vi tilnærmet arealet under Lorenz-kurven ved å trekke rette linjer mellom punktene, og regner ut arealet av hvert av trapesene vi da får.
Vi får S trapeser totalt.

Arealet av i-te trapes blir: $\frac{a_{i-1}+a_i}{2} \cdot (b_i - b_{i-1})$ (se figur 5)

$(b_0 = a_0 = 0,\ b_S = a_S = 1)$

1) Se Kakwani og Podder (1973) og (1976).

FIGUR 5.



Arealet av alle trapesene[1]: $A = \sum_{i=1}^{S} \frac{a_{i-1} + a_i}{2} (b_i - b_{i-1})$

Gini-koeffisienten blir da: $G = (\frac{1}{2} - A)/\frac{1}{2} = \underline{1-2A}$

Ved å estimere G på denne måten får vi et anslag som er for lite. Hvor stort avviket blir, avhenger av krumningen på Lorenz-kurven, og hvor mange inntektsintervaller vi har.

LOTTE gir oss ikke uten videre de punktene vi trenger. Vi vet at så og så mange har disponibel inntekt i det og det inntektsintervall. La oss definere:

N_i = antall inntektstakere i intervall i.

Det totale antall inntektstakere blir da:

$$N = \sum_{i=1}^{S} N_i$$

Dermed finner vi b-ene i figur 5 slik:

$$b_k = \frac{\sum_{i=1}^{k} N_i}{N} \qquad k=1,\dots, S.$$

Hvis vi hadde kjent gjennomsnittsinntekten i hvert intervall, kunne vi finne a-ene i figur 5 slik:

R_i = gjennomsnittsinntekt i intervall i $\qquad$ i=1,..., S.

Total disponibel inntekt blir:

$$R = \sum_{i=1}^{S} N_i R_i$$

---

1) Det er foretatt en normalisering slik at kvadratet har areal lik 1.

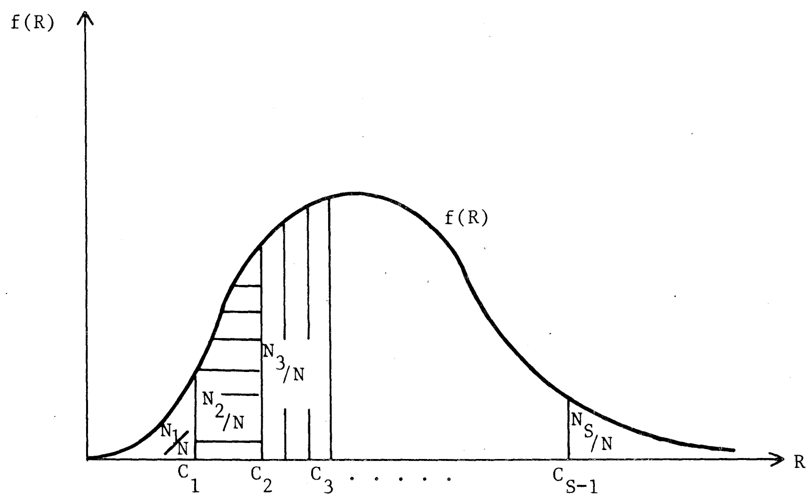
Da får vi:

$$a_k = \frac{\sum_{i=1}^{k} N_i R_i}{R} \qquad k=1,\ldots, S.$$

Det vi mangler er altså å få estimert R_i. Hvis vi har svært mange inntektsintervaller, kan vi tenke oss at vi bruker midtpunktet i hvert intervall som et anslag på gjennomsnittsinntekten. Da gjør vi en feil idet vi for de inntektsintervaller som ligger til venstre for typetallet gir et for lite anslag, og for de intervaller som ligger til høyre for typetallet gir et for stort anslag (under forutsetning av en en-toppet inntektsfordeling).

En bedre måte er å foreta en tilnærming via trapesmetoden. Dersom alle inntektsintervallene var like store kunne vi bruke Simpsons formel. Dette er imidlertid ikke tilfelle hos oss. Da kan vi gå fram på følgende måte:

FIGUR 6.



R = disponibel inntekt

$f(R)$ = tetthetsfunksjonen, $\int_0^{\infty} f(R)dR = 1$

c_i = øvre intervallgrense for i-te inntektsinterval1

$c_0 = 0$, C_S eksisterer ikke

Vi vet at:

