

Sykefravær og konjunktursvingninger – en prognosemodell

AV OLE CHRISTIAN LIEN

SAMMENDRAG

Sykefraværet i Norge har svingt kraftig de siste 15-20 årene. Vi hadde fallende sykefravær fra slutten av 1980-tallet og frem til 1995. Deretter hadde vi en periode med sammenhengende vekst frem til år 2003. I 2004 og 2005 falt sykefraværet kraftig, før vi igjen fikk en viss økning i 2006. Gjennom store deler av denne perioden har det vært sterk samvariasjon mellom sykefravær og forholdene på arbeidsmarkedet. Perioder med gode tider på arbeidsmarkedet har ofte falt sammen med stigende sykefravær, mens sykefraværet har falt i perioder med høy arbeidsledighet.

Når man lager prognoser for sykefraværet er det en utfordring å klare å forutsi eventuelle vendepunkter. Denne artikkelen presenterer en prognosemodell for sykefravær som trekker inn forhold på arbeidsmarkedet som en av forklaringsfaktorene. Med denne modellen har vi laget prognoser for sykefraværet de siste fem-seks årene basert kun på data forut for prognoseperioden. Når vi sammenligner disse prognosene med historiske budsjettprognoser for sykefravær som har blitt laget av NAV uten å bruke noen formell modell, finner vi at denne prognosemodellen ser ut til å øke treffsikkerheten på prognosene, og at modellen i større grad lykkes med å forutsi vendepunkter.

Teoriene bak sammenhengen mellom sykefravær og arbeidsmarked er omdiskuterte, og vi vet ikke med sikkerhet om denne samvariasjonen vil gjelde også i årene fremover. De siste årene har vært preget av omfattende og stadig hyppigere regelverksendringer og tiltak. Dette gjør det vanskelig å lage et sammenlignbart datagrunnlag over tid. Ofte er regelverksendringene av en slik art at det er vanskelig å måle effektene, fordi det er komplisert å skille effektene fra utviklingen på sykepengeområdet for øvrig. Det er da vanskelig å etterprøve hvor gode prognoser sykepengemodellen gir. Modellen brukes derfor som en av flere innfallsvinkler når NAV vurderer utviklingen i sykefraværet.

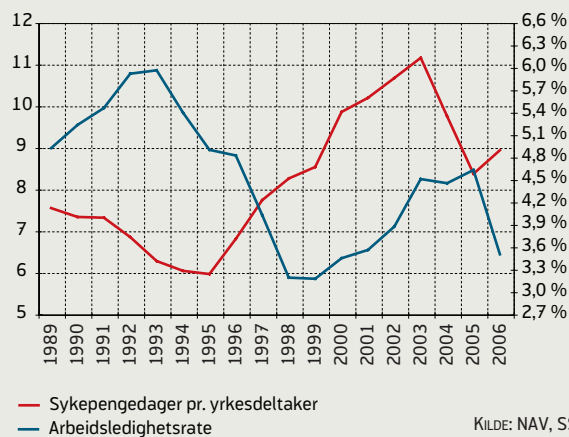
HØYT SYKEFRAVÆR I OPPGANGSTIDER

Historisk sett har sykefraværet i Norge svingt med konjunktorene. I oppgangstider har vi ofte observert høyt sykefravær, og omvendt i nedgangstider. Figur 1 og 2 viser hvordan sykefraværet har samvariert med viktige arbeidsmarkedsvariable som arbeidsledighetsraten og yrkesdeltakingsraten i perioden 1989–2006¹.

Vi observerer en klar negativ korrelasjon mellom sykefravær og arbeidsledighet og en klar positiv korrelasjon mellom sykefravær og yrkesdeltaking. Det ser videre ut til at det er en viss tidsforskyvning i denne korrelasjonen, slik at sykefraværet i et gitt år er sterkest korrelert med disse to variablene i året forut. Dette indikeres også av tabell 1

¹ Yrkesdeltakingsraten er et uttrykk for arbeidstilbudet og er definert som summen av antall sysselsatte og antall arbeidsledige i alderen 16-74 år som andel av befolkningen. Se ellers egen faktaboks om datagrunnlaget senere i artikkelen.

Figur 1.
Sykepengedager betalt av folketrygden pr. yrkesdeltaker og arbeidsledighetsraten 1989–2006.



KILDE: NAV, SSB

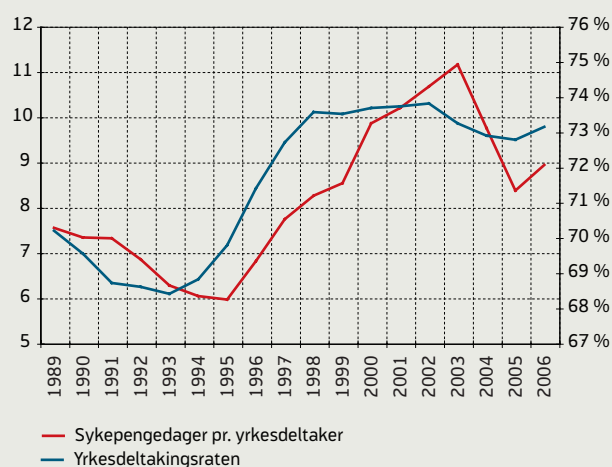
Tabell 1.

