

Søren Risbjerg Thomsen

Meningsmålingernes troværdighed

En undersøgelse af meningsmålingerne i valgkampen op til folketingsvalget i 1998 samt af de månedlige meningsmålinger 1990-98 viser en bemærkelsesværdig god overensstemmelse mellem de enkelte meningsmålingsinstitutter samt pæn stabilitet fra måned til måned. Dette må tilskrives det forhold, at alle institutterne er gode til at måle partiernes tilslutning blandt de vælgere, som har besluttet sig. Institutterne Gallup, GfK og Sonar klarer sig i den henseende tilsyneladende bedre end Vilstrup og Greens, som også kan kritiseres for deres stikprøvemeter. De ubeslutomme vælgere kan undertiden gøre det vanskeligt at forudsige valgets udfald, hvilket skete i 1998. GfK og Gallup har dog haft relativt større held end de øvrige institutter med at forudsige udfaldet af folketingsvalgene i 1990'erne. Et gennemsnit over samtlige meningsmålinger er imidlertid tilsyneladende lige så godt som GfK's og Gallups målinger til at forudsige valgenes udfald samtidig med, at det er mere stabilt fra måned til måned.

Op til folketingsvalget den 11. marts 1998 spillede meningsmålingerne om, hvordan vælgerne ville stemme, på godt og ondt en stor rolle i den offentlige debat. Fra valget blev udskrevet den 19. februar til og med selve valgdagen, blev der offentliggjort ikke mindre end 31 målinger, hvilket tangerede rekorden fra den lidt længere valgkamp i 1994. Ikke mindst politikerne syntes påvirkede af det forhold, at regeringspartierne efter en flyvende start viste en tydelig tendens til tilbagegang; og man sporede i avisernes læserbreve en irritation over den rolle, målingerne var kommet til at spille.

Det var derfor ikke så lidt af en sensation, da det viste sig, at selve valgresultatet kom til at ligge et godt stykke fra meningsmålingerne; bortset fra de målinger, der blev foretaget i de allersidste dage før valget. Man kunne derfor med rette spørge, om meningsmålingerne på et fejlagtigt grundlag havde været med til at påvirke valgets udfald.

Denne artikel handler ikke så meget om, hvorvidt meningsmålingerne udgør en politisk påvirkning, hvilket turde være indlysende (jf. Elklit, 1989: 106). Den handler først og fremmest om, hvorvidt

det billede, målingerne gav i løbet af valgkampen, var troværdigt. Det er i den forbindelse vigtigt at være opmærksom på, at de marketinginstitutter, som laver målingerne, selv lægger vægt på, at der indtil nogle få dage før valget ikke er tale om egentlige prognoser for, hvordan valget vil komme til at gå, men om øjebliksbilleder af vælgernes intentioner. Det giver derfor god mening at foretage en nærmere undersøgelse af, hvorvidt samtidige målinger også kommer frem til det samme resultat (jf. Voss *et al.*, 1995).

Efter en gennemgang af meningsmålingernes anvendelighed til - trods alt - at kunne forudsige valgresultatet vil der blive givet en oversigt over samt en vurdering af de forskellige institutters metoder. Der vil derefter blive anvendt en særlig udviklet metodik til at vurdere overensstemmelsen mellem samtidige meningsmålinger i løbet af valgkampen 1998. Endelig vil undersøgelsen blive udstrakt til de månedlige meningsmålinger og de tre folketingsvalg i hele perioden siden 1990. I den forbindelse vil meningsmålingernes troværdighed blive kommenteret i lyset af begivenheder i dansk politisk historie.

Søren Risbjerg Thomsen

Meningsmålingernes troværdighed

En undersøgelse af meningsmålingerne i valgkampen op til folketingsvalget i 1998 samt af de månedlige meningsmålinger 1990-98 viser en bemærkelsesværdig god overensstemmelse mellem de enkelte meningsmålingsinstitutter samt pæn stabilitet fra måned til måned. Dette må tilskrives det forhold, at alle institutterne er gode til at måle partiernes tilslutning blandt de vælgere, som har besluttet sig. Institutterne Gallup, GfK og Sonar klarer sig i den henseende tilsyneladende bedre end Vilstrup og Greens, som også kan kritiseres for deres stikprøvemeter. De ubeslutomme vælgere kan undertiden gøre det vanskeligt at forudsige valgets udfald, hvilket skete i 1998. GfK og Gallup har dog haft relativt større held end de øvrige institutter med at forudsige udfaldet af folketingsvalgene i 1990'erne. Et gennemsnit over samtlige meningsmålinger er imidlertid tilsyneladende lige så godt som GfK's og Gallups målinger til at forudsige valgenes udfald samtidig med, at det er mere stabilt fra måned til måned.

Op til folketingsvalget den 11. marts 1998 spillede meningsmålingerne om, hvordan vælgerne ville stemme, på godt og ondt en stor rolle i den offentlige debat. Fra valget blev udskrevet den 19. februar til og med selve valgdagen, blev der offentliggjort ikke mindre end 31 målinger, hvilket tangerede rekorden fra den lidt længere valgkamp i 1994. Ikke mindst politikerne syntes påvirkede af det forhold, at regeringspartierne efter en flyvende start viste en tydelig tendens til tilbagegang; og man sporede i avisernes læserbreve en irritation over den rolle, målingerne var kommet til at spille.

Det var derfor ikke så lidt af en sensation, da det viste sig, at selve valgresultatet kom til at ligge et godt stykke fra meningsmålingerne; bortset fra de målinger, der blev foretaget i de allersidste dage før valget. Man kunne derfor med rette spørge, om meningsmålingerne på et fejlagtigt grundlag havde været med til at påvirke valgets udfald.

Denne artikel handler ikke så meget om, hvorvidt meningsmålingerne udgør en politisk påvirkning, hvilket turde være indlysende (jf. Elklit, 1989: 106). Den handler først og fremmest om, hvorvidt

det billede, målingerne gav i løbet af valgkampen, var troværdigt. Det er i den forbindelse vigtigt at være opmærksom på, at de marketinginstitutter, som laver målingerne, selv lægger vægt på, at der indtil nogle få dage før valget ikke er tale om egentlige prognoser for, hvordan valget vil komme til at gå, men om øjebliksbilleder af vælgernes intentioner. Det giver derfor god mening at foretage en nærmere undersøgelse af, hvorvidt samtidige målinger også kommer frem til det samme resultat (jf. Voss *et al.*, 1995).

Efter en gennemgang af meningsmålingernes anvendelighed til - trods alt - at kunne forudsige valgresultatet vil der blive givet en oversigt over samt en vurdering af de forskellige institutters metoder. Der vil derefter blive anvendt en særlig udviklet metodik til at vurdere overensstemmelsen mellem samtidige meningsmålinger i løbet af valgkampen 1998. Endelig vil undersøgelsen blive udstrakt til de månedlige meningsmålinger og de tre folketingsvalg i hele perioden siden 1990. I den forbindelse vil meningsmålingernes troværdighed blive kommenteret i lyset af begivenheder i dansk politisk historie.

Sensationen i 1998

Tabel 1 viser udviklingen i meningsmålingerne op til folketingsvalget i 1998 med selve valgresultatet i den nederste række opgjort som procent af afgivne stemmer på de enkelte partier ordnet efter en venstre-højre skala. De øvrige rækker i tabellen viser de enkelte meningsmålinger med oplysninger om offentliggørelsesdato, dataindsamlingsperiode, den procentvise tilslutning til partierne m.v. De to nederste meningsmålinger er såkaldte exit polls, hvor der indgår interviews, der er foretaget på valgdagen, efter folk har stemt.¹ De øvrige målinger er ikke-overlappende, idet der kun er medtaget uafhængige udpluk fra Gallups daglige valgbarometer, som hver dag var en opsamling fra fire forudgående dage. Det skal endvidere nævnes, at de tre første målinger vedrører data, som blev indsamlet, før valget blev udskrevet den 19. februar.

De tre første målinger er interessante i den forstand, at de viser en fremgang til Socialdemokratiet (SD), der resulterede i en tilslutning på 33 pct. i Vilstrup-målingen, som var gennemført fra 15/2 til 18/2 og offentliggjort i dagbladet *Politiken* den 20/2. Den var med til at give indtrykket af en flyvende start for den socialdemokratiske ledede regering, som blev bekræftet i samtlige meningsmålinger, der blev gennemført i resten af februar måned. Dog med den undtagelse, at en AIM måling, som blev offentliggjort i TV2 den 28/2, kun viste 30.8 pct. til SD. Et noget lavere niveau på 30-32 pct. til SD blev derefter fastholdt helt frem til weekenden inden valget, der blev afholdt på en onsdag. Da støtten til det andet regeringsparti, Det radikale Venstre (RV), samt til venstrefløjspartierne Enhedslisten (EnhL) og Socialistisk Folkeparti (SF) ikke kom-

penserede for SD's tilbagegang, var der lagt op til en borgerlig regeringsovertagelse. På selve valgdagen blev der imidlertid i dagbladet *BT* offentliggjort en GfK måling fra tirsdag aften, som viste fornyet fremgang for SD, hvilket også blev bekræftet i de to exit polls samt - ikke at forglemme - i selve valgresultatet, der førte til en uventet videreførelse af regeringen.

I kontrast til SD var valgresultatet for de to store borgerlige oppositionspartier, Venstre (V) og Det konservative Folkeparti (KF), ringere end nogen enkelt måling, som blev foretaget under valgkampen. De mindre borgerlige partier klarede sig bedre. Således kunne Centrum-Demokraterne (CD) fastholde den forøgede tilslutning, partiet havde fået ved at erklære sig som støtteparti for en borgerlig regering i begyndelsen af valgkampen; og Kristeligt Folkeparti (KRF) viste en støt fremgang fra at være under til at komme over spærregrensen på to pct. Endelig kunne højrefløjspartierne, Dansk Folkeparti (DF) og Fremskridtspartiet (FRP), nogenlunde fastholde tilslutningen fra valgkampen.

Som det fremgår af ovenstående, var meningsmålingerne generelt ikke særligt gode til at forudsige det endelige valgresultat. Dette fremgår også af tredjesidste kolonne i tabel 1 samt af figur 1, som viser udviklingen i meningsmålingernes afvigelse fra det endelige resultat, målt ved, hvor mange procent der skal flyttes i en meningsmåling for at få det endelige valgresultat.²

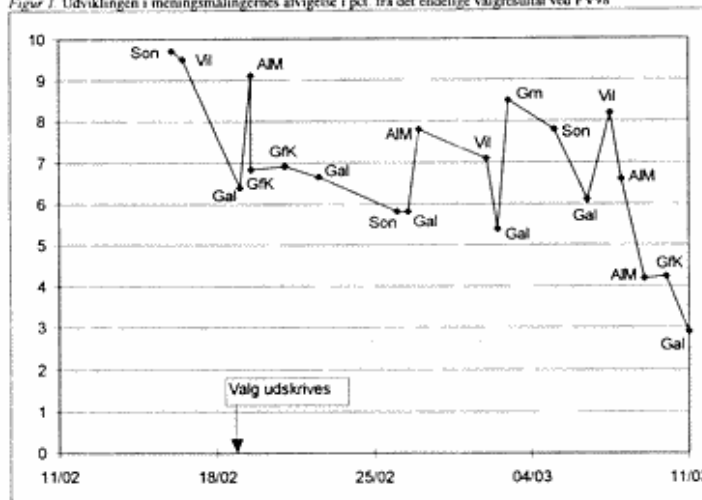
Det er erfaringen fra tidligere valgkampe helt tilbage til 1984, at meningsmålingernes afvigelse fra valgresultatet ligger på 3-5 pct. den sidste uges tid inden selve valget (Thomsen, 1995: 84-87). Denne gang var afvigelserne større, således at niveauet snarere var 5-8 pct., når

