

13. maj 2005

Modellering af benzin- og bilforbruget med bilstocken bestemt på baggrund af samlet forbrug

Resumé:

Dette tredje papir om en ny model for biler og benzin vil først prøve at samle trådene op og forklare, hvordan implementeringen i ADAM kunne tænkes at foregå, samt hvilke forbedringer der kunne tænkes at komme før den endelige version. En essentiel del af den nye struktur er, at det samlede bil- og benzinforbrug kan bestemmes som en andel af det samlede private forbrug eksklusivt boliger. I dette papir vises, at bilstocken faktisk kan forudsiges bedre på baggrund af det samlede private forbrug eksklusivt boliger, end den indkomstvariabel der tidligere har været brugt. Hermed er vejen banet for den omstrukturering, der er blevet lagt op til.

GRH130505

Nøgleord: Biler, benzin, bilstocken, fejlkorrektionsmodel, transportforbrug

Modelgruppepapirer er interne arbejdsrapporter. De konklusioner, der drages i papirerne, er ikke endelige og kan vFre Fndret inden opstillingen af nye modelversioner. Det henstilles derfor, at der kun citeres fra modelgruppepapirerne efter aftale med Danmarks Statistik.

1. Indledning

Dette papir er det tredje papir i en række vedrørende reformuleringen af bil- og benzinforbruget i ADAM. I det første GRH27105 undersøges om en fejlkorrektionsmodel kan forklare ændringen i det relative benzin- og bilforbrug ud fra de relative priser. Konklusionen var, at det ikke umiddelbart kunne lade sig gøre. På daværende tidspunkt så det ud til, at en alvorlig brist var modelformuleringen. Forbrugerne må antages at efterspørge biler og kørte kilometer og ikke biler og benzin. En årsag til problemerne kunne således være en ikke konstant benzineffektivitet. Andet papir GRH21205 viser, at ganske vidst har benzineffektiviteten ikke været konstant, og data kan bedre lide den alternative formulering, men der er stadig knas med residualerne. Problemet ser ud til at skyldes, at indkomstelasticiteten for biler og benzin ikke er ens.

Dette papir vil prøve at samle trådene. Afsnit 2 vil forklare den nye overordnede struktur for bil- og benzinrelationerne, mens afsnit 3 vil forklare, hvorledes problemerne med hensyn til det relative benzin- og bilforhold kan tænkes at blive løst. Det praktiske empiriske bidrag fra dette papir ligger i afsnit 4, hvor det forsøges at estimere bilstocken og hermed bilforbruget med $Cp4xhv1$, privat forbrug eksklusivt boliger, i stedet for Ycb , en sammenvejning af formue og forbrug. Endelig bliver der konkluderet i afsnit 5.

2. Struktureringen af bil- og benzinforbruget

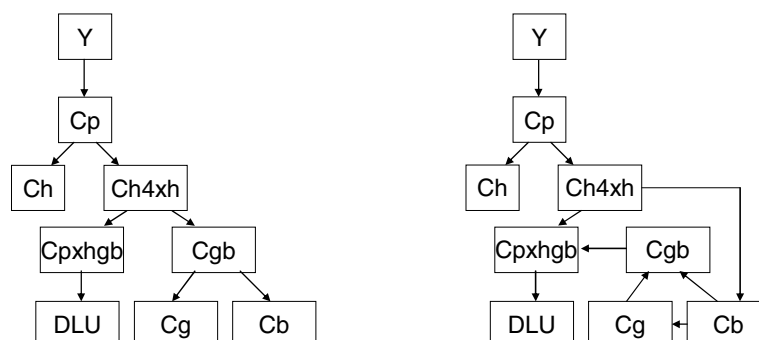
I den nuværende model-version bliver bilforbruget bestemt udenfor resten af forbrugssystemet påvirket af, Ycb , en sammenvejning af formue og indkomst. Det samlede transportforbrug bestemmes i DLU. Forbruget til benzin og kollektiv transport bestemmes residualt efter bilforbruget er fratrukket og fordeles via en CES-fejlkorrektionsmodel.

Et mål med den nye formulering er at få en mere naturlig strukturering. Umiddelbart er forestillingen, at efter boliger er blevet skilt fra forbruget skilles det samlede bil- og benzinforbrug fra inden resten af forbruget deles ud vha. DLU. En ulempe ved dette forslag er, at der skal antages svag seperabilitet mellem biler/benzin, Cgb , og andet forbrug, $Cpxhgb$. Endvidere er det nødvendigt at reestimere bilrelationen med $Cp4xhv1$ i stedet for Ycb - for at få en konsistent budgetbetingelse. En skitse for den valgte struktur er vist nedenfor. Estimeres biler direkte bestemmes Cgb nedefra. Dette er vist på skitsen til højre. Cb bestemmes ud fra $Ch4xh$. Herefter bestemms det relative bil/benzin forhold, og herved Cg i en model som den i GRH21205. Cgb bestemmes så som en simpel sum, og $Cpxhgb$ bestemmes residualt.

