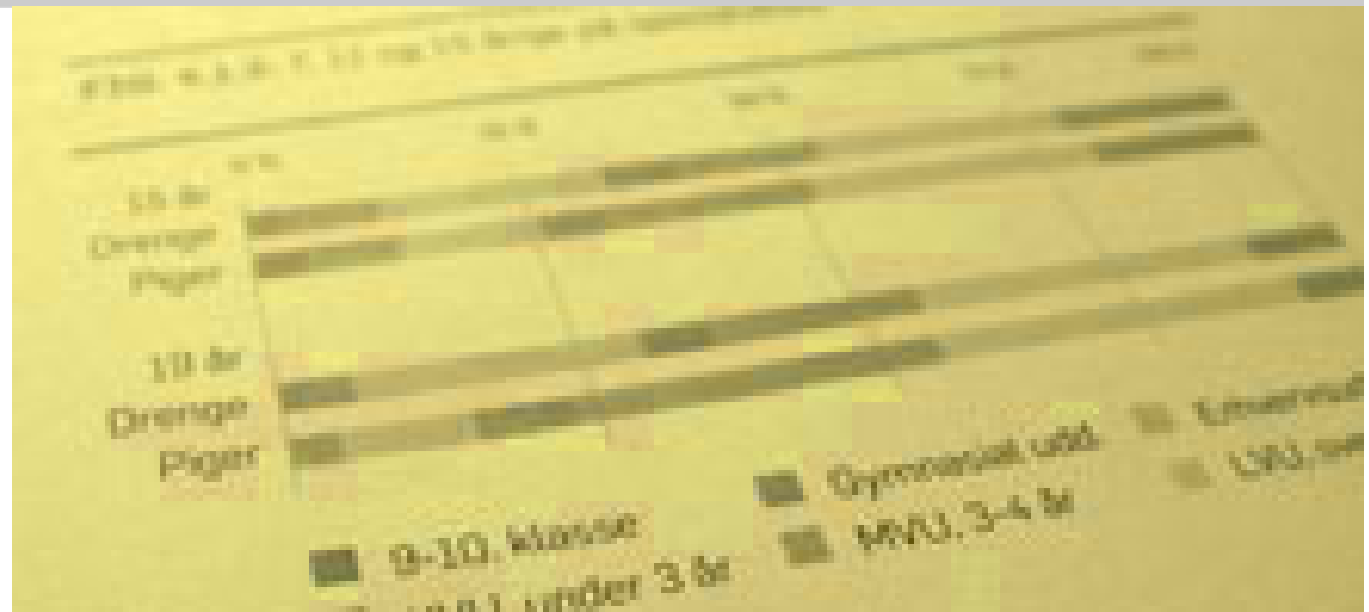


Mette Gørtz,
SFI og CSSR



Mulighederne i at anvende registerdata til evaluering

Dansk Evalueringssselskabs Årsmøde
22. september 2011, Nyborg Strand

Præsentation

- Mette Gørtz, cand. polit. 1993, Ph.D. 2006
- Joberfaring: Danmarks Statistik (93-95), DØRS (95-99), Økonomi- og Erhvervsministeriet (99-03), AKF (06-10)
- SFI (2010-): Seniorforsker og leder af Center for Survey og Survey/Registerdata (CSSR, www.cssr.dk)
- CSSR: Netværk af danske forskningsinstitutioner, som sammen opbygger database med survey- og registerdata
- Støttet af Forskningsråd for Samfund og Erhverv (FSE)
- Medlemmer: KU-Samf og KU-Sund, AU-Samf, AAU-Samf, SDU-Samf, CBC, NFA, SFI (formand og vært)
- Runde med præsentation af deltagere i seminaret

Plan for seminaret

1. Hvad er registerdata? Indsamling, indhold, fordele og perspektiver i forhold til evaluering.
2. Summemøde/gruppediskussion
3. Ulemper/mangler ved registerdata, surveys eller eksperimenter kombineret med registerdata, dataadgang
4. Øvelse i grupper: Beskriv evalueringsprojekt i din organisation, hvordan kunne registerdata, evt. i kombination med surveydata, bringes i spil?
5. Gennemgang: Eksempler på anvendelser af registerdata, evt. i kombination med andre data. Forskning og andre evalueringsprojekter.
6. Diskussion