Table 1. Meningsmålinger op til folketingsvalget den 11/3 1998

Måling	Off. dato	Start dato	Slut dato	Gns. dato	Enthl %	DeF %	SF %	SD %	RV %	CD %	KRF %	KF %	V %	DF %	FRP %	Uafh %	Sum %	Stik- prove N	Gnf. %	Gnf. N	Afg. %	Afg. N	Afv. f. FV %	Forv. afv. %	Fit %
Greens	06/02	02/02	04/02	03/02	3.6	0.2	7.4	31.3	5.2	2.5	1.3	12.9	25.5	7.2	2.9	0.0	100.0	2308	1154	50	(808)	(70)	8.2	3.7	95
Sonar	20/02	16/02	16/02	16/02	2.8	0.4	9.2	29.5	3.9	2.5	1.5	13.0	25.1	10.0	2.1	0.0	100.0	1479	1283	87	1222	95	9.7	3.0	93
Vilstrup	20/02	15/02	18/02	16/02	3.4	0.0	7.4	33.0	5.6	1.6	1.7	13.4	26.2	5.1	2.1	0.5	100.0	1610	(999)	(62)	859	86	9.5	3.6	94
Gallup	21/02	18/02	20/02	19/02	3.5	0.0	6.5	33.3	4.0	3.5	1.7	11.5	26.2	8.1	1.7	0.0	100.0	1535	(952)	(62)	693	73	6.4	4.0	98
AIM	20/03	19/02	20/02	19/02	2.5	0.8	8.7	33.4	3.4	1.8	1.4	11.4	27.7	6.2	1.4	1.3	100.0	2277	1300	57	958	74	9.1	3.4	94
GIK	21/02	19/02	20/02	19/02	2.2	0.6	7.9	34.8	3.4	2.4	1.4	12.4	26.3	6.8	1.4	0.4	100.0	1341	913	68	873	96	6.8	3.6	97
GIK	27/02	21/02	21/02	21/02	2.1	1.2	6.6	32.9	3.5	3.5	2.1	12.1	26.0	8.2	1.7	0.1	100.0	1475	1255	85	1221	97	6.9	3.0	96
Gallup	25/02	21/02	24/02	22/02	3.2	0.2	6.8	32.9	4.7	3.9	1.9	12.8	25.5	7.1	1.1	0.0	100.1	2535	(1572)	(62)	1114	71	6.7	3.2	96
Sonar	01/03	26/02	26/02	26/02	2.1	0.3	8.8	32.8	4.4	3.1	2.2	9.9	26.6	7.8	2.0	0.0	100.0	1433	1150	80	1082	94	5.8	3.2	97
Gallup	01/03	25/02	28/02	26/02	3.3	0.3	6.9	32.6	4.1	3.9	2.4	11.8	25.5	7.8	1.1	0.3	100.0	2298	(1425)	(62)	988	69	5.8	3.4	98
AIM	28/02	26/02	28/02	27/02	2.6	0.4	8.3	30.8	4.3	4.1	1.9	10.1	27.9	8.1	0.7	0.8	100.0	2324	1300	56	957	74	7.8	3.4	95
Vilstrup	05/03	01/03	03/03	02/03	2.9	0.2	6.9	31.2	4.5	5.2	1.9	12.0	26.3	7.4	1.5	0.0	100.0	1700	995	59	856	86	7.1	3.6	96
Gallup	05/03	01/03	04/03	02/03	2.8	0.0	7.9	31.8	4.4	4.3	3.1	11.1	24.9	8.1	1.6	0.0	100.0	2258	(1400)	(62)	1026	73	5.4	3.3	98
Greens	06/03	02/03	04/03	03/03	3.4	0.1	8.0	29.8	4.2	3.9	2.3	12.5	27.4	6.4	2.0	0.0	100.0	2341	1156	49	828	72	8.5	3.7	95
Sonar	08/03	05/03	05/03	05/03	2.2	0.1	9.6	31.2	3.6	3.2	2.2	11.3	26.5	8.2	1.9	0.0	100.0	1505	1205	80	1141	95	7.8	3.2	95
Gallup	09/03	05/03	08/03	06/03	3.1	0.5	6.8	31.7	5.1	4.5	2.5	10.6	26.3	6.9	1.8	0.2	100.0	2276	(1411)	(62)	1068	76	6.1	3.3	97
Vilstrup	10/03	07/03	08/03	07/03	2.9	0.3	7.3	30.7	4.5	3.9	3.0	13.6	26.2	5.8	1.8	0.0	100.0	1570	995	63	856	86	8.2	3.6	95
AIM	10/03	07/03	09/03	08/03	2.9	0.2	8.5	32.0	4.0	3.6	2.2	12.4	25.8	6.5	1.8	0.1	100.0	4088	2336	57	2035	87	6.6	2.4	96
GIK	11/03	10/03	10/03	10/03	3.0	0.4	6.9	34.4	3.8	2.9	2.3	11.0	25.8	7.3	2.2	0.1	100.0	1199	932	78	883	95	4.2	3.6	99
Exit Polls																									
AIM	11/03	07/03	11/03	09/03	3.1	0.5	7.8	33.4	4.2	3.9	2.1	11.5	24.4	7.1	1.9	0.1	100.0	4088	2336	57	2006	86	4.2	2.2	98
Gallup	11/03	11/03	11/03	11/03	3.1	0.3	6.8	36.9	3.8	5.0	2.4	9.8	22.6	7.2	2.1	0.0	100.0	3429	2947	86	2788	95	2.9	2.0	99
FV98					11/03	2.7	0.3	7.5	36.0	3.9	4.3	2.5	8.9	24.0	7.4	2.4	0.1	100.0				85	0.0		

Off., Start, Slut og Gns. dato angiver offentliggørelse, start, slut og gennemsnitlig dato for meningsmålingerne foretaget af Gallup Berlingske-DRTV, GIK/BT, Sonar Jyllands-Posten, Vilstrup Politiken, Greens Børsen, AIM/TV2. Stikprove angiver stikprovestørrelse. Gnf. er gennemførte interviews som antal (N) og pct. af stikproven (%). Afg. er partistemmer som antal (N) og pct. af gennemførte interv. Afv. f. FV angiver hvor mange procent der skal flyttes i en meningsmåling for at få det endelige valgresultat FV98. Forv. afv. er den teoretisk forventede afvigelse fra FV98 på grund tilfældigheder, hvis der ikke var sket nogen reel ændring. Fit % er faktisk overensstemmelse i pct. af forventet overensstemmelse med FV98. Tallene i parentes er skønnede af SRT.

Figur 1. Udviklingen i meningsmålingernes afvigelse i pct. fra det endelige valgresultat ved FV98



Gal: Gallup/Berlingske/DRTV, GfK: GfK/BT, Son: Sonar/Jyllands-Posten, Vil: Vilstrup/Politiken, Grn: Greens/Børsen, AIM: AIM/TV2

der ses bort fra den sidste GfK-måling fra aftenen inden valget og de to exit polls. Disse større afvigelser, sammenholdt med de seneste målingers bedre prognosekvalitet, tyder på, at der skete et faktisk holdningskift i valgkampens sidste dage, muligvis fordi en del vælgere var kommet i tvivl om de borgerlige partiers økonomiske ansvarlighed og derfor endte med at stemme socialdemokratisk (jf. Thomsen, 1998; *Ugebrevet Mandag Morgen*, nr. 12, 23. marts 1998). Dette understreger også af de to sidste kolonner i tabel 1 som viser dels den teoretisk forventede afvigelse fra FV98 på grund af rene stikprøvetilfældigheder, dels meningsmålingens faktiske overensstemmelse med FV98 i procent af den teoretisk forventede overensstemmelse.³ I alle tilfælde var overensstemmelsen dårligere end de 100 pct., man skulle forvente, hvis afvigelserne alene skyldtes stikprøveusikkerhed, selv om de tre sidste målinger var tæt på.

Når dette er sagt, fremgår det også af figur 1, at Gallup og GfK gennemgående lå nærmest på det endelige valgresultat. Dette kunne tyde på, at målingerne fra disse to institutter var mere gyldige end

de øvrige målinger. Denne vurdering bygger imidlertid på den diskutale forudsætning, at meningsmålingerne er prognoser for det endelige resultat. For at få en alternativ vurdering af meningsmålingernes troværdighed, skal vi derfor se nærmere på samtidige målingers gensidige overensstemmelse. Der vil dog først blive foretaget en mere kvalitativ vurdering af de enkelte institutters metoder.

Hvilke krav kan man stille?

Formålet med de meningsmålinger, som ikke er exit polls eller foretaget dagen før valget, er at give et troværdigt og repræsentativt billede af, hvordan vælgere ville handle, "hvis der skulle være folketingsvalg i morgen". Spørgsmålet er til trods for sin enkelhed lidt absurd, da der netop ikke skal være folketingsvalg i morgen; men idéen er naturligvis at kunne give offentligheden en fornemmelse af, hvad den aktuelle status er med henblik på et kommende folketingsvalg. Det vil føre for vidt at komme ind på en generel diskussion af repræsentative interviewundersøgelser forskningsmæssige berettigelse (jf. Olsen, 1998); men menings-

målinger har den lighed med videnskabelige undersøgelser, at man i hvert interview prøver at iscenesætte en situation, hvorfra man kan generalisere til "det virkelige liv", altså hvordan svarpersonen ville handle ved et rigtigt folketingsvalg. Ligeledes i lighed med videnskabelige undersøgelser prøver man endvidere ved hjælp af statistisk stikprøveteorien at finde et udvalg af svarpersoner, som gør det muligt at generalisere fra stikprøven til hele vælgerkorpsen. I denne sammenhæng er det også vigtigt at kende til teknikker, hvormed man kan reparere på de skævheder i udvalget, som uundgåeligt vil opstå. Tager man hensyn til det noget snævre formål med en meningsmåling, kan man altså i øvrigt stille de samme krav til en meningsmåling som til en videnskabelig undersøgelse.

Spørgsmålet er så, om institutterne lever op til disse krav. Her vil jeg med det samme sige, at det er mit indtryk, at kravene generelt set ikke er dårligere overholdt end i videnskabelige undersøgelser. De private marketinginstitutter har simpelt hen bundet så meget prestige i deres meningsmålinger, at de ikke kan tillade sig at slække for meget på kravene. På den anden side er det dog også muligt at anvende den videnskabelige tilgang til at pege på kontroversielle forhold ved de enkelte undersøgelser. Jeg vil her tage udgangspunkt i den tidligere citerede artikel (Thomsen, 1995), som fokuserede på meningsmålinger forud for 1994 valget. I denne sammenhæng vil jeg søge at skabe et overblik over de ret så forskellige anvendte metoder, som stort set var uændrede siden 1994.⁴

To typer undersøgelser

Helt overordnet er der en principiel forskel på de målinger, som foretages af AIM, Gallup, Greens og Vilstrup på den

ene side og GfK og Sonar på den anden side. De førstnævnte anvender nemlig telefoninterviews med en ny stikprøve, hver gang der foretages en ny meningsmåling, mens de sidstnævnte sender postspørgeskemaer til et *panel* af stort set de samme personer fra gang til gang.⁵

De to typer undersøgelser implicerer vidt forskellige metodiske problemer. For eksempel har man med stikprøveundersøgelserne ikke de store problemer med at konstruere statistisk repræsentative udvalg af svarpersoner, mens man med panelundersøgelserne hele tiden har problemer med at sikre sig, at panelet er repræsentativt. Særlige repræsentativitetsproblemer med panelundersøgelserne er, at det kan være vanskeligt at motivere bestemte grupper, så som unge og "karriere-personer", til at deltage i panelet; samt at selve opholdet i panelet udgør en påvirkning af paneldeltagerne i retning af, at de bliver mere interesserede i politik. Ved den jævnlige supplering med nye paneldeltagere kan man ikke blot foretage et tilfældigt udvalg, men man må bevidst gå efter at skaffe den type vælgere, som man mener mangler i panelet. Denne *kvote-udvælgelse* er kritiseret i den videnskabelige litteratur (Lynn and Jowell, 1996), da dens succes afhænger af, om man har udvalgt efter de rette kriterier.

Det er et problem for stikprøveundersøgelserne, at de i stigende grad har mange "nægttere", dvs. personer som ikke ønsker at deltage i undersøgelsen, samt mange, som enten ikke vil oplyse partivalg eller er i tvivl. Inden for panelerne er der derimod høje svarprocenter i de enkelte spørgerunder, og der er nærmest for mange, som oplyser partivalg. Disse forskelle fremgår klart af tabel 1, hvor GfK og Sonar har både højere procent gennemførte interviews og højere procent afgivne stemmer end de øvrige institutter.

I selve svarsituationen er der også stor forskel mellem stikprøvernes telefoniske interviews og paneldeltagernes besvarelse af et postspørgeskema. Som et konkret eksempel kan nævnes, at man ved et telefoninterview skal nævne et bestemt parti, mens man ved besvarelsen af et postspørgeskema i lighed med selve folketingsvalget blot kan sætte kryds ud for et bestemt parti. Dette kan muligvis forklare, hvorfor der i starten var flere som tilsluttede sig Dansk Folkeparti i panelundersøgelserne end i stikprøveundersøgelserne. Det må nemlig formodes at være nemmere at genkende et nyt partinavn end at genkalde det.⁶

En anden forskel mellem stikprøveundersøgelser og panelundersøgelser fremkommer i forbindelse med korrigerende for skævheder i de opnåede svar. En væsentlig procedure er *politisk vejning*, hvor der korrigeres ud fra, hvor godt oplysningerne om tidligere partivalg svarer til det tidligere valgresultat (jf. Hellevik, 1991: 332-335). Det er her en fordel for panelundersøgelsen, hvis man har rekrutteret paneldeltagerne umiddelbart efter det sidste valg; mens man i stikprøveundersøgelserne i reglen er afhængig af, om folk kan huske, hvad de stemte sidste gang.

Endnu en væsentlig forskel mellem stikprøve- og panelundersøgelser drejer sig om vurderingen af ændringer mellem de enkelte målinger. Ser man bort fra ændringer på grund af udskiftninger i panellet er det nemlig således, at ændringer i partiernes tilslutning skyldes, at paneldeltagerne har oplyst, at de har ændret deres partivalg, mens ændringer mellem to stikprøveundersøgelser kan skyldes statistiske tilfældigheder.

Sammenfattende vil jeg sige, at stikprøveundersøgelseernes styrke er repræsentativiteten, mens panelernes styrke er svarsituationen samt stabiliteten over tid.

Det ligger i den måde, paneldeltagerne rekrutteres på, at et panel i perioder kan være repræsentativt og give særdeles troværdige resultater, mens det i andre perioder kan have en skæv sammensætning og giver konsekvent utroværdige resultater. Et stort problem i stikprøveundersøgelserne er de forskellige former for bortfald af svarpersoner, som jeg vil kommentere nærmere i de følgende afsnit.