Som alternativ kan det forsøges, at estimere det samlede benzin/bil-forbrug ud fra den gamle bil-relation. Eller eventuelt fordelingen af $Ch4xh$ ud på $Cpxhgb$ og Cgb i et CES-nest a la den i GRH21205.

Kollektiv transport vil i den nye modelversion ikke være en selvstændig forbrugsgruppe og vil ikke indgå i det samlede transportforbrug.

Skitse over ny model:



3. Problemer og mulige løsninger ved at dele samlet bil- og benzinbudget ud på biler og benzin

Som nævnt i indledningen er et problem, at det ser ud til, at indkomstelasticiteterne er forskellige. Dette er ikke et problem på kort sigt, men vil på længere sigt betyde, at biler vil spise hele budgettet, hvilket ikke er en ønskværdig sammenhæng. Som vist i GRH21205 kan det ikke afvises, at indkomstelasticiteterne på lang sigt er ens. Faktisk kan stort set ingen hypoteser vedrørende det lange sigt afvises, når først de kortsigtede indkomstelasticiteter bliver frigivet. Noget tyder på, at det er de kortsigtede effekter, der dominerer.

Løsningen på dette problem vil være at binde de langsigtede pris- og indkomstelasticiteter til rimelige størrelser. Problemet bliver at finde disse rimelige størrelser. En mulighed er at tage udgangspunkt i mikrodata. En inspirationskilde til modellering kunne tages fra Bjørner (1997) eventuelt Bjørner og Leth-Petersen (2004). Man ville her eventuelt også kunne undersøge, hvor biased en aggregeret model er i forhold til individdata ved simulering. Eventuelt kunne man fastholde indkomstfordelingen og holde alle husholdningernes indkomst proportionalt i modellen og hermed udregne bias i en aggregering. Problemet her er at få data-udtræk. På flere punkter vil modellen adskille sig fra den givet af Bjørner (1997). Det bør undersøges om samlet forbrug eksklusivt bolig er en lige så god forklarende variabel som disponibel indkomst. Endvidere kan det være en idé at medtage usercost og lade udgifter til biler være et valg der afspejler bilens størrelse/komfort. På denne måde fås to dimensioner af bilvalget.

En yderligere mulighed er at opstille en fuld VAR-model for bedre at kunne undersøge modellens egenskaber, heriblandt hvilke indkomst- og priselasticiteter der kan bindes osv. Dette er muligt ved at opstille modellen på kvartaler eller måneder. Problemet er at få data langt nok tilbage i tid. Det er i GRH21205 vist, at der ikke er stor nok variation i data til at få plausible effekter, hvis ikke 1970'erne inkluderes. Denne konklusion vil højst sandsynlig også gælde for data med højere frekvens.

Det kan også være interessant at undersøge en endogen benzineffektivitet nærmere, selvom den rent statistisk så ud til at kunne betragtes som eksogen, jævnfør GRH21205.

4. Bestemmelse af bilstocken med privat forbrug eksklusiv boliger som indkomstvariabel

Idéelt set ønskes det at estimere det samlede benzin- og bilbudget ud af det samlede forbrug eksklusiv boliger og herefter dele det ud på biler og benzin. I dette afsnit vil det blive undersøgt om bilstocken, som nu forklares af Y_{cb} , kan forklares af $cp4xhv1$ i stedet.

I den nuværende model bliver bilstocken bestemt ved:

$$D \log K_t = \phi_1 D \log \left(\frac{Y_t}{p_{Y,t}} \right) + \phi_2 D \log \left(\frac{p_{b,t}}{p_{Y,t}} \right) + \beta_3 D d94_t - \gamma (\log K_{t-1} - \log K_{t-1}^o) + \phi_3 D \log K_{t-1} + \varepsilon_t \quad (1)$$

hvor $K_t \equiv fKcb$ er bilstocken, $Y_t = Ycb$ er et sammenvejet formue/pris-udtryk, $p_{Y,t} \equiv pcp4v2$ er prisudtryk for forbrugsfunktionen, $p_{b,t} \equiv ucbl$ er usercost, $d94_t$ er en dummy for skrotningspræmien i 1994 og K_t^o er den ønskede bilstock givet ved:

$$\log K_t^o = \beta_0 + \beta_1 \log \left(\frac{Y_t}{p_{Y,t}} \right) + \beta_2 \log \left(\frac{p_{b,t}}{p_{Y,t}} \right) + \beta_3 d94_t + \beta_4 \left(1 + \left(\frac{\log(Y_t / p_{Y,t})}{\lambda_1} \right)^{-\lambda_2} \right)^{-1} \quad (2)$$

