

## Bilag 1: Analysens datakvalitet

I dette bilag vil kvaliteten af de indsamlede data blive kontrolleret. En høj datakvalitet er en forudsætning for, at resultaterne fra analyserne kan bruges til at udtale sig om de danske virksomheder som helhed.

### Udvælgelse af stikprøve

Undersøgelsens målgruppe kan opdeles på henholdsvis IT-virksomheder, private virksomheder og offentlige virksomheder.

### Udvælgelse af stikprøve for IT-virksomheder

Virksomhederne i IT-erhvervene er grupperet i fire branchegrupper (jf. Informations-samfundet Danmark): IT-industri, IT-engroshandel, Telekommunikation og IT-konsulentvirksomhed.

I tabel 1.1 nedenfor er det samlede antal virksomheder med mindst 1 ansat i IT-erhvervene fordelt efter branche og størrelse. Tabellen er udarbejdet på baggrund af oplysninger fra Købmandsstandens Oplysningsbureau (KOB), februar 2001. Fordelingen af virksomheder viser, at langt størstedelen af virksomhederne inden for IT-erhvervene er beskæftiget med IT-konsulentytelser, idet 1.952 af de i alt 3.490 virksomheder i IT-erhvervene med mindst 1 ansat tilhører denne branchegruppe.

**Tabel 1.1. Virksomheder i IT-erhvervene fordelt efter branche og størrelse**

	(1-4)	(5-9)	(10-19)	(20-49)	(50-99)	(100+)	I alt
1. IT-industri (3001, ..., 3330)	135	72	44	54	30	41	376
2. IT-engroshandel (514320, ..., 516520)	460	242	164	128	36	33	1.063
3. Telekommunikation (6420)	29	16	17	18	9	10	99
4. IT-konsulentvirksomhed (7133, ..., 7260)	1151	336	197	148	64	56	1.952
I alt	1.775	666	422	348	139	140	3.490

Kilde: Købmandsstandens database.

På baggrund af tabel 1.1 vurderes det, at en simpel tilfældig udvælgelse af virksomheder fra IT-erhvervene ikke er hensigtsmæssig til en undersøgelse af behovet for kompetencer og kvalifikationer blandt de ansatte, da denne tilgang vil resultere i en stikprøve med mange virksomheder med få ansatte. Stikprøven vil derfor kun dække en relativt lille del af den samlede beskæftigelse i IT-erhvervene.

Stikprøven er i stedet sammensat således, at de store virksomheder overrepræsenteres. Tabel 1.2 indeholder en udvalgsbrøk, der sikrer, at der indgår forholdsvis mange store virksomheder og forholdsvis få små virksomheder i stikprøven.

**Tabel 1.2. Udvalgsbrøker fordelt på størrelsesgrupper**

	(1-4)	(5-9)	(10-19)	(20-49)	(50-99)	(100+)	I alt
Udvalgsbrøk	0,5	1,0	1,0	1,0	1,0	1,0	
Antal firmaer i population	1.775	666	422	348	139	140	3.490
Antal firmaer i bruttostikprøve	888	666	422	348	139	140	2.603

Samlet resulterer denne udvælgelse i en bruttostikprøve på knapt 75% af det samlede antal virksomheder inden for IT-erhvervene, nemlig 2.603 virksomheder. Disse virksomheder dækker til gengæld stort set alle ansatte i IT-erhvervene. Dette ses ved at sammenholde tabel 1.2 med nedenstående tabel 1.3, hvor antallet af ansatte i virksomhederne inden for IT-erhvervene er opgjort.<sup>1</sup>

**Tabel 1.3. Antal ansatte i virksomheder i IT-erhvervene fordelt efter branche og størrelse**

	(1-4)	(5-9)	(10-19)	(20-49)	(50-99)	(100+)	I alt
1. IT-industri (3001, ..., 3330)	293	491	608	1.722	2.091	11.772	16.977
2. IT-engroshandel (514320, ..., 516520)	1.013	1.650	2.289	3.929	2.301	9.645	20.827
3. Telekommunikation (6420)	67	108	238	564	632	2.555	4.164
4. IT-konsulentvirksomhed (7133, ..., 7260)	2.227	2.178	2.658	4.445	4.270	14.783	30.561
I alt	3.600	4.427	5.793	10.660	9.294	38.755	72.529

Kilde: Købmandsstandens database.

Ved at gange de respektive udvalgsbrøker fra tabel 1.2 på nederste række i tabel 1.3 findes det forventede antal ansatte i bruttostikprøven, nemlig 70.729, hvilket svarer til 98% af totalen. Der opnås derfor en næsten total dækning i forhold til undersøgelsens tema, de ansattes kompetencer og kvalifikationer, ved brug af denne metode.

Tilbage resterer branchedimensionen i stikprøven. S sammensætningen af stikprøven er foretaget på en sådan måde, at branchefordelingen inden for hver størrelsesgruppe er den samme som i den samlede population. Dette illustreres mest enkelt ud fra den endelige stikprøves sammensætning, der er givet i tabel 1.4 nedenfor.

