

Imputeret forbrug over livscyklussen

Stephanie Koefoed Rebbe

Danish Rational Economic Agents Model, DREAM

DREAM Arbejdsrapport 2014:1

Marts 2014

Abstract

Arbejdsrapporten beregner individers private forbrug ved anvendelse af registerdatainformation om deres disponible indkomst og nettoformue. Herefter beregnes det gennemsnitlige private forbrug for aldersgrupperne 16-årige til 90-årige. Det gennemsnitlige privatforbrug over alder er i rapporten beregnet for år 2000 til og med år 2009.

IMPUTERET FORBRUG OVER LIVSCYKLUSSEN

March 18, 2014

Stephanie Koefoed Rebbe
skr@dreammodel.dk
Tlf. nr.: 33 30 79 55

1 Indledning

En stor begrænsning ved at studere personers forbrugsadfærd over livet og over tid er manglen på data. Det skyldes, at det private forbrug ikke registreres, lige som for eksempel indkomst, formue, skatteindtægter og skatteudgifter registreres. Gennem stikprøver er det imidlertid muligt at finde privatforbruget for et repræsentativt udsnit af befolkningen fordelt på varegrupper. Forbrugsundersøgelsen, der hvert år undersøger forbruget for ca. 0,1 % af de i alt ca. 2,5 mio. danske husstande, er et eksempel på en sådan stikprøve. Denne tilgang til at finde forbruget kan være forbundet med meget usikkerhed.¹ Endvidere vil en opdeling på alder af de udvalgte husstandes hovedpersoner efterlade små stikprøver. En generations gennemsnitlige forbrug vil således være behæftet med stor usikkerhed, idet stikprøven vil være lille, ligesom forskelle i det gennemsnitlige forbrug på tværs af generationer vil være behæftet med meget usikkerhed.

Dette notat beskriver, hvordan det private forbrug over livscyklussen kan imputeres ved anvendelse af viden om individers formue og disponible indkomst. I notatet tages udgangspunkt i den regnskabsmæssige metode præsenteret i *Imputing Consumption from Income and Wealth Information* af Leth-Petersen og

¹Mere information om forbrugsundersøgelsen kan findes på <http://www.dst.dk/da/Statistik/emner/forbrug/forbrugsundersogelsen.aspx?tab=dok>

Browning (2003). Forbruget findes ved brug af registerdata fra Lovmodelservens FB33 datasæt, og årene 2000 til 2009 er undersøgt i dette notat.

For at få en ide om, hvorvidt den regnskabsmæssige metode præsenteret nedenfor tilnærmelsesvis rammer det eksakte individuelle forbrug, er det ønskeligt at sammenligne det imputerede individuelle forbrug med forbruget i forbrugsundersøgelsen. Det er dog på nuværende tidspunkt ikke muligt at sammenligne de to, idet individerne i forbrugsundersøgelsen ikke kan koples sammen med individerne i FB. I stedet sammenlignes det imputerede forbrug med makroforbruget i statistikbanken Tabel NATN01 i afsnit 4.

2 Metode²

Lad individ i påbegynde år t med nettoformuebeholdningen $A_{k,t-1}$, hvor $A_{k,t-1}$ er nettoformuebeholdningen af aktiv k i slutningen af år $t-1$. Nettoformuebeholdningen, $A_{k,t-1}$, giver kapitalindkomstafkastet, $i_{k,t}$, for komponent k i nettoformuebeholdningen. Individet opnår i løbet af år t indkomsten Y_t , betaler skatterne S_t og køber varebundet c_t . Lad den disponible indkomst være givet ved $Y_t^{Disp} = Y_t + \sum_k i_{k,t} - S_t$.

