

# Arbeidsnotater

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

D r o n n i n g e n s g t. 16, O s l o - D e p., O s l o 1. T l f. 41 38 20

IO 74/24

27. mai 1974

## METODEHEFTE NR. 12

Notater om sosial og demografisk statistikk  
om norske fødselsrater, og om befolkningsprognosearbeidet

### INNHold

	Side
Forord .....	2
Odd Aukrust: "Tanker om et 'Personregnskap for Norge' og bruken av det for analyse og forskning". (Utdrag) (Auk/IVG, 12/4-67) .....	3
Svein Nordbotten: "Merknader til notat Auk/IVG, 12/4-67: Tanker om et personregnskap ... ." (Utdrag) (SH/ea, 9/6-67) .	12
Svein Gåsemyr: "Bakgrunnen for innføring av fast identifi- kasjonsnummer for personer". (SH/eh, 3/1-73) .....	15
Jan M. Hoem og Eivind Hoffmann: "Strukturen i et System for Sosial og Demografisk Statistikk (SSDS). En momentliste". (JMH/EH/IH, 7/5-74) .....	17
Jan M. Hoem, Erling Berge og Liv Hansen: "Norske fødselsrater for ettårige aldersklasser, 1968-1972". .....	19
Erling Berge: "Analytisk glatting av fødselsrater for Norge for åra 1968 til 1972". .....	34

*Ikke for offentliggjøring. 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 Sentralbyrås oppfatning.*

## FORORD

Metodehefter i serien Arbeidsnotater.

I tilknytning til mange prosjekter i Statistisk Sentralbyrå utarbeides det mindre, upretensiøse notater for avklaring av spørsmål av metodisk interesse. Det kan dreie seg om utvalgsteknikk, alternative spørsmålsformuleringer, presentasjonsmetoder, begrepsavklaringer, diskusjon av "funn" i data, systemideer eller andre temaer. Selv om mange slike notater bare har begrenset interesse i ettertid, vil det blant dem være noen som kunne fortjene å bli alminnelig tilgjengelig. Det kan også være nyttig å ha dem registrert sentralt slik at det blir lettere å få oversikt over det stoffet som foreligger, og lettere å referere tilbake til det. Byrået publiserer derfor leilighetsvis et passende antall notater av dette slaget samlet i metodehefter i serien Arbeidsnotater.

Kontorlederne bes holde øynene åpne for denne nye publiseringsmuligheten.

Forsker Jan M. Hoem er redaktør av metodeheftene. Fullmektig Liv Hansen er redaksjonssekretær. Medarbeidere i Byrået som lager stoff som kan være aktuelt, bes sende dette til redaksjonen etter hvert som det blir ferdig. Retningslinjer for utformingen av inserater i metodeheftene finnes på side 46 til side 47 i Metodehefte nr. 9 (ANO IO 73/36).

TANKER OM ET "PERSONREGNSKAP FOR NORGE" OG BRUKEN AV DET  
FOR ANALYSE OG FORSKNING

(Utdrag)

av Odd Aukrust

I n n h o l d

- I. Innledning
- II. Elementer i en aksiomatikk for et personregnskap
- III. Begrepssystem og kontoplan
- IV. Register, dataarkiv, publikasjoner
- V. Personmodeller
- VI. Analytiske anvendelser av personregnskapet
- VII. Innholdet i persondataarkivene
- VIII. Klassifikasjons- og definisjonsproblemer

I. Innledning

1. Jeg tar utgangspunkt i tre viktige trekk ved de senere års utvikling:

- (i) Det er innført et system av faste personnummer, som i prinsippet gjør det mulig - via personnummeret - å knytte sammen informasjon om en bestemt person hentet fra ulike kilder.
- (ii) Vi har fått EDB-utstyr som gjør lagring og uttak av store mengder informasjon mye lettere enn før.
- (iii) Det har oppstått nye behov (forskning, regionaldata m.m.) som krever større fleksibilitet i statistikktilbudet.

Fra et teknisk synspunkt er det nå fullt mulig å registrere en persons hele livshistorie på bånd. Dette vil føre til en revolusjon i produksjonen av personstatistikk. Det gir også nye muligheter for å tilfredstille eiterspørselen etter statistikk som er skreddersydd for bestemte behov f.eks. for analysen av problemer hvor personer - eller grupper av personer - er observasjonsheter.

2. Jeg skal nedenfor spekulere litt over hvilke perspektiver som åpner seg for oppstilling av et "Personregnskap for Norge" og for bruk av et slikt personregnskap i analyse og forskning. Jeg skal særlig ha for øye hva som kan være å lære av erfaringer fra nasjonalregnskapsarbeidet

og bruken av nasjonalregnskapsdata i økonomisk analyse. Etter oppdrag skal jeg også ha i tanken behovet for å utvikle en terminologi i det arbeidet vi står foran.

## II. Elementer i en aksiomatikk for et personregnskap

3. Jeg skal bruke ordet personregnskapet i vid forstand i betydningen "all integrert informasjon vi sitter inne med om personer". Vekten ligger her på ordet "integrert". Det er derimot likegyldig om informasjonen gjelder enkeltpersoner eller grupper av personer og i hvilken form den foreligger (på bånd, på hullkort, i publikasjoner). Analyse bygd på informasjoner som inneholdes i et personregnskap, vil jeg forsøksvis kalle personanalyse (som motsetning til økonomisk analyse).

4. I nasjonalregnskapsarbeidet gjør vi bruk av en aksiomatikk hvor "økonomiske objekter" er de enheter som studeres. I prinsippet studerer nasjonalregnskapet "universet (massen) av alle økonomiske objekter som noen gang har eksistert". I praksis er vi riktignok bare interessert i slike objekter som i en viss forstand har tilknytning til Norge og som har hatt eksistens i det tidsrom nasjonalregnskapet omfatter. Andre grunnleggende kategorier i nasjonalregnskapet er "sektorer", "tid" og "verdi". De økonomiske objekter har en tidsbegrenset eksistens, de har visse "medfødte" egenskaper (de kan være realobjekter eller finansobjekter), og de kan være utsatt for begivenheter som kan tidfestes og som interesserer oss i nasjonalregnskapet (de kan kjøpes og selges, konsumeres, skifte eier, etc.).

5. De enheter vi studerer i personregnskapet, er de enkelte individer eller personer. Den masse vi observerer, er i prinsippet alle personer som noen gang har levd på jorden. Men i praksis interesserer vi oss bare for de personer som har "hatt tilknytning til Norge" en eller annen gang etter at det sentrale personregister ble opprettet, dvs. etter 1. oktober 1964. Det at en person senere er død, gjør ikke at han går ut av massen; vi er sikkert interessert i å lagre informasjoner også om de døde.

Problem: Hva skal være den konkrete tolkning av uttrykket "ha tilknytning til Norge"? Jfr. sontringen hjemmehørende/tilstedeværende personer.

6. Personene har egenskaper og kan være utsatt for begivenheter som vi er interessert i å studere. Disse egenskaper og begivenheter er

kjennemerker ved personene. Slike kjennemerker kan være:

- (i) Personens fødselstidspunkt og dødstidspunkt, som til sammen fastlegges personens tidseksistens.
- (ii) Medfødte eller permanente egenskaper som kjønn, foreldreopphav, lyter, intelligenskvotient, tvilling, nummer i barneflokk.
- (iii) Begivenheter som personen er utsatt for. Det kan være demografiske begivenheter, som giftemål og skifte av bosted, begivenheter som karakteriserer personens utdanningsstandard (avlagte eksamener), hans helsestandard (sykdom, sykehusopphold), hans forhold til rettssystemet (straffe-fellelse), hans holdning til politiske partier (valgdeltakelse), til presse og kringkasting (avisabonnement, lyttervaner), hans økonomiske og sosiale status (ligningsansettelser, forhold til sosiale institusjoner, inntektsanvendelse, tillitsverv). Alle slike begivenheter kan tidfestes.

Problem: (a) Her er antydnet en viss systematisk inndeling av det jeg har kalt "begivenheter". Er en systematikk mulig? Kan den nyttes for å trekke praktiske grenser mellom demografisk forskning, utdanningsforskning, medisinsk forskning, kriminologisk forskning, valgforskning, sosial eller sosiologisk forskning? (b) Jeg har ikke innført sontringen kjennemerke/kjennemerkeverdi som Nordbotten gjør bruk av. Er sontringen begrepsmessig nødvendig?

7. Foruten massen av personer eksisterer det masser bestående av andre typer enheter som vi også er interessert i å studere, f.eks. bedrifter, foretak, biler, jordstykker, kommuner. Mellom to enheter hentet fra to forskjellige masser kan det bestå bestemte tilknytninger som jeg vil kalle relasjoner mellom enheter fra ulike masser. En person kan f.eks. arbeide i en bestemt bedrift, være styremedlem i et bestemt foretak, være bosatt i en bestemt kommune, eie en bestemt bil eller et bestemt jordstykke. Noen av disse relasjoner mellom personer på den ene side og enheter fra andre masser på den annen kan vi være interessert i å studere i personregnskapet. Relasjoner som interesserer oss, skal vi også oppfatte som kjennemerker ved personene. (De kan samtidig være kjennemerker ved de enheter fra andre masser som tilknytningen gjelder. At person A eier bilen x - som er det samme som at x eies av A - kan oppfattes som kjennemerke både ved A og ved x.

8. Vi kan forestille oss personene utstyrt med merkelapper hvor alle kjennemerker (dvs. alle egenskaper og begivenheter som interesserer oss) er angitt, slik at de kan observeres. Vi bruker kjennemerkene til å avgrense og definere klasser eller delmasser i personregnskapet. En opptelling av antall personer i en klasse vil gi et tall for klassenes størrelse. Opplysningen om dødstidspunktet tillater en fundamental todeling av massen på ethvert tidspunkt i gjenlevende og døde. Beholdningsstørrelser(-begrep) defineres som klasser av personer som på et bestemt tidspunkt enten (i) har visse angitte egenskaper og/eller (ii) forut for tidspunktet har vært utsatt for angitte begivenheter og/eller (iii) står i angitte relasjoner til enheter fra andre masser. Strømningsstørrelser (begrep) defineres som klasser av personer som i en bestemt periode er (i) utsatt for angitte begivenheter og/eller (ii) på en angitt måte endrer relasjonsforhold til enheter fra andre masser. Mellom beholdningsstørrelser og strømningsstørrelser består en logisk sammenheng gitt ved at det korresponderende til en beholdningsstørrelse kan defineres to eller flere strømningsstørrelser som representerer henholdsvis tilgangen til og avgangen fra beholdningen i perioden mellom observasjonstidspunktene.

9. Mellom to personer kan det bestå tilknytninger som jeg vil kalle "relasjoner mellom enheter fra samme masse", i dette tilfelle relasjoner mellom personer. De to kan være gift, stå i en foreldre-barn-relasjon, på annen nærmere angitt måte være blodsbeslektet, bo i samme bolig, stå i forsørger-forsørget-relasjon. Slike relasjoner mellom personer kan brukes til å definere nye enheter hver bestående av en eller flere personer, og nye masser bestående av slike enheter. Jeg skal kalle slike enheter for persongrupper ("flokker"?) og de tilhørende masser for persongruppe-masser. Eksempler på persongrupper er ektepar og husholdninger. Sosiologene gjør - tror jeg - utstrakt bruk av persongrupper avgrenset på forskjellig vis i sine studier av gruppe-atferd ("syklubber", "de sinte, unge menn", "verkstedklubber", "regjeringskollegiet", "de politiske partier" og andre grupper som er knyttet sammen ved en følelse av samhörighet, ved et kommandoforhold eller på annen måte).

10. Jeg har her oppfattet enhetene "persongrupper" som enheter hentet fra andre masser enn massen av "personer". En persongruppe, slik oppfattet, er ikke det samme som en "klasse av personer" avgrenset ved visse kjennemerker. Men mellom en "person" (som er en enhet i massen av personer) og en "persongruppe" (som er en enhet i massen av persongrupper av denne bestemte type) kan det bestå tilknytninger av det slag som jeg

(punkt 7) har kalt "relasjoner mellom enheter fra ulike masser". Personen A kan f.eks. "tilhøre" - etter nærmere angitte kriterier - persongruppen Y, f.eks. en bestemt husholdning innenfor massen av husholdninger. (Han kan samtidig "tilhøre" en annen persongruppe Z innenfor en annen persongruppemasse, f.eks. massen av verkstedklubber). Slike relasjoner kan vi da oppfatte som kjennemerker ved A som beskrevet i punkt 7. De kan i prinsippet observeres og inngå som del av personregnskapet.

### III. Begrepssystem og kontoplan

11. Jeg har hittil diskutert "personer", "kjennemerker" og "klasser av personer" som abstrakte, logiske begreper. Når disse logiske kategorier skal gis et konkret motstykke i et personregnskap, må vi bl.a.:

(i) Erstatte de aksiomatiske definisjoner med operasjonelle definisjoner; vi må ha regler for hvordan vi identifiserer en person, ta konkret standpunkt til hvilke kjennemerker som skal observeres, fastlegge hvordan kjennemerkene operasjonelt skal defineres.

(ii) Ta standpunkt til hvordan innhentet informasjon skal ordnes, lagres og eventuelt publiseres.

Det første er et spørsmål om personregnskapets (operasjonelle) begrepssystem. Det annet er et spørsmål om personregnskapets kontoplan (begrepet tatt så vidt at det inkluderer maskinavdelingens arkiveringsplan<sup>1)</sup>). Begrepssystem og kontoplan vil naturlig måtte avstemmes etter hverandre, men det vil likevel være ulike hensyn å ta ved valget av de to.

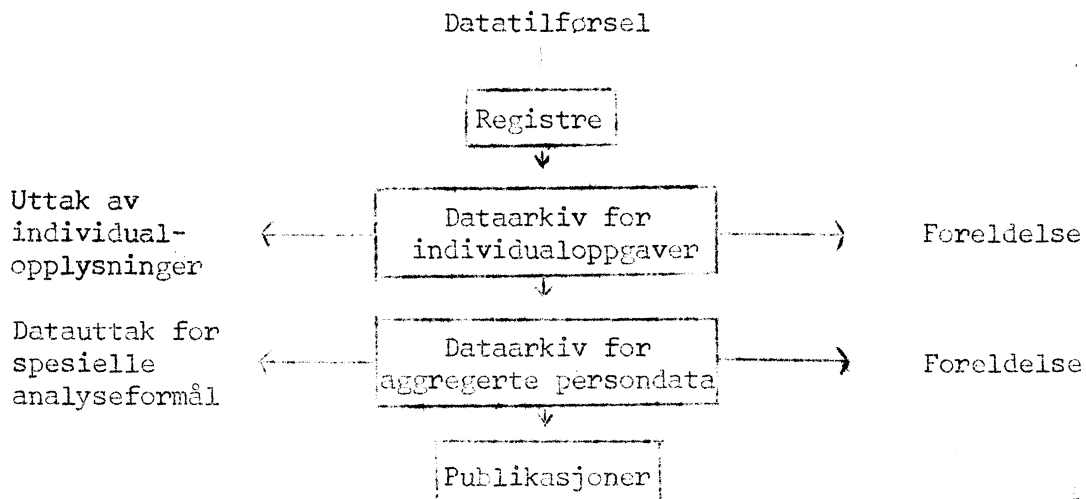
12. Valget av begrepssystem er et spørsmål om analytisk hensiktsmessighet, avveid - om nødvendig - mot observasjonstekniske problemer. Konkret dreier det seg om å avgjøre (i) hvilke arter av kjennemerker vi skal observere (ii) fastlegging av operasjonelle definisjoner som gjør observasjoner mulig (iii) fastlegging av klassifikasjonsstandarder. Drøfting av slike spørsmål må ta utgangspunkt i overveielser om hvilke analytiske og administrative anvendelser personregnskapet vil få.