Stikprøveundersøgelserne

Fælles for alle stikprøveundersøgelserne er, at de gennemføres ved hjælp af computerstøttede telefoninterviews. Man har altså på forhånd udelukket alle personer, som ikke bor i en husstand med telefon. Jf. TeleDanmark drejer det sig kun om en enkelt procent af samtlige husstande og udgør derfor ikke noget særligt problem.

Udvælgelsen af husstandene foregår i princippet ved, at computeren genererer tilfældige telefonnumre inden for *åbne serier* af mulige telefonnumre. AIM foretager denne udvælgelse simpel tilfældigt for hele landet, mens Gallup og Greens foretager udtrækningen tilfældigt inden for geografiske strata, hvilket giver lidt mindre statistisk usikkerhed. Vilstrup foretager udtrækningen ved systematisk fremtælling i regionale telefonnavnebøger, hvilket svarer til tilfældig udtrækning inden for geografiske strata; men proceduren har den uheldige konsekvens, at de hemmelige telefonnumre, der jr. TeleDanmark udgør 10-12 pct. af samtlige private telefonnumre, ikke kommer med.

Der, hvor vandene skilles, er ved udvælgelsen af en svarperson inden for husstanden. Den statistisk korrekte metode er at udvælge en tilfældig person med valgret i husstanden, uanset om vedkommende er hjemme eller ej. Dette kan i praksis gøres, ved at man blandt de valg-

berettigede i husstanden udvælger den, som har fødselsdag først.⁷ Er vedkommende ikke hjemme, må man ringe igen på et senere tidspunkt. Metoden anvendes af Gallup og AIM. Gallup er dog lidt mere ambitiøs end AIM, idet man anvender op til seks genopkald, mens AIM "ideelt set" nøjes med tre.

Man skulle tro, at Gallups mange genopkald kunne være med til at forsinke gennemførelsen af en given undersøgelse; men fidusen er, at proceduren anvendes vedvarende ved alle Gallups telefonundersøgelser dag efter dag (og i øvrigt også ved AIM's telefonundersøgelser, men altså med færre genopkald). Dette betyder, at man på en given dag har både nemme kontakter samt vanskelige kontakter med mange forudgående genopkald. Det bliver derved i praksis muligt at gennemføre en undersøgelse på en enkelt dag med den rette blanding af nemme og svære kontakter. Desværre giver denne fremgangsmåde ikke mulighed for at udregne den præcise procent gennemførte kontakter for en given undersøgelse. Gennemførelsesprocenten må enten udregnes for en længere periode eller skønnes med en vis usikkerhed.

Greens har valgt den nemme løsning med at interviewe den, som tager telefonen, hvis vedkommende har valgt. Vilstrup interviewer blandt valgberettigede medlemmer af husstanden ved to tredjedele af opkaldene den yngste tilstedeværende mand og i en tredjedel af opkaldene den yngste tilstedeværende kvinde for at kompensere for det forhold, at det især er unge og især er mænd, som er sværest at træffe hjemme. Problemet med både Greens' og Vilstrups personudvælgelse er, at personer, som sjældent er hjemme, bliver underrepræsenterede.

Bortfald

For samtlige institutter gælder det, at der er et betragteligt bortfald og dermed en fare for, at undersøgelserne ikke er repræsentative. Ved panelundersøgelserne sker bortfaldet især i forbindelse med rekrutteringen til panelet (hvor GfK oplyser, at kun 30-35 pct. vil medvirke), mens både gennemførelsen af kontakten og villigheden til at oplyse partivalg er høj. Det fremgår således af tabel 1, at de institutter, som anvender paneler, GfK og Sonar, både har høj procent gennemførte interviews (80-87 pct.; de noget lavere gennemførelsesprocenter for GfK 19-20/2 samt 10/3 skyldes, at der her var tale om telefonkontakt til panelmedlemmerne) samt meget høj procent med oplyst partivalg (94-97 pct.).

Ved stikprøveundersøgelserne er bortfaldet betragteligt både ved kontaktforsøgene, og når de udvalgte svarpersoner skal oplyse partivalg. Gennemførelsesprocenterne i tabel 1 (udregnet efter frasortering af erhvervsvirksomheder, faxnumre ol.) er lavest for Greens (49-50 pct.). Dernæst kommer AIM (56-57 pct.) og Vilstrup (59-63 pct.). Gallup udregner ikke gennemførelsesprocenten for den enkelte undersøgelse, men den oplyses til at være ca. 62 pct. for en længere periode.⁸ Det er bemærkelsesværdigt, at Greens og Vilstrup ikke kommer op på højere gennemførelsesprocenter, når de stiller sig tilfredse med at interviewe tilstedeværende personer, men det kan skyldes forskellig opgørelsespraksis. Den lidt højere gennemførelsesprocent for Gallup i forhold til Vilstrup kan måske tilskrives Gallups større antal genopkald. Under alle omstændigheder viser tallene, at alle institutter støder på en meget høj andel, som nægter at medvirke; og institutterne oplyser, at tendensen til at nægte er stigende.

Som det gjaldt for gennemførelsesprocenterne er procenterne med oplyst partivalg også lavere i stikprøveundersøgelserne end i panelundersøgelserne. Mærkeligt nok lå "stemmeprocenten" konstant høj hos Vilstrup på 86 pct., mens den for de tre øvrige institutter lå på 69-73 pct. i det meste af valgkampen. Det var først til allersidst, at stemmeprocenten hos AIM og Gallup kom i nærheden af den faktiske valgdeltagelse på 85 pct. Vilstrup afviser, at den højere stemmeprocent kan skyldes, at man giver mulighed for at nævne flere partier. Det gælder i øvrigt for alle stikprøveundersøgelserne, at man giver tvivlere en særlig tilskyndelse til at nævne et parti ved for eksempel at spørge, om der er et særligt parti, man hælder til at stemme på.

Det er ikke for meget sagt, at alle institutter har et alvorligt problem med et stort bortfald, enten fordi man ikke har fået fat i de relevante personer eller fordi svarpersonerne ikke kan eller vil oplyse partivalg. Man kommer derfor ikke uden om at skulle håndtere dette problem. Der foretages normalt to forskellige justeringer. Den første er *demografiske vejning* også kaldt *demografisk efterstratificering*, hvor man sørger for, at demografiske grupper fordelt efter køn, alder, geografi ol., som er underrepræsenterede i undersøgelsen, kommer til at veje tungere i den samlede opgørelse (Cochran, 1977: 134-135). Ikke mindre vigtig er den førnævnte politiske vejning, hvor man tager hensyn til oplysninger om tidligere partivalg. Her kan man støde på "forretningshemmeligheder"; men AIM oplyser dog, at man tager hensyn til mulige erindringsforskydninger angående tidligere partivalg. Det er mit klare indtryk, at meningsmålingerne ville give utroværdige resultater, hvis man ikke foretog den politiske vejning. Samtidig tyder den gode over-

ensstemmelse mellem de seneste meningsmålinger og selve valgresultatet både ved dette og tidligere folketingsvalg på, at de nævnte justeringer giver troværdige resultater. I 1998 var det dog kun GfK-målingen fra aftenen før valget, der lå tæt på det endelige resultat. Den kendsgerning, at de to exit polls også lå tæt på valgresultatet, er knap så anvendelig til at sige noget om kvaliteten af de øvrige meningsmålinger, da disse exit polls gennemføres på en anden måde end de øvrige målinger.

Konklusionen er, at meningsmålingernes troværdighed - på trods af det store bortfald - skyldes, at der stort set må foregå de samme forskydninger inden for den gruppe, det er nemt at få fat på, som inden for den gruppe, der sjældent er hjemme eller nægter at medvirke. Et særligt problem i 1998 i forhold til tidligere valg var nok, at de vælgere, som først besluttede sig i sidste øjeblik, stemte anderledes end dem, som havde besluttet sig tidligere i valgkampen. Dette bekræftes af en pressemeddelelse fra Gallup fra den 12/3 1998, hvoraf det fremgår, at næsten halvdelen af de 15 pct. af vælgerne, som først besluttede sig i de sidste dage inden valget, stemte på Socialdemokratiet. Gallups oplysning viser samtidig, at en stor del af disse tvivlere var til at få fat på, og således ikke var gemt i bortfaldet. Forekomsten af tvivlere kan altså belyses ved hjælp af meningsmålinger. Fra en prognosesynsvinkel er problemet imidlertid, at det kan være vanskeligt at forudsige, hvordan disse tvivlere i sidste instans vil opføre sig.

Stemmer meningsmålingerne overens?

Nu lægger institutterne som nævnt vægt på, at deres meningsmålinger ikke er pro-

Tabel 2. Gennemsnitlig parvis overensstemmelse mellem meningsmålinger i februar og marts 1998

a. Parvis forskel				b. Forventet parvis forskel				c. Parvis overensstemmelse			
	Febr.	Marts	Snit		Febr.	Marts	Snit		Febr.	Marts	Snit
Febr.	5.5	5.1	5.4	Febr.	4.9	4.8	4.8	Febr.	99.3	99.7	99.5
Marts	5.1	4.2	4.8	Marts	4.8	4.7	4.8	Marts	99.7	100.5	100.0
Snit	5.4	4.8	5.1	Snit	4.8	4.8	4.8	Snit	99.5	100.0	99.7

gnoser, men øjebliksbilleder af vælgerne meninger. Det må derfor være et rimeligt krav at stille til to forskellige meningsmålinger fra samme tidsrum, at de, bortset fra tilfældigheder, ligner hinanden så meget som muligt. Af denne grund er der udviklet en formel for den forventede forskel mellem to stikprøver, der er udtrukket fra den samme population.⁹ Forskellen måles ved, hvor mange procent, der skal flyttes i den ene stikprøve, for at få samme partifordeling som i den anden stikprøve.

For eksempel er forskellen mellem den AIM-måling, der er gennemført 19-20/2, og den GfK-måling, der er gennemført i samme tidsrum, lig med 3.7 pct., mens den forventede forskel er på 5.0 pct. Vi er altså i den lidt mærkelige situation, at de to meningsmålinger ligner hinanden "for meget". Dog er forskellen på 3.7 pct. ikke så lille, at den er signifikant forskellig fra den forventede forskel på 5.0 pct. Man kan også sammenligne den faktiske og forventede forskel mellem to meningsmålinger ved at beregne den faktiske overensstemmelse mellem de to meningsmålinger i procent af den teoretisk forventede overensstemmelse ved $100 \leftrightarrow (100 - 3.7) / (100 - 5.0) = 101$. Man får altså som ventet i dette eksempel en overensstemmelse som er lidt bedre end de 100 pct. man skulle forvente, hvis stikprøverne var simpel tilfældige stikprøver fra den samme population.

Med målingerne i tabel 1 er beregnet faktisk forskel, forventet forskel samt overensstemmelse for alle mulige parvise sammenligninger af meningsmålingerne (undtagen exit polls). For at få et overblik over alle disse tal er der i tabel 2 vist gennemsnitstal for samtlige indbyrdes parvise sammenligninger, som dog også er opdelt i henholdsvis februar og marts målinger.

Resultaterne i tabel 2 er overraskende. Det viser sig nemlig, at alle meningsmålingerne stemmer særdeles godt overens. Går man således ud fra, at partiernes tilslutning skulle have været helt uændret gennem hele valgkampen, ville den gennemsnitlige forventede parvise forskel mellem samtlige meningsmålinger i valgkampen jf. tabel 2 have været 4.8 pct., mens den faktiske forskel var på 5.1 pct., dvs. kun lidt større end den forventede. Den gennemsnitlige parvise overensstemmelse er udregnet til at være 99.7 pct., altså kun lidt mindre end de ideelle 100 pct. Tager man i betragtning, at det dels er en urealistisk antagelse, at partierne slet ikke ændrede tilslutning i løbet af valgkampen, samt at der er mange mulige fejlkilder på grund af bortfald og lignende, må man sige, at der er en imponerende grad af overensstemmelse mellem målingerne.