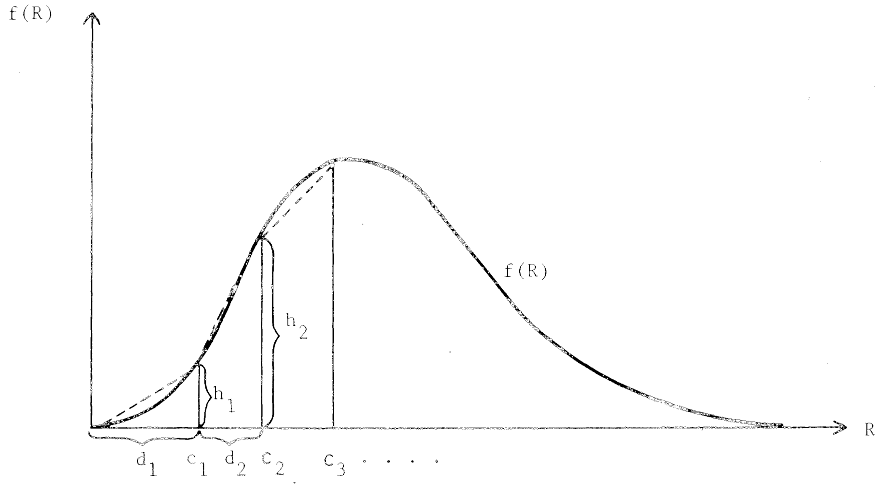
$$N \int_{c_{i-1}}^{c_i} f(R)\, dR = N_i \qquad N_i,\ N \text{ kjent}$$

Vi vil estimere forventet disponibel inntekt for hvert inntektsintervall (unntatt det øverste), og bruke dette som estimat for gjennomsnittsinntekten. Forventet disponibel inntekt for en person i intervall i er:

$$Y_i = \int_{c_{i-1}}^{c_i} Rf(R)dR$$

For å estimere Y_i tilnærmer vi inntektsfordelingen til trapeser:

FIGUR 7.



$d_i = c_i - c_{i-1}$ $\qquad\qquad (c_0 = 0,\ h_0 = 0)$

Første skritt er å beregne h-ene:

1. intervall: $\frac{1}{2}d_1h_1 = \frac{N_1}{N} \Rightarrow h_1 = \frac{2N_1}{d_1N}$

2. intervall: $\frac{1}{2}d_2(h_1+h_2) = \frac{N_2}{N} \Rightarrow h_2 = \frac{2N_2}{d_2N} - h_1 = \frac{2N_2}{d_2N} - \frac{2N_1}{d_1N}$

o.s.v.

På denne måte beregnes h-ene.

Likningen for den rette linjen vi tilnærmer til i hvert intervall blir:

$$\text{"rett linje i i-te intervall"} = (h_i + \frac{h_i - h_{i-1}}{d_i}(R-c_{i-1}))$$

Likningen for tettheten i i-te intervall blir da tilnærmet:

$$f(R) \approx (h_i + \frac{h_i - h_{i-1}}{d_i}(R-c_{i-1})) \qquad R\varepsilon\ (c_{i-1},\ c_i)$$

$$\text{La } \beta_i = \frac{h_i - h_{i-1}}{d_i}$$

Vi kan da beregne Y_i tilnærmet ved:

$$\tilde{Y}_i = \int_{c_{i-1}}^{c_i} R(h_i + \beta_i(R-c_{i-1}))dR$$

$$= \int_{c_{i-1}}^{c_i} (h_iR + \beta_iR^2 - \beta_ic_{i-1}R)dR$$

$$= \Big|_{c_{i-1}}^{c_i} (\tfrac{1}{2}h_iR^2 + 1/3\beta_iR^3 - \tfrac{1}{2}\beta_ic_{i-1}R^2)$$

$$= c_i^2(\tfrac{1}{2}h_i + 1/3\beta_i c_i - \tfrac{1}{2}\beta_i c_{i-1}) - c_{i-1}^2(\tfrac{1}{2}h_i - 1/6\beta_i c_{i-1})$$

Dette går bra for alle intervallene unntatt det øverste. Vi kan ordne intervallgrensene slik at svært få personer faller i det øverste intervallet, og så bruke skatteanslag i LOTTE for å beregne gjennomsnittsinntekten i dette ene intervallet. Når dette er gjort finner vi a-ene på tilsvarende måte som nevnt før:

$\tilde{Y}$ = total beregnet disponibel inntekt.

$$\tilde{Y} = \sum_{i=1}^{S} \tilde{Y}_i N_i$$

Vi får da:

$$a_k = \frac{\sum_{i=1}^{k} N_i \tilde{Y}_i}{\tilde{Y}} \qquad k=1,\dots, S$$

Dermed har vi alt vi trenger for å beregne Gini-koeffisienten.

3.5 χ^2-homogenitetstesten

Hvordan skal vi tolke de utslag vi får i Lorenz-kurver og Gini-koeffisienter ved de forskjellige skattealternativene? Er det bare tilfeldige utslag, eller er det utslag som tyder på mer dyptgående strukturelle endringer. Ved å bruke en χ^2-homogenitetstest kan vi komme et stykke på vei med dette spørsmålet.

χ^2-homogenitetstesten brukes til å sammenligne fordelinger, og vi kan bruke den til å teste om de inntektsfordelingene vi har under de ulike skattealternativene er signifikant forskjellige. For å bruke testen må vi ha mange observasjoner.

