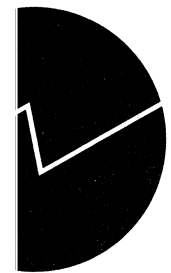


Hege Margrethe Bogstrand

**Analyse og metodeutvikling for
identifisering av enheter i
merverdiavgiftsmanntallet**

**Delprosjekt konvertering i
Enhetsregisterprosjektet**

Notater



Sammendrag

Hege Margrethe Bogstrand

Delprosjekt konvertering

Merverdiavgiftsmanntallet, Teststudie

Notater 94/7 • Statistisk sentralbyrå 1994

Delprosjektets konkrete resultatmål er å bringe Enhetsregisterets organisasjonsnummer inn i utvalgte etatsregistre slik at identiske enheter i hvert register får tildelt det samme nummeret. I praksis vil dette si nummer fra nummerserien i Foretaksregisteret i Brønnøysund (NyFor).

Skattedirektoratets merverdiavgiftsmanntall (MVA) er det registeret som først skal konverteres. Eksisterende relasjoner til NyFor må kontrolleres og godkjennes. Statistisk sentralbyrås Bedrifts- og Foretaksregister (BOF) er nøkkelen i dette arbeidet fordi det rent teknisk er via dette registeret at koblingene er etablert. Navnematch registre imellom må vurderes fordi det vil kunne gi flere koblinger til kontroll og på den måten forhindre at enheter som allerede har NyFor-nummer blir tildelt nytt nummer.

Hensikten med teststudiet var å få prøvd ut et kontroll- og godkjenningssystem for koblingene mellom enhetene i de tre involverte registrene.

Det viste seg at det maskinelle kontrollprogrammet godkjente ca. 91 og 94 prosent av koblingene mellom henholdsvis BOF-MVA og BOF-NyFor. Manuell gjennomgang av et utvalg av de godkjente linkene viste at bare i overkant av 1 prosent var faktiske feilkoblinger. Enhetsregisterprosjektet og Skattedirektoratet mente at dette var god nok kvalitet til at en ikke behøvde ytterligere kontroll av den maskinelt godkjente massen.

Alle potensielt gale linker, dvs. de som ikke ble godkjent i datakontroll-programmet, ble manuelt kontrollert. Omtrent 80 prosent av relasjonene mellom BOF og NyFor viste seg å være korrekte. Tilsvarende tall for BOF-MVA var om lag 95 prosent. Årsaken til at de ikke befinner seg i mengden med godkjente koblinger viste seg i hovedsak å være oppdateringsmangler i registrene på egenskaper ved enhetene. Det er derfor ikke mulig å skille ut disse som godkjente fra de faktisk gale ved å endre de maskinelle kontrollkriteriene.

Datakontrollprogrammet ble akseptert som tilfredsstillende og det vil bare være de potensielle feilkoblingene som må gjennomgås manuelt. På landsbasis utgjør dette omtrent 16 000 koblinger mot avgiftsmanntallet, når en legger en feilprosent på rundt tolv til grunn.

For å finne nye relasjoner ble det gjennomført navnematch av MVA mot NyFor og av MVA mot Løsreregisterets Foretaksregister (LøFor). En gjennomgang viste at de fleste av forslagene var korrekte, men ikke alle. Det er derfor nødvendig med en total manuell gjennomgang av slike relasjoner. Mengden ble anslått til å være vel 15 000 på landsbasis. Anslaget er imidlertid meget usikkert.

Konverteringen for hele landet gav 16 569 faktiske potensielle feilkoblinger. Resultatet av navnematchene gav 16 421 koblinger. Et resultat helt i tråd med teststudiets anslag.

Emneord: Bedriftsregister, Enhetsregister, Foretaksregister, Foretaksnummer, Identifikasjonsnummer, Registersamordning.

Innhold

1.	Innledning	7
2.	Grunnlaget	8
2.1.	Involverte registre	8
2.2.	Valg av geografisk område	8
2.3.	Formålstjenlig innhold i registrene	9
2.3.1.	BOF kontra NyFor	9
2.3.2.	BOF kontra MVA	9
2.3.3.	MVA kontra NyFor	10
3.	Valgte kontrollkriterier	11
3.1.	Eieform	11
3.1.1.	BOF kontra NyFor	11
3.1.2.	BOF kontra MVA	12
3.1.3.	MVA kontra NyFor	13
3.2.	Kommunenummer	13
3.2.1.	BOF kontra NyFor	13
3.2.2.	BOF kontra MVA	14
3.2.3.	MVA kontra NyFor.	14
3.3.	Navnelikhet	14
3.3.1.	BOF kontra NyFor	15
3.3.2.	BOF kontra MVA	15
3.3.3.	MVA kontra NyFor	15
4.	Mulige kontrollkriterier	16
4.1.	Tilstandskode	16
4.1.1.	BOF kontra NyFor.	16
4.1.2.	BOF kontra MVA	17
4.2.	Næringskode	17
4.3.	Adresse	17
4.4.	Dato	17
4.5.	Fødselsnummer	18
5.	MVA/NyFor enheter uten link via BOF	19
5.1.	Nyfor-nummer registrert etter 1/1-92 ved Fylkesskattekontorene	19
5.2.	Navnematch	19
5.2.1.	MVA kontra NyFor	19
5.2.2.	MVA kontra LøFor	19
6.	Datakontrollprogrammene	20
6.1.	Omfang til maskinell kontroll	20
6.2.	Resultat av maskinell kontroll	21
6.2.1.	Fordelingen av de ikke-godkjente koblingene etter kontrollkriterier	21
6.2.2.	Omfang av fellesmengden i MVA og NyFor	22
6.3.	Omfang til manuell kontroll i teststudiet	22

7.	Tilbakemelding fra Brønnøysundregistrene	23
7.1.	Maskinelt godkjente koblinger	23
7.1.1.	Resultater	23
7.1.2.	Feiltyper	23
7.1.3.	Konklusjon	24
7.2.	Maskinelt ikke-godkjente koblinger	24
7.2.1.	Resultater	24
7.2.2.	Beskrivelse av de som kom ut som feil i datakontrollprogrammet, men som viste seg å være riktige koblinger	25
7.2.3.	Koblinger som er gale fordelt på de tre kriteriene og tilstandsfeil	26
7.2.4.	Mulige årsaker til at koblingene er foretatt	26
7.2.5.	Konklusjon	26
7.3.	Navnematch mellom MVA og henholdsvis NyFor og LøFor	27
7.3.1.	MVA-NyFor koblingsforslag	27
7.3.2.	MVA-LøFor koblingsforslag	27
7.3.3.	Konklusjon	28
8.	Tilbakemelding fra fylkesskattekontoret i Sør-Trøndelag	29
8.1.	Maskinelt godkjente koblinger	29
8.1.1.	Resultater	29
8.1.2.	Feiltyper	29
8.1.3.	Konklusjon	30
8.2.	Maskinelt ikke-godkjente koblinger	30
8.2.1.	Resultater	30
8.2.2.	Beskrivelse av de som kom ut som feil i datakontrollprogrammet, men som viste seg å være riktige koblinger	31
8.2.3.	Koblinger som er gale fordelt på de tre kriteriene og tilstandsfeil	32
8.2.4.	Mulige årsaker til at koblingene er foretatt	33
8.2.5.	Konklusjon	33
9.	Hovedresultater og anbefalinger	34
9.1.	Omfang	34
9.2.	Kvalitet	34
Vedlegg		35

1. Innledning

Delprosjektets konkrete resultatmål er å bringe Enhetsregisterets organisasjonsnummer inn i utvalgte etatsregistre slik at identiske enheter i hvert register får tildelt det samme nummeret. I praksis vil dette si nummer fra nummerserien i Foretaksregisteret i Brønnøysund (NyFor).

Skattedirektoratets merverdiavgiftsmanntall (MVA) er det registeret som først skal konverteres. Eksisterende relasjoner til NyFor må kontrolleres og godkjennes. Statistisk sentralbyrås Bedrifts- og Foretaksregister (BOF) er nøkkelen i dette arbeidet fordi det rent teknisk er via dette registeret at koblingene er etablert. Navnmatch registre imellom må vurderes fordi det vil kunne gi flere koblinger til kontroll og på den måten forhindre at enheter som allerede har NyFor-nummer blir tildelt nytt nummer.

Hensikten med teststudiet var å få prøvd ut et kontroll- og godkjenningssystem for koblingene mellom enhetene i de tre involverte registrene. Systemet består av flere varianter av et datakontrollprogram og manuelle gjennomganger. Vi har prøvd dette ut for Sør-Trøndelag fylke. Dataprogrammet ble utviklet underveis i prosessen. Formålet med testen var å dokumentere omfanget på linkene (koblingene) og hvilken kvalitet på dem som konverteringen kunne forvente å gi. Dessuten ønsket vi å anslå hvor store mengder som vi måtte regne med å håndtere manuelt og hvilke krav til ressurser vi måtte sette.

Vi har gått igjennom de ulike mengdene kontrollsystemet har gitt i samarbeid med representanter for de involverte registrene. Et viktig mål var å se om vi maskinelt kunne redusere mengden uten link ved hjelp av navnesøk. Videre at kontroll og godkjenning i størst mulig grad kunne gjøres maskinelt slik at det manuelle arbeidet ble så lite som mulig.

2. Grunlaget

2.1. Involverte registre

MVA-registeret har første prioritet i konverteringen siden nummerserien i manntallet vil være oppbrukt i løpet av 1994. Derfor er det essensielt å få registrert merverdiavgiftsenhetene med identen i Enhetsregisteret, som er Brønnøysundregistrenes foretaksnummer. For å kunne sjekke koblingene mellom MVA og NyFor (Brønnøysunds foretaksregister) må vi gå om BOF (Statistisk sentralbyrås Bedrifts- og Foretaksregister). Dette fordi BOF-enheter har kobling mot MVA- og NyFor-enheter i Statistisk sentralbyrås database, men uten at det her foreløpig er en direkte teknisk relasjon mellom MVA og NyFor. Teststudiet vil i første omgang være todelt, gjennom å se på koblingene mellom:

- 1) BOF og NyFor
- 2) BOF og MVA

Direkte koblinger mellom MVA og NyFor er foretatt hos fylkesskattekontorene, men det gjelder kun enheter registrert etter 1.1. 1992. Det er ønskelig å fange opp de av koblingene her som ikke eksisterer i Statistisk sentralbyrås database. Vi velger å etablere et kontrollsystem også for denne koblingsvarianten. Mulige feilkoblinger må undersøkes mellom:

- 3) MVA og NyFor

Kopien av MVA-registeret i Statistisk sentralbyrås database inneholder kun de MVA-enheter som har tomåneders terminer. Dvs. at de fleste enheter innenfor primær-næringene ikke vil være å finne her fordi disse næringene leverer årsoppgaver dersom ikke annet er avtalt med SKD.

Brønnøysunds interne hjelperegister, Løsøreregisterets Foretaksregister (LøFor), ble i tillegg til NyFor benyttet i navnesøk (se kap. 6).