13. Valget av kontoplan er et spørsmål om teknisk hensiktsmessighet. Konkret dreier det seg om - når det er gitt hva som skal observeres - hvordan informasjonen mest hensiktsmessig skal ordnes og lagres for å være tilgjengelig ved behov. Jeg har lite å si om dette, fordi problemene i stor utstrekning er maskintekniske. Jeg skal likevel sette noen betraktninger ned på papiret, bl.a. for å antyde et forslag til terminologi.

1) Sondringen synes å være identisk med en sontring mellom "logiske" og "fysiske" begreper som har vært nyttet av Nordbotten.

#### IV. Register, dataarkiv, publikasjoner

14. Det integrerte personregnskap vil - i fysisk, håndgripelig forstand - bestå av fire hoveddeler: (i) Register, (ii) dataarkiv for individualoppgaver, (iii) dataarkiv for aggregerte persondata, (iv) publikasjoner. Sammenhengen mellom de fire deler kan diagrammatisk representeres slik:



15. Registerne inneholder bare identifikasjonsopplysninger, dvs. personnummer sammen med navn og adresse i klarskrift for hver person, såvel for levende ("registret over levende") som for døde og utvandrede ("registret over døde"). Det kan bestå av et sentralt personregister som omfatter alle personer. Det kan også tenkes opprettet delregistre for deler av massen. Delregister i klarskrift (f.eks. omfattende alle personer i en kommune) vil det være naturlig å referere til som utskrifter av det sentrale personregister (som bare vil foreligge på bånd). Et og samme register kan være ordnet på ulike måter (etter nummer, alfabetisk på kommuner etc.) og kan således foreligge i flere utgaver. Tallet på "utgaver" som det kan være hensiktsmessig å ha, er et spørsmål om hvilke statistiske og administrative anvendelser registrene vil få.

16. Dataarkivene for individualoppgaver inneholder opplysninger om den enkelte person, identifisert ved personnummer. Det vil bestå av et sentralt persondataarkiv som omfatter alle personer og et utvalg av opplysninger om hver person. I tillegg til det sentrale persondataarkiv kan det opprettes et eller flere spesialiserte delarkiv for persondata. Et delarkiv vil bare omfatte deler av massen; til gjengjeld kan det inneholde opplysninger om den enkelte som ikke finnes i det sentrale persondataarkiv. Et dataarkiv kan, som et register, være ordnet på ulike



måter og således foreligge i flere utgaver. Opplysninger som inneholdes i ulike arkiver vil kunne sammenholdes med hverandre ved kollering.

17. Dataarkivet for aggregerte persondata er et arkiv for statistikk ("statistikk på hånd"), fremkommet ved sortering og opptelling av opplysninger hentet fra dataarkivene for individualoppgaver. Det inneholder mål for de beholdningsstørrelser som inngår i personregnskapets begrepssystem og for de tilhørende strømningsbegrep. Det er slik ordnet at opplysningene kan tas ut i ulike former, f.eks. som krysstabeller, tidsserier o.l. Ved at statistikken vil bygge på opplysninger som er "integreert i mikro" vil dette automatisk resultere i en statistikk som også er integreert.

18. Vi kan finne det hensiktsmessig å gjøre deler av dataarkivet for aggregerte persondata tilgjengelig for publikum i form av publikasjoner. En kan forestille seg både en oversiktspublikasjon for hele riket ("Personregnskap for Norge") som i sterkt aggregert form gir et utdrag av den informasjon som dataarkivet for aggregerte persondata inneholder ("Det trykte personregnskap") og forskjellige "personstatistiske spesialpublikasjoner", smlg. "det trykte nasjonalregnskap" og de forskjellige spesialpublikasjoner for økonomisk statistikk. En viktig klasse av spesialpublikasjoner vil kunne være "regionale personregnskap". Det trykte personregnskap og de tilhørende spesialpublikasjoner vil til sammen utgjøre et (trykt) integreert personstatistisk system.

## V. Om personmodeller

19. I realvitenskapene kan vi med fordel sondre mellom de konkrete fenomener som er studiefeltet for vitenskapen, og de teoretiske modeller vi danner oss for å "forklare" eller "vinne innsikt" i den virkelighet vi studerer. Den teoretiske modell består av begreper - representert ved begrepsord som "inngåtte ekteskap", "flytninger" o.l. eller kort og godt bokstavsymboler - og relasjoner mellom disse. Vi identifiserer teoriens begrepsord og tegn med bestemte observerbare fenomener eller variable. Hvis vi ved iakttagelse kan påvise at de relasjoner som modellen fastlegger mellom begrepene har empiriske motstykker i de fenomener vi observerer, godtar vi modellen som en tankemessig representasjon av virkeligheten.

20. De konkrete fenomener som interesserer oss i sosialvitenskapene og som skal studeres ved hjelp av personstatistikk, er forskjellige utslag (uttryksformer) for menneskelig adferd eller livscyklus. Det kan

være mikro-variable, dvs. variable som beskriver den enkelte persons adferd eller gir bakgrunn for denne adferd. De mikrovariable er hva vi foran har kalt kjennemerker ved personene. Det kan også - og vil som oftest - være makrovariable, dvs. variable som på en mer eller mindre komplisert måte er definert som funksjoner av de mikrovariable, f.eks. størrelsen av visse klasser ("menn over 15 år") eller hyppigheter ("studiefrekvenser for artianere"). Det er personregnskapets oppgave å presentere i systematisk form de observasjoner vi gjør av disse makrovariable.

Problem: Er det riktig at alle makrovariable i personanalysen kan defineres som størrelser av klasser eller som enkle funksjoner (hyppigheter o.l.) av slike størrelsestall?

21. Vi må som en fundamental forutsetning anta at de variable ikke antar verdier helt på slump, men at verdiene er knyttet sammen på en måte som ikke bare er rent definisjonsmessig. Det er disse sammenhenger som det er sosialvitenskapenes mål å klarlegge. En sammenheng (antatt eller konstatert) mellom personregnskapets (makro-)variable vil vi kalle en relasjon. Noen av disse er definisjonsrelasjoner som alltid må være oppfylt, f.eks. relasjoner mellom beholdningsvariable og tilhørende strømningsvariable eller relasjoner som sier at størrelsen av en klasse er lik den algebraiske sum av størrelsen av de sub-klasser som klassen er inndelt i. Andre er strukturrelasjoner som beskriver menneskelig adferd. Analogt med hva vi gjør i sosialøkonomisk analyse kan vi sondre mellom ulike typer strukturrelasjoner også i personanalysen. En mulig inndeling kan være:

- (i) Institusjonelt betingete relasjoner som beskriver sammenhenger som springer ut av lov eller avtale eller andre institusjonelle forhold, f.eks. relasjoner som beskriver hvordan antall elever i folkeskolen avhenger av antall barn i skolepliktig alder (lovregler som fastsetter påbud eller forbud knyttet til aldersgrenseregler gir antakelig de viktigste eksempler).
- (ii) Livscyklus-betingete atferdsrelasjoner, som beskriver sammenhenger som vesentlig er betinget av den menneskelige livscyklus, f.eks. relasjoner som beskriver hvordan sykdom, giftemål, inntekt, etc. varierer med alderen.
- (iii) Teknisk betingete atferdsrelasjoner, som beskriver sammenhenger som vesentlig er betinget av naturgitte eller tekniske forhold, f.eks. relasjoner som uttrykker en sammenheng mellom yrke og bosted.

(iv) Psykologisk betingede adferdsrelasjoner, som beskriver sammenhenger som vesentlig er betinget av den menneskelige psykologi, f.eks. relasjoner som uttrykker en sammenheng mellom foreldres sosialstilling og barnas utdannelsesnivå.

(v) Arvemessig betingede adferdsrelasjoner, som beskriver hvordan visse former av adferd vesentlig er betinget av nedarvede anlegg.

En slik inndeling kan sikkert ikke bli skarp. Svært ofte vil vi få å gjøre med sammenhenger som forklarer bestemte former for adferd ved en blanding av institusjonelt, livscyklus, teknisk, psykologisk og arvemessig betingede forhold.

22. Ved en personmodell vil jeg forstå et system av (en eller flere) relasjoner for menneskelig adferd som antas å gjelde samtidig, forutsatt at minst en av relasjonene i systemet er en strukturel relasjon. (Et system bestående bare av definisjonsrelasjoner er en kontoplan, ikke en modell). Som i sosialøkonomien kan det være hensiktsmessig å sondre mellom generelle og spesielle modeller, sterkt aggregerte og sterkt disaggregerte modeller, åpne og lukkede modeller, statiske og dynamiske modeller.

(Notatet ble avsluttet med noen betraktninger om analytiske anvendelser av personregnskapet og om innholdet i persondataarkivene. Disse avsnitt er i bearbeid form publisert som deler av Artikler nr. 34).

## Merknader til notat Auk/IVG, 12/4-67:

## Tanker om et personregnskap ...

(Utdrag)

av Svein Nordbotten

Om begreper og betegnelser

Alle er trolig enige om at grunnenheten er person. Alle data som kan knyttes til en person er det naturlig å kalle persondata. Grupper av personer utgjør aggregate enheter, som hittil har dominert samfunnsanalyser. Slike enheter vil jeg kalle befolkninger. En befolkning beskrives ved befolkningsdata som vanligvis vil være aggregate persondata. Befolkningsdata vil derfor være identisk med hva vi vanligvis kaller statistikk. En må imidlertid være klar over at en person kan karakteriseres ved egenskapene hos den eller de befolkninger vedkommende hører til og omvendt. Betegnelsen masse kan det være hensiktsmessig å reservere for den samling av enheter som analyser tar sikte på å trekke slutninger om.

En type av endringer i befolkninger skjer ved at personer "flytter" fra en befolkning til en annen. Dette gjør det mulig å etablere et regnskap med dobbelt bokholderi for befolkningene. Et slikt vil jeg kalle et befolkningsregnskap med oppgjørsenhet person. Det er andre typer av endringer i befolkningene som ikke lar seg bygge inn i et regnskap. Et slikt ikke-økonomisk regnskap har, slik jeg ser det, ingen mening som et bokholderi mellom personenheter. For personene må en bygge opp definisjonssystemer for de kjennemerker som skal gi et databilde av personen.

Sondringen mellom person og befolkning har også konsekvenser for modell-betegnelsene. En personmodell må referere seg til strukturen og atferd hos enkeltpersoner og krever persondata for testing og estimering. Eksempler på slike personmodeller i Byrådet er Møglestues, Østbyes og Steinås modeller, alle basert på data om personer. En befolkningsmodell er etter mitt syn en modell som refererer seg til en befolknings struktur og atferd og som krever befolkningsdata for testing og estimering. Giljes prognose-modell er et typisk eksempel. Grensen mellom de to modell-typer er neppe skarp, og det er først og fremst kravet til data

som tillater en klassifisering.

Data-arkivet er det sted en oppbevarer person- og befolkningsdata. Når det er behov for en sontring kan en nytte betegnelsene person-data-arkiv og befolkningsdata-arkiv. Personregisteret er en fortegnelse med identifikasjon over alle enheter som vi på et bestemt tidspunkt har lagret data for. Tilsvarende er befolkningsregisteret en fortegnelse over alle befolkninger vi vedtar å avlede data om. Struktur og ordning av data-arkivet er vesentlig for utnytting. En struktur som tillater data for samme person eller befolkning over tiden er en forutsetning for testing og estimering av dynamisk formulerte modeller. Dette krever permanente identifikatorer både for personer og befolkninger.

Til slutt om data og informasjon. Betegnelsen data kan det være nyttig å reservere for det som holdes lagret i motsetning til informasjon som er det analyser eller beregningene tar sikte på å gi som resultat.

#### Merknader til de enkelte avsnitt i notatet

Jeg bruker i det følgende din avsnittsnummerering:

3. - 10. Jeg ville være varsom med å kalle dette aksiomatikk.

4. Jeg vil som jeg allerede har sagt, finne sammenlikningen med nasjonalregnskapet lettere om du sier at det er bedriftene eller foretakene som er enheter og som karakteriseres ved de økonomiske objekter som er gjenstand for deres transaksjoner. Analytisk er vel hovedinteressen økonomisk virksomhet og atferd og det er bedriftene (foretakene) som har slik virksomhet og atferd, ikke objektene. Begrepet "sektor" nevnes uten at en ser relasjonen til person- eller befolkningsregnskap. Om du sier at bedriften er enheten, vil sektor svare til befolkning slik jeg har brukt det ovenfor.

5. Ditt problem er allerede løst formelt i lov og forskrifter om folkeregistrering. En annen sak er om løsningen tilfredsstillerdagens krav.

6. Er ikke fødsel og død (i) begivenheter (iii)? Når det gjelder sontringen kjennemerke/kjennemerkeverdi er dette et spørsmål om hva som er praktisk. Jeg mener det er nyttig med grovklassifisering i kjennemerker, "utdannelse" og "yrke" oppfattes f.eks. som to kjennemerker, mens "økonomisk embetseksamen" "lektor" er de respektive "verdier" som er observert for en enhet. Dette er også i samsvar med god gammel statistisk teori.

8. Her har du hva jeg vil kalle "befolkninger". En del av de

begreper du nytter her har også anvendelse for person-enheten.

11. Er betegnelsen "konto-plan" god? Den kan nyttes som en spesifisering for det jeg har betegnet befolkningsregnskap, men gir den riktige assosiasjoner når det gjelder lagring av data om enkelt-personer?

14. I samsvar med tidligere merknader vil jeg foreslå at du nytter teksten "Arkiv for persondata" og "Arkiv for befolkningsdata" i de to midterste ruter, alternativt en rute med "Arkiv for person- og befolkningsdata".

20. Til problemet har jeg følgende synspunkt. Alle makro-variable som opptrer i en befolkningsmodell kan trolig defineres som antydnet. Men i en befolkningsmodell kan det også opptre mikro-variable. La problemet være andel av stemmer avgitt i en befolkning (stemmeberettigede i en kommune) for et bestemt parti. En valgforsker vil kanskje stille opp en modell som sier at den avhenger av befolkningens størrelse og næringsstruktur som er typiske makro-variable en finner i kommunestatistikk. Dessuten antas det at de personlige egenskaper ved partiets første mann i befolkningen som er mikrovariable som kan knyttes til befolkningen, også må bygges inn i modellen. Dette gir en befolkningsmodell (makromodell) som omfatter både makro- og mikrovariable.