Individets budgetbetingelse er givet ved ligning (1):

$$c_t = Y_t^{Disp} + \sum_k (A_{k,t-1} - A_{k,t}) p_{k,t} \quad (1)$$

Hvis Y^{Disp} , p og A er kendte, er det umiddelbart muligt at lave et skøn over det imputerede privatforbrug på baggrund af i ligning (1). Beholdningen af og priserne på nettoformuen er imidlertid ukendte, og det er kun værdien af nettoformuen, der er kendt, altså $W_{k,s} = p_{k,s} A_{k,s}$ for $s = t-1, t$. For at tage højde for dette omskrives ligning (1) som følger:

$$\begin{aligned} c_t &= Y_t^{Disp} - \sum_k p_{k,t} A_{k,t} + \sum_k p_{k,t} A_{k,t-1} + \sum_k p_{k,t-1} A_{k,t-1} - \sum_k p_{k,t-1} A_{k,t-1} \Leftrightarrow \\ c_t &= Y_t^{Disp} - \Delta W_t + \sum_k (p_{k,t} - p_{k,t-1}) A_{k,t-1} \end{aligned} \quad (2)$$

$$\Delta W_t = \sum_k p_{k,t} A_{k,t} - \sum_k p_{k,t-1} A_{k,t-1}.$$

²Hele afsnittet tager udgangspunkt i artiklen Imputing Consumption from Income and Wealth Information (2003).

$\sum_k (p_{kt} - p_{k,t-1})A_{k,t-1}$ i ligning (2) angiver kapitalværdiændringen i individets nettoformue, og ΔW_t angiver ændringen i værdien af individets nettoformue fra år $t - 1$ til t . Nettoformuen består af aktier, obligationer, pantebreve, indestående i pengeinstitutter, ejendomsaktiver, gæld til pengeinstitutter, pantebrevsgæld og gæld til realkreditinstitutter.

De komponenter i nettoformuen, der kan være udsat for kapitalværdiændringer er: aktier, obligationer og ejendomsaktiver. Ændringer i pantebreve/ pantebrevsgæld, indestående/ gæld i pengeinstitutter og gæld til realkreditinstitutter skal tilskrives en ændring i individers beholdning af disse og ikke kapitalværdiændringer. Det betyder, at $\sum_k (p_{kt} - p_{k,t-1})A_{k,t-1} = 0$ for pantebreve, pantebrevsgæld, indestående i pengeinstitutter, gæld til pengeinstitutter og gæld til realkreditinstitutter.

I data er kapitalværdiændringen på aktier og obligationer kendt, mens kapitalværdiændringen på ejendomsaktiver er ukendt. På Lovmodelserveren er kapitalværdiændringen på aktier og obligationer indeholdt i den disponible indkomst. Derfor defineres $Y_t^{Disp'}$, hvor det gælder, at $Y_t^{Disp'} = Y_t^{Disp} + \sum_j (p_{jt} - p_{j,t-1})A_{j,t-1}$ for $j = \text{aktier} \ \& \ \text{obligationer}$, således at kapitalværdiændringen er indeholdt i $Y_t^{Disp'}$. Ligning (2) kan derfor omskrives til:

$$c_t = Y_t^{Disp'} - \Delta W_t + (p_t^H - p_{t-1}^H)A_{t-1}^H \quad (3)$$

hvor p^H angiver prisen på ejendomsaktiver, og A^H angiver individets beholdning af ejendomsaktiver.

Der er tre faktorer, der kan påvirke individets ejendomsaktiver. Det er udvidelse af den eksisterende boligbeholdning, fra/tilflytning af ejerbolig og kapitalværdiændringer. Kapitalværdiændringer i forbindelse med boligaktiver er ikke indeholdt i den disponible indkomst i modsætning til tilfældet for aktier og obligationer. Det ville være ideelt at inkludere disse i indkomstbegrebet, men det er ikke muligt at lave et fornuftigt skøn over kapitalværdiændringer på bolig på baggrund af nuværende dataadgang og viden (det vil dog være en klar forbedring at lave dette skøn). Derfor anvendes en alternativ tilgang.

