



FFI Forsvarets
forskningsinstitutt

24/00690

FFI-RAPPORT

Hvordan har produktivitet i Forsvarsmateriell utviklet seg?

– modellering og analyse

Petter Fredrik Hemnes

Hvordan har produktivitet i Forsvarsmateriell utviklet seg? – modellering og analyse

Petter Fredrik Hennes

Emneord

Produktivitet
Effektivitet
Styring og kontroll
Rapportering

FFI-rapport

24/00690

Prosjektnummer

1625

Elektronisk ISBN

978-82-464-3535-0

Engelsk tittel

How has productivity in the Norwegian Defence Material Agency developed? – modeling and analysis

Godkjenner

Ane O. Presterud, *forskningsleder*
Steinar Gulichsen, *forskningssjef (fungerende)*

Dokumentet er elektronisk godkjent og har derfor ikke håndskreven signatur.

Opphavsrett

© Forsvarets forskningsinstitutt (FFI). Publikasjonen kan siteres fritt med kildehenvisning.

Sammendrag

For å måle effekten av forbedringsarbeid i en organisasjon må vi forstå forholdet mellom innsatsfaktorer og produksjon, altså produktivitet. I Forsvarsmateriell er det tre momenter som har gjort det vanskelig å analysere produktivitet: For det første er det ikke definert hva som er etatens produkter og tjenester. For det andre er leveransene Forsvarsmateriell står for, gjerne et resultat av komplekse prosesser. For det tredje omsettes ikke produksjonen i markeder, og produktene har derfor ikke priser som signaliserer hvilken verdi forskjellige leveranser har opp mot hverandre og for samfunnet.

Vi foreslår å bruke dataomhyllingsanalyse som metode for å imøtekomme disse utfordringene. Denne metoden gir fleksibelt valg av innsatsfaktorer og produkter og belager seg dessuten ikke på at det må være kjent hvordan priser eller mekanikker gjør innsatsfaktorer om til produkter. Dette gjør metoden ypperlig for å måle produktivitet i komplekse organisasjoner som opererer utenfor markedet, slik som Forsvarsmateriell.

I denne rapporten benytter vi dataomhyllingsanalyse til å estimere produktivitet på investeringsområdet i Forsvarsmateriell i perioden 2018 til 2023. Vi illustrerer hvordan denne modellen kan benyttes for å anslå utviklingen i produktivitet og objektive normer for produktivitet. Vi benytter detaljerte data på registrerte milepæler i investeringsprosjekter for å approksimere produksjonen. Vi benytter detaljerte data på registrerte milepæler i investeringsprosjekter for å beregne produksjonen. Spesifikt benytter vi milepælene som representerer fire leveranser som Forsvarsmateriell har ansvar for: produksjonen av sentrale styringsdokumenter, utarbeidede konkurransegrunnlag og kontraktsgjennomføringer, fremsendte materielloverføringsprotokoller og termineringsrapporter. På innsatsfaktorsiden benytter vi kostnader tilknyttet arbeidskraft (fastlønn, overtid, tillegg, annen lønn og konsulentkostnader), kostnader knyttet til kapital (hus- og maskinleie, avskrivninger, vedlikehold og drift) og arbeidstimer i prosjekter.

Resultater fra seks modeller antyder at Forsvarsmateriell har hatt en positiv produktivitetsutvikling i perioden 2018 til 2023. Dette stemmer med forståelsen vår av hvordan produktiviteten har utviklet seg i etaten, og vi diskuterer produktivitetsendringene i lys av annen forskning på og andre analyser av drivere av produktivitetsendringer i Forsvarsmateriell. Videre finner vi at produktivitetsendringene har liten sammenheng med rapportert resultatoppnåelse. Samtidig er estimatene på produktivitet usikre på grunn av lav kvalitet på dataene fra tidlige perioder og grovkornede data på innsatsfaktorbruk for deler av perioden. I tillegg er det heller ikke justert godt for kvalitetsforskjeller i produksjonen.

Vi anbefaler at Forsvarsmateriell viderefører måling og analyse av produktivitet slik som det er gjort i denne rapporten. Modellen bør utvides til å inkludere de andre områdene Forsvarsmateriell har ansvar for, spesielt forvaltningsområdet. Et første steg er å samle data og inkorporere tilsynsaktivitet og avdekkede avvik i modellen. Vi anbefaler også at modellen videreutvikles til å inkludere kvalitetsindikatorer for produkter og tjenester som leveres både på investeringsområdet og forvaltningsområdet.

Summary

To assess the impact of efficiency, quality, and effectiveness initiatives in an organization, it is crucial to grasp how inputs and outputs relate, defining productivity. In the Norwegian Defence Material Agency (NDMA) it has proven difficult to analyze this relationship. Three aspects make productivity analyses difficult: First, it is unclear which deliveries are to be considered final products or services. Second, NDMA delivers a spectrum of products and services, many of which require complex processes to be produced. Third, the value of each product or service must be known to derive the total factor productivity, but as NDMA operate outside of markets, prices—a natural approximation of value—is absent.

Data envelopment analysis, or DEA for short, can be used to overcome these issues. DEA provides considerable flexibility in selection of inputs and output. Moreover, the absence of prices or prior knowledge of the complexity of production is not an issue in DEA. Consequently, DEA can be used to assess productivity in NDMA.

In this report, we use DEA to estimate productivity in NDMA's investment activities from 2018 to 2023. We illustrate how a DEA model can be used to examine trends in productivity and identify best-practice periods. We use data on registered milestones in investment projects to approximate production. We focus on four milestones: production date of project initiation documents, date of contract signing, date of material deliveries, and production date of project termination reports. When it comes to input, we gather and use data relating to labor and capital costs as well as hours worked in investment projects. Inputs and outputs are varied in six different DEA models.

We find that NDMA has increased its productivity from 2018 to 2023. This aligns with our general understanding of the productivity trend in the organization. We discuss the changes in light of other research and analyses relating to productivity in NDMA. Our findings also indicate that productivity and goal attainment, as reported partially by NDMA, are largely unrelated. However, our findings should be interpreted with some caution due to limitations in the data. Most notably, our analysis suggest that the quality of the data is low prior to 2020. Moreover, due to high variance, several productivity estimates do not differ significantly from each other. Lastly, we do not have data that allow us to adjust for quality differences in production, which could mean that productivity is not being adequately measured.

We advocate the continuation of performance and efficiency measurement in NDMA, and further development of the model presented in this report. First and foremost, other deliverables that NDMA are responsible for, such as those relating to material readiness, should be included in the model. A first step would be to gather data on, and include in the model, the number of material inspections NDMA undertakes as well as any findings of non-compliance. We also view the lack of quality indicators as a severe limitation to our analysis. As such we recommend that quality indicators be developed and measured.

Innhold

Sammendrag	3
Summary	4
1 Innledning	7
2 En enkel produktivitetsmodell	9
3 Data	14
4 Analyse og resultater	20
4.1 Grafisk fremstilling av utviklingen i produksjon og innsatsfaktorbruk	20
4.2 Produktivitet	22
4.3 Kalibrering og usikkerhet	27
5 Hva skiller gode fra dårlige perioder?	30
6 Produktivitet og tertialrapportering	35
7 Videreutvikling	37
7.1 Måleparametere i investeringsprosessen	37
7.2 Måleparametere i forvaltningsprosessen	38
8 Konklusjon og videre studier	40
8.1 Videre studier	41
Referanser	42
Vedlegg	46
A Dataomhyllingsanalyse	47
A.1 Optimeringsproblemet	47
A.2 Fra ratio til lineær formulering	50
B Simar og Wilson (1998, 2007) – kort forklart	51

C	Prosjektkompleksitet	54
D	Alternative og numeriske produktivitetsskårer	57
E	Milepæler	59
F	Illustrasjon av datasett	61

1 Innledning

Forsvarsmateriell (FMA) har siden sin opprettelse i 2016 jobbet målrettet for profesjonalisering og effektivisering av virksomheten, men har manglet klare kvantitative indikatorer på verdien av dette arbeidet. Fraværet av kvantitative indikatorer har også gjort det vanskeligere å stadfeste etatens utvikling samt kommunisere denne utad, eksempelvis til Forsvarsdepartementet (FD). Gjentakende innsamling, evaluering og rapportering av informasjon på effektivitet er sentralt for å forstå forbedring og modernisering i FMA. Det kan også bidra til økt transparens, og aggregerte produktivitetsmål kan gi lavere behov for detaljstyring.

Det er krevende å kvantifisere verdien av vare- og tjenesteproduksjonen i FMA. Årsaken til dette er fraværet av velfungerende markeder for det som produseres. Dette fører blant annet til at produkter og tjenester sjeldent er godt definerte og at den relative verdien på produksjonen, som ellers ville vært reflektert i markedspriser, ikke eksisterer. Å få oversikt over produksjonsendringer og forstå sammenhengen mellom produksjon og innsatsfaktorbruk er derfor krevende. I denne rapporten gjør vi et forsøk på å løse denne utfordringen ved å foreslå en produktivetsmodell for FMA. Vi anvender også modellen empirisk på investeringsområdet og bruker historiske data fra FMA for å vurdere utviklingen i produktivitet over tid.

Vårt bidrag er nyttig for Forsvarsmateriell på to måter. For det første kan modellen og analysen benyttes til å skape objektive ankerpunkter for høy og lav produktivitet i etaten. Slike ankerpunkter kan videre benyttes for å stadfeste utvikling, forklare forskjeller i resultater og identifisere og evaluere tiltak for ytterligere prosessforbedring i Forsvarsmateriell. For det andre er modellen utformet på en måte som gjør den anvendelig for FMA. Den er lett å forstå, benytter data som er godt tilgjengelige, og er enkel å beregne. Dette gjør det mulig for etaten selv å samle inn og analysere sin egen produktivitet fremover. I tillegg gir modellen mulighet for å grave dypere i de spørsmålene som ikke besvares i denne rapporten. Herunder bidrar også modellen og analysen til å oppnå ambisjoner rundt effektivitetsanalyser nedfelt i forsvarssektorens instruks for økonomi- og virksomhetsstyring (Forsvarsdepartementet, 2019a). Dette er ambisjoner som virker like relevante i dag som for fem år siden og som har fått økt fokus som følge av prosjektet «Forsvarssektoren 2024» (F24) (Forsvaret, 2024). Til tross for å være et viktig utgangspunkt dekker ikke modellen hele virksomhetsområdet til Forsvarsmateriell. Forvaltningsområdet er, grunnet fraværet av data, holdt utenfor den empiriske analysen. Kvalitet i produksjonen er heller ikke godt fanget opp. Dette motiverer videre analyse og modellering, og vi bidrar her med forslag til hvordan modellen vi foreslår kan videreutvikles for å inkorporere fasetter ved forvaltningsområdet.

Strukturen i denne rapporten er som følger: I kapittel 2 beskriver vi produktivetsmodellen, valg av produkter og ressurser og valg knyttet til metode. I kapittel 3 forklarer vi hvilke data vi benytter i modelleringen, kvaliteten på datasettene og valg knyttet til datasettene. Vi presenterer hovedresultatene i kapittel 4 hvor produksjon og innsatsfaktorbruk i Forsvarsmateriell knyttes sammen for å forklare produktivetsutviklingen. I kapittel 5 diskuterer vi mulige forklaringer til endringene i produktivetsestimaterne over tid, og i kapittel 6 ser vi resultatene opp mot Forsvarsmateriell

sin egen rapportering på resultatoppnåelse. I kapittel 7 foreslår vi hvordan modellene kan videreutvikles, før vi konkluderer og foreslår videre studier i kapittel 8.

Målgruppen for rapporten er ledelse i FMA og ansatte i virksomheten som jobber med, eller har interesse for, virksomhetsstyring, forbedring og effektivisering. Vi har forsøkt å gjøre rapporten leselig for alle og forutsetter derfor minimalt med forkunnskaper om produktivitetsanalyse.

2 En enkel produktivetsmodell

I rapporten benytter vi en modell som gjør det mulig å skille mellom perioder med høy produktivitet og lav produktivitet og observere utviklingen over tid. I dette kapitlet introduserer vi denne produktivetsmodellen nærmere. Vi forklarer begreper, forklarer hvilke modelleringsvalg vi har tatt, hva disse innebærer, og hvorfor våre valg er fornuftige. Vi begynner med å definere produktivitet.

Produktivitet defineres ved forholdet mellom produkter og innsatsfaktorer (Charnes et al., 1994; DFØ, 2013). For en virksomhet med p produkter ($y \in \mathbb{R}_+^p$) og m innsatsfaktorer ($x \in \mathbb{R}_+^m$) kan vi formulere produktiviteten i virksomheten som,

$$P_t \stackrel{\text{def}}{=} \sum_{i,j}^{p,m} \mu_i y_i / v_j x_j,$$

hvor P_t er produktiviteten i tid t , y_i er produksjon av produkt i , x_j er innsatsfaktorbruk av innsatsfaktor j , og μ_i og v_j er vekter for henholdsvis produktene og innsatsfaktorene. For virksomheter som produserer flere produkter og tjenester og benytter flere innsatsfaktorer i produksjonen er altså produktivitet definert som den vektete kombinasjonen av alt som produseres over den vektete kombinasjonen av alle ressursene som brukes.

Vektene som benyttes er sentrale for å beregne produktivitet. De lar oss samle produksjonen til én og samme enhet og lar oss fastslå den totale effekten av produksjon og innsatsfaktorbruk. En vekt for et produkt eller en innsatsfaktor kan tolkes som verdien til produktet eller innsatsfaktoren. Etter at alle vekter er fastslått vil vektene altså fortelle oss hvordan produkter og innsatsfaktor skal verdsettes opp mot hverandre. På produksiden vil en lav vekt (relativt til vekten på andre produkter) tilsi at produktet verdsettes lavt og en høy vekt vil tilsi at produktet verdsettes høyt. I velfungerende markeder er informasjon om verdi gitt ved priser. Priser er derfor naturlig å benytte som vekter. Ved fraværet av priser, slik som ofte er tilfellet for produksjon i offentlige virksomheter, må vi benytte andre metoder for å bestemme vektene.

I denne rapporten benytter vi dataomhyllingsanalyse (Data Envelopment Analysis – DEA) for å definere fornuftige vekter. Dataomhyllingsanalyse er en metode som innebærer å gi hver enhet, eller i vårt tilfelle hver periode, egne vekter. Metoden er noe teknisk og vi henviser derfor den interesserte leser til vedlegg A for en mer detaljert gjennomgang av metoden. For den generelle leseren holder det derimot å ha kjennskap til noen få egenskaper ved metoden. Den første er at metoden er spesielt egnet for produktivetsanalyser med flere produkter og innsatsfaktorer, hvor informasjon om priser på produkter er utilgjengelige, og det er lite informasjon om hvordan innsatsfaktorer transformeres til produkter (Charnes et al., 1978, 1994; Emrouznejad & Yang, 2018; Liu et al., 2013). Dette gjør metoden spesielt egnet for analyser av offentlige virksomheter og mer spesifikt FMA hvor produksjonsfunksjonen er vanskelig å modellere. Den andre egenskapen er at vektene til produktene og innsatsfaktorene velges ut ifra hva som maksimerer produktivitet i hver enhet eller periode. Disse vektene sies derfor å være de optimale vektene for hver periode. Dette betyr også at de periodene eller enhetene som vurderes som ineffektive

domineres på tvers av alle dimensjoner; det finnes ingen mulige vekter som gjør disse enhetene mer produktive enn alle andre enheter. Den siste og, for vårt formål, viktigste egenskapen ved DEA er at resultatene fra analysen er en produktivitetsskår mellom 0 og 1 for hver periode. Verdien 1 (=100 prosent) tilsier at perioden er effektiv og jo nærmere man er 1, desto mer effektiv er man. Resultatene er sådan lette å tolke selv for de uten kjennskap til DEA.

Produktene og innsatsfaktorene må også defineres for å regne ut produktivitet. Vi benytter de to aggregerte innsatsfaktorgruppene arbeidskraft og kapital, og vi avgrensner produksjonen til beslutningsunderlag i forprosjektfasen¹, konkurransegrunnlag og kontrakter, leveranser av materiell, termineringsrapporter, forvaltningsregelverk og tilsyn. Denne produksjonen dekker en stor del av Forsvarsmateriell sin virksomhet innen investeringer og forvaltning² og dekker sentrale ansvarsområder som FMA har i rollen som prosjektansvarlig i materiellinvesteringsprosjekter (Forsvarsdepartementet, 2019b). Vi gjennomførte også samtaler med kapasiteter i FMA og fant at denne produksjonen var mest fremtredende. Dette er produksjon som FMA også i stor grad har ansvaret for, og hvor produksjonen i liten grad påvirkes av andre aktører. FMA sitt arbeid som i forvaltningen søker å effektivisere drift og vedlikehold i Forsvaret velger vi å holde utenfor modellen av hensyn til endringer som vil følge av F24 (Forsvaret, 2024). For tjenesteproduksjon i denne delen av forvaltningen har data også vært utilgjengelige. Beskrivelse og deskriptiv statistikk for alle produkter og innsatsfaktorer vi benytter finnes i tabell 3.1.

For å benytte produktene i produktivitetsanalyser må vi gjøre antagelser eller korrigeringer for forskjeller i kvalitet og kompleksitet mellom produktene. Vi antar at samtlige produkter omfatter omtrent det samme over tid. Eksempelvis antar vi at det ikke er en fundamental forskjell mellom hva et sentralt styringsdokument (SSD) var i 2018 og hva en SSD var i 2023 for lignende prosjekter. For å justere for kvalitetsforskjeller inkluderer vi en kvalitetsindikator, og for å skille mellom kompleksitet i produktene deler vi dem inn i to grupper. Kompleks produksjon i FMA strekker seg over lengre tid enn enkel produksjon. For å sammenstille enkel og kompleks produksjon i tid inkluderer vi delproduksjon i kompleks produksjon. Eksempelvis teller vi kun produksjonen av SSD som beslutningsgrunnlag for enkle prosjekter, men produksjonen av både oppdragsanalyse og analyse av livssyklus kostnader sammen med SSD som beslutningsgrunnlag for komplekse prosjekter. Vi kommer tilbake til dette i kapittel 3.

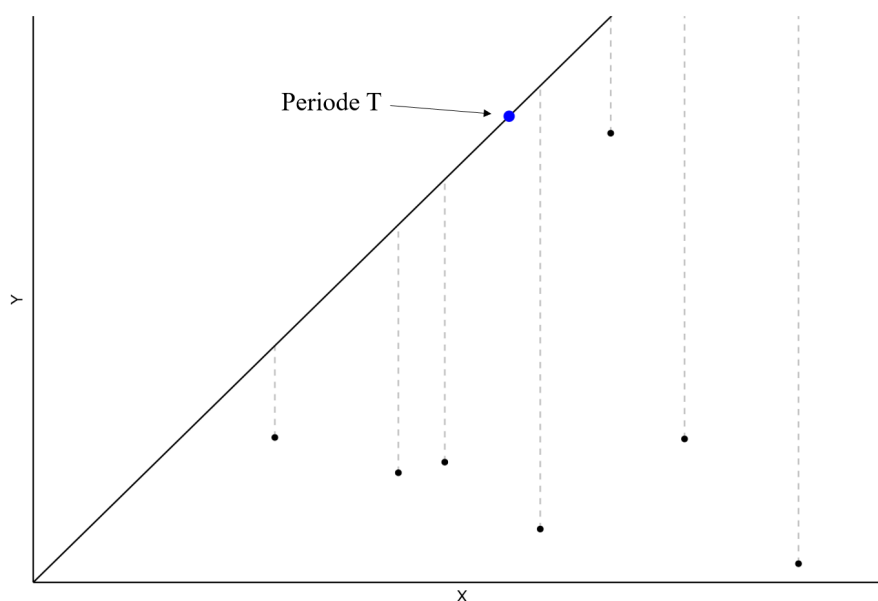
Vi kjenner ikke til produksjonsteknologien i Forsvarsmateriell. Med andre ord vet vi ikke på hvilken måte ressurser kan transformeres til produkter på en ideell måte. Følgelig vet vi heller ikke hva den forventede produksjonen er til et gitt ressursnivå – det som kalles produksjons- eller effektivitetsfronten. Produksjonsfronten er viktig om vi ønsker å forstå hva som er relativt gode produksjonsnivåer og relativt lave produksjonsnivåer. For å definere denne antar vi at produksjonen i FMA følger konstant skalautbytte. Vi antar altså at en endring i innsatsfaktorer i FMA vil føre til en proporsjonal endring i produksjonen. Dette er en rimelig antagelse da det ikke er store skalaforskjeller i perioden. Eksempelvis har antall årsverk økt med 11 prosent i perioden. I

¹ Se <https://www.fma.no/prinsix/prosjektfaser/forprosjektfase>.

² FMA er også en viktig aktør i annen produksjon. Innen investeringer støtter FMA eksempelvis produksjonen av konseptvalgutredningen (KVU). Dersom vi hadde inkludert KVU-produksjon i modellen måtte vi imidlertid hatt oversikt over Forsvarets produktivitet i perioden for å trekke slutninger om FMA sin produktivitet.

litteraturen anbefales det også å kun benyttes variabelt skalautbytte en tydelig kan vise til slike egenskaper (Dyson et al., 2001).³

Under forutsetningen om at produksjonen følger konstant skalautbytte kan vi estimere en bestepraksis produksjonsfront og observere hvilke perioder FMA presterer relativt godt og hvilke perioder FMA har prestert relativt dårlig. Som eksempel plotter vi en slik front i Figur 2.1. I denne figuren er forholdet mellom produksjonen av produktene $Y = \mu^T y$ og bruk av innsatsfaktorene $X = v^T x$ regnet ut for en bedrift over åtte perioder. Produktiviteten for hver periode er vist ved punkter i figuren. Verdiene av Y og X er i dette eksempelet uinteressante og aksene er følgende ikke nummererte.



