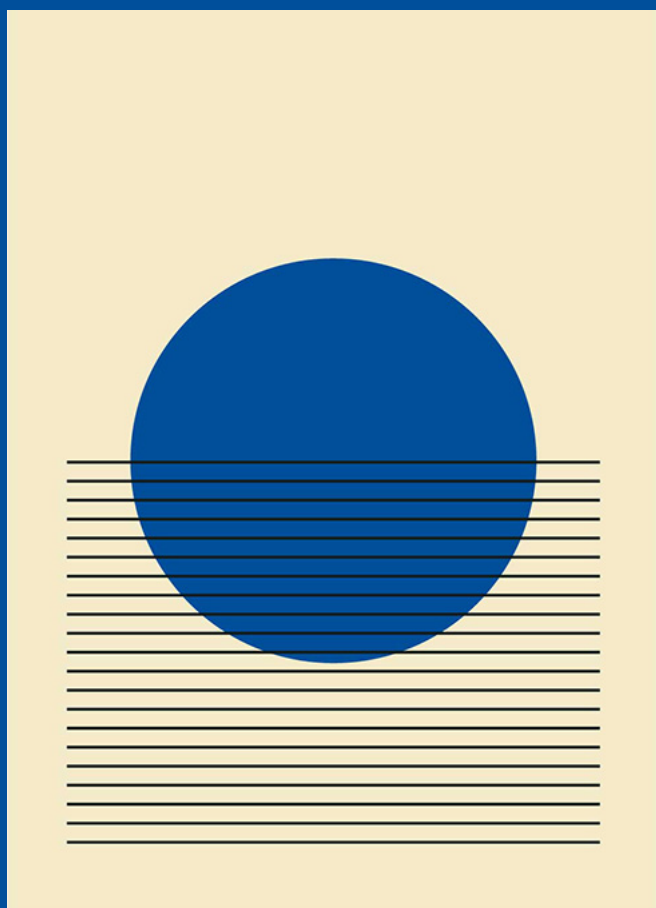


Udvikling af regnemodel for uddannelsesinvesteringer (REFUD)

Teknisk dokumentationsrapport



*Udvikling af regnemodel for uddannelsesinvesteringer (REFUD)
Teknisk dokumentationsrapport*

© VIVE og forfatterne, 2022

e-ISBN: 978-87-7582-008-5

Projekt: 302060

Modelfoto: Ricky John Molloy/VIVE

Finansiering: Børne- og Undervisningsministeriet

VIVE

Det Nationale Forsknings- og Analysecenter for Velfærd

Herluf Trolles Gade 11

1052 København K

www.vive.dk

VIVEs publikationer kan frit citeres med tydelig kildeangivelse.



Forord

Dette er en teknisk dokumentationsrapport udarbejdet i forbindelse med udvikling af en generisk prototype afregnemodel for uddannelsesinvesteringer (REFUD) for indsatsen to-voksen-ordning. Arbejdet udspringer af et indledende arbejde, hvor VIVE afdækkede mulighederne for udbredelse af SØM til indsatser på børne- og undervisningsområdet. Arbejdet og rapporten er finansieret af Børne- og Undervisningsministeriet.

REFUD-prototypen på børne- og undervisningsområdet består af et beregningsmodul udviklet i R Shiny; inputdata i beregningsmodulet er udarbejdet via programmer på Danmarks Statistiks forskermaskine, mens empirien vedrørende effekter af to-voksen-ordning er hentet fra et studie i forskningslitteraturen.

Rapporten er skrevet af senioranalytiker Rune Vammen Lesner, forsker Anne Brink Nandrup samt seniorforsker Karl Fritjof Krassel.

Rapporten har været læst og kommenteret af to eksterne reviewere, ligesom den har gennemgået VIVEs interne kvalitetssikring. Både eksterne og interne reviewere takkes for gode og brugbare kommentarer.