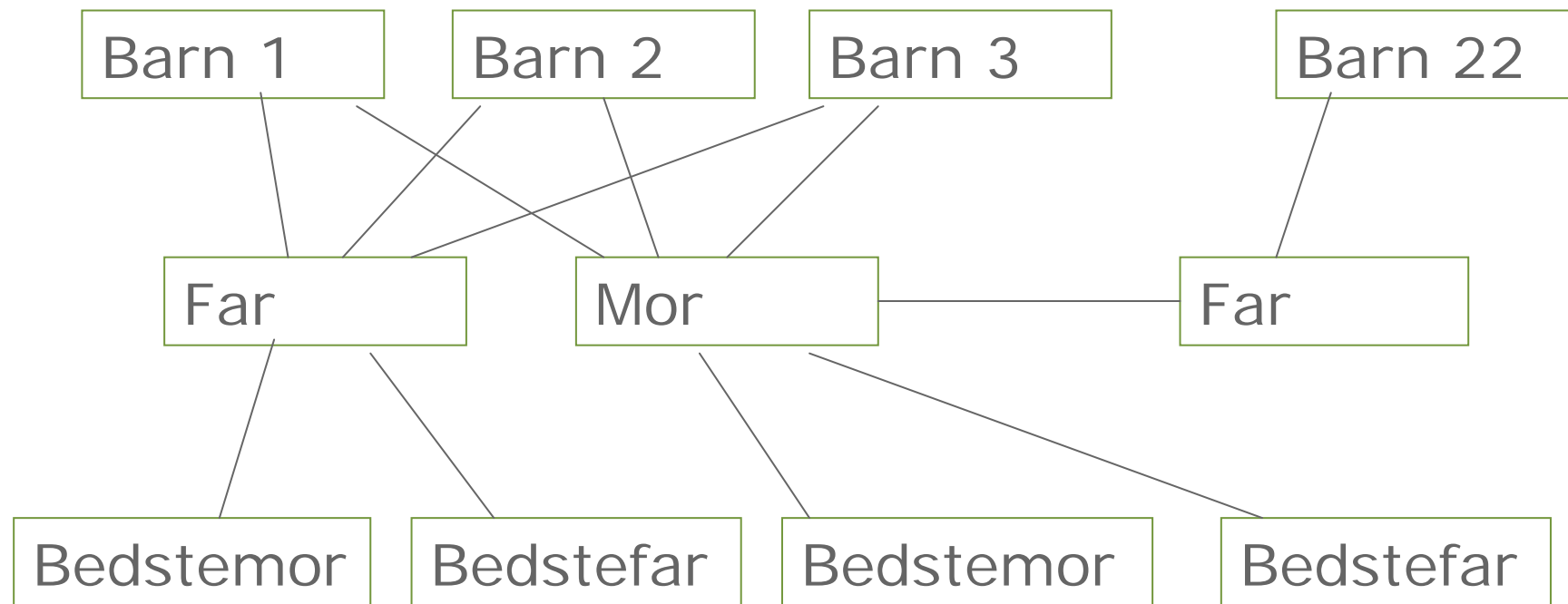
Hvad er registerdata?

- Data indsamlet med administrativt formål.
- Indsamles af myndigheder.
- Fx: Skat, sundhedssystem, uddannelse, udbetaling af sociale ydelser, virksomheders juridiske oprettelse etc.
- Oplysninger registreres på personer eller virksomheder, dvs. én observation pr. person eller pr. firma.
- Danmarks Statistik samler oplysninger fra administrative registre:
 - Anvendes til at producere statistik. www.statistikbanken.dk
 - Stilles til rådighed – **anonymiseret form** - for forskere under Forskerordning. www.dst.dk/TilSalg/Forskningservice.aspx

Hvilke typer af registeroplysninger?

- Registeroplysninger på individniveau i Danmarks Statistik:
 - Familieforhold: Danmarks Statistiks familierelationer - "E-familie".
www.dst.dk/tilsalg/forskningservice/hkt4forsker/hkt4_variabel_liste_forsker/hkt4_variabel.aspx?fk=37407
 - Kobling af CPR-numre mellem børn og forældre for de fleste personer født efter 1965 → familierelationer, generationer,
www.dst.dk/TilSalg/Forskningservice/Databaser/FertilitetsDB.aspx
 - Børn som gennemgående individ i data
www.dst.dk/TilSalg/Forskningservice/Databaser/Bornedatabasen.aspx

Familierelation



Hvilke typer af registeroplysninger?

- Uddannelse
 - Uddannelse: Igangværende og fuldført
 - Kurser under voksen-efteruddannelse
 - Folkeskoleelever: Oplysninger om klassetrin siden 2007. Før 2007 kun for de ældste klassetrin.
 - Ungdomsuddannelser.
 - Faglige og videregående uddannelser.
 - Karakterer.
- Kriminalitet