Figur 2.1 Empirisk produktivetsfront for en tenkt bedrift med produksjon i åtte perioder.

Den best presterende perioden i utvalget er periode T som er et blått punkt i figuren. I denne perioden omgjorde bedriften innsatsfaktorbruk til produksjon på den mest effektive måten som til nå er observert. Perioden definerer også produksjonsfronten, som grunnet antagelsen om konstant skalautbytte, er en heltrukken linje gjennom origo og produktivitetsskåren til periode T. Alle de resterende periodene er ineffektive i forhold til periode T. Akkurat hvor ineffektive de er defineres ved den vertikale avstanden fra produksjonsfronten. Perioder med lang avstand fra produksjonsfronten er dårlige prestasjoner og kalles teknisk ineffektive. Perioder med kort avstand fra produksjonsfronten er gode prestasjoner, og perioder akkurat på produksjonsfronten kalles teknisk effektive (se boks 2.1). Antagelsen er altså at bedriften i enhver periode bør kunne prestere like godt som periode T. Ved å analysere prosessforskjeller og variasjon i faktorer

³ Et argument for å benytte modeller med variabelt skalautbytte, selv der vi ikke er sikre på om skalaegenskapene er variable, er at slike modeller kan la oss transformere innsatsfaktorbruk og produksjon uten at det påvirker resultatene. Dette er særlig gunstig for å måle produktivitet i virksomheter som kan ha negative eller nullverdier i produksjonen. Som vi skal se senere er produksjonssiden i FMA i visse tertiær og for visse produkter forbundet med nullproduksjon. De overordnede resultatene endres imidlertid ikke stort om vi benytter en VRS-modell isteden (se figur D.2).

mellom periodene opp mot produktivitetsskårene kan vi derfor identifisere måter å forbedre produktiviteten i virksomheten. Den samlede effekten av forbedrings og effektiviseringsinitiativ kan også observeres. På denne måten kan produktivetsanalyse og videreutvikling av modellene bidra til mer struktur i forbedring og effektiviseringsarbeid og forbedret styringsinformasjon.

Både vertikal og horisontal avstand fra produktivetsfronten kan benyttes som (in)effektivitetsmål; effektivisering kan skje enten ved å produsere mer for en gitt ressursinnsats eller like mye med færre ressurser (se boks 2.1). Som de stiplede linjene i Figur 2.1 viser, velger vi å benytte avstand mellom beste-praksis produksjon og faktisk produksjon til samme innsatsfaktorbruk som definisjon på god og dårlig. Dette fordi FMA «innenfor disponible ressursrammer [skal] realisere de målene og resultatkravene som Forsvarsdepartementet har fastsatt i tildelingsbrevet» (Forsvarsmateriell, 2023). Per i dag når ikke Forsvarsmateriell opp til egne produksjonsmål og resultatkrav. Forsvarsmateriell har altså i mindre grad mulighet til å redusere ressurser samtidig som produksjonen holdes konstant, men må isteden forsøke å øke produksjonen gitt de eksisterende ressursene. At oppmerksomheten rettes mot økninger i produksjon heller enn reduksjon i innsatsfaktorbruk kalles output-modellering eller output-orientering i faglitteraturen for dataomhyllingsanalyse.

Oppsummert benytter vi en output-orientert DEA-modell med konstant skalautbytte. Som vi kommer tilbake til i kapittel 4 benytter vi forskjellige varianter av denne grunnmodellen, men attributtene om output-orientering og konstant skalautbytte ligger fast. Ellers oppsummerer vi de mest sentrale begrepene vi benytter videre i rapporten i tabellen på neste side (tabell 2.1). Med modellen og begrepene i mente redegjør vi i neste kapittel for hvilke data vi benytter for å måle produksjon og innsatsfaktorbruk i Forsvarsmateriell.

Boks 2.1 – Effektivitetsbegrepet

Effektivitet er et ord som benyttes ofte og om mye forskjellig. Et par definisjoner av effektivitet er 1) gjøre de riktige tingene (Hanson, 2013), 2) gjøre tingene riktig (Kvalvik & Berg-Knutson, 2009), og 3) bruke minst mulig ressurser for å oppnå best mulig resultater.

Sammenhengen mellom definisjonene er at effektivitet beskriver avstanden mellom noe faktisk og noe optimalt. I den første definisjonen benyttes eksempelvis effektivitet for å beskrive avstanden mellom ønsket virkeeffekt og faktisk virkeeffekt. For å presisere hvilke avstander det er snakk om, knyttes derfor typisk et prefiks til effektivitetsbegrepet. Eksempler på dette er formålseffektivitet (ytre effektivitet), kostnadseffektivitet (indre effektivitet), produksjonseffektivitet og teknisk effektivitet. I tillegg finnes allokeringseffektivitet, prioriteringseffektivitet, planeffektivitet, styrings- og omstillingseffektivitet, m.m. Effektivitet i denne rapporten refererer til avstanden mellom best mulig produktivitet og faktisk produktivitet. Dette kalles teknisk effektivitet.

Tabell 2.1 Oversikt over sentrale begrep.

Begrep	Definisjon
Ressurser, innsatsfaktorer eller input	Det som anvendes for å skape produkter eller tjenester
Tjenester, produkter eller output	En gjenstand eller handling som tilbys for å fylle et behov hos bruker
Produktivitet	Summen av vektete tjenesteprodukter over summen av vektete ressurser
Produktivitets- og effektivitetsfront	En kurve som definerer best mulig (optimal) produksjon for et gitt nivå av innsatsfaktorer
Skalaegenskaper, skautbytte eller skalaavkastning	Beskrivelser av hvordan produksjonen endres som følge av endringer i (alle) innsatsfaktorene
Teknisk effektivitet	Avstand mellom faktisk produktivitet og best mulig produktivitet

3 Data

For analysen av produktivitet må vi benytte de dataene som FMA selv har samlet inn i perioden. Dette legger begrensninger for hvor godt vi kan utrede hva den historiske produktiviteten i FMA har vært og hvordan organisasjonen har utviklet seg. Det finnes ikke data på alt som vi mener det er hensiktsmessig å inkludere i produktivitetsmodellen. I dette kapittelet redegjør vi for hvilke data vi benytter og kvaliteten på disse dataene.

Vårt datasett består av 18 observasjoner av produksjon og innsatsfaktorbruk på etat-tertial-nivå. Det vil si at hver linje i datasettet angir produksjonen og innsatsfaktorbruk for Forsvarsmateriell i et gitt år og tertial. Utenom variabler som beskriver tid inneholder datasettet totalt 12 variabler, hvorav 9 er produksjonsvariabler (se vedlegg F) og 4 er innsatsfaktorvariabler. En oversikt over datasettets variabler og deskriptiv statistikk for variablene finnes i tabell 3.1. Data på variablene er hentet fra forskjellige kilder, og vi beskriver disse datakildene samt datavalg knyttet til variablene kort i avsnittene som følger.

Produksjon i investeringer. Forsvarsmateriell har i oppdrag å fremskaffe materiell og tjenester til avtalt tid, kostnad og ytelse (Forsvarsmateriell, 2023). Dette innebærer at FMA, ut ifra Forsvarets behov, planlegger for, lyser ut og kontraherer materiell fra leverandører. FMA faser også materiellet inn til drift i Forsvaret. Hver enkelt anskaffelse organiseres i prosjekter, og Forsvarsmateriell har rollen som prosjektansvarlig i gjennomføringen av prosjektet.

Vi fokuserer på fire produkter FMA produserer i investeringsprosjekter: beslutningsgrunnlag i forprosjektfasen, konkurransegrunnlag og kontrakter, materielloverføringsprotokoller, og termineringsrapporter som markerer prosjektets avslutning. Dette er leveranser som dekker store deler av aktivitetene i investeringsprosessen og som Forsvarsmateriell har hovedansvar for. De fire produktene oppfyller dermed de viktigste kriteriene for seleksjon av produkter i produktivitetsanalyser (Dyson et al., 2001, s. 248). For å forsikre oss om at dette er produksjon som er viktig for FMA gjennomførte vi også samtaler med Luftkapasiteter og Maritime kapasiteter. De fire produktene ble konsekvent trukket frem og fremhevet i disse samtalene.

Fra 2016 har FMA samlet tidsstempelt data på produksjonen av disse produktene for alle investeringsprosjekter i et datasett over milepæler. Milepælene representerer generelt viktige punkter eller hendelser i prosjektene, og er etablert for måling og oppfølging av prosjektenes fremdrift (Presterud et al., 2022). Milepælene vi benytter er først og fremst obligatoriske milepæler (OMP) som alle prosjekter skal benytte. Blant disse benytter vi fremsendt anbefalt sentralt styringsdokument (ASSD), alle signerte kontrakter, all materielloverføring utført i SAP og fremsendt termineringsrapport for å måle produksjonen av de fire produktene. Vi teller også milepæler som representerer samme produksjon, men er navngitt annerledes. Eksempler er der fremskaffelsesløsning (FL) benyttes istedenfor SSD eller der en milepæl tilhører en samlegruppe for øvrige milepæler, heller enn forprosjekt og gjennomføringsfase, men likevel passer beskrivelsen til milepælene nevnt ovenfor. For komplekse prosjekter teller vi de obligatoriske milepælene: oppdragsanalyse gjennomført, LCC-analyse gjennomført, alle konkurransegrunnlag utarbeidet, alle tilbud evaluert. I tillegg inkluderer vi prosjektmilepæler for kontraktinngåelser, materiellover-

føringer og sikkerhetsgodkjenninger. Milepæler som går på enkelte kontraktinngåelser, materiell-overføringer og sikkerhetsgodkjenninger er ikke-obligatoriske milepæler eller såkalte prosjekt-milepæler. Prosjekt-milepæler er milepæler som prosjektet selv anser som nødvendige å definere (Presterud et al., 2022) og de varierer følgelig mellom prosjekter. En naturlig antagelse er at de er registrert hyppigere for prosjekter med særegent behov for kontroll og fremdriftsrapportering. De er derfor mer relevante for komplekse prosjekter enn enkle prosjekter. Vi kommer tilbake til hvorfor vi teller delleveranser kun for komplekse prosjekter i avsnittet om kompleksitet. For å teste resultatenes sensitivitet til bruk av prosjekt-milepæler analyserte vi også produktivitet kun ved bruk av de obligatoriske milepælene. Vi finner at resultatene ikke blir nevneverdig annerledes.

Vi benytter milepældataene fra første tertial i 2018 til siste tertial i 2023 for prosjekter som har gjennomført milepæler i forprosjekt- og gjennomføringsfasen. Vi utelater prosjekter som er kansellerte fra analysen. Dataene dekker om lag 157 prosjekter.⁴ Medianprosjektet har om lag 8 registrerte milepæler, men kun 1 registrert milepæl for de 4 produktene vi er interessert i. Gjennomsnittsprosjektet har imidlertid om lag 12 registrerte milepæler og nærmere 2 milepæler som treffer de 4 produktene.

Presterud et al. (2022) har påpekt flere utfordringer knyttet til bruk av disse milepælene. Mest alvorlig for våre analyser er at prosjektene i liten grad har tatt i bruk de obligatoriske milepælene før Forsvarsmateriell får ansvar for anskaffelsen i det som kalles gjennomføringsfasen. I gjennomføringsfasen er milepælene tatt i bruk, men ikke alltid oppdatert. Som vi kommenterer på senere (i kapittel 5) kan dette også å være et problem for våre analyser.

Produksjon i materiellforvaltning. I tillegg til å anskaffe materiell har FMA også som hovedoppgave å forvalte Forsvarets materiell. Det vil si utvikle egne regelverk der sivile regelverk er utilstrekkelig, og gjennomføre tilsyn på materiellet. Vi har ikke identifisert data som gjør det mulig å tallfeste produksjon av regelverk og tilsyn på en hensiktsmessig måte og har derfor måttet utelate denne produksjonen fra analysen. Vi kommer tilbake til produksjon i materiellforvaltning i kapittel 7.

Kvalitet. FMA har ingen strukturerte kvalitetsindikatorer knyttet til enkeltleveranser. Det nærmeste vi kommer er en årlig brukerundersøkelse FMA sender ut til brukere i Forsvaret hvor de får tilbakemeldinger på opplevd kvalitet i leveranser og samarbeid med FMA. Spørreundersøkelsen er standardisert og benytter samme spørsmålsbatteri hvert år og en firepunktsskala for de fleste spørsmålene. Vårt inntrykk fra samtaler med kapasitetene og støtteavdelingene i FMA er at brukerundersøkelsene er uegnet for å fange opp kvalitet i produksjonen. I faglitteraturen påpekes det også at slike kvasi-kvantitative mål bør benyttes varsomt (Dyson et al., 2001). Vi inkluderer svar fra brukerundersøkelsene i én av modellene for å analysere hvorvidt dette påvirker resultatene for øvrig. Kun ett spørsmål fra brukerundersøkelsen inkluderes i denne modellen: «Materielleleveranser fra Forsvarsmateriell har ytelse iht. til forventninger». Skåren for en periode regner vi som den vektete summen av svarene fra Hæren, Sjøforsvaret, Luftforsvaret, Cyber-

⁴ I rådataene er det om lag 300 prosjekter, men et stort antall mangler registrerte milepæler for produksjonen vi er interessert i.

forsvaret etter at firepunktsskalaen er gjort om til tall fra 1 til 4. Dataene har dekning fra 2018–2022.

Kompleksitet. Produktene og tjenestene som Forsvarsmateriell leverer er av forskjellig kompleksitet. Det er mer krevende å utrede og beslutte fremskaffelsesstrategi for å anskaffe nye ubåter i samarbeid med Tyskland enn det er å gjøre det samme for nye uniformer. Det førstnevnte stiller høyere krav til koordinering på tvers av flere etater og landegrenser, og koordinering på tvers av mange fagområder. Det krever også mer planlegging for hvordan materiellet skal passe inn i en eksisterende struktur, og en langsiktighet i planleggingsprosessen ettersom utvikling av materiellet vil ta lang tid. Det samme kan gjelde konkurranseprosessen, leveranse av materiellet, og regelverksutvikling og tilsyn.

For å skille prosjekter og tilhørende produkter i kompleksitet deler vi prosjekter inn i to grupper basert på planlagt arbeidstid og arbeidsintensitet. Antagelsen er at komplekse prosjekter krever koordinering blant flere mennesker med forskjellig kompetanse, at de vil være prioritert ved at det er tildelt flere årsverk per år og at de tar lang tid å gjennomføre. I inndelingen benytter vi data fra Waage og Lien (2020) på bemanningsplaner. Waage og Lien (2020) kartlegger bemanningsbehovet beskrevet i SSD-ene per 2017 og fordeler behovet på kompetansekategorier. Vi definerer komplekse prosjekter som de prosjektene som særlig høy investeringskostnad og lang varighet og arbeidsintensitet i gjennomføringsfasen (se vedlegg C). Datasettet har kun dekning for prosjekter som hadde en SSD i 2019.⁵ For prosjekter som ikke omfattes av kartleggingen til Waage og Lien utnytter vi at forventet investeringskostnad (P50) og varighet er svært korrelert. Vi benytter forventet investeringskostnad per prosjekt fra investeringsplanen og definerer komplekse prosjekter som de med en P50 over 88-persentil. Færre prosjekter defineres som komplekse enn det som defineres som enkle etter denne inndelingen. Av de 158 prosjektene vi inkluderer i analysen er 28 prosjekter definert som komplekse og de resterende 130 prosjektene er definert som enkle.

Det finnes mange alternativer til kompleksitetsinndelingen vi har foretatt. Eksempelvis kunne vi gjennomført en spørreundersøkelse eller en kvalitativ inndeling i samarbeid med domeneeksperter. Vi har ingen formening om hvilken tilnærming som er bedre enn den andre, men valgte å benytte en kvantitativ inndeling for å redusere subjektivitet i analysen, øke transparens, og gjøre analysene våre reproduserbare. Alternative kvantitative inndelinger finnes også. Eksempelvis kunne vi benyttet faktisk bemanning i prosjekter istedenfor planlagt tall på bemanning. Vi benyttet det sistnevnte for å unngå støy i faktisk bemanning. Eksempelvis kan støy introduseres ved at prosjektlederturnover eller leverandørforsinkelser. Slike hendelser kan drive opp bemanningstid og varighet (og kostnad) opp uten at prosjektets kompleksitet egentlig har endret seg. Vi henviser den interesserte leser til vedlegg C for mer om hvordan vi har skilt mellom enkle og komplekse prosjekter.

Ved å differensiere prosjekter i kompleksitet deler vi beslutningsgrunnlag, konkurransegrunnlag og kontrakter, materielloverføringer og termineringer i to grupper – én gruppe for enkel produksjon og én gruppe for kompleks produksjon. For eksempel skiller vi produksjon av enkle

⁵ Vi bruker her fremskaffelsesløsning (FL) og SSD synonymt.

beslutningsgrunnlag fra produksjon av komplekse beslutningsgrunnlag. Dette gjør at disse fire produksjonsvariablene blir til åtte produksjonsvariabler.

For å tilskrive produktivitetsendringer til periodene der de oppstår må vi også sammenstille produksjon med frekvensen vi legger til grunn i analysen. Vi analyserer produktivitet per tertial og ønsker at produksjonen vi observerer per tertial skal kunne tilskrives ressursbruk i om lag samme tertial. For enkle prosjekter, som har en gjennomsnittlig og median varighet 5 år i gjennomføringsfasen, vil dette i stort være tilfelle. Om produksjonen fordeler seg uniformt i tid er det ikke urimelig å anta at om lag én leveranse for ett enkelt prosjekt forekomme hvert år. Sådan er det sannsynlig at noe av innsatsfaktorbruken i periode $t - 1$ og $t - 2$ påvirker produksjonen i periode t . Komplekse prosjekter, slik vi har definert dem, er få og har typisk en (planlagt) varighet på mellom 7 og 17 år. Antar vi på samme måte her at leveransene fordeler seg jevnt kan vi kun forvente én leveranse per 4 år. Dette skaper usikkerhet knyttet til hvor produktivitetsendringen oppstod. For å sammenstille produksjon med frekvensen for analysen deler vi derfor opp produksjonen i komplekse prosjekter. Mer spesifikt teller vi også delleveranser for komplekse prosjekter sammen med hovedleveransen som produksjon. For beslutningsunderlag innebærer det at vi for komplekse prosjekter teller både ferdigstilling av oppdragsanalysen og LCC-analyse i tillegg til fremsendt SSD som produksjon av beslutningsunderlag. For kontrakter og konkurransegrunnlag teller vi kontraktinngåelse og utarbeidelse av kontraktgrunnlag sammen med den obligatoriske milepælen, og for materielloverføringer teller vi også prosjektmilepæler for materielloverføringer sammen med den obligatoriske milepælen. En bekymring kan være at produktivitetsanalyser med tertial som frekvens blir meningsløse når produksjonssyklusene strekker seg lengre ut i tid. Enkelte produktivitetsendringer kan da reflektere at slutføringen av et produkt av tilfeldigheter faller ned på en periode heller enn en annen. Dette er en reell bekymring uten en åpenbar løsning. Dataene tillater ikke en nærmere oppstyking av produksjonen, og en aggregering fører til at mye informasjon kan forsvinne. For vårt formål utgjør de lange produksjonssyklusene derimot ikke et stort problem ettersom vi i hovedsak er ute etter å tyde en trend heller enn å fastslå produktivitetsskåren knyttet til et spesifikt tertial.

Innsatsfaktorer. Produksjonen i FMA krever, som stor del av øvrig tjenesteproduksjon, store mengder arbeidskraft. Om lag 80 prosent av alle driftskostnader i FMA er relatert til arbeidskraft i form av fastlønn, overtid, tillegg eller konsulenttimer. De resterende kostnadene er til kapital slik som husleie, vedlikehold, drift av IKT-systemer, leie av maskiner og avskrivninger.

Vi benytter to datakilder for innsatsfaktorbruk. Den første er det periodiserte virksomhetsregnskapet til Forsvarsmateriell. Her rapporteres periodiserte driftskostnader per tertial knyttet til arbeidskraft og realkapital. Vi har dekning i dataene tilbake til 2018. Fra regnskapene henter vi kostnader knyttet til arbeidskraft (fastlønn, overtid, tillegg, annen lønn og konsulenttjenester) og kapital (husleie og annen leie, vedlikehold og avskrivninger). En svakhet ved denne datakilden er at kostnadene ikke fordeles per aktivitet, og vi kan ikke isolere hvor stor andel av kostnadene som

tilhører investeringsaktiviteter utover grove estimater. Vi justerer kostnadene for generell prisvekst ved bruk av KPI.⁶

Den andre datakilden er registrerte arbeidstimer (arbeidstimeregistrering – ATR). Disse dataene inneholder registrerte timer på aktivitetskategorier. Ettersom vi på produksiden kun inkluderer investeringsproduksjon i den empiriske analysen benytter vi kun arbeidstimer ført på prosjekter. Denne datakilden gjør det mulig å isolere innsatsfaktorbruk til investeringsprosessen. ATR-dataene går tilbake til 2020, men er først av god kvalitet fra 2021. Vi benytter kun dataene fra 2021 og utover. For ATR-dataene har Forsvarsmateriell også beregnet en årlig timespris som inkluderer fordelingsnøkler for faste kostnader. Timesprisene kan benyttes for å knytte faste kostnader til aktivitet i prosjekter.

All analyse og databehandling er gjort i programmet R. For DEA-modellering benyttet vi pakken rDEA (Simm, 2014/2022) og Benchmarking (Bogetoft & Otto, 2011).

⁶ I denne analysen har prisvekstjusteringer liten konsekvens for resultatet ettersom variasjon i produksjonen i stort har bestemt utfallet. En vektning ved forvarsindeksen istedenfor KPI påvirker ikke resultatene nevneverdig.

Tabell 3.1 Aggregert oversikt over variabler og deskriptiv statistikk.