Hans Hummelgaard

Forsknings- og analysechef for VIVE Effektmåling



Indholdsfortegnelse

Sammenfatning	6
---------------	---

1	Indledning	9
---	------------	---

2	Modelramme	11
2.1	Afsæt	11
2.2	Operationalisering	14
2.3	Aggregering	18

3	Indsatsen to-voksen-ordninger	20
3.1	Indsatsvarianter	20
3.2	Helårseffekter	20
3.3	Målgruppe og depreciering	21

4	Beregningsmodul	23
4.1	Programstruktur og filbeskrivelser	23
4.2	UI-værktøj og brugerrejse	30
4.3	Nye indsatser, nye data og ny evidens	34

5	Beregninger på Danmarks Statistiks forskermaskine	37
5.1	Modelstrukturen og beregninger på Danmarks Statistiks forskermaskine	38
5.2	Link-outcomes	43
5.3	Konsekvenser	44
5.4	Priser	49

Litteratur	55
------------	----

Bilag 1 Ordforklaringer	56
-------------------------	----

Bilag 2 Versioner af R-pakker	57
<hr/>	
Bilag 3 Kodestruktur til databearbejdelse på Danmarks Statistiks forskermaskine	58
<hr/>	
Bilag 4 Kodestruktur til beregning af model-input på Danmarks Statistiks forskermaskine	60
<hr/>	
Bilag 5 Versioner af R-pakker	61
<hr/>	
Bilag 6 Kodestruktur til databearbejdelse på Danmarks Statistiks forskermaskine	62
<hr/>	
Bilag 7 Kodestruktur til beregning af model-input på Danmarks Statistiks forskermaskine	64

Sammenfatning

Formålet med dette projekt er at udvikle en prototype af regnemodel for uddannelsesinvesteringer (REFUD). Arbejdet har taget udgangspunkt i indsatsen to-voksen-ordningen og beror på tilgangen anbefalet i VIVE rapporten "Afdækning af muligheder for udbredelse af SØM til indsatser på børne- og undervisningsområdet" (Jacobsen et al., 2022). Projektet indeholder i alt fire deliverancer:

1. Denne tekniske dokumentationsrapport
2. Et beregningsmodul udviklet i R Shiny
3. Programmer til bearbejdning af inputdata på Danmarks Statistiks forskermaskine
4. En workshop med henblik på vidensoverdragelse til BUVM.

Rapportens indhold

Rapporten dokumenterer selve beregningsmodulet og beregningsmodulets forskellige inputs. Rapporten indledes dog med at beskrive modelrammen og afsættet for den praktiske operationalisering af REFUD. Afsættet sker fra Jacobsen et al. (2022), men der diskuteres også elementer, som er nødvendige at få præciseret for den praktiske operationalisering:

- Diskontering
- Depreciering
- Persistens
- Timing
- Aggregering.

Indsatsen to-voksen-ordning

Det empiriske grundlag for rapporten er Andersen et al. (2020), som dokumenterer et RCT med to-voksen-ordninger i danske 6. klasser. På baggrund af studiet og i dialog med Børne- og Undervisningsministeriet (BUVM) er det besluttet at præsentere brugeren for to forskellige indsatsvarianter i modellen:

- En ekstra ikke-læreruddannet
- En ekstra læreruddannet.

For at lette beregningerne i modellen foretages en omregning af resultaterne i Andersen et al. (2020), således at de korresponderer med det tilfælde, hvor indsatsen er implementeret hele året. Omregnet til helårseffekter betyder det, at *effect size* ved fuld implementering af indsatsvarianten "en ekstra læreruddannet" er 0,244, mens *effect size* ved fuld implementering af indsatsvarianten "en ekstra ikke-læreruddannet" er 0,209.

Som beskrevet i Andersen et al. (2020) var målgruppen for lodtrækningsforsøget 6. klasse. Ud fra dette empiriske grundlag vurderer VIVE, at modellens beregninger på indsatser af to-voksen-ordninger vil være mest retvisende for elever på mellemtrinnet (4.-6. klasse). I mangel af empiri og for at imødekomme brugere, der måtte ønske at foretage beregninger på folkeskolens øvrige klassetrin, tillader modellen dog også beregninger på indskoling (0.-3. klasse) og udskoling (7.-9. klasse). Modellen antager desuden, at effekten af en to-voksen-indsats vil være den samme i første og efterfølgende år ved en eventuel gentagen indsats. Det understreges, at denne linearitet i effekter over år udelukkende er en antagelse, da evidensen alene baserer sig på den 1-årige indsats i Andersen et al. (2020).