21. Som før nevnt tror jeg inndelingen kan vries i retning av:

- (1) Institusjonelle relasjoner
- (2) Biologisk - fysiologiske relasjoner, herunder medisinske reaksjonsrelasjoner og arvebetingede relasjoner
- (3) Atferdsrelasjoner, herunder psykologiske, sosiologiske og økonomiske relasjoner
- (4) Menings- eller opinions-relasjoner

SG/eh, 3/1-73

## BAKGRUNNEN FOR INNFORING AV FAST IDENTIFIKASJONSNUMMER FOR PERSONER

av Svein Gåsemyr

1. Forslag om et fast løpenummer for alle skatteyttere ble i sin tid framsatt av et samarbeidsutvalg for effektivisering i statsforvaltningen som næringsorganisasjonene hadde satt ned etter initiativ fra regjeringshold. Forslaget ble fremmet overfor Rasjonaliseringsdirektøren.
2. Finansdepartementet sendte dette forslaget til skattedirektøren til uttalelse. Skattedirektøren framholdt at spørsmålet hadde en bredde og et omfang som gikk langt ut over det skattemyndighetene hadde befatning med.
3. I brev av 27. februar 1961 fra Finansdepartementet fikk Byrådet i oppdrag å lede arbeidet med å etablere et fast løpenummer etter fødselsdata. Byrådet skulle samarbeide med de institusjoner saken hadde interesse for.
4. Byrådet tok på seg arbeidet med å lede utredningene og samarbeidet særlig med skattemyndighetene, Rikstrygdeverket, de militære myndigheter og Rasjonaliseringsdirektøren. Det var et samstemmig ønske om å etablere et permanent identifikasjonsnummer for personer på landsbasis. Bakgrunnen for dette ønsket var at flere myndigheter allerede hadde tatt i bruk løpenummer (personnummer) av forskjellig slag i sine administrative rutiner. Nummersystemene var ulik fra etat til etat. Ofte hadde samme etat ulike systemer i de enkelte kommuner og ofte skiftet en nummersystem fra et år til et annet. Myndighetene krevde at nummersystemene ble tatt i bruk av næringslivet, særlig arbeidsgiverne, i oppgavegivingen til myndighetene. Etter hvert som flere nummersystemer kom i bruk, ble det mer og mer problematisk for oppgavegiverne å holde orden på de enkelte nummerserier og faren for å forveksle nummersystemene ved innberetninger til myndighetene var stor. Mange bedrifter hadde også egne nummersystemer i sin egen personalforvaltning. Siden ulempene ved at flere nummersystemer

var i drift først ble merkbar i næringslivet var det naturlig at forslag om et nummersystem for personidentifikasjon som var felles for alle myndigheter og private institusjoner og som var permanent over tid, kom fra næringslivet.

5. Når myndighetene på dette tidspunkt hadde innført løpenummer som personidentifikasjon, var årsaken at en var gått over til å bruke hullkort i databehandlingsrutinene. Bruk av navn eventuelt med fødselsdata og adresseopplysning i tillegg som intern personidentifikasjon i hullkortrutiner, er særdeles upraktisk. I dag brukes det for det meste magnetbånd som datamedium. Også for magnetbåndteknikken er det store fordeler med nummer som personidentifikasjon. I prinsippet kan EDB-rutiner nytte navn som personidentifikasjon, men det er mange og til dels betydelige ulemper ved dette:
  - navn endrer seg over tid
  - det oppstår langt flere skrivefeil i navn enn i et nummer, et nummer kan dessuten ha innebygget kontroll-siffer slik at en stor del av skrivefeil kan avsløres før nummeret blir tatt i bruk i en konkret rutine
  - et nummer er nesten alltid kortere enn navn
  - det vil alltid forekomme at to eller flere personer har samme navn, dette gjelder også om en i tillegg til navn tar med fødselsdata og adresse i identifikasjonsbegrepet.
  
6. Allerede i oppdraget fra Finansdepartementet het det at det faste nummer for personidentifikasjon skulle bygge på fødselsdata. Byrådet sa seg enig i dette. Opplysninger om fødselsdata vil svært ofte kreves i alle tilfelle. Folk kjenner fødselsdato, bygger identifikasjonsnummeret på dette er det tilstrekkelig med 3 siffer i tillegg for å kunne tildele nummer til alle personer med samme fødselsdato. Fødselsnummeret inneholder dessuten 2 kontroll-siffer og er på 11 siffer i alt
  
7. Det sentrale personregister ble etablert for å gi en fortegnelse over tildelte fødselsnummer. Dette er fortsatt personregisterets primær-oppgave, men registeret nyttes også for statistiske og administrative formål. Enheten i registeret er alle tildelte fødselsnummer. Foruten fødselsnummer inneholder personregisteret opplysninger som kan identifisere den enkelte person, i første rekke navn og adresse og også noen andre opplysninger fra folkeregistreringen som en har funnet det praktisk at registeret inneholder.



JMH/EH/IH, 7/5-74

Strukturen i et  
System for Sosial og Demografisk Statistikk (SSDS)

En momentliste

av Jan M. Hoem og Eivind Hoffmann

I. Innledning

Denne momentlisten tar sikte på å strukturere diskusjonen i Byrådet om arbeidet med SSDS. Momentene gjelder dels slikt som må være med i SSDS-arbeidet, dels slikt som en må ta standpunkt til om en skal gjennomføre eller forbedre, og der en må bestemme seg for hvordan det eventuelt skal gjøres. Momentene knytter seg til Hoffmanns diskusjon i notatet "Arbeidet med et system for sosiodemografisk statistikk (SSDS) i Statistisk Sentralbyrå - dets framtidige omfang og plassering", (EH/TF, 17/4-74).

II. Momentliste

1. Systemets funksjon (formål)
  - 1.1. Generelt
  - 1.2. Forholdet til anvendelser av data
    - 1.2.1. Beskrivelser og analyser av mikro-fenomener (enkelt-personers handlinger og situasjon)
    - 1.2.2. Beskrivelser og analyser av makro-fenomener. (Eksempler; Samlet sysselsetting, samlet utdanningsetterspørsel.)
2. Systemets innhold
  - 2.1. Datamasse
    - 2.1.1. Individualoppgaver
    - 2.1.2. Oppgaver om grupper av personer
    - 2.1.3. Upubliserte tabeller
    - 2.1.4. Publiserte tabeller
    - 2.1.5. Analyser
  - 2.2. Dokumentasjonssystem (variabelkatalogen)
    - 2.2.1. Teknisk dokumentasjon om data
    - 2.2.2. Brukerveiledning
  - 2.3. Programsystem
    - 2.3.1. Dokumentasjon av ferdige programmer
    - 2.3.2. Beredskap for produksjon av nye programmer

3. Datamassens omfang og dekning
  - 3.1. Fagområder og temaer
  - 3.2. Tidaaspektet
  - 3.3. Befolkningsgrupper og geografiske områder, og institusjonelle tilknytninger (særlig for tilknytningen til nasjonalregnskapsdata)
  - 3.4. Spesifikasjonsgrad
4. Virksomheter
  - 4.1. Vedlikehold (av datamasse, variabelkatalog og programsystem)
  - 4.2. Betjening av brukere (i bruk av variabelkatalog, framskaffing og analyse av data, inklusive utarbeidelse av nye programmer)
  - 4.3. Forbedring av systemets kvalitet
    - 4.3.1. Omfang og dekning (kfr. pkt. 3)
    - 4.3.2. Produksjonstid for tabeller og analyser
    - 4.3.3. Pålitelighet (av oppgaver) og relevans (av problemstillinger, variable, befolkningsgrupper som er dekket)
    - 4.3.4. Kostnader ved ekspedering av forespørsler
    - 4.3.5. Presentasjonsform for statistikken (kfr. pkt. 2.1.)
5. Virkemidler
  - 5.1. Produksjonsrutiner
    - 5.1.1. For vedlikehold og brukerbetjening
    - 5.1.2. For kvalitetskontroll
    - 5.1.3. For dokumentasjon
  - 5.2. Samordning av de ulike statistikkområder
    - 5.2.1. Utarbeidelse av standarder
    - 5.2.2. Rådgivende samordningsutvalg
    - 5.2.3. Planlegger med samordningsoppgaver
  - 5.3. Planleggingsrutiner
    - 5.3.1. For utvidelse av omfang og dekning
    - 5.3.2. For forbedring av de øvrige sider av systemets kvalitet  
I begge tilfeller: a) Gjennom enkeltstående prosjekter  
b) Gjennom løpende statistikkproduksjon
6. Personell
  - 6.1. Kontinuerlig beskjeftiget med SSDS i snever forstand (kfr. pkt. B nedenfor)
  - 6.2. Sporadisk beskjeftiget (gjennom rådgivende utvalg o.l.)

### III. Kommentarer

A. Betegnelsen på systemet bør endres fra System for Sosiodemografisk statistikk til System for sosial og demografisk statistikk. Systemet skal omfatte mye mer enn det vi i Byrådet vil kalle sosiodemografiske statistikkområder.

B. På hvilken måte dreier SSDS seg om noe mindre enn Byrådets samlede arbeid med statistikk om sosiale og menneskelige forhold? Skal SSDS i virkeligheten være systemet for organiseringen av denne virksomhet? (Dette kan vi kalle SSDS-arbeid i vid forstand.) Skal det bare omfatte pkt. 1, 2.1., 2.2. (?) 3, 5.2.2., 5.2.3. (SSDS-arbeidet i snever forstand).

## Norske fødselsrater for ettårige aldersklasser, 1968-1972

av Jan M. Hoem, Erling Berge og Liv Hansen<sup>x</sup>

## INNHOLD

	Side
1. Innledning .....	19
2. Beregningsmetoder .....	20
3. Karakteristiske trekk ved kurvene .....	23
Referanser .....	28
Figurer .....	29

1. Innledning

I et annet notat (Hoem, 1974) minner en av oss om at aldersspesifikke befolkningsrater kan beregnes på en rekke forskjellige måter. Noen regneeksempler der viser at beregningsmåten har betydning for resultatet.

I inneværende notat skal vi illustrere dette ytterligere ved å gjengi diagrammer over norske fødselsrater for kvinner i ettårige aldersklasser for hvert av årene 1968 til 1972, beregnet etter ulike metoder. Ratene er blitt regnet ut i Statistisk Sentralbyrå ved forskjellige anledninger og for ulike formål. Ingen av dem er beregnet spesielt med samlet publisering for øye. Etter vår mening får man imidlertid frem interessante og karakteristiske likheter og forskjeller når man ser de ulike fødselskurvene under ett, så vi gjengir dem samlet i figurene 1 til 5.

---

x) Vi vil takke Helge Brunborg for kommentarer.

## 2. Beregningsmetoder

2A. På alle figurene representerer den heltrukne kurven fødselsrater pr. 1 000 kvinner i middelfolkemengden etter morens alder i fylte år ved nedkomsten. Disse ratene er beregnet etter formelen

$$(1) \quad \frac{F_x^u(n)}{\frac{1}{2}\{L_x(n-1)+L_x(n)\}}$$

De finnes publisert for årene 1966-1971 hos Brunborg (1973a), tabell 3.4.

Den stiplede kurven representerer fødselsrater pr. 1 000 kvinner i middelfolkemengden etter morens alder i fylte år ved utgangen av året. De er beregnet etter formelen

$$(2) \quad \frac{F_x^p(n)}{\frac{1}{2}\{L_{x-1}(n-1) + L_x(n)\}}$$

Siden ratene ikke har vært publisert før, gjengir vi tallene for årene 1968-1972 i tabell 1.

På figur 1 representerer den prikkede kurven fødselsrater pr. 1 000 kvinner ved årets begynnelse etter morens alder i fylte år ved nedkomsten. Formelen er

$$(3) \quad \frac{F_x^a(n)}{L_{x-1}(n-1)}$$

For noen formål er det hensiktsmessig å regne ut ratene etter formelen

$$(4) \quad \frac{F_x^p(n)}{L_{x-1}(n-1)}$$

For årene 1971 og 1972 er ratene beregnet etter formlene (2) og (4) så like at det ikke er mulig å se noen forskjell på de tilsvarende kurvene i et diagram. Vi har derfor bare plottet kurvene beregnet etter (2) og utelatt kurvene beregnet etter formel (4).

I alle tellerne har en regnet med levendefødte barn av begge kjønn.

2B. Etter det system for symbolisering av beregningsmetoden som er gjengitt i underkapittel 8D hos Hoem (1974), er de fire beregningsmetodene altså følgende:

Formel (1):



Formel (2):



Formel (3):



Formel (4):



Formel (3) er å betrakte som et substitutt for formel (4). Da beregningene etter formel (3) ble utført, forelå ikke fødsels-dataene for 1968 organisert etter morens alder ved årets utgang.

Tabell 1. Norske fødselsrater pr. 1 000 kvinner i middelfolkemengden etter morens alder i fylte år ved utgangen av året (formel (2)).

Alder	1968	1969	1970	1971	1972
15 .....	.7	.7	.7	.6	.7
16 .....	3.3	4.2	4.9	4.7	4.8
17 .....	19.5	18.6	20.0	22.6	23.8
18 .....	50.8	54.4	54.1	55.5	57.6
19 .....	87.8	91.6	92.2	92.5	92.3
20 .....	128.4	125.2	124.3	127.6	124.4
21 .....	157.2	155.1	150.9	150.3	141.0
22 .....	175.8	178.9	167.3	170.8	161.6
23 .....	191.7	185.9	174.5	177.5	172.0
24 .....	195.7	191.6	178.8	184.6	174.4
25 .....	192.2	189.3	177.3	173.5	173.6
26 .....	188.9	178.7	165.9	168.6	166.4
27 .....	172.6	172.0	158.5	157.0	153.4
28 .....	161.8	155.7	144.8	143.7	143.1
29 .....	148.4	140.9	129.5	129.1	125.2
30 .....	130.6	126.8	116.7	114.0	110.0
31 .....	114.6	114.9	107.8	103.1	96.6
32 .....	103.9	101.7	94.7	88.9	84.0
33 .....	94.3	86.2	80.0	76.5	71.5
34 .....	79.5	76.8	67.4	63.6	59.7
35 .....	68.8	68.0	60.5	60.4	52.6
36 .....	61.2	62.0	52.9	49.9	43.0
37 .....	52.8	48.8	44.5	44.3	39.3
38 .....	42.6	43.3	36.9	34.9	30.6
39 .....	37.4	36.4	29.1	28.5	23.9
40 .....	29.8	27.1	22.6	22.8	19.4
41 .....	22.2	21.2	16.7	15.4	14.6
42 .....	16.0	16.4	11.0	12.1	9.8
43 .....	10.9	10.3	8.3	6.9	6.0
44 .....	6.9	6.3	5.5	4.5	3.9
45 .....	4.0	4.2	2.9	2.3	2.2
46 .....	1.8	1.6	1.5	1.0	1.6
47 .....	.9	.7	.6	.5	.3
48 .....	.5	.3	.2	.2	.2
49 .....	.1	.0	.1	.1	.1

### 3. Karakteristiske trekk ved kurvene

3A. De kurvene vi gjengir, har enkelte iøyenfallende trekk som vi skal nevne. Vi skal da

- (i) sammenlikne de kurvene vi har for ett og samme observasjonsår,
- (ii) sammenlikne utviklingen av den enkelte kurve fra ett observasjonsår til et annet, og
- (iii) sammenlikne kokort og periodeeffekter i kurvene.