Det sidste led er en logistisk trend, der giver mulighed for, at indkomstelasticiteten i en periode kan have været større end dens "naturlige" konstante niveau. Den økonomiske baggrund for at benytte denne trend er, som følger. Bilkøbet er et diskret valg, og før folk opnår en vis indkomst, vælger de ikke at have bil. Når folk først når en vis indkomst vælger de at investere i en bil. Den indkomstgrænse, hvor folk vælger at investere i en bil, vil selvfølgelig variere, og nogle vil vælge aldrig at investere i en bil uanset indkomst. Når den aggregerede indkomst er meget lav, vil stort set ingen købe bil, og selv om indkomsten stiger, vil det ikke rykke særlig mange over tærsklen. En meget høj

aggregeret indkomst vil betyde, at stort set alle dem, der potentielt vil købe en bil, har købt en bil, og selv om indkomsten stiger, vil det ikke rykke særlig mange over grænsen. For en mellemhøj aggregeret indkomst vil der dog være mange, der rykker over grænsen, og indkomstelasticiteten til bilejerskab vil være høj. Dette er altså et klassisk logistisk mætningsproblem. Hvis dette var hele historien ville indkomstelasticiteten gå mod nul, men dette vil ikke være tilfældet, eftersom bilstocken ikke kun udtrykker om man ejer en bil, men også hvor dyr en bil¹ man ejer. Det forventes, at folk med bil køber ”mere” bil, når deres indkomst stiger.² Denne effekt forventes dog ikke at være lige så stor som springet fra ingen bil til en bil. Slutresultatet er en bilmodel med en positiv indkomstelasticitet selv for en meget høj indkomst, men med en ekstra høj indkomstelasticitet i overgangsperioden fra ”bil-ejerskabs-take-off” til mætning.

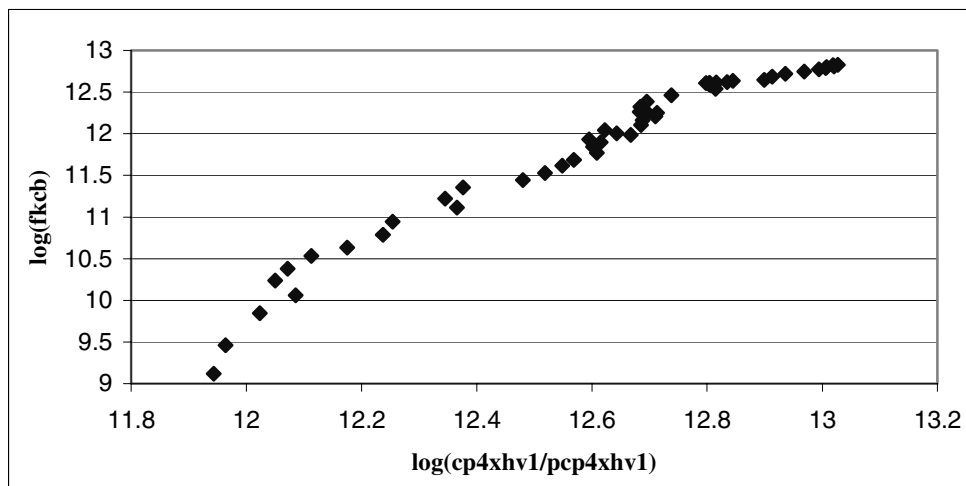
Rent teknisk er bidraget fra den logistiske trend 0, når indkomsten er meget lav, og β_4 når indkomsten er meget høj. Værdierne for λ_1 og λ_2 estimeres ikke direkte. λ_2 vælges for at få β_1 så tæt på 1 som muligt. Det statistiske belæg for dette kan diskuteres, men alternativet i langsigtede fremskrivninger er ikke kønt. Hvis den langsigtede indkomstelasticitet for biler i det samlede budget er større end en, vil biler til sidst udgøre hele budgettet – og dette er ikke en reel mulighed, så på lang sigt må indkomstelasticiteten på biler være en eller mindre. λ_1 vælges ved grid search for at maksimere ML-værdien. En idé om størrelsen af λ_1 kan findes ved at tage værdien for $\log(Y_t / p_{Y,t})$, hvor indkomstelasticiteten er størst, hvilket kan ses grafisk ved at plotte $\log K_t$ mod $\log(Y_t / p_{Y,t})$.