Vi tenker oss at vi har σ hyppighetsfordelinger for en variabel (disponibel inntekt) som er delt inn i s klasser. Klasseinndelingen må være den samme i alle σ fordelingene, og samtlige observasjoner forutsettes innbyrdes uavhengige. Vi definerer:

n_{ij} = antall observasjoner i den i-te klassen i den j-te fordelingen. $i=1,\dots, S,\ j=1,\dots, \sigma.$

$$n_{\cdot j} = \sum_{i=1}^{S} n_{ij} \qquad j=1,\dots, \sigma.$$

$$n = \sum_{j=1}^{\sigma} n_{\cdot j}$$

Videre må vi ha at antall observasjoner $n_{\cdot j}$ og dermed totaltallet n er gitte ikke-stokastiske størrelser. Vi tenker oss at det er en a priori ukjent sannsynlighet p_{ij} for en observasjon i klasse nr. i fra fordeling nr. j.

Da er:

$$\sum_{i=1}^{S} p_{ij} = 1 \qquad \text{for } j=1,\dots, \sigma.$$

Vi skal teste:

$$\begin{aligned} H_0:: p_{11} &= p_{12} = \dots = p_{1\sigma} \\ p_{21} &= p_{22} = \dots = p_{2\sigma} \\ &\vdots \\ p_{S1} &= p_{S2} = \dots = p_{S\sigma} \end{aligned}$$

mot en uspesifisert klasse av alternativ hvor minst to av likhetstegnene er byttet ut med "forskjellig"-tegnet $\neq$. Testen er videre beskrevet i H.T. Amundsen (1973). Hvis alle $n_{.j}$ er like store, lik $\frac{n}{\sigma}$, kan en vise at vi får en test med tilnærmet nivå ε hvis vi forkaster H_0 når:

$$Z = \sigma \sum_{i=1}^{S} \frac{1}{n_{i.}} \sum_{j=1}^{\sigma} n_{ij}^2 - n \geq Z_{1-\varepsilon,\ (\sigma-1)(S-1)}$$

der $Z_{1-\varepsilon,\ (\sigma-1)(S-1)}$ er fraktilen i χ^2-fordelingen med $(\sigma-1)(S-1)$ frihetsgrader.

4. Datamaterialet

Datamaterialet baserer seg på inntekts- og formuesundersøkelsen i 1973, og er hentet fra den del av undersøkelsen som er basert på opplysninger fra den ordinære skattelikning i 1973 for et utvalg av personlige skattytere.

Den statistiske enhet er "inntektstaker". Som inntektstaker regnes alle personer som ble behandlet som selvstendige enheter ved skattelikningen i 1973 (selv om de hadde null inntekt og formue). Ektefeller som hadde særskilt likning, regnes som to inntektstakere. Ektefeller som ble liknet felles, men som leverte separat oppgave, regnes også som to inntektstakere. Forsørgede personer regnes som inntektstakere dersom de ble særskilt eller selvstendig liknet i 1973[1]).

Utvalget av personer er sammensatt av to deler. Den ene delen består av alle inntektstakere som inngår i et tilfeldig utvalg av husholdninger. Dette er det samme utvalget som ble brukt i forbruksundersøkelsen i 1973, og omfatter i alt ca. 8 200 inntektstakere. Den andre delen består av alle inntektstakere som i 1972 hadde netto inntekt ved ordinær statsskattelikning pluss særfradrag større enn 150 000 kr, eller som hadde netto formue ved kommuneskattelikningen i bostedskommunen større enn 900 000 kr. Denne delen omfatter ca. 6 500 inntektstakere.

For de enheter som er med i utvalgene har likningskontoret ført over opplysningene fra selvangivelsene for 1973 til egne skjemaer. Alle disse skjemaene er lest inn på et bånd som så danner grunnlaget for regnemaskinprogrammet LOTTE.

5. Programmet LOTTE

LOTTE beregner skatt på individuelle inntekter. Utgangspunktet er at vi har opplysninger fra en rekke selvangivelser for 1973 som beskrevet i punkt 4.

Ved hver kjøring kan programmet operere med opptil fire skattesystemer; det første kan betraktes som referansesystemet og de andre som alternativer. Ved hvert alternativt system kan en fritt definere hvilke inntekter som skal beskattes, hvilke utgifter som kan brukes til fradrag og hvilke skatteregler som skal anvendes. For hvert alternativ beregner programmet gjennomsnittlig skatt for ulike grupper skattytere, og totalt skattebeløp for hele befolkningen. Det blir også beregnet disponibel inntekt for hver skattyter ved hvert skattealternativ, og dette blir skrevet ut i tabeller. Ellers foretas det en sammenligning med referansesystemet. Vi får skrevet ut tabeller der skattyterne er fordelt etter den endring de har fått i skatten. Noen av tabellene er meget detaljerte; vi får gruppert etter både inntekt, næring, familietype og sosioøkonomisk gruppe (lønnstakere, selvstendige og trygdede)[2]).

Vi er først og fremst interessert i tabellene for fordeling av disponibel inntekt. Enheten her er hva vi kan kalle en "skatteenhet", dvs. at ektefeller som blir liknet særskilt, regnes som to enheter. Ektefeller som blir liknet felles, regnes som en enhet, enten de har levert egen selvangivelse eller ikke. Forsørgede personer som blir liknet særskilt regnes som selvstendig enhet. Svakheten ved et slikt enhetsbegrep er diskutert i punkt 3.3.