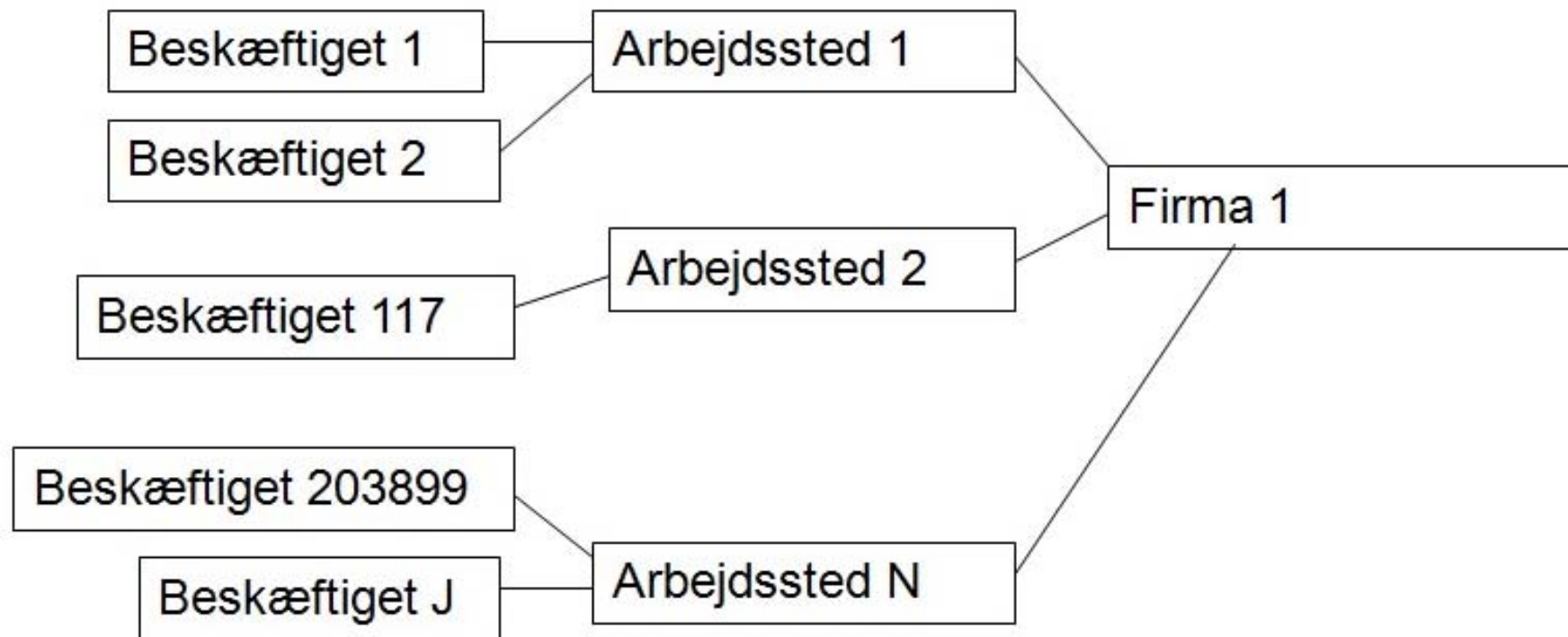
Hvilke typer af registeroplysninger?

- Registeroplysninger på individniveau i Danmarks Statistik:
 - Arbejdsmarkedssituation
www.dst.dk/TilSalg/Forskningservice/Databaser/IDA.aspx
 - Sociale ydelser på ydelsestyper
 - DREAM: Beskæftigelsesoplysninger på ugeniveau, se
www.dst.dk/TilSalg/Forskningservice/Databaser/Dream.aspx
 - Arbejdssted via arbejdsstedskode
 - Indkomster og skat

Hvilke typer af registeroplysninger?

- Oplysninger om **virksomheder**
 - Oplysninger på arbejdssteder
 - Oplysninger på juridiske enheder, dvs. firmaer
 - **Sammenhæng** mellem firmaer, arbejdssteder og personer
 - Oplysninger om antal ansatte, omsætning, køb af varer, lønsum etc.
 - Oprettelse og lukning af firmaer
 - Firmaer registreret med branchekode
 - Oplysninger om ejerform

Personer, arbejdssteder og firmaer



Sundhedsregistre

- Sundhedsstyrelsen indsamler registre vedr. sundhed, fx
 - Indlæggelser på somatiske hospitalsafdelinger
 - Psykiatriske indlæggelser
 - Besøg hos praktiserende læger og speciallæger
 - Dødsårsag

www.sst.dk/publ/Publ2010/DOKU/Registre/Doedsaarsagsregisteret2009.pdf

- Lægemiddelstyrelsen bestyrer:

- Lægemiddelregistret

www.dst.dk/TilSalg/Forskningservice/Databaser/LMDB.aspx

- Disse registre leveres videre til Danmarks Statistik, som stiller data til rådighed via Forskerordningen

Fordele ved registerdata

- Kæmpe datamateriale: Hele befolkningen.
- Målemetoder: Konsistente målemetoder over tid. Myndigheder indsamler samme oplysninger hvert år på samme måde.
- Paneldata: Longitudinale data. Muligt at følge personer over tid.
- Følge personer geografisk vha. CPR-nummer-system.
- Kobling af personer i forskellige netværk:
 - Familien
 - Arbejdspladsen
 - Uddannelsessted, skole etc.
- Muligt at koble med andre datakilder via CPR-nummer.