Korrelasjoner mellom sykefravær og arbeidsmarkedsvariable i perioden 1989–2006 (tall for «året før» gjelder 1988–2005).

	Arbeidsledighets-rate	Arbeidsledighets-rate året før	Yrkesdeltakings-rate	Yrkesdeltakings-rate året før
Sykepenge-dager pr. yrkes-deltaker	-0,65	-0,74	0,82	0,92
Arbeids-ledighets-rate	1,00	0,75	-0,92	-0,83

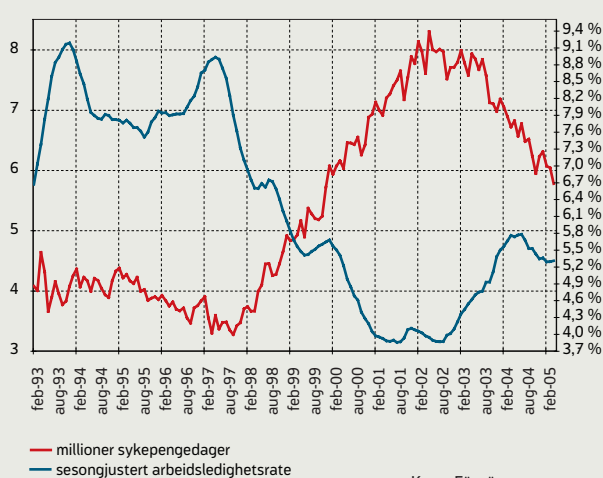
KILDE: NAV

Figur 2.
Sykepengedager betalt av folketrygden pr. yrkesdeltaker og yrkesdeltakingsraten 1989–2006.



KILDE: NAV, SSB

Figur 3.
Antall sykepengedager og arbeidsledighetsraten i Sverige feb. 1993–mar. 2005.



KILDE: FÖRSÄKRINGSKASSAN

som viser tall for korrelasjonen mellom sykefravær og arbeidsmarkedsvariable². Vi ser også at arbeidsledighet og yrkesdeltaking er minst like sterkt korrelert med hverandre som med sykefraværet, noe som kan gjøre det vanskelig å skille disse to effektene fra hverandre når det gjelder påvirkningen på sykefraværet.

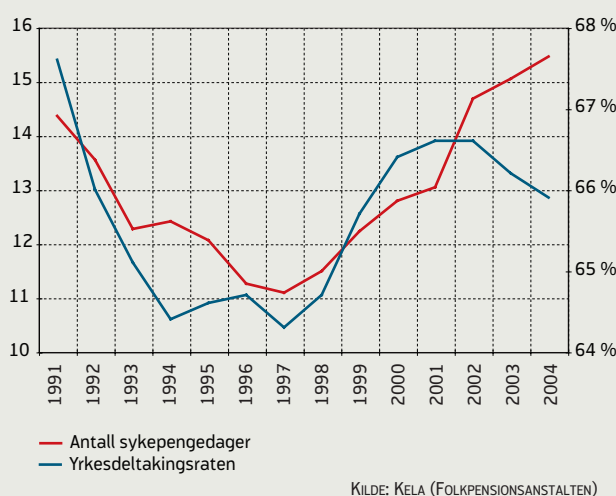
Sammenhengen mellom sykefravær og forhold på arbeidsmarkedet kan observeres også i andre nordiske land. Figur 3 og 4 viser utviklingen i antall sykepengedager og arbeidsledighetsraten i Sverige i perioden 1993–2005 og antall sykepengedager og yrkesdeltakingsraten i Finland i perioden 1991–2004.

ER ARBEIDSTAKERE MER DISIPLINERTE I NEDGANGSTIDER?

Det er ofte to hypoteser som brukes for å forklare korrelasjonen mellom sykefravær og konjunkturer – disiplineringshypotesen og sammensetningshypotesen. I Shapiro og Stiglitz (1984) påpekes det at sykefravær som skyldes skoft er kostbart for den enkelte, på grunn av risikoen for å bli oppsagt dersom dette oppdages. Kostnaden avhenger videre av arbeidsledighetsnivået, for jo høyere arbeidsledigheten er, jo vanskeligere er det å finne en ny jobb. Denne innfallsvinkelen til disiplineringshypotesen virker mest rimelig å anvende på korttidsfraværet.

Men som nevnt i Askildsen m.fl. (2002) kan også langtidsfravær ha indirekte kostnader for den enkelte. Sykefraværet kan påvirke lønnsomheten i bedriften og

Figur 4.
Antall sykepengedager og yrkesdeltakingsraten i Finland 1991–2004.