<sup>1</sup> Det skal bemærkes, at der samlet er færre ansatte i IT-erhvervene i ovenstående opgørelse på baggrund af KOB, end opgjort i 'Informationssamfundet Danmark', der er baseret på oplysninger i Danmarks Statistiks 'Firmastatistik'. Forskellen skyldes først og fremmest forskelligt enhedsniveau, hvor tabel 1.3 kun omfatter virksomhedernes hovedafdelinger. Dette vurderes at være uden betydning ved beskrivelse af metoden, hvor det alene er relative sammenligninger, (stikprøvens dækning i forhold til totalen), der vurderes. I forbindelse med selve undersøgelsen spørges der ind til antal ansatte, og forskellene er derfor uden betydning ved den egentlige analyse.

I stikprøven indgår der f.eks. 68 virksomheder med mellem 1 og 4 ansatte inden for IT-industri, hvilket er 7,7% af samtlige virksomheder i denne størrelsesgruppe. I den samlede population af virksomheder er andelen ligeledes 7,7%, nemlig 135 af i alt 1.775 med 1 til 4 ansatte.

**Tabel 1.4. Bruttostikprøve af virksomheder inden for IT-erhvervene.**

	(1-4)	(5-9)	(10-19)	(20-49)	(50-99)	(100+)	I alt
1. IT-industri	68	72	44	54	30	41	309
2. IT-engroshandel	230	242	164	128	36	33	833
3. Telekommunikation	15	16	17	18	9	10	85
4. IT-konsulentvirksomhed	576	336	197	148	64	56	1377
I alt	888	666	422	348	139	140	2.603

Da der i stikprøven er en overrepræsentation af store virksomheder, skal analysen af resultaterne korrigeres herfor.

### Udvælgelse af stikprøve med private virksomheder

Undersøgelsen af behovet for IT-kompetencer og kvalifikationer suppleres med en undersøgelse af behovet uden for IT-erhvervet, herunder såvel private som offentlige virksomheder. I det følgende beskrives stikprøven af private virksomheder.

Der er – ikke overraskende – langt flere virksomheder uden for IT-erhvervene end inden for IT-erhvervene. For at sikre en tilstrækkelig dækning i forhold til undersøgelsens tema, nemlig de ansattes kvalifikationer og kompetencer, er det derfor nødvendigt at målrette undersøgelsen mod relevante virksomheder. Det er valgt at begrænse populationen af virksomheder med hensyn til brancher og antal ansatte.

Private virksomheder afgrænses i nærværende undersøgelse til virksomheder i byerhvervene, hvilket er den private sektor minus de primære erhverv, dvs. landbrug, fiskeri m.v. Det vurderes, at behovet for IT-kompetencer og kvalifikationer inden for de primære erhverv er så begrænset, at omkostningerne ved at medtage denne gruppe i undersøgelsen langt overstiger værdien af at kende behovene i disse erhverv.<sup>2</sup>

De private byerhverv omfatter følgende hovedbrancher: Industri, Bygge- og anlægsvirksomhed, Handel og reparationsvirksomhed, Hotel- og restaurationsvirksomhed,

<sup>2</sup> Af tabel 6.4.1. i 'Informationssamfundet Danmark' fremgår det, at der i alt er beskæftiget 108.655 personer i hhv. Landbrug og Fiskeri, hvoraf kun 410 har enten en IT-uddannelse eller er beskæftiget i en IT-stilling.

Transportvirksomhed m.v., Pengeinstitutter, finansierings- og forsikringsvirksomhed, Forretningsservice m.v., samt Offentlige og personlige tjenesteydelser.

I virksomhederne uden for IT-erhvervene vil IT-medarbejderne typisk være knyttet til virksomhedens kerneaktivitet, men til administrative funktioner, vedligeholdelse af systemer og programmel m.v. Det vurderes, at det kun er virksomheder af en vis størrelse, der har et behov for IT-medarbejdere til varetagelse af disse opgaver. For at målrette undersøgelsen er det derfor valgt at sætte en generel grænse på minimum 20 ansatte som dog, efter ønske fra opdragsgiver, suppleres af en mindre stikprøve blandt virksomheder med minimum 10 ansatte for brancherne Industri, Pengeinstitutter m.v. og Forretningsservice.

De valgte branche- og størrelsesmæssige afgrænsninger er sat som en afvejning mellem på den ene side, at opnå et nuanceret og detaljeret billede af behovet for IT-kompetencer og kvalifikationer i dansk erhvervsliv, og på den anden side at sikre, at stikprøven, undersøgelsen baseres på, udgør et tilstrækkeligt stort udsnit af danske virksomheder til, at undersøgelsens resultater er statistisk valide.

Med de valgte afgrænsninger består population i denne del af undersøgelsen af virksomheder i de private byerhverv med mindst 20 ansatte, jf. tabel 1.5 nedenfor.

**Tabel 1.5. Virksomheder uden for IT-erhvervene fordelt efter branche og størrelse**

	(20-49)	(50-99)	(100+)	I alt
1. Industri (15-37)	1.630	643	662	2.935
2. Bygge- og anlægsvirksomhed (45)	879	167	94	1.140
3. Handel og reparationsvirksomhed (50-52)	1.781	484	306	2.571
4. Hotel- og restaurationsvirksomhed (55)	251	54	31	336
5. Transportvirksomhed m.v. (60-64)	556	170	131	857
6. Pengeinst., finansierings- og forsikringsvirk. (65-67)	354	122	79	555
7. Forretningsservice m.v. (70-74)	652	202	145	999
8. Off. og pers. tjenesteydelser (80-85, 92 og 93)	806	212	187	1.205
I alt	6.909	2.054	1.635	10.598

Kilde: Købmandsstandens database.