De personer, der er flyttet fra en ejerbolig eller til en ejerbolig fra et givent år $t - 1$ til år t , ekskluderes fra datasættet. Når disse personer slettes, vil de fluktuationer, der er tilbage i ejendomsaktiverne, være skabt af husprisvariationer, dvs. kapitalværdiændringer, eller ved udvidelser i den eksisterende boligbeholdning. De personer, der har fortaget udvidelser i deres eksisterende boligbeholdning fra et givent år $t - 1$ til år t , ekskluderes ligeledes fra datasættet.³ Således er det nu kun

³ Disse personer identificeres ved at undersøge, hvorvidt antallet af kvadratmeter, personen bor

kapitalværdiændringer på bolig, der kan skabe udsving i ejendomsaktiverne. Det ønskes ikke at kapitalværdiændringer i ejendomsaktiverne mekanisk skal påvirke opsparingen og dermed det samlede forbrug, hvorfor ejendomsaktiver udelades fra nettoformueudtrykket. Det er muligt at fjerne ejendomsaktiverne fra formueudtrykket, idet vi ikke er interesserede i niveauet for nettoformuen, men blot ændringen i nettoformuen fra et år til et andet.

Når personer, der flytter til/fra en ejerbolig eller foretager udvidelser af den eksisterende boligbeholdning, slettes, bliver det oplagte skøn over forbruget som i ligning (4):

$$c_t = Y_t^{Disp'} - \Delta W_t \quad (4)$$

hvor $W_t = \sum_m p_{m,s} A_{m,s}$ for $s = t - 1, t$
 og $m =$ aktier, obligationer, pantebreve,
 indestående i pengeinstitutter, gæld til pengeinstitutter,
 pantebrevsgæld og gæld til realkreditinstitutter.

3 Data

På Lovmodelserveren findes oplysninger om formue, indkomster og skattebetalinger for den danske befolkning fra år 1993 til 2009. I dette notat er årene 2000 til 2009 undersøgt. Da der 'kun' eksisterer data tilbage til 1993, er det ikke muligt, at følge et specifikt individs disponible indkomst og formue over hele dets livscyklus. I stedet findes individernes gennemsnitlige forbrug over livet ved at foretage en tværnitanalyse. Data i notatet er baseret på en repræsentativ stikprøve af den danske befolkning på 33 procent, altså datasættet FB33.

Der tages ikke højde for størrelsen af den husstand, et individ lever i. Det betyder, at et individs forbrug kan indeholde forbrug til eventuelle børn i husstanden. Individder, der dør, udvandrer, indvandrer og nyfødte, fjernes, og således følger tværnitanalysen den samme stikprøve af befolkningen over tid. På den baggrund er det muligt at finde ændringen i individers formue fra et tidspunkt til et andet.

De individer, der er flyttet fra en ejerbolig eller til en ejerbolig fra et givent år $t - 1$ til år t , er udeladt fra stikprøven, ligesom individer, der har foretaget boligudvidelser er trukket ud af stikprøven. Endvidere er de individer for hvilke, der er oplyst en negativ disponibel indkomst fjernet.

I Tabel 1 ses det undersøgte antal individer fra år 2000 til 2009.

på, har ændret sig fra år $t-1$ til år t .

TABEL 1 ANTAL INDIVIDER I STIKPRØVEN, år 2000 til 2009

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Antal personer i FB33	1 708 849	1 711 396	1 715 966	1 721 109	1 726 267	1 722 921	1 724 057	1 732 295	1 738 447	1 746 624
Antal flyttede personer og personer, der har foretaget boligudvidelse	184 536	183 625	182 839	203 414	208 346	121 160	125 894	140 385	100 906	90 323
Antal personer tilbage i stikprøven	1 524 313	1 527 771	1 533 127	1 517 695	1 517 921	1 601 761	1 598 163	1 591 910	1 637 541	1 656 301

Kilde: Egne beregninger på registerdata.