Først ønskes det at undersøge, hvordan modellen opfører sig, hvis alt bibeholdes og budgetvariablen redefineres så $Y_t \equiv cp4xhv1$ og $p_{Y,t} \equiv pcp4xhv1$. For at få et indtryk af indkomstelasticiteten plottes $\log K_t$ mod $\log(Y_t / p_{Y,t})$. Det ser ikke ud til at hverken en lineær eller lineær samt logistisk trend vil kunne frembringe nedenstående, jf. figur 1, dog er der ikke her taget højde for priselasticiteten.

¹ Eller eventuelt, hvor mange biler.

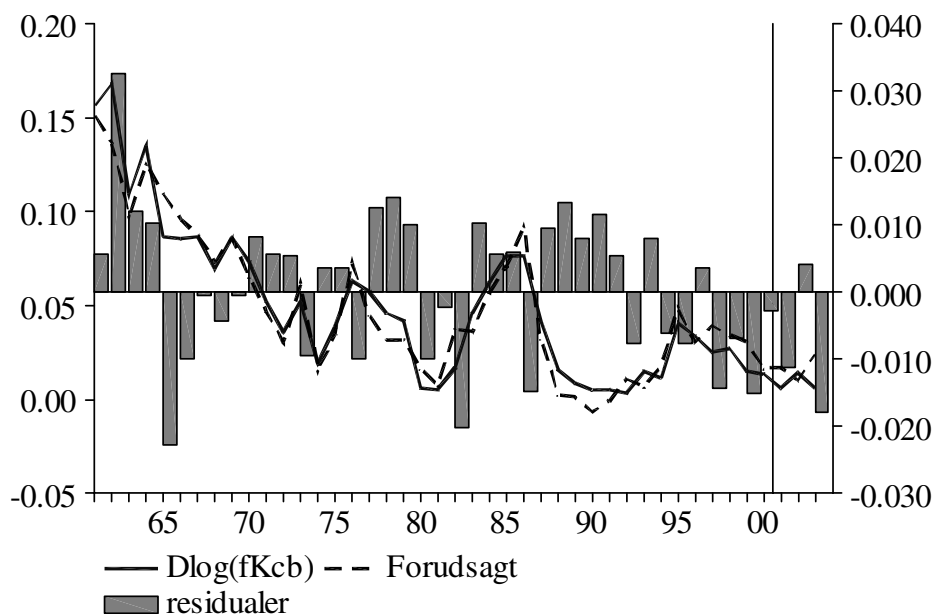
² Mere bil kan tænkes på som at købe en Mercedes i stedet for en Skoda og/eller en bil nummer 2.

Figur 1. Forhold mellem bilstock og samlet forbrug eksklusiv boliger fra 1952 til 2003.



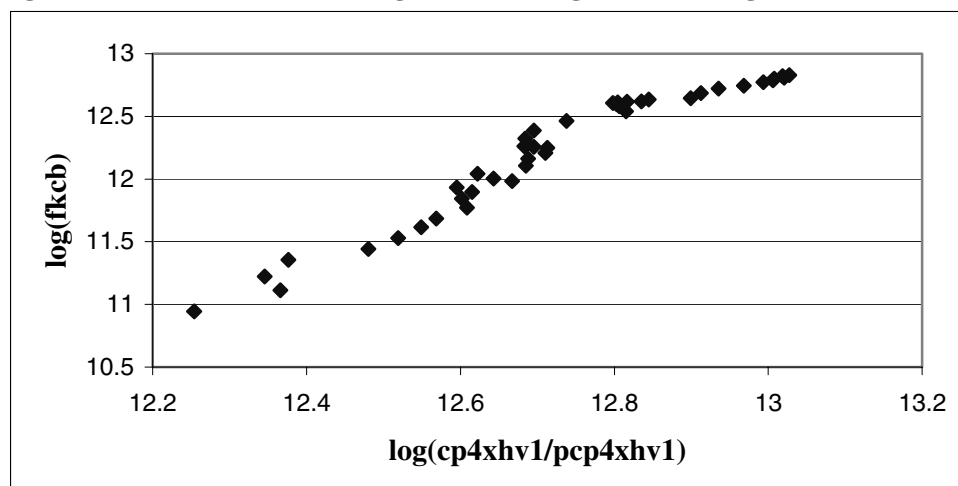
Ved estimation af modellen over hele dataperioden er det umuligt at fitte en logistisk trend til data, hvilket ikke er overraskende figur 1 taget i betragtning. Figur 2 viser hvor godt denne model fitter data, og tabel 1 angiver de estimerede parametre. Alt i alt giver denne model et dårligere fit end den gamle.

