

# Oppdragsbeskrivelse: Analyse av hvilken effekt tildelinger av FRIPRO-midler har på forskernes vitenskapelige produksjon

## Prosjektidé og –formål:

Formålet med dette prosjektet er å måle hvilken effekt tildeling av prosjektmidler i Forskningsrådets ordning for fri prosjektstøtte (FRIPRO) har på de tildelte forskernes vitenskapelige produktivitet og kvalitet. Dette innebærer ikke bare å måle de direkte resultatene ("output") i form av vitenskapelige publikasjoner generert i de konkrete prosjektene, men å se om slik tildeling har langsiktig effekt på forskernes framtidige vitenskapelige resultater. To analyser skal utføres: En som måler effekt i vitenskapelig produktivitet i form av antall fagfelleverderte vitenskapelige publikasjoner, og en som måler bruk (som uttrykk for kvalitet) gjennom siteringer i slike.

Det er mange faktorer utover tildeling av en bestemt type forskningsmidler som potensielt påvirker forskernes vitenskapelige resultater. Hensikten med denne analysen er å prøve å isolere effekten av tildelinger fra FRIPRO fra disse øvrige faktorene ved å sammenlikne forskere med og uten slik tildeling.

## Bakgrunn:

FRIPRO er en åpen, nasjonal konkurransearena som favner alle fag og disipliner uten at det legges tematiske føringer eller settes krav om anvendelse og umiddelbar nytteverdi. FRIPRO skal fremme vitenskapelig kvalitet i internasjonal forskningsfront, dristig og nyskapende forskning og karriere for unge forskertalenter. FRIPRO er et av Forskningsrådets sentrale virkemidler for å fremme fri, forskerinitiert grunnleggende forskning og har i sin nåværende form, organisert samlet innenfor Divisjon for vitenskap, eksistert siden utlysningen i 2004. Det var også dette året Forskningsrådet innførte elektronisk søknadsskjema som betyr at datakvaliteten er betydelig bedre fra og med dette året sammenliknet med årene før.

I hele denne perioden har søknadstypen *forskerprosjekt* vært gjennomgående brukt i FRIPRO. Det er en søknadstype som er rettet mot etablerte forskere. Fra og med utlysningen i 2013 har det vært et krav at det for prosjektleder må ha gått minst seks år siden disputas eller oppnådd førstestillingskompetanse. Prosjektlederen og -gruppen må være i internasjonal forskningsfront.

I perioden 2004 – 2015 har Forskningsrådet tildelt prosjektstøtte til totalt 1159 forskerprosjekter med 855 forskjellige prosjektledere innenfor FRIPRO.

## Datakilder:

Det er tre hoveddatakilder som er nødvendige for å gjennomføre denne studien:

- Oversikt over alle norske forskere
- Oversikt over forskerprosjekter i FRIPRO med navn på prosjektledere
- Oversikt over publikasjoner med norske forfatteradresser

### Forskerpersonalregisteret:

Forskerpersonalregisteret (FPR) er en database over forskerpersonalet ved universiteter, høyskoler, helseforetak, forskningsinstitutter og andre institusjoner med FoU, og inngår som en del av den nasjonale FoU-statistikken. Registeret omfatter ikke forskere i næringslivet. Registeret oppdateres i oktober hvert år og har følgende opplysninger: Fødselsnr., Navn, Stilling (-kode, -tekst, -prosent), Institusjon, Fagtilhørighet, Grad, Fag, År og sted for embetseksamen, Dr.grad (evt.) og år for disputas. Det er Nordisk institutt for studier av innovasjon, forskning og utvikling (NIFU) som samler inn og vedlikeholder registeret som et ledd i arbeidet med nasjonal FoU-statistikk. I følge denne statistikken for 2016 er det registrert 31.844 forskere pr. 01.10.2014 fordelt på 20.090 i UoH-sektoren, 7.491 i instituttsektoren og 4.263 i helseforetakene.