Beregningsmodulet

Beregningsmodulet er helt central i REFUD-prototypen og har til formål at omsætte data fra Danmarks Statistiks forskermaskine, evidens fra forskningslitteraturen og bruger-input til konsekvensberegninger for en given indsats. Til det formål har VIVE udviklet et beregningsmodul i R med pakken Shiny, som rapporten dokumenterer.

Beregningsmodulet er kodet med stor fleksibilitet, og alle input-parametre kan ændres af brugeren. Beregningsmodulet er ligeledes kodet generisk, så det er muligt let at udvide med nye data, nye indsatser og ny evidens.

Beregninger på Danmarks Statistik

Inputdata til REFUD-prototypen baseres på de detaljerige registre, der er tilgængelige på Danmarks Statistiks forskermaskine. Med udgangspunkt i den opstillede model i afdækningsrapporten (Jacobsen et al., 2022) omhandler beregningerne på forskermaskinen tre konkrete output-dimensioner:

1. Lokale baseline-standardafvigelser på link-outcomes
2. Betingede korrelationsmatricer for sammenhængen mellem link-outcome og konsekvenser rensset for individspecifik baggrundskaraktistika i en op til 4-årig horisont
3. Lokale baseline-standardafvigelser på modellens konsekvenser.

Herudover er enkelte priser beregnet ved hjælp af registerdata.

1 Indledning

Den Socialøkonomiske Investeringsmodel (SØM) blev i 2016-2017 udviklet af VIVE (daværende KORA) og Incentive for Socialstyrelsen på baggrund af en bevilling på satspuljeaftalen i efteråret 2015. Den første modelversion blev offentliggjort i januar 2018, og for nuværende er version 2.3 tilgængelig til download på Socialstyrelsens hjemmeside.

Modellens formål er give et bud på de økonomiske konsekvenser over tid af sociale indsatser over for afgrænsede målgrupper af socialt udsatte voksne og børn og unge. Det har først og fremmest været tanken, at modellen kunne anvendes som vidensbase og til at understøtte beslutninger om iværksættelse af specifikke indsatser over for specifikke lokale målgrupper.

Det er dog rimeligt at antage, at det ikke kun er velfærdsindsatser på det specialiserede socialområde, der har afledte økonomiske konsekvenser. Indsatser inden for børn og undervisning må forventes – ud over den direkte effekt på fx læring eller trivsel – at have en effekt på efterfølgende brug af offentlige ydelser og tilbud som fx specialundervisning, tilbøjeligheden til at starte på en ungdomsuddannelse og eventuelt også sundhedstilbud. Derudover er der en klar påvist sammenhæng mellem et succesfuldt skole- og uddannelsesforløb og individers senere tilknytning til og fastholdelse på arbejdsmarkedet. Derfor giver det god mening at udvikle en socialøkonomisk investeringsmodel til børne- og undervisningsområdet. Målet med modellen er at beregne økonomiske nettogevinster af en række udvalgte brede indsatser lokalt i kommunerne (modsat den eksisterende SØM, som er målgruppebaseret).

Således er en vigtig forskel mellem den nærværende REFUD på børne- og undervisningsområdet og Socialstyrelsens SØM for udvalgte grupper af socialt udsatte individer. Hvor Socialstyrelsens SØM har en målgruppebaseret tilgang til beregning af indsatseffekter på forbrug af offentlige ydelser, har REFUD en indsatsbaseret tilgang og måler på den generelle population af normalklasse-elever i folkeskolen. Det er derfor tanken, at REFUD løbende skal tilføjes en lang række indsatser, som brugeren herefter kan sammenligne indsatsgevinster for.