Som for virksomhederne inden for IT-erhvervene vurderes det på baggrund af tabel 1.5, at en simpel tilfældig udvælgelse blandt virksomheder med 20 eller flere ansatte ikke er hensigtsmæssig, da denne tilgang som tidligere beskrevet vil resultere i en stikprøve med mange virksomheder med få ansatte og dermed en forholdsvis lav dækning i forhold til det samlede antal ansatte. I stedet sammensættes stikprøven, så de store virksomheder bliver overrepræsenteret.

Tabel 1.6 indeholder udvalgsbrøker der sikrer, at der indgår relativt mange store virksomheder og relativt få mellemstore virksomheder i stikprøven. Disse udvalgsbrøker afspejler de indledende overvejelser om sammenhængen mellem behovet for IT-kompetencer og kvalifikationer og virksomhedens størrelse, nemlig at behovet vokser med virksomhedens størrelse. Ved siden af de opstillede udvalgsbrøker udtrækkes der efter ønske fra opdragsgiver 100 industrivirksomheder, 59 pengeinstitutter m.v. og 100 forretningservicevirksomheder med mellem 10 og 19 ansatte.

**Tabel 1.6. Udvalgsbrøker fordelt på størrelsesgrupper**

	(20-49)	(50-99)	(100+)	I alt
Udvalgsbrøk	0,1	0,2	0,4	
Antal firmaer i population	6.909	2.054	1.635	
Antal firmaer i bruttostikprøve	691	411	654	1.756

Samlet resulterer denne udvælgelse i en bruttostikprøve på ca. 1,5% af det samlede antal virksomheder uden for IT-erhvervene, nemlig 1.756 virksomheder af i alt 104.664 virksomheder med mindst 1 ansat. I forhold til populationen af virksomheder med 20 ansatte og derover dækker stikprøven i ca. 16%.<sup>3</sup> I forhold til antallet af ansatte dækker stikprøven ca. 20% af det samlede antal ansatte. Dette ses ved at sammenholde tabel 1.6 med nedenstående tabel 1.7, hvor antallet af ansatte i virksomhederne uden for IT-erhvervene er opgjort.

**Tabel 1.7. Antal ansatte i virksomheder uden for IT-erhvervene fordelt efter branche og størrelse**

	(1-4)	(5-9)	(10-19)	(20-49)	(50-99)	(100+)	I alt
1. Industri	11.199	17.358	27.346	51.085	45.209	195.157	347.354
2. Bygge- og anlægsvirksomhed	15.816	19.374	21.903	25.297	11.089	25.988	119.467
3. Handel og reparationsvirksomhed	42.117	42.232	44.778	56.829	35.237	86.887	308.080
4. Hotel- og restaurationsvirksomhed	8.370	4.946	5.547	7.845	3.559	9.508	39.775
5. Transportvirksomhed m.v.	9.079	8.618	10.567	17.877	12.279	68.303	126.723
6. Pengeinst., finansierings- og forsikringsvirk.	3.988	7.212	10.012	10.366	8.057	23.999	63.634
7. Forretningservice m.v.	20.157	16.420	17.825	24.532	18.082	61.465	158.481
8. Offentlige og personlige tjenesteydelser.	19.256	14.199	16.846	24.386	14.642	47.686	137.015
I alt	129.982	130.359	154.824	218.217	148.154	518.993	1.300.529

Kilde: Købmandsstandens database.<sup>4</sup>

<sup>3</sup> Opgjort på baggrund af KOB.

<sup>4</sup> Se fodnote 1.

Ved at gange de respektive udvalgsbrøker fra tabel 1.6 på nederste række i tabel 1.7 findes det forventede antal ansatte i bruttostikprøven, nemlig 259.049, hvilket svarer til 20% af totalen. Der opnås derfor en relativt stor dækning i forhold til undersøgelsens tema; de ansattes kompetencer og kvalifikationer.

Tilbage resterer branchedimensionen i stikprøven. Stikprøven sammensættes, så branchefordelingen inden for hver størrelsesgruppe er den samme som i den samlede population. Den endelige stikprøves sammensætning er givet i tabel 1.8 nedenfor.

**Tabel 1.8. Bruttostikprøve af virksomheder uden for IT-erhvervene**

	(10-19)	(20-49)	(50-99)	(100+)	I alt
1. Industri (15-37)	100	163	129	265	656
2. Bygge- og anlægsvirksomhed (45)	0	88	33	38	159
3. Handel og reparationsvirksomhed (50-52)	0	178	97	122	397
4. Hotel- og restaurationsvirksomhed (55)	0	25	11	12	48
5. Transportvirksomhed m.v. (60-64)	0	56	34	52	142
6. Pengeinst., finansierings- og forsikringsvirk. (65-67)	59	35	24	32	150
7. Forretningsservice m.v. (70-74)	100	65	40	58	264
8. Off. og pers. tjensteydelser (80-85, 92 og 93)	0	81	42	75	198
I alt	259	691	411	654	2.015

Da der i stikprøven er en overrepræsentation af store virksomheder, skal analysen af resultaterne korrigeres herfor.