Resultaterne er endnu mere overraskende, når man ser på opdelingen i henholdsvis februar-målinger og marts-målinger i tabel 2. Sammenligner man fe-

Tabel 3. Parvis grad af overensstemmelse mellem de enkelte målinger i valgkampen 1998

	Gms	Son	Vilst	Gall	AIM	GfK	GfK	Gall	Son	Gall	AIM	Vilst	Gall	Gms	Son	Gall	Vilst	AIM	GfK	
	3/2	16/2	16/2	19/2	19/2	19/2	21/2	22/2	26/2	26/2	27/2	2/3	2/3	3/3	5/3	6/3	7/3	8/3	10/3	Snit
Greens	3/2	100	101	101	97	99	99	101	99	100	98	101	99	100	99	100	101	100	100	100
Sonar	16/2	100	97	98	96	97	97	98	98	98	97	98	98	99	99	96	98	98	97	98
Vilstrup	16/2	101	97	100	99	100	98	100	98	99	96	99	97	99	97	99	100	99	99	99
Gallup	19/2	101	98	100	99	100	103	102	101	103	99	102	101	100	101	101	100	101	102	101
AIM	19/2	97	96	99	99	101	99	98	99	98	99	98	97	98	99	97	97	99	99	98
GfK	19/2	99	97	100	100	101	101	100	99	100	98	99	98	99	99	98	99	100	101	99
GfK	21/2	99	97	98	103	99	101	101	100	102	98	101	100	99	100	100	99	100	101	100
Gallup	22/2	101	98	100	102	98	100	101	99	103	98	102	100	100	98	101	101	101	101	100
Sonar	26/2	99	98	98	101	99	99	100	99	100	101	100	100	99	102	100	99	100	101	100
Gallup	26/2	100	98	99	103	98	100	102	103	100	100	102	102	100	100	101	100	101	102	101
AIM	27/2	98	97	96	99	99	98	98	101	100	100	100	100	100	100	100	99	99	98	99
Vilstrup	2/3	101	98	99	102	98	99	101	102	100	102	100	101	101	100	102	102	101	100	101
Gallup	2/3	99	98	97	101	97	98	100	100	102	100	101	100	101	101	100	101	100	100	100
Greens	3/3	100	99	99	100	98	99	99	100	99	100	100	101	100	100	100	102	101	99	100
Sonar	5/3	99	99	97	101	99	99	100	98	102	100	100	100	101	100	99	99	100	100	100
Gallup	6/3	100	96	99	101	97	98	100	101	100	101	100	102	101	100	99	101	100	101	100
Vilstrup	7/3	101	98	100	100	97	99	99	101	99	100	99	102	100	102	99	101	101	99	100
AIM	8/3	100	98	99	101	99	100	100	101	100	99	101	101	101	100	100	101	100	100	100
GfK	10/3	100	97	99	102	99	101	101	101	101	102	98	100	100	99	100	101	99	100	100
Snit		100	98	99	101	98	99	100	100	100	101	99	101	100	100	100	100	100	100	100

Den anførte dato er den gennemsnitlige indssamlingsdato.

bruar-målingerne indbyrdes, får man ganske vist en ringere end ideel overensstemmelse mellem målingerne; men sammenligner man marts-målingerne indbyrdes får man det resultat, at målingerne stemmer "for godt" overens. Den gennemsnitlige forskel mellem marts-målingerne på 4.2 pct. er markant mindre end den forventede gennemsnitlige forskel på 4.7 pct., hvilket svarer til en overensstemmelse på 100.5 pct.

I tabel 3 er vist graden af overensstemmelse for samtlige parvise sammenligninger mellem alle meningsmålingerne fra valgkampen i 1998. Der er nogle enkelte tilfælde, hvor man faktisk skulle forvente højere overensstemmelse end 100 pct. Det er for panelerne, hvor der ikke er uafhængighed mellem de enkelte målinger, da det er de samme personer, man udspørger fra gang til gang. Det viser sig da også, at de tre GfK-målinger i løbet af valgkampen har en indbyrdes overensstemmelse på 101 pct. Der er også en overensstemmelse

mellem Sonars måling fra den 16/2 og målingen fra den 5/3 på 102 pct.; men overensstemmelsen mellem Sonars måling fra den 16/2 og den 26/2 er kun på 98 pct. Da de øvrige institutter ikke har overlappende meningsmålinger i tabel 1, skulle man ikke umiddelbart forvente mange overensstemmelser over 100 pct., men specielt Gallup har i 8 ud af 9 mulige parvise sammenligninger en overensstemmelse på over 100 pct. I øvrigt er der også mange sammenligninger med høj overensstemmelse på tværs af institutter, især i marts måned.

Der kan være flere forklaringer på den høje grad af overensstemmelse mellem de enkelte meningsmålinger. Den mest optimistiske er, at de anvendte formler gælder for simpel tilfældigt udvalgte stikprøver, og således ikke tager højde for en meget vellykket efterstratificering hos alle institutter. En mere pessimistiske forklaring ville være en påstand om, at institutterne hele tiden skæver til hinandens

resultater og løbende ændrer deres procedurer for ikke at skille sig for meget ud og samtidig dæmper kraftige udsving i forhold til tidligere målinger. Med det kendskab jeg har til institutterne, finder jeg dog ikke den sidstnævnte forklaring særlig sandsynlig. Det eneste, jeg kunne forestille mig i den retning, er, at hvis et institut i en længere periode har målinger, som adskiller sig markant fra alle andre institutter, vil man begynde at overveje sine procedurer; hvilket vil være helt legitimt (jf. note 6). Jeg har derimod svært ved at forestille mig, at alle over en kam skulle ændre procedurer fra februar til marts måned 1998, og dermed skabe en falsk højere overensstemmelse mellem samtlige målinger.

Ikke desto mindre synes jeg, at man har et problem, når man skal forklare den relativt dårlige overensstemmelse mellem valgkampens meningsmålinger og det endelige valgresultat. I lyset af den tidligere diskussion omkring bortfaldsproblemerne og tvivlerne vil jeg mene, at en mulig forklaring er, at de danske institutter er dygtige til at måle meningene hos de vælgere, som har besluttet sig endeligt for deres valg. Forekomsten af vælgere, som ikke har besluttet sig eller kunne finde på at ændre partivalg, kan imidlertid undertiden gøre meningsmålingerne uegnede til prognoseformål. Det skal dog påpeges, at GfK og Gallup i lighed med sidste valg for det meste lå tættere på det endelige valgresultat end de øvrige institutter. Disse institutter er altså tilsyneladende kommet nærmere på en løsning af problemet med de ubesluttsomme vælgere end de øvrige institutter. Specielt lå Greens og Vilstrup, der også har de mindst tilfredsstillende statistiske procedurer for udvælgelse af svarpersoner, med størst afstand til det endelige valgresultat.

Meningsmålingerne i 1990'erne

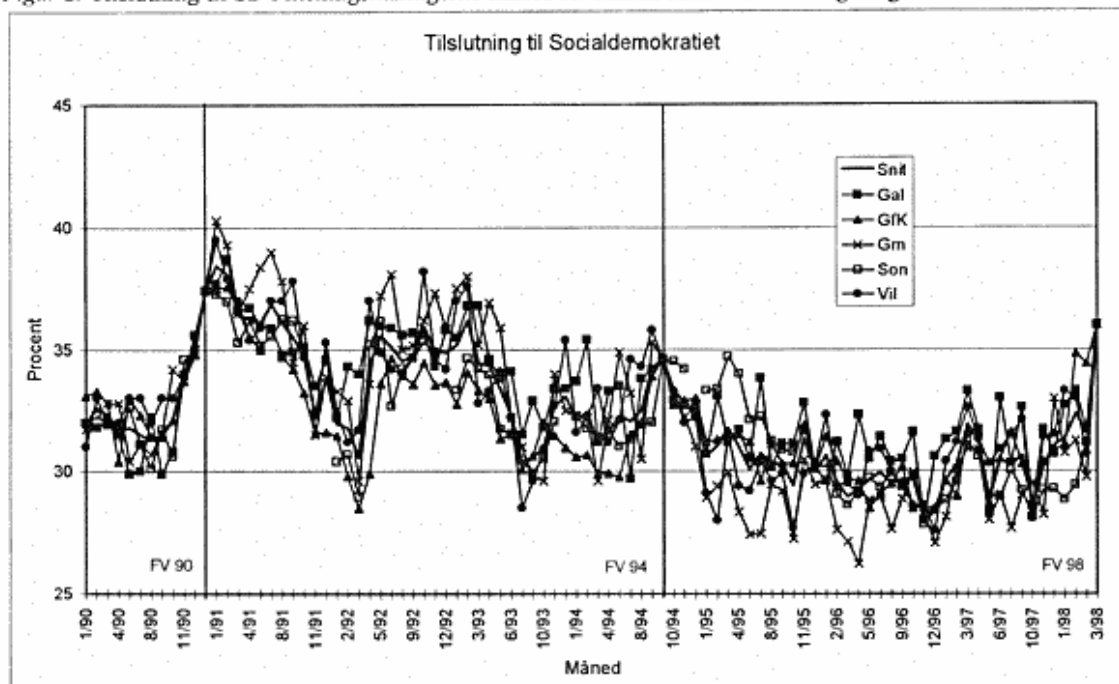
Med udgangspunkt i ovenstående diskussion vil jeg i det følgende foretage en nærmere analyse af overensstemmelsen mellem institutternes månedlige målinger i 1990'erne samt overensstemmelsen med de faktiske valgresultater. Der er tale om de samme institutter, som er omtalt ovenfor, bortset fra, at AIM ikke er medtaget, da de ikke offentliggjorde månedlige meningsmålinger i 1990'erne. Indledningsvis vil jeg demonstrere udviklingen i meningsmålingerne for de tre største partier, Socialdemokratiet, Venstre og Det konservative Folkeparti, samt for det nye Dansk Folkeparti.¹⁰

Figur 2 viser udviklingen i tilslutningen til Socialdemokratiet i meningsmålingerne, måned for måned (bortset fra feriemånedens juli, hvor der sjældent foretages meningsmålinger), både for det enkelte institut og for gennemsnittet over alle institutter.

I de måneder, hvor der blev afholdt et folketingsvalg, er der for hvert institut anført den sidste meningsmåling forud for selve valgdagen (dvs. ingen exit polls). For at man kan sammenligne med selve valgresultatet, er dette indføjet som en ekstra måned efter valgmanden, og alle kurverne er ført gennem valgresultatet. På denne måde kan man for eksempel se, at alle institutter forud for valget både i 1990 og 1998 undervurderede den endelige tilslutning til Socialdemokratiet. Derimod var der i 1994 målinger som lå over såvel som under SD's tilslutning ved valget.

Det fremgår af figur 2, at de forskellige institutter var nogenlunde enige om tilslutningen og udviklingen i tilslutningen til SD. I enkelte tilfælde er der dog betragtelige afvigelser. For eksempel havde Sonar i marts 1995 en høj tilslutning

Figur 2. Tilslutning til SD i meningsmålingerne måned for måned samt ved folketingsvalgene i 90'erne



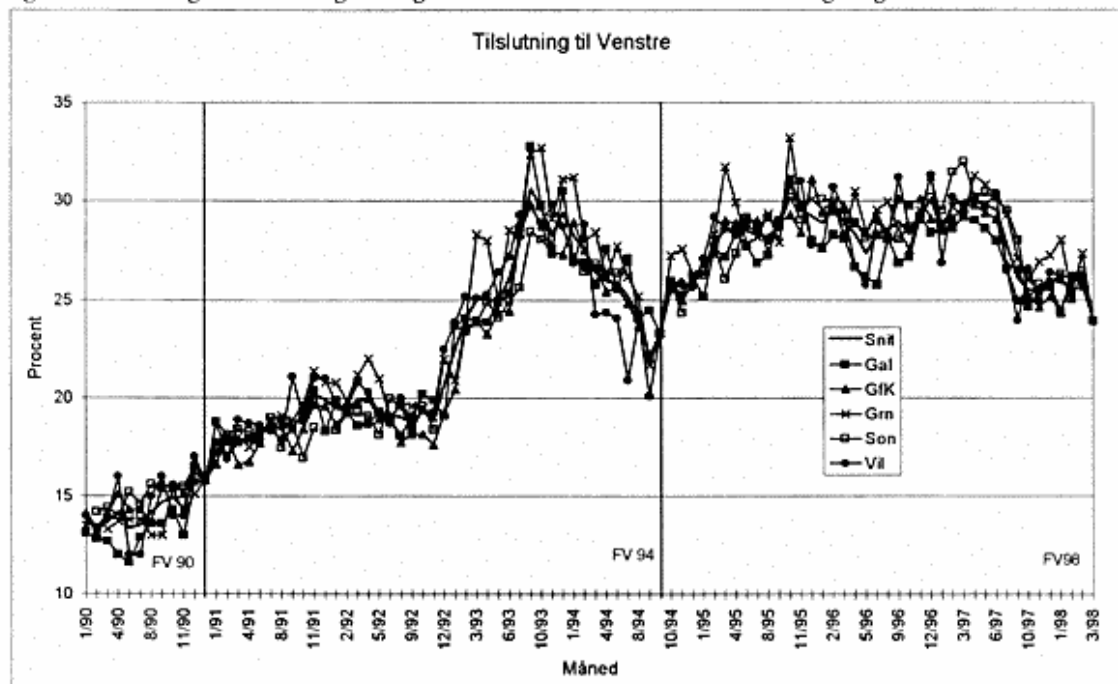
Faktiske resultater ved folketingsvalgene er indsat som en ekstra måned i måneden efter valgets afholdelse.

til SD på omkring 35 pct., men de øvrige institutter lå på 30-32 pct. Omvendt havde Greens særlig lav tilslutning til SD i april 1996 på omkring 26 pct. I øvrigt kan der i perioder være en tendens til, at det samme institut vedvarende har enten en særlig høj eller særlig lav tilslutning til SD. I perioden juni 1995 til november 1997 var der således en tendens til, at Gallup viste højere tilslutning til SD end de andre institutter i de fleste målinger, men Greens i den samme periode ofte viste lavere tilslutning til SD end de øvrige institutter. Det er altså ikke kun paneleterne, som i længere perioder kan afvige systematisk fra de andre institutter. Her kan forklaringen ikke være afhængige stikprøver, men afvigende procedurer for stikprøveudtagning, vejning, mv. Det er imidlertid ikke sådan, at Greens gennem hele 1990'erne havde lavest tilslutning til SD - fra 1991 til 1993 var det nærmest omvendt.