3B. Når antall fødsler organiseres etter morens alder i fylte år ved nedkomsten, slik at en får frem observasjoner av  $F_x^{\square}(n)$ , vil mødrenes eksakte alder gjennomsnittlig være et halvt år høyere ved nedkomsten enn når fødsleene organiseres etter morens alder i fylte år ved årets utløp, slik at en får frem observasjoner av  $F_x^{\diamond}(n)$ . Dette gjenspeiles tydelig i figurene ved at den heltrukne kurven ligger over den stiplede for de lavere aldre, mens den stiplede ligger høyest i de høyere aldre i den fruktbare perioden. Den prikkede kurven løper tilsvarende. Legg merke til at den prikkede kurven for 1968 er beregnet etter formel (3), og at den oppfører seg helt i tråd med den forklaring vi nettopp ga.

Brunborg (1973b, §§ 4. 3 og 6.2.3.) har før diskutert momenter som disse.

3C. Ratene i tabell 1 er altså beregnet etter formel (2). Tallene kan lett omorganiseres etter fødselskull, og en kan da undersøke om det store fallet i fruktbarheten fra 1968 til 1972 ser ut til å være dominert av en kohorteffekt eller av en periodeeffekt. Kohorteffekten vil være viktigst dersom endringene i fødselsratene i de enkelte aldre først og fremst er knyttet til fødselskullene. Periodeeffekten er viktigst hvis endringene i første rekke følger kalenderåret.

I tabell 2 har vi stilt opp endringene i den aldersspesifikke fødselsraten siden forrige kull for en del kohorter og aldre. Vi kan f.eks. lese der at fødselsraten pr. 1 000 23-årige kvinner sank med 11.4 fra kohorten født i 1946 til fødselskullet fra 1947. (I 1969 var fødselsraten for alder 23 i tabell 1 lik 185.9 I 1970 var den 174.5. Differansen er 11.4.)

Dersom det nå var noen merkbar kohorteffekt på fruktbarheten over disse årene, burde tallene i hver linje i tabell 2 være av noenlunde samme størrelsesorden. I hvert fall bør de ha samme fortegn. Det er mulig at

en kan spore en slik effekt for de eldste kullene her, men tallene for de yngste kohortene følger ikke noe slikt mønster. Det er vanskelig å se noen slik kohorteffekt for kullene født etter ca. 1945.

Det er mulig at kohorteffektene er mer langsiktige utviklings-tendenser, slik at en ikke kan vente å finne noen særlig utvikling fra en kohort til den neste, mens et mønster bør komme fram om en sammenlikner endringer kohortene i mellom for flere par av kohorter. I tabell 2 skulle dette i så fall gi seg utslag i at tallene i en og samme kolonne var av noenlunde samme størrelsesorden, og i hvert fall at de hadde fast fortegn. Igjen er det mulig at en kan spore et slikt mønster for kohortene fra før ca. 1945, men for de yngre fødselskullene finner en det ikke.

Diagonalene i tabell 2 gir oss aldersspesifikke endringstall fra ett kalenderår til det neste. Hvis det er noen utpreget periodeeffekt til stede, bør en kunne se et mønster i disse diagonalene. Et slikt mønster framtrer også ganske tydelig. Med unntak for de aller laveste aldersgruppene har det vært fruktbarhetsnedgang hvert år. Den sterke nedgangen fra 1969 til 1970 skiller seg klart ut ved at tallene i den tilsvarende diagonalen er sterkt negative.

Over årene 1968 til 1972 har det altså vært en dominerende periodeeffekt i fruktbarhetsutviklingen. I den utstrekning en her kan tale om en merkbar kohorteffekt overhodet, gjør denne seg bare gjeldende for kohortene født før ca. 1945.



Tabell 2. Endringer i aldersspesifikke fødselsrater siden foregående fødselskull

Fødsels- år	Alder																		
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1953 ...	0,9	1,4	1,4	-0,2															
1952 .....	-0,9	-0,3	0,3	-3,2															
1951 .....		3,6	0,6	3,3	-9,3														
1950 .....			3,8	-0,9	-0,6	-9,2													
1949 .....				3,2	-4,2	3,5	-5,5												
1948 .....					-2,1	-11,6	3,0	-10,2											
1947 .....						3,1	-11,4	5,8	0,1										
1946 .....							5,8	-12,8	-3,8	-2,2									
1945 .....								-4,1	-12,0	2,7	-3,6								
1944 .....									-2,9	-12,8	-1,5	-0,6							
1943 .....										-10,2	-13,5	-1,1	-3,9						
1942 .....											-0,6	-10,9	-0,4	-4,0					
1941 .....												-6,1	-11,4	-2,7	-6,5				
1940 .....													-7,5	-10,1	-4,7	-4,9			
1939 .....														-3,8	-7,1	-5,8	-5,0		
1938 .....																0,3	-7,0	-3,5	-3,9
																1968	1969	1970	1971
																-69	-70	-71	-72

K i l d e: Tabell 1.

3D. La oss til slutt sammenlikne utviklingen av de heltrukne kurvene over disse observasjonsårene. (De stiplede kurvene har en analog utvikling). Dette er år da en har hatt et kraftig fall i fruktbarheten her i landet. Det samlede fruktbarhetstall har utviklet seg som vist i tabell 3. (Se også diagram 1 hos Rideng og Tønnesen (1974)). Det var et jevnt fall i fruktbarheten fra 1966 til 1969, et kraftig fall i 1970, liten endring i 1971, og et nytt kraftig fall i 1972. (Foreløpige fall for 1973 tyder på et fortsatt fall).

Tabell 3. Samlet fruktbarhetsfall for Norge, 1966-72

	(1)	$\Delta$	(2)	$\Delta$
1966 .....	2 900			
1967 .....	2 804	96		
1968 .....	2 749	55	2 754	
1969 .....	2 693	56	2 696	58
1970 .....	2 498	195	2 504	192
1971 .....	2 488	10	2 489	15
1972 .....	2 384	104	2 378	111

(1) Fødselsrater pr. 1 000 kvinner i middelfolkemengde etter morens alder ved nedkomsten, summert f.o.m. alder 15 t.o.m. alder 49, dvs.

$$1\ 000 \sum_{x=15}^{49} F_x^{\square}(n) / \left\{ \frac{1}{2} L_x(n-1) + \frac{1}{2} L_x(n) \right\}.$$

Kilde for 1966-71: Brunberg (1973a), tabell 3.4.

(2) Tilsvarende summerte fødselsrater pr. 1 000 kvinner i middelfolkemengde etter morens alder ved utgangen av året, dvs.

$$1\ 000 \sum_{x=15}^{49} F_x^{\square}(n) / \left\{ \frac{1}{2} L_{x-1}(n-1) + \frac{1}{2} L_x(n) \right\}.$$

Kilde: Tabell 1.

Brunborg (1973a, § 3.4) har imidlertid vist at utviklingen er forskjellig i de ulike aldersklasser, og dette gjenspeiles i våre kurver.

(i) De yngste kvinnene, 20 år og under: Fødselsratene for disse aldre har vært forbausende stabile over årene 1968-1972. For de aller yngste (15-16 år) har det allikevel vært en liten oppgang.

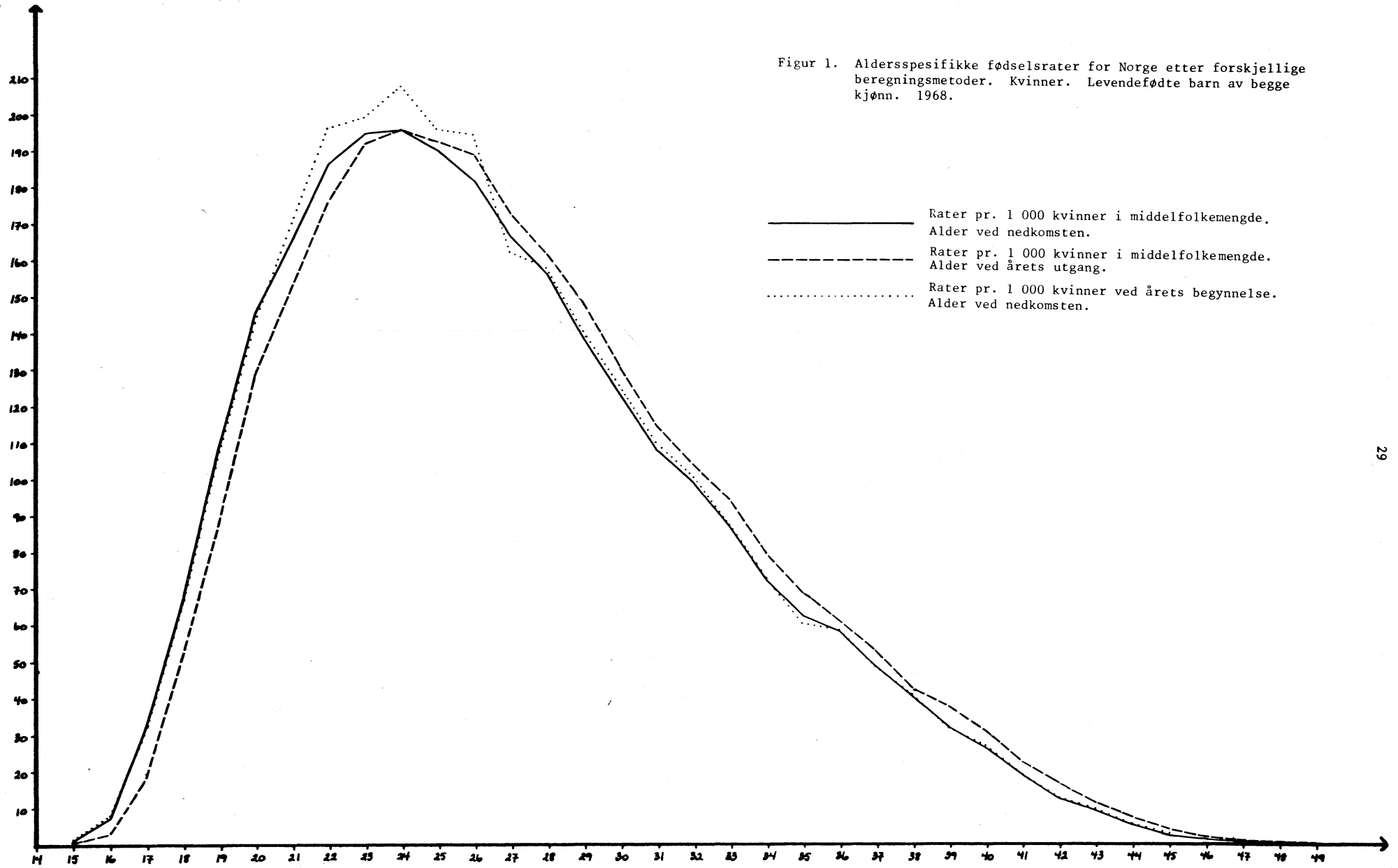
(ii) Kvinnene i de sentrale aldre, fra ca. 21 år til ca. 30 år: Fødselsraten gikk stort sett svakt ned fra 1968 til 1969 og sterkt ned i 1970. De var stort sett stabile i 1971, og gikk så ned igjen i 1972, mest i aldre under 25.

(iii) De eldre kvinnene, ca. 30 år og over: Fødselsratene har stort sett vist en jevn nedgang over perioden, men med relativt små endringer fra 1968 til 1969 og fra 1970 til 1971.

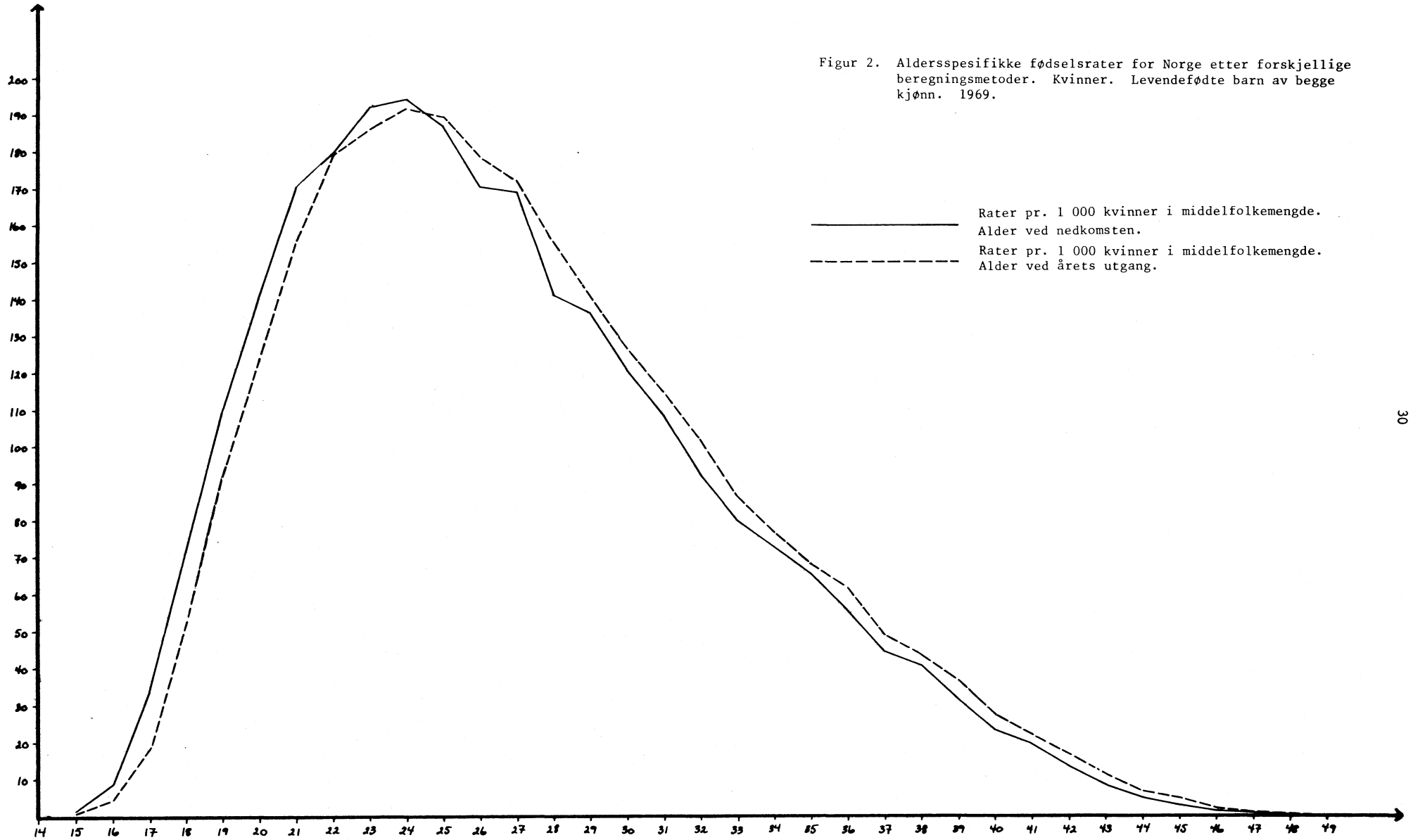
Referanser.