---

1) Se Bjørklund (1975) og "Formuesstatistikk 1970" s. 9. 2) Se Garaas (1973) s. 2 og s. 4.

Programmet har sine klare begrensninger. En viktig forutsetning er at inntektene før skatt tas som gitte og upåvirket av skattesystemet. Det innebærer at vi forutsetter at det ikke vil komme noen reaksjon fra skatteyternes side på endringer i skattereglene. Spesielt antas arbeidstilbud og inntektsfastsettelse som upåvirket av en endring i marginalskattene.

Vi vil heller ikke få en korrekt fordeling av disponibel inntekt for personlige skattytere. Alle beregninger er basert på inntekter registrert ved Skattelikningen. Inntekter som er ulovlig unndratt beskatning, skattefrie ytelser fra folketrygden og andre skattefrie ytelser blir ikke tatt med i beregningene.

Begrensningene vi har omtalt, gjør at svært omfattende skatteendringer neppe vil kunne behandles tilfredsstillende ved dette programmet.

En siste svakhet jeg vil nevne, er at datamaterialet delvis er en utvalgsundersøkelse. Husholdningsutvalget "blåses opp" og legges til de med "høy inntekt/formue" slik at vi får uttrykk for totale skatter for personlige skattytere. Oppblåsningsfaktorene blir beregnet ut fra totaltall for personlige skattytere, netto inntekt i alt og netto formue ved kommuneskattelikningen i 1973.

For å få et overblikk over hvor gode skatteanslag LOTTE gir, har jeg tatt med beregningene for 1973 for de skatter vi bruker i denne undersøkelsen:

1973:	Beregnet i mill.kr	Faktisk i mill.kr	Avvik i %
Inntektsskatt til kommune og fellesskatt til Skattefordelingsfond	10249	10287	- 0.37
+ Inntektsskatt til stat	3097	3063	+ 1.11
+ Folketrygd, pensjonsdel	3494	3497	- 0.09
+ Folketrygd, sykedel	2051	2037	+ 0.69
- Barnetrygd	1418	1361	+ 4.19
- Forsørgerstønad	101	85	+18.82
+ Formuesskatt til kommune	366	377	- 2.92
+ Formuesskatt til stat	159	149	+ 6.71
= Totalt 1973-proveny	17897	17964	- 0.37

Avvikene skyldes delvis oppblåsningen og utvalget, men det er også mulig at det kan forekomme enkelte programfeil.

6. Beregningsmetoder og beregningsresultater

Utgangspunktet, situasjon 0, er fordeling av disponibel inntekt for personlige skattytere i 1973 under dagjeldende skattesystem. Vi har med følgende skatter: inntektsskatt til kommune, fellesskatt til skattefordelingsfond, inntektsskatt til stat, pensjonsdelen av folketrygden, sykedelen, barnetrygd, forsørgerstønad, formuesskatt til kommune og formuesskatt til stat. Utfra situasjon 0 foretar vi tre tenkte skatteendringer som alle forutsettes skal gi samme proveny som i situasjon 0. Vi skal så studere fordeling av disponibel inntekt under hvert alternativ og sammenligne alternativene med hverandre.

I situasjon 1 tenker vi oss at 73-reglene opprettholdes, men kommuneskatten til stat og kommune faller bort og erstattes med at inntektsskatten til staten øker med a%. Vi har før bestemt a(s. 7):

$$a = \frac{\text{formuesskatt til stat + formuesskatt til kommune}}{\text{inntektsskatt til stat}} = \frac{525 \text{ mill.kr}}{3097 \text{ mill.kr}} = 0{,}17 \text{ dvs. } 17\%.$$

Situasjon 2 tilsvarer situasjon 1, men her tenker vi oss at alle satsene i inntektsskatten til staten får et like stort tillegg på 2.42%, dvs. at laveste sats øker fra 10% til 12.42%, nest laveste sats fra 15% til 17.42% osv. De skattefrie beløp på 18 000 kr i kl. 1 og 27 000 kr i kl. 2 beholdes. For å finne fram til tallet 2.42% må vi foreta en del kjøringer med LOTTE. Først kan vi gi et vilkårlig tillegg på 5% og se hvilket proveny-utslag det gir, prøve på nytt, interpolere og foreta nye kjøringer til vi får et resultat som er tilfredsstillende.