Variabel	Beskrivelse	N	Snitt	Standard- avvik	Minimum	Maksimum
<i>Produkter og tjenester¹</i>						
Beslutningsgrunnlag	Antall fremsendte FL/SSD samt oppdragsanalyse, LCC for komplekse prosjekter	18	3,833	2,256	0	8
Konkurransgrunnlag og kontrakter	Antall kontraktmilepæler gjennomført samt konkurransegrunnlag for komplekse prosjekter (se vedlegg E)	18	1,889	1,641	0	5
Materielloverføringer	Antall OMP materiellføringer samt enkeltoverføringer og sikkerhetsgodkjenninger for komplekse prosjekter	18	6,333	5,688	0	17
Terminering	Antall fremsendte termineringsrapporter	18	3,611	3,534	0	11
Brukertilfredshet	Brukers vurdering av ytelse på levert materiell (2018–2022)	15	3,05	0,194	2.75	3.25
<i>Innsatsfaktorer</i>						
Lønnskostnader (arbeidskraft)	I mill. 2023-NOK	18	1077,094	235,927	663,202	1498,564
Kostnader til kapital (kapital)	I mill. 2023-NOK	18	117,114	118,375	34,500	516,753
Arbeidstimer ²		9	151418,6	17118,01	127980,7	172519,2
Arbeidstimer*priser	I mill. 2023-NOK	9	136,178	16,401	113,629	162,343

* Aggregeringen tilsvare datasettet for modell 2 i kapittel 4.

¹ Produksjonsvariablene utenom brukertilfredshet lar seg også skille i kompleksitet. For hver av de fire andre produksjonsvariablene finnes det i datasettet en enkel og en kompleks variabel. Eksempelvis kan vi skille mellom enkle beslutningsgrunnlag og komplekse beslutningsgrunnlag. Til sammen utgjør disse $4 \times 2 = 8$ produksjonsvariabler. Vi oppgir kun nøkkeltallene for de to gruppene samlet i tabellen.

² Arbeidstid, overtid og fleksitid.

4 Analyse og resultater

I dette kapittelet beskriver vi hvordan produksjonen og innsatsfaktorbruken i FMA har utviklet seg. Deretter ser vi på hvordan disse har utviklet seg sammenlignet med hverandre ved å estimere produktivitetsskår for hvert tertial fra tertial 1 i 2018 til tertial 3 i 2023. For å øke troverdigheten til analysen benytter vi varianter av den enkle produktivitetsmodellen til å trekke slutninger om når Forsvarsmateriell har vært mest produktive og hvordan produktivitetsutviklingen har vært.

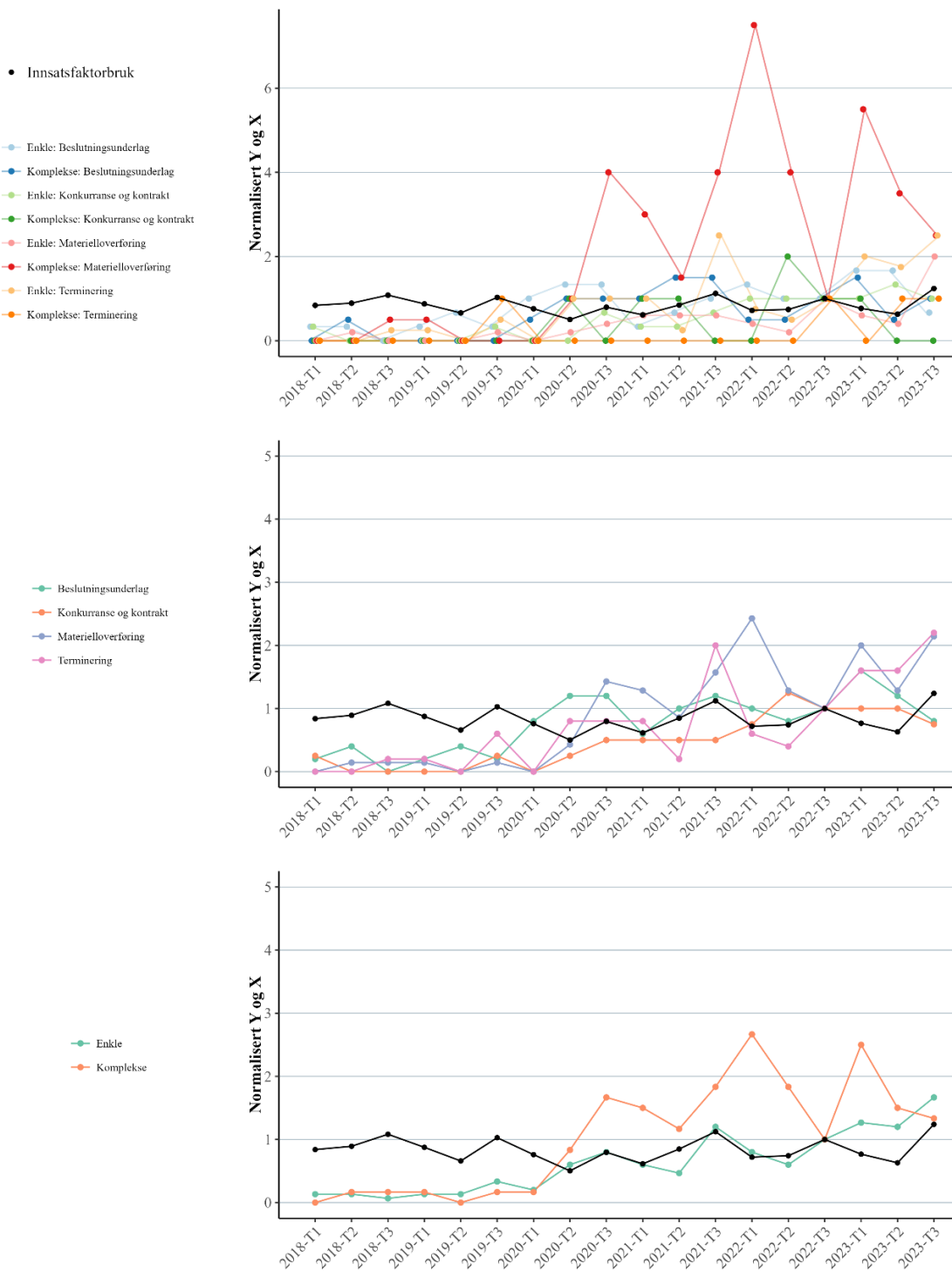
4.1 Grafisk fremstilling av utviklingen i produksjon og innsatsfaktorbruk

I figur 4.1 plotter vi tre spredningsdiagram som viser produksjonen av beslutningsgrunnlag, konkurransegrunnlag, materielloverføringer og termineringsrapporter per tertial fra 2018 til 2023 aggregert på forskjellige måter. I figurene er også total innsatsfaktorbruk illustrert ved en svart linje. Alle tallene er normalisert til 2022-T3.

I først diagram er produksjonen disaggregert per produkt og delt etter kompleksitet. Diagrammet viser at produksjonen varierer substansielt mellom perioder. Spesielt er dette tilfellet for komplekse materielloverføringer. Det kan også antydes at produksjonen har vært økende. Flere punkter ligger lavere tidligere i diagrammet enn antallet punkter som ligger lavere senere i diagrammet. Samtidig har ikke produksjonen vært monotont økende og flere produkter har sine produksjonstopper før siste periode.

Aggregerer vi produksjonen til fire produkter (diagram 2 i figuren) eller til å kun skilles i kompleksitet (diagram 3) blir trendene tydeligere. Til tross for at det er stor variasjon mellom tertialene virker produksjonen overordnet å ha økt. Det har vært en økning i perioden for samtlige produkter og økningen gjelder både for enkle og komplekse prosjekter.

Det er kun for diagram tre at vi kan observere klare forbedringer mellom perioder. Eksempelvis øker både enkel og kompleks produksjon fra tertial 1 til tertial 2 i 2021. For de andre diagrammene er det vanskeligere å tyde klare forbedringer. Hadde vi aggregert dataene til årssnivå hadde dette vært enklere. For oppstykkningen i åtte produkter hadde FMA i 2020 eksempelvis en henholdsvis enkel og kompleks (antall i parentes) produksjon på 11 (5) beslutningsgrunnlag, 2 (1) konkurransegrunnlag og kontrakter, 3 (10) materielloverføringer og 8 (0) prosjekttermineringer. Ett år senere var produksjonen endret seg med: -5 (+3) beslutningsgrunnlag, +1 (+1) konkurransegrunnlag og kontrakter, +6 (+7) materielloverføringer og +7 (0) termineringer. I 2021 var altså produksjonen lik eller høyere enn i 2020 for samtlige produkter foruten enkle beslutningsunderlag. Med mindre produksjonen av enkle beslutningsgrunnlag verdsettes svært høyt vil derfor 2021 være en forbedring over 2020. Gjør vi imidlertid det samme for årene 2021 og 2022 er endringene i produktene positive og negative om hverandre, og det er vanskelig å si om 2022 er en forbedring fra 2021 eller ikke.



Figur 4.1 Normalisert produksjon og innsatsfaktorbruk (2022-T3=1). Første diagram viser produktene delt inn etter kompleksitet. Andre diagram viser produksjon av milepæler. Det tredje diagrammet viser produksjonen fordelt kun på kompleksitet, altså summen av produkter fordelt på de to kompleksitetsgruppene.

Utviklingen i lønnskostnader og kostnader knyttet til kapital er vist samlet i figur 4.1. Av figuren kan vi se at kostnadene fluktuerer fra tertial til tertial, men langt mindre enn produksjonen. Kostnadene svinger også uten en klar trend.

Lønnskostnader utgjør mesteparten av kostnadene. Disse har ligget svært stabilt i perioden. I samme perioden vet vi at antall årsverk i Forsvarsmateriell har vært stigende.⁷ Ettersom lønnskostnadene ikke har hatt en lik trend kan det tyde på at årsverksjusteringer ikke har hatt stor innvirkning på kostnader tilknyttet arbeidskraft. Kostnader til kapital har utviklet seg annerledes. Disse kostnadene har ligget lavt frem til 2022 og stiger kraftig deretter. Kostnader tilknyttet vedlikehold har vært dominerende for den sterke økningen i kostnadsnivået i 2022 og 2023.

Til tross for at kostnadene ligger nokså stabilt over hele perioden observerer vi en form for sesongvariasjon. Spesifikt ser kostnadene ut til å følge et årlig mønster – tertial 3 er alltid høyere enn de foregående tertialene i samme år. Tertial 3 har gjennomgående høyere kostnader til fastlønn og konsulenter, og som regel er overtidskostnader og kostnader knyttet til tillegg også høyere i dette tertialet. At kostnader knyttet til arbeidskraft svinger såpass mye tyder på at investeringer blir særlig prioritert i tertial 3.

4.2 Produktivitet

For å vurdere samlet produktivitet i Forsvarsmateriell vektet vi produktene og tjenestene som produseres. Vi benytter den enkle DEA-modellen vi presenterte i kapittel 2 for å definere teknisk effektivitet og finne optimale vekter for produksjonen. Innsatsfaktorene vektet vi likt, ettersom disse er kostnader i kroner. For å øke troverdigheten til resultatene analyserer vi også modellspesifikasjoner med forskjellige kombinasjoner av input og output og forskjellige aggregeringsnivå (se tabell 4.1). Vi forklarer hva som inngår på produksiden og innsatsfaktorsiden under, men henviser også til vedlegg F for en visualisering av hvordan datasettene på produksiden ser ut for modell 1, 2 og 3. Produksiden i Modell 1, 2 og 3 korresponderer også med produksiden vist i henholdsvis diagram 3 2 og 1 fra figur 4.1.

⁷ Årsverk i Forsvarsmateriell har økt fra 1374 i 2018 til 1524 i 2023.

Tabell 4.1 Oversikt over modeller med forskjellige input- og outputvariabler. Kolonnene beskriver en modell fra 1 til 6. Produktene og innsatsfaktorene som inngår i en modell er gitt ved et pluss tegn i de korresponderende cellene. Modell 3, som er baseline-modellen, benytter eksempelvis kombinasjonen av kompleksitet og milepæler som produkter og lønn og realkapitalkostnader som innsatsfaktorer.

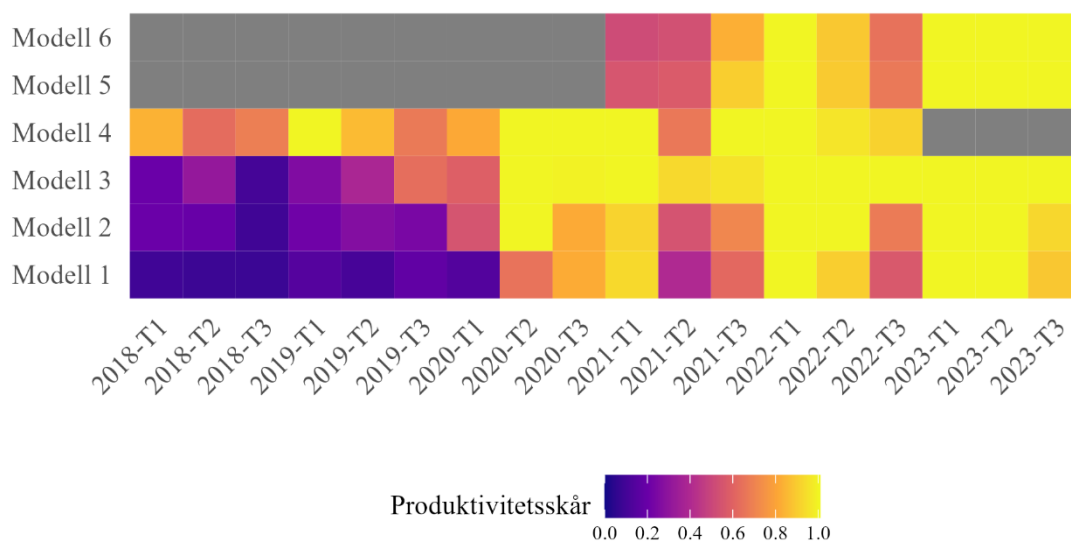
	Modeller					
	1	2	3	4	5	6
<i>Produkter og tjenester</i>						
Kompleksitet		+		+	+	+
Milepæler	+					
Kompleksitet*Milepæler			+			
Brukertilfredshet				+		
<i>Innsatsfaktorer</i>						
Lønnskostnader	+	+	+	+		
Kostnader til realkapital	+	+	+	+		
Arbeidstimer					+	
Arbeidstimer*Timepris						+
Tidsrom	18–23	18–23	18–23	18–22	21–23	21–23
Observasjoner	18	18	18	15	9	9

Modellresultater og trend. Vi plottet resultatene fra modellene i figur 4.2. En rad viser produktivitetsskåren for en gitt modell på tvers av alle periodene. Produktivitetsskårene er bundet mellom 0 og 1, hvor 1 definerer beste praksis. Vi forenkler visualisering ved å fargelegge skårene basert på et kontinuerlig fargespekter fra mørkeblå til lysegul der mørkeblå er 0 og lysegul er 1. Dataene som ligger til grunn for figuren finnes i vedlegg D.

I modell 1 benytter vi lønn og kostnader tilknyttet realkapital som innsatsfaktorer (input) og aggregert enkel og kompleks produksjon som produkter (output). Med aggregert produksjon mener vi at vi verdsetter de fire produktene – beslutningsgrunnlag, konkurransegrunnlag og kontrakter, materielloverføring og terminering – likt, men skiller enkel og kompleks produksjon. Med andre ord er det kun to produkter i denne modellen: summen av enkel produksjon og summen av kompleks produksjon. Resultatene fra denne tilsier at FMA er blitt mer produktive. Modellresultatene beveger seg fra blå i tidligere perioder og mot gul i senere perioder. Vendepunktet virker å være i 2020, da produktivitetsskårene for første gang befinner seg om lag midt på skalaen. Etter 2020 flater den positive utviklingen noe ut – det er med andre ord mer utydelig om FMA har forbedret produktiviteten sin i perioden 2021–2023.

Ettersom det i modellen skiller mellom kompleks og enkel produksjon tyder resultatene på at Forsvarsmateriell har blitt bedre på både enkel og kompleks produksjon. Husk at en periode kun er dårligere enn en annen om den (strengt) domineres. At Forsvarsmateriell har en blå skår i tidligere perioder og en gul skår i senere perioder betyr altså at forholdet mellom enkel produksjon

og innsatsfaktorbruk samt forholdet mellom kompleks produksjon og innsatsfaktorbruk er høyere i senere perioder enn i tidligere perioder. At produktivitet har vært økende i både enkle og komplekse prosjekter kan tyde på at store prosjekter ikke fortrenger små prosjekter i like stor grad som tidligere antatt (se Presterud & Øhrn, 2015).



Figur 4.2 Produktivitetsestimater. Kolonnene beskriver perioder og radene viser modeller. En celle viser produktivitetsestimatet for et gitt tertial og en gitt modell. Skårene er fargelagt fra mørkeblå til lysegul, hvor lysegul definerer best mulig produktivitet. Der data mangler er cellene farget grå.

I modell 2 og 3 benytter vi to alternative definisjoner av produksjon. Vi benytter de samme innsatsfaktorene som i den første modellen, men varierer produksjonssiden. I modell 2 benytter vi produktene uten en kompleksitetsgruppering. Denne modellen forteller oss hvilke perioder som er mest produktive om vi antar at produksjon av beslutningsgrunnlag, konkurransegrunnlag og kontrakter, materielloverføringer og termineringer kan verdsettes forskjellig, men at det ikke finnes noe kompleksitetsskille (for mer om dette se vedlegg A). Modell 2 har flere produkter, og det er derfor større sannsynlighet for at en periode blir vurdert å være helt effektiv (produktivitetsskår lik 1). I modellen er det fem perioder som er helt effektive, i kontrast til tre fra modell 1. Som for den foregående modellen faller produktiviteten markant i starten av 2021 og 2022, men FMA er stort sett effektive i det siste året.

I modell 3 setter vi modell 1 og 2 sammen, slik at kombinasjoner av produkter og kompleksitet kan verdsettes forskjellig. I modellen er det altså åtte produkter på output-siden (se vedlegg F).⁸ Ettersom det er relativt mange produkter i denne modellen er det enda mer sannsynlig enn de foregående modellene at perioder vurderes å være effektive. Som ved modell 1 og 2 virker periodene før 2020 å være lite effektive. I modell 3 er derimot nærmest alle periodene etter 2020 effektive.

⁸ Enkle beslutningsgrunnlag, enkle konkurransegrunnlag og kontrakter, enkle materielloverføringer, enkle termineringer, komplekse beslutningsgrunnlag, komplekse konkurransegrunnlag og kontrakter, komplekse materielloverføringer og komplekse termineringer.

ive, med noen få og små unntak i tertial 1 og 2 i 2021. Dette innebærer at nesten alle periodene i dette intervallet skårer best på minst ett produkt relativt til innsatsfaktorbruk.

Modell 1 til 3 taler alle for at Forsvarsmateriell har blitt mer produktive over tid og at perioden før og etter 2020 markerer et skille. Modell 1 indikerer også at det har vært en produktivitetsutvikling etter 2020, men dette er vanskeligere å se ut ifra modellene med flere produkter. Dette kommer blant annet av at forholdet mellom antall produkter og innsatsfaktorer og antall observasjoner blir for skjevt.⁹ I de resterende modellene bygger vi derfor videre på modell 1 ettersom modellen virker rimelig samtidig som den benytter et fåtall produkter og innsatsfaktorer.

I modell 4 bygger vi på modell 1 og undersøker hvordan produktivitetsestimatene påvirkes om vi inkluderer brukertilfredshet. Brukertilfredshet er den uniformt vektete summen av 4-punkts Likert-skalaesponser fra grenene og FST på den årlige spørreundersøkelsen sendt ut av FMA. Resultatene for modell 4 skiller seg noe fra de tidligere modellene. Perioden før 2020 er fremdeles lite effektiv, men langt mindre enn i de andre modellene. Brukertilfredsheten har vært høy, men fallende i perioden 2018–2021. I denne modellen (4) kan perioder som i modell 1–3 var ineffektive, kompensere for lav produksjon av andre produkter med høy brukertilfredshet. Modellresultatene understreker hvor viktig det er med gode kvalitetsindikatorer. Samtidig er ikke brukertilfredshet, slik den er definert i denne rapporten, en særlig troverdig indikator på kvalitet i produksjonen, og det er behov for bedre måleindikatorer for kvalitet i produksjonen.

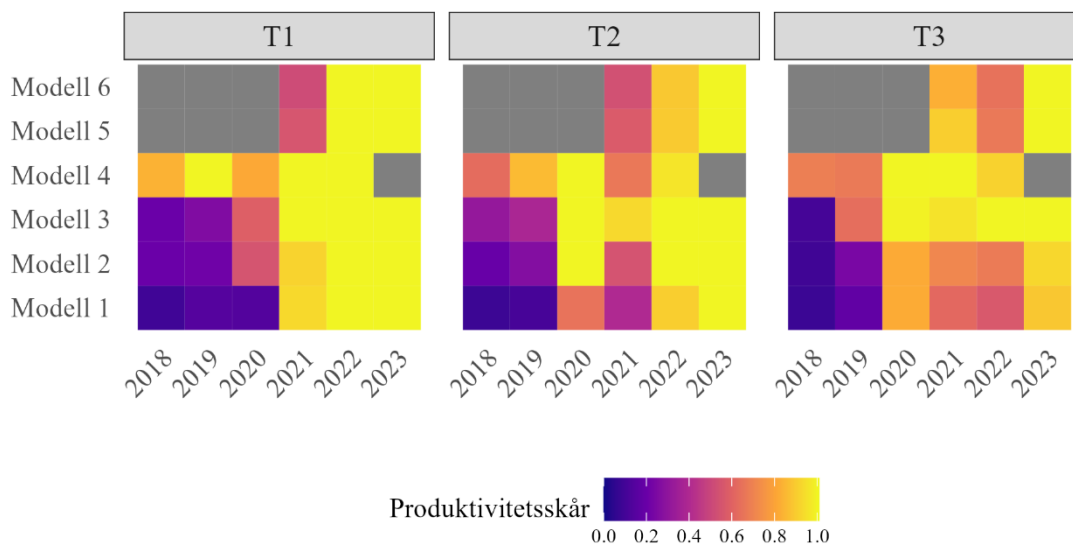
De neste modellene (5–6) benytter kompleksitet på produksiden og isolerer innsatsfaktorbruk i prosjekter ved å benytte arbeidstimer (5) og arbeidstimer vektet ved ATR-timeprisen som innsatsfaktorer (6). Disse innsatsfaktorene er knyttet til aktiviteten kun i prosjekter til forskjell fra innsatsfaktorene i de tidligere modellene som reflekterte ressursbruk i hele FMA. ATR-timeprisen inneholder også fordelingsnøkler for faste kostnader slik at realkapitalkostnader i noen grad ivaretas. Merk at arbeidstimerregistreringsdata kun finnes fra 2021, og vi kan derfor ikke estimere produktivitet for tidligere perioder i disse modellene. Resultatene fra modell 5 og 6 er svært like som de foregående modellene, med unntak av 2021-T1 og 2022-T3 som blir relativt ineffektive. Dette tyder på at arbeidskraftproduktiviteten for disse periodene er lavere enn den generelle produktiviteten. Ellers tyder modellene på at total innsatsfaktorbruk i FMA er proporsjonal til innsatsfaktorbruk i prosjekter.