### Udvælgelse af stikprøve for offentlige institutioner

Den tredje stikprøve af virksomheder i undersøgelsen består af offentlige virksomheder og institutioner. Der er her foretaget en lignede afgrænsning af populationen i forhold til funktion og antal ansatte. Det er valgt at udelade offentlige virksomheder inden for 'Pleje og pasning', dvs. børnehaver, plejehjem, fritidshjem m.v., da tidligere undersøgelser har vist, at behovet for medarbejdere med IT-kompetencer stort set er ikke eksisterende i denne type virksomhed.<sup>5</sup> Som for den private sektor er det endvidere valgt at sætte en nedre grænse på minimum 20 ansatte.

Den samlede population af offentlige virksomheder er med disse begrænsninger opgjort i tabel 1.9 nedenfor, mens udvalgsbrøkerne er givet i tabel 1.10. Udvalgsbrøkerne er voksende med virksomhedernes størrelse. Samlet resulterer dette i en stikprøve på 728 virksomheder, hvor de store virksomheder er overrepræsenteret.

<sup>5</sup> Se 'Flaskehalsundersøgelsen', PLS RAMBØLL Management, 2000.

**Tabel 1.9. Offentlige virksomheder fordelt efter funktion og størrelse**

	(1-4)	(5-9)	(10-19)	(20-49)	(50-99)	(100+)	I alt
1. Forsyningsvirksomhed m.v.	241	139	166	173	48	18	785
2. Transport og kommunikation	113	23	15	11	2	10	174
3. Kontrol, rådgivning og forskning	193	73	39	42	12	30	389
4. Offentlig adm.	419	116	178	155	107	473	1.448
5. Undervisning	300	166	550	824	640	66	2.546
6. Sundhedsvæsen	123	132	81	54	12	80	482
7. Foreninger og kultur	2.373	433	174	95	35	12	3.122
I alt	3.762	1.082	1.203	1.354	856	689	8.946

Kilde: Købmandsstandens database.

**Tabel 1.10. Udvalgsbrøker fordelt på størrelsesgrupper**

	(20-49)	(50-99)	(100+)	I alt
Udvalgsbrøk	0,125	0,25	0,5	
Antal firmaer i population	1.354	856	689	
Antal firmaer i bruttostikprøve	169	214	345	728

Samlet resulterer denne udvælgelse i en bruttostikprøve på godt 8% af det samlede antal offentlige virksomheder eksklusiv pasnings- og plejeområdet. Disse virksomheder dækker 286.756 ansatte, eller det der svarer til 43% af de ansatte.

**Tabel 1.11. Antal ansatte i offentlige institutioner fordelt efter funktion og størrelse. Oplysninger fra Købmandsstandens database<sup>6</sup>**

	(1-4)	(5-9)	(10-19)	(20-49)	(50-99)	(100+)	I alt
1. Forsyningsvirksomhed m.v.	497	973	2.490	6.055	3.600	6.340	19.955
2. Transport og kommunikation	187	161	225	375	145	12.265	13.358
3. Kontrol, rådgivning og forskning	329	511	587	1.456	887	10.006	13.776
4. Offentlig adm.	685	812	2.668	5.446	7.942	410.190	427.743
5. Undervisning	604	1.163	8.250	28.831	47.967	30.515	117.330
6. Sundhedsvæsen	305	924	1.215	1.875	895	56.800	62.014
7. Foreninger og kultur	3.881	3.031	2.610	3.325	2.625	3.525	18.997
I alt	6.488	7.575	18.045	47.363	64.061	529.641	673.173

Bruttostikprøven for de offentlige institutioner er angivet nedenfor.

<sup>6</sup> Se fodnote 1.

**Tabel 1.12. Bruttostikprøve for offentlige institutioner**

	(20-49)	(50-99)	(100+)	I alt
1. Forsyningsvirksomhed m.v.	22	12	9	43
2. Transport og kommunikation	1	1	5	7
3. Kontrol, rådgivning og forskning	5	3	15	23
4. Offentlig adm.	19	27	237	283
5. Undervisning	103	160	33	296
6. Sundhedsvæsen	7	3	40	50
7. Foreninger og kultur	12	9	6	27
I alt	169	214	345	728

## Sammenfatning

Spørgeskemaundersøgelsen til belysning af danske virksomheders behov for IT-kompetencer og kvalifikationer udarbejdes på baggrund af en stikprøve på 5.087 virksomheder,<sup>7</sup> fordelt med 2.603 virksomheder i IT-erhvervene med mindst 1 ansat, 1.756 virksomheder i de private byerhverv uden for IT-erhvervene med mindst 20 ansatte<sup>8</sup> og endelig 728 offentlige virksomheder, ligeledes med mindst 20 ansatte. Udvalgsbrøkerne er fastsat, så der indgår relativt mange store virksomheder i undersøgelsen: Inden for IT-erhvervene er stikprøven stort set totaldækkende med hensyn til ansatte, i de øvrige grupper er det ca. 20% af de ansatte, der er dækket af stikprøven.