Tabellen nedenfor viser antall enheter i Sør-Trøndelag fylke registrert i hvert av de tre registrene; BOF, MVA og NyFor:

REGISTER	AKTIVE	ANDRE	TOTALT
BOF pr. 7.10. 1993	10 758	11 820	22 578
MVA pr. 6.10. 1993	9 479	18 659 (slettete)	28 136
NyFor pr. 7.8. 1993	8 902	1 882	10 784

2.2. Valg av geografisk område

Sør-Trøndelag fylke ble valgt av flere årsaker:

- Området er rimelig representativt for landet, dvs. at en kan anta at alle næringene, by og landsbygd vil være representert.
- Grunnlaget er stort nok til å kunne trekke generelle slutninger, men ikke så stort at det blir uoversiktlig og uhåndterlig.

Med BOF som utgangspunkt er det ca. 7 300 linker mot NyFor og ca. 18 600 mot MVA. Et viktig poeng for oss var dessuten at vi allerede hadde etablert kontakt med fylkesskattekontoret i Sør-Trøndelag gjennom arbeidet i Enhetsregisterprosjektet.

2.3. Formålstjenlig innhold i registrene

I det følgende listes opp hvilken informasjon som ble funnet å kunne være nyttig i kontrolløyemed av linkene og som er tilgjengelig i Statistisk sentralbyrås database.

2.3.1. BOF kontra NyFor

BOF	NYFOR
BOFs foretaksnummer	NyFors foretaksnummer
NyFors foretaksnummer	navn
foretaksnavn	postboks
postboks	gateadresse
gateadresse	postnummer
postnummer	kommunennummer
kommunennummer	
sektorkode	
næringskode	meldingsstatus
tilstandskode	eierforhold og foretakstype
eieform	
foretakstype	
foretaksnummerdato	stiftelsesdato
navnedato	

Foretakstype i BOF og NyFor er ikke det samme. I BOF er det en kode for om det er et en- eller flerbedriftsforetak. Mens foretakstype i NyFor beskriver organisasjonsformen.

2.3.2. BOF kontra MVA

BOF	MVA
BOFs foretaksnummer	BOFs foretaksnummer
NyFors foretaksnummer	MVA nummer
foretaksnavn	firmanavn
postboks	postboks
gateadresse	gateadresse
postnummer	postnummer
kommunennummer	kommunennummer
sektorkode	
næringskode	bransje
	hovednæring, binæring
tilstandskode	status, sist innbetalte termin, merknader
eieform	ansvarsforhold
omsetning	omsetning
sysselsetting	
foretakstype	
foretaksnummerdato	registreringsdato, ajourholdsdato
	eieres fødselsnummer (gj. for inntil 4 eiere)
navnedato	

2.3.3. MVA kontra NyFor

MVA	NYFOR
BOFs foretaksnummer	
MVA nummer	NyFors foretaksnummer
firmanavn	navn
post- og gateadresse	gateadresse
postnummer	postnummer
kommunennummer	kommunennummer
bransje	
hovednæring, binæring	
status, sist innbetalte termin,	
merknader	meldingsstatus
ansvarsforhold	eierforhold og foretakstype
registreringsdato, ajourholdsdato	stiftelsesdato
eiers fødselsnummer	eiers fødselsnummer

Dette er, som nevnt innledningsvis, data som er tilgjengelig i Statistisk sentralbyrås database. En ser at det ikke eksisterer noen felles ident her. Kopien av MVA inneholder BOFs foretaksnummer, mens BOF inneholder NyFors foretaksnummer og det er dette som indirekte utgjør koblingene mellom MVA og NyFor i Statistisk sentralbyrås database. Direkte relasjoner mellom MVA og NyFor har ikke Statistisk sentralbyrå i sin database.

Imidlertid har Fylkesskattekontorene i sitt system fra og med 1992 begynt å føre inn NyFors foretaksnummer for de MVA-enheter der dette er oppgitt. I avsnitt 6.1. fremkommer det hvordan vi har tatt hensyn til Fylkesskattekontorenes direkte relasjoner mellom MVA og NyFor.

3. Valgte kontrollkriterier

I dette kapitlet beskriver vi kontrollkriteriene som i flere varianter inngår i datakontrollprogrammet. Hensikten er å kontrollere de etablerte relasjonene mellom registrene slik at vi får skilt ut hvilke som godkjennes og hvilke som må gjennomgås manuelt.

3.1. Eieform

Avvikende eieform i registrene kan være tegn på at det ikke er de samme enheter som er koblet. Men at koblinger kommer ut med feil på eieformkriteriet er ikke ensbetydende med at de er gale. Årsaken vil ofte kunne være at registrene ikke håndterer enhetene likt ved skifte av eieform og/eller at vedlikehold er mangelfullt i ett eller flere av registrene.

I det følgende beskrives hva den maskinelle kontrollen godkjenner som samme eieform/ansvarsforhold/foretakstype/eierforhold i de tre registrene.

3.1.1. BOF kontra NyFor

Her er hva vi anser som sammenfallende eieform i BOF og NyFor.

BOFs	NYFORs	
Eieform	Eierforhold	F.type
01 = ansvarlig selskap	4,5	ANS, DA,PRE
02 = aksjeselskap	1	AS
04 = andelslag, samvirke-, borettslag, P/L osv.	12,13,15	BRL,BBL,BA,GFS,
05 = forening, sparebank, stiftelse	8,14,17	STI,SPA,FLI,NÆR
06 = statlig		
07 = kommunal		
08 = fylkeskommunal		
09 = uoppgitt, ukjent		
10 = enpersonforetak	9	ENK
11 = kommandittselskap	3	KS
12 = annet	(6)	NUF,SÆR,FLI
13 = delt ansvar (ny kode)	5	DA

Eierforhold og Foretakstype er definert i vedlegg s. 36.

I NyFor og BOF registreres ny enhet med nytt identifikasjonsnummer dersom enheten går over til ny eieform. Ved navneendring, endring av komplementar i KS og endring av aksjonær i AS vil enheten stå uendret og beholde sitt identifikasjonsnummer. Det er kun ved utskifting av ansvarlige deltakere i et ANS/DA/PRE at de to registrene ikke har samme praksis. NyFor tildeler ikke nytt nummer selv om samtlige

deltakerne byttes ut. BOF regner det som et eierskifte dersom minimum 50 prosent av de ansvarlige deltakerne går ut til fordel for nye. I praksis har denne regelen ikke alltid blitt fulgt fordi BOF ikke har den nødvendige informasjon om hvor mange som blir byttet ut.

Boligbyggelag er enheter som ikke føres inn i BOF, men dersom det ligger noen gamle inne vil de ha eieform 04.

Enkeltmannsforetak i NyFor er ofte registrert som ansvarlig selskap i BOF. Enhetene i BOF registreres i hovedsak på grunnlag av MVAs registreringsmeldinger. Det kan være flere årsaker til at MVA ofte registrerer ansvarlige selskap der NyFor registrerer enkeltmannsforetak. En årsak er at det skjer endringer som ikke blir registrert til samme tid i de ulike registre. Hovedårsaken er imidlertid at avgiftsmyndighetene ønsker mer enn en skyldner. Dette innebærer den praksis at en virksomhet hvor både mann og kone er involvert registreres i MVA-registret som ansvarlig selskap uavhengig av om det er registrert etter Foretaksloven som enkeltmannsforetak.

Eieform 12 i BOF vil inneholde mye forskjellig, men skulle inneholde alle de eieformer som ikke passer i noen av de øvrige kodene. En del utenlandske med stedlig representant kommer inn her.

Statlige, kommunale og fylkeskommunale enheter registreres ikke i NyFor, dersom de da ikke har en eller annen form for forretningsmessig virksomhet og er organisert iht. selskapslovgivningen.

3.1.2. BOF kontra MVA

Nedenfor listes opp hva vi betrakter som samme eieform i BOF og MVA og dermed kriteriet for likhet. Eieformkodene i NyFor er tatt med for å illustrere sammenhengen mellom de tre registrene.

BOFs	MVAs	NyFors
Eieform	Ansv.forh.	F.type
01 = ansv.selskap	2=ansv.selskap	ANS,DA,PRE,(4,5)
02 = aksjeselskap	3=begr.ansvar	AS, (1)
04 = andelslag,samvirkelag, borettslag,P/L osv.	3=begr.ansvar	BRL,BBL,BA,GFS, (12,13,15)
05 = forening,sparebank, stiftelse	3=begr.ansvar	STI,SPA,FLI,NÆR (8,14,17)
06 = statlig	7=velferdsinst.	
07 = kommunal	5=offentlig virks.	
08 = fylkeskommunal	5=offentlig virks.	
09 = uoppgitt, ukjent		
10 = enpersonforetak	1=eneinnehaver	ENK,(9)
11 = kommandittselskap		KS,(3)
12 = annet	4=utlending, reg. ved representant	NUF,SÆR,FLI,(6)
13 = delt ansvar (ny kode)	2=ANS	DA

Reglene for å etablere ny enhet ved skifte av eieform er i prinsippet de samme i BOF og MVA. Det gjelder også dersom et enkeltmannsforetak får ny eier eller dersom alle ansvarlige deltakere i et ANS/DA/PRE byttes ut. De to registrene er også på linje når de ikke etablerer ny enhet ved: endring av aksjonær i AS, endring av selskapsnavn, endring av et mindretall av de ansvarlige deltakerne i et ANS/DA/PRE og endring av komplementar i KS.

BOF vil alltid etablere en ny foretaksenhet ved skifte av eieform. For MVA gjelder dette i prinsippet, men ikke alltid i praksis. Avgiftsmyndighetene vil følge opp personen bak avgiftsplikten, dvs. et fødselsnummer er ofte knyttet opp til ett og samme MVA-nummer selv om selskapet ikke er det samme som ved tildelingen av

MVA-nummeret. Derfor vil f.eks. et eierskifte fra ANS til ENK ofte ikke gi en ny momsenhet. Og en som starter ny virksomhet etter å ha avvirket en annen virksomhet vil kunne beholde det gamle MVA-nummeret. Det er videre noe forskjellig praksis ved de ulike Fylkesskattekontorene.

MVA vil imidlertid ikke etablere noen ny enhet dersom de får beskjed om eierskiftet lenge etter at det har kommet inn nye eiere. Det er også eksempler på at eierskifte registreres i rett tid, men at MVA-nummeret likevel beholdes. Dette gjelder f.eks. ANS hvor en av eierne går ut og det da gjenstår som et ENK.

Kode 0 i feltet for ansvarsforhold i MVA-registeret gir ikke informasjon om eieform. Enheter med denne koden velger vi å kjøre ut særskilt slik at de kan kontrolleres hos fylkesskattekontorene. Mange av disse er sannsynligvis eldre enheter, fra før 1973-74, dvs. fra før kodene for ansvarsforhold ble innført i MVA.

3.1.3. MVA kontra NyFor

Her er hva vi aksepterer som tilsvarende eieform i MVA og NyFor.

MVAs Ansvarsforhold	NYFORs Eierforhold	Foretakstype
0 = dersom ansv.forh. er ukjent		
1 = eneinnehaver	9	ENK
2 = ansvarlig selskap	4,5	ANS, DA, PRE
3 = begrenset ansvar	1, 8, 12, 13, 14, 15, 17	AS, BRL, BBL, BA, GFS, STI, SPA, FLI, NÆR
4 = utlending, registrert ved representant	6	NUF, SÆR, FLI
5 = offentlig virksomhet		
6 = konkursbo	<i>tilsv. meldingsstatus = S, P, M eller B*</i>	
7 = velferdsinstitusjoner	8, 14, 17	STI, SPA, FLI, NÆR

* S = slettet, P = meldt oppløst, M = meldt konkurs, B = meldt oppløst og konkurs.