Overordnet tyder estimatene på at FMA har blitt mer produktiv over tid. Resultatene passer godt med vår forståelse av FMA sin virkelige utvikling. Siden 2016 har FMA investert kraftig i verktøy, spesialisering og kompetanseheving i tillegg til forbedring og effektiviseringsarbeid for å bygge opp en effektiv investeringsorganisasjon. Samtidig har FD lagt økende vekt på gjennomføringsevne og transparens i roller, ansvar og myndighet i investeringsprosessen.

Tertial over tertial. På grunn av sesongvariasjon kan det være hensiktsmessig å se like perioder opp mot hverandre. Eksempelvis vil alltid sommerferien falle i tertial 2. Lange pauser mellom arbeidsoppgaver, slik som sommerferie, kan føre til dupp i produktivitet da det er forbundet

⁹ Generelt anbefales det å ha et antall observasjoner som er minst to ganger produktet av produkter og innsatsfaktorer (Charnes et al., 1978, 1994; Dyson et al., 2001).

kostnader ved å starte dem opp igjen. Det er derfor grunn til å tro at produktiviteten i tertial 2 vil være gjennomgående lavere enn de andre tertialene. Vi grupperer resultatene etter tertial i figur 4.3. Av figuren kan vi ikke se at produktivitetsskårene er gjennomgående dårligere i tertial 2. Dette kan tyde på at oppstartskostnadene ikke er betydelige nok til å gi utslag i produktivitet, eller at det finnes lignende fenomen i tertial 1 og 3 som gjør at tertial 2 ikke skiller seg ut til tross for betydelige oppstartskostnader.



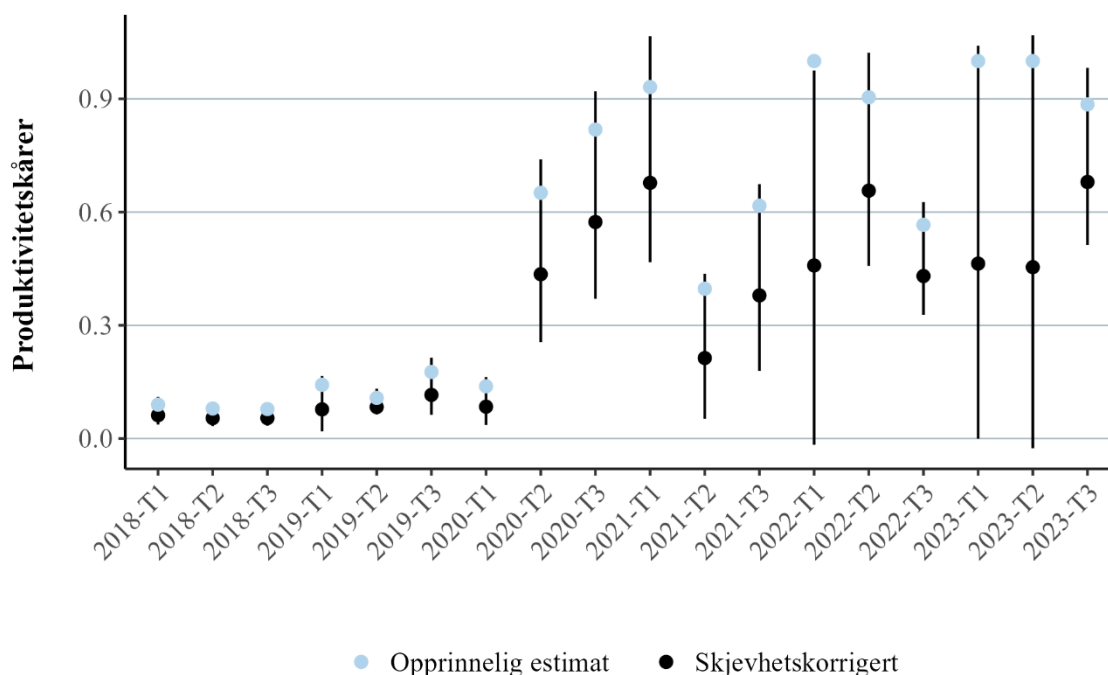
Figur 4.3 Produktivitetsestimat per tertial. Figuren er delt inn i tre paneler hvor hvert panel markerer ett tertial. Kolonnene innad i et panel beskriver perioder for et gitt tertial og radene viser modeller. En celle viser produktivitetsestimatet for et gitt tertial i et gitt år og for en gitt modell. Skårene er fargelagt fra mørkeblå til lysegul hvor lysegul definerer best mulig produktivitet. Grå celler markerer perioder der det mangler data for å estimere produktivitet gitt en modell.

Produktiviteten i tertial 3 virker å være gjennomgående lavere enn i tertial 1 og 2. Som nevnt tidligere i kapittelet er kostnader knyttet til arbeidskraft også gjennomgående høyere i tertial 3 enn i de andre tertialene. Det produseres som regel også mer i tertial 3, men ettersom produktiviteten er noe lavere i tertial 3 enn i de andre tertialene indikerer dette at det innsatsfaktorbruket vokser raskere enn produksjonen i disse periodene. En forklaring på hvorfor innsatsfaktorbruk er høy i siste tertial er at FMA i denne perioden jobber hardt for å komme så nærme årlige målsettinger som mulig. Resultatene kan altså tyde på at FMA har en skippertakskultur, og at årsslutt oppleves som en tidsfrist. At produktiviteten i perioden ikke er den høyeste kan tyde på at den n -te arbeidstimen på kort sikt ikke er like produktiv som den første arbeidstimen.

4.3 Kalibrering og usikkerhet

Til nå har vi estimert en beste-praksis produksjonsfront og funnet at FMA har økt sin produktivitet siden etaten ble opprettet. Vi har også adressert muligheten for målefeil i dataene. Vi ønsker videre å adressere innbyrdes usikkerhet ved å se nærmere på variasjonen i produktivitetsskårene mellom periodene.

For å estimere usikkerheten i skårene følger vi Simar og Wilson (1998) sin foreslåtte metode for sensitivitetsanalyse av DEA-estimer. De foreslår at perioder, ettersom de antas å være homogene og ha homogen teknologi, bør oppnå samme produktivitetsskår om vi fjerner all usikkerhet. Avstand fra fronten kan således ses på som residualer, og gjenvælg over residualene kan benyttes til å reskalere produksjonen og til å anslå usikkerheten i de originale estimatene. Simar og Wilson (1998) argumenterer også for at DEA-estimer er utsatt for systematiske skjevheter ettersom vi kun observerer et utvalg av den mulig realiserede produksjonen. Simar og Wilson (1998) foreslår å benytte gjenvælg over residualene til å også korrigere estimatene for disse skjevhetene. Simar og Wilson kaller dette for en sensitivitetsanalyse. Vi henviser den interesserte leseren til vedlegg B for mer om hvordan vi estimerer usikkerheten og korrigerer for skjevheter.



Figur 4.4 Resultater fra sensitivitetsanalyse for DEA foreslått av Simar og Wilson (1998). De vertikale heltrukne linjene viser et 95 prosent konfidensintervall sentrert ved de skjevhetsskorrigerte (svarte) punktene. De originale estimatene er gitt ved lyseblå punkter.

I Figur 4.4 viser vi resultatene fra en sensitivitetsanalyse med utgangspunkt i modell 1. Usikkerheten i estimatene er vist ved 95 prosent konfidensintervaller som i figuren er illustrert ved svarte

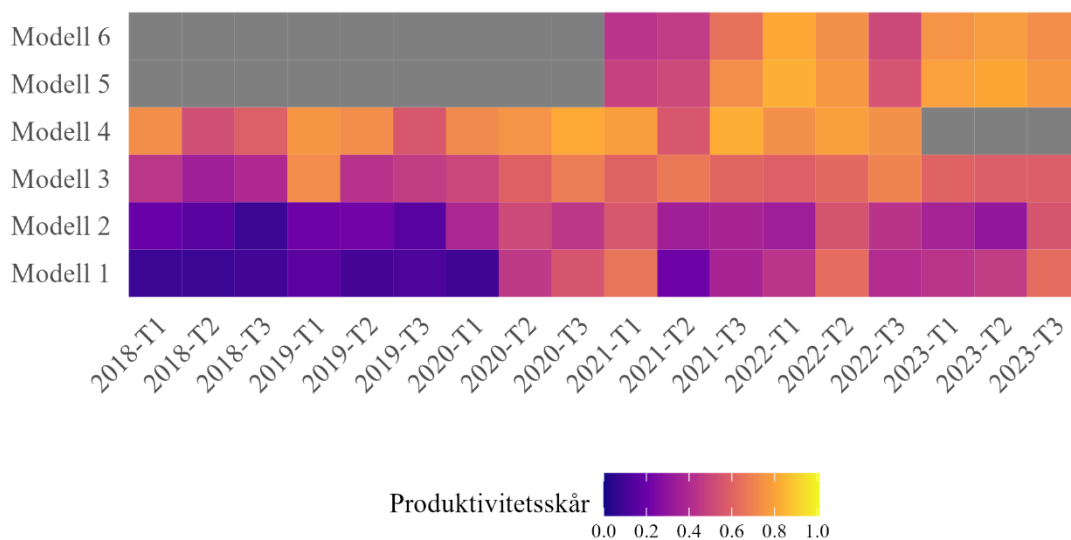
vertikale linjer for hver periode. De originale estimatene er gitt ved lyseblå punkter og de skjevhetsskorrigerte estimatene er gitt ved svarte punkter. Det er lett å se at flere av de opprinnelige estimatene, særlig de fra senere perioder, er forbundet med høy usikkerhet. Dette viser at effektiviteten til periodene i stor grad er drevet av en spesifikk faktorproduktivitet heller enn at produksjonen har vært effektiv for alle produktene og tjenestene i perioden (se vedlegg B). Konfidensintervallene for estimatene i disse periodene er også i stor grad overlappende, hvilket betyr at forskjellene mellom periodene ikke er statistisk signifikante. Periodene før tertial 2 i 2020 er forbundet med svært lite usikkerhet. For disse periodene er dårlig produksjon forbundet med dårlig faktorproduktivitet langs all produksjon. Konfidensintervallene for disse estimatene er i liten grad overlappende med perioden etter tertial 2 i 2020 som indikerer at forskjellene mellom periodene er statistisk signifikante. Analysen av usikkerheten i estimatene viser at vi skal være varsomme med å trekke sterke konklusjoner om den relative effektiviteten mellom enkelte perioder. Analysen gjør oss imidlertid mer sikre på at perioden før og etter 2020 er forskjellige.

Av de skjevhetsskorrigerte estimatene kan vi se at samtlige estimater var overestimerte. De skjevhetsskorrigerte estimatene faller alltid under de originale estimatene. Rangeringen blant de mest effektive enhetene endrer seg også for de skjevhetsskorrigerte estimatene. Våre originale estimater viser at tre perioder var helt effektive med $\theta = 1$. Dette var 2022-T1, 2023-T1 og 2023-T2. Alle disse periodene faller langt ned etter skjevhetsskorrigeringen. Etter skjevhetsskorrigeringen er det også flere andre tertial som oppnår om lag like høy eller høyere skår enn disse periodene. Periodene som i de originale estimatene var helt effektive er også forbundet med særdeles høy usikkerhet. Usikkerheten for disse estimatene er så høy at produktivitetsskåren ikke kan sies å være statistisk signifikant forskjellig fra 0.

Det er også mulig å benytte den estimerte variasjonen til å trekke inferens om faktorer som påvirker produktivitet. Simar og Wilson foreslår i en artikkel fra 2007 hvordan metoden fra 1998 kan bygges på for å gjøre nettopp dette (Simar & Wilson, 2007). Metoden innebærer å benytte trunkert regresjon for å anslå forholdet mellom forklaringsfaktorer og produktivitetsskårene og bootstrapping for å lage konfidensintervall for å anslå signifikans i estimatene. Eksempelvis kunne vi samlet data på de faktorene som nevnes i kapittel 5 og studert samvariasjonen mellom dem og produktivitetsskårene. I denne rapporten har vi ikke hatt som mål å forklare produktivitetsskårene. Vi forklarer likevel hvordan dette kan gjøres i vedlegg B.

Vi estimerer usikkerhet og skjevhetsskorrigerte estimatene for hver modell og viser resultatene på likt vis som tidligere i figur 4.4. Vi synliggjør ikke hvor mye usikkerhet hvert estimat er forbundet med i denne figuren, men forskjellen mellom de originale og de skjevhetsskorrigerte skårene gir en indikasjon på dette. Slik som for modell 1 er estimatene for de andre modellene korrigert nedover hvilket innebærer at de er forbundet med betydelige usikkerhet. At selv de beste presterende periodene etter skjevhetsskorrigeringen befinner seg et stykke unna produktivitetsskåren tyder på at Forsvarsmateriell fremdeles har stort rom for å øke sin produktivitet. Trendene vi observerte tidligere før skjevhetsskorrigeringen endrer seg noe. De store forskjellene mellom perioden før 2020 og etter 2020 forsvinner ikke. Tendenser i perioden 2020 – 2023 er blitt mindre tydeligere. Skårene i denne perioden er mindre spredt enn før og det er ikke lenger perioder i 2022

og 2023 som er de best presterende periodene. Ut ifra de korrigerede estimatene kan vi ikke observere at det har skjedd substansielle produktivetsendringer etter 2020.



Figur 4.5 Skjevhetekorrigerede produktivetsestimater. Kolonnene i figuren beskriver perioder og radene viser modeller. En celle viser produktivetsestimater for et gitt tertial og en gitt modell. Skårene er fargelagt fra mørkeblå til lysegul, hvor lysegul definerer best mulig produktivitet. Der data mangler er cellene farget grå.

5 Hva skiller gode fra dårlige perioder?

Produktivitetmålinger er mest nyttige om vi kan forklare hvorfor produktivitet er høy i noen perioder og lav i andre. Det gir meningsfull innsikt for forbedringsarbeid ved å sette søkelys på de endringene som gir effekt, og for eksempel ikke endre prosesser der det har vært redusert produktivitet på grunnlag av forhold som ikke kan kontrolleres. I dette kapitlet diskuterer vi mulige forklaringer til endringer i produktivitetsestimatene. Å bidra med fullverdige forklaringer til produktivitetendringene er utenfor omfanget av denne rapporten. Vi nevner noen mulige forklaringer som motiveres fra tidligere forskning og analyser av investeringer og FMA sine tertialrapporter.

Dataregistrering og -kvalitet. Den første mulige forklaringen på produktivitetstrenden er knyttet til datakvalitet. Dataene viser at det er svært få registrerte milepæler før 2020 for all produksjon (se figur 4.1). Dette gjelder særlig for konkurransegrunnlag og kontrakter og materielloverføring. Dette kan enten skyldes at FMA produserte svært lite i disse årene, eller at den faktiske produksjonen ikke ble registrert i disse årene. Det er derfor mulig at det vi kaller produktivitetstrender egentlig kan forklares av endringer i registreringshyppighet. Dette vil være en form for ikke-tilfeldig målefeil hvor faktisk produksjon avviker signifikant fra det vi observerer og har målt. Dette reduserer troverdigheten til resultatene. Som nevnt i kapittel 3 påpekte Presterud et al. (2022) at milepældataene manglet registrering og oppdatering, men det ble ikke påpekt hvorvidt dette har variert over tid. Samtaler med FMA tyder på at registrering kan ha vært lav før 2020 og har blitt bedre i årene etter. For å teste dette videre ser vi til FMA sine tertialrapporter, hvor blant annet antall produserte beslutningsgrunnlag rapporteres. Vi finner at antall produserte SSD-er som rapporteres produsert i FMA sine tertialvise rapporter til FD skiller seg fra hvor mange som er registrert i milepældataene. Eksempelvis rapporterte FMA å ha fått godkjent ni SSD-er i første tertial av 2018. I milepældataene er det derimot kun registrert én SSD i dette tertialet. Riktignok benytter vi fremsendt SSD til forskjell fra godkjent SSD, men dette virker ikke å forklare forskjellene. Avviket mellom tertialrapportene og milepældataene er langt mindre for årene etter 2020 som også tyder på at registreringshyppigheten har endret seg. For å teste hvor sensitive resultatene er for endringer i registreringshyppighet fjerner vi årene før 2020 og kjører modellene på nytt. Vi observerer at produktivitetstrenden fortsatt er økende, men stigningstallet er lavere enn før. Modell 5 og 6 benyttes allerede kun data fra årene etter 2020. Fra kun disse modellene, som består av seks produktivitetsestimater, er det imidlertid vanskelig å tyde en trend.

Tilsynelatende lav kvalitet på dataene og få datapunkter gjør at vi bør være varsomme med å trekke sterke konklusjoner om utviklingen i produktivitet. Dette innebærer at funn må tolkes med omhu og at rapportens funksjon blir mer illustrerende for hvordan en produktivitetsanalyse kan gjennomføres, enn en autoritet på utvikling i produktivitet i FMA. Andre analyser må supplere funnene i denne rapporten før sterke konklusjoner kan trekkes.

Prosjektlederrotasjon. Utskifting av prosjektledere kan føre til mindre kontinuitet i prosjekter og at dette kan negativt påvirke kvalitet, kostnader og tid brukt i prosjekter. Utskifting av prosjektledere var blant annet den tredje mest frekvente årsaken til forsinkelser (etter leverandørforsink-

elser og samtidighetsutfordringer) i en gjennomgang av forsinkelsesårsaker i investeringsporteføljen av 2022.¹⁰ Dersom prosjektledere skiftes ut i mindre grad enn før kan dette være med på å forklare produktivitetsveksten.

De seneste studiene vi finner på området er Waage (2021) som analyserer avgang og slutttrater i FMA i perioden 2013–2019. Waage (2021) finner at separasjonsrater var spesielt høye i 2018 og 2019. I studien vises det at militært tilsatte utgjorde en stor del av prosjektledere. I studien forventes det at slutttrater blant militært tilsatte vil være høye i perioden fremover mot 2025. Dette kan tyde på at prosjektlederturnover i perioden har vært høy.

Hyllewareandel. I anskaffelser oppstår tidvis valget mellom å utvikle eller tilpasse materiell fremfor å anskaffe eksisterende materiell, såkalt *hylleware*. Anskaffelse av hylleware antas å være raskere og mindre kostnadsdrivende – det medgår mindre tid til kravspesifikasjon, det kan føre til flernasjonalt konsolidering av behov og dermed kvantumsrabatt, ingen utviklingstid og lavere produksjonstid. Det finnes noe empirisk støtte for denne hypotesen. Berg et al. (2019) finner ved å studere investeringsplanen 2015–2022 at differansen mellom planlagt og faktisk gjennomføringstid i snitt er 1,35 år lavere for utviklingsprosjekter enn hyllewareprosjekter. Med andre ord forventes det at utviklingsprosjekter er om lag 1,35 år mer forsinket enn et hyllewareprosjekt, når investeringssum og organisatorisk tilhørighet kontrolleres for. Tilpassingsprosjekter er om lag 0,56 år mer forsinket enn hylleware, men dette estimatet er ikke statistisk signifikant. Forsvarssektoren har tidligere blitt kritisert for at de i for stor grad velger å utvikle eller tilpasse materiell heller enn å kjøpe ferdigutviklet. I nyere tid har det derfor vært et mål om å øke andelen hylleware i forsvarssektorens investeringer. Om hyllewareandelen har økt ytterligere kan dette være med på å forklare produktivitetsendringene.

Presterud og Øhrn (2015) utviklet en metode for å kategorisere prosjekter som hylleware, og anvendte metoden på den daværende investeringsplanen. Berg og Waage (2019) fulgte samme metode og gjennomførte en ny kartlegging for investeringsplanen i 2019. Voldhaug, Prebensen og Presterdud (2024) oppdaterte kartleggingen fra 2019 til investeringsplanen av 2023. Studiene viser at hyllewareandelen har steget fra om lag 36 prosent i perioden til om lag 41 prosent i perioden 2019 til 2023 (Voldhaug et al., 2024).

Presiseringer, endringer og tillegg. Selv om anskaffelser er godt definerte før de gjennomføres kan omstendigheter oppstå som fordrer endringer i anskaffelsen. Materiellet kan eksempelvis ha hatt en omfattende prisøkning som krever økte bevilgninger eller reduksjoner i omfang. Behovet for endringer kan føre til forsinkelser og lavere produktivitet.¹¹ Dersom prosjekter endres skjer dette gjennom presiseringer, endringer og tillegg (PET). Endringer i sentrale rammebetingelser og oppdraget blir også utgitt som PET. Ved å observere hvordan antallet PET-er varierer sammen

¹⁰ Forskere ved FFI gikk i februar 2022 gjennom prosjektledernes statusoppdatering av alle de 106 prosjektene i gjennomføringsfasen som grupperte fritekstforklaringene for forsinkelse i ikke-ekskluderende grupper. Ressursmangel i prosjektavdelingene ble oppgitt som årsak til forsinkelser i 5 prosent av prosjektene. Analysen er ikke publisert.