Figur 2. Bilmodel med cp4xhv1 som indkomstvariabel (sample 1953-2000)
Uden logistisk trend i indkomstelasticiteten.



I den gamle model er der kun brugt data fra 1961. Plottes $\log K_t$ mod $\log(Y_t / p_{Y,t})$ for perioden 1961 og frem fås et noget andet billede, jf. figur 3. Denne sammenhæng kan noget bedre ligne en sammensat lineær og logistisk trend. Problemet skyldes, at alle biler købt før 1947 ikke er inkluderet i stocken, og at der derfor vil komme et stort catch-up i denne periode. Hovedparten af denne catch-up er dog forekommet allerede inden 1960, men dette er en entydig årsag til kun at benytte data fra efter 1960.

Figur 3. Forhold mellem bilstock og samlet forbrug eksklusiv boliger fra 1961 til 2003.



Estimeres en model fra perioden 1961 til 2000 bliver den logistiske trend stærkt signifikant og nem at bestemme, og modellen giver en meget bedre tilpasning, jf. figur 4. Sammenlignet med den gamle model, jf. figur 5, er fittet, når cp4xhv1 benyttes øjesynligt bedre. Parameterestimerer er angivet i tabel 1.

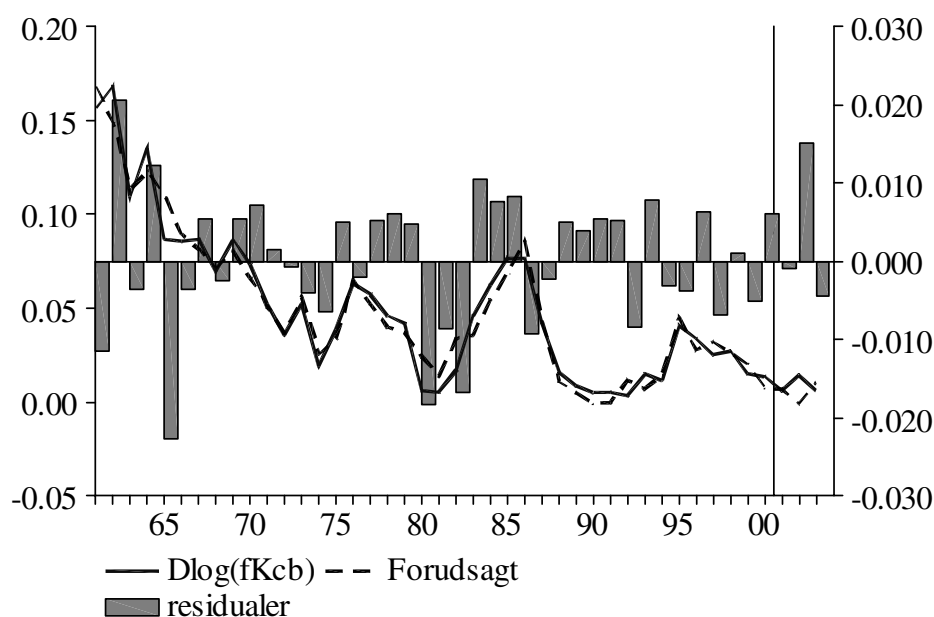
Tabel 1. Parameterestimerer for bilstockmodel med lagget endogen variabel.

	Gammel model	Ny model 1	Ny model 2
Indkomstvariabel	Ycb	cp4xhv1	cp4xhv1
Periode	1961-2000	1953-2000	1961-2000
β_0 (konstant)	-2.80741 (3.79462)	-8.98659*** (1.56703)	-2.09796 (2.40331)
β_1 (lang indkomstelast.)	1.03616*** (0.296736)	1.54859*** (0.134661)	1.00036*** (0.190844)
β_2 (lang priselast.)	-0.599264*** (0.120037)	-0.722008*** (0.103918)	-0.651173*** (0.071187)
β_3 (dummy)	-0.015827* (0.009318)	-0.026322* (0.015289)	-0.023195*** (0.008195)
β_4 (logistisk trend)	0.434068 (0.276784)	NA	0.367298*** (0.131185)
ϕ_1 (kort indkomstelast.)	0.245873*** (0.056901)	0.796343*** (0.134672)	0.609409*** (0.078674)
ϕ_2 (kort priselast.)	-0.05 (bundet)	-0.066123 (0.062556)	-0.037568 (0.039089)
γ (fejlkorrektion)	0.178332*** (0.023388)	0.217940*** (0.027607)	0.218943*** (0.048961)
ϕ_3 (lagget variabel)	0.40 (bundet)	0.165785* (0.089893)	0.252072*** (0.084316)
λ_1	12.87	NA	12.63
λ_2	-110	NA	-135