### Forskningsrådets datavarehus:

Datavarehuset gir oversikt over alle tildelinger i Forskningsrådet. Relevant informasjon som kan hentes ut fra denne basen er: Navn på prosjektleder, Søknadstype, Prosjekttype, Fagområde, Fag, Oppstartsår, Tildelt beløp pr. år. For perioden 2004 – 2015 er det registrert totalt 1174 forskerprosjekter i FRIPRO. Av disse er 15 en annen prosjekttype enn 'Prosjektstøtte', f.eks. 'Personlig stipend'. Ser vi bort fra disse er det totalt 1159 prosjekter med 855 forskjellige prosjektledere (705 fra UoH-sektoren, 156 fra instituttsektoren – noen få kan være registrert begge steder). Det er viktig å merke seg at det i de fleste slike prosjekter er en gruppe av forskere som inngår, og at flere enn prosjektlederen har sentrale roller i prosjektene, f.eks. stipendiater og andre forskere ansatt ved prosjektansvarlig institusjon eller samarbeidende institusjoner. Disse forskerne vil også kunne ha en potensielt målbar effekt av at prosjektet har blitt tildelt midler. Imidlertid gir ikke datavarehuset tilgang til informasjon om disse forskerne, så studien må begrense seg til prosjektledere.

### Publikasjonsdatabaser:

De to store databasene for vitenskapelig publisering og siteringsindekser er Scopus og Web Of Science (WoS).

For perioden 2005 – 2016 inneholder Scopus totalt 189.185 publikasjoner med minst én norsk forfatteradresse, hvorav 126.740 er 'Article'. Av disse er f.eks. 44.711 innenfor 'Medicine', 13.985 innenfor 'Social Science' og 4.107 innenfor 'Arts and Humanities'.

For samme periode inneholder WoS 146.640 publikasjoner med minst én norsk forfatteradresse, hvorav 116.396 er 'Article'. Her er det en mye finere inndeling av fagområder, så det er ikke direkte sammenliknbart med Scopus.

Begge disse databasene har en lavere dekning av publikasjoner innenfor humaniora og samfunnsvitenskap enn innenfor medisin, matematikk og naturvitenskap/teknologi. Scopus har noe høyere dekning i medisin enn WoS, mens forholdet er motsatt for naturvitenskap og teknologi. Scopus, som er europeisk basert, dekker 22.000 tidsskrifter og har en overvekt av europeiske artikler og journaler. WoS er amerikansk og har derfor en overvekt av amerikanske journaler. Databasen dekker ca. 12.000 tidsskrifter. Av disse to databasene dekker WoS en betydelig lengre historisk periode, men Scopus har god dekning fra 1996, og har etter hvert også tilfredsstillende dekning før dette.

CRISTin (Current Research Information System in Norway) registrerer publikasjoner og andre forskningsresultater i Norge og administrerer Norsk vitenskapsindeks (NVI) som er en felles publikasjonsdatabase for universiteter og høyskoler, instituttsektoren og de regionale helseforetakene. Denne har ikke fullstendig dekning, spesielt ikke for instituttsektoren før 2011.

Dette betyr at dersom produktivetsanalysen skal baseres på mer komplette publikasjonslister enn de som finnes i WoS og Scopus, må NVI-basen suppleres med informasjon fra andre kilder.

### **Indikatorer for vitenskapelig effekt:**

Å måle effekt av forskning er hverken enkelt eller entydig. Det gjelder heller ikke når man vil måle effekten av tildelte midler til grunnleggende, fri forskning. I denne studien ønsker vi å bruke publiseringsvolum og siteringer som mål for vitenskapelig effekt fordi dette er målbare størrelser som gjenspeiler grunnleggende aktiviteter knyttet til forskningen, nemlig å dokumentere resultatene gjennom fagfellevurderte publikasjoner og å referere til andres eller egen forskning. Samtidig er det omstridt hvorvidt f.eks. siteringer er det beste mål for vitenskapelig kvalitet og om det måler hvilken forskning som på lang sikt vil ha største effekt.

På grunn av store forskjeller mellom fag- og fagområder når det gjelder publiserings- og siteringstradisjon, og ulik dekning av disiplinene i de tilgjengelige publikasjonsdatabasene, bør analysen basere seg på indikatorer som er mest mulig fag-/disiplinnøytrale eller på en metodikk som justerer for slike skjevheter.

Publiseringsvolum: Vitenskapelig produktivitet måles ved å telle antall publikasjoner for hver enkelt forfatter. Publikasjoner som er poenggivende i CRISTin-systemet teller, dvs. vitenskapelige artikler i monografier, antologier, periodika, serier og nettsteder som oppfyller kravene til vitenskapelig publisering (se nettsidene til Norsk publiseringsindikator ([npi.nsd.no](http://npi.nsd.no)) for detaljer). Publikasjonene kan telles enten ved *heltelling*, der hver forfatter gis ett poeng for hver publikasjon, ved *deltelling*, der hver forfatter får poeng lik sin andel av totalt antall forfattere (forfatterandel) eller ved en kombinasjon av disse. Det bør videre tas hensyn til lengden på perioden det måles over for eksempel ved å dele på antall år siden (første) tildeling.