Summe-session

- Diskuter muligheder for at gennemføre evaluering ved brug af registerdata i forhold til problemstillinger på din arbejdsplads
- Hvilke konkrete problemstillinger kunne evalueres vha. registerdata?
- Hvad skal der til for at kunne gennemføre en sådan evaluering?
- Hvilke forhindringer og potentialer ser I?

- 5-10 minutter, derefter diskussion i plenum

Ulemper

- Data indsamlet med administrative formål for øje
- Data ikke indsamlet for at svare på forsknings- eller evalueringsspørgsmål
- Data præcise på indkomst (registreret), men ikke forbrug
- Data fx ringe dækning vedr. formue, fordi formueskat er væk
- Data viser brug af sundhedsydelser, men ikke helbredstilstand
- Data fortæller om arbejde i arbejdsmarkedet, men ikke om omfang af sort arbejde, husarbejde, børneomsorg, fritid
- Data beskriver ikke præferencer, holdninger, tilfredshed etc.
- Registerdata kan anvendes til at beskrive *effekter*, men ikke *hvorfor* en indsats virker

Andre datakilder

- Administrative datasæt i kommuner el. lign.
 - Data indsamlet i forbindelse med særligt projekt eller eksperiment, fx Rask i Job
 - Surveydata
- Det er ofte muligt at kombinere registerdata med andre data.
- Det kræver, at
 - Det er muligt at koble via CPR-nr. eller virksomhedsnr.
 - Der er mange observationer i alternative datakilde

Eksperiment: Rask i Job

- Eksperiment i 12 kommuner
 - Fokus på KRAM-faktorer
 - Uddannelse af helsepiloter i hver enhed
 - Fulgt op på sygefravær via samtaler
-
- ➔ Evaluering så på indsats- og kontrolgruppen.
 - ➔ Indsamlede oplysninger om bl.a. sygefravær, KRAM-faktorer
 - ➔ Koblet til registeroplysninger for yderligere information om fx brug af sundhedsydelser og baggrundskarakteristika
-
- Fandt, at sygefraværet *steg* efter indførelse af programmet!

Surveys

- Surveys er i mange lande eneste kilde til viden om personer eller virksomheder
- I DK dækker registre mange af de oplysninger, som tidligere blev indsamlet vha. surveys.
- Surveys kan imidlertid give svar på spørgsmål, som ikke dækkes af registre, fx
 - Selvvurderet helbred og trivsel
 - Livssituation, hvordan/hvorfor indretter man sig
 - Præferencer, holdninger, forventninger, planlægning
 - Forbrug og opsparing

Surveys

- Surveys kan kobles til registerdata via CPR-nummer
- Kræver, at survey har adgang til CPR-numre for respondenter
- Potentiale:
 - Surveydata kan valideres: Hvor repræsentative er respondenter, hvad betyder bortfald for repræsentativitet?
 - Giver adgang til et hav af baggrundsoplysninger, fx beskæftigelse, uddannelse, familieforhold
 - Muliggør longitudinale analyser: Historie eller follow-up vedr. surveypopulationen
 - Survey- og registerdata kan supplere hinanden, fx helbred

Eksempler på surveys, der kobles til registre

- Børneforløbsundersøgelsen (SFI)
- Ældredatabasen (SFI og AKF)
- Ungdomsforløbsundersøgelsen (SFI og DPU)
- Forbrugsundersøgelserne (Danmarks Statistik)
- Arbejdsmiljøkohorten (Nationale Forskningscenter for Arbejdsmiljø, NFA)
- Bedre Sundhed for Mor og Barn (Seruminstituttet)
- Virksomhedernes Sociale Ansvar (SFI)
- European Values Survey (KU og SFI, den danske del)
- PISA og PIAAC (SFI og DPU mv.)
- European Social Survey (SFI, den danske del)
- SHARE (SDU, den danske del)
- International Social Survey programme (ISSP, dansk del)

Eksempler på andre typer data, der (kan) kobles til registre

- Nationale tests i folkeskolen. Aktuelle projekter i SFI:
 - For Horsens kommune.
 - Betydningen af daginstitutioner for resultater i folkeskolen.
- Survey blandt skoleledere og skolelærere (SFI)
 - Målinger af ledelse og undervisningskvalitet. Kobles til 9. klasses karakterer.

Gennemførelse af ny survey, der kan kobles med registerdata

- Ny survey, fx gennemført af SFI-Survey eller Danmarks Statistik
- Survey-indsamler udtrækker population baseret på CPR-udtræk
- Population kan indsnævres vha. registerdata, fx population i bestemt kommune, i bestemt branche, personer med bestemt uddannelse mv.
- Organisationer kan også gennemføre surveys i deres medlemsskare
- Kobling af surveys og registre kræver godkendelse hos Datatilsynet

Geografiske netværk

- Belyse geografiske relationer mellem individer og virksomheder
- GIS-koder (Geografisk InformationsSystem)
- Ruter og geografiske afstande. Korteste eller hurtigste.
- Eksempler:
 - Pendling, dvs. afstande mellem bolig og job
 - Afstande inden for kommuner, fx fra bolig til institutioner, skoler mv.
 - Boligområder, netværk, ghettodannelse
- Kan kobles til registerdata via CPR eller virksomhedsnumre (adressekoder)
- Serviceopgave, der bestilles i Danmarks Statistik. Dyrt!