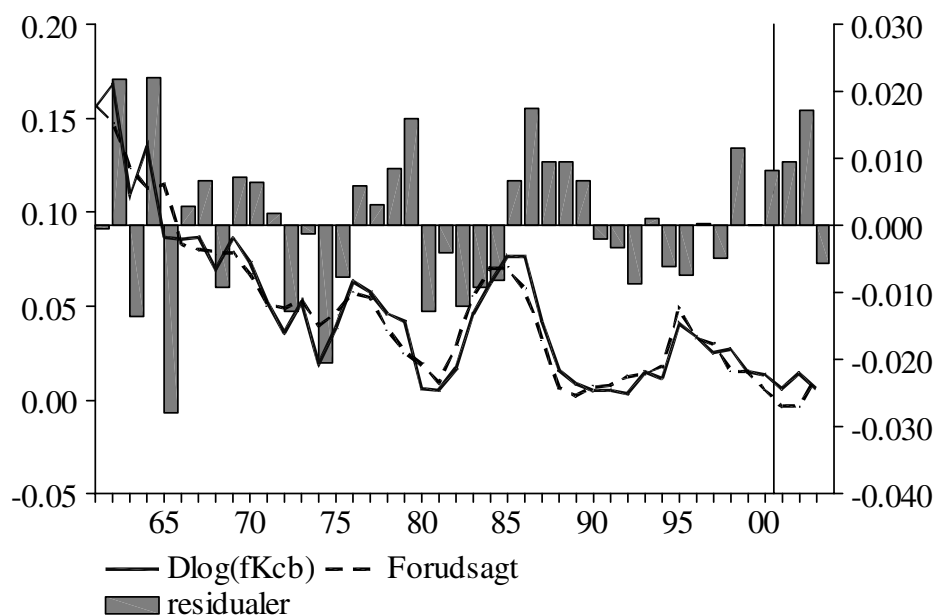
Note: Standardafvigelser i parentes. * signifikant på 10% niveau. ** signifikant på 5% niveau. *** signifikant på 1% niveau.

Sammenlignes parameterestimerne i tabel 1 fås følgende konklusioner. I og med en logistisk trend ikke kan fittes ind, når den fulde periode benyttes, fås en langsigtet indkomstelasticitet forskellig fra 1 ved ”ny model 1”. Denne forskel er yderst signifikant, og parameteren ligger ca. 4 standardafvigelser over. Når 1961-2000 bruges som periode og cp4xhv1 som indkomstvariabel fås ret pæne egenskaber. Modellen minder meget om den gamle, dog adskiller den sig på 3 punkter. Den logistiske trend er stærkt signifikant, hvilket ikke var tilfældet før. Endvidere er den kortsigtede indkomstelasticitet steget betydeligt, hvilket indikerer, at bilstocken i højere grad følger privat forbruget end den vægtede formue/indkomst-variabel, Y_{cb} . Dette betyder ikke nødvendigvis noget for den overordnede konjunkturfølsomhed, da de to ”indkomster” ikke er direkte sammenlignelige. Endeligt var det i den gamle model nødvendigt at binde ϕ_3 , ellers ville den være større end 0,6, hvilket ville give uheldige dynamiske egenskaber ved fremskrivning. Dette er ikke nødvendigt med den nye indkomstvariabel. Den laggede endogene variabel er stadig signifikant, men er af langt mindre størrelsesorden. Alt i alt må det konkluderes, at den nye model baseret på årene 1961-2000 faktisk har pænere egenskaber end den gamle, hvilket både undebygges af parameterestimerne givet i tabel 1, men også af de fittede værdier vist i figur 4 i forhold til dem fra den gamle model vist i figur 5.

Figur 4. Bilmodel med cp4xhv1 som indkomstvariabel og lagget endogen variabel (sample 1961-2000)