I øvrigt er alle institutter stort set enige om den generelle tendens: At SD efter en stærk fremgang ved 1990-valget havde en vigende tendens helt frem til årsskiftet 1996/97, hvorefter der var fremgang igen. Der var endvidere et dyk i tilslutningen til partiet i forbindelse med partiets lederkrise i foråret 1992 og igen i efteråret 1993 i forbindelse med megen kritik af statsminister Nyrup Rasmussens ledelsestil. Fremgangen for SD fra de sidste meningsmålinger til selve valgresultatet i marts 1998 ses længst til højre i figur 2. Det skal dog bemærkes at GfK lå tæt på valgresultatet både i februar og marts målingen.

Det fremgår af figur 3, at institutterne var mere enige om Venstres tilslutning end om Socialdemokratiets. Af systematiske afvigelsestendenser kan der dog peges på, at partiet gennemgående har klarer sig relativt godt i Greens' målinger, men relativt dårligt i Gallups.

Figur 3. Tilslutning til V i meningsmålingerne måned for måned samt ved folketingsvalgene i 90'erne



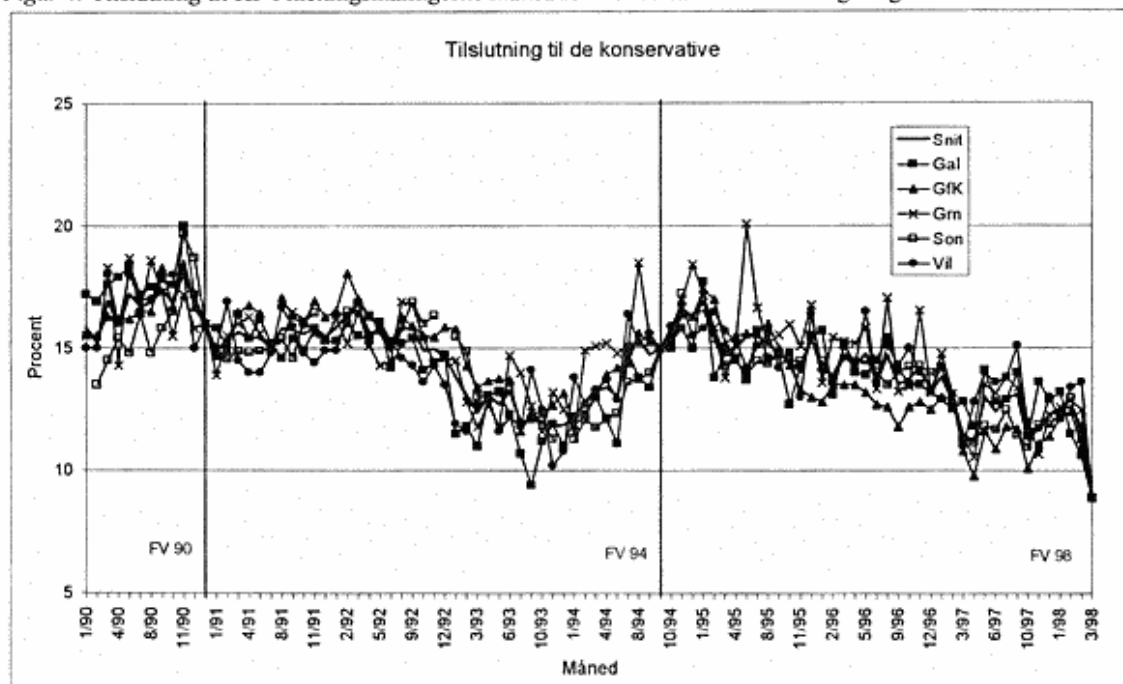
Faktiske resultater ved folketingsvalgene er indsat som en ekstra måned i måneden efter valgets afholdelse.

Udviklingen for partiet Venstre i figur 3 viser, at partiet under Uffe Ellemann-Jensens ledelse har klaret sig godt i 1990'erne, bortset fra ved valgene. Venstre var således tæt på Socialdemokratiets tilslutning i meningsmålingerne i en kortvarig periode i slutningen af 1993 og i en længere periode omkring 1996, men oplevede en markant tilbagegang både forud for valget i 1994 og forud for valget i 1998. Den markante fremgang for Venstre fra årsskiftet 1992/93 faldt sammen med krisen hos de konservative i forbindelse med afgørelsen i Tamil-sagen og den borgerlige Schlüter-regerings afgang, men også med den nye regerings ledelsesproblemer. Her profiterede Venstre sikkert af Ellemann-Jensens markante ledelsesstil (Borre og Andersen, 1997: 86-113). Som mulige forklaringer på tilbagegangen for V frem til 1994-valget kan fremhæves dels fremgangen for de konservative under Hans Engells ledelse, dels

den upopulære alliance mellem V, K og FRP forud for valget. Den markante tilbagegang for Venstre i løbet af 1997 falder sammen med fremgangen for Dansk Folkeparti, der var mere kompromisløs i flytninge- og indvandrerspørgsmålet end Venstre.

Ligesom det gælder for Venstre, fremgår det af figur 4, at institutterne var mere enige om de konservatives tilslutning end om Socialdemokratiets. Enkelte målinger fra Greens var dog stærkt afvigende i positiv retning. Det mest markante eksempel var maj 1995, hvor de konservative fik 20 pct. hos Greens, mens de øvrige institutter lå omkring 15 pct. I hele 1996 havde GfK konsekvent relativ lav tilslutning til KF. GfK målingerne i denne periode illustrerer i øvrigt ganske godt, at der i et panel er mindre kraftigt tilfældigt udsving fra måned til måned, end det gælder for institutter med statistisk uafhængige stikprøver.

Figur 4. Tilslutning til KF i meningsmålingerne måned for måned samt ved folketingsvalgene i 90'erne



Faktiske resultater ved folketingsvalgene er indsat som en ekstra måned i måneden efter valgets afholdelse.

Jeg har tidligere været inde på tilbagegangen for de konservative i forbindelse med afgivelsen af regeringsmagten i januar 1993 og fremgangen i 1994, efter ledelsesproblemerne var overvundet. Partiets mere selvstændige politik i forhold til Venstre i efteråret 1996 gav ikke fremgang, og Hans Engells sammenstød med en betonklods i februar 1997 kan aflæses ved et markant fald i den følgende måned. I forhold til marts 1997 kom troværdigheds- og ledelsesproblemerne under den nye partileder, Per Stig Møller, ikke til udtryk før på selve valgdagen den 11. marts 1998, hvor partiet kun fik 8.9 pct. af stemmerne, hvilket som nævnt måske kan forklares med den rolle, de ubeslutsomme vælgere spillede i meningsmålingerne.

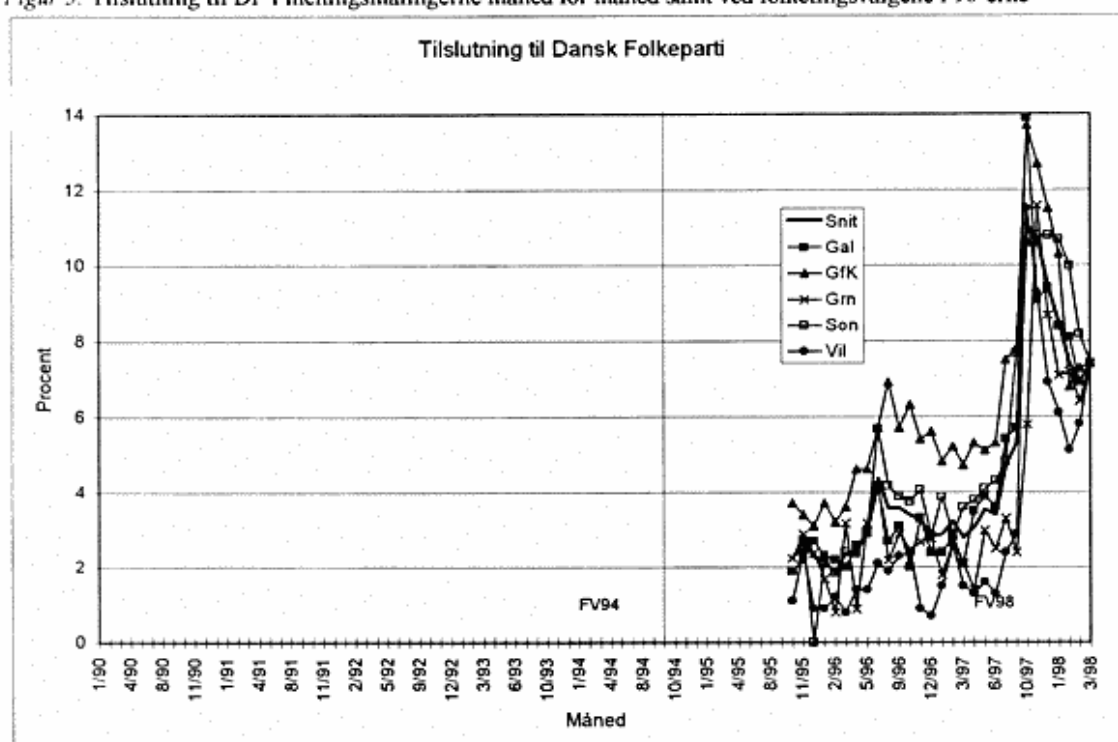
Tilslutningen til Dansk Folkeparti i figur 5 er taget med på grund af de tidligere omtalte systematiske forskelle mellem institutterne. Den højere tilslutning

til DF i en længere periode i de to paneler, GfK og Sonar (efter at partiet var blevet godkendt og dermed var kommet med i Sonars målinger), kan som nævnt forklares med spørgeformen, hvor partiets navn krydses af på en liste. DF har gennemgående haft lavest tilslutning i Vilstrup-målingerne.

Den stærke fremgang for DF i efteråret 1997 hænger uden tvivl sammen med indvandrer- og flygtningepolitikens dominerende position på dagsordenen i denne periode; tilbagegangen i foråret 1998 kan tilsvarende forklares med inddæmningspolitikken i det samme spørgsmål fra de øvrige partier.

Figurene 2-5, der handler om enkelte partier, er bragt for at illustrere nogle karakteristiske former for overensstemmelser eller afvigelser mellem de enkelte institutter, som har betydning for vurdering af meningsmålingernes troværdighed. Det institut, som skiller sig mest ud, er Gre-

Figur 5. Tilslutning til DF i meningsmålingerne måned for måned samt ved folketingsvalgene i 90'erne



Faktiske resultater ved folketingsvalgene er indsat som en ekstra måned i måneden efter valgets afholdelse.

ens, som har flest markante afvigelser. Troværdigheden er endvidere blevet belyst ved at se på, om kurvernes samlede forløb giver mening i forhold til vigtige begivenheder i dansk politisk historie. Selv om mine fortolkninger i høj grad har karakter af efterrationaliseringer, mener jeg alligevel, at de er så plausible, at de styrker tiltroen til meningsmålingerne. I resten af denne artikel vil jeg søge at vurdere meningsmålingernes troværdighed ud fra analyser, som går på tværs af samtlige partier.

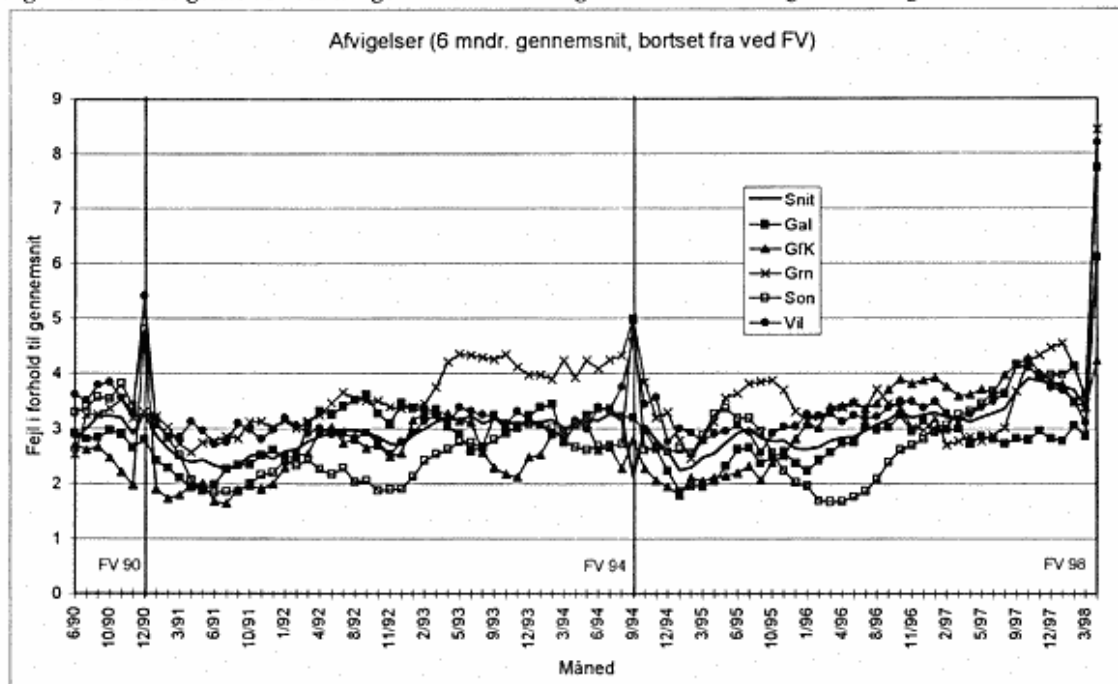
De enkelte institutters afvigelser

I forbindelse med valgkampen i 1998 blev overensstemmelsen mellem de enkelte målinger belyst ved parvise sammenligninger. En sådan analyse vil være uoverskuelig at gennemføre for alle de månedlige målinger i 1990'erne. I stedet er valgt at se på det enkelte instituts samlede af-

vigelse i hver måned i forhold til gennemsnittet for alle institutter i den samme måned. Afvigelsesmålet for det enkelte institut viser, hvor mange procent, der skal flyttes i instituttets måling, for at få gennemsnittet. Hvis man tegner kurver for de enkelte institutters afvigelse fra gennemsnittet måned for måned, får man et meget kaotisk billede, da tilfældigheder her har et stort spillerum. For at få et mere stabilt billede har jeg derfor for hver måned valgt at udregne gennemsnittet af de enkelte institutters afvigelser (fra de enkelte måneders gennemsnit) over de sidste seks måneder, inklusive den aktuelle måned. Udviklingen i disse *glidende gennemsnit* er vist i figur 6.