I situasjon 3 tenker vi oss først at inntektsskatten til staten gjøres "flat" på 10%, dvs. at skattefritaking på 18 000 kr i kl. 1 og 27 000 kr i kl. 2 beholdes, og at det overskytende beløp beskattes med 10%. Dette reduserer ifølge LOTTE inntektsskatten til staten med 934 mill. fra 3097 mill til 2163 mill. Vi skal så tenke oss at dette blir kompensert ved at formuesskatten til staten øker. Skal vi ha samme proveny som i situasjon 0 må formuesskatten til staten da øke fra 159 mill. til 1093 mill. Jeg prøvde med flere forskjellige regler for formuesskatten. Det var ønskelig å justere formuesskatten slik at satsene ikke ble urimelig høye og at progresjonen i formuesskatten ble noenlunde opprettholdt. De skattefrie fradragene måtte derfor reduseres betraktelig. Til slutt ble jeg stående med en formuesskatt til staten der de skattefrie fradrag var på 10 000 kr i kl. 1 og 15 000 kr i kl. 2. Av det øvrige ble det trukket 2% av de første 250 000 kr, 3.4% av de neste 250 000 kr og 4.8% av det overskytende beløp. Forøvrig ble 1973-reglene opprettholdt i situasjon 3, men for å slippe problemet med skatt større enn skattepliktig inntekt innfører vi en 100%-regel, dvs. at de samlede skatter ikke må overstige nettoinntekten ved statsskattelikningen pluss særfradrag.

Ved å bruke LOTTE får vi følgende proveny-anslag for de 4 alternativene vi har skissert ovenfor:

	Proveny med LOTTE
Situasjon 0	17897 mill.kr
Situasjon 1	17898 "
Situasjon 2	17898 "
Situasjon 3	17899 "

Så skal vi gå over til å se på fordelingen av disponibel inntekt ved hvert alternativ. Ved hjelp av LOTTE får vi vite hvor mange inntektstakere som har disponibel inntekt i det og det inntektsintervall, dvs. kolonne X i Tabell I s. 16. Det kumulerte antall inntektstakere i %, tallene i kolonne B, finnes da direkte fra formelen:

$$b_k = \frac{\sum_{i=1}^{k} N_i}{N} \qquad k=1,\ldots, S$$

som vi har vist i punkt 3.4.

For å finne kumulert disponibel inntekt i %, tallene i kolonne A, var det meningen å bruke tilnærming ved den trapesmetoden som er vist i punkt 3.4. Dette førte imidlertid galt av sted, idet det var så mange som hadde disponibel inntekt lik null. Den feilen vi dermed gjør i 1. intervall ved trapesmetoden vil forplante seg videre i hvert av de andre intervallene. En mulig løsning er å la inntektsfordelingen starte et stykke opp på den loddrette aksen i stedet for i origo. Vi skal imidlertid velge en enklere løsning. Da vi har såpass mange intervaller og "fin" oppdeling, bruker vi midtpunktet i hvert intervall som et estimat på gjennomsnittsinntekten i intervallet. For det øverste intervallet, inntekter over 125 000 kr, har vi satt gjennomsnittsinntekten lik 200 000 kr, et cirka anslag ut fra de skatteanslag LOTTE gir. La:

$\bar{R}_i$ = midtpunktet i i-te intervall $\qquad i=1,\ldots, S.$

Tabell I

Disponibel inntekt i kr	Situasjon 0			Situasjon 1			Situasjon 2			Situasjon 3		
	X	B	A	X	B	A	X	B	A	X	B	A
0	67443	3.053	0.000	66594	3.015	0.000	66594	3.015	0.000	72102	3.264	0.000
0 - 2 999	105394	7.824	0.376	104490	7.745	0.373	104490	7.745	0.373	105150	8.024	0.376
3 000 - 5 999	112080	12.898	1.575	111183	12.778	1.565	111183	12.778	1.564	114165	13.192	1.601
6 000 - 8 999	114653	18.088	3.620	112850	17.886	3.581	112850	17.886	3.579	119636	18.608	3.741
9 000 - 11 999	300710	31.701	11.127	298319	31.391	11.043	298319	31.391	11.037	311267	32.698	11.536
12 000 - 14 999	239510	42.543	18.815	239467	42.231	18.745	239763	42.244	18.743	235983	43.381	19.133
15 000 - 17 999	199418	51.570	26.639	207268	51.614	26.891	210783	51.786	27.023	204958	52.659	27.198
18 000 - 19 999	127579	57.345	32.402	128699	57.440	32.717	126941	57.533	32.765	122564	58.207	32.752
20 000 - 21 999	137814	63.584	39.283	137781	63.677	39.609	138662	63.810	39.698	141021	64.591	39.815
22 000 - 24 999	201469	72.704	50.540	201385	72.793	50.883	200503	72.886	50.916	196135	73.470	50.807
25 000 - 27 999	161286	80.005	60.702	161246	80.093	61.062	160952	80.172	61.070	149531	80.239	60.257
28 000 - 30 999	131153	85.943	69.902	131666	86.063	70.315	134301	86.252	70.502	134014	86.305	69.685
31 000 - 33 999	95066	90.246	77.248	97068	90.447	77.830	94430	90.526	77.809	87699	90.275	76.483
34 000 - 36 999	66076	93.237	82.826	66893	93.475	83.487	66599	93.541	83.438	62196	93.091	81.748
37 000 - 39 999	43197	95.193	86.780	42886	95.417	87.421	40543	95.377	87.154	41090	94.951	85.521
40 000 - 54 999	86117	99.091	96.606	81798	99.119	96.677	82650	99.118	96.501	86791	98.880	95.353
55 000 - 69 999	12056	99.637	98.297	13126	99.714	98.631	11840	99.654	98.262	14373	99.531	97.495
70 000 - 84 999	5810	99.900	99.268	4083	99.898	99.385	4534	99.859	99.099	5032	99.758	98.425
85 000 - 99 999	1305	99.959	99.655	1118	99.949	99.631	1757	99.939	99.486	2740	99.882	99.030
100 000 - 124 999	420	99.978	99.767	801	99.985	99.846	620	99.967	99.652	1289	99.941	99.376
125 000 og over	489	100.000	100.000	324	100.000	100.000	731	100.000	100.000	1309	100.000	100.000
Sum:	2209045			2209045			2209045			2209045		

X = Antall inntektstakere som har disponibel inntekt i angitte intervall.