Anm.: Rækkerne angiver antal personer i stikprøven, hvor døde, indvandrede, udvandrede og nyfødte er fratrukket. Endvidere er personer med negativ disponibel indkomst ekskluderet.

Ved imputeringen af det private forbrug anvendes som nævnt viden om individernes disponible indkomst og nettoformue. Den disponible indkomst og nettoformuebegrebet er udledt som følger:

DISPONIBEL INDKOMST, DISP5N

Den disponible indkomst er udledt som summen af personlig indkomst⁴, overførselsindkomst⁵ og kapitalindkomst fratrukket skattebetalinger og fradrag til pensionsindbetalinger⁶. Variablen DISP5N i grunddata er benyttet. Indeholdt i DISP5N er variablen INKS, der anvendes ved dannelsen af PI5PO. Kapitalværdiændringer på obligationer og aktier ligger i INKS.

⁴Personlig indkomst består af: Lønindkomst mv., virksomhedsindkomst, indkomst fra private pensioner og anden privat indkomst.

⁵Overførselsindkomst består af: Pensionsydelser, ydelser fra A-kasse, orlovsordninger, ydelser under tidlig tilbagetrækning, kontanthjælp, statens uddannelsesstøtte mv., boligstøtte, børnefamilieydelse og børnetilskud.

⁶Fradrag ved pensionsindbetalinger består af: Fradrag for præmier vedr. forsikringer med fuld fradragsret og begrænset fradragsret ved privattegnede kapitalpensioner.

NETTOFORMUE, AKT_FIN - PASSIVER

Nettoformuen er sammensat af samlet finansiel formue i indenlandske og udenlandske aktiver⁷ fratrukket de samlede passiver⁸. Det vil sige, at nettoformuen i notatet er givet ved AKT_FIN (fra grunddata) fratrukket PASSIVER (fra grunddata).

3.1 Den disponible indkomst

Tabel 2 viser den minimale, maksimale og gennemsnitlige disponible indkomst, samt medianen til den disponible indkomst fra år 2000 til 2009 i løbende priser, hvor henholdsvis alle personer i FB33 er medtaget og de personer, der er flyttet fra/til en ejerbolig eller har foretaget boligudvidelser, er slettet.

TABEL2: DISPONIBEL INDKOMST I LØBENDE PRISER, ÅR 2000 TIL 2009

		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Alle personer i FB33	Minimum	-	-	-	-	-	-	-	-	-	-
	Maksimum	31 178 313	29 034 290	25 038 393	27 244 112	27 294 094	39 823 759	29 720 388	35 754 944	31 550 910	29 428 676
	Median	112 689	117 071	120 558	124 239	128 801	133 498	138 635	145 840	152 181	151 771
	Gennemsnit	113 626	117 393	120 740	123 856	129 927	135 659	141 435	150 358	155 340	154 340
Flyttede personer og personer, der har foretaget boligudvidelse, er slettet	Minimum	-	-	-	-	-	-	-	-	-	-
	Maksimum	26 812 540	29 034 290	23 905 845	27 244 112	27 294 094	22 591 818	28 431 617	26 287 323	31 550 910	23 962 689
	Median	111 995	116 474	119 927	123 354	128 109	130 324	136 521	142 485	150 007	150 400
	Gennemsnit	112 543	116 474	119 768	122 652	128 759	132 370	139 142	147 075	153 322	152 936

Kilde: Egne beregninger på registerdata

Stort set alle år ligger medianen til den disponible indkomst under den gennemsnitlige disponible indkomst, hvilket indikerer, at indkomstfordelingen er højreskæv. Der er således få personer, der trækker den gennemsnitlige disponible indkomst op. Dette er tilfældet for begge stikprøver.

Tabel 2 viser endvidere, at både medianen til den disponible indkomst og den gennemsnitlige disponible indkomst falder, når de personer, der flytter og de personer, der foretager boligudvidelser, slettes. Der er således tendens til, at mere

⁷Finansielle aktiver består af: Obligationer, aktier, indestående i pengeinstitutter og pantebreve.