Siteringer: For de fleste publikasjoner er det slik at de blir sitert hyppigst i løpet av de første årene etter utgivelse. For de yngste utgivelsene vil det likevel kunne representere en ulempe at de har kortere "virketid" enn de eldre. Det bør derfor benyttes et siteringsvindu på f.eks. tre år slik at det er siteringer i de tre påfølgende kalenderår etter utgivelsesår som teller. Også her foreslår vi at to ofte benyttede bibliometriske indikatorer vurderes. Det er *relativ siteringsindeks*, altså antall siteringer for en forfatters artikler som andel av verdensgjennomsnittet innenfor samme fag og samme utgivelsesår, eller *andel artikler blant de 10 prosent mest siterte* innenfor samme fag og samme utgivelsesår. Det må gjøres en vurdering om kun en av disse indikatorene bør benyttes eller om en eventuell kombinasjon er mest hensiktsmessig. En utfordring kan være at andel ofte siterte artikler blir en ustabil indikator når den benyttes på individ- og ikke på institusjonsnivå.

### **Effektanalyse:**

Potensielle prosjektledere for forskerprosjekter i FRIPRO er alle norske forskere (registrert i FPR) som har vært aktive i den perioden analysene skal dekke. Av denne gruppen er det en betydelig mindre gruppe som faktisk har fått tildeling. Metodisk kan tildeling av forskerprosjekt ses på som 'behandling', og de med slik tildeling er da gruppen av 'behandlede', mens de øvrige (potensielle prosjektledere) utgjør en 'kontrollgruppe'. Formålet med analysene er dermed å estimere den gjennomsnittlige behandlingseffekt i form av uttelling på de valgte indikatorene for vitenskapelig produktivitet og kvalitet.

For å måle effekter må tildelingene gis en rimelig tid til å virke. Det vil normalt sett gå noen år fra en tildeling til effekten av denne gir seg utslag i publiserte artikler, både fordi forskningen faktisk må utføres og at utgivelsesprosessen tar tid. I tillegg vil det gå ytterligere noe tid før publikasjonen blir

sitert. Populasjonen som benyttes i analysen for produktivitet (publikasjonsvolum) defineres derfor til prosjektledere med tildelte forskerprosjekter i FRIPRO i perioden 2004 – 2013 (1018 tildelinger fordelt på 755 prosjektledere), og publikasjoner utgitt til og med 2016 inngår i tellingen. I analysen for kvalitet (siteringer) begrenses populasjonen av forskere med tildeling til perioden 2004 - 2010 (723 tildelinger fordelt på 554 prosjektledere). Siteringer telles til og med 2016, men de publikasjonen det refereres til i siteringsanalysen begrenses til utgivelser til og med 2013 ((for å gi rom for et tre års siteringsvindu). Tilsvarende begrensninger gjøres for kontrollgruppene.

Ettersom gruppen av 'behandlede' og kontrollgruppen ikke kan antas å ha samme utgangspunkt for sine vitenskapelige resultater, oppstår det en utvalgsskjevhet ('selection bias') som det må kontrolleres for. Det kan være grunn til å anta at forskere som når opp i den harde konkurransen om FRIPRO-tildelinger i utgangspunktet har et høyere nivå av vitenskapelige resultater enn gjennomsnittet i kontrollgruppen, og at de derfor vil skåre høyt også uten slik tildeling. Behandlingseffekten kan altså ikke bare måles ved å se på forskjellen mellom de to gruppene etter behandling, og heller ikke bare i forskjell i utvikling over en periode fordi det kan være underliggende forskjeller mellom gruppene som forklarer ulik utvikling.

For å kontrollere for utvalgsskjevhet må de variablene som kan tenkes å både påvirke om forskeren får tildeling av midler og han eller hennes skår på de vitenskapelige indikatorene identifiseres. Slike variabler kan være indikatorer for vitenskapelig produktivitet og kvalitet *før* effekten av tildeling slår inn, alder, antall år siden disputas, stilling, institusjonstilhørighet og fag-/disiplintilhørighet.