¹¹ Omfangsendringer og interne og eksterne beslutningsprosesser ble oppgitt som årsak til forsinkelser i henholdsvis 1, 8 og 5 prosent av IKT-prosjektene i en FFI-analyse fra 2022 (Voldhaug & Presterdud, 2022). Berg et al. (2019) finner at prosjekter med omfangsendringer i snitt er assosiert med mindre forsinkelse, men estimatet er ikke signifikant forskjellig fra 0. Omfangsendringer er her definert som en absolutt endring i P50-verdi (på godkjenningstidspunktet) på mer enn 50 prosent.

med produktivitetsskårene kan vi muligens forklare hvordan endringsbehov påvirker produktivitet. Vi finner ingen tilgjengelige data som kan godt informere hvordan antall PET-er har variert mellom periodene.

Teknologi og verktøy. Innføring av nye verktøy og teknologi hylles som en uendelig brønn for økt produktivitet for alle virksomheter. I Forsvarsmateriell har også ny teknologi og nye verktøy blitt tatt i bruk for å effektivisere prosesser. Samtidig er det mulig at innføringen av (i hvert fall på kort sikt) fører økt til økte kostnader og dermed redusert produktivitet. Disse kostnadene er også inkludert i denne produktivetsmodellen ettersom avskrivninger for anleggsmidler og IKT-drift inkluderes i realkapitalkostnader. Det kan også tenkes at innføringen av prosjekttimeføring og andre datainnsamlingssystemer kan ha redusert teknisk effektivitet i FMA da ansatte har måttet dedikere mer tid enn før til å kode arbeidstimer under antagelse om at denne tiden før ble benyttet til produktivt arbeid. Denne kostnaden vil fanges opp både i lønnskostnader og i arbeidstimer ført på prosjekter. Vi har ikke grunnlag til å teste hypotesen, men nærmere analyser av dette kan være hensiktsmessig. Dette gjelder særlig for langsiktige investeringer i teknologi hvor nytten kan påløpe lenger frem i tid enn hva vår analyse dekker. I videre studier kan man vurdere å utelate kostnader knyttet til slike investeringer eller sammenstille kostnadene bedre med der verdien av investeringen faller.

Kompetanse- og organisasjonsutvikling. FMA har et økende fokus på profesjonalisering, samarbeid mellom kapasitetene, kompetansebygging og reduksjon i dobbeltarbeid. Dette kan ha medført produktivitetssøkninger. I 2019 gikk eksempelvis antall leveranseavdelinger fra 5 til 4, da tidligere Felleskapasiteter ble innlemmet i Landkapasiteter. FMA har også utviklet kurs og seminarer som samler beste praksis mellom kapasitetene og de har utredet kompetansebehov og muligheter for bedre samarbeid og koordinering mellom kapasitetene (se blant annet A. O. Presterud et al., 2021; Røtvold et al., 2018). I perioden har FMA også endret sin kompetansebeholdning. Waage (2021) viser blant annet at det i perioden 2013–2019 har vært en økning av sivil kompetanse. FMA har med andre ord i større grad enn før hentet inn personell med erfaring fra andre typer anskaffelser. Enkelte har også erfaring fra mer moderne organisasjoner. Dette kan ha hjulpet FMA å etablere mer effektive prosesser og arbeidsvaner. Samtidig har FMA redusert andelen militære tilsatte og dermed sin sektorspesifikke kompetanse. Dette kan ha påvirket produktiviteten i leveransene på måter vi ikke fanger opp i våre analyser, eksempelvis gjennom ytelsen på materiellet som anskaffes.

Leverandørforsinkelser. Forsinkelser hos leverandøren har vist seg å være den mest hyppige oppgitte grunnen for forsinkelser i prosjekter.¹² I årene 2020 til 2021 førte COVID-pandemien til store forsinkelser i flere forsyningsstrømmer da landegrenser ble stengt og folkemasser befant seg på sykehus eller i karantene. Ettervirkningene av COVID førte også til forsinkelser i 2022. Idet pandemien så ut til å være over angrep Russland Ukraina. Dette medførte et positivt etterspørselsjokk i forsvarsgoder og til at råvareknappheten vedvarte (se blant annet Hodnesdal & Hvinden, 2024). Etterspørselsjokket kan også ha ført til andre typer forsinkelser enn leverandørforsinkelser

¹² Kartleggingen av Berg og Voldhaug (2022) viste at prosjektene oppga leverandørforsinkelser som en av årsakene til prosjektforsinkelsene i 56 prosent av tilfellene. Dette var den største gruppen. Pandemien ble nevnt som en forsinkelsesårsak i 41 prosent av prosjektene.

for FMA. Blant annet har prisene på materiell steget kraftig, noe som fører til at estimater på anskaffelseskostnader må oppdateres og i noen tilfeller må alternativer revurderes eller nye bevilgninger må innhentes. I FMA sitt tildelingsbrev av 2023 (og 2024) skrives det blant annet at «økte priser og at krigen i Ukraina har lagt press på forsvarsindustrien og leverandører, gjør samlet at det i 2023 ikke lar seg gjøre å øke investeringstakten like mye som langtidsplanen forutsetter» (Forsvarsdepartementet, 2023). Markedet ble også nevnt som årsak til forsinkelser i 5 prosent av prosjektene med forsinkelser i 2022 (Voldhaug & Presterud, 2022). Leverandørforsinkelser i et prosjekt kan redusere produktivitet i FMA om ressurser ikke enkelt kan forflyttes til like produktive arbeidsoppgaver. Vi har ikke presise data på hvordan leverandørforsinkelser har variert i perioden, men en rimelig antagelse er at de har økt som følge av COVID og krigen i Ukraina. Dette kan være med på å forklare hvorfor årene 2021 og 2022 bryter med den ellers positive trenden.

Parallele oppgaver. En siste mulig forklaring på produktivitetsendringer vi gir i denne rapporten, kan være variasjon i, og tilgangen på, parallelle oppgaver. Et visst antall parallelle oppgaver kan være gunstig for produktivitet om oppgavene involverer betydelige mengder forventet og asynkron ventetid. I disse situasjonene kan perioder som ville vært lite produktive av årsaker utenfor Forsvarsmateriell sin kontroll forflyttes til mer produktive aktiviteter. Eksempler på dette kan være anbudsperioder, behandlingstid på forespørsler til leverandører, eller forventet ventetid ved faseoverganger, ambisjonsendringer eller søknader om nye bevilgninger. På den annen side vil positive effekter av høyere antall parallelle oppgaver forventes å avta og skifte fortegn om antallet parallelle oppgaver blir høyt eller disse må utføres samtidig. Det er krevende for mennesker å bytte oppgaver hurtig. Den negative effekten av å bytte oppgaver hurtig og ofte på produktivitet (byttekostnader) er godt dokumentert i forskning på individnivå (Adler & Benbunan-Fich, 2012; American Psychological Association, 2006). For at parallelle oppgaver skal være gunstige for produktivitet må altså kostnader av ventetid være høyere enn byttekostnadene. Det kan også tenkes at det finnes synergier mellom parallelle oppgaver som gjør det lettere å utføre sammen enn separat, selv ved hyppige bytter mellom oppgavene (se også Fischer & Plessow, 2015).

Vi har ikke mulighet til å konstatere om antallet parallelle oppgaver har økt, og sådan heller ikke hvorvidt ventetider har vært asynkrone eller om det har vært sterke synergier i oppgavene. Vi kan likevel kommentere noen relevante trender i dataene. Ifølge ATR-dataene har arbeidstid tilknyttet prosjekter i de seneste periodene gått ned og aktivitet knyttet til avhending, program, og samleposten *annet* gått opp. Investeringer (i budsjettermer) har også økt i perioden. Samtidig har antall årsverk ikke variert mye. Det er derfor mulig at antall arbeidsoppgaver per person også har økt. Det kan innebære at personell som tidligere jobbet med prosjekter har fått andre arbeidsoppgaver og at personell som før jobbet på ett prosjekt nå jobber på flere. Produktiviteten har økt sammen med antallet arbeidsoppgaver. Trolig er det derimot en grense for hvor mange arbeidsoppgaver personell kan ha uten at det går utover produktiviteten. En mulighet implementering av endringer i F24 er at FMA rendyrkes som anskaffelsesorganisasjon (Forsvaret, 2024). Dette kan innebære en reduksjon i spredningen av arbeidsoppgaver, og mer spesialisering og med dette ytterligere økninger i produktivitet på sikt. På kort sikt kan det imidlertid også føre til en nedgang

i produktivitet ettersom omstillinger ofte innebærer nye roller og arbeidsmåter som tar noe tid å bli vant med.

6 Produktivitet og tertialrapportering

Hvert tertial rapporterer FMA resultatoppnåelse til FD. I denne rapporteringen ser FMA sin produksjon opp mot egne planer som legges i starten av hvert år. Ettersom produksjon ses opp mot planer gir tertialrapporteringen innsikt i en annen form for effektivitet enn teknisk effektivitet (produktivitet). Vi kaller dette for planeffektivitet. I dette kapitlet ser vi på hvorvidt planeffektivitet korrelerer med teknisk effektivitet i etaten.

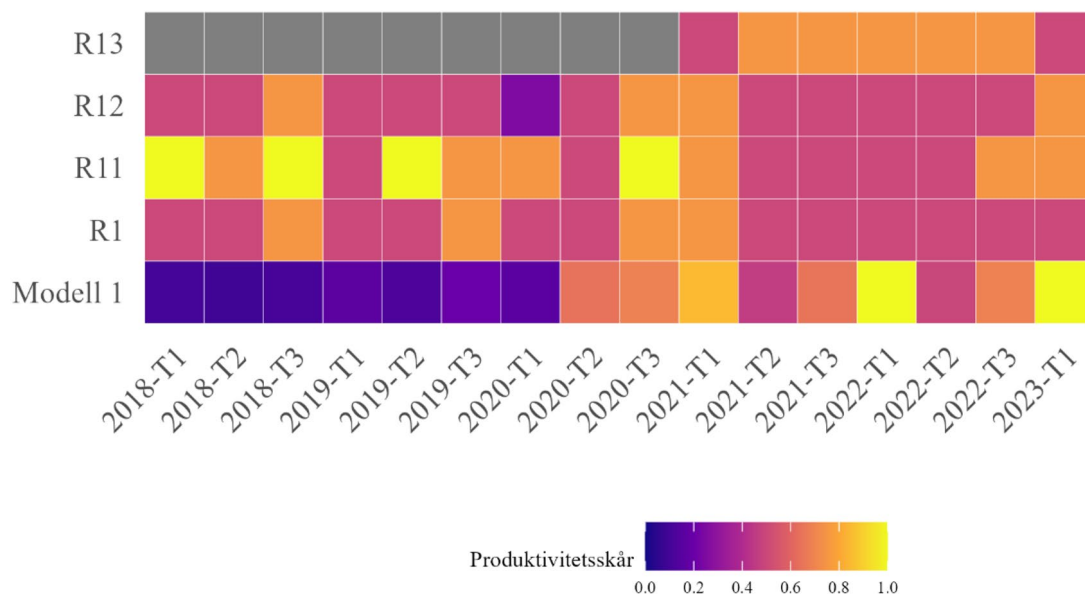
I tertialrapporteringen vurderer FMA hvor godt de har oppnådd resultatmål 1 (R1) – «Fremskaffelse av materiell og tjenester til avtalt tid, kostnad og ytelse». Dette resultatmålet består av styringsparameterne: (R11) «Bidra til og utarbeidelse av beslutningsunderlag i henhold til planlagt fremdrift og enkelte sentrale milepæler underveis», (R12) «Milepælsoppnåelse for godkjente investeringsprosjekter i henhold til planlagt og avtalt fremdrift», og (R13) «Tilfredsstillende forpliktelsesgrad for år 0 og 4 år frem i tid».¹³ Disse har vært nokså like i perioden 2016–2023, foruten R13 som først ble introdusert 2021. FMA gjør både kvantitative og kvalitative vurderinger av måloppnåelse på disse styringsparameterne. Summen av disse vurderingene danner grunnlaget for en av fire farger fra rød til grønn, hvor grønn er best. Oppnåelse på resultatmålet gis lignende en av fire farger basert på oppnåelse på styringsparameterne.¹⁴

Vi transformerer skårene fra tertialrapportene til å passe fargeskalaen vi benyttet i kapittel 4 og plottet utviklingen sammen med modell 1 i figur 6.1. Oppnåelse på styringsparameterne og resultatmålet har svinget rundt 0,5 og vi greier ikke lese et tydelig mønster verken stigende eller fallende for noen av styringsparameterne eller resultatmålet. I perioden er det kun R11 som får beste skår, og denne oppnås i tertialene før 2021. I kontrast er dette periodene som FMA har hatt lavest produktivitet i henhold til våre analyser. Periodene der FMA skårer dårligst langs styringsparameterne virker altså å være periodene FMA har høy produktivitet og vice-versa.

At utviklingen i planeffektivitet og estimert teknisk effektivitet er forskjellig kan forklares på flere måter. En forklaring er at utviklingen i planeffektivitet reflekterer reell produktivitetsutvikling bedre enn våre estimater (se boks 6.1). Dette kan være tilfellet om nivået på innsatsfaktorene har vært hensyntatt i planleggingen og ambisjonen i planene har vært konstant. Vi mener at dette ikke er tilfellet og at planene er frikoblet fra innsatsfaktorbruk og at ambisjonen er varierende. Vår vurdering er derfor at utviklingen i måloppnåelse på resultatmålene ikke godt reflekterer produktivitet eller nødvendigvis forbedring og effektivisering i FMA. Dette poenget understreket behovet for å både måle og analysere produktivitet og måloppnåelse.

¹³ Se f.eks. Vedlegg 1 til Direktør Forsvarsmateriells Plan– Mål og Resultatstyring, 2023.

¹⁴ Fargekoden som gis til R1 er derimot ikke det aritmetiske gjennomsnittet av skårene på styringsparameterne. Skåren fastsettes isteden kvalitativt med tett forankring i tidligere skårer.



Figur 6.1 Sammenligning av planeffektivitet og teknisk effektivitet. En celle viser skåren for et gitt tertial og en gitt modell, styringsparameter eller resultatmål 1.

Boks 6.1 – Teknisk effektivitet og gjennomføringsevne

Vi finner at FMA har økt sin produktivitet innen investeringsområdet i perioden 2018 til 2023. Samtidig er det en sterk opplevelse i forsvarssektoren om at gjennomføringsevnen i forsvarssektorens investeringsplaner er lav. Hvordan henger dette sammen?

Som vi har sett er det uproblematisk å både mene at gjennomføringsevnen er lav og samtidig at FMA har hatt stigende produktivitet. Gjennomføringsevne relaterer seg til hvor mye som ønskes gjennomført og er en annen type effektivitet enn teknisk effektivitet. Om det er stor avstand mellom målsettinger og det som faktisk blir gjennomført vil gjennomføringsevnen være lav. I teknisk effektivitet og produktivitet derimot, knyttes ressursbruk opp mot produksjon og målsetting utelates. Produktivitet kan sies å være lav (høy) dersom det produseres mindre (mer) enn forventet til et gitt antall ressurser. Under antagelsen om konstant skautbytte er det enkelt å se sammenhengen mellom de to. Ved klart definerte mål om gjennomføring og en oversikt over produksjonsfronten kan prosessforbedringer iverksettes og innsatsfaktorene opp- eller nedjusteres slik at produksjonsmålet nås.

Merk at gjennomføringsevne likevel kan være vedvarende lav til tross for at produktiviteten øker. Mulige forklaringer i disse tilfellene er blant annet eksempelvis at produktiviteten vokser saktere enn det målsettinger vokser, 2) innsatsfaktorene er utilstrekkelige, og tilgrensende, 3) målet for gjennomføring er for høyt.

7 Videreutvikling

Modellen(e) for å analysere produktivitet i FMA som foreslås i denne rapporten er ikke alene ment å være førende for forbedringsarbeidet. Likevel mener vi at det er viktig å begynne arbeidet med å måle på produktivitet der det er mulig – og der data er tilgjengelig – for å demonstrere nytten av måleindikatorer og for å gi innsikt. Modellering er også viktig for å synliggjøre de områdene der det er behov for mer data. Ettersom modellene benyttes og videreutvikles, og troverdigheten til produktivitetsanalysene øker, kan de enklere anvendes for å danne retning i virksomhetsutviklingen. Hva som bør inngå i modellen vil også bli klarere og et spesifikt sett av måleindikatorer for hver prosess kan defineres. Vi mener det viktigste er at produktivitetsmodeller tas i bruk, og at FMA bygger en evne til å lære og forbedre seg gjennom strukturert innsamling og analyse av data på produksjon og innsatsfaktorbruk. I dette kapittelet gir vi noen forslag til hvordan modellene kan videreutvikles for å gi bedre innsikt i produktivitet i hele FMA.

7.1 Måleparametere i investeringsprosessen

I analysen av produktivitet la vi til grunn det FMA samler av data i dag, og i stort var dette kun det som FMA rapporterer til FD. Vår vurdering er at disse dataene alene er utilstrekkelige for å gi klare signaler om produktivitet og på den måten legge grunnlaget for internt forbedrings- og moderniseringsarbeid.

En stor utfordring er at vår analyse ikke inkluderte kvalitetsrelaterte variabler knyttet direkte til produksjonen (se eksempelvis Hanson, 2010, som understreker viktigheten av kvalitetsindikatorer). Årsaken til at kvalitetsrelaterte variabler uteblir er at FMA i liten grad gjør vurderinger av kvaliteten i produksjonen, og at det ikke samles inn systematiske vurderinger fra bruker som også kan knyttes til de konkrete leveransene. Brukertilfredshet, slik definert i denne rapporten, fanger heller ikke opp kvalitet godt nok. Blant annet er den for generell og frikoblet fra produksjon og ressursbruk i perioden. Ved å utelate kvalitetsvariabler i modellene dyttes produksjonsfronten nedover. Dette kommer av at det, alt annet likt, alltid er mulig å produsere like mengder produkter og tjenester, men med lavere kvalitet (Besstremyannaya, 2013). Følgen er at vi overestimerer produktivitetsskårene og at produktivitetsendringer som følger av kvalitetsendringer ikke synliggjøres. Det betyr også at det er særdeles vanskelig å estimere produktivitetsfronten troverdig og derigjennom danne en forståelse av rotårsaker til produktivitetsendringer. Vi mener derfor det er behov for å dokumentere kvaliteten på leveransene gjennom interne vurderinger, og at vurderinger fra bruker klarere knyttes opp til leveranser i prosjekter.

Økt datainnsamling kommer ofte med en kostnad, og kvalitetsvurderinger kan være krevende. Det er viktig at vurderinger av kvalitet gjøres på en slik måte at verdien man får av å synliggjøre kvalitetsendringer ikke spises opp av økte kostnader knyttet til datainnsamling. Forsvarssektoren investerer derimot for svært store summer og kostnadene knyttet til kvalitetsvurderinger kan til sammenligning være små. Å synliggjøre kvalitet i prosjektets produksjon kan også være viktig for å skape transparens rundt hvordan offentlige midler disponeres. Nytteverdien av kvalitetsvurderinger vil også skille seg mellom prosjekter. For hyllewareanskaffelser vil kvalitetsvurderinger

være mindre relevante, og det kan være rimelig å anta konstant kvalitet. For anskaffelse av særlig avansert materiell, slik som ofte fører til komplekse prosjekter, er nytteverdien av å synliggjøre kvalitet i leveransene særlig stor.

Indikatorene for produksjon bør også videreutvikles. I denne rapporten benyttet vi milepæler for å måle mengder produsert innen et tidsrom. Milepælene er ikke utviklet for dette formålet, og benyttes i hovedsak for å måle fremdriften i prosjekter. En følge er at enkelte milepæler er for aggregerte for å måle produksjon innad i et tertial. Eksempelvis registreres det kun én obligatorisk milepæl for overføring av materiell når alt materiell er overført. Materielloverføringer i samme prosjekt kan derimot strekke seg over flere tertial og til og med flere år. Det samme gjelder kontrakter hvor den obligatoriske milepælen kun fanger opp når samtlige kontrakter er inngått. Inngåelse av enkeltkontrakter er derfor vanskelig å fange opp. For å bedre fange opp produksjonen innad i perioder, og dermed danne bedre innsikt i produktiviteten i perioden, kan det derfor være hensiktsmessig med mer hyppig datafangst for produksjon. Herunder er det viktig at innsamling av data automatiseres i størst mulig grad slik at gevinsten ved å få bedre oversikt over produksjon og produktivitet ikke gjennomgående forsvinner som følge av økte kostnader til datafangst.

7.2 Måleparametere i forvaltningsprosessen

I arbeidet med denne studien hadde vi en ambisjon om å måle produktivitet i forvaltningsprosessen. Vi vurderte det dithen at dataene som eksisterer i dag er utilstrekkelige, og det er behov for flere indikatorer og datainnsamling før produktivitet kan måles. For å vurdere produktivitet i hele FMA og for å få innsikt i hvorvidt økninger i produktivitet i én prosess påvirker produktiviteten i andre prosesser, vil det være nødvendig å etablere og samle inn data på indikatorer i forvaltningsprosessen.