I prinsippet skal også enheter som skifter eieform etableres som nye enheter. Begge registrene etablerer ny enhet med nytt identifikasjonsnummer til enkeltmannsforetak (ENK) som får ny eier. Men praksis ved fylkesskattekontorene er oftere å la MVA-nummer følge personen ved dennes fødselsnummer. Dette betyr at eiere ofte tar med seg gammelt MVA-nummer til ny virksomhet og/eller ved endring av eieform.

Registrene etablerer ikke ny enhet ved endring av :

- enhetens navn,
- aksjonær i aksjeselskap (AS),
- et mindretall av deltakerne i ANS/DA/PRE,
- komplementar i kommandittselskap (KS).

Dersom alle ansvarlige deltakere i et ANS/DA/PRE byttes ut vil fylkesskattekontorene se på dette som ny enhet, mens Foretaksregisteret vil beholde foretaket.

Når det gjelder ansvarsforhold 6 i MVA, konkursbo, er sammenligningen de to registrene imellom noe problematisk. Konkursbo er ikke med i NyFor, men selskaper som er under konkursforhandlinger vil være der. Siden det ikke er noen eierforholdskode eller foretakstype for konkursbo i NyFor benyttes meldingsstatus i stedet.

3.2. Kommunenummer

Dersom relaterte enheter ikke har adresse til den samme kommunen kan det være et signal om at koblingen er gal. *Er imidlertid kommunekoden den samme i begge registrene, er dette kontrollkriteriet oppfylt.*

Det kan være tilfeller hvor kommunenumrene er ulike uten at det er snakk om feil link :

- privatadressen kan være ført opp i et av registrene slik at kommunenumrene er ulike.
- det kan være snakk om adresseendringer som ikke er registrert i begge registrene.
- postboks og beliggenhetsadresse kan være i ulik kommune og således være årsaken til at samme enhet er registrert med forskjellig kommunenummer i ulike registre.

3.2.1. BOF kontra NyFor

Dersom verken foretaket eller en av foretakets bedrifter i BOF er i samme kommune som den tilknyttede NyFor-enhet må linken undersøkes nærmere.

Hvis BOF-foretaket har ulikt kommunenummer leter dataprogrammet etter NyFor-enhetens kommunenummer blant bedriftene i BOF-foretaket. (Bedriftens navn må i tillegg stemme, se avsnitt 4.3.1. og 4.3.2.). Fordi hvert BOF-foretak kun kan ha ett NyFor-nummer knyttet til seg i Statistisk sentralbyrås database stilles det tilleggskrav ved bedriftssøk. Kun aktive NyFor-foretak (meldingsstatus I) mot aktive bedrifter (tilstand 9) kan oppfylle kriteriet for lik kommune.

Det er tilfeller hvor et aktivt BOF-foretak er koblet til et slettet NyFor-foretak som igjen stemmer overens med en slettet bedrift, men ikke BOF-foretaket. Uten de ovenfor nevnte tilstandskrav ville en slik kobling blitt godkjent. Når det finnes et annet aktivt NyFor-foretak som burde vært koblet til BOF-foretaket, er det avgjørende at den eksisterende koblingen kommer ut som feil i kontrollen slik at den aktive enheten blir koblet rett.

3.2.2. BOF kontra MVA

På samme måte som i 4.2.1. sjekkes kommunenummeret først i BOF-foretaket, og dersom det ikke blir klaff her forsøkes det med bedriftssøk.

I motsetning til hva som i dag gjelder for NyFor-enheter i Statistisk sentralbyrås database, kan *mange* MVA-enheter knyttes til *ett* BOF-foretak. Det er derfor ingen grunn til å ta med tilstandskrav ved søk på bedrift fordi vi ikke risikerer å miste informasjon om "bedre" eller "ferskere" MVA-enheter som skulle vært koblet (Jf. 3.2.1.).

3.2.3. MVA kontra NyFor

MVA-enhetens kommunenummer kontrolleres mot NyFor-foretakets. Ved avvik skal linkene kontrolleres manuelt.

3.3. Navnelikhet

At navnet på to relaterte enheter er forskjellig kan være en indikasjon på at det ikke er samme enhet. Men ulike registre har ofte ikke eksakt identisk navn på samme enhet. Det kan være fordi den som har meldt inn navnet ikke har oppgitt helt det samme til ulike registre, det kan være ulik standard for hvordan navn skal registreres eller rett og slett skrivefeil.

I Statistisk sentralbyrå blir navn redigert noe ved en egen navnesorteringsrutine. Denne kjøres på tapene med tilganger og endringer som kommer regelmessig fra Brønnøysundregistrene (NyFor) og Skattedirektoratet (MVA). Den fjerner tegnene; apostrof, stjerne og punktum, legger en blank posisjon mellom elementer som består av bokstaver og tall, gjennomfører store bokstaver og påfører en SOUNDEX-kode (dvs. lydkode).

Det er gjort følgende tilpasninger i den maskinelle kontrollen for å unngå at svært små ulikheter skal slå ut som navnefeil:

- Programmet sørger for at BOF-navn med bindestrek også leses uten bindestrek og/eller som et sammendratt ord dersom det ikke er identisk med det andre registerets navn.
- I tillegg vil de første to ordene i BOF-navnet leses som ett ord for å se om de to navnene da blir like.
- A/S, AS og v/ tas ikke i betraktning i navnesammenligningen.
- Uansett plassering av leddene i navnet vil de telles med som like ledd dersom de er å finne i MVA-navnet eller NyFor-navnet, alt etter hvilken kobling som sjekkes.

Navneendringer som ikke er registrert i ett av de to registrene en ser på koblingene mellom, vil her komme ut som feil. Dersom endringen er utført i BOF vil man manuelt kunne se det i databasen til Statistisk sentralbyrå. Men i kopien av MVA-registeret og NyFor-registeret er det ikke knyttet historiske data til enhetene.

Kravet som stilles er at ett ledd i navnet skal være likt, uansett leddets plassering i navnet.

3.3.1. BOF kontra NyFor

"Navn", og ikke "firmanavn", er det i Brønnøysundregistrene som er mest sammenlignbart med BOFs navneform. Praksis for registrering av "navn" er blant annet:

- genitiv s faller bort i etternavn og stedsnavn, men beholdes i fornavn.
- "studio" og "galleri" kommer oftest først i navnet.

På samme vis som for kommunenummerkriteriet søkes det også etter navnelikhet på bedriftsenhetene. Dette for å fange opp de tilfellene hvor BOF-foretaket har annet navn enn NyFor-enheten, men hvor en av bedriftene har tilsvarende navn som den tilkoblede enheten. De to kriteriene kan ikke sees uavhengig av hverandre. Både kommune og navn på bedriften må stemme overens med NyFor-enheten. Dvs. en kan ikke godkjenne en kobling hvor bedriften er plassert i en annen kommune enn NyFor-enheten selv om navnet skulle være det samme, og BOF-foretaket, med annet navn, plassert i samme kommune.

3.3.2. BOF kontra MVA

I MVA-navnet kan siste ledd være en tilleggsopplysning i form av visse koder, f.eks. KB=konkursbo. En slik kode blir i Statistisk sentralbyrås navnesorteringsrutine flyttet til et eget felt utenfor navnefeltet.

En ønsker også her å fange opp de tilfellene hvor BOF-foretaket har et annet navn enn moms-enheten, men hvor en av bedriftene har tilsvarende navn som den tilkoblede enheten. Klaffer ikke foretakets navn med den merverdiavgiftspliktige, leter programmet gjennom foretakets bedrifter. Kommune og navn som stemmer overens med MVA-enheten, må være å finne på samme BOF-enhet for å kunne aksepteres.

3.3.3. MVA kontra NyFor

Det er "firmanavn" i MVA og "navn" fra Brønnøysundregistrene som sammenlignes.

4. Mulige kontrollkriterier

Kapitlet omtaler andre opplysninger som ble vurdert tatt med i den maskinelle kontrollen, men som er utelatt blant kriteriene. Disse vil imidlertid være til hjelp i den manuelle gjennomgangen.

4.1. Tilstandskode

Er en enhet opphørt i et register, men ikke i et annet kan dette være en indikasjon på feil. Det kan være at enheten skulle vært opphørt i begge registrene, at den er opphørt ved en feil eller at det ikke er rette "makkere".

Forskjellige tidspunkt for oppdatering vil i perioder gi ulike tilstander for en og samme enhet alt etter hvilket register man ser i. Men tilstandskode vil kunne være satt ut ifra ulike formål, slik at det kan være riktig at samme enhet har ulik tilstand i ulike registre. Det er f.eks. ikke nødvendigvis galt at en slettet moms enhet er knyttet til en aktiv BOF-enhet. Forklaringen kan være så enkel som at enheten har sluttet med den momspliktige virksomheten og derfor er slettet i MVA-registeret, men har annen virksomhet som gjør at den er aktiv i BOF. Derfor har vi valgt ikke å benytte tilstandskoden som ett av kontrollkriteriene.

Imidlertid ble mengdene til manuell kontroll, BOF-NyFor og BOF-MVA, fordelt etter om tilstanden var tilsvarende eller ikke. Dette ble gjort bl.a. for å forsikre oss om at vi hadde gjort rett i å utelate tilstand som kontrollkriterium.

Koblinger hvor enhetene var ute av drift i begge registrene ble utelatt av mengden som skulle kontrolleres for å unngå å bruke tid på disse (siden de ikke skulle omfattes av konverteringen). Disse linkene bruker vi ikke tid på. Dette gjaldt kun de som har tilstand 4(opphørt) i BOF og som har S(slettet) i MVA eller eventuelt S/P/M/B (se definisjoner under punkt 5.1.1.) i NyFor.

Nedenfor kommer oversikter over tilstandskode/meldingsstatus i de tre registrene, NyFor og MVA sett i forhold til BOF.

4.1.1. BOF kontra NyFor

BOFs tilstandskode	NYFORSs meldingsstatus
0 = under oppbygning	
2 = foretak med alle bedrifter ute av drift	S, M, B, P
3 = proforma foretak	
4 = opphørte foretak	S, M, B, P
7 = fjernet som feilinnført	
8 = fjernet som dublett	
9 = i ordinær aktivitet	I

NyFors meldingsstatus:

I = Innført
 S = Slettet
 P = meldt oppløst
 M = meldt konkurs
 B = meldt oppløst og meldt konkurs

4.1.2. BOF kontra MVA

BOFs

tilstandskode

0 = under oppbygning
 2 = foretak med alle bedrifter ute av drift
 3 = proforma foretak
 4 = opphørte foretak

 7 = fjernet som feilinnført
 8 = fjernet som dublett
 9 = i ordinær aktivitet

MVAs status, sist innbetalte termin og merknader

S=slettet, siste 2 md. termin dato, A*

 S=slettet, Siste 2 md. termin dato,
 DB=dødsbo, KB=
 konkursbo, EG=engangsinnført

 A=aktiv, R=reaktivisert

A* - tilstand 2 aksepteres som aktiv i MVA dersom momsnetten ikke har omsetning.