KILDE: KELA (FOLKPENSIONSANSTALTEN)

dermed risikoen for at virksomheten skal overleve. Videre kan sykefraværet påvirke den enkeltes karrieremuligheter. I tider med lav arbeidsledighet er det lettere å finne en alternativ jobb, og disse indirekte kostnadene er derfor lavere. Dette kan forklare at terskelen for å melde seg syk heves i dårligere tider, noe som kan gi et mønster i sykefraværet som er negativt korrelert med arbeidsledigheten og positivt med yrkesdeltakingen, slik vi ser i figur 1 og 2. I mange tilfeller er det vanskelig å definere en klar grense mellom hva som bør regnes som sykdom (og som rimelig grunn til å melde seg syk) og hva som ikke bør det, uansett om sykemeldingen skjer gjennom egenmelding eller sykemelding. På hvilket tidspunkt arbeidsevnen igjen er

² Korrelasjonskoeffisientene som er vist i tabellen er pr. definisjon alltid mellom -1 og 1, hvorav ytterpunktene betyr hhv. perfekt negativ og perfekt positiv lineær sammenheng mellom to variable. En korrelasjon på 0 tyder på at det ikke er noen sammenheng mellom to variable.

tilstrekkelig til at fraværet kan avsluttes, kan også til en viss grad være en subjektiv vurdering. Det kan dermed oppstå konjunktursvingninger i sykefraværet, selv om fraværet er helt legitimt og ikke relatert til skoft.

En annen forklaring på dette mønsteret kan være at marginale grupper som ellers har problemer med å komme inn på arbeidsmarkedet, har større muligheter for dette i perioder hvor det er mangel på kvalifisert arbeidskraft. Disse kan tenkes å ha høyere sykefravær enn andre, på grunn av dårligere helse eller sosiale problemer. Dersom disse gruppene utgjør en større andel av arbeidsstyrken i oppgangstider, vil dette isolert sett føre til høyere sykefravær i slike perioder. Dette kalles sammensetningshypotesen. Askildsen m.fl. (2002) har studert sykefraværdata for perioden 1990–1995, og finner at sykefraværet varierer mer med konjunktorene blant stabil arbeidskraft sammenlignet med sykefraværet blant marginale grupper som befinner seg i arbeidsmarkedet over kortere tidsrom. Sammensetningshypotesen ser altså ikke ut til å finne særlig støtte i empiri. Også andre undersøkelser tyder på det samme, bl.a. Kolstad (2005).

I den svenske debatten på dette området nevnes det også at innstramminger i regelverket for velferdsordninger som oftest skjer når kostnadene har økt mye, eller i en nedgangskonjunktur når det er press på statsfinansene. Regelverksendringene i seg selv kan dermed også skape et konjunkturmønster i sykefraværet (Ahonen m.fl. 2006). Disiplineringshypotesen og sammensetningshypotesen fokuserer på hvordan forhold på arbeidsmarkedet påvirker sykefraværet. Vi kan imidlertid ikke utelukke at årsaksammenhengen går i motsatt retning, i den forstand at sykefraværet påvirker yrkesdeltakingen. Dersom sykefraværet øker, kan det tenkes at virksomheter henter inn ytterligere arbeidskraft for å opprettholde produksjonen. I og med at sykemeldte også regnes med blant de sysselsatte, vil dermed økt sykefravær forårsake økt yrkesdeltaking. I figur 2 ser vi at endringer i yrkesdeltakingen ligger i forkant av endringer i sykefraværet, og det ser dermed ut til at vi finner mer støtte til den motsatte årsaksretningen i våre data.

NAVS PROGNOSEMODELL FOR SYKEPENGES

Selv om disiplineringshypotesen og sammensetningshypotesen er anerkjente hypoteser, er årsakssammenhengene bak disse hypotesene omdiskuterte. Det er uansett utvilsomt at det historisk sett har vært en sterk korrelasjon mellom sykefraværet og forholdene på arbeidsmarkedet. Som vist i figur 1 og 2, ser det også ut til at arbeidsmarkedet ligger noe i forkant av hva som skjer med sykefraværet. Tidligere har NAV ikke anvendt noen formell modell for å forutsi sykefraværsutviklingen i Norge. Prognoser for sykefraværet har blitt laget ved å studere den historiske veksten i sykefraværet, men uten å ha tallfestede sammenhenger mellom sykefraværet og sannsynlige forklarings-

Sykepengemodellen - modellbeskrivelse:

Modellen består av åtte separate regresjoner basert på data for følgende demografiske grupper: Begge kjønn 16–19 år, begge kjønn 20–24 år, kvinner 25–39 år, kvinner 40–59 år, kvinner 60–66 år, menn 25–59 år, menn 60–66 år og begge kjønn 67–74 år. Den avhengige variabelen er antall sykepengedager betalt av folketrygden pr. yrkesdeltaker innenfor hver undergruppe. Tidsenheten er kvartal. Modellen er spesifisert som en autoregressiv lineær regresjonsmodell. Med autoregressiv mener vi her at den avhengige variabelen i forrige periode inngår som en av forklaringsvariablene. Dessuten har alle variable i modellen har blitt logaritmisk transformert. På denne måten kan regresjonskoeffisientene tolkes som elastisiteter.