## Undersøgelsens metode

Undersøgelsen er gennemført som en kombination af telefoninterview og Internet-besvarelser, og dataindsamlingen er gennemført i perioden fra den 6. maj til den 30. maj 2001.

Den 6. maj blev der sendt et brev til de personaleansvarlige/virksomhedsledere i de ca. 5.300 private og offentlige virksomheder i stikprøven. Respondenterne blev i brevet opfordret til at besvare spørgeskemaet elektronisk over Internettet på webadressen [www.survey.pls.dk/internet/it.asp](http://www.survey.pls.dk/internet/it.asp), hvor de skulle angive en personlig kode (IP nr.) samt et password, som blev oplyst i brevet.

<sup>7</sup> Hertil kommer 259 Industrivirksomheder (100), Pengeinstitutter m.v. (59) og Forretningsservicevirksomheder (100) udenfor IT-branchen udtrukket særskilt blandt virksomheder med 10-19 ansatte, som behandles særskilt i forhold til den øvrige gruppe i undersøgelsen.

<sup>8</sup> Eksklusiv pasnings- og plejeområdet.



Virksomheder, som selv besvarede spørgsmålene via Internettet, blev ikke ringet op efterfølgende. Mens de virksomheder, som endnu ikke havde besvaret spørgeskemaet en uge efter udsendelsen af brevene, blev ringet op af en af PLS RAMBØLLs telefoninterviewere med henblik på at gennemføre et telefoninterview.

Telefoninterviewene er gennemført i perioden den 14. maj til den 30. maj 2001 i tidsrummet kl. 8.00 til 16.30 ugens fem hverdage. Telefoninterviewene er gennemført ved hjælp af CATI (Computer Assisted Telephone Interviewing), dvs. at hver interviewer sidder ved en PC og indtaster respondenternes svar direkte. Spørgsmålene viser sig ét efter ét på skærmen, idet programmet sørger for, at den enkelte respondent ledes rigtigt gennem spørgeskemaet.

Computeren administrerer spring i spørgeskemaet, som er betinget af baggrundsoplysninger eller af de svar, der bliver givet. Det vil sige, hvis en respondent f.eks. har svaret 'uændret' i spørgsmål 10, sørger programmet for, at intervieweren næste gang stiller spørgsmål 12 til den pågældende respondent, og således overspringer de spørgsmål, som ikke er relevante for netop denne respondent. Endvidere sørger programmet for, at spørgsmål 7 f.eks. kun stilles til IT-virksomheder og private virksomheder, som har svaret 'nej' på spørgsmål 6, mens spørgsmål 9 stilles til offentlige virksomheder, som har svaret 'nej' på spørgsmål 6.

## **Deltagelse og svarprocent**

Tabel 1.13 viser den samlede deltagelse i undersøgelsen. Som det fremgår af tabellen, er der alt i alt gennemført interview med 2.016 virksomheder, mens 2.907 virksomheder, som havde mulighed for at deltage i undersøgelsen, ikke gjorde det af den ene eller anden grund. Den samlede svarprocent er således 41, hvilket må betragtes som tilfredsstillende i en undersøgelse som denne.

Ser vi på svarprocenterne opdelt på virksomhedstyper, fremgår det af tabel 1.13, at svarprocenten er størst blandt de offentlige virksomheder (44%) og mindst blandt de private virksomheder (39%), mens svarprocenten for IT-virksomhederne (42%) ligger mellem svarprocenten for henholdsvis de offentlige og de private virksomheder.

De 2.907 respondenter, som ikke deltog i undersøgelsen, men som havde mulighed for at deltage, ønskede ikke at deltage eller kunne ikke træffes i undersøgelsesperioden. Det skal i den forbindelse understreges, at alle disse virksomheder har haft mu-

lighed for at deltage i undersøgelsen, idet virksomhederne har modtaget et brev med en opfordring til at deltage i undersøgelsen via Internettet. Endvidere er alle virksomheder blevet ringet op (mindst 3-4 gange) med henblik på gennemførelsen af et telefonisk interview.

**Tabel 1.13. Deltagelse og svarprocent i undersøgelsen**

	IT-virksomheder		Private virksomheder		Offentlige virksomheder		Samlet	
	Antal	Procent	Antal	Procent	Antal	Procent	Antal	Procent
<b>Bruttostikprøve</b>	<b>2.604</b>		<b>2.014</b>		<b>729</b>		<b>5.347</b>	
Virksomheden eksisterer ikke	137	50,2	38	29,0	11	55,0	186	43,9
Falder uden for målgruppen	50	18,3	21	16,0	3	15,0	74	17,5
Tidligere ringet op	10	3,7	3	2,3	0		13	3,1
Filial	76	27,8	69	52,7	6	30,0	151	35,6
<b>Nettostikprøve</b>	<b>2.331</b>	<b>100,0</b>	<b>1.883</b>	<b>100,0</b>	<b>709</b>	<b>100,0</b>	<b>4.923</b>	<b>100,0</b>
Ønsker ikke at deltage <sup>1)</sup>	1.351		1.156		400		2.907	
<b>Gennemførte interview</b>	<b>980</b>	<b>100,0</b>	<b>727</b>	<b>100,0</b>	<b>309</b>	<b>100,0</b>	<b>2.016</b>	<b>100,0</b>
<b>Svarprocent</b>		<b>42,0</b>		<b>38,6</b>		<b>43,6</b>		<b>41,0</b>

Anm. Svarprocenten er beregnet i forhold til nettostikprøven.