I rapporten "Afdækning af muligheder for udbredelse af SØM til indsatser på børne- og undervisningsområdet" (Jacobsen et al., 2022) afdækker VIVE mulighederne for at udbrede principperne bag SØM til indsatsområder inden for børne- og undervisningsområdet. For hver indsats undersøges det, hvordan den kan afgrænses, hvilke mekanismer der kan drive de økonomiske konsekvenser, om der er evidens for, at indsatsen virker, og hvordan modelberegninger kan gennemføres. Med udgangspunkt i denne afdækning er målet at udvikle en generisk prototypemodel, der dels kan benyttes til konkret beregning af de socialøkonomiske

effekter af en given indsats, dels kan danne udgangspunkt for udarbejdelse af REFUD for andre indsatser på børne- og undervisningsområdet.

Prototypen vil tage udgangspunkt i indsatsen to-voksen-ordningen og tage afsæt i anbefalingerne i VIVE-rapporten "Afdækning af muligheder for udbredelse af SØM til indsatser på børne- og undervisningsområdet" (Jacobsen et al., 2022). Projektet indeholder i alt fire delleverancer:

1. Denne tekniske dokumentationsrapport
2. Et beregningsmodul udviklet i R Shiny
3. Programmer til bearbejdning af inputdata på Danmarks Statistiks forskermaskine
4. En workshop med henblik på vidensoverdragelse til BUVM.

Hovedleverancerne i projektet er delleverance 2 og 3. Den tekniske dokumentationsrapport kan læses selvstændigt, men skal primært fungere som støtte til punkt 2-4. I tilfælde, hvor den tekniske dokumentationsrapport synes mindre grundig, skyldes det, at dokumentation er givet som kommentarer direkte ved koderne i under punkt 2 og 3.

Rapporten er organiseret som følger: I kapitel 2 beskrives modelrammen og den praktiske operationalisering med afsæt i Jacobsen et al. (2022). Dernæst følger i kapitel 3 en beskrivelse af indsatsten to-voksen-ordning, og hvordan den implementeres i modellen, mens kapitel 4 beskriver selve beregningsmodulet. Kapitel 5 beskriver udarbejdelsen af data-inputtet til beregningsmodulet. Dette kommer fra bearbejdede data på Danmarks Statistiks forskermaskine.

Som en hjælp til læseren findes i Bilag 1 ordforklaringer af rapportens ord og begreber.

2 Modelramme

Den økonomiske model har til formål at give et estimat på den monetære nutidsværdi af udvalgte indsatser på BUVM's område inden for et 4-årigt perspektiv. Modellen er udviklet – ligesom SØM – med inspiration fra Washington State Institute for Public Policy (WSIPP)'s cost-benefit-model¹. En af styrkerne ved modellen er, at den tilbyder cost-benefit-beregninger af en række indsatser inden for den samme ramme. Det giver derved et rimeligt sammenligningsgrundlag til budgetlægning/konsekvensvurdering på tværs af indsatser.

Modellen er designet til alene at se på budgetøkonomiske konsekvenser, dvs. konsekvenser for de offentlige aktører. Der er altså ikke tale om en fuld samfundsøkonomisk model, hvor også privatøkonomiske aktører (private husholdninger og virksomheder) indgår. Derudover er modellen også på forhånd udvalgt til at alene at måle på økonomiske outcomes – altså det, der kan måles i kroner og øre. Det betyder, at modellen ikke inddrager ikke-monetære effekter på fx trivsel eller psykisk velbefindende (undtagen som eventuel mellemregning) og forsøger ikke at værdisætte disse områder.

Nedenfor følger et kort resumé af modellen fra Jacobsen et al. (2022), hvor den mere praktiske operationalisering og aggregering beskrives. Der henvises til hertil for nærmere diskussion af modellen. Afsnit 2.1 er altså udelukkende inkluderet for at vise, hvordan den praktiske operationalisering og aggregering er i forhold til den eksisterende ramme i REFUD.

Som diskuteret i ovennævnte rapport kan modellen tilbyde information om præcision i beregningerne. Som det bliver afdækket, er der en række problemstillinger forbundet hermed. Som anbefalet i rapporten håndteres usikkerhed ikke i øjeblikket.