B = Kumulert antall inntektstakere i %.

A = Kumulert disponibel inntekt i %.

Total disponibel inntekt for hvert alternativ blir:

$$\hat{R} = \sum_{i=1}^{S} N_i \bar{R}_i$$

Vi beregner da kumulert disponibel inntekt, tallene i kolonne A, slik:

$$a_k = \frac{\sum_{i=1}^{k} N_i \bar{R}_i}{\hat{R}} \qquad k=1,\ldots, S$$

For å få et innblikk i hvor stor feil vi gjør når vi bruker denne metoden, kan en kjøre ut total disponibel inntekt i situasjon 0 ved hjelp av LOTTE. Total disponibel inntekt blir:

TD = 41705 mill.kr

Når vi bruker midtpunktet i hvert intervall som anslag på gjennomsnittsinntekten får vi:

	$\hat{R}$	Avvik fra TD i %	Avvik fra sit.0 i %
Situasjon 0	42058 mill.kr	+ 0.85	0.00
Situasjon 1	41978 "	+ 0.65	- 0.19
Situasjon 2	42003 "	+ 0.71	- 0.13
Situasjon 3	41931 "	+ 0.54	- 0.31

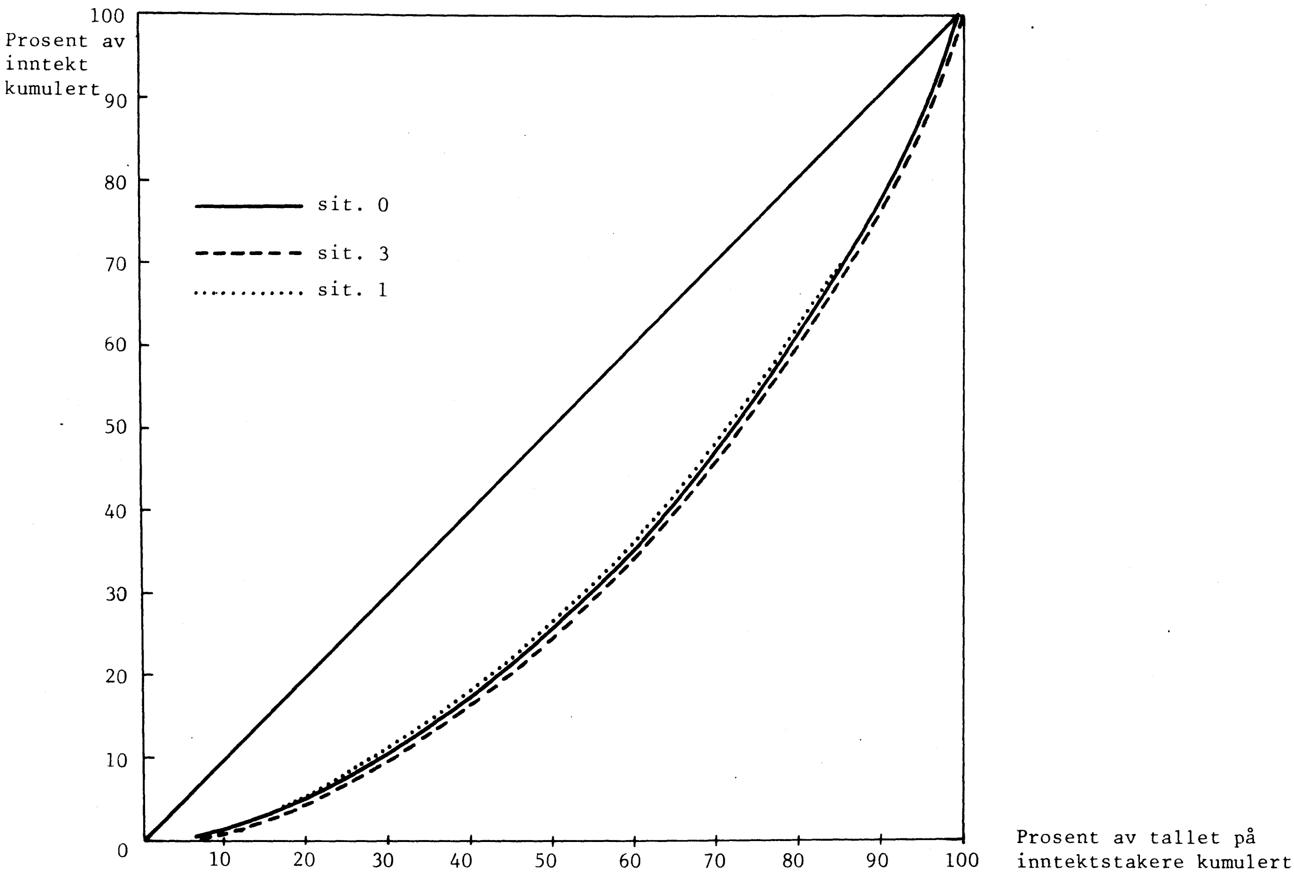
Selv om avvikene er så små (under 1%), kan avvikene innenfor hvert intervall være mye større ved at feilene totalt opphever hverandre.