- [1] Brunborg, Helge (1973a): "Statistisk Sentralbyrås befolkningsprognosemodell VIII. Framskrivninga 1971-2000. Bakgrunnsmateriale og kommentarer til resultatene". Arbeidsnotat IO 73/14.
- [2] Brunborg, Helge (1973b): "Statistisk Sentralbyrås befolkningsprognosemodell IX. Framskrivninga 1971-2000. Teknisk dokumentasjon. Arbeidsnotat IO 73/15.
- [3] Hoem, Jan M. (1974): "Beregning av befolkningsrater". Arbeidsnotat IO 74/22.
- [4] Rideng, Arne og Bjørn Lied Tønnesen (1974): "Regionale befolkningsframskrivninger. Nåværende opplegg og utviklingsplaner. 1974". Statistisk Sentralbyrå, Artikler, nr. 62.

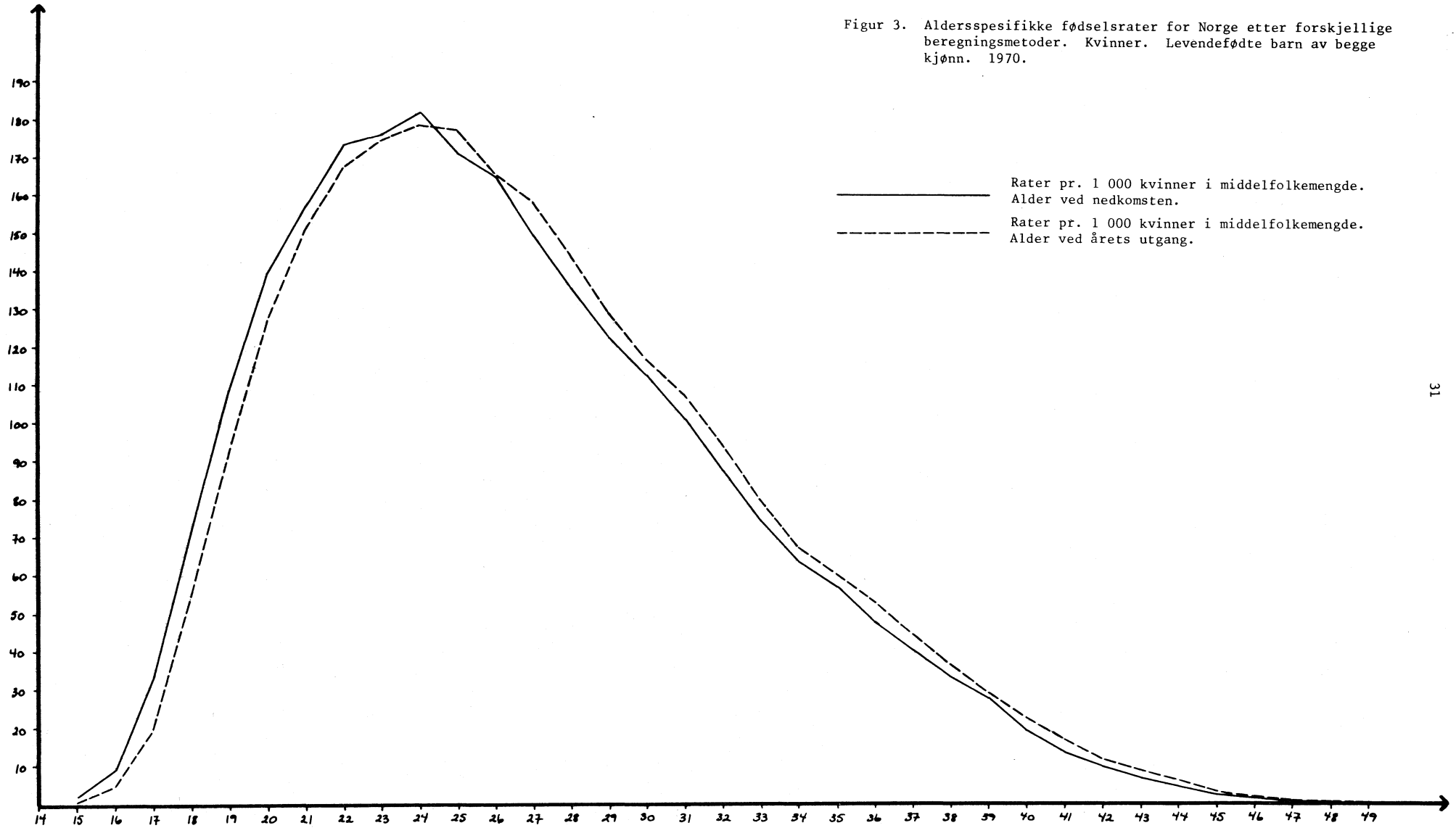
Figur 1. Aldersspesifikke fødselsrater for Norge etter forskjellige beregningsmetoder. Kvinner. Levendefødte barn av begge kjønn. 1968.



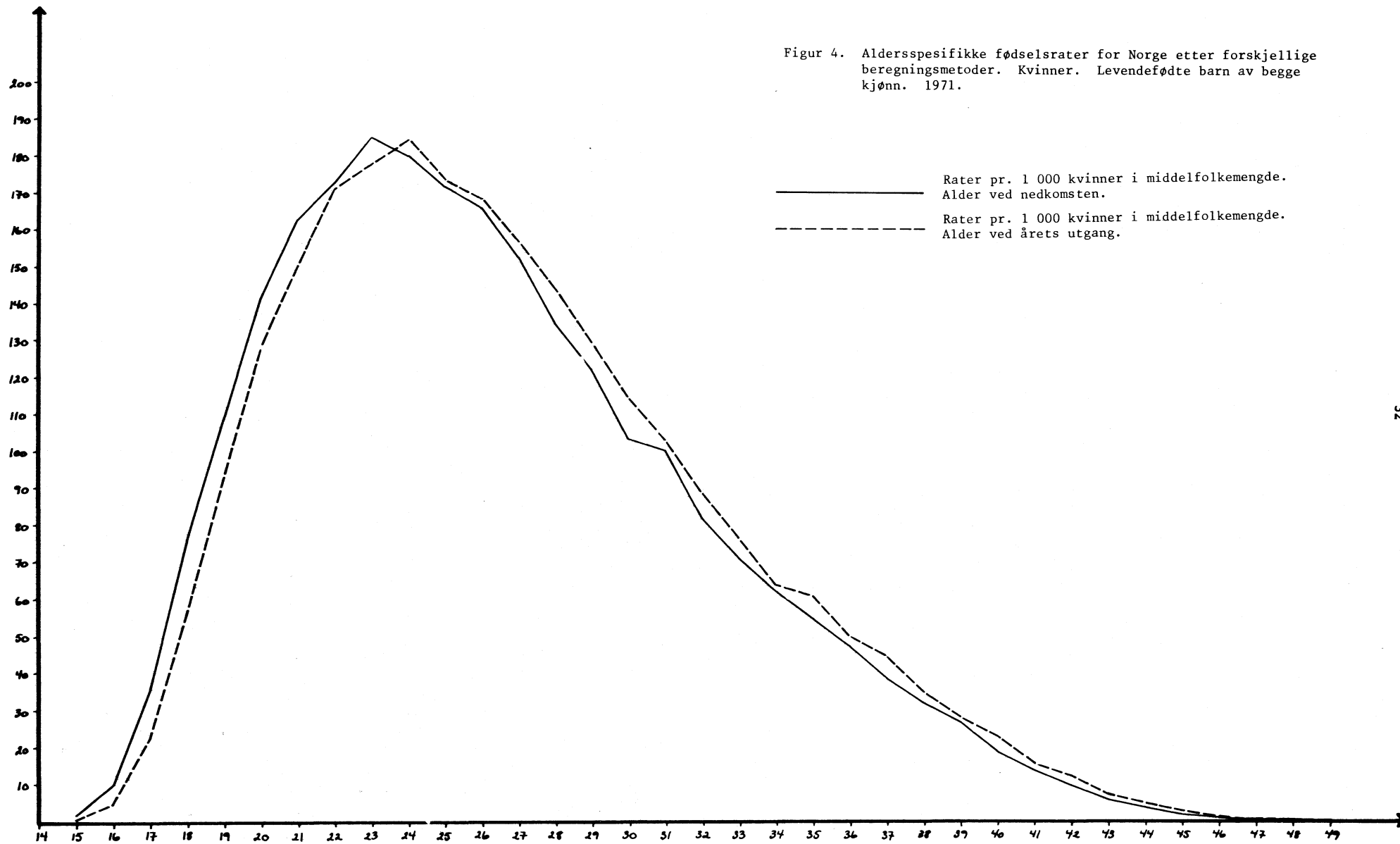
Figur 2. Aldersspesifikke fødselsrater for Norge etter forskjellige beregningsmetoder. Kvinner. Levendefødte barn av begge kjønn. 1969.



Figur 3. Aldersspesifikke fødselsrater for Norge etter forskjellige beregningsmetoder. Kvinner. Levendefødte barn av begge kjønn. 1970.

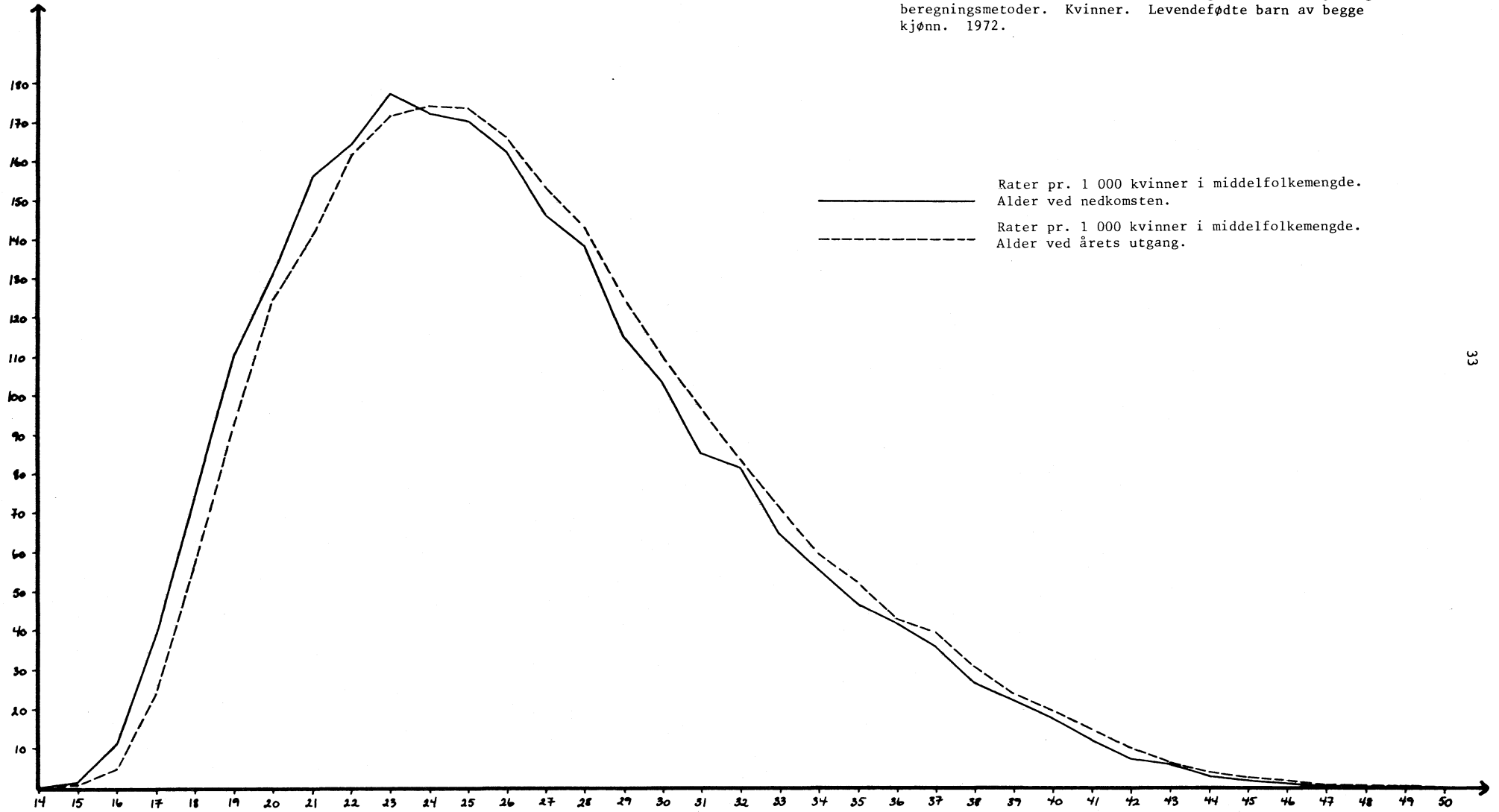


Figur 4. Aldersspesifikke fødselsrater for Norge etter forskjellige beregningsmetoder. Kvinner. Levendefødte barn av begge kjønn. 1971.





Figur 5. Aldersspesifikke fødselsrater for Norge etter forskjellige beregningsmetoder. Kvinner. Levendefødte barn av begge kjønn. 1972.



Analytisk glatting av fødselsrater for Norge for åra  
1968 til 1972

av

Erling Berge<sup>\*</sup>

	Side
1. Innleiing .....	35
2. Glattingane .....	36
3. Utviding av modellen .....	45
4. Sluttord .....	49
Litteraturliste .....	50

---

\* Takk til Jan M. Hoem for veiledning og samarbeid under prosjektet.

## 1. Innleiing.

1.1. Vi har glatta dei nasjonale fødselsrater for Norge for åra frå 1968 til 1972, analytisk av to grunner: Først og fremst har vi vore interessert i å studere utviklinga av parametrane i glattingsfunksjonen over tida, men vi var og interessert i den empiriske testen av modellane vi fikk på dette viset.

1.2. Modellane og glattingsmetoden er beskrevne i tidlegare notat (Hoem og Berge 1974a og b, Berge 1974). Vi har glatta med både Hadwigerfunksjonen og GG-funksjonen (ein glattingsfunksjon basert på gamma-tett-heta). Fødselsratene vart rekna ut etter mors alder i fylte år ved utgangen av året. I tillegg til ratene for dei einskilde åra 1968 til 1972 rekna vi og ut gjennomsnittlege rater for 1968-71 og for 1969-72. (Sjå også Hoem, Berge og Hansen, 1974).

1.3. I figurane 1 til 7 har vi framstilt dei glatta ratene saman med dei observerte. Avvika er store og systematiske. I eit tidlegare notat (Hoem og Berge, 1974b) har vi og rapportert relativt store kjikvadratverdier som eit tekn på systematisk avvik frå modellen. Vi kan reagere på dette på to måter. Vi kan anten forkaste modellen totalt, og gjere noko anna, eller vi kan freiste å utvide den med nye relasjonar. Vi skal i dette notatet peike på ei mogleg utviding av modellane.

1.4. I ein freistnad på å utvide modellen fører vi inn den aldersspesifikke kjønnsproporsjonen mellom  $x+i$  år gamle menn og  $x$  år gamle kvinner ( $i=1, 2, 3$ ). Dette grunnar seg på følgande resonnement.