2.1 Afsæt

Som diskuteret i Jacobsen et al. (2022) kan modellen bygges på:

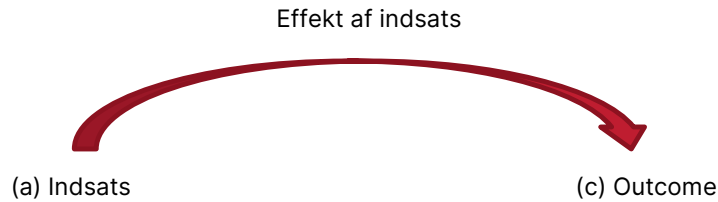
1. direkte evidens
2. indirekte evidens
3. beregnet sammenhæng

¹ Nyeste version af den tekniske manual til WSIPP findes her: <http://www.wsipp.wa.gov/TechnicalDocumentation/WsippBenefitCostTechnicalDocumentation.pdf>

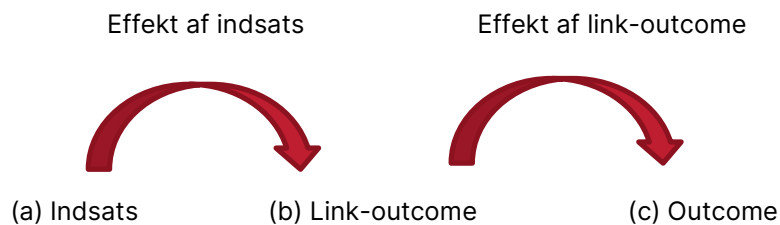
eller kombinationer heraf for forskellige outcomes². Figur 2.1 illustrerer de tre forskellige tilfælde, hvor "a" angiver indsats, "b" angiver link-outcome og "c" angiver outcome.

Figur 2.1 Modellens trin i de tre tilfælde

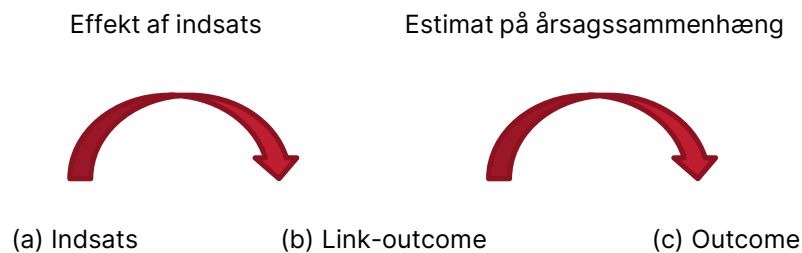
Tilfælde 1: Direkte evidens



Tilfælde 2: Indirekte evidens



Tilfælde 3: Statistisk beregnet sammenhæng



Tabel 2.1 viser modelstrukturen i de tre tilfælde med parametre. β^{ac} er effekt-estimatet mellem indsats og outcome, mens β^{ab} og β^{bc} inddrager et link-outcome. Sidste kolonne i tabellen dækker de sidste to tilfælde, da forskellen mellem de to tilgange blot er, hvorvidt effekten hentes i litteraturen eller beregnes som en del af modelarbejdet.

² Outcome og konsekvens bruges som ækvivalenter i rapporten.

Tabel 2.1 Modelstruktur ved direkte og indirekte evidens

	Effektstørrelse ved direkte evidens	Effektstørrelse ved indirekte evidens
Effekt-estimat	β^{ac}	β^{ab} og β^{bc}
Populationsgennemsnit	$Base^{ac}$	$Base^{ab}$ og $Base^{bc}$
Effektstørrelse	$f(\beta^{ac}, Base^{ac})$	$g(\beta^{ab}, Base^{ab})$ og $h(\beta^{bc}, Base^{bc})$
Total effektstørrelse (Q)	$f(\beta^{ac}, Base^{ac})$	$g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc})$
Monetær effektstørrelse	$Q(\beta^{ac}, Base^{ac}) * P$	$Q(\beta^{ab}, \beta^{bc}, Base^{ab}, Base^{bc}) * P$

Proceduren til at foretage beregninger afhænger af typen af evidens. Ved direkte evidens foretages beregningen af effektstørrelser ved:

1. Find i forskningslitteraturen estimerer på effekten af hver indsats målt ved de relevante outcome-mål
2. Find enhedspriser for outcome-mål

Mens beregningen af effektstørrelse ved brug af links foretages ved:

1. Find i forskningslitteraturen estimerer på effekten af hver indsats på link-outcome
2. Find estimerer på forholdet mellem link-outcome og de relevante outcome-mål
3. Find enhedspriser for outcome-mål.