Analyser af reformer, ændringer i administrativ praksis mv.

- Registerdata som "hukommelse"
- Fx belyse ændringer ved reformer:
 - Skolereformer i 50'erne
 - Reform af barselsorlov i 80'erne og igen i 2002
 - Efterlønsreform 1998/99
 - Kommunalreformen 2007
- Forskelle i administrativ praksis på tværs af kommuner
 - Benchmarking af integrationsindsats
 - Forskelle i praksis for tilkendelse af førtidspension
 - Beskæftigelsesindsats i kommunerne

Dokumentation af registerdata

- Generelt dokumenteret på Danmarks Statistiks hjemmeside
- Høj kvalitetsdata: Data med særlig fokus på god dokumentation
www.dst.dk/TilSalg/Forskningservice/hkt4Forsker.aspx
- Høj kvalitetsdata indeholder pt. over 600 variabler

Hvordan får man adgang til registerdata?

- Alle analyser på Danmarks Statistiks forskermaskine
- Forskningsinstitutioner, analyseinstitutioner, ministerier mv.
- Brugere af forskermaskinen godkendes af Danmarks Statistik
- Skrive projektbeskrivelse. Godkendes af Danmarks Statistik.
- Beskrive formål, analyser, population, variabler, årgange mv.
- Statistisk analyse i SAS, SPSS, Stata el. lign. programmer.
- Startomkostninger er store:
 - Tidsforbrug med at sætte sig ind i data, programmering mv.
 - Økonomisk side. Betaling for dataudtræk.
- Kan betale sig at alliere sig med institut, der kender data og analysemetoder.

Survey- og registerdata via CSSR

- Medlemmer af CSSR kan benytte CSSR's survey- og registerdatabase
- Betingelse: Survey er omdrejningspunkt, ikke rene registeranalyser
- CSSR har købt egen server på Danmarks Statistik
- Adgang til høj kvalitetsdata for alle år, hele befolkningen
- De mest almindelige programmer, herunder SAS, Stata, SPSS, Gauss etc.
- CSSR vejleder og er sparringspartner i forbindelse med udvælgelse af relevante registerdata, kontakt med Danmarks Statistik mv.

Praktisk øvelse: Evaluering med registerdata

- Grupper 2-3 personer, 30 minutter i grupper, 20 minutter til præsentation, brug evt. posters
- Vælg case i jeres organisation: Evaluering
- Beskriv indsatsen og evalueringsopgaven
- Forventninger og hypoteser (evt. programteori)
- Beskriv, hvordan casen kunne evalueres vha. registerdata
- Suppleres med andre typer undersøgelser?
 - Fx survey, eksperiment, yderligere data fra organisationen
- Metode: Statistisk metode, forløbsanalyse el. lign.
- Proces: Hvordan og hvem gennemfører evaluering?
- Hvad skal evalueringen bruges til? Implementering?