"Status", "SKDs merking", "siste tomåneders termin" og "omsetning" er felter i MVA-registeret som kan si oss noe om momsnettenens tilstand. Statusen er "aktiv", "reaktivisert" eller "slettet". De to førstnevnte passer med "i ordinær drift" i BOF. "Slettet" tilsvarer som regel "opphørt"/"alle bedrifter ute av drift" i BOF. Dersom det er dato i feltet "siste 2 md. termin" er momsinnbetalingene opphørt og en kan regne denne som en slettet enhet. Konkursbo og dødsbo er avgjort opphørte foretak. Engangsinnført skal ikke registreres i BOF.

Koden for reaktivisert ble først tatt i bruk for tre til fire år siden, tidligere ble en reaktivisert påført A for aktiv. Dato for første tomåneders termin er første termin etter reaktiviseringen.

4.2. Næringskode

Både fylkesskattekontorene og Statistisk sentralbyrå har næringskode i registrene, noe som kunne være et mulig sammenligningspunkt mellom MVA og BOF. Hvis næringskoden er den samme i begge registrene, kunne dette stadfeste at det var samme enhet. Er derimot næringskoden ulik behøver ikke det å bety at linken er feil. Det kan like gjerne være at samme enhet har virksomhet innenfor flere næringer og at den største virksomheten ikke er momspliktig. Næringskoder er derfor bare et hjelpemiddel i manuell kontrolløymed.

4.3. Adresse

Adressene i NyFor og MVA blir omformet, så godt som det lar seg gjøre, til Statistisk sentralbyrås standard. (Dette blir gjort fra oppdateringsfilene fra Skattedirektoratet og Brønnøysundregistrene). Likevel er verken gateadresse eller postboksadresse brukbare som maskinelle kontrollkriterier. Små skriveforskjeller vil medføre at de samme adressene ikke aksepteres som de samme i en maskinell kontroll. Registrering av adresser blir gjort ved hjelp av ulike kilder i de ulike registrene. Dette er i tillegg manuelle rutiner. Dermed er det ikke til å unngå at det oppstår flere varianter av samme adresse. En maskinell kontroll vil ikke klare å skille mellom ulike skrivemåter og faktiske adresseforskjeller. Det er ikke godt å si om endringer av gateadresse eller postboksadresse blir registrert, og om det i tilfelle blir gjort til samme tid i de ulike registrene.

I den manuelle kontrollen vil gateadressen komme til nytte. Postboks er ikke like greit å benytte seg av. Men dersom postboksnummeret skulle være det samme, kan en være temmelig sikker på at koblingen er riktig.

4.4. Dato

I BOF vil en finne dato for når enheten, navnet, tilstanden, adressen osv. ble registrert eller sist endret. MVA har, foruten en ajourholdsdato, også dato for når den avgiftspliktige ble registrert. NyFor har informasjon om når foretaket ble stiftet. I tillegg er det dato for når enheten er gitt navn og adresse i Foretaksregisteret, eller eventuelt dato for siste endring på navn og/eller adresse. (Statistisk sentralbyrå burde hatt dato for når registreringen er foretatt i NyFor, men i det feltet får vi dato for hvilken måned enhetens endring/tilgang er kjørt ut på tape for utsendelse til Statistisk sentralbyrå).

Datoene er ofte ikke sammenlignbare og det er i praksis vanskelig å gi standard retningslinjer. F.eks. er stiftelsesdatoen i NyFor den dagen beslutningen tas om å stifte et selskap. Den kan imidlertid være lik navneendringsdato fra de gamle handelsregistrene. Her var praksis å registrere ny enhet ved navneendring. Foretaket med det gamle navnet ble slettet i registeret.

Registreringsdatoen i BOF er dato for når enheten blir registrert i BOF og får tildelt BOF-nummer. Et foretak kan være stiftet lenge før det blir registrert i BOF. Oppdateringer og nyregistreringer i BOF kan komme fra forskjellige kilder slik at datoer kan sprike begge veier i tid i forhold til andre registre. De ulike registrene oppdateres heller ikke nødvendigvis til samme tid. Nyten av datoer må derfor vurderes i hvert enkelt tilfelle, og vil således bare være til hjelp i gjennomgangen av de linkene som ikke blir godkjent i den maskinelle kontrollen. En bør f.eks. sjekke at det ikke er kobling mellom en enhet som er stiftet i NyFor lenge etter at den er registrert i BOF. Dette kan imidlertid være tilfelle for enheter som var registrert i de gamle handelsregistrene (jf. forrige avsnitt om praksis ved navneendring i Handelsregistrene).

4.5. Fødselsnummer

Statistisk sentralbyrås database har kopi av MVAs eierregister, hvor eiere er ført opp med navn og fødselsnummer. Disponent, styreformann, styremedlem, representant og fullmektiget er også registrert i dette registeret. Egne koder identifiserer hvilken funksjon vedkommende har. Eier, som ikke er personlig, kan være registrert med MVA-nummer, men det gjelder ikke mange.

Registerkopien blir tilført nye eiere fra Skattedirektoratet hver fjortende dag. Dessuten gir Brønnøysundregistrene (NyFor) månedlig informasjon om nyregistrerte eiere av personlig eide selskap med inntil fire eiere og ansvarlige selskap. Enkeltmannsforetak i NyFor er påført innehaverens fødselsnummer.

Forsøk på koblinger blir gjort maskinelt i Statistisk sentralbyrå hver måned på oppdateringene som kommer fra Brønnøysund. En av de maskinelle forsøkene går om eierregisteret. Dersom det er samme eier bak enheter i to ulike registre er det sannsynlig at det dreier seg om identiske enheter. Det har vist seg at denne metoden gir meget liten gevinst av såkalte helt sikre koblinger.

I konverteringen vil bruk av fødselsnummer alene, som selvstendig kontrollkriterium, trolig behandle mange riktige linker som feil pga. sprik i oppdateringen av de to registrene. Det som kunne vært en mulighet var å bruke fødselsnummer på eiere av enkeltmannsforetak til å godkjenne linker som er kommet ut med feil i datakontrollprogrammet. Kriteriet benyttes da som et godkjenningskriterium dersom det ikke blir klaff på eieform, kommune eller navn. Dette ble forsøkt på den potensielle feilmengden i koblingene mellom BOF og NyFor. Ca. 23 prosent av disse fikk klaff på fødselsnummer. Det ser ut for at de fleste av disse er riktige koblinger, men de har alle oppdaterings-mangler. Dersom fødselsnummer tas med som et godkjenningskriterium vil vi miste informasjon om opprettinger av datainnhold som burde vært gjort. Dessuten vil ikke alle linkene som slik sett blir godkjent være riktige og dermed øker feilprosenten på de vi godkjenner maskinelt. Delprosjektet har derfor valgt ikke å benytte eiers fødselsnummer som kriterium i den maskinelle kontrollen.

5. MVA/NyFor-enheter uten link via BOF

Det er en del MVA-enheter som ikke har NyFors foretaksnummer knyttet til seg via BOF. Mange av disse skal heller ikke være registrert i BOF, mens andre skulle vært det. I tillegg kan en ikke være sikker på at alle mulige linker mellom registrene er opprettet.

I dette kapitlet beskrives hva som er gjort for å fange opp flest mulig av MVA-enhetene i Sør-Trøndelag fylke som er registrert i Foretaksregisteret (NyFor). I tillegg til å få testet ut metoden er hovedmålet selvsagt å unngå dobbelnummerering, siden alle MVA-enheter som en ikke finner i NyFor vil bli tildelt nytt nummer.

5.1. NyFor-nummer registrert etter 1.1. 1992 ved Fylkesskattekontorene

Statistisk sentralbyrå har mottatt MVA-enheter som er registrert etter 1.1. 1992 med NyFors foretaksnummer. Uttaket er gjort ved Skattedirektoratets Grimstadkontor pr. 27.11. 1993. Her er tatt vare på og kontrollert de koblingene i Sør-Trøndelag fylke som ikke er etablert via BOF. Resultatene av den maskinelle kontrollen blir presentert under kapittel syv.

5.2. Navnematch

5.2.1. MVA kontra NyFor

Aktive og slettede MVA-enheter (i fylke 16) som ikke hadde link til NyFor verken via BOF eller direkte med foretaksnummer fra Fylkesskattekontoret, ble forsøkt funnet i NyFor-kopien i Statistisk sentralbyrås database. MVA-enheter med link til BOF ble ikke tatt med. NyFor-enhetene det ble søkt mot var alle i fylket og det ble ikke satt noe krav til tilstand for disse.

Navnesøket baserer seg på lyd-koder som ble knyttet til de aktuelle MVA- og NyFor-enhetene. Dersom lyd-koden er identisk tolker programmet det som likt navn. I tillegg kreves det at enhetene er i samme kommune før de kommer ut som koblingsforslag.

5.2.2. MVA kontra LøFor

Etter at fylkesskattekontorenes foretaksnummer og navnematchforslagene overfor NyFor var tatt hensyn til, var det fortsatt mange *aktive* MVA-enheter uten noen form for link. De som ikke hadde forslag til foretaksnummer ble forsøkt matchet mot en tape fra Brønnøysundregistrene med LøFor-enheter. Som for NyFor opererte vi fortsatt i fylke 16, og stilte ingen krav til tilstand for LøFor-enhetene. Men kriteriet for likt kommunenummer tok vi imidlertid ikke med her.

6. Datakontrollprogrammene

De maskinelle programmene omfatter kontrollkriteriene *navn, eieform og kommune*. For at en kobling skal godkjennes (konverteres) må *ett ledd i navnet, kommune og eieform være tilsvarende for de to linkete enhetene*. Som omtalt i kapittel fire prøver programmene med en del utvalgte redigeringer dersom kriteriene ikke oppfylles i utgangspunktet, f.eks. sammendragning av ord i navnet og søk i bedriftene til foretaket i BOF.

Koblingene kontrolleres mellom to og to registre av gangen. Dataprogrammet kjøres på koblingene mellom BOF og NyFor, mellom BOF og MVA og mellom MVA og NyFor.

I teststudiet gjelder spesielt at MVA-enhetene og NyFor-enhetene skal ha adresse i Sør-Trøndelag fylke før de inngår i kontrollen.

6.1. Omfang til maskinell kontroll

Koblinger der enhetene er opphørt i begge registrene er holdt utenfor. Disse linkene bruker vi ikke tid på fordi de ikke er en del av konverteringen. Dette gjelder enheter som er ute av drift, som pga. tilstanden heller ikke vil være oppdaterte i BOF og derfor ville vært vanskelige å finne ut av.

ANTALL KOBLINGER			
REGISTER	TOTALT	OPPHØRT I BEGGE REGISTRE	TIL KONTROLL
BOF OG NYFOR	7 342	1 047	6 295
BOF OG MVA	18 641	9 000	9 641
MVA OG NYFOR	951	0	194*

Datoene for kontrollkjøringene: 23.9.93 (BOF-NyFor), 19.8.93 (BOF-MVA) og --/11.93(MVA-NyFor).