⁸Passiverne består af: Gæld til pengeinstitutter, pantebrevsgæld, realkreditlån og samlede udenlandske passiver.

velhavende individer fjernes fra stikprøven. Dette er uhensigtsmæssigt, idet det kan betyde, at det imputerede gennemsnitsforbrug undervurderes. Det ses dog af afsnit 4, at der ikke er en klar indikation af, hvorvidt det imputerede makroforbrug undervurderes eller overvurderes.

4 Imputeret privatforbrug

Som forklaret i afsnit 2 udregnes forbruget i år t på baggrund af viden om den disponible indkomst i året fratrukket ændringen i nettoformuen fra år $t - 1$ til t . Et konsistentstjek af det imputerede privatforbrug kunne i særdeleshed være en sammenligning med det private forbrug i forbrugsundersøgelsen. Som nævnt indledningsvist er dette desværre ikke muligt. Derfor sammenlignes det imputerede makroforbrug i stedet med makroforbruget i henhold til Tabel NATN01 fra Danmarks Statistik i Tabel 3. Det imputerede makroforbrug er fundet som følger:

$$C_t = \sum_a \bar{c}_{a,t} * N_{a,t}$$

hvor C_t angiver det imputerede makroforbrug, $\bar{c}_{a,t}$ angiver det imputerede gennemsnitlige privatforbrug for aldersgruppe a , og $N_{a,t}$ angiver antallet af personer i aldersgruppen.

Tabel 3 illustrerer, at det imputerede makroforbrug er tilnærmelsesvist tæt på makroforbruget ifølge DST for årene 2000, 2001, 2003, 2004, 2005, 2006, og 2009.

TABEL3: IMPUTERET MAKROFORBRUG OG MAKROFORBRUG IFØLGE NATN01, LØBENDE PRISER, ÅR 2000 TIL 2009

Privat forbrug	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
NATN01, mia. kr.	616.7	631.7	652.3	666.9	707.2	745.1	786.6	820.4	840.0	822.0
Imputeret, mia. kr.	604.8	638.4	691.4	652.5	697.6	726.4	777.6	856.4	1 003.7	811.4

Kilde: Egne beregninger på registerdata. Samt Tabel NATN01 og FOLK2 fra DST. $N_{a,t}$ stammer fra FOLK2.