Kvinner gifter seg jamnt over med menn eldre enn dei sjølve. Når vi har svingingar i størrelsen på fødselskulla, vil dette føre til svingingar i tilbodet av ektemenn. Dersom størrelsen av fødselskulla er veksande, vil det vere eit underskot på menn eldre enn kvinnene slik at fleire kvinner må vente lengre på å bli gift enn ein normalt ville vente med jamstore kohortar. Dette bør gi seg utslag i ein lågare fødselsrate enn normalt for desse årsklassane. Det omvendte gjeld sjølvsgt når størrelsen på fødselskulla er avtakande. Dette resonnementet bygg sjølvsgt på den premiss at familiemønsteret i hovudsak er monogamt.

1.5. Resultatet av denne freistnaden på å utvide modellen synest å vere relativt vellukka. For den glattinga vi omtaler her, er kjikvadratsummen omlag halvert. Men fortsatt er avvika for store til å kallast tilfeldige. Ei mogleg vidareutvikling kan vere å nytte ein lineærkombinasjon av alle dei tre kjønnsproporsjonane.

## 2. Glattingane

2.1. Vi glatta fødselsratene frå 1968 til 1972 og gjennomsnittsratene for 1968-71 og 1969-72 med både Hadwiger-funksjonen og GG-funksjonen. (På figurane har vi berre teikna inn Hadwigerglattingane.) Vi nytta heile tida kjikvadratminimeringsmetoden med estimerte grenser for den reproduktive perioden slik det er beskrive av Berge og Hoem (1974). For åra 1968 til 1971 var lågaste alder som kom med i glattinga 15 år, høgaste var 48 år. I 1972 var lågaste og høgaste aldrane 15 år og 47 år, mens vi for gjennomsnittstalla nytta ratene frå 15 år til 49 år.

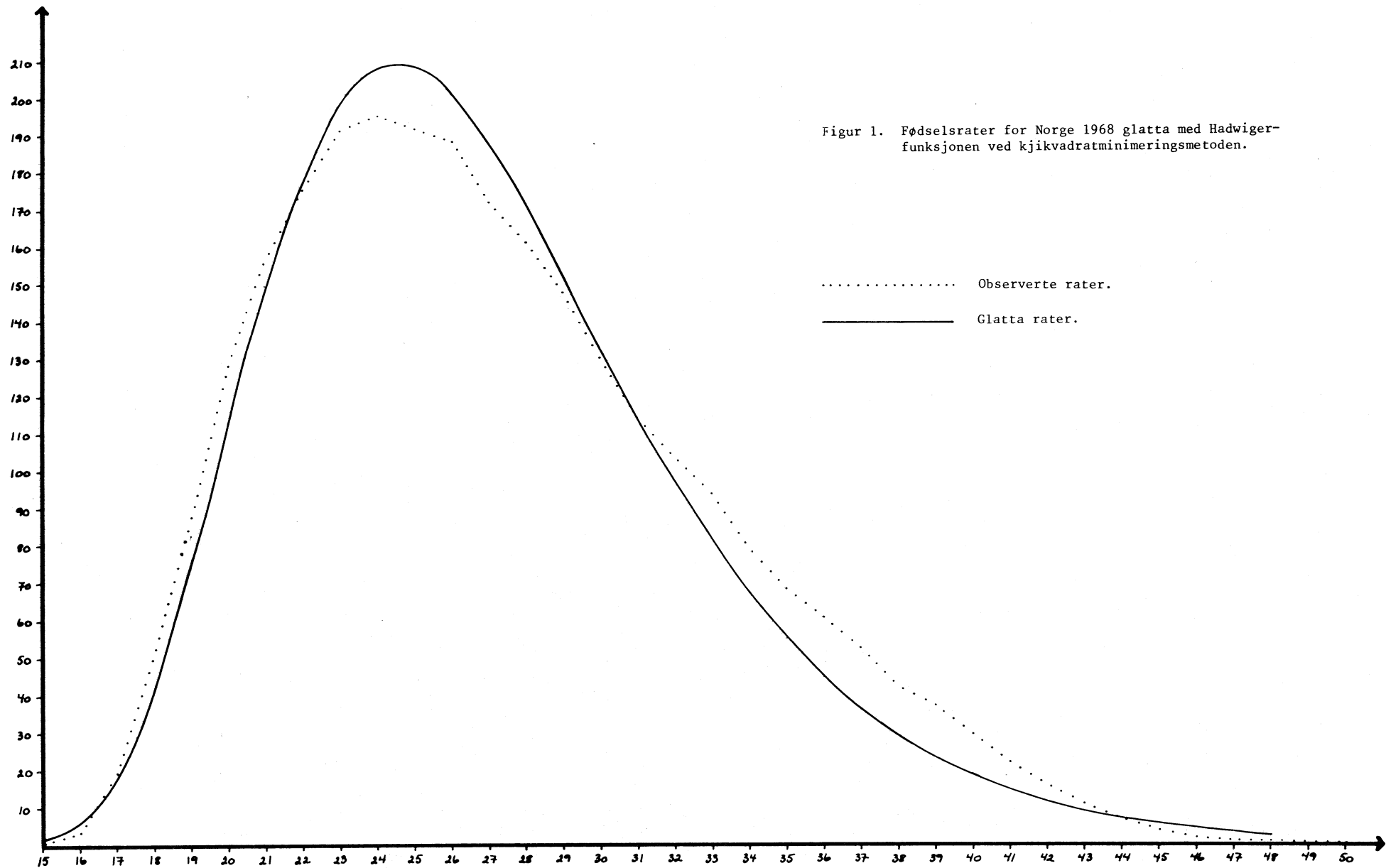
2.2. I tabell 1 finn vi estimat av fertilitetsparametrane for dei ulike åra. Estimatet av det samla fertilitetstallet ligg under det observerte, men viser sjølvsagt same utviklingstendens. Sjå tabell 3 hos Hoem, Berge og Hansen (1974). Både modal og gjennomsnittleg fødealder har i femårsperioden gått nedover med omlag eit halvt år. Samtidig finn vi skilnader i utviklinga. Frå 1971 til 1972 stoppa nedgangen i gjennomsnittleg fødealder, mens modal fødealder fortsatte å falle. Når det samla fertilitetstallet minkar og denne minken særleg skuldast nedgang i fødselsratene for kvinner eldre enn modal fødealder (sjå tabell 1 hos Hoem, Berge og Hansen (1974)), er det ein naturleg konsekvens at begge parametrane vil gå mot lågare verdi og også at avstanden mellom dei vil bli mindre. Dersom eit fortsatt fall i det samla fertilitetstallet fører til aukande avstand mellom modal og gjennomsnittleg fødealder, kan dette forklarast ved at reduksjonen i fødselsratene no ikkje lenger i hovudsak skjer mellom dei eldste kvinnene, men er flytta til dei aller yngste eller til kvinnene i aldrane omkring modal fødealder. Ser vi på tabell 1 hos Hoem, Berge og Hansen (1970), synest dette siste å vere tilfelle.

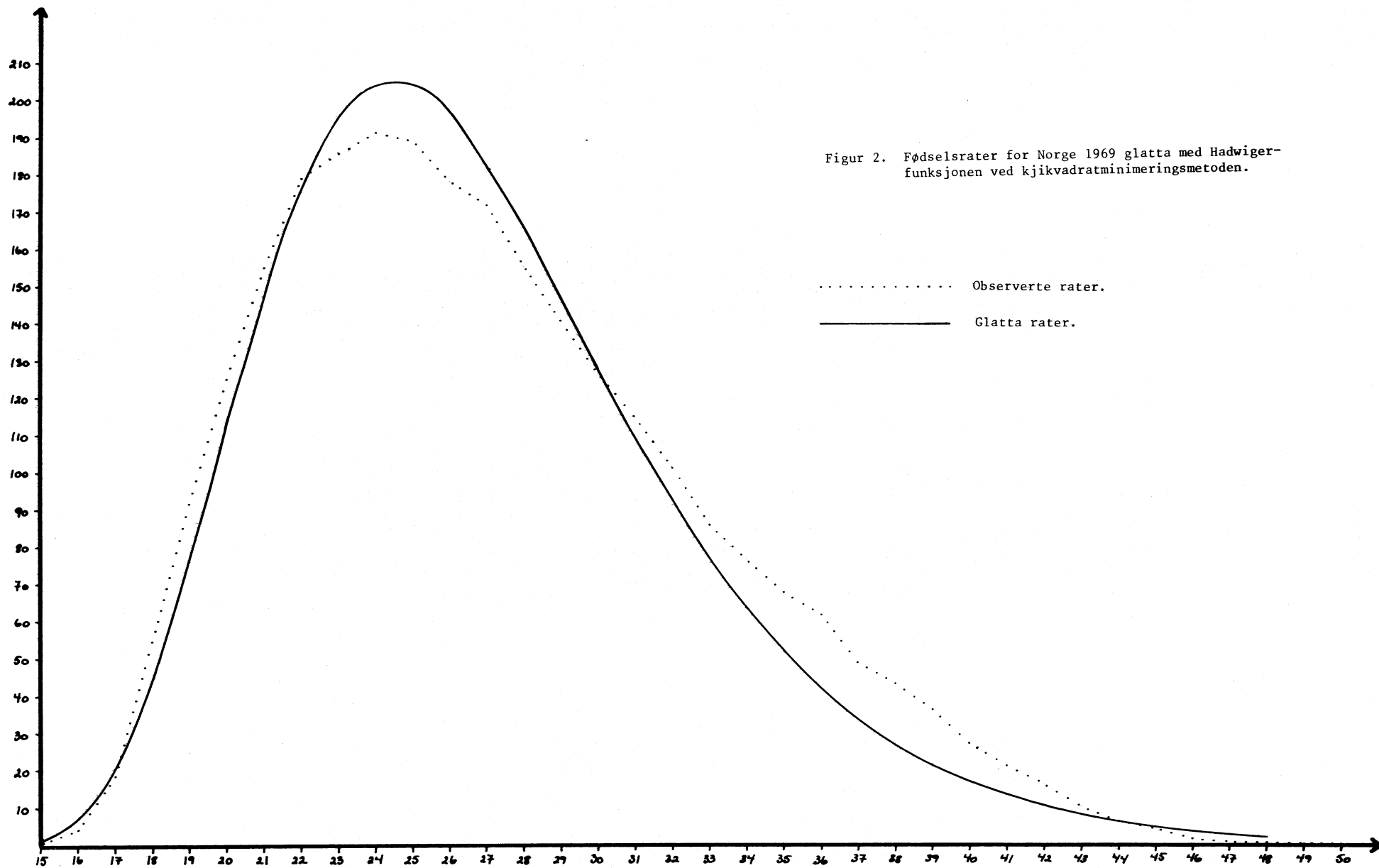
2.3. Samanliknar vi parametrane frå Hadwiger-glattingane med parametrane frå GG-glattingane, finn vi at GG-funksjonen estimerer det samla fertilitetstallet til å vere noko høgare enn Hadwiger-funksjonen. For modal og gjennomsnittleg fødealder er skilnadane nesten overraskande små når ein tenker på dei systematiske skilnadene mellom funksjonane som vi før har dokumentert (Berge og Hoem, 1974). Det er tydeleg at desse skilnadane først og fremst gir seg utslag i variansparameteren. Men også her er skilnadane relativt små.

Tabell 1. Estimerte<sup>\*</sup> fertilitets-parametrar for nasjonale data 1968-72

År	Samla fertilitets-tall	Modal føde-alder	Gjennomsnittleg føde-alder	Varians
HADWIGER				
1968 .....	2.682	24,54	27,04	32,58
1969 .....	2.603	24,50	26,90	31,62
1970 .....	2.417	24,30	26,70	31,14
1971 .....	2.393	24,15	26,52	30,15
1972 .....	2.327	24,07	26,53	31,10
1968-71 .....	2.491	24,50	26,66	29,07
1969-72 .....	2.406	24,37	26,54	28,90
GAMMA				
1968 .....	2.699	24,55	27,06	32,66
1969 .....	2.622	24,50	26,92	31,82
1970 .....	2.436	24,30	26,72	31,34
1971 .....	2.414	24,16	26,55	30,42
1972 .....	2.342	24,07	26,54	31,11
1968-71 .....	2.511	24,46	26,68	29,55
1969-72 .....	2.426	24,33	26,56	29,37

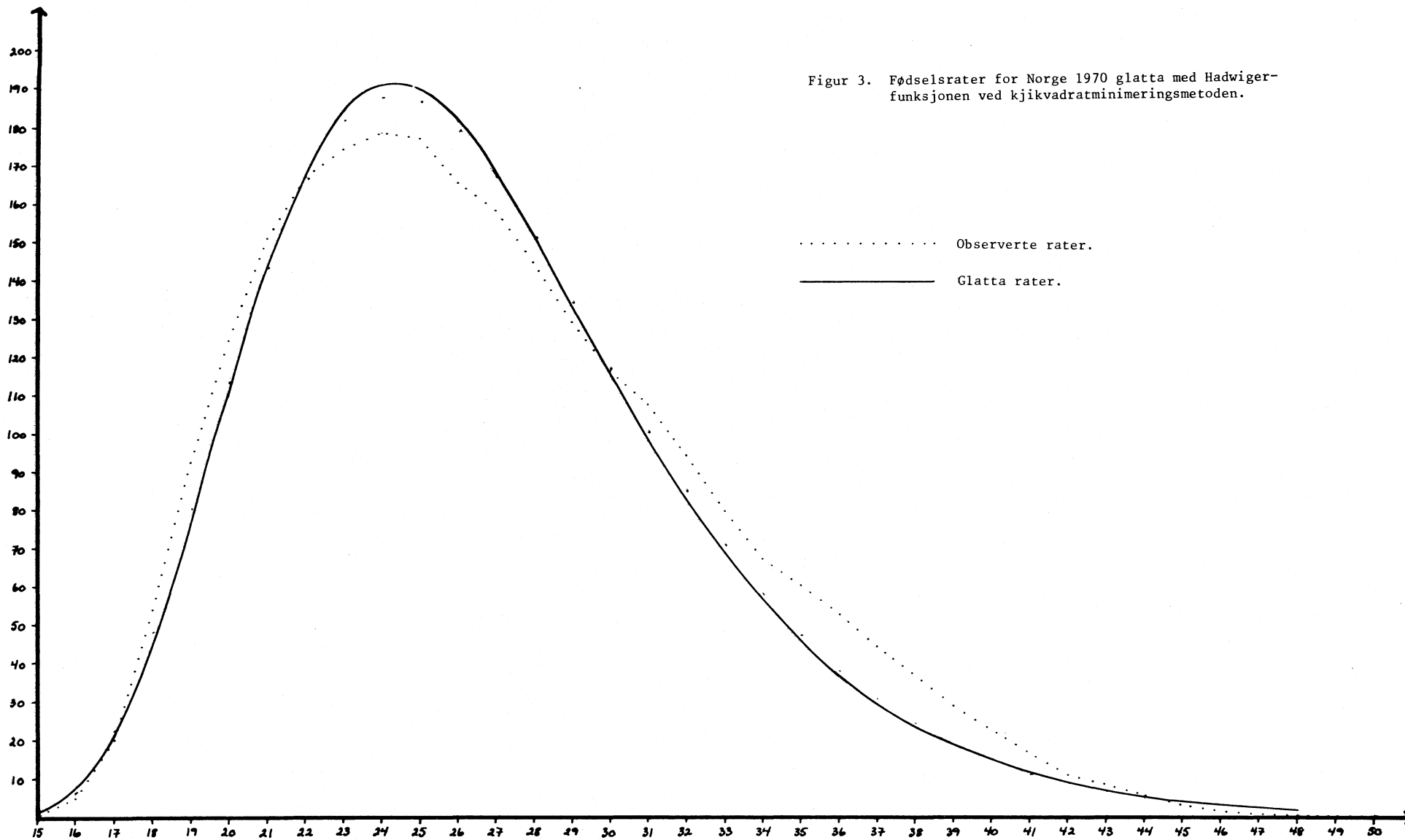
\* Det er nytta estimerte grenser for den reproduktive perioden.