<sup>1)</sup> Virksomheder som ikke ønsker at deltage i undersøgelsen eller ikke kan træffes i undersøgelsesperioden.

Tabel 1.14 viser svarprocenterne opdelt på brancher, og det fremgår heraf, at svarprocenten varierer fra 32,6 for transportvirksomhed m.v. til 66,7 for offentlig transport og kommunikation (denne svarprocent er dog kun baseret på besvarelserne fra 4 ud af 6 virksomheder).

Vi har testet, om der er signifikant forskel på svarprocenterne i de enkelte brancher indenfor de tre virksomhedstyper (IT-erhvervene, private virksomheder og offentlige virksomheder). Ud fra et chi-square test viser det sig, at der ikke er grundlag for at tro, at enkelte brancher er under- henholdsvis overrepræsenterede i en grad, at det gør det til et problem i analyserne.

**Tabel 1.14. Svarprocenter opdelt på brancher**

	Bruttostikprøve	Nettostikprøve	Gennemførte interview	Svarprocent
<b>Total</b>	<b>5.347</b>	<b>4.923</b>	<b>2.016</b>	<b>41,0</b>
IT-virksomheder	IT-industri	309	123	43,6
	IT-engroshandel	833	304	40,4
	IT-Telekommunikation	85	75	34,7
	IT-konsulentvirksomhed	1377	1.222	43,1
<b>Total (IT-virksomhed)</b>	<b>2.604</b>	<b>2.331</b>	<b>980</b>	<b>42,0</b>
Private virksomheder	Industri	656	236	38,2
	Bygge- og anlægsvirksomhed	159	149	38,9
	Handel og reparationsvirksomhed	397	363	36,6
	Hotel og restaurationsvirksomhed	48	46	41,3
	Transportvirksomhed m.v.	142	129	32,6
	Pengeinst, finansierings- og forsikringsvirk.	150	134	36,6
	Forretningservice m.v.	264	251	41,8
	Offentlige og personlige tjenesteydelser	198	193	44,0
<b>Total (Private virksomheder)</b>	<b>2.014</b>	<b>1.883</b>	<b>727</b>	<b>38,6</b>
Offentlige virksomheder	Offentlig forsyningsvirksomhed m.v.	43	15	38,5
	Offentlig transport og kommunikation	7	4	66,7
	Offentlig kontrol, rådgivning og forskning	23	23	43,5
	Offentlig administration	283	278	43,2
	Offentlig undervisning	296	289	43,3
	Offentlig sundhedsvæsen	50	47	46,8
	Offentlig foreninger og kultur	27	27	48,1
<b>Total (Offentlige virksomheder)</b>	<b>729</b>	<b>709</b>	<b>309</b>	<b>43,6</b>

Anm. Svarprocenten er beregnet i forhold til nettostikprøven.

## Repræsentativitet

Repræsentativiteten skal være til stede, for at det er muligt at generalisere ud fra undersøgelsens resultater. For at sikre repræsentativitet på udvalgte baggrundsvariable, er der foretaget en løbende kontrol af repræsentativiteten med hensyn til virksomhedernes størrelse og branche. Disse repræsentativitetstest er indarbejdet i indtastningsprogrammet, således at det for den enkelte interviewer har været muligt at se hvilke grupper, der har været de mest kritiske grupper. Tabel 1.15 viser fordelingen på antallet af ansatte i henholdsvis IT-erhvervet, private virksomheder og offentlige virksomheder.

**Tabel 1.15. Fordeling af virksomheder i populationen og i undersøgelsen**

	Gennemførte interview		Populationen		Vægtning til population		Vægtning til stikprøve		
	N	Pct.	N	Pct.	Vægt	Antal	Vægt	Antal	
IT-erhvervet	(1-4)	313	31,9	1.775	50,9	5,67	1.775	1,59	498
	(5-9)	268	27,3	666	19,1	2,49	666	0,70	187
	(10-19)	161	16,4	422	12,1	2,62	422	0,74	118
	(20-49)	129	13,2	348	10,0	2,70	348	0,76	98
	(50-99)	58	5,9	139	4,0	2,40	139	0,67	39
	(100+)	51	5,2	140	4,0	2,75	140	0,77	39
	Total	980	100,0	3.490	100,0		3.490		980
Private	(20-49)	259	41,3	6.909	65,2	26,68	6.909	1,58	409
	(50-99)	136	21,7	2.054	19,4	15,10	2.054	0,89	122
	(100+)	232	37,0	1.635	15,4	7,05	1.635	0,42	97
	Total <sup>1)</sup>	627	100,0	10.598	100,0		10.598		627
Offentlige	(20-49)	64	20,7	1.354	46,7	21,16	1.354	2,26	144
	(50-99)	98	31,7	856	29,5	8,73	856	0,93	91
	(100+)	147	47,6	689	23,8	4,69	689	0,50	73
	Total	309	100	2.899	100,0		2.899		309

<sup>1)</sup> Hertil kommer 100 gennemførte interview med virksomheder med 10 - 19 ansatte.