I de måneder, hvor der blev afholdt folketingsvalg, er der for hvert institut på normal vis vist det glidende gennemsnit over de sidste seks måneder. Endvidere er der herefter, som en ekstra måned, indsat afvigelsen mellem det enkelte insti-

Figur 6. Glidende gennemsnit af afvigelser fra månedens gennemsnit samt afvigelse fra valresultatet



For valgmanden er både vist de glidende gennemsnit samt afvigelse fra det faktiske valgresultat (indsat som en ekstra måned).

tuts sidste måling inden valgdagen og selve valgresultatet (altså udpluk af de samme tal, som for 1998-valget blev vist i tredjesidste kolonne i tabel 1).

Lad os for et øjeblik antage, at alle meningsmålingerne var ulastelige, forstået således, at hvert institut måned for måned viste den faktiske partitilslutning i en simpel tilfældig stikprøve på 1.000 vælgere blandt alle de vælgere, der havde besluttet deres partivalg i den pågældende måned. I så fald skulle kurverne for de glidende gennemsnit i figur 6 udvise svingninger omkring et gennemsnit på ca. tre pct., som normalt ikke afviger mere end højst én pct. til hver side.¹¹

Imponerende nok er det faktisk det mønster, man ser i figur 6. Det er endda således, at især panelerne i perioder (GfK i 1991 og omkring årsskiftet 1994/95 samt Sonar i 1992 og 1996) har meget lave afvigelser fra gennemsnittet på omkring to pct. På den anden side ligger GfK

panelet i 1996 på en stabil måde ret tæt på den øvre grænse. Men dette mønster er netop, hvad man skulle forvente for panelerne, som har *afhængige* stikprøver. I øvrigt skyldes den store afvigelse for GfK overrepræsentationen af DF, der som nævnt *kunne* være en fejl hos de øvrige institutter. Blandt institutter med uafhængige stikprøver er det især Greens, som i perioder skiller sig ud med relativt store afvigelser over fire pct. Gallup ligger ofte under det forventede afvigelsesgennemsnit på tre pct., mens Vilstrup ofte ligger lidt over. For hele perioden fra 1990 til 1998 er gennemsnittet af samtlige afvigelser for de enkelte institutter i prioriteret rækkefølge som følger: Sonar (2.78), GfK (2.80), Gallup (2.81), Vilstrup (3.19) og Greens (3.45). Dette betyder altså, at Sonar med en gennemsnitlig afvigelse på 2.78 pct. gennemgående har ligget tættest på det fælles månedlige gennemsnit og i den forstand klarer sig bedst.

GfK og Gallup har imidlertid næsten samme gennemsnit som Sonar. Derimod har Vilstrup en noget større gennemsnitlig på 3.19 pct. og Greens den største afvigelse på 3.45 pct. Den nævnte prioriterede rækkefølge betyder ikke nødvendigvis, at de to paneler og Gallup er mere troværdige end Vilstrup og Greens, da et institut har mulighed for at ændre sine procedurer, hvis det kommer for langt væk fra de øvrige institutter. Specielt de to paneler har gode muligheder for dette ved hjælp af udskiftning af deltagere, som derefter fastholdes gennem længere tid i panelet.

Figur 6 giver som nævnt også mulighed for at vurdere, hvor meget de sidste meningsmålinger inden valgdagen afviger fra selve valgresultat. Hvis vi igen forestiller os den helt ulastelige situation, at meningsmålingerne var simpel tilfældige stikprøver på 1.000 vælgere blandt alle de vælgere, som faktisk afgiver deres stemme på selve valgdagen (og vælgerne i øvrigt ærligt fortalte, hvilket parti de ville stemme på), skulle de enkelte institutters afvigelse på de lodrette streger for hvert folketingsvalg ligge omkring 3.4 pct. med en afvigelse på ca. to pct. til hver side.¹² Her kniber det dog noget. Afvigelserne er gennemgående for store. Da tallene kan være vanskelige at aflæse i figur 6, er de også vist i tabel 4.

Tabel 4 viser, at afvigelserne gennemgående var større end de ca. 3.4 pct., man skulle forvente ud fra fornævnte betragt-

ning. Endvidere var afvigelserne som tidligere nævnt større i 1998 end ved de tidligere valg. Blandt institutterne lå Gallup i 1990 tættest på valgresultatet med en afvigelse på 2.8 pct., mens GfK både i 1994 og 1998 var tættest på valgresultatet med afvigelser på henholdsvis 2.8 pct. og 4.2 pct. GfK's bedre resultat i 1998 hænger nok, som nævnt, sammen med, at målingen som den eneste blev foretaget aftenen før valget. Udregner man gennemsnitlig afvigelse på tværs af valgene, har GfK gennemgående klaret sig bedst (3.9 pct.), med Gallup lige i hælene (4.0 pct.). Der er derefter et spring til Greens, Sonar og Vilstrup med henholdsvis 5.6, 5.9 og 6.1 pct. gennemsnitlig afvigelse fra det faktiske valgresultat. Det bør dog nævnes, at der er ret få tal i tabel 4, så rene tilfældigheder kan have spillet ind på, hvor godt de enkelte institutter har klaret sig ved de tre seneste folketingsvalg. Endelig er der ikke taget højde for længden af perioden fra den sidste måling til selve valgdagen, hvor der kan ske reelle ændringer.

Det viser sig i øvrigt, at hvis man ser på gennemsnittet over samtlige institutters sidste måling inden valget, kaldet "Snit" i tabel 4, får man det allerbedste resultat med en afvigelse fra det endelige valgresultat på kun 3.8 pct. Ved at danne et fælles gennemsnit over samtlige målinger har man altså tilsyneladende ikke "blandet god vin med dårlig vin" i et sådant omfang, at man får et slutprodukt,

Tabel 4. Sidste målings afvigelse fra faktisk valgresultat

Valg	Gallup	GfK	Grns	Sonar	Vilstr	Snit
1990	2.8	4.7	3.3	4.8	5.4	3.3
1994	3.2	2.8	5.0	5.0	4.6	2.1
1998	6.1	4.2	8.5	7.8	8.2	6.1
Gens.	4.0	3.9	5.6	5.9	6.1	3.8

Exit polls undtaget

som ikke kan konkurrere med de enkelte mærker. Specielt var forudsigelsen ved hjælp af Snit særdeles god i 1994 med en afvigelse på kun 2.1 pct. fra valgresultatet.

De enkelte institutters stabilitet

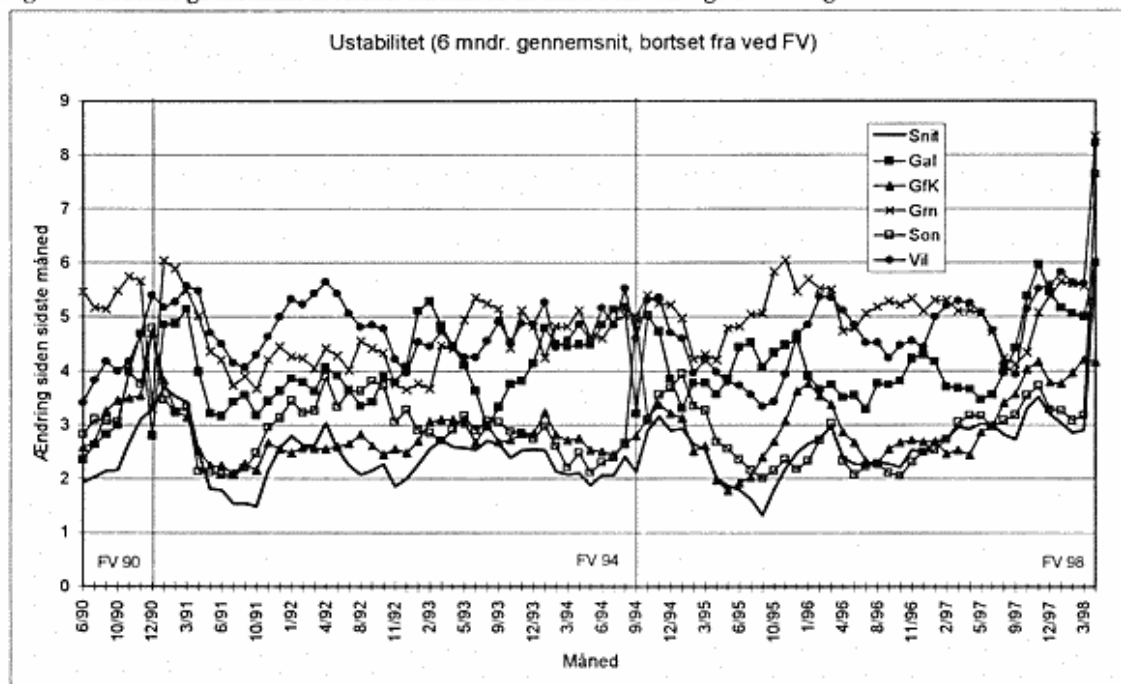
Som en sidste "kvalitetskontrol" af de enkelte institutter samt blandingsproduktet Snit vil jeg undersøge graden af stabilitet over tid. Synspunktet er her, at hvis et institut har meget store udsving fra måned til måned, kan det være udtryk for manglende troværdighed. Af denne grund er for hver måned beregnet instituttets afvigelse i forhold til forrige måned ved det tidligere benyttede mål for forskellen mellem to stikprøver, omtalt i note 9 (hvor mange procent der skal flyttes i den ene stikprøves partifordeling for at få den anden stikprøves fordeling). For at undgå et for kaotisk billede er der ligesom i sid-

ste afsnit udregnet seks måneders glidende gennemsnit af denne forskel, vist i figur 7.

Ligesom i figur 6 er der efter hver valg-måned indsat en ekstra måned, som viser forskellen mellem sidste måling inden valget og selve valgresultatet (altså det samme som i figur 6).

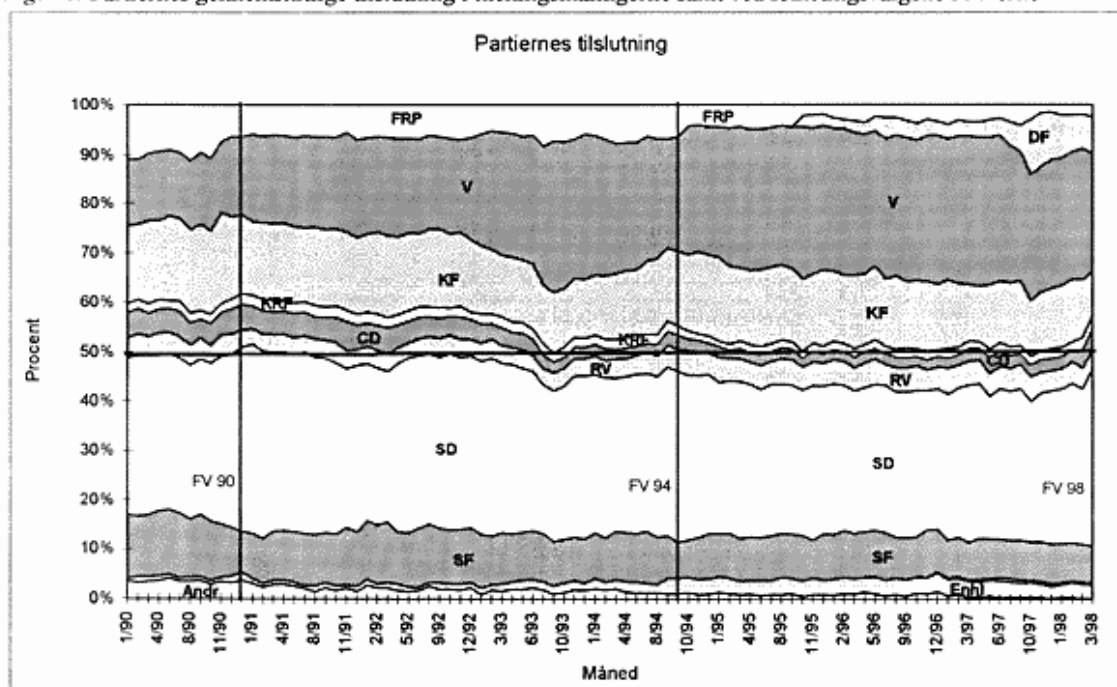
Hvis vi opstiller den ekstreme antagelse, at partiernes reelle tilslutning (blandt alle vælgere, der har besluttet sig) er uændret fra måling til måling, skulle kurverne for de glidende gennemsnit i figur 7 udvise svingninger omkring et gennemsnit på ca. fem og ikke afvige mere end en enkelt pct. til hver side.¹³ Det viser det sig imidlertid, at dette mønster kun gælder for Greens og Vilstrup. For de øvrige målinger gælder det i lighed med, hvad vi fandt for valgkampen i 1998, at forskellene fra måned til måned er mindre, end man skulle forvente ud fra antagelsen om uafhængige, simpel tilfældigt ud-

Figur 7. Glidende gennemsnit af forskel fra måned til måned samt afvigelse fra valgresultatet



For valgmåned er både vist de glidende gennemsnit samt afvigelse fra det faktiske valgresultat (samme som i figur 6 og indsat som en ekstra måned).