Generelt kan modellen uttrykkes slik⁵:

$$\ln(sp_{i,t}) = \beta_0 + (1 - \delta) \ln(sp_{i,t-1}) + \beta_1 \ln(y_{i,t-4}) + \beta_2 \ln(al_{i,t-4}) + \beta_3 \Delta \ln(y_{i,t-4}) + \beta_4 \Delta \ln(al_{i,t-4}) + \text{dummyvariable} + \varepsilon_{i,t}$$

der t angir periode (kvartal), der $sp_{i,t}$ betyr antall sykepengedager pr. yrkesdeltaker i undergruppe i , $y_{i,t-4}$ betyr yrkesdeltakingsraten for gruppe i for ett år siden (4 kvartaler tilbake i tid), $al_{i,t-4}$ betyr den aggregerte arbeidsledighetsraten for ett år siden, der β_1 - β_4 og $1 - \delta$ er regresjonskoeffisienter, der $\varepsilon_{i,t}$ er et feilledd for gruppe i og der Δ betyr førstedifferensen til en variabel (vi bruker årsdata for arbeidsledighet og yrkesdeltaking, og det er derfor snakk om endring fra året før). Merk at vi bruker den gruppespesifikke yrkesdeltakingsraten, men den aggregerte arbeidsledighetsraten. Grunnen til dette er at vi antar at alle arbeidstakere konkurrerer i samme arbeidsmarked, uavhengig av alder. Den aggregerte arbeidsledighetsraten blir derfor mer relevant ved forsøk på å estimere effekten av disiplineringshypotesen. Det er imidlertid vesentlig å bruke gruppespesifikke yrkesdeltakingsrater, i og med at endringer i befolkningsstrukturen med økende antall eldre isolert sett vil medføre at den aggregerte yrkesdeltakingsraten reduseres, helt uavhengig av konjunktorene. Utviklingen i de gruppespesifikke ratene påvirkes i hovedsak av forhold på arbeidsmarkedet (selv om de også kan påvirkes av demografiske endringer innad i gruppen).

Denne typen modell kalles i blant *partial adjustment model*, og den kan estimeres med ordinær minste kvadraters metode⁶. δ kalles the *coefficient of adjustment* («tilpasningskoeffisienten»). Når yrkesdeltakingsraten eller arbeidsledighetsraten endrer seg, vil antall sykepengedager pr. yrkesdeltaker gradvis tilpasse seg til sitt nye langsiktige nivå. δ uttrykker hastigheten til denne tilpasningen og er alltid

⁵ Det varierer mellom de ulike regresjonene hvilke forklaringsvariable vi har latt inngå (se vedlegg 1). I hovedsak har vi valgt forklaringsvariable i de ulike modellene basert på Mallow's $C(p)$ -kriteriet, se Mallow (1973), men vi har også gjort visse skjønsmessige vurderinger.

⁶ Gitt standardforutsetninger i denne typen regresjon. Det særlig vesentlig i denne typen modell at den laggede avhengige variabelen (som inngår som forklaringsvariabel) er ukorrelert med restleddet i regresjonen. Vi har testet for autokorrelasjon i restleddene ved å beregne den såkalte Durbins h for de ulike regresjonene, og konklusjonen er at nullhypotesen om ingen autokorrelasjon opprettholdes i samtlige tilfeller (se vedlegg). Se ellers f.eks. Gujarati (1995), kap. 17.10 for en nærmere metodebeskrivelse for denne typen modell.

mellom 0 og 1. $\partial = 0$ betyr at tilpasningen skjer umiddelbart, mens 1 betyr at tilpasningen skjer uendelig sakte. Merk at β_1 - β_4 kan tolkes som kortsiktselastisiteter, dvs. de umiddelbare effektene av endringer på arbeidsmarkedet. Det kan videre vises at langsiktselastisitetene kan uttrykkes som β_i / ∂ ($i=1,2,3,4$), og dette uttrykker effekten på lang sikt av en endring i forholdene på arbeidsmarkedet, alt annet likt.

I tillegg har vi inkludert dummyvariable for å ta høyde for sesongvariasjoner og for større endringer i lovverket (tidsseriene er prekorrigert for antatt effekt av enkelte lovendringer, mens andre er estimert i selve modellen). Yrkesdeltakingsraten og arbeidsledighetsraten er kraftig korrelert (jf. tabell 1), og dette kan føre til problemer med multikollinearitet. For hver av de åtte regresjonene, har vi derfor sørget for at kun én av disse variablene inngår på nivåform. Man kan lage prognoser for sykefraværet frem i tid ved å legge inn prognoser for arbeidsledighet og yrkesdeltaking i modellen. En ulempe er at prognosefeil for disse arbeidsmarkedsvariablene, kan forplante seg videre og gi prognosefeil for sykefraværet. Dette skaper ytterligere usikkerhet utover selve modellusikkerheten.

Om datagrunnlaget i modellen

Sykepengedager: Sykepengedataene i modellen gjelder sykepengedager betalt av folketrygden til arbeidstakere. Dette innbefatter dermed sykefravær utover arbeidsgiverperioden på 16 dager (før 1998: utover 14 dager). Dataene er splittet opp etter 8 demografiske grupper. Kilde: NAV.

Arbeidsledighetsraten: Vi har her brukt aggregerte arbeidsledighetstall fra Arbeidskraftundersøkelsen. Kilde: SSB.

Antall yrkesdeltakere / yrkesdeltakingsraten: Vi har her også brukt tall basert på fra Arbeidskraftundersøkelsen. Tallene er her splittet opp etter 8 demografiske grupper. Kilde: SSB.