I beregningen for et repræsentativt udsnit af virksomheder for henholdsvis IT-erhvervet, private virksomheder og offentlige virksomheder er det nødvendigt at veje datamaterialet. For eksempel kan stikprøveandelen og populationsandelen af IT-virksomheder med 1 - 4 ansatte, på grundlag af tallene i tabel 1.15, bestemmes til henholdsvis 0,319 og 0,509.

IT-virksomheder med 1 - 4 ansatte er altså underrepræsenteret i stikprøven, og de bør derfor vægtes relativt meget i forhold til de øvrige IT-virksomheder. Beregnes vægten, viser det sig, at hver enkelt IT-virksomhed med 1 - 4 ansatte skal vægte 1,59  $(1.775 \times 980) / (313 \times 3.490)$  for at vægte relativt i stikprøven og 5,67  $(1.775 / 313)$ , når stikprøven skal opskaleres til populationsniveau. Tilsvarende vægte er beregnet for de øvrige IT-virksomheder samt for de private og de offentlige virksomheder. Samtlige beregninger i rapporten er gennemført på baggrund af det vægtede materiale.

## *Bilag 2: Metodebilag vedr. Linear Structural Relationships*

En vigtig problemstilling i undersøgelsen er at afdække de forhold, som er af betydning for virksomhedernes fremtidige efterspørgsel efter IT-arbejdskraft. For at svare på dette, er der foretaget en kausalanalyse (jf. afsnit 4.4 i rapporten).

De fleste er nok bekendte med kausalanalyse i form af regressionsmodeller. I situationer som denne med mange variable målt på en ordinal skala (1,2,3,4,5) og behæftet med usikkerhed, har almindelige regressionsmodeller dog en række afgørende svagheder, der gør det problematisk at anvende disse. Som nogle af de vigtigste kan nævnes, at regressionsmodeller forudsætter:

- at alle forklarende variable er målt på interval-/ratioskala
- at forklarende variable er kendte størrelser, der ikke er behæftet med måleusikkerhed
- at de forklarende variable ikke er stærkt indbyrdes korrelerede (multikolinearitet)
- at i modeller, der består af mere end én regressionsligning, kan disse estimeres uafhængigt af hinanden.

I spørgeskemaundersøgelser vil flere – om ikke alle – af ovenstående forudsætninger næsten altid være brudt. På den baggrund er der gennem de seneste år udviklet metoder, der er specielt velegnede til kausalanalyse af spørgeskemaundersøgelser og lignende, hvor ovenstående forudsætninger i reglen ikke kan forventes at være opfyldt.

Den mest generelle af disse metoder er uden tvivl LISREL (Linear Structural Relationships), hvor den ordinære regressionsmodel optræder som et simpelt specialtilfælde.

## Opstilling af målemodeller

Som nævnt i indledningen til dette afsnit, er en afgørende forudsætning i den ordinære regressionsmodel, at de forklarende variable ikke er behæftede med målefejl eller andre former for usikkerhed. I LISREL bryder vi med denne forudsætning ved opstilling af målemodeller for bagvedliggende latente variable.

Vi kan for eksempel betragte spørgsmål 5.6 og 22.3 i spørgeskemaet (jf. bilag 3):

5.6 Outsourcing af IT-opgaver i Danmark ( $X_1$ )

22.3 Øget brug af outsourcing af IT-opgaver ( $X_2$ )

Disse to spørgsmål udtrykker begge i en eller anden grad noget om outsourcing af IT-opgaver. På den baggrund må det forventes, at disse variable er rimelig tæt forbundet, således at hvis en virksomhed A generelt outsourcer flere IT-opgaver i Danmark end virksomhed B, vil virksomhed A også i højere grad være kendetegnet af øget brug af outsourcing af IT-opgaver end virksomhed B.

I en almindelig regressionsmodel ville dette skabe problemer, fordi vi ville have svært ved at skille effekten af den ene fra effekten af den anden.

I LISREL anvender vi i stedet det faktum, at de to variable må forventes at hænge rimeligt tæt sammen, således at vi kan danne en ny *latent* variabel ( $\xi$ ), som vi passende kunne kalde *Outsourcing af IT-opgaver*. Antagelsen er nu, at det i virkeligheden er værdien af denne *ikke observerede* (latente) variabel, der er bestemmende for virksomhedernes besvarelse af de to spørgsmål.

For at undersøge i hvilken udstrækning dette er tilfældet, opstiller vi en målemodel:

$$X_1 = \lambda_1 \xi + \delta_1$$

$$X_2 = \lambda_2 \xi + \delta_2$$

Det afgørende nye i forhold til en regressionsanalyse er, at vi tillader en målefejl  $\delta$  i vores observerede variable. Dette betyder, at vi for hver enkelt af vores indikatorvariable  $X_1$  og  $X_2$  kan vurdere reliabiliteten.