Figur 2. Fødselsrater for Norge 1969 glatta med Hadwiger-funksjonen ved kji kvadratminimeringsmetoden.

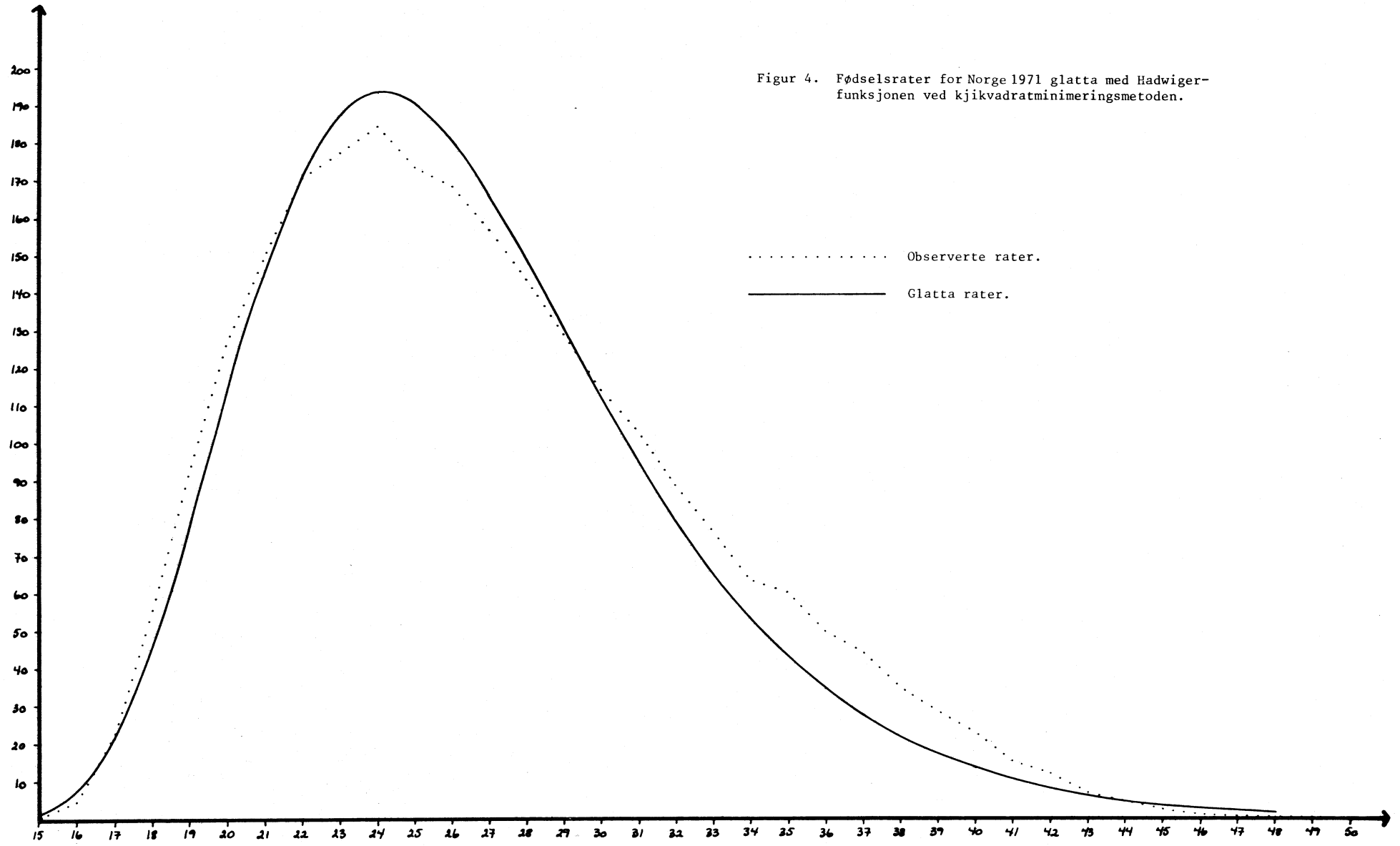
..... Observerte rater.  
 — Glatta rater.



Figur 3. Fødselsrater for Norge 1970 glatta med Hadwiger-funksjonen ved kjkvadratminimeringsmetoden.

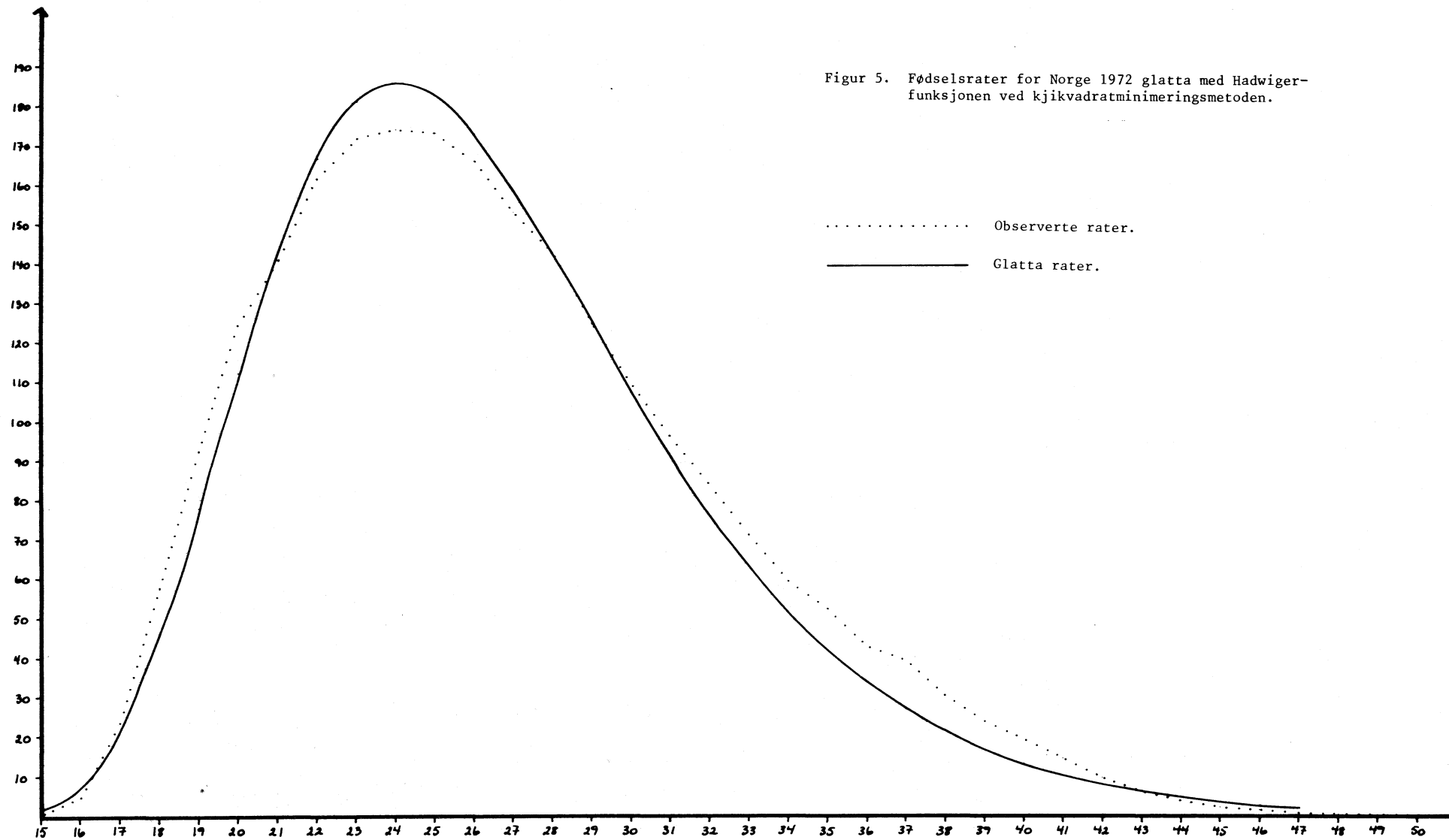
..... Observerte rater.  
 — Glatta rater.





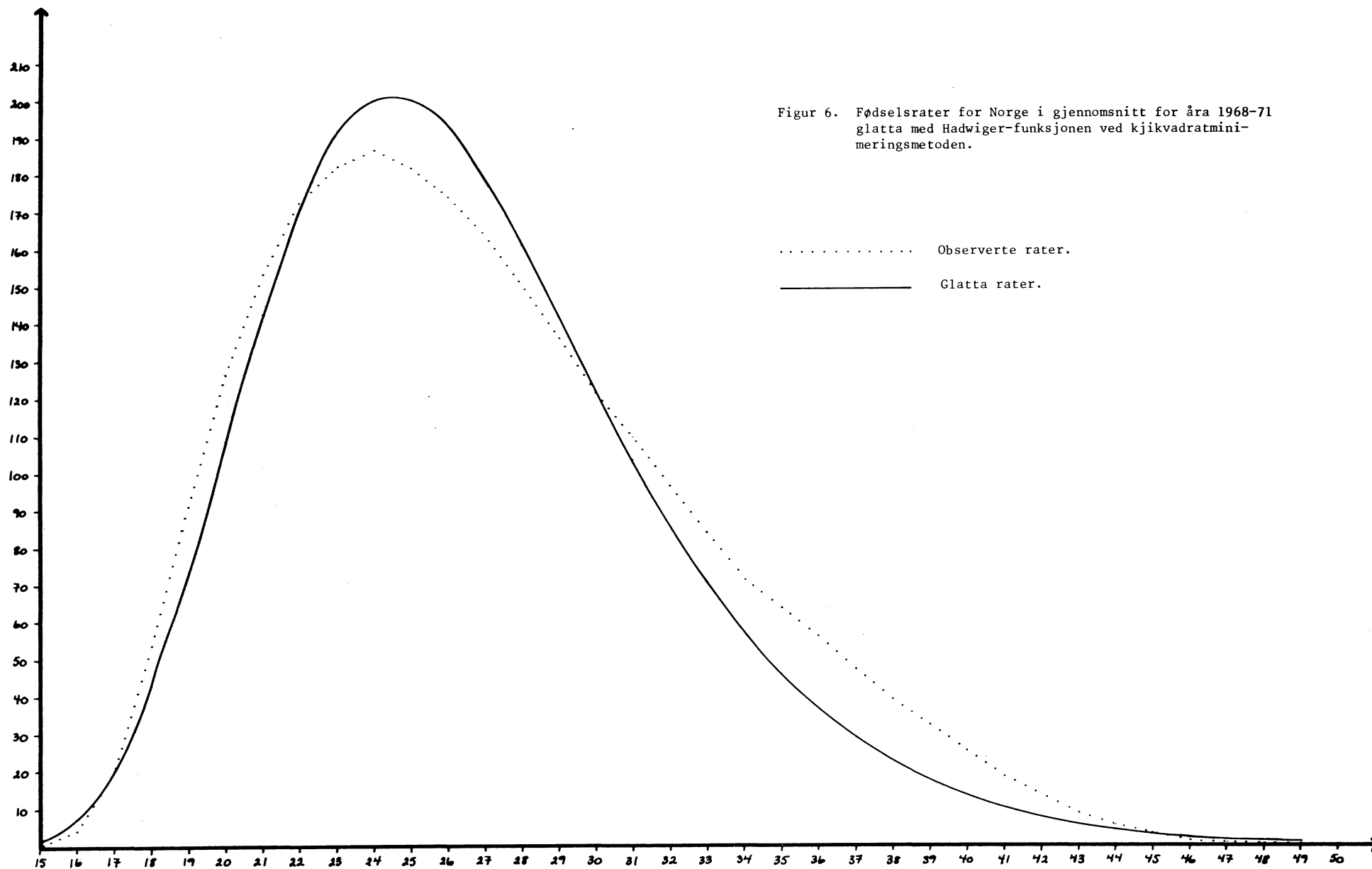
Figur 4. Fødselsrater for Norge 1971 glatta med Hadwiger-funksjonen ved kjkvadratminimeringsmetoden.

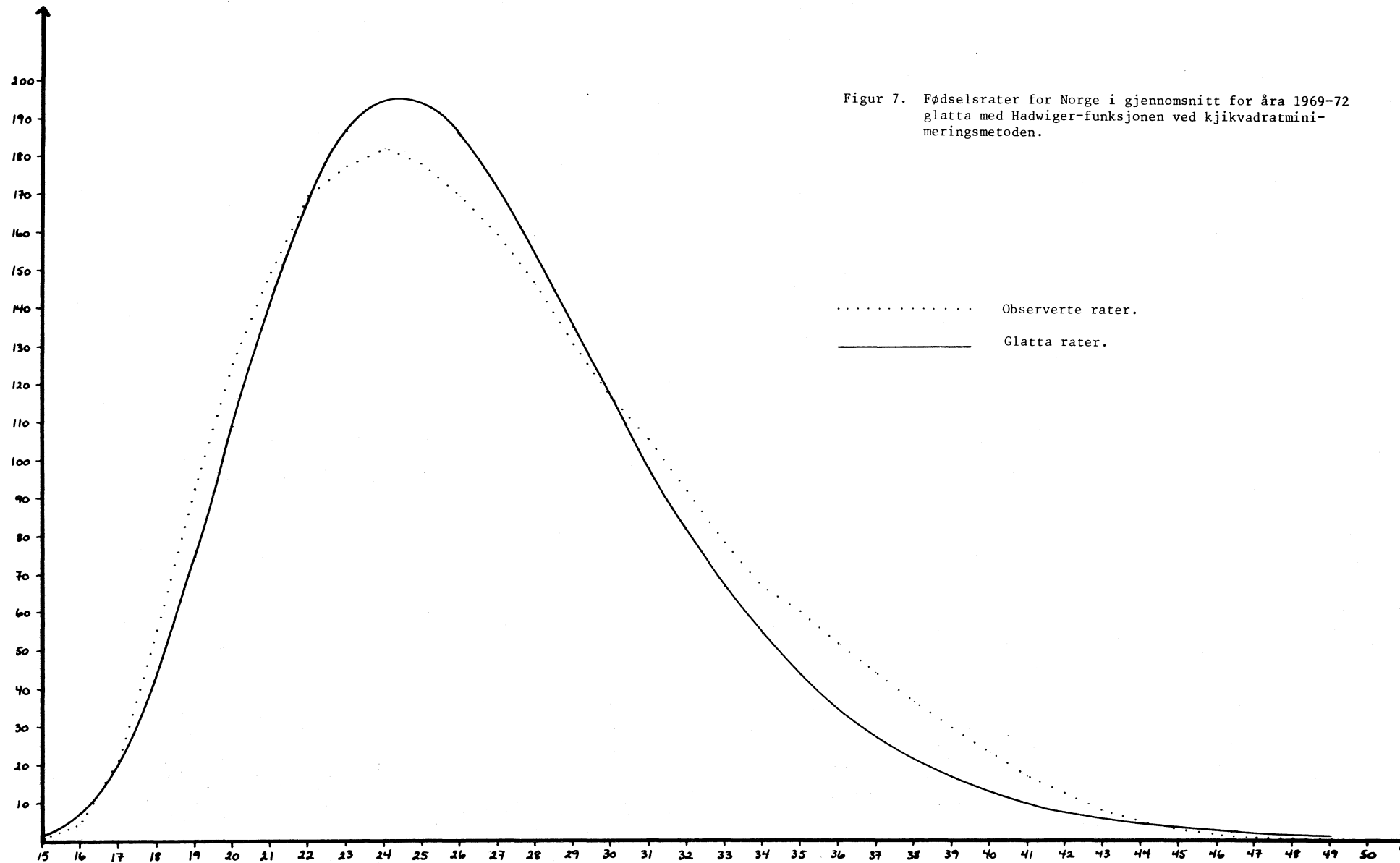
..... Observerte rater.  
 — Glatta rater.