Sykepengedager pr. yrkesdeltaker: Da det som hovedregel er en forutsetning å være i arbeid for å motta sykepenger, er det rimelig å se antall sykepengedager i forhold til utviklingen i antall sysselsatte. Ideelt sett burde derfor den avhengige variabelen i modellen vært *antall sykepengedager pr. sysselsatt*. Fordelt på de 8 demografiske gruppene som modellen er inndelt i, har vi imidlertid kun hatt tilgang til gruppespesifikke historiske tall og prognoser over yrkesdeltakingen og ikke over sysselsettingen. For å bruke et sammenlignbart datagrunnlag gjennom hele artikkelen, har vi derfor gjennomgående brukt *antall sykepengedager pr. yrkesdeltaker* som et mål på sykefravær. Antall yrkesdeltakere er et uttrykk for arbeidstilbudet, dvs. summen av sysselsatte og arbeidsledige. Merk også at sykepengedataene vi har brukt, kun gjelder arbeidstakere (ikke selvstendig næringsdrivende), mens tallene over yrkesdeltakere også innbefatter selvstendige. Dersom andelen selvstendig næringsdrivende endrer seg systematisk over tid, gir dette en ytterligere feilkilde i metodene som er brukt.

faktorer. Dette har gitt dårligere grunnlag for å forutsi vendepunkter i sykefraværet før de faktisk har oppstått. Vanskeligheter med å forutsi vendepunkter bidrar også til store prognosefeil i slike perioder. En formell modell gir ikke nødvendigvis bedre treffsikkerhet, men den gir bedre mulighet til å dekomponere utviklingen og dermed også prognosefeilen.

Sollie (2000) utviklet på oppdrag fra Finansdepartementet en enkel tidsseriemodell for sykefravær der forhold på arbeidsmarkedet inngikk som forklaringsvariable³. NAV har senere utvidet og videreutviklet denne modellen, som heretter omtales som sykepengemodellen (se faktabokser). Fra 2004 har NAV brukt denne modellen som en av flere kilder i sitt arbeid med å vurdere sykefraværsutviklingen, og i sine innspill knyttet til Regjeringens budsjettprognose på sykepengeområdet.

HVOR GODT VILLE VI TRUFFET OM VI HADDE BRUKT SYKEPENGEDELLEN DE SISTE ÅRENE?

Hvor god føyning en modell gir til dataene den er estimert på, er ikke nødvendigvis noen god indikasjon på hvor gode prognoser modellen vil gi, i og med at prognoser innebærer å forlenge modellen utenfor den tidsserien man har data for. Som en alternativ metode har vi derfor sløytet de siste årenes observasjoner i dataene våre og forsøkt å lage prognoser for disse årene kun basert på data for tidligere år, såkalte «out-of-sample prognoser». Dette er vist i figur 5, der vi har laget prognoser for sykefraværet kun basert på tilgjengelige data t.o.m. utgangen av år 2000, 2001, 2002, 2005 og 2006⁷. På grunn av de betydelige endringene i sykepengereguleringen i 2004, har vi ikke laget prognoser basert på data t.o.m. utgangen av 2003.

I tabell 2 har vi oppsummert treffsikkerheten ved out-of-sample prognosene. Tabellen viser en oversikt over den faktiske veksten i sykefraværet, og vi sammenligner denne med den predikerte veksten i følge sykepengemodellen. Hvor godt modellen treffer i seg selv, er i og for seg ikke det mest interessante her, men heller hvor godt modellen treffer i forhold til andre metoder for å forutsi sykefraværet. For hvert år har vi derfor funnet frem til to ulike historiske prognoser utarbeidet av Regjeringen i budsjetteringssammenheng etter innspill fra NAV og sett hvilken vekst i sykefraværet som ble lagt til grunn den gang. Disse prognosene har ikke blitt basert på noen formell modell.

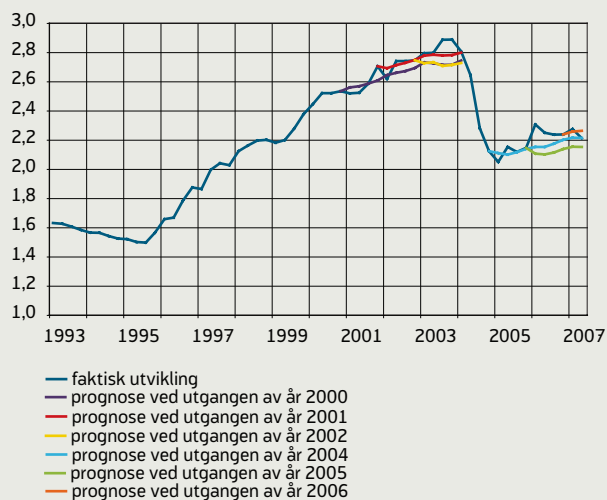
Når det gjelder prognoser ett år frem i tid, ser vi at sykepengemodellen treffer godt på veksten i 2001, 2002 og 2005. Treffsikkerheten i modellen er bedre enn i begge de

³ Se nærmere omtale av denne modellen bl.a. i Bjørnstad (2006).

⁷ Vi har hatt tilgang til historiske arbeidsledighetsprognoser fra SSB, men ikke historiske yrkesdeltakingsrater for de åtte gruppene som inngår i modellen. Vi har derfor brukt faktiske data for yrkesdeltaking selv om disse dataene ikke var tilgjengelig tilbake i tid. Dette er en viss svakhet ved metoden.