Figur 5. Gammel bilmodel med Ycb som indkomstvariabel

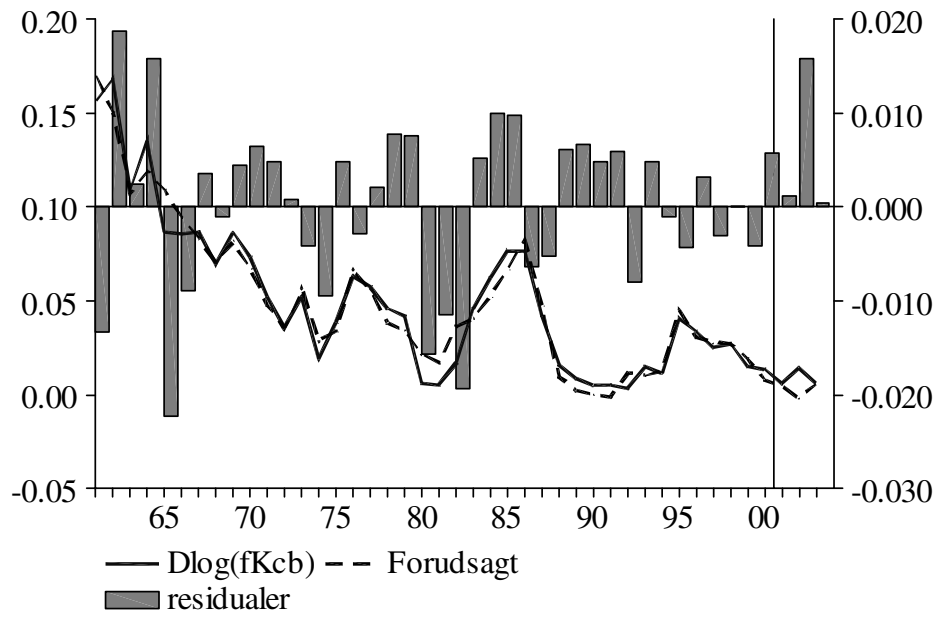


Laggede endogene variable kan give uheldige egenskaber i tilpasningsdynamikken især ved høj persistens. Persistensen er jf. tabel 1 ikke umådelig høj, så at inkludere en lagget endogen variabel er muligvis acceptabelt. Alternativt kan den forudsagte laggede endogene variabel inkluderes ved en såkaldt rho-konstruktion. Figur 6 viser den forudsagte og faktiske bilstock ved denne konstruktion. Konklusionen må være, at det ud fra fittet kan være næsten hip som hap om rho-konstruktionen eller en lagget endogen variabel inkluderes. Endeligt kan ϕ_3 bindes til at være 0 – dvs. den laggede endogene variabel fjernes. Eftersom ϕ_3 var af forholdsvis beskeden størrelse, giver det ikke anledning til et væsentligt dårligere fit, jf. figur 7. Dog skal det bemærkes, at ϕ_3 var signifikant selv på 1 % signifikansniveau. Estimerne for de øvrige parametre er ikke særligt følsomme overfor om specifikationer af ϕ_3 -leddet, jf. tabel 2.

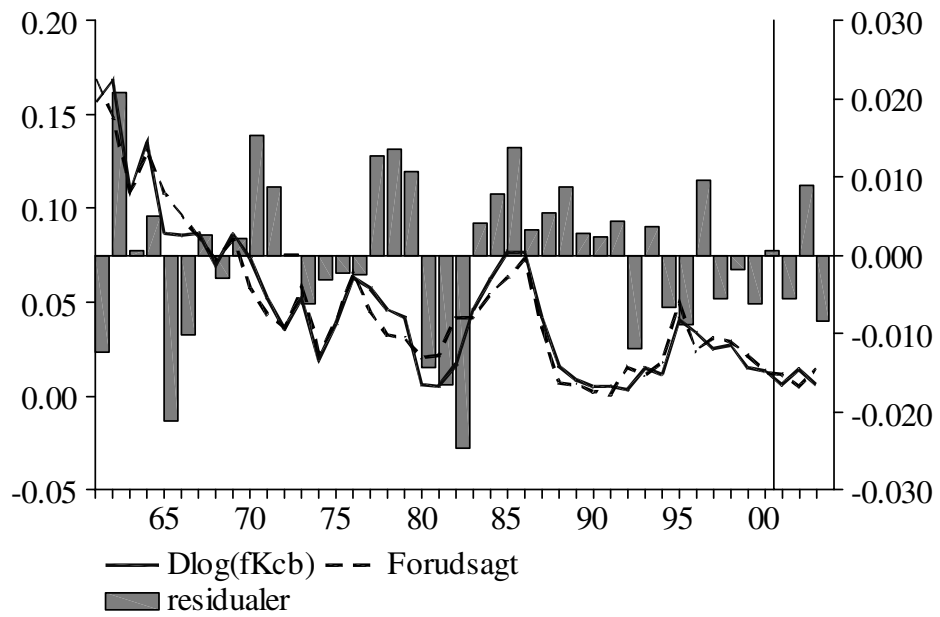
Umiddelbart har rho-konstruktionen flere attraktive træk. Den er nem at forklare; det er blot en AR-proces i fejlleddet. Endvidere vil et stød til J-leddet med en rho-konstruktion betyde en gradvis dæmpning, mens en lagget endogen vil have svingninger, der gradvist dør ud. Dog bør en endelig konstruktion vælges på baggrund af J-ledsundersøgelser og multiplikatoreffekter ved en ADAM-kørsel.