Ved å bruke tallene i kolonne A og B i tabell I kan vi så tegne Lorenz-kurvene direkte. I figur 8 har jeg tegnet Lorenz-kurver for situasjon 0, 1 og 3. Utslagene blir svært små, men Lorenz-kurven for situasjon 3 ligger i sin helhet under Lorenz-kurven for situasjon 0 og 1. Lorenz-kurven for situasjon 1 ligger nærmest 45^o-linjen i det området den er tegnet. Der den ikke er tegnet ligger den så tett opptil kurven for situasjon 0 at vi ikke kan skilne dem på tegningen. Vi har ikke tegnet kurven for situasjon 2. Det ser ut som om den ligger mellom kurvene for situasjon 1 og 0, men skjærer kurven for situasjon 0 nær endepunktene.

På basis av de punktene vi bruker når vi tegner Lorenz-kurvene, beregner vi så Gini-koeffisienten. Framgangsmåten her er vist i punkt 3.4. Vi tegner rette linjer mellom punktene på Lorenz-kurven og tilnærmer med trapeser. Da får vi:

	Gini-koeffisient	Avvik fra sit.0 i %
Situasjon 0	0.349414	0.00
Situasjon 1	0.346430	- 0.85
Situasjon 2	0.347504	- 0.55
Situasjon 3	0.358082	+ 2.48

FIGUR 8. Lorenz-kurver for situasjon 0, 1 og 3.



Til slutt skal vi bruke en χ^2-homogenitetstest for å teste om de utslag vi har fått i fordelingene er tilfeldige eller av mer signifikant karakter. Testen er beskrevet i punkt 3.5. Vi kan sette alle fordelingene opp mot hverandre samtidig, eller teste to og to mot hverandre. Vi skal gjøre det siste. Det vi gjør er at vi tester om observasjonene fra to alternativer kan tenkes å være observasjoner fra samme fordeling, mot at de ikke er det. Vi forkaster nullhypotesen vår hvis observatoren:

$$Z = \sigma \sum_{i=1}^{S} \frac{1}{n_{i\cdot}} \sum_{j=1}^{\sigma} n_{ij}^2 \qquad -n \geq Z_{1-\varepsilon,\ (\sigma-1)(S-1)}$$

der $Z_{1-\varepsilon,(\sigma-1)(S-1)}$ er fraktilen i χ^2-fordelingen med $(\sigma-1)(S-1)$ frihetsgrader. σ er antall hyppighetsfordelinger og S er som før antall inntektsklasser. Når vi tester to og to fordelinger mot hverandre, er $\sigma=2$. Antall inntektsklasser er 21. Bruker vi 5%-nivå blir da:

$$Z_{1-\varepsilon,\ (\sigma-1)(S-1)} = Z_{0,95,20} = 31.41$$

Vi får:

H_0:	Z:	Konklusjon:
Obs. fra sit.0 og sit.1 fra samme fordeling	11.72	Ingen
" " " 0 " " 2 " " "	10.86	Ingen
" " " 0 " " 3 " " "	35.20	Forkastning av H_0
" " " 1 " " 2 " " "	8.36	Ingen
" " " 1 " " 3 " " "	41.02	Forkastning av H_0
" " " 2 " " 3 " " "	29.20	Ingen

Denne testen forutsetter at samtlige observasjoner er innbyrdes uavhengige. Dette vil nok ikke helt gjelde i vårt tilfelle. Vi starter med situasjon 0, og foretar så skatteendringer utfra denne utgangssituasjonen. Det medfører at disponibel inntekt for en person ved et alternativ er avhengig av disponibel inntekt for samme person i utgangssituasjonen.

7. Kommentar til resultatene

Vi har foretatt en vurdering av tallmaterialet på tre måter; ved Lorenz-kurver, Gini-koeffisienten og χ^2- homogenitetstesten. De forutsetninger som ligger bak disse metodene, og forutsetninger for bruk av LOTTE, er ikke fullstendig oppfylt hele veien. En må derfor utvise forsiktighet når en trekker konklusjoner. Resultatene fra de tre måtene å vurdere tallmaterialet på strider ikke mot hverandre på noe felt.

Ved Gini-koeffisientene får vi en fullstendig rangering av inntektsfordelingene ved de fire alternative skattesystemene. Den jamneste inntektsfordelingen har vi i situasjon 1. Så følger situasjon 2, situasjon 0, og vi ser også at den mest ujamne inntektsfordelingen får vi i situasjon 3.