Figur 8. Partiernes gennemsnitlige tilslutning i meningsmålingerne samt ved folketingsvalgene i 90'erne



Faktiske resultater ved folketingsvalgene er indsat som en ekstra måned i måneden efter valgets afholdelse.

tagne stikprover. Det er ikke så mærkeligt, at dette er tilfældet for de to paneler, GfK og Sonar, der jo er afhængige stikprover, eller for Snit, der er et gennemsnit af flere målinger. Det mest bemærkelsesværdige er nok, at Gallups målinger har markant mindre forskelle, end man skulle forvente ud fra uafhængige, simpel tilfældige stikprover. Gallup-institutet har bekræftet, at dette især skyldes efterstratificering ud fra tidligere partivalg: den politiske vejning. Den kraftige vækst i ustabiliteten hos Gallup i efteråret 1997 skyldes den kraftige fremgang for DF.

Det er også interessant at konstatere, at selv om alle fem institutters målinger indgår i Snit, har dette månedlige gennemsnit højere stabilitet end selv de to paneler. Dette må skyldes, at de tre uafhængige stikprøveundersøgelser, Gallup, Greens og Vilstrup, også er indbyrdes uafhængige, og at deres noget større til-

fældige udsving dermed ophæver hinanden. Den prioriterede rækkefølge efter stabilitet for hele perioden 1990-98 er: Snit (2.50), Sonar (2.88), GfK (2.90), Gallup (4.04), Vilstrup (4.67) og Greens (4.86). Når det tages i betragtning, at der også sker reelle ændringer over tid, er det imponerende, at samtlige institutter mindst har den stabilitet, som ville gælde for simpel tilfældigt udtagne stikprover fra en stationær population.

Partiernes tilslutning i 1990'erne

Da det månedlige gennemsnit både klarede sig bedst ved forudsigelse af valgene i 1990'erne og samtidig udviser den største stabilitet fra måned til måned, bliver det mit bedste bud på et troværdigt mål for partiernes tilslutning.¹⁴ I figur 8 er vist den månedlige, gennemsnitlige tilslutning til samtlige partier ordnet efter en venstre-højre inddeling.

I valgmanden er vist gennemsnittet for de sidste målinger inden valget, og selve valgresultatet er herefter indsat som en ekstra måned. Det bliver dermed muligt at sammenligne de sidste meningsmålinger med selve valgresultatet. For eksempel fremgår det længst til højre i figur 8, at den samlede tilslutning til de socialistiske partier i de sidste meningsmålinger var mindre end ved selve valget i marts 1998.

Figur 8 viser et relativt stabilt billede af udviklingen i tilslutningen til partierne opdelt i en socialistisk og borgerlig gruppe. Derimod kan der inden for hver gruppe være noget større forskydninger, for eksempel i forbindelse med Tamil-sagen og Schlüter-regeringens afgang, hvor Venstre gjorde et stærkt indhug i de konservatives tilslutning. Samtidig var der dog også en noget mindre tilbagegang for Socialdemokratiet. En lignende situation opstod i forbindelse med den stærke fremgang for Dansk Folkeparti i efteråret 1997, som direkte eller indirekte gjorde indhug i både Venstres og de konservatives tilslutning, men kun i mindre grad i Socialdemokratiets tilslutning. I begge situationer var den samlede tilslutning til venstrefløjen til venstre for Socialdemokratiet tilsyneladende upåvirket af disse bølger.

Det er i øvrigt et særligt kendetegn ved perioden efter 1990-valget, at det traditionelle mønster med, at en fremgang for Socialdemokratiet betyder en tilbagegang for de øvrige socialistiske partier og vice versa, ophører. Fra 1990 startede en jævn tilbagegang for Socialdemokratiet, som ikke blev opvejet af en tilsvarende fremgang for venstrefløjen. Dette viser sig i figur 8 ved, at den samlede tilslutning til de socialistiske partier gradvis synker under den vandrette linje, som angiver 50 pct. tilslutning. Efter 1994-valget var det

sågar således, at heller ikke de radikales bidrag kunne bringe den samlede tilslutning i meningsmålingerne op på 50 pct., hvilket styrkede CD's og måske også de konservatives potentielle muligheder for at for at få en afgørende rolle i dansk politik. Det overraskende resultat ved 1998 valget gav imidlertid lige akkurat igen regeringens parlamentariske grundlag føling med de 50 pct.

Konklusion

Var det helt utroværdigt, da meningsmålingerne under valgkampen op til folketingsvalget i 1998 først viste, at regeringen havde fået en flyvende start, dernæst at vi nok alligevel skulle have en borgerlig regering, for endelig, i sidste øjeblik, at fortælle den undrende offentlighed, at partierne i regeringens parlamentariske grundlag havde genvundet alt det, de havde tabt siden 1994?

Jeg tror det ikke. For det første bør man tage meningsmålingsinstitutterne på ordet, når de gør opmærksom på, at meningsmålinger ikke er prognoser for det endelige valgresultat, men øjeblikkelige målinger af, hvad vælgerne mener. Men her til vil jeg også tilføje: Ja, og vel at mærke øjeblikkelige målinger blandt de vælgere, som har besluttet sig, og i øvrigt ikke er modstandere af at lade sig interviewe eller meget vanskelige at komme i kontakt med.

Til gengæld må man sige, at til trods for meget forskellige metodiske principper er de forskellige institutter meget enige om, hvad denne særlige population af vælgere mener. Der var således i den sidste halvdel af valgkampen mindre forskelle mellem de enkelte målinger fra de forskellige institutter, end man skulle forvente ud fra den antagelse, at alle målingerne var simpel tilfældige stikprøver fra

den samme population. Denne træfsikkerhed kan især forklares med, at alle institutter anvender efterstratificering ud fra tidligere partivalg: den politiske vejning. Dette kan også både forklare den store overensstemmelse mellem de forskellige institutter og stabiliteten for de enkelte institutter fra måned til måned i løbet af 1990'erne. Det skal dog nævnes, at de to institutter, som anvender de mindst tilfredsstillende stikprøvemethoder, Vilstrup og Greens, både har den laveste overensstemmelse med de øvrige institutter og den laveste stabilitet fra måned til måned. Det gælder endvidere for de to institutter, der anvender paneler, GfK og Sonar, at de har lettere ved at justere afvigelser i forhold til de andre institutter, samt at de på grund af faste svarpersoner i høj grad undgår tilfældige udsving fra måned til måned.

Ser man specielt på, hvorfor Gallups exit polls kom meget tæt på det endelige valgresultat både i 1994 og 1998, må dette blandt andet skyldes, at der på selve valgdagen ikke er noget problem med vælgere, som ikke har besluttet sig. Man må dog samtidig antage, at der i lighed med andre meningsmålinger fortsat er en del nægtede samt folk, som på anden måde er svære at komme i kontakt med. Men de sidstnævnte problemer kan der åbenbart kompenseres for, blandt andet ved den politiske vejning.

Dermed har vi også fået indkredset, hvad der kan være problemet med de øvrige målinger: de ubeslutsomme vælgere. Ganske vist betød dette problem ikke så meget ved valgene fra 1984 til 1994, men ved 1998-valget adskilte denne gruppe sig åbenbart meget fra de øvrige vælgere, muligvis fordi en del vælgere var kommet i tvivl om de borgerlige partiers økonomiske ansvarlighed og derfor endte med at stemme socialdemokratisk. Det for-

hold, at Gallup og GfK i 1990'erne i de seneste målinger kom tættest på det endelige resultat, tyder endvidere på, at disse to institutter bedre end Sonar, Vilstrup og Greens kunne kompensere for problemet med de ubeslutsomme vælgere. Kvaliteten af de tre sidstnævnte institutters målinger er imidlertid så god, at gennemsnittet over samtlige fem institutter kom lige så tæt som de to førstnævnte institutter på det endelige valgresultat, samt var mere stabilt end noget enkelt institut. De gennemsnitlige målinger giver endvidere et plausibelt indtryk af udviklingen i partiernes tilslutning i 1990'erne, herunder afmatningen i tilslutningen til de socialistiske partier indtil feberredningen i 1998 samt effekten af diverse lederkriser i partierne.

Noter

1. Ved en exit poll forstår man normalt en undersøgelse, hvor vælgerne interviewes, når de forlader særligt udvalgte valgsteder. Det, man kaldte exit polls i 1994 og 1998, var imidlertid undersøgelser, hvor man foretog telefoninterviews med vælgerne på valgdagen, efter de havde stemt. Dette princip overholdt Gallup konsekvent, mens AIM i 1994 også inkluderede interviews fra dagen før. I 1998 konstruerede AIM en slags exit poll, ved at man aftalte med de svarpersoner fra undersøgelsen 7-9/3, som var i tvivl, at man kunne ringe til dem på selve valgdagen den 11/3. AIM's exit poll forelå kl. 19.30, mens Gallups endelige exit poll forelå kl. 20.20.
2. I en given stikprøve betegnes andelen, som vælger parti nr. i med P_i . Den tilsvarende andel i populationen, dvs. det faktiske valgresultat, betegnes med p_i . Over alle n partier beregnes et indeks for absolut afvigelse fra det faktiske valgresultat ved

$$A = (\sum |P_i - p_i|) / 2$$

dvs. halvdelen af summen af forskellene mellem stikprøvens og populationens partiandele uanset fortegn. Dette svarer til den andel af

stikprøven, der skulle flyttes mellem partierne for at give samme resultat som i populationen. Ved fuldstændig overensstemmelse mellem population og stikprøve antager indekset værdien 0, mens værdien er 1 (100 pct.) ved fuldstændig uoverensstemmelse. I tabel 1 betegnes A ved "Afv. f. FV" (Afvigelse fra folketingsvalget) i tredjesidste kolonne.

- For et givet parti nr. i betegnes differencen mellem andelen i en stikprøve P_i og den tilsvarende andel og populationen p_i ved

$$A_i = P_i - p_i$$

I en simpel tilfældig stikprøve er denne difference tilnærmet normalfordelt med forventning lig med 0 og varians lig med $p_i(1-p_i)/N$, hvor N er stikprøvens størrelse (antal der har oplyst partivalg). Forventningen på den absolutte værdi af A_i er

$$E(|A_i|) = \text{sqrt}[2p_i(1-p_i)/(\pi N)]$$

da forventningen på den absolutte værdi af en standardiseret normalfordelt variabel er $\text{sqrt}(2/\pi)$ (Johnson *et al.*, 1994: 91). Den forventede værdi for den absolutte afvigelse mellem en population og en stikprøve, defineret i note 2, bliver hermed

$$E(A) = E[(\sum |A_i|)/2] \\ = \sum \text{sqrt}[p_i(1-p_i)]/\text{sqrt}(2\pi N)$$

der er beregnet for de enkelte meningsmålinger og vist i den næstsidste kolonne i tabel 1. Samplingfordelingen for A er beregnet ved computersimulation (1000 simulationer af stikprøver på 1000 personer med valgresultatet som population), og det viser sig, at fordelingen med denne stikprøvestørrelse er en smule højreskæv med en standardafvigelse på 0.95 pct. og et gennemsnit på 3.39, hvilket er tæt på den forventede værdi 3.37. Som et mål for den faktiske overensstemmelse mellem en stikprøve og en population i procent af den teoretisk forventede overensstemmelse beregnes

$$\text{Fit}(A) = 100(1-A)/[1-E(A)]$$

der er vist i den sidste kolonne i tabel 1.

- Kontaktpersoner i de enkelte institutter var fra A.C. Nielsen, AIM: Jens Rom Jensen, Gallup A/S: Troels Palner, GfK Danmark A/S: Marcus Schmidt, Greens Analyseinstitut: Jens Nicolaisen, Sonar: Gunnar Sehested Larsen samt Vilstrup Research: Kasper Vilstrup. Alle har været imødekommende med redegørelse for metoder bortset fra meget de-

taljerede procedurer, der har karakter af forretningshemmeligheder. Alle har endvidere haft mulighed for at kommentere artiklen, hvilket gav anledning til nogle mindre rettelser. Dette betyder dog ikke, at alle institutterne er enige i mine konklusioner. Greens har oplyst, at man som en konsekvens af kritikken i denne artikel har ændret sin procedure for udvælgelse af personer i husstanden til den AIM og Gallup bruger.