Vi mener et fornuftig første steg er å samle inn og strukturere data på inspeksjoner, tilsyn og avvik. Et eksempel til etterfølgelse kan være Finanstilsynets produktivetsmodeller og rapportering. Finanstilsynet samler data på alle inspeksjoner de gjennomfører, og rapporterer blant annet på antallet inspeksjoner per industri, fagområde og inspeksjonsart (Finanstilsynet, 2023). De rapporterer også på antallet avvik som identifisert og handlet på. På samme måte kan det være hensiktsmessig for FMA å kategorisere avvik og tiltak i grupper som deler fellestrekk. I denne rapporten grupperte vi prosjekter analytisk etter kompleksitet ved bruk av data på prosjektnivå (se vedlegg C). Det kan derfor være behov for å gjøre mer detaljert datainnsamling på aktiviteter i forvaltningen utover hva som vil inngå i produktivetsmålingene, nettopp for å etablere hensiktsmessige kategorier. FMA har begynt innsamlingen av slike data, og er i prosess med å implementere et verktøy og system for avviksrapportering for hele FMA. Enkelte kapasiteter mener også selv at de har oversikt over antallet avvik, inspeksjoner og fagmyndighetskontroller. Datapunktene har vært for få og av for lav kvalitet til å benyttes i denne rapporten. Videreutvikling av systemene vil være viktig for å lykkes i å måle produksjonen på forvaltningssiden.

Å inkludere tilsyn, f.eks. i form av avvik og inspeksjoner, kan kreve endringer i modellene vi benytter for estimeringen. I denne rapporten har vi benyttet en enkel DEA-modell for å estimere produktiviteten på investeringsområdet. Å legge til materiellavvik i disse modellene kan gi

uheldige utfall og øke risiko for å trekke feilslutninger om produktivitet utvikling. Definerer vi eksempelvis identifisering av avvik som en positiv tjeneste risikerer vi at perioder der antall avvik reduseres fører til lavere produktivitet i FMA, alt annet likt. Dette behøver ikke reflektere faktisk produktivitet, da FMA kan ha konstant kvalitet og mengde i inspeksjonene, samtidig som avvik reduseres. Om vi på den annen side definerer avvik som en innsatsfaktor i produksjonen, eller som negativ eller uønsket produksjon, risikerer vi at perioder der FMA identifiserer mange avvik defineres som særlig ineffektive perioder for FMA. Dette virker også uheldig. Årsaken til problemene er at avvik oppstår til dels som en tilfeldig prosess og er til dels beroende på forhold utenfor organisasjonen. Etter at datainnsamling på avvik og inspeksjoner er gjennomført, og disse er hensiktsmessig kategorisert, vil det derfor være naturlig å gjøre endringer i estimeringsmetoden for å tilrettelegge for enklere tolkning av resultatene. En mulighet er å nærmere definere (egentlig anta) strukturen i problemet og modellere hvor mange avvik som oppstår hver periode. Det kan tenkes at det mest hensiktsmessige er å modellere avvik som noe som oppstår delvis tilfeldig og som noe som oppstår basert på adferd i organisasjonene. Da kan det også være relevant å benytte varianter av DEA som er designet for å håndtere situasjoner der deler av produksjonen er underlagt tilfeldigheter slik som Chance-Constrained DEA (Land et al., 1993).

8 Konklusjon og videre studier

I denne rapporten har vi foreslått en produktivetsmodell for FMA i prosessene investeringer og forvaltning. Vi har videre gjennomført en empirisk vurdering av produktivetsutviklingen for investeringer i perioden 2018 til 2023. Vår analyse tyder på at FMA har økt sin produktivitet i perioden. Dette stemmer godt med vår forståelse av forbedring og effektiviseringsarbeidet i FMA og inntrykket ansatte i FMA har av produktivetsutviklingen. Vi finner også at produktivetsutviklingen ikke har en klar sammenheng med resultatoppnåelse slik den er rapportert i tertialrapportene. Samtidig er det sannsynligvis målefeil i dataene som gjør produktivetsmålingene mindre troverdige. Det er også stor usikkerhet forbundet med estimatene. Det er derfor behov for ytterligere analyser og mer data for å øke troverdigheten til resultatene.

Modellen vi har foreslått i rapporten må videreutvikles slik at den i størst mulig grad kan benyttes til å gi beslutningsrelevant innsikt i forbedringsarbeid. Vi har i denne sammenheng foreslått at FMA utreder hvordan kvalitet i leveransene kan inkorporeres i modellen utover årlige brukerundersøkelser, som er et lite hyppig og svært overordnet mål på ekstern kvalitet. Responsen på brukerundersøkelsene er også kun løst knyttet til produksjonen i samme periode. Innen forvaltningsprosessen var det umulig å anvende modellen empirisk, ettersom data på produksjonen var utilgjengelige. Vi har anbefalt at FMA viderefører arbeidet med å samle inn og strukturere data på forvaltningssiden. Det vil også være behov for å nansere produksjonen på forvaltningssiden for å fange opp kvalitet og kompleksitet i de produktene og tjenestene som produseres. Dette kan gjøres kvalitativt samtidig med datainnsamling, eller kvantitativt slik det er gjort for investeringsprosessen i denne rapporten.

Til tross for at modellen bør videreutvikles kan modellen tas i bruk allerede i dag. Modellen forutsetter data som i stort allereide samles inn og struktureres, og rapportering i henhold til modellen er derfor lett å automatisere. Produktivetsmodellen skaper også et godt grunnlag for å etablere objektive normer for produktivitet, skape struktur i diskusjoner rundt produktivitet, og danne og teste hypoteser rundt produktivetsendringer. For å identifisere tiltak for ytterligere prosessforbedringer kan det imidlertid være behov for supplerende analyser av mer finkornede data.

Oppsummert anbefaler vi at FMA:

1. Tar i bruk modellen. Dette bør innebære en delvis automatisering av datainnsamling og estimering, og at analysekompetanse bygges opp og anvendes til å tolke resultatene.
2. Videreutvikler modellen. Herunder bør FMA inkorporere kvalitet knyttet til enkeltleveranse og annen produksjon, slik som produksjon i forvaltningsprosessen, i modellen.

8.1 Videre studier

Underveis i rapporten har vi foreslått retninger for videre studier. Blant annet bør mulige kvalitetsindikatorer for både produksjonen i investeringsprosessen og forvaltningsprosessen studeres. Herunder er det også relevant å tallfeste Forsvarsmateriell sin evne til å få best mulig materiell til billigst mulig investeringskostnad. Eksempelvis kan momenter ved kravmoderering eller markedsrekkevidde tallfestes. Vi har også diskutert forklaringer på produktivitetsendringer. Videre studier bør formulere mer konkrete hypoteser knyttet til disse og teste dem. Forutsetninger om skalaegenskaper og homogen produksjon og teknologi som vi har forutsatt bør også testes.

I tillegg til disse er det en rekke andre studier knyttet til produksjonen i Forsvarsmateriell som er relevant og viktig å studere. Vi fremhever to: bærekraft og insentiver. I denne studien ble bærekraft ikke trukket opp. Det er imidlertid ikke uvanlig at en organisasjon sitt klimaavtrykk inkorporeres i produktivetsmodeller enten som negativ produksjon (*undesirable output*) eller som en innsatsfaktor. Fokuset på bærekraft og klima er viktig for Forsvarsmateriell og forventes å bli enda viktigere over tid. Videre studier bør derfor se på hvordan bærekraft og klima hensiktsmessig kan inkorporeres i produktivetsmodeller for virksomheten.

I denne rapporten har vi fokusert på produktivetsmodellering og -analyse som er verktøy for å stadfeste utvikling, strukturere diskusjoner rundt utvikling og generere hypoteser rundt forbedring og effektiviseringstiltak. Vi har derimot fokusert lite på implikasjonen av å benytte produktivetsmodeller i styring og kontroll. Hanson (2010) påpeker at produktivetsmålinger som benyttes i styring kan påvirke adferd i organisasjonen. Bare kjennskap til at produktivitet måles og observeres av ledere kan til og med i seg selv føre til endret adferd (Celhay et al., 2024). Samtidig må det tenkes nøye gjennom hvilken produksjon og innsatsfaktorbruk som inkorporeres i produktivetsmodeller. Feil valg av disse kan føre til at organisasjonen i modellen belønnes om de handler mot sin hensikt. I vårt tilfelle kan det være i form av at svært komplekse prosjekter velges bort til fordel for mindre komplekse prosjekter. I kapittel 7 var vi også kort innom hvordan inkorporering av produksjon i forvaltningsprosessen kan motivere lignende oppførsel. Videre studier bør derfor se nærmere på hvordan produktivetsmodeller for styring i Forsvarsmateriell bør utformes slik at modellene ikke incentiverer utilsiktet og uønsket adferd.

Referanser

Adler, R. F., & Benbunan-Fich, R. (2012). Juggling on a high wire: Multitasking effects on performance. *International Journal of Human-Computer Studies*, 70(2), 156–168.

<https://doi.org/10.1016/j.ijhcs.2011.10.003>

American Psychological Association. (2006). Multitasking: Switching costs.

<https://www.apa.org/topics/research/multitasking>

Berg, H., Ofstad Presterud, A., & Øhrn, M. (2019). Military Off the Shelf Procurements: A Norwegian Case Study. *Defence and Peace Economics*, 30(1), 98–110.

<https://doi.org/10.1080/10242694.2017.1342182>

Besstremyannaya, G. (2013). The impact of Japanese hospital financing reform on hospital efficiency: A difference-in-difference approach. *The Japanese Economic Review*, 64(3), 337–362.

<https://doi.org/10.1111/j.1468-5876.2012.00585.x>

Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA, and R* (Bd. 157). Springer.

<https://doi.org/10.1007/978-1-4419-7961-2>

Celhay, P. A., Gertler, P., Olivares, M., & Undurraga, R. (2024). *How Managers Can Use Purchaser Performance Information to Improve Procurement Efficiency* (Working Paper 32141). National Bureau of Economic Research.

<https://doi.org/10.3386/w32141>

Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1994). *Data Envelopment Analysis: Theory, Methodology, and Applications*. Springer Netherlands.

<https://doi.org/10.1007/978-94-011-0637-5>

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444.

[https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)

Direktoratet for økonomistyring. (2013). *Bruk av informasjon om produktivitet i styringen—Et praktisk eksempel fra Statens innkrevingsentral*.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., & Shale, E. A.

(2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132(2), 245–259.

[https://doi.org/10.1016/S0377-2217\(00\)00149-1](https://doi.org/10.1016/S0377-2217(00)00149-1)

Emrouznejad, A., & Yang, G. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, 61, 4–8.

<https://doi.org/10.1016/j.seps.2017.01.008>

-
- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3), 253–290. <https://doi.org/10.2307/2343100>
- Finanstilsynet. (2023). *Årsrapport 2022*. <https://www.finanstilsynet.no/contentassets/3ae8ad709df043aa9c2635c2b511b66b/arsrapport-2022.pdf>
- Fischer, R., & Plessow, F. (2015). Efficient multitasking: Parallel versus serial processing of multiple tasks. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01366>
- Forsvaret. (2024, mars 5). *Forsvars-sektoren 2024*. Forsvarssektoren 2024. <https://www.forsvaret.no/soldater-og-ansatte/modernisering-og-effektivisering-i-forsvarssektoren/forsvarssektoren-2024>
- Forsvarsdepartementet. (2019a). *Instruks for Økonomi- og Virksomhetsstyring i Forsvarssektoren*. https://www.regjeringen.no/contentassets/437925d854d14d06b94df1731fef9355/instruks-for-okonomi-og-virksomhetsstyring-i-forsvarssektoren_1.februar-2019.pdf
- Forsvarsdepartementet. (2019b). *Retningslinjer for investeringer i forsvarssektoren*.
- Forsvarsdepartementet. (2023). *Tildelingsbrev for Forsvarsmateriell 2023*.
- Forsvarsmateriell. (2023). *Årsrapport 2022*.
- Hanson, T. (2010). *Produktivitetsmålinger i Forsvaret – metode og anvendelsesområder* (FFI-rapport 2010/01495). Forsvarets forskningsinstitutt (FFI).
- Hanson, T. (2013). *Produktivitetsmålinger i Forsvaret – eksperimentell modell for prioriteringseffektivitet* (FFI-rapport 2012/02265). Forsvarets forskningsinstitutt (FFI).
- Hanson, T. (2016). Efficiency and productivity in the operational units of the armed forces: A Norwegian example. *International Journal of Production Economics*, 179, 12–23. <https://doi.org/10.1016/j.ijpe.2016.05.016>
- Hodnesdal, S. A., & Hvinden, E. S. (2024). *Blir forsvarssektoren kompensert for den generelle prisveksten?* (FFI-rapport 24/00082). Forsvarets forskningsinstitutt (FFI).
- Kvalvik, S., & Berg-Knutsen, E. (2009). *Kostnadseffektiv drift av Forsvaret – Teoretisk fundament* (FFI-rapport 2009/00305). Forsvarets forskningsinstitutt (FFI). <https://ffi-publikasjoner.archive.knowledgearc.net/bitstream/handle/20.500.12242/1845/09-00305.pdf?sequence=1&isAllowed=y>
- Land, K. C., Lovell, C. A. K., & Thore, S. (1993). Chance-Constrained Data Envelopment Analysis. *Managerial and Decision Economics*, 14(6), 541–554.

-
- Liu, J. S., Lu, L. Y. Y., Lu, W.-M., & Lin, B. J. Y. (2013). A survey of DEA applications. *Omega*, 41(5), 893–902. <https://doi.org/10.1016/j.omega.2012.11.004>
- Presterud, A., Lien, B., & Voldhaug, J. E. (2022). *Porteføljestyling i forsvarssektoren – status i leveranseoppfølgingen* (Unntatt Offentlighet) (FFI-rapport 22/01167). Forsvarets forskningsinstitutt (FFI).
- Presterud, A. O., Lien, B., Waage, K., Halvorsen, O. K., Voldhaug, J. E., & Øverseth, H. P. (2021). *Forbedring og effektivisering i Forsvarsmateriell – tiltak og gevinster i perioden 2021–2024*. (Unntatt offentlighet) (FFI-rapport 21/00555).
- Presterud, A. O., & Øhrn, M. (2015). *Effektive materiellanskaffelser i Forsvaret – En studie av insentiver i investeringsprosessen* (FFI-rapport 2015/00555). <https://www.ffi.no/publikasjoner/arkiv/effektive-materiellanskaffelser-i-forsvaret-en-studie-av-insentiver-i-investeringsprosessen>
- Røtvold, A., Presterud, A. O., & Øhrn, M. (2018). *Personell og kompetanse i Forsvarsmateriell – Kartlegging av basislinje og analyser av fremtidig kompetansebehov*. (BEGRENSET) (FFI-rapport 18/01190).
- Simar, L., & Wilson, P. W. (1998). Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models. *Management Science*, 44(1), 49–61.
- Simar, L., & Wilson, P. W. (1999). Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research*, 115(3), 459–471. [https://doi.org/10.1016/S0377-2217\(97\)00450-5](https://doi.org/10.1016/S0377-2217(97)00450-5)
- Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1), 31–64. <https://doi.org/10.1016/j.jeconom.2005.07.009>
- Simm, J. (2022). *RDEA* [R]. <https://github.com/jaak-s/rDEA> (Opprinnelig verk utgitt 2014)
- Voldhaug, J. E., Prebensen, F. W., & Presterud, A. O. (2024). *Materiellinvesteringer i forsvarssektoren – Når vi målene?* (FFI-rapport 24/00554).
- Voldhaug, J. E., & Presterud, A. (2022). *Analyse av forsinkelsesårsaker i investeringsporteføljen* [Upublisert analyse]. Forsvarets forskningsinstitutt (FFI).
- Waage, K. (2021). *Personell og kompetanse i Forsvarsmateriell – Avgang og intern rotasjon i perioden 2013–2019* (Unntatt Offentlighet) (FFI-rapport 21/01159). Forsvarets forskningsinstitutt (FFI).

Waage, K., & Lien, B. (2020). *Personell og kompetanse i Forsvarets materiellinvesteringsprosjekter – Analyser av personellbehovet i bemanningsplaner* (Unntatt Offentlighet) (FFI-rapport 20/02132). Forsvarets forskningsinstitutt (FFI).

Vedlegg

Gjennomgående i rapporten har vi henvist interesserte lesere til vedleggene for mer detaljerte forklaringer av metoden og empiriske valg vi har tatt. Velkommen, interesserte lesere. I vedleggene gir vi en teknisk innføring i DEA og modellen vi har spesifisert (A), samt Simar og Wilson's foreslåtte metoder for å estimere usikkerhet og trekke inferens fra produktivitetsestimatene (B). Videre forklarer vi i mer detalj, og gjennom figurer, hvordan vi har valgt og definert kompleksitet og skilt investeringsprosjekter inn i enkle og komplekse prosjekter (C). Deretter viser vi de numeriske modellresultatene og hvordan de ville sett ut om dataene var aggregert til årnivå (D) og en oversikt over hvilke milepæler som ble benyttet til å approksimere produksjonen. Til slutt viser vi hvordan produksjonsvariablene i datasettet ser ut i modell 1 til 3 (F).

A Dataomhyllingsanalyse

Dataomhyllingsanalyse («Data Envelopment Analysis» – DEA) ble først foreslått av Charnes, Cooper og Rhodes (1978) som en metode for å måle produktivitet ved flere output og input basert på en tidligere artikkel av Farrell (1957). Cooper og Rhodes (1978) foreslo å måle teknisk effektivitet under flervariabel produksjon og innsatsfaktorbruk ved å konstruere *virtuelle* output og *virtuelle* input gjennom vektete summer av produkter og innsatsfaktorer (Charnes et al., 1994). I studien av Charnes mfl. (1978) ble metoden benyttet til å evaluere et utdanningsprogram for vanskeligstilte studenter som var satt i gang ved et utvalg amerikanske offentlige skoler. I utdanningsprogrammet var det flere utfall og flere innsatsfaktorer som benyttet. Både utfallene og innsatsfaktorene var vanskelig å verdsette etter tradisjonelle mål slik som priser. Ett eksempel på utfall som ble vurdert er *selvsikkerhet hos studenter* målt ved psykologiske tester og et eksempel på en innsatsfaktor er *tid brukt av mødre på høytlesning for studentene*. Ved bruk av DEA kunne forskerne likevel måle programmenes relative effektivitet. Siden den gang er DEA blitt brukt av en rekke forskere og aktører som søker å få innsikt i produktivitet i andre virksomheter der markeder for utfall og innsatsfaktorer ikke eksisterer eller er sviktende og der produksjonsfunksjonen er kompleks (Emrouznejad & Yang, 2018; Liu et al., 2013).

Fra 1978 til i dag har et stort antall varianter av DEA blitt foreslått og tatt i bruk. Heriblant er det varianter som tillater å modellere fronten som stykkevis log-lineær, stykkevis Cobb-Douglas eller stykkevis lineær (Charnes et al., 1994). I denne rapporten har vi benyttet en enkel modell med konstant skalautbytte hvor effektivitet defineres ved avstanden mellom best mulig og faktisk produksjon.¹⁵ I de neste delkapitlene ser vi nærmere på optimeringsproblemet for denne modellen og hva løsninger av problemet sier om effektivitet i enhetene som analyseres.

Vi har løst og reskalert det lineære programmet som finnes i A.2, men begynner likevel med å forklare problemet slik det er stilt opp i A.1. Problemet i A.1 mener vi er enklere å forstå ettersom det er klart koblet opp til definisjonen av produktivitet som forholdet mellom produksjon og ressursinnsats.

A.1 Optimeringsproblemet

I DEA benyttes matematisk programmering for å gi hver enhet de vektene som er optimale og for å samle alle enheters produksjon og innsatsfaktorbruk på samme skala. I våre modeller løses en linearisering av minimeringsproblemet,

¹⁵ En slik modell kalles også en output-orientert CCR-modell (Charnes et al., 1994).

$$\min_{v, \mu} \theta_b \stackrel{\text{def}}{=} \frac{\sum_i v_{ib} x_{ib}}{\sum_j \mu_{jb} y_{jb}} \quad \forall b \in B$$

$$\text{s.t.} \quad \frac{\sum_i v_{id} x_{id}}{\sum_j \mu_{jd} y_{jd}} \geq 1 \quad \forall d \in D$$

$$v/vY_b \geq \vec{1}$$

$$\mu/\mu Y_b \geq \vec{1}$$

hvor x_{it} er mengden forbruk av innsatsfaktor i i periode t , y_{jt} er produksjon av produkt j i periode t . Vektene for henholdsvis produksjon og innsatsfaktorbruk er μ og v . Minimeringsproblemet løses for hver periode t som finnes i settet av alle perioder T . Dette innebærer at hver periode t har egne vekter for produksjon og innsatsfaktorbruk. Disse vektene er de verdiene som for periode t minimerer brøken som ved konvensjon kalles θ gitt de betingelsene som følger under. Den første betingelsen er sentral. Betingelsen er egentlig flere betingelser i én, ettersom den må gjelde for alle perioder (betegnet ved $d \in D$). Denne sier at de vektene som velges aldri kan være slik at forholdstallet for de andre periodene (inkludert seg selv) er mindre enn 1. Følgen er at θ aldri (i minimeringsproblemet) kan bli mindre enn 1 og at θ for en periode b kun blir akkurat 1 om alle andre enheter er mindre effektive eller akkurat like effektive som b under de samme vektene.¹⁶ Sagt på en annen måte vil de periodene som skårer $\theta = 1$ være de periodene som er best i klassen gitt en viss kombinasjoner av vekter. Foruten å være løsningen på et optimeringsproblem er dette årsaken til at vektene kalles optimale. Det er også grunnen til at alle perioder som oppnår en $\theta_b \neq 1$ for alle mulige kombinasjoner av μ, v kalles dominerte – en slik produktivitetsskår innebærer at det minst er én annen periode som er bedre uansett hvilke vekter som velges. Perioder med en θ forskjellig fra 1 er også Pareto-ineffektive ettersom produksjonen kan økes uten å benytte flere innsatsfaktorer, eller innsatsfaktorene kan reduseres uten at produksjonen påvirkes (se figur 2.1 fra kapittel 2).