Lorenz-kurvene gir oss ikke så klare skiller. Det er klart at vi i situasjon 3 har den mest ujamne inntektsfordelingen. Lorenz-kurvene for de tre andre alternativene ligger alle hele veien nærmere 45^o-linjen, men de ligger så tett inntil hverandre at det er vanskelig å foreta noen rangering mellom dem. Det ser også ut som om vi får skjæring flere steder.

χ^2-homogenitetstesten støtter opp om dette bildet. Våre H_0-hypoteser om at observasjoner fra situasjon 0 og fra situasjon 3 er fra samme fordeling, og at observasjoner fra situasjon 1 og fra situasjon 3 er fra samme fordeling, må forkastes. Tester mellom situasjon 0, 1 og 2 fører ikke til noe resultat; situasjon 2 mot situasjon 3 gjør heller ikke det[1].

På forhånd kunne vi ventet oss jamnere inntektsfordeling i situasjon 1 enn i situasjon 2. I begge alternativene er formuesskatten sløyfet og inntektsskatten til staten økt for å kompensere dette. I situasjon 1 øker inntektsskatten for alle skatteytere med en gitt prosent. Det er det samme som at satsene for hvert progresjonstrinn får en like stor prosentvis økning. I situasjon 2 får alle satsene et like stort tillegg. Dette vet vi er mindre inntektsutjamnende enn å gi hver sats en like stor prosentvis økning (når endringene skal tilsvare hverandre provenymessig). Gini-koeffisienten gir et utslag i overensstemmelse med dette, men ved Lorenz-kurvene og χ^2-homogenitetstesten er ikke denne effekten sterk nok til å gi noe resultat.

Konklusjonen det skulle være mulig å trekke etter denne undersøkelsen, er at overgang til et skattesystem der inntektsskatten til staten blir redusert og formuesskatten til staten økt tilsvarende slik som beskrevet i situasjon 3 i 1973, ser ut til å føre til at inntektsfordelingen blir mer ujamn. Når en sammenlikner situasjon 3 med situasjon 0, er det to momenter som er av betydning. Det ene er selvsagt i hvilken grad formuen er korrelert med inntekten. Det andre som spiller inn, er hvor sterk progressiviteten i inntektsskatten til staten er i forhold til progressiviteten i formuesskatten til staten. I vårt tilfelle ser det ut som om progressiviteten i inntektsskatten til staten har vært av avgjørende betydning når det gjelder å skape den jamneste inntektsfordelingen.

For å belyse disse sammenhengene nærmere, kunne det være aktuelt å bringe inn i analysen forholdet mellom inntekt og formue. Ved hjelp av LOTTE kunne en kjøre ut tabeller som viser sammenheng mellom disponibel inntekt eller bruttoinntekt og formue. Det kunne også være av interesse å se på fordelingsvirkningene av skatteendringer for ulike grupper av inntektstakere; selvstendige, lønnstakere og trygdede.

1) Testen gir imidlertid ikke noen rangering av fordelingene.

8. Litteraturhenvisninger:

Amundsen, H.T. (1973): "Statistisk metodelære II".

Atkinson, A.B. (1970): "On the Measurement of Inequality" Journal of Economic Theory, no. 2, 1970.

Atkinson, A.B. (1974): "Poverty and income inequality in Britain". Fra "Poverty, Inequality and Class Structure". Cambridge University Press 1974.

Bigsten, A. (1976): "Att mäta inkomstojämlikhet". Ekonomisk Debatt nr. 4, 1976.

Bjørklund, O. (1975): "Inntekts- og formuesundersøkelsen 1973". Notat.

"Den personlige inntektsfordeling 1958, 1962 og 1967". Statistiske analyser nr. 2.

Engebretsen, J.D. (1974): "Om 80 prosent-regelen i statsskattevedtaket". Arbeidsnotat fra SSB. IO 74/19.

"Formuesstatistikk 1970". NOS A 542.

Garaas, E. (1973): "LOTTE - en revidert verson av programmet for beregning av skatt under ulike inntektsdefinisjoner og skatteregler". Arbeidsnotat fra SSB. IO 73/32.

Gastwirth, J.L. (1972): "The Estimation of the Lorenz Curve and Gini Index". Review of Economics and Statistics 1972.

"Historisk oversikt over skattesatser m.v.". Arbeidsnotat fra SSB. IO 76/17.

Kakwani, N.C. og Podder, N. (1976): "Efficient estimation of the Lorenz Curve and associated inequality measures from grouped observations". Econometrica No. 1, 1976.

Kakwani, N.C. og Podder, N. (1973): "On the estimation of Lorenz Curves from grouped observations". International Economic Review No. 2 1973.

Paglin, M. (1975): "The Measurement and Trend of Inequality: A Basic Revision". American Ec. Rev. Vol. 65.

Sen, A (1973): "On Economic Inequality".

Serck-Hanssen, J. (1967): "Trekk av fordelingslæren". Haavelmo's forelesninger høsten 1954.