Figur 5.

Sykepengeprognoser basert på data t.o.m. 2000K4, 2001K4, 2002K4, 2004K4, 2005K4 og 2006K4. Sesongjusterte kvartalstall.



to historiske prognosene for alle disse årene. Det må riktignok nevnes at de historiske prognosene som ble laget for 2002 ble laget på et tidspunkt da det ble forventet store effekter av IA-avtalen, og at man derfor la til grunn lavere vekst i 2002 enn i årene forut. På veksten i 2003 traff de to historiske prognosene best, mens verken prognosen fra sykepengemodellen eller fra Statsbudsjettet traff særlig godt i 2006, i og med at vi fikk et vendepunkt og en betydelig økning i sykefraværet dette året. Økningen i sykefraværet i 2006 skjedde tidlig på året, og denne informasjonen

var dermed kjent da prognosen i Revidert nasjonalbudsjett for 2006 ble laget. Denne prognosen traff derfor rimelig bra. Selv om sykepengemodellen ikke klarte å predikere den betydelige veksten i 2006, så predikerte modellen basert på data ved utgangen av 2005 i stedet at det ville bli en vekst i 2007, basert på bedringen på arbeidsmarkedet i 2006 og 2007. Modellen forutså altså et vendepunkt her, men vendepunktet kom ett år tidligere enn forventet. Når det gjelder prognoser to år frem i tid, viser tabellen bare den faktiske veksten og prognosen fra sykepengemodellen. Vi finner her naturlig nok noe dårligere treffsikkerhet, men retningen på prognosene er riktige og avvikene fra faktisk vekst er ikke enormt store.

Tabell 2 gir for lite sammenligningsgrunnlag til å si noe sikkert om hvor gode prognoser sykepengemodellen gir. Noe av hensikten med modellen var å kunne forutsi vendepunkter. Vi hadde et vendepunkt i sykefraværet i 2004, men dette var for en stor grad forårsaket av regelverksendringer. Det er derfor ikke mulig å teste om modellen klarer å forutsi dette vendepunktet.

PASSER MODELLEN TIL SVENSKA DATA OGSÅ?

På grunn av korte tidsserier og stadige endringer i regelverk, er det ofte vanskelig å lage tilstrekkelig mange out-of-sample prognoser når man skal teste en tidsseriemodell. En alternativ mulighet kan da være å teste modellen på andre lands data.

Vi har hatt tilgang til månedlige data over antall sykepengedager, arbeidsledighetsraten og yrkesdeltakingsraten i Sverige i perioden 1993–2005¹⁰. Sykepengemodellen

¹⁰ Kilde: Försäkringskassan (takkt til Johan Ekstedt for fremskaffing av disse dataene).

Tabell 2.

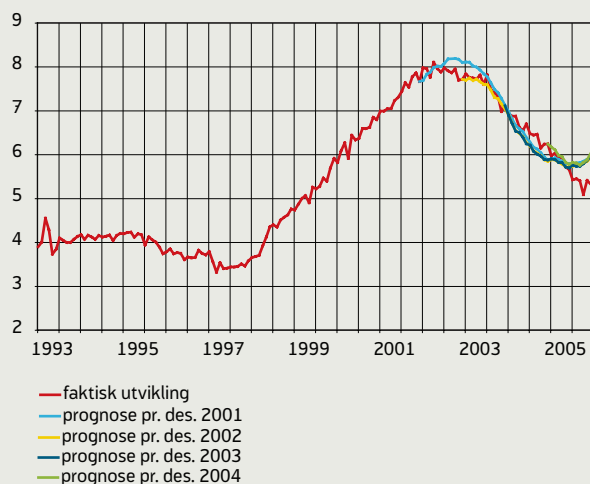
Vekst i antall sykepengedager pr. yrkesdeltaker – sammenligning av faktiske tall med prognoser fra sykepengemodellen og NAVs historiske budsjettprognoser.

År	Vekst siden året før				Samlet vekst siste 2 år ⁸	
	Faktiske tall	Modell-prognose	Historisk prognose ⁹		Faktisk vekst	Modell-prognose
			Fra Statsbudsjettet	Fra Revidert nasjonalbudsjett		
2001	3,1 %	3,0 %	10,0 %	2,0 %	-	-
2002	4,9 %	5,3 %	2,0 %	3,0 %	8,2 %	6,5 %
2003	4,8 %	0,3 %	3,0 %	2,0 %	9,9 %	7,6 %
2005	-14,2 %	-14,1 %	0,0 %	-7,0 %	-	-
2006	6,7 %	-0,1 %	0,0 %	8,0 %	-8,4 %	-12,0 %

⁸ Vi har ikke vist prognoser 2 år frem i tid basert på data t.o.m. utgangen av 2002. Dette skyldes de betydelige endingene i sykepengereguleringen som trådte i kraft fra juli 2004.

⁹ Prognosene er ikke helt sammenlignbare med prognosene fra sykepengemodellen fordi de ikke er basert på tall ved utgangen av året. Prognosene laget i forbindelse med Statsbudsjettet (publisert i St. prp. nr. 1 (2005-2006) og tidligere utgaver) er basert på tall t.o.m. juli året før, mens prognosene laget i forbindelse med Revidert Nasjonalbudsjett (publisert i St. prp. nr. 14 (2006-2007) og tidligere utgaver) er basert på tall t.o.m. februar-mai inneværende år.