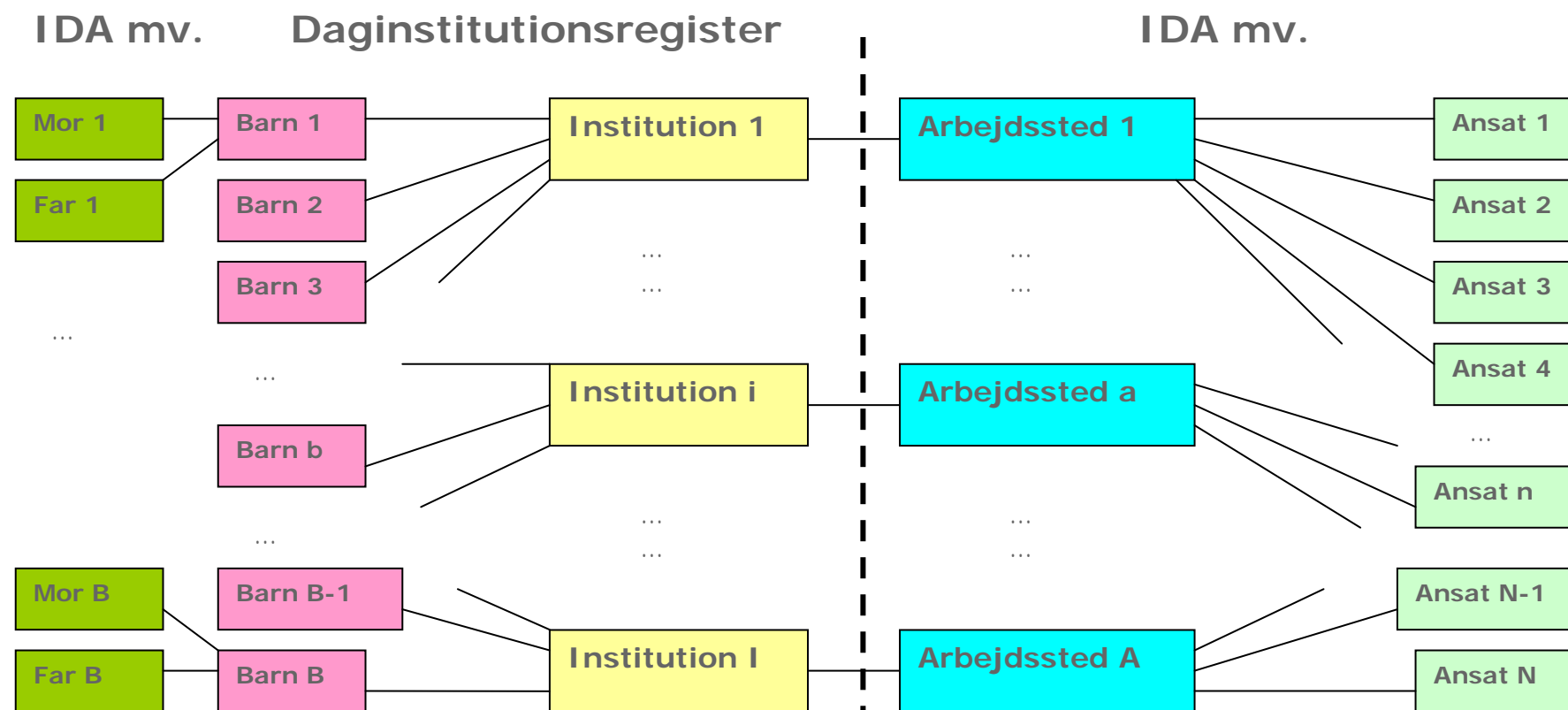
Adgang til Danmarks Statistiks forskermaskine