Et problem ved ovenstående estimation er som pointeret af Heino Bohn risikoen for simultaneitetsbias. Det er ikke rimeligt at betragte forbruget eksklusivt bolig som en eksogen variabel, idet denne bliver bestemt simultant med bilforbruget. En løsning på dette kan være at estimere relationen med indkomst og formue som instrument for forbruget eksklusiv bolig.

Figur 6. Bilmodel med cp4xhv1 som indkomstvariabel og Rho-konstruktion (sample 1961-2000)



Figur 7. Bilmodel med cp4xhv1 som indkomstvariabel uden lagged endogen variabel (sample 1961-2000)



Tabel 2. Parameterestimer for bilstockmodel for perioden 1961-2000 og med cp4xhv1 som indkomstvariabel.

	Lagget endogen	Rho-konstruktion	Udeladt lagget
β_0 (konstant)	-2.09796 (2.40331)	-2.06698 (1.97278)	-2.02782 (2.33044)
β_1 (lang indkomstelast.)	1.00036*** (0.19084)	1.00058*** (0.15692)	1.00086*** (0.18565)
β_2 (lang priselast.)	-0.65117*** (0.07119)	-0.64892*** (0.05782)	-0.67482*** (0.05312)
β_3 (dummy)	-0.02320*** (0.00820)	-0.02262** (0.00947)	-0.02151** (0.00951)
β_4 (logistisk trend)	0.36730*** (0.13119)	0.33701*** (0.10084)	0.27084*** (0.08903)
ϕ_1 (kort indkomstelast.)	0.60941*** (0.07867)	0.58916*** (0.08354)	0.58347*** (0.08980)
ϕ_2 (kort priselast.)	-0.03757 (0.03909)	-0.02271 (0.03962)	-0.03407 (0.04325)
γ (fejlkorrektion)	0.21894*** (0.04896)	0.20478*** (0.04675)	0.25558* (0.03933)
ϕ_3 (lagget variabel)	0.25207*** (0.08432)	0.33744** (0.14596)	0 (bundet)
λ_1	12.63	12.63	12.70
λ_2	-135	-140	-138

Note: Standardafvigelser i parentes. * signifikant på 10% niveau. ** signifikant på 5% niveau. *** signifikant på 1% niveau.

5. Konklusion

I papiret blev kort skitseret en plan for den nye bilmodel. En vigtig del af modellen er at bestemme budgettet til biler og benzin ud af det samlede private forbrug eksklusiv boliger, cp4xhv1. Det blev i dette papir undersøgt, om dette er muligt. Den klare konklusion må være, at det er det. Ikke blot er forudsigelserne ved at benytte cp4xhv1 bedre end ved at benytte den gamle indkomstvariabel Ycb, men det er også muligt at frigive tidligere bundne parameterestimer, samt at opnå mindre persistens og hermed pænere egenskaber under fremskrivninger.

Allerede nu er resultatet noget, der ligner en funktionel bilmodel. Bilstocken bestemmes på baggrund af det samlede private forbrug. Det relative benzin/bilforbrug bestemmes af CES-modellen givet i GRH21205. Når både det relative benzin/bilforbrug og bilforbruget er bestemt, burde det være nemt at bestemme benzinformbruget. Slutteligt bliver det samlede benzin- og bilbudget fundet ved simpel addition, og det resterende private forbrug findes og fordeles via DLU eller tilsvarende udgiftssystem. Det der mangler for at have en acceptabel bil- og benzinmodel, er rimelige parametre i den langsigtede relation for det relative benzin- og bilforbrug. Endvidere kunne det af rent kosmetiske årsager være rart at estimere det samlede bil- og benzinbudget og dele dette ud i stedet for at estimere bilstocken.

Litteratur

- Bjørner, Thomas Bue (1997), En økonomisk model for bilejerskab og bilkørsel, *Nationaløkonomisk Tidsskrift* 135 (1997), p. 233-250.
- Bjørner, Thomas Bue og Let-Petersen (2004), Dynamic Models of Car Ownership at the Household Level, *AKF working paper*.
- Høegh, Grane (2005), "Biler og benzin I CES-nests genbesøgt", GRH27105.
- Høegh, Grane (2005), "Modellering af benzin- og bilforbruget med benzineffektivitet", GRH21205.