For hver forsker i behandlingsgruppen finnes det informasjon om når første tildeling fant sted. Noen forskere har fått flere tildelinger, det kan derfor vurderes om det også kan måles om ytterligere tildelinger forsterker en eventuell effekt av tildeling, dersom datagrunnlaget er godt nok for en slik analyse.

Dersom datagrunnlaget tillater det, vil det være interessant å se om det er forskjeller mellom de forskjellige fagene eller faggruppene med hensyn til effekt av tildelte prosjektmidler. Dette kan eventuelt undersøkes på overordnet plan mellom de tre hovedgruppene innenfor FRIPRO: Humanistiske og samfunnsvitenskapelige fag (FRIHUMSAM), naturvitenskapelige og teknologiske fag (FRINATEK) og medisinske og biologiske fag (FRIMEDBIO).

### **Metodiske utfordringer:**

Begrensninger i datagrunnlaget gir noen metodiske utfordringer som det må tas hensyn til i utformingen av analysen:

Utfordring 1: En kompliserende faktor kan være at det finnes noen 'falske ikke-behandlede' i kontrollgruppen. Ettersom det kun er prosjektledere som identifiseres er det trolig at det blant kontrollgruppen finnes forskere som har hatt en effekt av FRIPRO-tildeling fordi de har vært sentrale i prosjektgjennomføringen, men ikke hatt status som prosjektledere. Det kan også tenkes at forskere som tidligere har vært mottakere av midler uten å være prosjektledere har blitt prosjektledere på et senere tidspunkt. Disse vil da bli registret som 'behandlet' senere enn de i praksis oppnådde en eventuell effekt av behandling.

Utfordring 2: Det kan være vanskelig å avgjøre om en publikasjon er utarbeidet før eller etter tildeling av prosjektmidler selv om den blir utgitt etter tildeling. Dette skyldes at det kan være en lang periode fra artikkelen skrives til den offisielt blir utgitt. Det vil da si at den kan komme til å telle med som en effekt av tildeling selv om arbeidet faktisk ble utført før. Dette problemet må løses på en pragmatisk måte.

**Leveranse:**

Resultatene av effektanalysene skal presenteres i en rapport på norsk som utarbeides av oppdragstaker. Det må redegjøres på en oversiktlig og systematisk måte for eventuelle forskjeller i publiserings- og/eller siteringsprofil mellom forskere med tildeling og kontrollgruppen. Det samme gjelder eventuelle forskjeller mellom fagområder eller andre funn som analysene avdekker. Oppdragstaker må også gi en vurdering av de hovedfunnene som avdekkes og trekke konklusjoner av analysen. Rapporten må redegjøre for tolkning og avgrensning av problemstillingene, oppgi hvilke datakilder som er benyttet, beskrive metoder som er brukt og hvilke forutsetninger som ligger bak. Det skal også skrives et ikke-teknisk sammendrag som inneholder viktigste funn og konklusjoner. Resultatene fra analysen skal i tillegg presenteres muntlig i et sluttseminar hos Forskningsrådet.

**Eventuelt oppfølgingsprosjekt:**

Dette prosjektet kan ses på som en første fase for å undersøke om det å måle vitenskapelig effekt av tildeling av forskningsmidler på en kvantitativ måte er hensiktsmessig. Vi ønsker derfor at dette prosjektet skal legge praktisk til rette for at lignende analyser kan gjøres i senere faser gjennom bruk av samme eller noe modifiserte metoder og verktøy. En mulig oppfølging er å utvide analysen til forskerprosjekter innenfor flere av Forskningsrådets aktiviteter, for eksempel de tematiske og handlingsrettede programmene.

**Tidsramme:**

Oppdraget må utføres i løpet av fire måneder fra kontrakt mellom oppdragsgiver og oppdragstaker er inngått.

**Krav til tilbudet:**

Tilbyder må redegjøre tydelig for hvordan oppdraget er forstått. Tilbudet må beskrive hvordan oppdraget skal løses i form av metodevalg, bruk av data, hvilke avgrensninger som må gjøres og hvilke forutsetninger som tas. Det må også redegjøres for mulige feilkilder og hvor pålitelige resultatene er. Det må argumenteres for de valg av modell og metode som gjøres, gjerne stilt opp mot andre mulige alternativer.

Tilbudet bør henvise til litteratur som underbygger den metoden som velges og som har relevans for dette prosjektets problemstilling.