- <https://sslvpn.dst.dk>

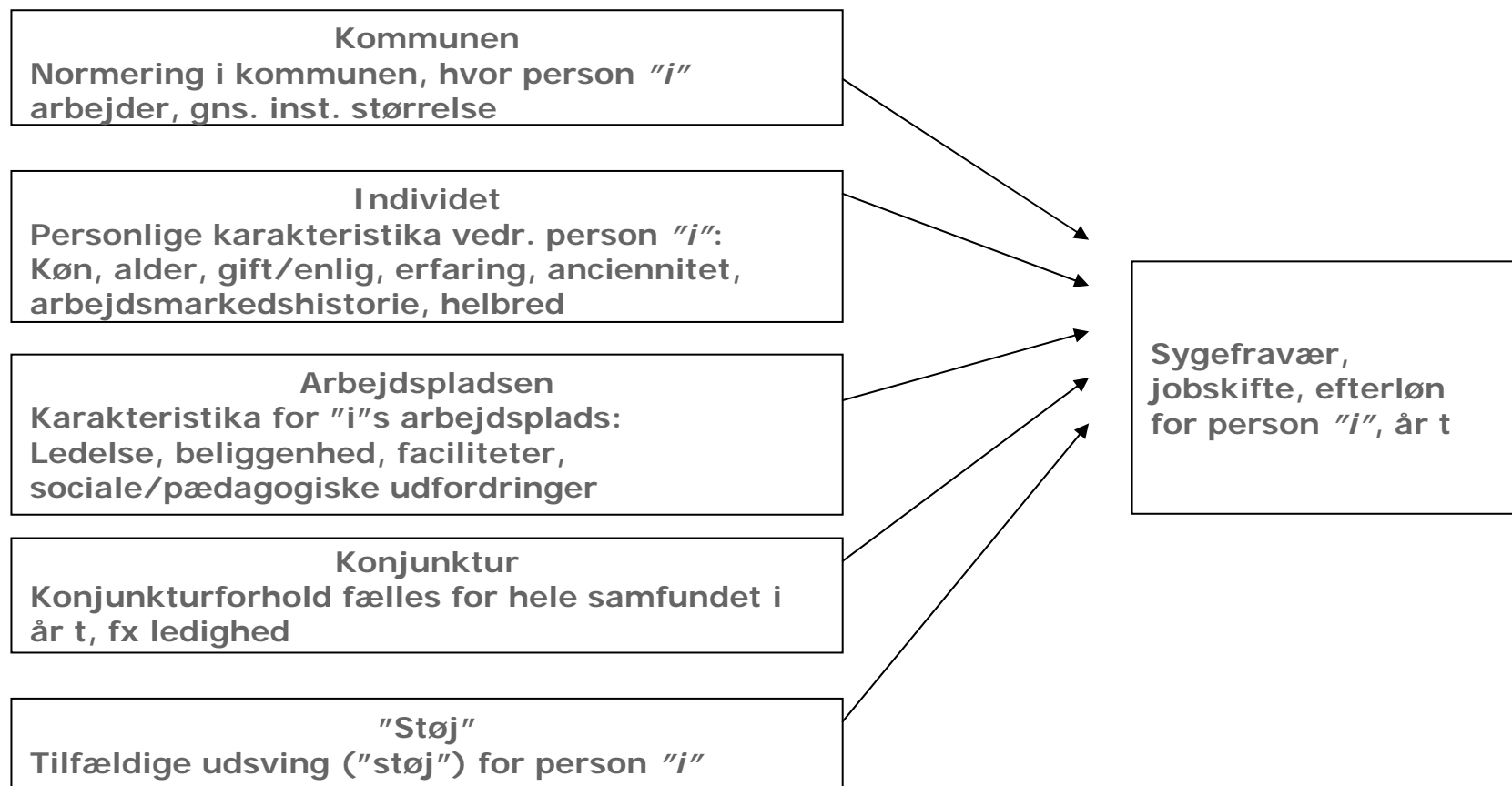
Case 1: Sygefravær i daginstitutioner

- Spørgsmål: Hvad betyder normeringer for sygefravær, efterløn og jobskifte i kommunale daginstitutioner?
- Registerdata fra Danmarks Statistik
- Population:
 - Ansatte i børnehaver og vuggestuer
 - Panel 1995 – 2006, dvs. følge personer over tid
- Langvarigt sygefravær for 2001-2006 (dagpengestatistikken)
- Kort sygefravær i 2005-2006
- Helbred, herunder besøg hos egen læge, speciallæge, hospitalsindlæggelser mv.
- Børn i daginstitutioner via Daginstitutionensregistret

Case 1: Sygefravær i daginstitutioner



Case 1: Sygefravær i daginstitutioner



Case 1: Sygefravær i daginstitutioner

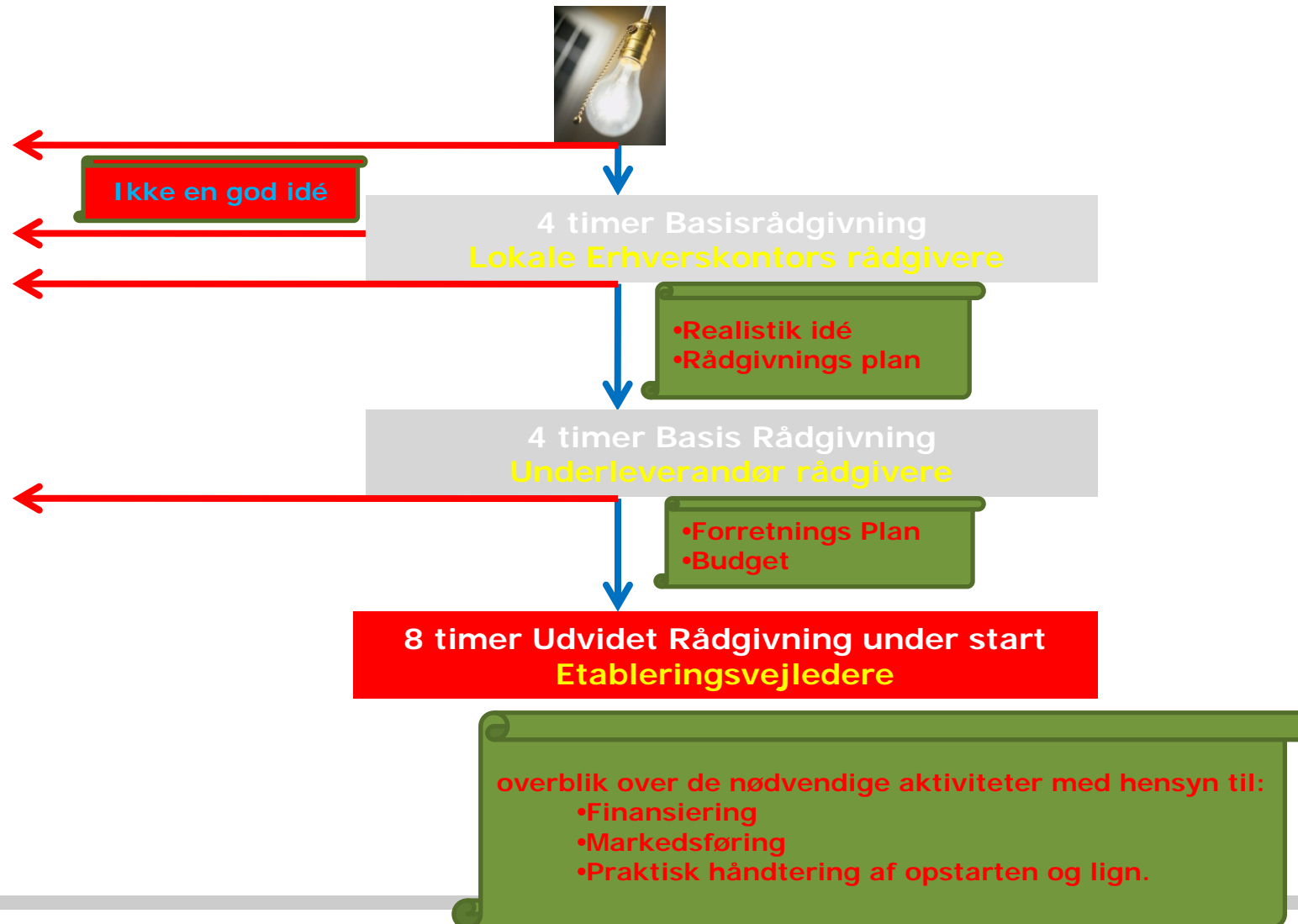
Hovedresultater:

- Blandet billede, men generelt positiv sammenhæng mellem normeringer og sygefravær i delperioder og for grupper af medarbejdere → arbejdspress øger sygefravær.
- Normeringer også forbundet med jobskifte, så flere søger væk fra kommuner med ringe normeringer.
- Normeringer ikke nogen tydelig (signifikant) effekt på tilgang til efterløn.

Case 2: NiN – Nordjysk Iværksætter Netværk

- Nordjysk Iværksætter Netværk (NiN) yder støtte i form af rådgivning før og under start af ny virksomhed
- Netværket er en del af regionale erhvervsprogrammer
- AKF evaluerede effekter af rådgivningsforløb for Erhvervs- og Byggestyrelsen
- Outputmål: Overlevelse, jobskabelse og omsætning (vækst) i nye iværksætter-virksomheder
- Problem: Finde passende kontrolgruppe, dvs. iværksættere med lignende karakteristika, som ikke modtog støtte fra NiN

Case 2: NiN rådgivningstyper



Case 2: Data

- Administrativt datasæt fra NiN med virksomheder registreret under deres CVR-nummer
- NiN-oplysninger: Type af rådgivning, ønsket branche
- Både indsats- og kontrolgruppe stammer fra NiN's egne registre
- Koblet med data fra Danmarks Statistik via CVR-nummeret
- Oplysninger om:
 - Overlevelse
 - Omsætning
 - Beskæftigelse

Case 2: Metode og hovedresultater

- Metode: Matching – virksomhed, der modtog rådgivning, matches med lignende virksomheder, der ikke modtog rådgivning på samme trin.
- Resultat: Deltagelse i NiN's rådgivningsforløb øger chancen for overlevelse
- Også positiv effekt på omsætning og beskæftigelse
- Erhvervs- og Byggestyrelsen ønsker nu at evaluere flere af deres programmer. NiN-evalueringen standard for fremtidige evalueringer.

Case 3: Analyse af den socioøkonomiske struktur i Fredericia kommune

- Analyse gennemført af AKF for Fredericia kommune
- Data: ECO Nøgletal og registerdata
- Fredericia sammenlignet med Vejle, Kolding og Middelfart.
- Demografi: flere førskolebørn, færre skolebørn, flere ældre.
- Overførselsudgifterne høje
- Flere kun folkeskole, få med videregående uddannelse
- Tilflyttere og fraflyttere – skæv uddannelsesprofil
- Relativt mange arbejdspladser, men mange indpendlere og dermed få skattekroner
- Relativt mange brancher med høj grad af nedslidning

Case 4: Virkninger af kommunalreformen for de ansatte i kommuner og regioner

- Observation: Sygefraværet stigende omkring kommunalreformen for kommuner, der fusionerede
- Analyse: Udvikling i sygefravær i kommuner, der blev sammenlagt, sammenlignes med udvikling i kommuner, der ikke blev sammenlagt (men lignende størrelse og sammensætning)
- Metode: Diff-in-diff
- Data: Registerdata for sygefravær, tilbagetrækning, jobskifte
- Kombineres med sygefravær fra FLD (Fælleskommunale Løndata)
- Kombineres med survey til kommunale ledere vedr. håndtering af organisationsforandringer

Case 4: Virkninger af kommunalreformen for de ansatte i kommuner og regioner

- Spørgsmål: Hvorfor er det en fordel at anvende individdata (registerdata)
 - Mange observationer – muliggør opdeling på undergrupper, fx se, om bestemte faggrupper klarer sig bedre end andre igennem organisationsforandringer
 - Følge personer over tid, også længere tilbage end reformen

Afslutning: Diskussion

- Kan registerdata anvendes til evaluering i jeres organisation
- Hvilke perspektiver ser I i muligheden for at kombinere surveys, administrative data, data fra eksperimenter mv. med registerdata?
- Muligheder for praktisk gennemførelse af evaluering?
- Behov for mere viden?
- Metoder til effektevaluering, herunder statistisk metode – behov for støtte og rådgivning?