Figur 6.
Sykepengeprognoser for svenske data basert på data t.o.m. utgangen av 2001, 2002, 2003 og 2004. Sesongjusterte månedstall.



har deretter blitt estimert basert på data t.o.m. utgangen av 2001, 2002, 2003 og 2004, og det har blitt laget prognoser for årene fremover. Disse prognosene er vist i figur 6. Akkurat som på norske data, finner vi i denne modellen at sykefraværet avhenger negativt av arbeidsledighetsraten og positivt av yrkesdeltakingsraten. Vi har dessverre ikke hatt tilgang til historiske prognoser for arbeidsledigheten

og yrkesdeltakingen i Sverige. I out-of-sample prognosene har vi derfor lagt inn faktiske tall for disse to variablene i årene prognosene gjelder for. Prognosefeil i arbeidsledighet og yrkesdeltaking kunne derfor ha ført til større avvik mellom prognosene og faktiske tall for perioden 2002–2005.

Det kom et vendepunkt i sykefraværet i Sverige i 2002. Out-of-sample prognosen basert på data t.o.m. utgangen av 2001 klarer å forutsi dette vendepunktet (selv om vendepunktet antas å komme rundt et halvt år senere enn det faktisk gjorde). Figur 6 viser ellers at out-of-sample prognosene treffer godt for samtlige år, med unntak av 2005. Merk at svenske arbeidsgivere fra 2005 ble pålagt å dekke 15 prosent av kostnadene til sykepenge utover arbeidsgiverperioden på 2 uker (se f.eks. Brage 2007). Denne regelverksendringen er ikke tatt hensyn til i modellen og kan forklare hvorfor samtlige prognoser overestimerer utviklingen i 2005.

Det har vært mange og stadig hyppigere regelverksendringer på sykepengeområdet i Norge de siste årene, og dette kan være noe av årsaken til at modellen ser ut til å treffe enda litt bedre på svenske data enn på norske data. At datagrunnlaget ikke er sammenlignbart over tid, er et problem uansett hva slags prognosemodell man tar i bruk. Likevel er kanskje sykepengemodellen noe mer utsatt for dette, ettersom det kreves en relativt lang tidsserie for å estimere effekten av forhold på arbeidsmarkedet. Sykepengemodellen brukes i dag derfor ikke som den eneste innfallsvinkelen, men sammen med mer uformelle metoder når NAV vurderer utviklingen i sykefraværet.

VEDLEGG 1. ESTIMERINGSRESULTATER

Tabell 3.

Sykepengemodellen estimert på norske kvartalsdata for perioden 1989K1-2007K2.

Populasjon	16-19 år	20-24 år	Kvinner 25-39 år	Kvinner 40-59 år	Kvinner 60-66 år	Menn 25-59 år	Menn 60-66 år	67 år og over
Parameter								
- konstantledd	-3,319***	0,089*	-1,970*	-5,901***	-0,135	0,178***	0,341***	1,139***
- sykepengedager pr. yrkesdeltaker i forrige periode	0,713***	0,912***	0,932***	0,769***	0,772***	0,942***	0,870***	0,890***
- 2. kvartal	0,777***	0,125***	-0,105***	-0,086***	-0,149***	-0,082***	-0,144***	-1,335***
- 3. kvartal	0,782***	0,100***	0,197***	-0,198***	-0,217***	-0,085***	-0,173***	-1,520***
- 4. kvartal	1,053***	0,222***	0,033**		-0,101***	0,031***	-0,086***	-1,613***
- yrkesdeltakingsraten for ett år siden	0,592***		0,475*	1,408***	0,166*			
- arbeidsledighetsraten for ett år siden		-0,125***				-0,079***	-0,045*	
- endring i yrkesdeltakingsraten siste år		0,437**						
- endring i arbeidsledighetsraten siste år								
- regelverksendringer i 1993 ¹¹	-0,114*	-0,101**	-0,059	-0,103***	-0,108**	-0,080**	-0,152***	
R (adj)	96,7 %	97,3 %	98,0 %	96,6 %	81,0 %	98,0 %	91,7 %	97,1 %
Durbins h								
(p-verdi)	0,447	-0,461	0,334	0,192	1,099	1,150	-0,100	0,625
	(0,655)	(0,645)	(0,739)	(0,848)	(0,272)	(0,250)	(0,920)	(0,532)

* **, *** = signifikansnivå på hhv. 10 %, 5 % og 1 %.

¹¹ Bl.a. innstrammning i de medisinske vilkårene for å få sykepenger.