- GfK gennemførte dog telefoninterviews med paneldeltagerne 19-20/2 samt 10/3 1998.
- Vilstrup oplyste i Politiken den 20/7 1997, at man fra juli 1997 havde ændret procedure fra den sædvanlige regel om, at partinavn (Fremskridtspartiet eller Dansk Folkeparti) har forrang frem for personnavn (Pia Kjærsgaard) til den modsatte regel.
- Gallup bruger endvidere datoen for fødselsdagen til at fravælge nogle en-persons husstande, så disse ikke bliver overrepræsenterede.
- Ændringen i forhold til den tidligere viderebragte oplysning om, at Gallup i 1994 havde en gennemførelse på godt 80 pct. (Thomsen, 1995: 73), forklares af Gallup med, at man nu normalt gennemfører væsentligt længere interviews samt en almindelig træthedseffekt blandt respondenter. Derimod fremgår det af tabel 1, at man ved det væsentligt kortere interview i forbindelse med exit poll'en i 1998 kunne opnå en markant højere gennemførelse på 86 pct.
- I to stikprøver fra den samme population betegnes andelen, som stemmer på parti nr. i, i den første stikprøve ved P_{1i} og i den anden stikprøve ved P_{2i} . Den absolutte forskel mellem de to stikprøver beregnes ved

$$D = (\sum |P_{1i} - P_{2i}|)/2 = (\sum |D_i|)/2$$

dvs. halvdelen af summen af forskellene mellem de to stikprøvers andele uanset fortegn. Dette svarer til den andel, der skal flyttes i den ene stikprøve, for at få samme partifordeling som i den anden stikprøve. Hvis de to stikprøver er udtrukket simpel tilfældigt, er differencen D for parti nr. i tilnærmet normalfordelt med forventning lig med 0 og varians lig med $p_i(1-p_i)(1/N_1 + 1/N_2)$, hvor N_1 og N_2 er de to stikprøvers størrelser (antal der har oplyst partivalg i hver stikprøve) og p_i er partiets andel i populationen. Forventningen på den absolutte værdi af D_i er

$$E(|D_i|) = \text{sqrt}[2p_i(1-p_i)(1/N_1 + 1/N_2)/\pi]$$

da forventningen på den absolutte værdi af en standardiseret normalfordelt variabel, som nævnt i note 3, er $\sqrt{2/\pi}$. Den forventede værdi af den absolutte forskel mellem de to stikprøver bliver hermed

$$E(D) = E[(\sum|D_i|)/2] = \sum \sqrt{p_i(1-p_i)} \sqrt{(1/N_1 + 1/N_2)/2\pi}.$$

Samplingfordelingen for D er beregnet ved computersimulation (1000 simulationer af 2 stikprøver på hver 1000 personer med valgresultatet som population), og det viser sig, at fordelingen med disse stikprøvestørrelser er en smule højreskæv med en standardafvigelse på 1.29 pct. og et gennemsnit på 4.66, hvilket er tæt på den forventede værdi 4.77. Som et mål for den faktiske overensstemmelse mellem de to stikprøver i procent af den teoretisk forventede overensstemmelse beregnes

$$\text{Fit}(D) = 100(1-D)/[1-E(D)]$$

der er vist for alle mulige parvise sammenligninger af valgkampens meningsmålinger i tabel 3.

10. De viste tal er ikke altid præcis de samme, som i sin tid blev offentliggjort. Det hænger sammen med, at nogle institutter kun offentliggør tilslutningen til opstillingsberettigede partier (GfK indtil FV94, Greens og Sonar). Efter at Jacob Haugaard blev indvalgt i Folketinget uden for partierne (og i den periode hvor Kristeligt Folkeparti og Dansk Folkeparti ikke var opstillingsberettigede) kan man derfor ikke foretage en meningsfuld sammenligning af de forskellige institutter, med mindre man foretager en justering af de målinger, som mangler de pågældende oplysninger. Endvidere er der det problem, at Gallup grupperede tilslutningen til alle de partier, som fik under 2 procents tilslutning. I praksis foregår justeringen ved, at man indsætter en manglende oplysning ud fra gennemsnittet af de af månedens målinger, som har oplysningen med. Denne proces er ikke helt simpel, da en given måling undertiden både kan modtage manglende oplysninger fra andre målinger samt bidrage til andre målingers manglende oplysninger. I praksis optimeres justeringen ved en iterativ procedure. Justeringen betyder i øvrigt, at man ved undersøgelsen af den indbyrdes overensstemmelse mellem målingerne giver en lille fordel til de målinger, som er blevet justeret, fordi de indeholder gennemsnitstal.

11. Da den enkelte måling indgår i det fælles gennemsnit, er den forventede værdi af afvigelsen \hat{A} af det enkelte institut i forhold til det fælles gennemsnit lidt mindre end forventningen af afvigelsen A i forhold til populationen, jf. note 3. Mere præcis er

$$E(\hat{A}) = \sqrt{(k-1)/k} E(A) = \sqrt{4/5} \leftrightarrow 3.37 = 3.01$$

hvor k er antal institutter, som indgår i det fælles gennemsnit. Den anførte talværdi gælder for k = 5 og stikprøver på 1000 vælgere. Anvender vi simulationsresultatet fra note 3 bliver standardafvigelsen på et gennemsnit over 6 måneder = $\sqrt{4/(5 \leftrightarrow 6)} \leftrightarrow 0.95 = 0.35$. Strengt taget er 95 pct. konfidensintervallet dermed ca. 3.0 ± 0.7 , men da antal vælgere med oplyst partivalg ofte er lidt mindre, forekommer 4.0 at være en rimelig øvre grænse.

12. Jf. note 3.
13. Jf. note 9 er den forventede værdi på forskelsmålet 4.8 med en standardafvigelse på 1.29 pct. Dermed bliver standardafvigelsen på et gennemsnit over 6 måneder = $1.29/\sqrt{6} = 0.53$. Med den samme reservation som i note 11 forekommer 6.0 at være en rimelig øvre grænse.
14. Dermed være ikke sagt, at det ikke ville være muligt at konstruere et endnu bedre mål for partiernes månedlige tilslutning. For eksempel kunne man overveje at lade de forskellige målinger indgå i gennemsnittet med forskellig vægt, afhængig af deres kvalitet. Endvidere kunne man ved hjælp af *a priori* antagelser indbygge et forsigtigt tendens-skøn. Jeg har haft gode erfaringer med at gøre noget lignende i forbindelse med DR TV's valgaftensprognoser.

Litteratur

- Borre, Ole og Jørgen Goul Andersen (1997). *Voting and Political Attitudes in Denmark: A Study of the 1994 Election*, Aarhus: Aarhus University Press.
- Cochran, William G. (1977). *Sampling Techniques*, 3rd edition, New York: Wiley.
- Elklit, Jørgen (1989). "Opinionsmålinger og valgresultater", i Jørgen Elklit og Ole Tonsgaard, *To folketingsvalg: Vælgerholdninger og vælgeradfærd i 1987 og 1988*, Aarhus:

da forventningen på den absolutte værdi af en standardiseret normalfordelt variabel, som nævnt i note 3, er $\sqrt{2/\pi}$. Den forventede værdi af den absolutte forskel mellem de to stikprøver bliver hermed

$$E(D) = E[(\sum|D_i|)/2] = \sum \sqrt{p_i(1-p_i)} \sqrt{(1/N_1 + 1/N_2)/2\pi}.$$

Samplingfordelingen for D er beregnet ved computersimulation (1000 simulationer af 2 stikprøver på hver 1000 personer med valgresultatet som population), og det viser sig, at fordelingen med disse stikprøvestørrelser er en smule højreskæv med en standardafvigelse på 1.29 pct. og et gennemsnit på 4.66, hvilket er tæt på den forventede værdi 4.77. Som et mål for den faktiske overensstemmelse mellem de to stikprøver i procent af den teoretisk forventede overensstemmelse beregnes

$$\text{Fit}(D) = 100(1-D)/[1-E(D)]$$

der er vist for alle mulige parvise sammenligninger af valgkampens meningsmålinger i tabel 3.

10. De viste tal er ikke altid præcis de samme, som i sin tid blev offentliggjort. Det hænger sammen med, at nogle institutter kun offentliggør tilslutningen til opstillingsberettigede partier (GfK indtil FV94, Greens og Sonar). Efter at Jacob Haugaard blev indvalgt i Folketinget uden for partierne (og i den periode hvor Kristeligt Folkeparti og Dansk Folkeparti ikke var opstillingsberettigede) kan man derfor ikke foretage en meningsfuld sammenligning af de forskellige institutter, med mindre man foretager en justering af de målinger, som mangler de pågældende oplysninger. Endvidere er der det problem, at Gallup grupperede tilslutningen til alle de partier, som fik under 2 procents tilslutning. I praksis foregår justeringen ved, at man indsætter en manglende oplysning ud fra gennemsnittet af de af månedens målinger, som har oplysningen med. Denne proces er ikke helt simpel, da en given måling undertiden både kan modtage manglende oplysninger fra andre målinger samt bidrage til andre målingers manglende oplysninger. I praksis optimeres justeringen ved en iterativ procedure. Justeringen betyder i øvrigt, at man ved undersøgelsen af den indbyrdes overensstemmelse mellem målingerne giver en lille fordel til de målinger, som er blevet justeret, fordi de indeholder gennemsnitstal.

11. Da den enkelte måling indgår i det fælles gennemsnit, er den forventede værdi af afvigelsen \hat{A} af det enkelte institut i forhold til det fælles gennemsnit lidt mindre end forventningen af afvigelsen A i forhold til populationen, jf. note 3. Mere præcis er

$$E(\hat{A}) = \sqrt{(k-1)/k} E(A) = \sqrt{4/5} \leftrightarrow 3.37 = 3.01$$

hvor k er antal institutter, som indgår i det fælles gennemsnit. Den anførte talværdi gælder for k = 5 og stikprøver på 1000 vælgere. Anvender vi simulationsresultatet fra note 3 bliver standardafvigelsen på et gennemsnit over 6 måneder = $\sqrt{4/(5 \leftrightarrow 6)} \leftrightarrow 0.95 = 0.35$. Strengt taget er 95 pct. konfidensintervallet dermed ca. 3.0 ± 0.7 , men da antal vælgere med oplyst partivalg ofte er lidt mindre, forekommer 4.0 at være en rimelig øvre grænse.

12. Jf. note 3.
13. Jf. note 9 er den forventede værdi på forskelsmålet 4.8 med en standardafvigelse på 1.29 pct. Dermed bliver standardafvigelsen på et gennemsnit over 6 måneder = $1.29/\sqrt{6} = 0.53$. Med den samme reservation som i note 11 forekommer 6.0 at være en rimelig øvre grænse.
14. Dermed være ikke sagt, at det ikke ville være muligt at konstruere et endnu bedre mål for partiernes månedlige tilslutning. For eksempel kunne man overveje at lade de forskellige målinger indgå i gennemsnittet med forskellig vægt, afhængig af deres kvalitet. Endvidere kunne man ved hjælp af *a priori* antagelser indbygge et forsigtigt tendens-skøn. Jeg har haft gode erfaringer med at gøre noget lignende i forbindelse med DR TV's valgaftensprognoser.

Litteratur

- Borre, Ole og Jørgen Goul Andersen (1997). *Voting and Political Attitudes in Denmark: A Study of the 1994 Election*, Aarhus: Aarhus University Press.
- Cochran, William G. (1977). *Sampling Techniques*, 3rd edition, New York: Wiley.
- Elklit, Jørgen (1989). "Opinionsmålinger og valgresultater", i Jørgen Elklit og Ole Tonsgaard, *To folketingsvalg: Vælgerholdninger og vælgeradfærd i 1987 og 1988*, Aarhus:

- Politica.
- Hellevik, Ottar (1991). *Forskningsmetode i sociologi og statsvitenskap*, 5. udgave, Oslo: Universitetsforlaget.
- Johnson, Norman, Samuel Kotz and N. Balakrishnan (1994). *Continuous Univariate Distributions*, vol. 1, 2nd edition, New York: Wiley.
- Lynn, Peter and Roger Jowell (1996). "How Might Opinion Polls be Improved?: the Case for Probability Sampling", *Journal of the Royal Statistical Society. Series A - Statistics in Society*, vol 159, part 1: 21-28.
- Olsen, Henning (1998). *Tallenes talende tavshed*, København: Akademisk Forlag.
- Thomsen, Søren Risbjerg (1995). "Politiske meningsmålinger", *Politica*, 27. årg., pp. 71-90.
- Thomsen, Søren Risbjerg (1998). "FV98: Mobilisering mod en borgerlig regering i valgkampens slutspurt", <http://www.monday.dk/ugebrevet/bgmat/default.htm>.
- Voss, D. Stephen, Andrew Gelman and Gary King (1995). "The Polls - A Review. Preelection Survey Methodology: Details from Eight Polling Organizations, 1988 and 1992", *Public Opinion Quarterly*, vol. 59, pp. 98-132.