Det originale minimeringsproblemet legger opp til at alle produktene og innsatsfaktorene som inkluderes har frie vekter. Det er altså ingen struktur på hvorvidt ett produkt er verdt mer enn et annet og lignende for innsatsfaktorer. I vår rapport lot vi ikke vektene for innsatsfaktorbruk variere fritt, men definerte dem isteden. Dette fordi innsatsfaktorene vi benyttet allerede inkorporerte informasjon om priser fra nokså velfungerende markeder, herunder arbeidsmarkedet og markedet for husleie, maskiner mm. Vektene for produksjon μ knyttet til produktene $j = \{1 \dots m\}$ lot vi variere fritt, og vi hadde altså ingen formening om beslutningsgrunnlag er verdt mer enn materielloverføringer, eller om produksjon i enkle prosjekter er verdt mindre enn produksjon i komplekse prosjekter.

¹⁶ Utenfor minimeringsproblemet omgjør vi verdiene til å ligge mellom 0 og 1 ved å la $\theta_{ny} = \frac{1}{\theta_{gamlet}}$

At vektene for produksjon får variere helt fritt innebærer at periodene som skårer best på minst ett produkt relativt til ressursinnsats vil være effektive. Dette kan sees av minimeringsproblemet. Anta eksempelvis at vi ønsker å studere to bedrifter som måles på produksjonen av to produkter, epler og bananer. Den første bedriften produserer 100 epler, men ingen bananer. Den andre bedriften produserer 100 bananer, men ingen epler. Bedriftene benytter samme mengde innsatsfaktorer, og det finnes ingen priser på eplene og bananene. Under forutsetning om at epler verdsettes relativt høyt, dvs. $\mu_1 \leq \mu_2$ i minimeringsproblemet, vil bedrift 1 være effektivt. På den annen side vil bedrift 2 være effektiv under forutsetning om at bananer verdsettes relativt høyt. Det eksisterer altså vekter som gjør bedrift 1 mest effektiv og bedrift 2 mest effektiv og begge bedrifter vil ha en θ lik 1. Anta videre at en tredje bedrift oppdages som produserer både 100 bananer og epler, men er ellers lik de to andre bedriftene. Under alle vekter er bedrift 3 minst like effektiv som de andre bedriftene samtidig som den for noen vekter er bedre enn én av bedriftene. Slik vi har satt opp problemet vil likevel bedrift 1, 2 og 3 være like og helt effektive. Dette kommer av at vektene kan ta verdien 0. For å unngå situasjoner der svakt dominerende enheter vurderes som like effektive som andre enheter, erstattes ikke-negativitetsbetingelsene ofte med strengere betingelser (Dyson et al., 2001). Det minst påtrengende er å legge til ε på høyresiden av ikke-negativitetsbegrensningene hvor ε er uendelig liten positiv verdi. Alternativt kan vi definere et sett av absolutte eller relative betingelser for vektene (ibid.). Ettersom vektene fungerer som substitutter for informasjonen vi vanligvis ville fått gjennom priser i velfungerende markeder, forutsetter nærmere spesifisering av begrensninger på vekter at man har informasjon og produkter og innsatsfaktorerens verdi. Det er altså en fare for å havne inn på et spor der subjektive oppfatninger om verdi blir styrende om man fastsetter absolutte eller relative begrensninger for verdiene. Likevel vil styrken ved å benytte DEA også der slike begrensninger legges på vektene være at subjektive vurderinger fremstilles på en objektiv måte. I vårt tilfelle kunne det vært relevant å utforske absolutte og relative vektbegrensninger for brukertilfredshet – det er meningsløst å ha høy brukertilfredshet uten at noe annet produseres. Det kunne også vært relevant å sette relative begrensninger på verdien av kompleks produksjon opp mot enkel produksjon.

Går vi tilbake til optimeringsproblemet ser vi at det ikke har en beslutningsvariabel ω som tillater justering av periodens posisjon basert på skala i produksjonen. Ved fraværet av en slik skalavariabel tvinges effektivitetsfronten til å ligge på en rett linje gjennom origo. Med andre ord defineres fronten slik at det er et konstant forhold mellom skala og produktivitet (konstant skala-utbytte). For eksempel, hvis vi øker innsatsfaktorbruk for et punkt som ligger langs fronten må all produksjon også ha en proporsjonal økning for at det nye punktet også skal ligge langs fronten. Alternative varianter av DEA-modeller legger gjerne til en skalavariabel som tillater variabelt skala-utbytte eller ikke-økende skala-utbytte (Charnes et al., 1994). Implikasjonen for vår analyse er at alle periodene vurderes som skalauavhengige og at vi ser bort ifra skalaforskjeller som en forklaring på produktivitetsendringer. I litteraturen regnes konstant skala-utbytte som den mest naturlige antagelsen. Med andre ord må vi ha særskilt grunnlag til å modellere ut ifra andre skalaegenskaper. Som nevnt i kapittel 2 mener vi at konstant skala-utbytte er passende for periodene vi studerer ettersom det i dataene er liten variasjon i skala. Skalaegenskapene kan også statistisk testes f.eks. ved Simar og Wilson (2002, 2011) test for skala-utbytte i output-orienterte DEA-modeller.

A.2 Fra ratio til lineær formulering

Optimeringsproblemet vi gjennomgikk i delkapittel A.1 gir uendelig mange optimale løsninger (Charnes et al., 1994, s. 41), og det kan være vanskelig å løse. Charnes og Cooper (1962 som referert til av samme forfattere i 1994) foreslår derfor å sette $\mu_t Y_t = 1$ ved en betingelse, slik at problemet kan omformuleres til,

$$\begin{aligned} \min_{v, \mu} \theta_b &\stackrel{\text{def}}{=} \sum_i v_{ib} x_{ib} \quad \forall b \in B \\ \text{s.t.} \quad &\sum_j \mu_{jb} y_{jb} = 1 \\ &\sum_i v_{id} x_{id} - \sum_j \mu_{jd} y_{jd} \geq 0 \quad \forall d \in D \\ &v, \mu \geq \vec{1} \end{aligned}$$

Ettersom nevneren i flere av brøkene settes lik 1 kan vi fjerne den både i optimeringsproblemet og i det som før var de to siste betingelsene. Optimeringsproblemet går da fra å være et lineært forholdsproblem («*Linear Fractional Problem*») til å bli et lineært problem. Dette gjør problemet langt enklere å løse da et lineært program kan løses ved enhver lineær programløsningsmetode slik som f.eks. simpleks-algoritmen.

B Simar og Wilson (1998, 2007) – kort forklart

Simar og Wilson (1998, 2007) foreslår å benytte en metode for å anslå usikkerhet og trekke inferens fra DEA-estimer. I kapittel 4 benyttet vi metoden foreslått i Simar og Wilson (1998) (heretter SW98) for å anslå usikkerhet. I kapittel 4 kunne nevnte vi også en utvidelse av metoden foreslått i Simar og Wilson (2007) (heretter SW07) for hypotesetesting, men benyttet den ikke ettersom vi manglet tilstrekkelige data. I de neste avsnittene redegjør vi for disse metodene.

SW98. Produksjonen antas å være en usikker realisering fra homogene perioder med homogen teknologi. Med andre ord antas det i metoden at alle periodene skal produsere akkurat likt (justert for innsatsfaktorbruk) om vi kunne fjernet all usikkerhet. I SW98 blir usikkerhet anslått ved og gjentatt reskalere produksjon¹⁷ og reestimere produktivitetsskårer basert på simulerte data. Fordelingen av de nye estimatene på produktivitetsskårer benyttes som estimat på usikkerheten. Metoden kan oppsummeres i fem steg (Simar og Wilson, 1998):

1. Først, estimer produktivitetsskårer $\hat{\theta}_i \in [0,1]$ ved bruk av DEA for alle periodene $i \in T$.
2. Simuler et nytt datasett ved å plukke n tilfeldige θ_i^* fra (en glattet¹⁸) fordelingen av $\hat{\theta}$ hvor n er kardinaltallet til settet T .
3. Lag n simulerte realiseringer av produksjonen $y_i^* = \frac{\theta_i^*}{\hat{\theta}_i} y_i$ ¹⁹.
4. Estimer nye produktivitetsskårer $\hat{\theta}_i^*$ hvor produksjonen til periode i er erstattet med y_i^* .
5. Gjenta steg 2 til 5 B antall ganger, hvor B er et stort tall (ofte 1000).

Beregn snittet av de bootstrappede estimatene for periode i , $\bar{\theta}_i^*$ og den skjevhetsskorrigerede estimatoren ved $\tilde{\theta}_i = 2\hat{\theta}_i - \bar{\theta}_i^*$. Konfidensintervallet til en produktivitetsskår defineres så ved å sortere differansene mellom bootstrap-estimatene og de originale estimatene ($\theta_i^* - \hat{\theta}_i$) og fjerne en prosentandel ($\alpha/2 \times 100$) av verdiene på hver side av det sorterte settet (Hanson, 2016).

Av metoden kan vi se hvorfor det er store forskjeller mellom konfidensintervallene som vises i Figur 4.3. Fra steg 3 ser vi at produksjonen reskaleres basert på et tilfeldig trekk av de estimerte produktivitetsskårene. Trekker vi en verdi som er større enn det opprinnelige estimatet oppskalerer vi produksjonen. Trekker vi en verdi som er mindre enn det opprinnelige estimatet nedskalerer vi produksjonen. For perioder før 2020 – som har gjennomgående lave skårer, men er like seg imellom – innebærer det at vi nærmest alltid oppskalerer produksjonen. For senere perioder med høye skårer innebærer det at vi nærmest alltid oppskalerer produksjonen. Videre ser vi fra steg 4 at sannsynligheten for at en bootstrap-produktivitetsskår er svært forskjellig fra det opprinnelige estimatet avhenger av hvorvidt skåren i utgangspunktet er bestemt av effektivitet i et fåtall faktorer. Eksempelvis vil en periode som utmerker seg produksjonen av ett produkt (relativt til innsatsfaktorbruk), men produserer relativt lite av et annet, kunne få en langt lavere skår om produksjonen av førstnevnte nedskaleres kraftig. For at en periode som utmerker seg i produksjonen av flere produkter og tjenester må en kraftig nedskalering inntreffe samtidig for at skåren skal påvirkes. Konfidensintervallene sier altså noe om hvor sensitiv den opprinnelige produktivitetsskåren er for endringer i faktorer og hvorvidt skåren drives av generell produktivitet

¹⁷ Ved input-orientert modellering reskaleres innsatsfaktorbruken isteden.

¹⁸ Se Simar og Wilson, 1998, s. 55.

¹⁹ Se funksjonen *bias.correction.98* og *dea.output.rescaling* i RDEA-pakken.

eller ikke. Den korrigerede produktivitetsskåren kondenserer denne informasjonen i ett punkt heller i et intervall. Problemene ved at effektive perioder kan svakt dominere hverandre, som ble trukket opp i vedlegg A, blir sådan synliggjort i konfidensintervallene.

SW27. For å teste hypoteser om hvordan forklaringsfaktorer korrelerer med produktivitetsskårene foreslår Simar og Wilson (2007) å kombinere bootstrapping og fremgangsmåten i SW98 med trunkert regresjon. Simar og Wilson foreslår herunder to fremgangsmåter (algoritmer). Vi gjennomgår det Simar og Wilson kaller «Algoritme 2».

Algoritme 2 er en 7-stepsprosess. Som i SW98, begynner med (1) å beregne produktivitetsskårer ved DEA for hver periode. Deretter (2) estimerer $\hat{\beta}$ og $\hat{\sigma}_\varepsilon$ ved bruk av trunkert regresjon. I neste steg (3) benyttes $\hat{\sigma}_\varepsilon$ for å simulere n bootstrap-sett ved residualgjenvalg. Denne prosessen består av fire steg som gjentas L_1 ganger: først trekk feilledd fra en normalfordeling $N(0, \hat{\sigma}_\varepsilon^2)$ trunkert på venstreside ved $(1 - z_i \hat{\beta})$. Regresser $\theta_i^* = z_i \hat{\beta} + \varepsilon_i$ og benytt estimatet til å definere $x_i^* := x_i$, og $y_i^* := y_i \hat{\theta}_i / \theta_i^*$ for alle i . Deretter beregn $\hat{\theta}_i^*$ for alle i ved bruk av DEA og $Y^* = [y_1^* \dots y_n^*]$, $X^* = [x_1^* \dots x_n^*]$. Etter disse stegene har vi simulert L_1 bootstrap-estimer, $\{\hat{\theta}_{ib}^*\}_{b=1}^{L_1}$. Bootstrap-estimatene benyttes til å (4) beregne skjevhetsskorrigerte DEA-estimer $\hat{\theta}_i = \hat{\theta}_i - BIAS(\hat{\theta}_i)$. De skjevhetsskorrigerte estimatene benyttes til å (5) estimere $(\hat{\beta}, \hat{\sigma})$. I det neste steget (6) simulerer vi igjen nye bootstrap-sett i en løkke med L_2 iterasjoner, men denne gangen for å finne bootstrap-estimer for β og σ : Igjen trekk ε_i , men fra fordelingen $N(0, \hat{\sigma})$ trunkert på venstreside ved $(1 - z_i \hat{\beta})$. Beregn igjen $\theta_i^{**} = z_i \beta + \varepsilon_i$, og estimer $(\hat{\beta}^*, \hat{\sigma}^*)$ ved regresjon ($\hat{\theta}_i$ på z_i). Til slutt (7) benytt bootstrap-verdiene fra steg 6 og de originale estimatene fra steg 5 for å konstruere konfidensintervaller for hvert element av β og σ_ε .

Metoden som foreslås av Simar og Wilson (2007) er kompleks, men nyttig. Den lar oss estimere sammenhengen mellom produktivitetsskåren og forklarende faktorer. Den kan benyttes til å besvare spørsmål slik som: Hvor mye avhenger produktivitetsskåren av variasjon i prosjektlederturnover, eller hvor mye påvirkes produktivitetsskåren av markedsrisiko. Vi kan også fastslå hvor sikre vi er i konklusjonene vi trekker. Samler vi data over disse forholdene kan vi altså teste hypoteser rundt hva som driver og forårsaker produktivitetsskårene, gitt at modellene våre ellers korrekt fremstiller produktivitet i FMA.

Samtidig påpeker Simar og Wilson (1999, 2007) at metodene over kan benyttes for å estimere usikkerhet og trekke inferens under visse antagelser om datagenereringsprosessen. Enkelte av disse antagelsene stemmer ikke for vårt utvalg. Simar og Wilson (1998) forutsetter blant annet at prosessen som genererer ineffektivitet i produktivitetsskårene følger en identisk og uavhengig fordeling, altså at,

$$(\theta_1, \dots, \theta_n) \sim i. i. d F,$$

hvor F er en tetthetsfunksjon bundet mellom 0 og 1 (Simar & Wilson, 1998, s. 54). I vårt tilfelle vet vi at dette ikke stemmer. Produktivitetsskåren i en periode har en klar sammenheng med produktivitetsskårer i tidligere perioder. En materielloverføring kan ikke skje før en kontrakt er

etablert, og en kontrakt kan ikke etableres før en SSD er skrevet. En periode med flere SSD-er innebærer at neste periode kan produsere flere kontrakter, og en periode med flere kontrakter produsert øker sannsynligheten for overføringer av materiell i fremtidige perioder. Med andre ord er produktivitetsskårene korrelerte i tid eller såkalt seriekorrelerte. Følgen er at standardfeilen vil være underestimert. Om metoden skal benyttes for å trekke inferens fra våre modellresultater bør den derfor tilpasses at vi studerer produktivitet over tid. Se Simar & Wilson (1999) for et forslag til hvordan pseudosampling metoden fra 1998 kan tilpasses for å analysere produktivitet over tid.²⁰

²⁰ Simar og Wilson benytter i denne artikkelen begrepet temporal correlation istedenfor serial correlation eller det mer generelle begrepet autocorrelation.

C Prosjektkompleksitet

Forsvarsmateriell administrerer til enhver tid om lag 300 prosjekter. For alle disse prosjektene skal FMA ha produsert eller produsere beslutningsgrunnlag, konkurransegrunnlag, materielloverføringer og til slutt termineringsrapporter. Prosjektene er ikke like, og kan omfatte alt fra uniformer til feltvogner og fra arkivsystem til kampfly. For å skape troverdige sammenligninger må vi sammenligne epler med epler og pærer med pærer – vi med andre ord må sammenligne like produkter. Følgelig må vi justere for forskjellene mellom prosjekter.

Til tross for at materiellet (eller tjenestene) som anskaffes i investeringsprosjekter er tydelig forskjellige er det ikke like åpenbart at leveransene innebærer like store forskjeller. Ubåter og kampfly er forskjellige, men de prosessene som inngår i å produsere beslutningsgrunnlag, konkurransegrunnlag og kontrakter, overføringer av materiellet til drift og prosjektterminering kan være likt mellom prosjektene. I denne rapporten har vi valgt å benytte tid og kostnad som proksier for kompleksitet. Vi gjør ikke en kontinuerlig vektning, men forsøker heller å identifisere grupper som virker å skille seg fra resten av porteføljen. Fra tidligere analyser var vi bevisste på at et fåtall investeringsprosjekter er særlig ressurskrevende og dominerer investeringsbudsjettet. Disse prosjektene virket naturlige å behandle separat fra den resterende massen og vi valgte å dele prosjektene inn i to grupper. I rapporten har vi referert til dette som enkle og komplekse prosjekter. For å finne nøyaktige og numeriske grenser å dele prosjektene inn etter så vi på de empiriske kumulative fordelingene («Empirical Cumulative Distribution Function» – ECDF) til prosjektporteføljens varighet i tid, arbeidsintensitet (årsverk per år) og kostnad ved P50 slik de er i bemanningsplanene i fremskaffelsesløsningene. Dataene stammer fra kartleggingen gjennomført av Waage og Lien (2020).

I figur C.1 plotter vi tre paneler med ECDF-ene til tid, arbeidsintensitet og kostnad. I panelene sorterer vi prosjektene fra minst til størst etter variabelen vi er interessert i og plotter persentilen prosjektet tilhører langs x-aksen. Y-aksen viser verdien som en persentil står til. Eksempelvis varer et prosjekt i 50-persentil 5 år.

Av figur C.1 kan vi se at samtlige fordelinger er skjevt fordelt. Om fordelingene hadde vært normalfordelt ville den tatt formen av en S-kurve.²¹ Isteden ligner fordelingene på eksponentielle kurver med et grunntall høyere enn 1. At fordelingene krummer sterkt i høyere persentiler betyr at de er høyreskjeve. Dette samsvarer med tidligere analyser som viser at prosjektporteføljen domineres av noen få store prosjekter. Det er altså noen få prosjekter som skiller seg særlig ut på både kostnad, tid og arbeidsintensitet.

Vi ønsker å skille prosjektene i to grupper i kompleksitet. De få prosjektene som koster langt mer, er langt mer arbeidsintensive og/eller er bemannet langt lenger i gjennomføringsfasen er gode kandidater for komplekse prosjekter. Prosjekter som koster mye, fordrer ofte hyppigere kontrollposter og flere interessenter. Blant annet er Forsvarsdepartementet, og ikke Forsvaret, eier av prosjekter som har en kostnad over 500 millioner (eller 300 millioner for IKT-prosjekter), og

²¹ Eller en liggende S, ettersom vi plotter persentilene langs x-aksen heller enn y-aksen.

endringer for prosjektet må sanksjoneres av Stortinget. Arbeidsintensive prosjekter er også ofte bemannet av personer fra forskjellige enheter og etater. Flere interensenter nevnes ofte som en driver av kompleksitet. Prosjekter med lang varighet er gjerne utviklingsprosjekter med lang produksjonstid. Variablene er også korrelert. Ofte er det slik at prosjektene som er bemannet lenge også er prosjektene med høy arbeidsintensitet og høy kostnad. Samtidig finner vi at noen prosjekter er bemannet lenge, men har ikke særlig stor kostnad eller arbeidsintensitet. Andre prosjekter er bemannet i kort tid, men med svært mange årsverk per år. Prosjekter som har lav varighet, men høy arbeidsintensitet kaller vi for prioriterte, heller enn komplekse.

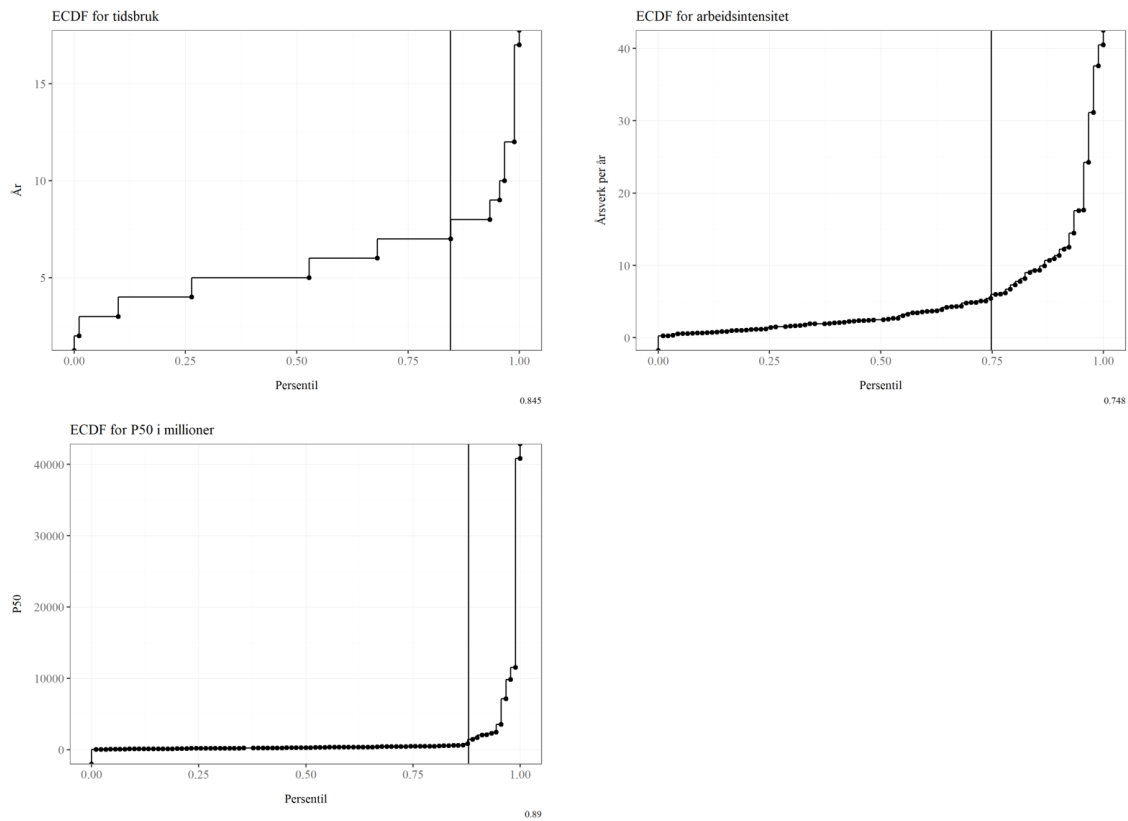
At et prosjekt er en uteligger i varighet, arbeidsintensitet eller kostnad trenger likevel ikke å bety at prosjektet er komplekst. Eksempelvis kan et prosjekt som koster mye kan være omfattende i antall (f.eks. en stor anskaffelse av hjelmer), men trenger ikke å være svært komplekst. Vi er altså bevisste på at dette ikke nødvendigvis er helt presise tilnærminger, men vi mener det er mer sannsynlig at prosjektene er komplekse når de er uteligger langs en av dimensjonene enn at de ikke er det.