6.295 relasjoner mellom Statistisk sentralbyrås og Brønnøysunds foretaksregister kommer til maskinell kontroll.

9.641 relasjoner mellom Statistisk sentralbyrås foretaksregister og Skattedirektoratets avgiftsmanntall kommer til maskinell kontroll.

Koblingene mellom MVA og NyFor som vi har fått fra Skattedirektoratet er alle MVA-enheter registrert etter 1.1. 1992 der NyFor-nummeret er påført av fylkesskattekontorene. 757 av de 951 var allerede linket via BOF og inngår derfor i de 6 295 som skal til kontroll.

*Av de 194 MVA-enhetene som har NyFor-nummer fra fylkesskattekontoret i Sør-Trøndelag, er det 25 som ikke befinner seg i Statistisk sentralbyrås database. Dermed blir det 169 som går videre til maskinell kontroll.

6.2. Resultater av maskinell kontroll

Resultatene av de maskinelle kontrollene er oppsummert i tabellene nedenfor:

	ANTALL KOBLINGER		
REGISTER	KONTROLLERT	GODKJENT	IKKE-GODKJENT
BOF OG NYFOR	6 295 100 prosent	5 914 93,9 prosent	381 6,1 prosent
BOF OG MVA	9 641 100 prosent	8 735 90,6 prosent	906 9,4 prosent
MVA OG NYFOR	169 100 prosent	159 94,1 prosent	10 5,9 prosent

Alle prosenttall er avrundet til en desimal.

En ser at henholdsvis ca. seks, ni og seks prosent av koblingene ikke blir godkjent av datakontrollprogrammet. Dette betyr at disse har enten ulikt navn, ulik kommune eller ulik eieform - eller en kombinasjon av to eller tre av feiltypene. De aller fleste viste seg å ha bare en type feil. Kun en relasjon til BOF slo ut med feil på alle tre kriteriene, mens det tilsvarende gjelder for tre av koblingene mellom MVA og NyFor. Disse tre koblingene blir omtalt nærmere under neste punkt.

Det er grunn til å merke seg at for en del av de godkjente koblingene har enhetene ulik tilstand. 15,0 prosent (889) av de godkjente koblingene mot NyFor har ulik tilstand. Tilsvarende tall mot MVA er 18,6 prosent (1 624). Viser til avsnitt 5.1. om tilstand som kriterium.

6.2.1. Fordelingen av de ikke-godkjente koblingene etter kontrollkriterier

Tabellen beskriver fordelingen av mulige feilkoblinger fordelt etter de tre kontrollkriteriene:

	ANTALL KOBLINGER			
REGISTER	IKKE-GODKJENT	NAVN	KOMMUNE	EIEFORM
BOF OG NYFOR	381 100 prosent	142 37 prosent	155 41 prosent	127 33 prosent
BOF OG MVA	906 100 prosent	280 31 prosent	291 32 prosent	541 60 prosent
MVA OG NYFOR	10 100 prosent	6 60 prosent	9 90 prosent	4 40 prosent

Alle prosenttall er avrundet til nærmeste hele prosent.

En del av de maskinelt ikke-godkjente koblingene hadde mer enn ett kriterium som ikke stemte. Dvs. at koblingene oppfyller ikke kravene til ett *eller flere* av kriteriene. Den prosentvise fordelingen av koblingene på kriteriene vil dermed naturlig nok ikke summere seg opp til hundre. Dette gjelder også for de tabellene i kapittel 8 og 9 hvor kontrollkriteriene inngår.

I forholdet mellom NYFOR og BOF er det kommune som har den høyeste feilprosenten, men navn og eieform kommer ikke langt etter. Eieform ser ut til å være et større problem for BOF i forhold til MVA enn det det er i forhold til NYFOR. Den største andelen av de ikke-godkjente koblingene mellom MVA og BOF utgjøres av de med ulik eieform. For relasjonene mellom MVA og NyFor kan det se ut som at kommunenummer er den største feilkilden, men her vil det være vanskelig å trekke generelle slutninger siden det er så få koblinger.

MVA og NyFor har tre koblinger som kommer ut med feil på alle tre kriteriene. Disse er i realiteten ikke kontrollert fordi NyFor-enhetene ikke er å finne i registerkopien i Statistisk sentralbyrås database, dermed er det ingen data å kontrollere mot. Årsaken er at en enda ikke er kommet med fra Brønnøysund og at to er enheter som er nektet registrert i NyFor. I det sistnevnte tilfellet vil det være snakk om LøFor-nummer og koblingene vil da mest sannsynlig være fanget opp i navnematchen mot LøFor og bli kontrollert som en del av det manuelle arbeidet.

Navnekriteriet er det minst strenge av de tre kriteriene. Vi krever som tidligere nevnt ikke mer enn ett likt ledd i navnet. Imidlertid er det hele 64 prosent (6 128 stk.) av de kontrollerte BOF-MVA-koblingene som har eksakt likt navn. BOF-navnet er ofte hentet fra MVA-meldinger og vil da være helt likt (forutsatt at det ikke er foretatt noen navneendringer i mellomtiden). Det er bare 27 prosent (1 684) av de kontrollerte BOF-NYFOR-koblingene som har eksakt likt navn.

6.2.2. Omfang av fellesmengden i MVA og NyFor

Opptellingene som er nyttet i beregningene i dette avsnittet er status i registrene pr. juli 1993.

I Sør-Trøndelag fylke var det :

6 295	Nyforenheter med kobling til BOF
- 274	Nyforenheter med kobling til BOF-enheter uten link til MVA
<hr/>	
6 021	Nyforenheter med kobling til BOF-enheter med link også til MVA

M.a.o. er det 96 prosent av NyFor-enhetene i Sør-Trøndelag fylke som er knyttet til BOF, som også har kobling til MVA via BOF. Det betyr at storparten av NyFor-enhetene, som er linket til BOF i Statistisk sentralbyrås database, har videre tilknytning til MVA-enheter. Men det innebærer også at det kun er 6 021 MVA-enheter som her har NyFor-nummer. Dette utgjør 62 prosent av BOF-MVA koblingene av enheter i Sør-Trøndelag fylke.

Det totale antallet enheter i MVA er betydelig større enn i NyFor, fordi ikke alle skal være registrert i Brønnøysund. Nærmere 90 prosent av enhetene i MVA-registeret har link til BOF-foretak. Ca. 66 prosent av NyFor-foretakene har link til BOF-foretak. På denne bakgrunn må en kunne anta at opplegget sørger for at storparten av MVA-enhetene tildeles NyFors foretaksnummer via BOF.

6.3. Omfang til manuell kontroll i teststudiet

Alle ikke-godkjente og et utvalg av de godkjente koblingene mellom BOF og henholdsvis NyFor og MVA ble gjennomgått manuelt. I tillegg kom navnematchmengdene. Arbeidet ble delt mellom Brønnøysundregistrene og Fylkesskattekontoret i Sør-Trøndelag.

Til Brønnøysund ble det sendt i alt 450 godkjente koblinger og til Sør-Trøndelag totalt 498. I tillegg fikk hver av dem de maskinelt ikke-godkjente linkene relatert til "eget" register. Koblingsforslagene basert på navnematchene, totalt 773, ble imidlertid alle sendt til Brønnøysund.

7. Tilbakemelding fra Brønnøysundregistrene

7.1. Maskinelt godkjente koblinger

7.1.1. Resultater

Tre utvalg av maskinelt godkjente linker mellom BOF og NyFor ble kontrollert i Brønnøysund. 300 koblinger gjaldt kun aktive enheter (etter definisjonene under avsnittet "tilstand"), og 150 koblinger omfattet enheter som hadde ulik tilstand. Utvalgene fordelte seg slik :

- 1) 150 koblinger med mer enn ett likt ledd i navnet.
- 2) 150 koblinger med kun ett likt ledd i navnet.
- 3) 150 koblinger mellom enheter med ulik tilstand :

Disse tre mengdene fordelte seg slik på riktig og feil kobling:

MENGDE	1)	2)	3)	TOTALT	PROSENT
RIKTIG	148	150	146	444	98,7 prosent
FEIL	2	0	4	6	1,3 prosent
TOTALT	150	150	150	450	100,0 prosent

Mengde 2 inneholdt ingen gale koblinger. De fleste her har navn med bare ett ledd. Dette betyr at ved eksakt likt navn, lik eieform og lik kommune vil de to enhetene være en og samme enhet.

Dersom det er flere enn ett likt ledd i navnet vil 1,3 prosent av relasjonene være feil. Her vil det også være mange med identiske navn, men i motsetning til de med bare ett ledd i navnet er det feilkoblinger her. Nå er det bare snakk om to feil så det kan være snakk om en ren tilfeldighet at de er i mengde 1 og ikke i mengde 2.

Når enhetene har ulik tilstand er usikkerheten størst. Tilnærmet vil 2,7 prosent av disse være gale koblinger. Imidlertid, som tidligere nevnt, er det bare ca. 15 prosent av de godkjente koblingene hvor BOF- og NyFor-enheterne ikke har lik tilstand.

7.1.2. Feiltyper

I mengde 1 er det to feilkoblinger. Den ene skyldes at det i BOF ikke er registrert en fusjon, mens den er registrert i Brønnøysund. To selskaper har gått sammen til ett med et av de opprinnelige selskaperens navn. Navne-likheten fører i dette tilfellet til godkjenning av en feil kobling.

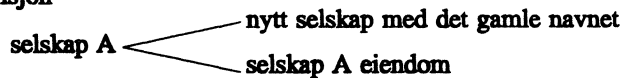
Den andre skyldes en fisjon. BOF har ikke registrert at det opprinnelige selskapet er skilt ut som et eiendoms-selskap, mens resten av driften er lagt til et nytt selskap. Det nye selskapet har fortsatt med det gamle navnet og dermed fører navnelikheten til godkjenning av koblingen. Dvs. at det i BOF ikke har vært endret navn på det gamle selskapet, slik at det har blitt koblet til det nye i NyFor.

Det følgende er ment og illustrere de to feiltypene i mengde 1 :

Fusjon



Fisjon



I mengde 3 var det fire feilkoblinger. Tre av dem kan konkluderes som feil fordi enhetene i BOF er registrert før de sammenkoblede NyFor-enhetene er stiftet.

Den siste feilkoblingen består av et opphørt BOF-foretak og et aktivt NyFor-foretak. Det er registrert et annet BOF-foretak som passer med NyFor-foretaket. Den opphørte BOF-enheten er registrert i LøFor (Løsøreregisterets foretaksregister).

7.1.3. Konklusjon

Det er ingen enkelte typer feil som skiller seg ut. Derfor er det heller ikke noen måte vi får fanget opp de gale koblingene uten å måtte ta med mange flere riktige til manuell kontroll. En feilprosent på 1,3 er akseptabel for de maskinelt godkjente koblingene.