Ved reliabiliteten forstår vi, i hvor høj grad den konstruerede latente variabel  $\xi$  er i stand til at forklare variationen i hver enkelt af indikatorvariablene ( $X$ 'erne), dvs.:

$$R_{X_1}^2 = 1 - \frac{\text{var}(\delta_1)}{\text{var}(X_1)}$$

En betingelse for, at det er forsvarligt at antage, at de to spørgsmål alle er indikatorer for den samme bagvedliggende faktor *Outsourcing af IT-opgaver*, er, at reliabiliteten er rimeligt høj (omkring 0,6) for alle indikatorvariablene. Da dette er tilfældet for begge variable vedrørende *Outsourcing af IT-opgaver*, er det rimeligt at antage, at disse dækker over den samme bagvedliggende faktor.

På samme måde sammensættes resten af analysevariablene i overordnede faktorer med udgangspunkt i følgende to kriterier:

- Faktorerne skal være logiske, således at spørgsmålene under en given faktor meningsfuldt kan navngives samlet.
- Reliabiliteten af de enkelte indikatorer for faktoren skal være så høj, at det er rimeligt at antage, at indikatorerne er udtryk for den samme bagvedliggende faktor.

### **Strukturelle sammenhænge**

Efter at have opstillet målemodeller for de enkelte faktorer hver for sig såvel som samlet, og testet om disse på tilfredsstillende vis beskriver sammenhængene i datasættet, kan vi gå et skridt videre og analysere sammenhænge imellem de latente variable (strukturelle sammenhænge).

Dette skridt svarer i det store hele til en regressionsmodel, bortset fra at den strukturelle model her er udtrykt i de ikke-observerede latente variable (faktorer) frem for de direkte observerede variable i spørgeskemaet.

Ud over at give et langt større overblik over datamaterialet, har dette den store fordel, at multikolaritet (stærkt korrelerede forklarende variable) ikke er noget stort problem mellem de latente variable (da vi jo har sammensat stærkt korrelerede variable til faktorer), mens mange af de originale variable vil være så tæt forbundet, at de er svære at skille ad.

Resultatet af LISREL-analysen bliver således en model, der på baggrund af de overordnede faktorer forklarer respondenternes vurderinger af det pågældende område. Disse strukturelle sammenhænge estimeres simultant med alle målemodellerne i én samlet model, hvilket sikrer en optimal udnyttelse af informationen i datasættet.

For hver sammenhæng bestemmes endvidere forklaringsgraden ( $R^2$ ), der udtrykker den andel af variationen i den afhængige variabel, som modellen kan forklare.

Ikke signifikante sammenhænge (koefficienter, der ikke statistisk signifikant kan påvises at være forskellige fra 0) fjernes fra modellen, hvorefter de resterende kan anvendes til en rangordning af betydningen af forskellige faktorerens påvirkning af vurderingen af den overordnede tilfredshed.

Der vil generelt være en statistisk usikkerhed på mellem 0,05 og 0,10 på modelkoefficienter. Koefficienter med værdier under dette niveau er således ikke signifikante. Dette betyder også, at man generelt ikke kan konkludere, at der er forskel på to faktorerens betydning, med mindre disse koefficienter afviger fra hinanden med mere end 0,10 - 0,15.

### **Standardisering og anvendelse af ordinalt skalerede variable**

I det ovenstående er beskrevet, hvorledes vi ved anvendelse af LISREL kan tage højde for målefejl og undgå multikolaritet. Som nævnt i indledningen til dette afsnit forudsætter traditionel kausalanalyse dog også, at variable i analysen er målt på en interval- eller ratioskala. Det vil sige en veldefineret kontinuert skala, hvor 4 er dobbelt så meget som 2.

I spørgeskemaundersøgelser med ordinalt skalerede svarkategorier, som fx

(1) Slet ikke (2) I ringe grad (3) I nogen grad (4) I høj grad (5) I meget høj grad

er det spørgsmålet, om det er rimeligt at antage, at en respondent, der i høj grad er enig i et givent spørgsmål (4), er lige nøjagtigt dobbelt så enig som én, der i ringe grad er enig i et givent spørgsmål (2).

Under opstillingen af kausalmodellen ved LISREL, anvender vi ikke ordinalt skalerede variable, som om de var målt på en intervallskala. Det vil sige, at vi kun antager, at den ordinale skalering udtrykker en rangordning, således at en person, der svarer 4, er mere enig end én, der svarer 3.

Alternativt antager vi, at der bag de ordinalt skalerede variable, vi observerer, ligger kontinuerte normalfordelte variable. Ved antagelse om, at disse bagvedliggende variable parvist følger todimensionale normalfordelinger, har vi mulighed for at estimere



korrelationer mellem disse (polykoriske korrelationer). Analyserne baseres således på korrelationer mellem disse ikke-observerede variable frem for korrelationer mellem de faktisk observerede variable (pearson's korrelationer). Eksperimenter har vist, at de polykoriske korrelationer i forbindelse med analyse af ordinale variable er betydeligt bedre estimatorer for den "sande" korrelation end ordinære pearson's korrelationer.

Ovenstående fremgangsmåde betyder, at alle variable standardiseres, hvorved alle koefficienter i modellen (generelt) skal ligge i intervallet -1 til 1.