KILDE: NAV

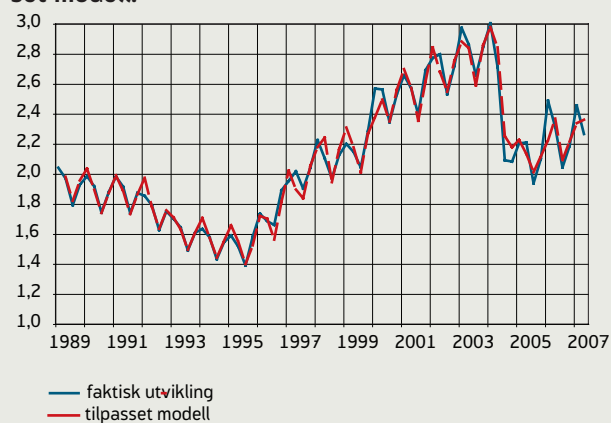
Tabell 3 viser estimeringsresultater når sykepengemodellen tilpasses norske kvartalsdata for perioden fra 1. kvartal 1989 til 2. kvartal 2007. Selv om modellen er estimert på årsdata, har vi kun hatt tilgang til årsdata for arbeidsledighet og yrkesdeltaking. En regelverksendring er estimert som dummyvariabel i modellen, men i tillegg er tidsserien prekorrigert for antatt effekt av en del andre regelverksendringer/statistikken- dringer¹². Som tidligere nevnt kan koeffisientene i modellen tolkes som elastisiteter. Eksempler: Arbeidsledighetsraten for ett år siden har en koeffisient på -0,079 for menn i alderen 25–59 år. Dette betyr at dersom arbeidsledigheten øker med 1 prosent (merk at dette gjelder prosent, ikke prosentpoeng), så vil sykefraværet for denne gruppen gå ned med 0,079 prosent på kort sikt. Regelverksendringene i 1993 har en koeffisient på -0,103 for kvinner i alderen 40–59 år. Dette tyder på at regelverksendringene førte til en nedgang i sykefraværet for denne gruppen på 10,3 prosent.

Figur 7 viser en tidsserie over faktiske data og en modelltilpasset tidsserie for samme periode som i tabel-

len. Vi ser at modellen kan forklare utviklingen svært godt, men dette er også vanlig i autoregressive modeller (der den avhengige variabelen inngår som en av forklaringsvariablene). Figur 7 gir derfor ikke alene noen god indikasjon på hvor gode prognoser denne modellen gir.

Figur 7.

Sykepengedager pr. yrkesdeltaker 1. kvartal 1989-2. kvartal 2007. Faktisk utvikling og tilpasset modell.



¹² Dette gjelder utvidelse av arbeidsgiverperioden fra 14 til 16 dager fra 1. april 1998, inkludering av statsansatte i statistikken fra 1. januar 2000 (sykepenger til statsansatte ble tidligere finansiert direkte av arbeidsgiver), samt de betydelige endringene i sykepengereguleringen fra 1. juli 2004 (bl.a. innføring av aktivitetskrav fra 8 uker, samt krav om at gradert sykemelding alltid skal vurderes først). Effekten av disse endringene er anslått utenfor modellen, og tidsserien er deretter korrigert slik at alle observasjoner blir «omregnet» til gjeldende regelverk.

REFERANSER

- Ahonen, Antero, Johan Ekstedt, Simon Gault, Linda Hauge, Sarah Hawken, Ole Chr. Lien, Markku Ryyänänen og Angela Smith (2006), «Sickness benefit models in Finland, Norway, Sweden and the United Kingdom», upublisert arbeidsgrupperapport i nettverket Improving Confidence in Forecasting (ICF)
- Askildsen, Jan Erik, Espen Bratberg og Øyvind Anti Nilsen (2002), «Unemployment, Labour Force Composition and Sickness Absence: A Panel Data Study», IZA Discussion paper no. 466, Forschungsinstitut zur Zukunft der Arbeit, Bonn
- Bjørnstad, Roger (2006), «Er det økte sykefraværet tegn på et mer inkluderende eller ekskluderende arbeidsliv?», Økonomiske analyser 6/2006 (Statistisk sentralbyrå)
- Brage, Søren (2007), «Trender i de skandinaviske sykefraværsordningene», NAV-publikasjonen Arbeid og velferd 2/2007
- Gujarati, Damodar (1995), «Basic Econometrics», McGraw-Hill
- Kolstad, Arne (2005), «Sykefravær og konjunkturer», RTV-rapport 1/2005 (Rikstrygdeverket)
- Mallow, Colin (1973), «Some comments on C», Technometrics vol. 15, no. 4 (nov. 1973), side 661-675
- Shapiro, Carl og Joseph E. Stiglitz (1984), «Equilibrium unemployment as a worker discipline device», the American economic review, vol. 74, no. 3 (jun. 1984), side 433–444
- Sollie, Marte (2000), «Utviklingen i folketrygdens utgifter til sykepengen», upublisert notat fra Statistisk sentralbyrå¹³
- St. prp. nr. 1 (2005–2006), Arbeids- og inkluderingsdepartementet, samt tidligere utgaver av denne tilbake til (2000–2001)
- St. prp. nr. 14 (2006–2007), «Om endringer av løyvingar på statsbudsjettet for 2006 under Arbeids- og inkluderingsdepartementet» (Revidert nasjonalbudsjett), samt tilsvarende utgaver fra tidligere år tilbake til (2001–2002)

¹³ Se evt. også Bjørnstad, A.F. og M. Sollie (2006), «Utviklingen i folketrygdens utgifter til sykepengen», Rapporten 2006/40, Statistisk sentralbyrå, som er en oppdatert og publisert versjon av dette notatet.