4.1 Gennemsnitligt imputeret privatforbrug og gennemsnitlig disponibel indkomst over alder

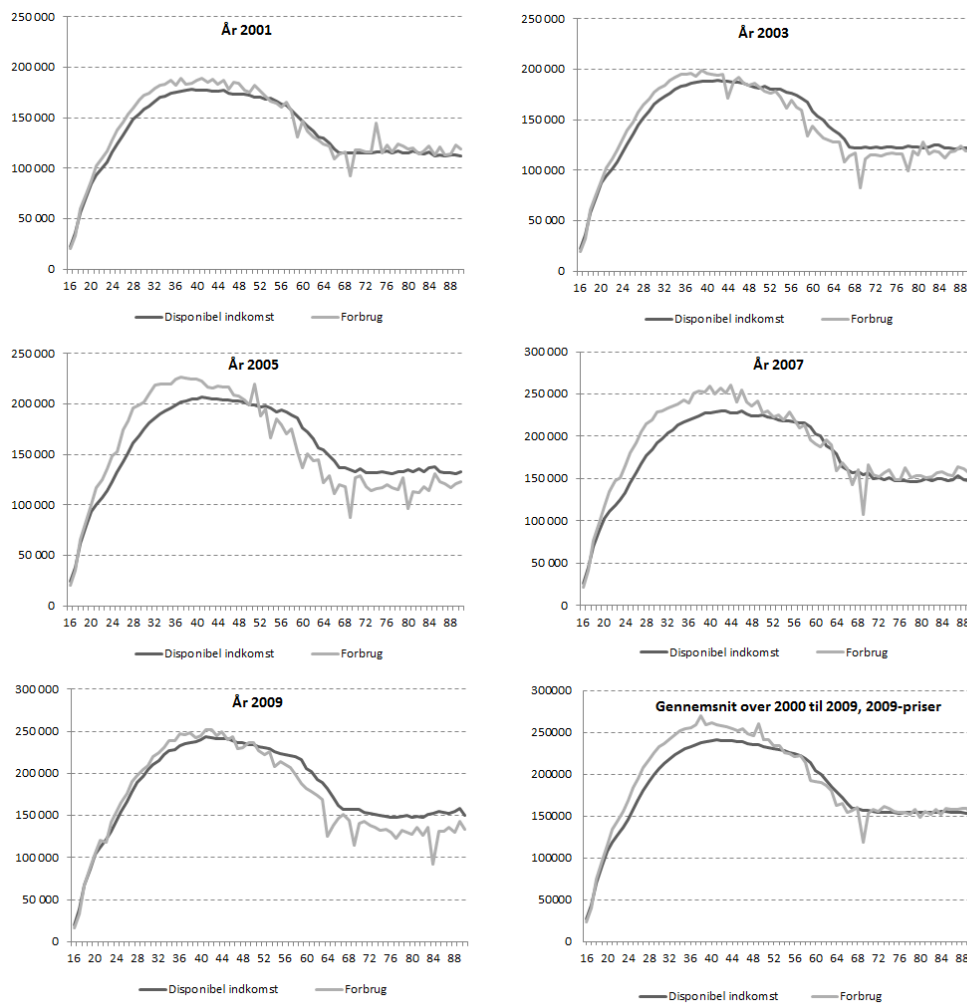
Figur 1 på næste side viser den gennemsnitlige disponible indkomst og det gennemsnitlige privatforbrug for 16 til 90-årige for år 2001, 2003, 2005, 2007 og 2009

i løbende priser. Endvidere ses det gennemsnitlige forbrug over årene 2000 til 2009 i 2009 priser. Figurene viser, at det gennemsnitlige privatforbrug for 16-årige til omtrent 20-årige er lig den gennemsnitlige disponible indkomst for kohorterne. Herefter er gennemsnits privatforbruget alle årene højere end den gennemsnitlige disponible indkomst indtil omtrent 50-årsalderen, hvilket indikerer en stor grad af utålmodighed hos de yngre generationer. I pensionsalderen er det gennemsnitlige imputerede privatforbrug tilnærmelsesvist konstant.

4.2 Manglende håndtering af personer, der køber varige forbrugsgoder

Ved imputeringen af det gennemsnitlige private forbrug er der ikke taget højde for køb af varige forbrugsgoder som bil, sommerhus, andelsbolig mv. Køb af disse forbrugsgoder kommer til at have stor betydning for det imputerede privatforbrug det år, købet finder sted, mens forbruget vil være upåvirket året efter købet. Køb af eksempelvis andelsbolig kommer på den måde til kun at påvirke forbruget ét år, nemlig det år andelsboligen er købt. Det ville være mere hensigtsmæssigt, hvis andelsboligkøbet blev spredt ud over en række år. Det er imidlertid ikke muligt, at identificere de personer, der har købt bil, sommerhus, andelsbolig mv. et givent år, ligesom prisen på disse forbrugsgoder er ukendte.

FIGUR1: IMPUTERET GENNEMSNITLIGT FORBRUG OG GENNEMSNITLIG DISPONIBEL INDKOMST OVER ALDER, LØBENDE PRISER, ÅR 2001, 2003, 2005, 2007 OG 2009, SAMT GENNEMSNIT OVER ÅRERNE 2000-2009 I 2009-PRISER



Kilde: *Egne beregninger på registerdata*