Figur 5. Fødselsrater for Norge 1972 glatta med Hadwiger-funksjonen ved kjikvadratminimeringsmetoden.

..... Observerte rater.  
 — Glatta rater.





### 3. Utviding av modellen.

3.1. I tabell 2 finn vi tre ulike mål på tilpasninga av glattingane omtala i avsnitt 2.

Både for Hadwiger-glattingane og for GG-glattingane er kjikvadratverdien rett og slett enorm. Vi merkar oss likevel at GG-glattingane passar noko betre enn Hadwiger-glattingane.

3.2. I avsnitt 1.4 vart det framstilt eit resonnement som indikerte at kjønnsproporsjonen kan gi opphav til systematiske avvik fra ein modell som ville vere riktig i ei stasjonær befolkning. Vi rekna difor ut kjønnsproporsjonen

$$iK_x = L_{x+i}^m / L_x^k$$

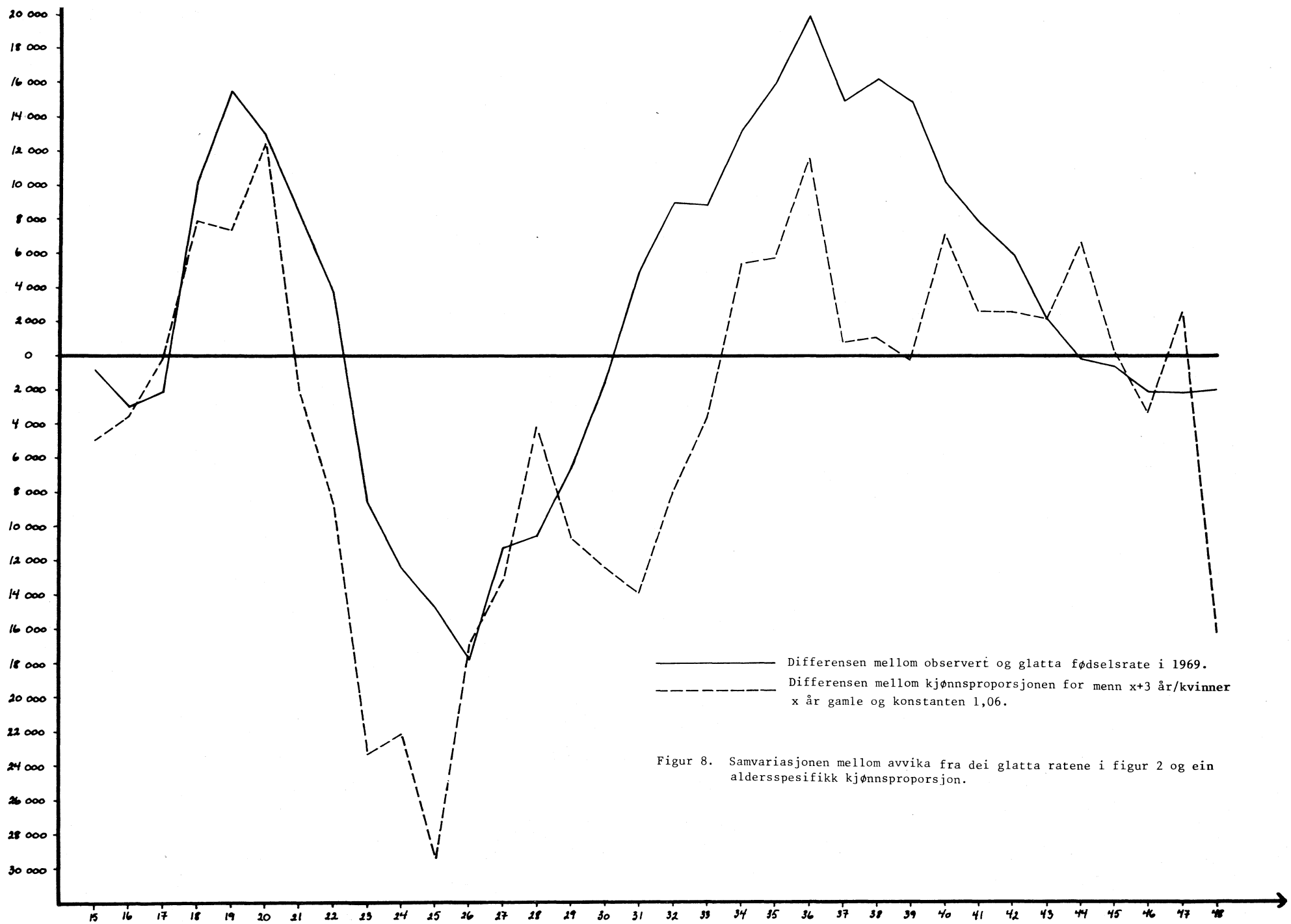
for  $i=1, 2$  og  $3$  og for  $x=15, 16, \dots, 50$ . Her er  $L_x^m$  antall  $x$ -årige menn og  $L_x^k$  er antall  $x$ -årige kvinner.

På figur 8 har vi for 1969 framstilt grafisk residualen fra glattinga, dvs.  $\hat{\lambda}_x - h(x, \hat{\theta})$ , saman med  $({}_3K_x - 1.06)$ . Samvariasjonen er tydeleg.

Tabell 2. Tilpasninga til nasjonale data for estimerte fødselsrater

År	$\chi^2(\hat{\alpha}-\hat{\beta})$	$SQ(15-50)^x$	$SQ(20-30)^x$
<b>HADWIGER</b>			
1968 .....	1631.5	3 185	1 282
1969 .....	1866.3	3 605	1 296
1970 .....	1578.8	3 037	1 123
1971 .....	1659.3	2 922	974
1972 .....	1073.2	1 647	595
1968-71 .....	9031.0	4 427	1 684
1968-72 .....	8024.8	3 677	1 301
<b>GAMMA</b>			
1968 .....	1242.1	2 186	821
1969 .....	1467.1	2 578	877
1970 .....	1186.3	2 086	746
1971 .....	1241.2	1 981	614
1972 .....	743.5	966	315
1968-71 .....	7184.3	3 240	1 147
1969-72 .....	6251.8	2 610	849

x I tusendeler.



Figur 8. Samvariasjonen mellom avvika fra dei glatta ratene i figur 2 og ein aldersspesifikk kjønnsproporsjon.

3.3. Vi utvidar modellen med kjønnsproporsjonen slik at vi får glattingsfunksjonen

$$f(x, \theta) = h(x, \theta_0) + A_3 K_x + B$$

til å glatte fødselsratene frå 1969 med. Her er  $h$  Hadwiger-funksjonen med den vanlege parametervektoren, som vi no kallar  $\theta_0$ . Parametervektoren  $\theta$  i den nye glattingsfunksjonen får vi ved å legge dei to nye parametrane  $A$  og  $B$  til  $\theta_0$ . Alder kallar vi  $x$  og kjønnsproporsjonen altså  $3K_x$ .

Estimatet av fødselsraten til ei  $x$ -årig kvinne skulle da ifølge den nye modellen bli

$$\hat{\lambda}_x = f(x, \hat{\theta}) + \varepsilon_x,$$

der  $\varepsilon_x$  er eit tilfeldig feilledd med forventning null.

Programmeringa for å glatte med funksjonen  $f$  er beskriven i Berge (1974), Appendiks D. Estimatet av parameteren  $\theta$  for 1969 finst i tabell 3.

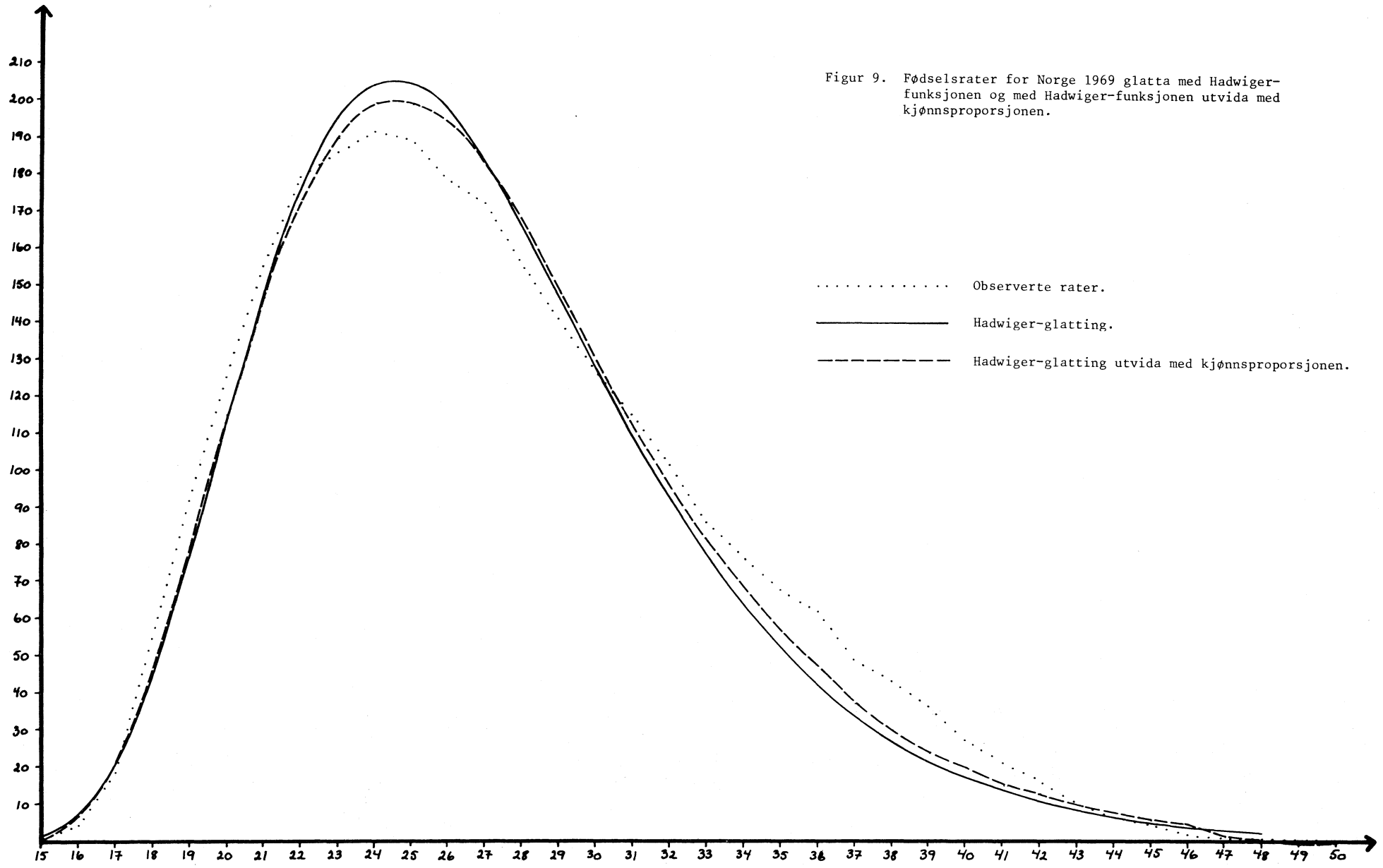
Tabell 3. Estimat av fertilitetsparametrar for 1969 i utvida modell.

Samla fertilitetstall .....	2.683
Modal fødealder .....	24.56
Gjennomsnittleg fødealder .....	27.10
Varians .....	34.00
A (i tusendeler) .....	14.71
B (i tusendeler) .....	-16.48

Vi merkar oss at det estimerte samla fertilitetstallet no ligg svært nært det observerte, som er på 2.696. I modal og gjennomsnittleg fødealder er endringane relativt små i høve til dei i tabell 1. Like interessant er det å sjå på kjikvadratverdien. Den er no redusert til 911.2, eller omlag halvparten av det opprinnelege. (Sjå tabell 2). Til samanlikning fann vi at GG-funksjonen reduserte kjikvadratverdien til 1467,1. (Sjå tabell 2). Ei utviding av Hadwiger-modellen gir altså tydeleg betre resultat enn eit skifte av modell frå Hadwiger- til GG-funksjonen.

Utvidinga av modellen burde gå like greitt for GG-funksjonen. Utforskinga av dette må imidlertid utsettast inntil vidare.

3.4. I figur 9 har vi teikna inn både den opprinnelege glatta kurva og kurva fra den utvida modellen. Som vi ser er forbetringa tydeleg, men ho er enno langt ifra å vere god nok.



Figur 9. Fødselsrater for Norge 1969 glatta med Hadwiger-funksjonen og med Hadwiger-funksjonen utvida med kjønnsproporsjonen.

..... Observerte rater.  
 ——— Hadwiger-glatting.  
 - - - - - Hadwiger-glatting utvida med kjønnsproporsjonen.



#### 4. Sluttord.

4.1. Verken Hadwiger-funksjonen eller GG-funksjonen synest å vere rette modellen for å glatte nasjonale fødselsrater med for Norge i dag. Sjølv med den mogelege utvidinga av modellen vi har peika på her, gir ikkje modellen ein tilstrekkeleg god representasjon av data. Det er tydeleg at ein både må finne fram til ein betre grunnmodell og studere vidare utbygging av den slik det har vore freista her.

Litteraturliste

Berge, Erling og Jan M. Hoem (1974): "Nokre praktiske erfaringar med analytisk glatting". Statistisk Sentralbyrå, ANO IO 74/23.

Berge, Erling (1974): "MINSYS; eit reknemaskinprogram for analytisk glatting av befolkningsrater". Statistisk Sentralbyrå, ANO IO 74/11.

Hoem, Jan M. og Erling Berge (1974a): "Some problems in Hadwiger fertility graduation". Statistisk Sentralbyrå, ANO IO 74/5.

Hoem, Jan M. og Erling Berge (1974b): "Theoretical and empirical results on the analytic graduation of fertility rates". I Hoem et.al.: "Two papers on analytic graduation". Statistisk Sentralbyrå, ANO IO 74/17.

Hoem, Jan M., Erling Berge og Liv Hansen (1974): "Norske fødselsrater for ettårige aldersklasser 1968-72". Side 19 til 33 i dette metodeheftet.