7.2. Maskinelt ikke-godkjente koblinger

7.2.1. Resultater

De koblingene som kom ut som potensielt gale i datakontrollprogrammet ble først manuelt gjennomgått i Statistisk sentralbyrå. 214 av disse så ut til å være riktige og ble sendt på egen liste til Brønnøysund (liste 1). Etter en manuell kontroll der, ble fordelingen av disse seende slik ut:

LISTE 1	ANTALL KOBLINGER	PROSENT
RIKTIG	185	86,4 prosent
FEIL	29	13,6 prosent
TOTALT	214	100,0 prosent

De resterende 167 potensielle feilkoblingene (liste 2) fant Brønnøysundregistrene å kunne fordele slik :

LISTE 2	ANTALL KOBLINGER	PROSENT
RIKTIG	117	70,1 prosent
FEIL	50	29,9 prosent
TOTALT	167	100,0 prosent

En ser at av de riktige relasjoner Statistisk sentralbyrå har plassert i liste 1 er Brønnøysundregistrene ikke enig i 13,6 prosent av tilfellene. Årsaken til avviket er nok at Statistisk sentralbyrå har vært bevisst at dette skulle sjekkes i etterhånd og dermed "tatt godt i". Dessuten har Brønnøysundregistrene tilgang til annen informasjon enn Statistisk sentralbyrå (og motsatt).

I liste 2 var det nærmere 30 prosent som ikke skulle vært knyttet sammen, men størstedelen var også her riktig til tross for at Statistisk sentralbyrå ikke fant å kunne plassere disse i liste 1. En stor andel av de relasjonene som det maskinelle kontrollprogrammet ikke godkjenner består altså av BOF-enheter med riktig NyFor-foretaksnummer.

I alt vil koblingene som kontrollprogrammet ikke godkjente fordele seg som følger :

LISTE 1 og 2	ANTALL KOBLINGER	PROSENT
RIKTIG	302	79,3 prosent
FEIL	79	20,7 prosent
TOTALT	381	100,0 prosent

Opp mot 80 prosent av de 381 potensielle feilkoblingene viste seg faktisk å være riktige. Et sentralt spørsmål var derfor hva som kunne kjennetegne disse, m.a.o. er det noen fellestrekk? Ville det være mulig å fange opp noen av disse som godkjente i den maskinelle kontrollen uten å miste de som ikke skulle vært knyttet sammen? De ca. 20 prosent som skulle hatt andre foretaksnummer er de vi ideelt sett ønsker skal stå igjen som feil etter en maskinell kontroll. Det vil derfor være interessant å vite hvilke av kontrollkriteriene disse er kommet ut på og hvilke trekk det er ved enhetene som gjør at de har blitt linket sammen. De følgende fire punktene behandler disse spørsmålene.

7.2.2. *Beskrivelse av de som kom ut som feil i datakontrollprogrammet, men som viste seg å være riktige koblinger*

Nedenfor følger en oversikt over riktige koblinger i liste 1 og 2 :

1 og 2	ikke oppd. navn	ikke oppd. adr.	skrivefeil	ulik skrivemåte	ulik eieform	privat adr.	ulik tilst.	SUM
navn ikke oppdatert	56	29			4	5	12	106
adr. ikke oppdatert		80			2		12	94
skrivefeil			7					7
ulik skrivemåte				27				27
ulik eieform					58	1	1	60
privat adresse						7	1	8
SUM	56	109	7	27	64	13	26	302

Det ser ut for at hovedmengden av de riktige som den maskinelle kontrollen fanget opp som mulige feil var oppdateringsmangler ved navn og adresse (adresse omfatter også kommunenummer). Det hender at NyFor og BOF skriver litt ulikt, enten fordi de får oppgitt ulikt eller at de har ulik praksis ved registrering av navn og/eller adresse. I tillegg hadde over 20 prosent ikke oppdatert eieform.

Skrivefeil og at privat (eiers) adresse er ført opp i stedet for firmaets var det noen tilfeller av.

Av 302 riktige koblinger var det 8,6 prosent som hadde ulik tilstand. Det var altså ikke mange, men det hender at BOF- og NyFor-enheter ikke har lik tilstand uten at linken er feil. (Ulike oppdateringsrutiner, tidsaspektet. Se avsnitt 5.1.).

7.2.3. *Koblinger som er gale fordelt på de tre kriteriene og tilstandsfeil*

Lister	Antall foretak	Navn	Eieform	Kommune	Tilstandsfeil
liste 1	29 (100 prosent)	6 (21 prosent)	12 (41 prosent)	11 (38 prosent)	17 (59 prosent)
liste 2	50 (100 prosent)	9 (18 prosent)	25 (50 prosent)	21 (42 prosent)	28 (56 prosent)
1 og 2	79 (100 prosent)	15 (19 prosent)	37 (47 prosent)	32 (41 prosent)	45 (57 prosent)

En ser at flestparten av de som kommer ut som potensielle feil i det maskinelle kontrollprogrammet, og som er faktiske feil, er relasjoner mellom enheter med ulik eieform og/eller ulikt kommunenummer. Over halvparten har tilstandsfeil. De som er funnet som feil pga. navneulikheter er her i mindretall.

7.2.4. Mulige årsaker til at koblingene er foretatt

LIKHETER MELLOM ENHETENE I DE GALE KOBLINGENE:

Lister	Antall foretak	Navn	Post-nummer	Eieform	Gate-/postboksadr.
liste 1	29 (100 prosent)	24 (83 prosent)	11 (38 prosent)	18 (62 prosent)	10 (34 prosent)
liste 2	50 (100 prosent)	42 (84 prosent)	19 (38 prosent)	23 (46 prosent)	13 (26 prosent)
1 og 2	79 (100 prosent)	66 (84 prosent)	30 (38 prosent)	41 (52 prosent)	23 (29 prosent)

Navnelikhet gjelder for størstedelen av mengden. Hele 84 prosent har likt eller tilnærmet helt likt navn. Mange er nok her blitt koblet maskinelt i Statistisk sentralbyrås månedlige koblingsrutine. Eksakt likt navn har tidligere vært tilstrekkelig som kriterium for maskinell kobling. Dette ser ut til å være årsaken til en stor andel av eksisterende feilkoblinger.

Også i den manuelle behandlingen i Statistisk sentralbyrå vil navnelikhet veie tungt ved kobling. Men da skal det i tillegg gjøres andre vurderinger bl.a. om enhetene er på samme sted og om de har samme eieform. En må regne med at det blir gjort feilvurderinger, særlig fordi en del enheter kan se identiske ut uten å være samme enhet.

7.2.5. Konklusjon

Navneforskjeller er en hyppigere foreteelse blant de riktige koblingene enn de som faktisk er feil. Dette tyder på at navnet kan være tilnærmet eller eksakt det samme selv om linken er feil, og motsatt at navnet gjerne kan være ulikt for samme enhet i de to registrene BOF og NyFor. Det siste kan forklares ut fra oppdateringsmangler eller små skriveforskjeller. Likt eller tilnærmet likt navn som koblingsgrunnlag er en sannsynlig årsak til feilkobling både i maskinelle og manuelle rutiner. Navn er ergo ikke noe entydig kriterium for om en kobling er korrekt eller ikke, men en enhet bør logisk sett ha samme navn i alle registre.

Det er en stor andel enheter med ulik eieform både blant de riktige og de gale av de potensielle feilkoblingene. Men andelen er størst for de faktiske feilkoblingene, hvor hele 46 prosent består av enheter som ikke har samme eieform i BOF og NyFor. Derfor kan en trygt si at eieform er anvendelig som kontrollkriterium.

Avvik i kommunenummer utgjør omtrent en like stor andel av de riktige som av de gale koblingene. Mange adresseendringer er ikke registrert i begge registrene, dermed kan samme enhet se ut for å være en annen i ett av registrene.

Nesten 80 prosent riktige koblinger blant de som kommer ut som feil i den maskinelle kontrollen er mye. Årsaken til dette er imidlertid i hovedsak oppdateringsfeil i registrene, noe som gjør det vanskelig å få skilt ut maskinelt de riktige koblingene. Siden det dessuten er viktig med et mest mulig riktig datainnhold i Enhetsregisteret, er det nødvendig å fange opp slike inkonsistenser for å få dem rettet opp senere.

Vi vil anbefale en manuell gjennomgang av alle potensielle feilkoblinger. Å la være vil innebære en feilprosent på rundt 20, noe som ikke er akseptabel kvalitet. Grunnen til dette er i første omgang ikke knyttet til nødvendigheten av å få rettet linkene, siden 20 prosent ikke utgjør mer enn om lag 2 000 koblinger på landsbasis. Hovedpoenget vil være å få fanget opp enheter med inkonsistens i datainnholdet.

7.3. Navnematch mellom MVA og henholdvis NyFor og LøFor

Vi har antatt at et maskinelt navnesøk som koblingsverktøy sannsynligvis er mer usikkert enn et system som i hovedsak er bygd på flere manuelle rutiner. Vi har derfor valgt *ikke* å benytte et maskinelt kontrollprogram på disse relasjonene, men forutsatt manuell gjennomgang.

Navnesøket mot NyFor og LøFor er basert på den samme metoden, når en ser bort fra at kommunekriteriet er utelatt i forhold til LøFor. De to registrene har felles registerfører og er underlagt de samme rutiner i utgangs-

punktet. NyFor er det registeret av de to som er best oppdatert. I teststudiet nøyde vi oss med å undersøke ca. en femtedel av mengden fra navnematch mellom MVA og NyFor, mens MVA-LøFor mengden ble gjennomgått i sin helhet.

7.3.1. MVA-NyFor koblingsforslag

Resultatet ble 755 forslag til link. Av disse var 238 NyFor-enheter knyttet til andre MVA-enheter via BOF. Dette er koblinger som kontrolleres i det øvrige opplegget og vi har derfor ikke brydd oss om disse i denne omgang. De resterende 517 må vurderes som mulige nye koblinger. Imidlertid vil mange MVA-enheter få foreslått opptil flere NyFor-enheter. Derfor vil det være færre enn 517 MVA-enheter i mengden. Tallet angir koblingsforslag, en og samme MVA-enhet kan altså være involvert i flere av koblingsforslagene.

Av de 517 enhetene som fikk koblingsforslag ble 110 kontrollert og 5,5 prosent viste seg å være feil.

7.3.2. MVA-LøFor koblingsforslag

Resultatet ble 256 forslag til link. Også her vil en MVA-enhet kunne få foreslått flere mulige foretak til kobling.

Antall koblingsforslag, MVA-LøFor	Antall riktige koblinger	Antall gale koblinger
256 (100 prosent)	189 (73,8 prosent)	67 (26,2 prosent)

En stor andel av navnematchens koblingsforslag gir riktige relasjoner.

Av de 67 gale koblingene er det 22 med ulikt kommunenummer, mens det er seks av de godkjente relasjonene som ikke har samme kommune. Feilprosenten hadde sunket med kanskje så mye som seks prosent dersom man kun hadde plukket ut kandidater i samme kommune. På den annen side hadde vi også mistet en del riktige forslag.

7.3.3. Konklusjon

En streng navnematchrutine, som den valgte, gir ikke mange forslag til link mellom MVA- og NyFor-/LøFor-enheter. Men antallet vil øke om vi også tar med de aktive MVA-enheter som har link til BOF uten videre kobling til NyFor. Det vil være fornuftig å ta med disse fordi vi da vil finne flere MVA-enheter som er registrert i Brønnøysund og således unngå dobbeltnummerering av disse. Antallet vil imidlertid reduseres dersom MVA-enheter med status slettet blir utelatt. Om nettoresultatet gir en økt, redusert eller uendret mengde i forhold til hva metoden i teststudiet ville gitt er dermed ikke lett å forutsi.