Grensene for komplekse prosjekter finner vi ved visuelt å inspisere ECDF-ene. Vi setter grensene der vi først observerer en stor økning i y for en liten økning i x og som ligger etter 50-persentil. Disse grenseverdiene er vist ved vertikale linjer i figur C.1. Vi setter grensene slik at komplekse prosjekter enten må ha en kombinasjon av lang varighet og arbeidsintensitet eller særdeles høy kostnad. I datasettet vårt koder vi altså kompleksitet for prosjekt p som,

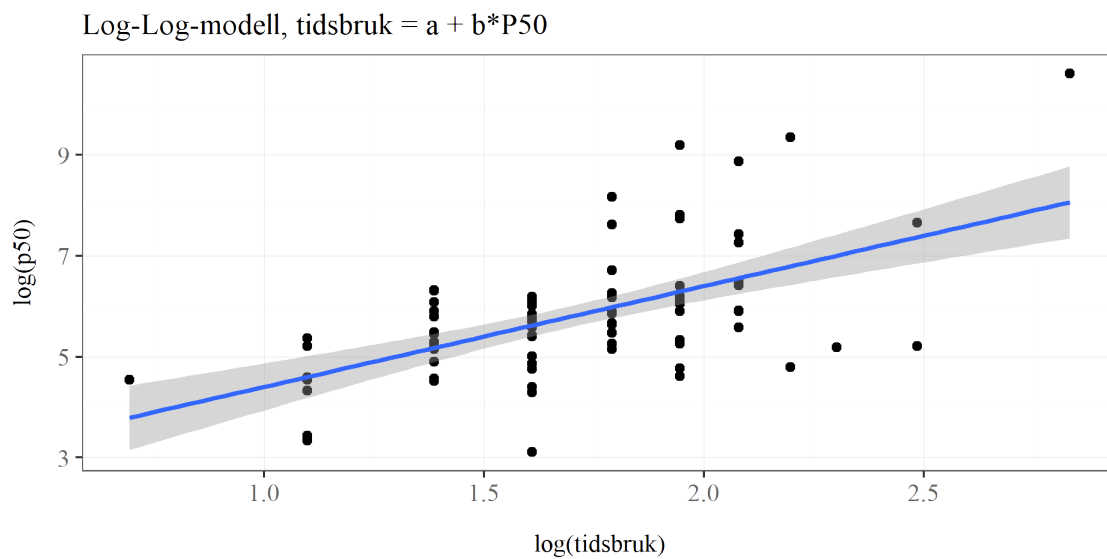
$$Kompleksitet_p = \begin{cases} 1, & \text{hvis } (t \geq 0.845 \wedge i \geq 0.784) \vee c \geq 0.88 \\ 0, & \text{ellers} \end{cases},$$

hvor t er tidspersentilen, i er arbeidsintensitetspersentilen og c er kostnadspersentilen. Vi tester også hvor sensitive resultatene våre er for ulike definisjoner av kompleksitet ved å variere persentilene. Lar vi persentilene variere enhetlig mellom 0.75 og 0.9, med 0.05 i mellomrom, finner vi at resultatene i stort er de samme. Estimatene skiller seg altså ikke signifikant fra hverandre når vi varierer definisjonen på denne måten.

For prosjekter som ikke fantes i kartleggingen til Waage og Lien (2020) har vi kun forventet anskaffelseskostnad og ikke data på forventet varighet og arbeidsintensitet. For disse prosjektene benytter vil kompleksitet kun bestemmes ut ifra kostnadene. Likevel er det slik at variablene i stort er godt korrelert. Faktisk korresponderer som regel prosjekter som har en varighet over 85-persentil i varighet om lag 89-persentil i kostnadsdataene. Korrelasjonen mellom kostnad og varighet er illustrert i figur C.2.

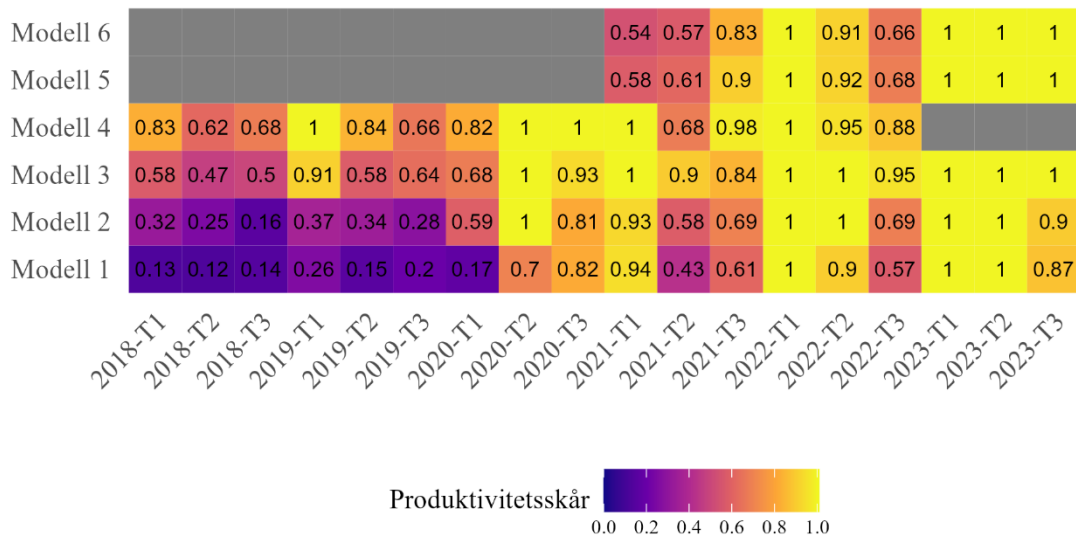


Figur C.1 Empiriske kumulative fordelinger for (fra topp) tidsbruk, arbeidsintensitet og forventet kostnad.

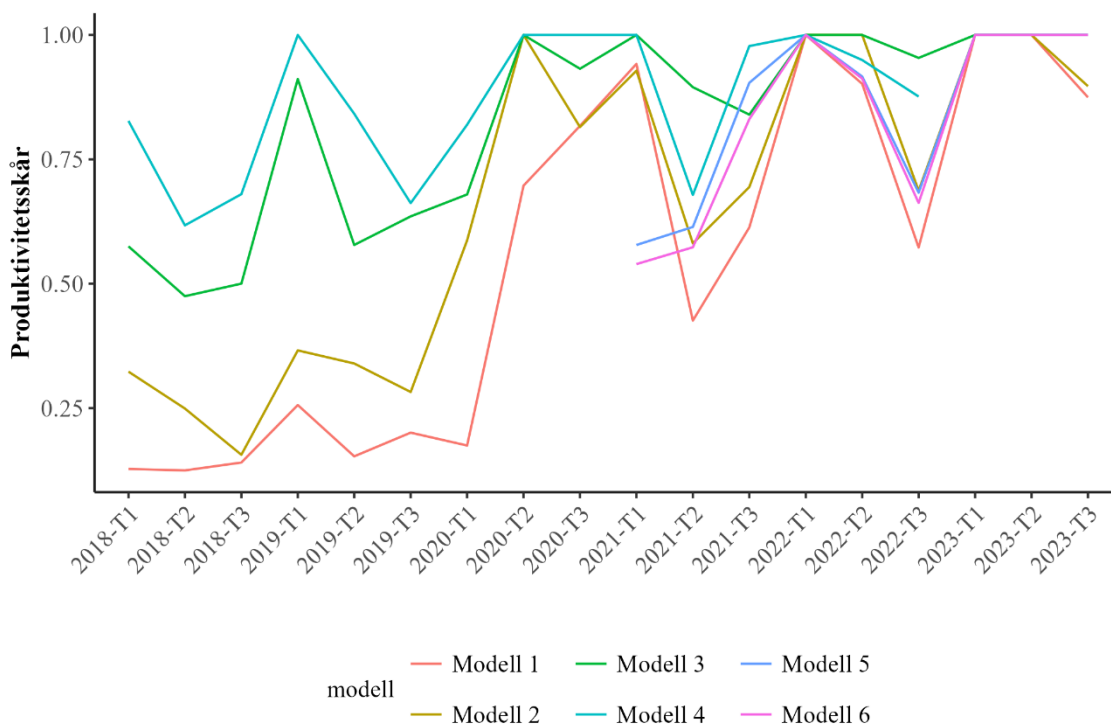


Figur C.2 Sammenhengen mellom prosjekters varighet (tidsbruk) og P50.

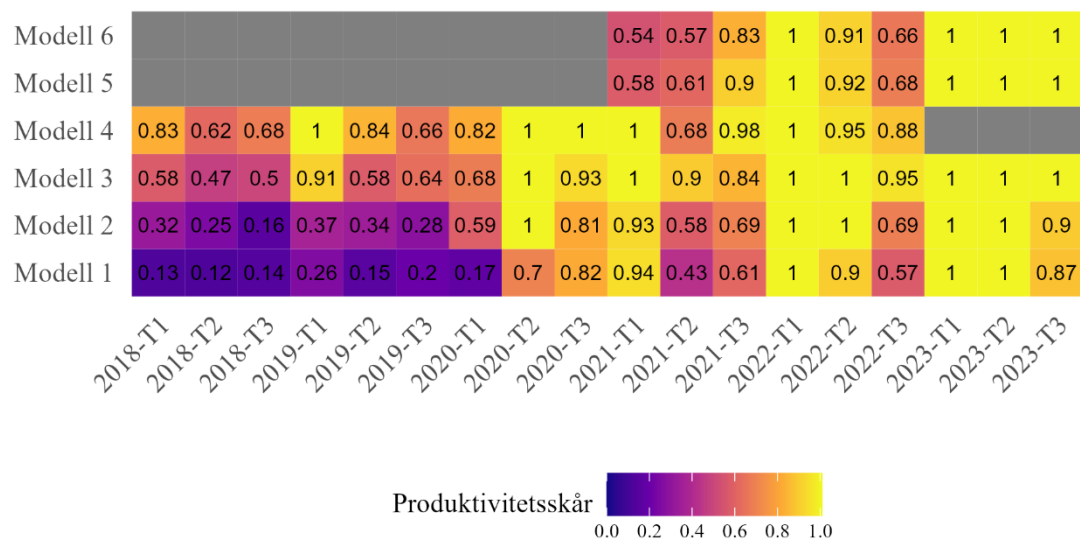
D Alternative og numeriske produktivitetsskårer



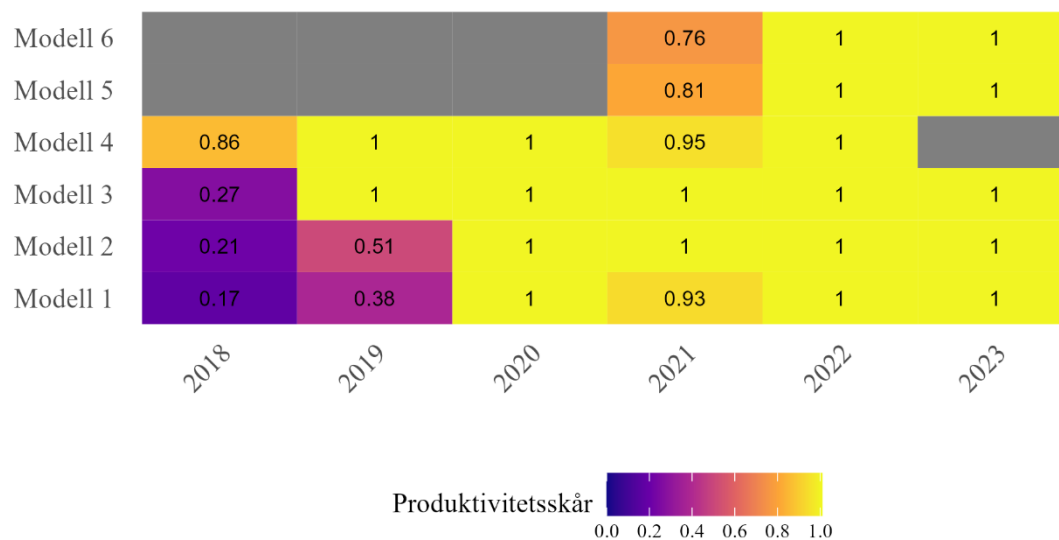
Figur D.1 Produktivitetsskårer fra figur 4.3 med numeriske verdier. Som ellers er produktive periode definert ved en skår nær eller lik 1.



Figur D.2 Utvikling i produktivitetsskårer fra de 6 modellene illustrert som linjediagram.



Figur D.3 Produktivitetskårer under variabelt skautbytte og output-transformering ved $y = y+1$. Output-transformasjon gjøres for å unngå null-produksjon. Modellen som er benyttet for å finne resultatene vist i denne figuren er transformasjonsuavhengig (translation invariant).



Figur D.4 Produktivitetskårer ved aggregering til årlig produktivitet. Aggregering til årlig nivå unngår noe av svakhetene ved hvordan produksjonen er periodisert i modellene (se blant annet variasjonen i figur 4.1). Det reduserer også antall observasjoner og derfor evnen til å skille mellom perioder.

E Milepæler

For å måle produksjonen i FMA benytter vi datoer for registrert fremdrift i prosjekter (milepæler). Fortrinnsvis benytter vi milepæler som er obligatoriske. Disse er like på tvers av prosjekter og skal registreres i likt omfang for alle prosjekter. I tabell E.1 viser vi hvilke obligatoriske milepæler vi benytter og hvordan de benyttes.

Tabell E.1 Oversikt over obligatoriske milepæler. Milepælene vi har benyttet for enkle prosjekter er fargelagt i blått og milepælene som er benyttet for komplekse prosjekter er fargelagt i grønt. Milepæler, som vi benytter for begge kompleksitetskategoriene, er fargelagt gult. Tilpasset fra Presterud et al. (2022)

PRINSIX-faser		
Konseptfase	Forprosjektfase	Gjennomføringsfase
OKVU Sendt	OSSD sendt	GO sendt
OKVU Mottatt	OSSD mottatt	GO mottatt
Oppdragsanalyse gjennomført	Prosjektet er kvalifisert	Oppdragsanalyse gjennomført
KVU-produksjon planlagt	SSD produksjon planlagt	Gjennomføringen planlagt
Oppstartsmøte gjennomført	Oppstartsmøte gjennomført	Alle konkurransegrunnlag utarbeidet
Konseptvalg valgt	SSD produsert	Alle konkurransegrunnlag godkjent
Løsning valgt	SSD høring gjennomført	Alle konkurransegrunnlag produsert
Neste fase planlagt	SSD presentert for kapasitetssjef	Alle tilbud evaluert
Hoveddokument produsert	SSD godkjent på ledermøte	Alle kontrakter forhandlet
KVU godkjent ledermøte	ASSD fremsendt PE	Alle kontrakter signert
KVU fremsendt	SSD godkjent	PLP oppdatert
KVU godkjent		Leverandør fulgt opp

		Alt materiell mottatt FMA
		All materielloverføring utført i SAP
		Kontrakt(er) avsluttet
		Termineringsrapport fremsendt
		Termineringsrapport godkjent

I tillegg til obligatoriske milepæler benytter vi prosjektmilepæler for deler av produksjonen som skjer i komplekse prosjekter. Disse milepælene bestemmes av prosjektet selv og er derfor forskjellige mellom prosjekter og registreres også i ulik hyppighet mellom prosjekter. Prosjektene bestemmer behovet for prosjektmilepæler ut ifra behovet for å måle fremdrift. Sånn sett er det mer sannsynlig at komplekse prosjekter har flere prosjektmilepæler enn enkle prosjekter. I tabell E.2 beskriver vi hvilke prosjektmilepæler vi benytter. Eksakt hvilken tekst som er tilknyttet den enkelte milepælen vil variere, og tabellen viser derfor essensen vi har forsøkt å trekke ut fra prosjektmilepælene heller enn av mange konkrete formuleringer. Enkelte vil bekymre seg for at bruk av analyse vil være svært sensitiv til bruk av prosjektmilepæler. Vi tester dette ved å filtrere vekk prosjektmilepæler fra dataene og estimere produktivitetsskårene på nytt. Vi finner at resultatene for perioden vi analyserte ikke er særlig sensitiv til bruk av prosjektmilepæler.

Tabell E.2 Oversikt over prosjektmilepæler som er benyttet til å approksimere produksjonen av komplekse prosjekter.

PRINSIX-faser		
Konseptfase	Forprosjektfase	Gjennomføringsfase
-	-	KUA/LCC-analyse gjennomført
-	-	Kontrakt inngått/signert
-	-	Materiell overført i SAP
-	-	Materiell sikkerhet og forvaltningsmessig godkjent

F Illustrasjon av datasett

I modell 1–3 varierer vi innsatsfaktorsiden ved å aggregere produksjonen på forskjellige måter. I figur F.1 viser vi hvordan produksiden ser ut i henholdsvis modell 1, 2 og 3.

År	Tertial	Kompleks Produksjon	Enkel Produksjon			
2018	1					
2018	2					
2018	3					
2019	1					
2019	2					
2019	3					
År	Tertial	Beslutningsunderlag	Konkurranses grunnlag og Kontrakter	Materielloverføringer	Termineringsrapporter	
2018	1					
2018	2					
2018	3					
2019	1					
2019	2					
2019	3					
År	Tertial	Enkle Beslutningsunderlag	Komplekse Beslutningsunderlag	...	Enkle Termineringsrapporter	Komplekse Termineringsrapporter
2018	1					
2018	2					
2018	3					
2019	1					
2019	2					
2019	3					

Figur F.1 Illustrasjon av produksiden i datasettet for modell 1–3.

Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan. Med særskilte fullmakter underlagt Forsvarsdepartementet.

FFIs formål

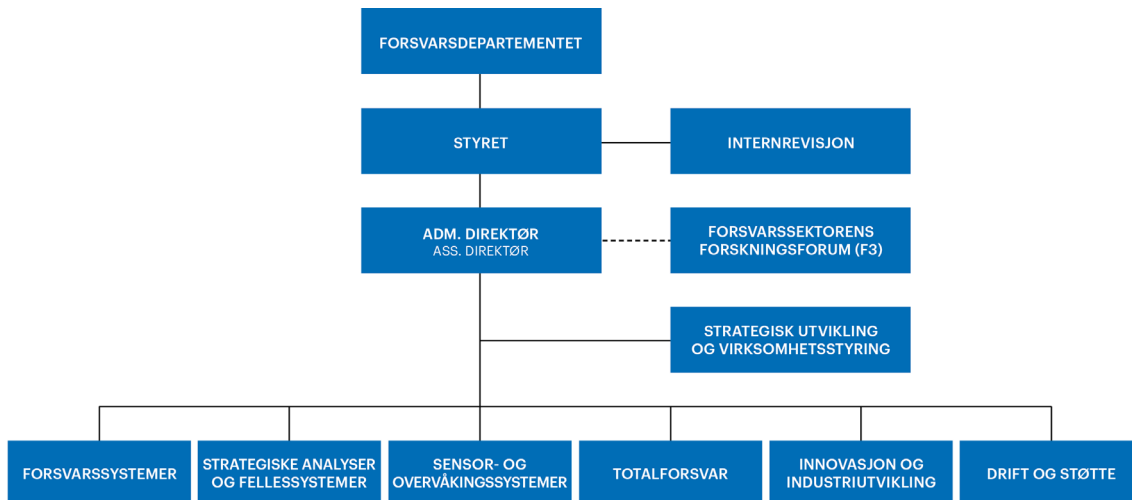
Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

FFIs visjon

FFI gjør kunnskap og ideer til et effektivt forsvar.

FFIs verdier

Skapende, drivende, vidsynt og ansvarlig.



Forsvarets forskningsinstitutt (FFI)
Postboks 25
2027 Kjeller

Besøksadresse:
Kjeller: Instituttveien 20, Kjeller
Horten: Nedre vei 16, Karljohansvern, Horten

Telefon: 91 50 30 03
E-post: post@ffi.no
ffi.no

Norwegian Defence Research Establishment (FFI)
PO box 25
NO-2027 Kjeller
NORWAY

Visitor address:
Kjeller: Instituttveien 20, Kjeller
Horten: Nedre vei 16, Karljohansvern, Horten

Telephone: +47 91 50 30 03
E-mail: post@ffi.no
ffi.no/en