Fordelen med å bruke en såpass strikt metode er at det gir en stor andel riktige koblinger, særlig i forhold til NyFor. Dermed fanger konverteringen opp allerede nummererte enheter som ellers ville fått tildelt nytt nummer. Jo flere dubletter en kan unngå desto bedre vil det være.

Vi har valgt å beholde lik kommune som betingelse for navnematch mellom MVA og NyFor. Dersom denne betingelsen blir utelatt vil antall koblingsforslag øke, men vi har ikke ønske om å øke dette antallet uten å vite at det vil gi stor gevinst i form av mange flere riktige koblinger.

I forhold til LøFor derimot, satte vi ikke noe krav til likt kommunenummer i navnematchen. Her er kvaliteten på datainnholdet noe ujevnt og antallet metoden gir oss ikke avskrekkende stort med tanke på en manuell gjennomgang. I håp om å få fanget opp flest mulige riktige koblinger har vi valgt å ta med de ekstra enhetene som ikke har likt kommunenummer, men som har lik lydkode (navnematch).

Det vil være nødvendig å gå manuelt gjennom samtlige koblingsforslag. Det er antatt at det for de fleste av disse koblingene sin del vil være forholdsvis raskt og enkelt å avgjøre om de er riktige eller ei.

8. Tilbakemelding fra Fylkesskattekontoret i Sør-Trøndelag

8.1. Maskinelt godkjente koblinger

8.1.1. Resultater

Tre utvalg av maskinelt godkjente linker mellom BOF og MVA ble kontrollert ved Fylkesskattekontoret i Sør-Trøndelag. Utvalgene er tatt blant de eldste relasjonene, fordi en antar at koblingene blir mer usikre over tid.

- 1) 231 koblinger med aktive enheter
- 2) 117 koblinger med aktive enheter som har kun ett likt ledd i navnet
- 3) 150 koblinger med tilstandsfeil

De tre mengdene fordeler seg slik på riktig og feil kobling:

	Mengde 1)	Mengde 2)	Mengde 3)	Totalt	Prosent
Riktig	230	115	148	493	99,0 prosent
Feil	1	2	2	5	1,0 prosent
Totalt	231	117	150	498	100,0 prosent

Av de 117 i mengde 2) var det 61 koblinger som hadde mer enn ett ledd i begge navn. Eksakt likt navn her, som mellom BOF og NyFor, ser fortsatt ut til å være et godt kriterium, men en kan ikke være helt sikker dersom det kun er ett likt ledd i flerleddsnavn. De to gale koblingene hadde begge to ledd i navnene.

Flesteparten i mengde 1) hadde flere enn ett ledd i navnet, hele 214 av 231. Her var det kun en gal link. Å stille krav om bare ett likt ledd i navnet, når de andre to kriteriene samtidig er oppfylt, ser derfor ikke ut til å være problematisk.

Mengde 3) inneholdt 62 koblinger mellom enheter som er opphørt i BOF og aktive i MVA. Denne varianten skal ikke kunne eksistere i teorien. Siden det bare var to av disse som var gale koblinger må de resterende være opp-dateringsmangler.

De øvrige (86) av koblingene i mengde 3) er relasjoner mellom slettede MVA-enheter og aktive BOF-enheter. En enhet må drive merverdiavgiftspliktig virksomhet for å være registrert som aktiv i MVA. Slettede enheter i MVA som er aktive i BOF kan derfor være en reell situasjon. Likevel er det sannsynlig at en del av disse BOF-enhetene skulle vært opphørt.

8.1.2. Feiltyper

Mengde 1

Det var koblet sammen to enheter her med feil navn, men ett likt ledd i navnet gjorde at den maskinelle kontrollen ikke fanget opp denne feilen. Det fins en annen MVA-enhet som er den samme som BOF-enheten.

Mengde 2

Det er fisjoner som er årsaken til at det her har blitt to gale forbindelser. Begge er utskillelse av eiendomsdrift som eget selskap, hvor det nye selskapet har fortsatt med det gamle navnet. (Dette er tilsvarende som for feilkoblingen mellom BOF og NyFor beskrevet under punkt 8.1.2.). Det nye selskapet er registrert i MVA med nytt MVA-nummer og det opprinnelige selskapets navn. Dette var koblet med det opprinnelige selskapet som var registrert i BOF med det samme navnet. Navnelikhet førte til at koblingene ikke ble fanget opp i kontrollen som gale koblinger.

Når det gjelder den ene av koblingene, er det ikke eiendomsselskapet som er registrert i avgiftsmanntallet. Årsaken til dette er at eiendomsselskap ikke er avgiftspliktige. De har imidlertid anledning til å registrere seg dersom de leier ut til andre som er avgiftspliktige. Det er det nye selskapet som betaler inn på det gamle MVA-nummeret. Dermed skal ikke det gamle selskapet som nå kun har eiendomsdrift ha kobling til den MVA-enheten det tidligere var knyttet til.

Mengde 3

Den ene linken her ser ut til å være en ren feiltakelse fordi enheten i BOF er slettet før den er registrert i MVA. Den andre linken skyldes en dublett i BOF. Det er opprettet et nytt foretak i BOF, mens enheten i MVA er blitt reaktivisert.

8.1.3. Konklusjon

Utfisjonerte eiendomsselskap fører til gale linker dersom fisjonene ikke blir registrert i BOF. Prosedyren for behandling av disse er imidlertid der, men Statistisk sentralbyrå er avhengig av å få melding om disse reorganiseringene. Det vil derfor være enkelte tilfeller av feilkoblinger av denne typen. Imidlertid er det ikke noen måte å fange opp disse på uten å måtte sjekke alle eiendomsselskapene. Det vil være en uforholdsmessig stor jobb for å fange opp så få feiltfeller.

De andre tre feilene har ikke noen fellestrekk, og vil ikke kunne fanges opp uten samtidig å få med altfor mange riktige koblinger. Jf. hvor godt navnekriteriet har vist seg å fungere og tidligere drøfting om dato som kriterium.

8.2. Maskinelt ikke-godkjente koblinger

8.2.1. Resultater

De maskinelt ikke-godkjente koblingene ble fordelt på tre mengder før de ble oversendt Fylkesskattekontoret i Sør-Trøndelag :

Mengde 1:

Liste 1.1. og 1.2. inneholdt henholdsvis enkle (en til en) og multiple (en til mange og/eller mange til en) relasjoner. Alle enheter som i MVA har ansvarsforhold forskjellig fra 0 er med.

Mengde 2 :

Liste 2.1. og 2.2. ble delt inn på samme vis som listene i mengde 1, bortsett fra at kun de MVA-enheter som har ansvarsforhold lik 0 er med. Alle disse vil dermed ikke ha oppfylt eieformkriteriet.

Mengde 3 :

Også liste 3.1. og 3.2. hadde en inndeling i enkle og multiple koblinger. Dette er koblinger som er manuelt gjennomgått i Statistisk sentralbyrå og vurdert å være enheter som er riktig linket sammen.

Gjennomgangen ved Fylkesskattekontoret i Sør-Trøndelag fordeler koblingene etter riktig og feil slik tabellen under viser :

LISTE	1.1.	1.2.	2.1.	2.2.	3.1.	3.2.
RIKTIG PROSENT	195 97,0 prosent	51 92,7 prosent	29 90,6 prosent	7 58,3 prosent	371 95,4 prosent	35 92,1 prosent
FEIL PROSENT	6 3,0 prosent	4 7,3 prosent	3 9,4 prosent	5 41,7 prosent	18 4,6 prosent	3 7,9 prosent
TOTALT PROSENT	201 100,0 prosent	55 100,0 prosent	32 100 prosent	12 100 prosent	389 100 prosent	38 100 prosent

I en multippel kobling bryr vi oss ikke med slettede MVA-enheter dersom det også er en aktiv og riktig MVA-enhet knyttet til BOF-enheten. Fra liste 1.2. og 3.2. ble henholdsvis 106 (av 161 utsendte) og 51 (av 89 utsendte) koblinger ikke manuelt kontrollert. Bortsett fra enkelte dubletter var disse slettede MVA-enheter med en annen aktiv MVA-enhet knyttet til samme BOF-enhet, hvor koblingen mellom BOF-enheten og den aktive MVA-enheten var riktig. Linken mellom den aktive MVA- og BOF-enheten inngår i kontrollsystemet på vanlig måte.

På liste 2.2. ble av praktiske årsaker ikke koblingene av slettede MVA-enheter og aktive BOF-enheter sett bort fra, selv om det i tillegg fantes en annen aktiv og riktig MVA-enhet knyttet til. Alle de faktisk riktige koblingene her var av denne typen.

Liste 3.1. inneholdt 396 koblinger, men syv av disse viste seg å være dubletter, slik at det i realiteten er 389 koblinger som ble kontrollert.

Multiple relasjoner innenfor hver av de tre mengdene hadde gjennomgående høyere feilprosent i sine koblinger enn de koblingene som kun hadde en til en-relasjoner. Det er logisk at en til en-relasjoner er lettere å holde orden på.

Mengde 2 hadde betraktelig høyere feilprosent enn de to andre mengdene. Alle hadde eieform 0 og måtte i alle fall komme ut som potensielle feilkoblinger pga. eieformkriteriet. En skulle da tro at mange av disse var riktige koblinger bare eieformen ble oppdatert, men slik var det altså ikke. Dette kan tyde på at de som blir registrert ved fylkesskattekontorene med ansvarsforhold 0 er enheter uten tilstrekkelig opplysninger til å plassere enhetene innenfor rett ansvarsforhold. Mangelfull informasjon i registrene gjør det verre for Statistisk sentralbyrå å koble rett.

Koblingene fra de seks listene i alt fordelte seg som følgende :

	ANTALL	PROSENT
RIKTIG	688	94,6 prosent
FEIL	39	5,4 prosent
TOTALT	727	100,0 prosent

Nærmere 95 prosent av de potensielle feilkoblingene viste seg å være riktige linker. Det kan være interessant å merke seg at dersom mengde 2 hadde vært utelatt ville feilprosenten vært 4,5 og ikke 5,4. Dette betyr at MVA-enheter med eieform 0 peker seg ut som problemenheter i koblingssammenheng.

Som i forrige kapittel beskrives i det følgende de som faktisk er gale og riktige av de linkene som kom ut som potensielt gale. Her er det relativt færre reelle feilkoblinger og det vil derfor være en desto større besparelse om vi klarer å isolere flere av de som faktisk er riktige koblinger.

8.2.2. Beskrivelse av de som kom ut som feil i datakontrollprogrammet, men som viste seg å være riktige koblinger

Nedenfor følger en oversikt over riktige koblinger i mengde 1, 2 og 3 :

Mengdene : 1, 2 og 3.	ikke oppd. navn	ikke oppd. adr.	skrive- feil	ulik skrive- måte	ulik eie- form	privat adr.	ulik tilst.	SUM
navn ikke oppdatert	49	8		1	10		42	110
adr. ikke oppdatert		69		1	8		51	129
skrivefeil			2				1	3
ulik skrivemåte				23	5		4	32
ulik eieform					177	1	211	389
privat adresse						21	4	25
SUM	49	77	2	25	200	22	313	688

41 av de 211 som har ulik tilstand og ulik eieform hadde også forskjellig navn. Dermed var det i alt 151 koblede enheter som hadde oppdateringsmangler ved navnet. Sammen med adresseforskjeller utgjorde dette om lag 40 prosent av de manuelt godkjente koblingene. Det var nesten 60 prosent som hadde ulik eieform og mange av disse hadde i tillegg ulik tilstand. Ergo var forskjellig eieform hovedårsaken til at riktige relasjoner mellom MVA og BOF ble fanget opp som gale i den maskinelle kontrollen. Dette i motsetning til BOF-NyFor-koblingene der hovedproblemet var ulik navn og adresse.

Mange av de som hadde ulik eieform var offentlig virksomhet. Fylkesskattekontorene registrerer ofte offentlig virksomhet under begrenset ansvar selv om det i MVA er en egen kode for offentlig virksomhet. Dermed vil samme enhet kunne opptre i BOF som f.eks. en kommunal enhet og i MVA som selskap med begrenset ansvar.

Av de 688 riktige koblingene var det 45,6 prosent som hadde ulik tilstand. Dersom tilstandsforskjellene ikke skyldtes oppdateringsfeil må det tolkes som at den MVA-pliktige virksomheten var opphørt mens annen virksomhet fortsatte.

8.2.3. Koblinger som er gale fordelt på de tre kriteriene og tilstandsfeil

Liste	Antall foretak	Navn	Eieform	Kommune	Tilstandsfeil
1.1.	6 (100 prosent)	4 (67 prosent)	4 (67 prosent)	0	5 (83 prosent)
1.2.	4 (100 prosent)	1 (25 prosent)	3 (75 prosent)	1 (25 prosent)	2 (50 prosent)
2.1.	3 (100 prosent)	0	3 (100 prosent)	0	3 (100 prosent)
2.2.	5 (100 prosent)	5 (100 prosent)	5 (100 prosent)	1 (20 prosent)	5 (100 prosent)
3.1.	18 (100 prosent)	5 (28 prosent)	12 (67 prosent)	3 (17 prosent)	9 (50 prosent)
3.2.	3 (100 prosent)	3 (100 prosent)	1 (33 prosent)	3 (100 prosent)	0
sum	39 (100 prosent)	18 (46 prosent)	28 (72 prosent)	8 (21 prosent)	24 (62 prosent)

Mellom BOF og MVA var det så få faktisk gale koblinger at det er vanskelig å trekke noen generelle konklusjoner om hvilke kriterier som vil fange opp flest mulig gale koblinger. Imidlertid ser det ut for at ulik eieform er et viktig kriterium. En stor andel hadde også tilstandsforskjell. Begge disse to momentene utgjorde også den største mengden av de faktisk riktige koblingene.

8.2.4. Mulige årsaker til at koblingene er foretatt

Nedenfor følger en oversikt over likheter mellom enhetene i de gale koblingene:

Liste	Antall foretak	Navn	Postnummer	Eieform	Gate-/ postboksadr.
1.1.	6 (100 prosent)	3 (50 prosent)	3 (50 prosent)	2 (33 prosent)	2 (33 prosent)
1.2.	4 (100 prosent)	3 (75 prosent)	1 (25 prosent)	1 (25 prosent)	1 (25 prosent)
2.1.	3 (100 prosent)	3 (100 prosent)	3 (100 prosent)	0	0
2.2.	5 (100 prosent)	0	3 (60 prosent)	0	0
3.1.	18 (100 prosent)	15 (83 prosent)	12 (67 prosent)	13 (72 prosent)	12 (67 prosent)
3.2.	3 (100 prosent)	0	1 (33 prosent)	2 (67 prosent)	0
sum	39 (100 prosent)	24 (62 prosent)	23 (59 prosent)	18 (46 prosent)	15 (38 prosent)

Antallet var såpass lite at de relative fordelingene på de fire valgte alternativene i tabellen ovenfor ikke nødvendigvis sier noe allmenngyldig om feilkoblinger mellom MVA- og BOF-enheter. Navnelikhet gjaldt imidlertid for over halvparten.

8.2.5. Konklusjon

Storparten av de riktige koblingene hadde ikke tilsvarende eieform/ansvarsforhold. En del av disse var oppdateringsmangler, mens andre skyldtes ulik praksis ved fylkesskattekontorene og Statistisk sentralbyrå. Oppdateringsmangler på navn og adresse var også en hyppig foreteelse.

Mengden av faktiske feilkoblinger mellom MVA og BOF var liten. De fleste av de potensielle feilkoblingene viste seg kun å være resultater av oppdateringsmangler. Fordi det er viktig med et mest mulig riktig datainnhold i Enhetsregisteret, er det nødvendig å fange opp slike inkonsistenser for å få dem rettet opp. Anbefalingen er derfor at samtlige potensielle feilkoblinger blir manuelt gjennomgått.

I og med at det er så få faktiske feilkoblinger er det ikke lett å forutsi hvordan de på landsbasis fordeler seg på kriteriene og mulige årsaker til at koblingene er foretatt.

9. Hovedresultater og anbefalinger

9.1. Omfang

Over 90 prosent av koblingene ble godkjent i datakontrollprogrammene. Dvs. at færre enn ti prosent av koblingene kommer ut som potensielle feil. Dette gjelder for både MVA-BOF og NyFor-BOF. Totalt sett vil den potensielle feilmengden være større fordi samme enhet ikke nødvendigvis er feilkoblet både mot MVA og NyFor. Noen få, et tilnærmet betydningsløst antall relativt sett, feilrelasjoner mellom MVA og NyFor kommer i tillegg. I dårligste fall kan resultatet bli summen av feilprosentene, men dette er ikke videre realistisk. MVA har høyere feilprosent i relasjonene til BOF enn NyFor har, henholdsvis ca. 9 mot 6. Summen av disse er teoretisk den maksimale feilprosenten. Et sted mellom 9 prosent og 15 prosent vil være realistisk, og vi har landet på at 12 prosent kan være et sannsynlig anslag.

Om vi velger å kontrollere bare enheter som er i alle tre registrene i forhold til BOF, vil antall koblinger til kontroll reduseres. En kan regne med at 96 prosent av de NyFor-enhetene som har kobling til BOF kommer med. Men bare 62 prosent av MVA-BOF linkene har videre tilknytning til NyFor. Dersom en bare konsentrerer seg om MVA-enheter som har NyFor-nummer vil mengden på landsbasis utgjøre om lag 130 000 koblinger. Antar en at 12 prosent av disse vil være potensielle feilkoblinger vil antallet til manuell gjennomgang bli ca. 16 000 koblinger. Dette innebærer at om lag 115 000 linker blir maskinelt godkjent, hvorav ca. 1 200 uriktig godkjent.

I tillegg kommer koblingsforslagene basert på navnematch. Hvis vi går ut fra at de ulike fylkene i gjennomsnitt vil få foreslått om lag samme mengde som Sør-Trøndelag (dvs. 773), vil et totalt anslag for hele landet være på 15 000 koblinger. Hovedmengden vil bestå av korrekte forslag og det vil stort sett være enkelt å finne frem til de som er gale.

Opplegget vil innebære at alle maskinelt ikke-godkjente koblinger og navnematchbaserte koblinger må gjennomgås manuelt. På landsbasis utgjør dette i alt ca. 31 000 koblinger. Etter manuelle gjennomganger og opprettinger vil MVA-registeret være konvertert av Statistisk sentralbyrå. På landsbasis syns dette opplegget å konvertere om lag 144 000 MVA-enheter. I tillegg skal det konverteres en mindre del av de slettede MVA-enheter som ikke er med i det øvrige opplegget. Omfanget vil først bli avklart senere (av Skattedirektoratet).

9.2. Kvalitet

I utvalget av de maskinelt godkjente er 98,7 prosent BOF-NyFor-koblinger og 99,0 prosent BOF-MVA-koblinger riktige. Dette betyr at rundt en prosent av koblingene vil være feil, noe som aksepteres som helt tilfredsstillende.

Alle de maskinelt ikke-godkjente koblingene ble manuelt kontrollert. Vel 50 prosent av disse er skilt ut på grunnlag av eieformkriteriet. 79,3 prosent av BOF-NyFor-koblingene og 95,6 prosent av BOF-MVA-koblingene viste seg å være korrekte relasjoner. Å få til et opplegg som maskinelt skiller ut disse som riktige er ikke mulig siden årsaken til at de ikke godtas hovedsakelig er oppdateringsmangler. Det å la være å ta hensyn til dette vil føre til at de faktisk gale går igjennom. Det kan ikke aksepteres. Det er også viktig å fange opp koblinger med oppdateringsmangler for å få rettet opp registrene på dette området.

Navnmatchene gir forslag til mange riktige relasjoner. Særlig mellom MVA og NyFor, hvor den manuelle gjennomgangen viser at kun 5,5 prosent er feil. Overfor LøFor ligger feilprosenten på 26,2. Ved å benytte denne metoden vil færre allerede nummererte enheter få tildelt nytt foretaksnummer når Enhetsregisteret opprettes. Imidlertid vil det være nødvendig å gå manuelt igjennom samtlige koblingsforslag, fordi det her ikke gjelder linker som er opprettet gjennom et etablert system som er underlagt rutiner for oppdatering. Dessuten er de maskinelle kontrollkriteriene til dels implisert i selve navnmatchmetoden, og dermed ikke like aktuelle som kontrollgrunnlag.

Over halvparten av de som faktisk er uriktig knyttet sammen har ulik tilstand. Det ser ut som at feilkoblinger mellom BOF og NyFor oftest skjer fordi navnene er like eller tilnærmet like.

Vedlegg

NyFors foretakstyper:

- AS - aksjeselskap
- DA - delt ansvar
- ANS - ansvarlig selskap
- PRE - partrederi
- ENK - enkeltmannsforetak
- KS - kommandittselskap
- STI - stiftelse
- SPA - sparebank
- NÆR - næringsdrivende forening
- FLI - forening/lag/innretning
- BBL - boligbyggelag
- BRL - borettslag
- GFS - gjensidige forsikringsselskap
- BA - begrenset ansvar
- SÆR - annet foretak iflg. særskilt lov
- NUF - norsk avdeling av utenlandsk foretak

LøFors eierforholdskoder:

- 1 - aksjeselskap
- 2 - VPS-foretak
- 3 - kommandittselskap
- 4 - ansvarlig selskap / partrederi
- 5 - pro rata selskap / delt ansvar / partrederi
- 6 - utenlandsk næringsdrivende
- 8 - stiftelse
- 9 - enkeltmannsforetak
- 11 - LøFor-foretak
- 12 - boligbyggelag/borettslag
- 13 - andelslag
- 14 - sparebank
- 15 - gjensidige forsikringsselskap
- 16 - annet foretak iflg. særskilt lov
- 17 - forening/lag/innretning
- 18 - revisor

Statistisk sentralbyrå

Oslo
Postboks 8131 Dep.
0033 Oslo

Tlf.: 22 86 45 00
Fax: 22 86 49 73

Kongsvinger
Postboks 1260
2201 Kongsvinger

Tlf.: 62 88 50 00
Fax. 62 88 50 30



Statistisk sentralbyrå
Statistics Norway