I tilfælde, hvor det ikke er muligt at finde estimerer på forholdet mellem link-outcome og de relevante outcome-mål, er det nødvendigt at estimere β^{bc} ved beregning af effektstørrelser ved statistisk beregnet sammenhæng. Dette er dog kun muligt, i det omfang data på individer, der har modtaget indsatsen og de relevante outcomes på disse individer, er tilgængelig. Ved en model alene baseret på statistiske beregninger på data vil modellen følge sølje 2 i Tabel 2.1, blot ved at effektestimaterne er beregnet på data og ikke baseret på eksisterende evidens. En model baseret alene på beregnede statistiske sammenhænge vil som udgangspunkt ikke kunne betegnes som en kausal model, og resultater fra en sådan model vil derfor ikke kunne benyttes til at konkludere på årsagssammenhænge. Modellen kan til gengæld vise det økonomiske potentiale for de konsekvenser, som optræder i modellen, hvis en given indsats implementeres. Som det uddybes senere i rapporten, baseres REFUD-prototypen på indsatsen to-voksen-ordning, hvor der hverken er direkte eller indi-

rette evidens. Der må derfor foretages beregninger af statistiske sammen-
gænge. I tillæg hertil antages der at være fuld linearitet i modellen, hvorfor ne-
derste celle til højre i Tabel 2.1 blot reduceres til:

$$\beta^{ab} \cdot Base^{ab} \cdot \beta^{bc} \cdot Base^{bc} \cdot P \quad (1)$$

2.2 Operationalisering

Beregningsmodulet kan relativt let bygges med afsæt i (1) ovenfor. Dog er der en række forhold, som en praktisk operationalisering skal tage højde for – alle forhold, som i et vist omfang spiller sammen med, hvordan brugeren specificerer den indsats, det ønskes at regne på. Det drejer sig om diskontering, depreciering, persistens, timing og relevante beregningsvarianter, der behandles i det følgende.

2.2.1 Diskontering

I beregningerne er det nødvendigt at kunne diskontere forskellige perioder, hvorfor der både opereres med fra j og til h . Det vil sige, at det eksempelvis er muligt at diskontere over årene 2-4. Diskontering bliver derved specificeret som:

$$disk_{j,h} = \left(\frac{1}{(1+r)^j} \quad \dots \quad \frac{1}{(1+r)^h} \right)$$

hvor r er diskonteringsrenten. I modellen anvendes en diskonteringsrente på 3,5 % i overensstemmelse med Finansministeriets anbefalinger (Finansministeriet, 2021).

2.2.2 Depreciering

Med depreciering menes det forhold, at en indsats' effektivitet kan være afhængig af antallet af gentagelser af indsatsen. Det skal modellen kunne tage højde for. Mens det eksempelvis må forventes, at gentagelse af det samme læsetræningsforløb har begrænset effekt på elevernes læsefærdigheder, synes det på samme tid rimeligt, at elever drager nytte af en lav klassekvotient i alle årene. For at være så fleksibel som mulig opererer modellen med forskellige deprecieringsfaktorer for forskellige antal gentagelser. Konkret skaleres beregningerne med en d_i -faktor, hvor i er antallet af gentagelser, $d_1 = 1$ (og er derved redundant), mens d_i afhænger af evidensen bag den implementerede indsats. For to-voksen-ordninger diskuteres evidensen i afsnit 3.3.

2.2.3 Persistens

Effekten af en indsats kan med rimelighed forventes ikke at være permanent over tid. For at tage højde for dette skal modellen kunne håndtere persistens, dvs. hvor vedholdende en effekt af en given indsats er over tid. Det er muligt at implementere dette ved at måle persistens på link-outcomes eller ved at måle på konsekvenserne.

I ligning (2) vises et udsnit af (1), hvor persistens måles direkte på konsekvenserne og tilføjes som ekstra faktor β^{pers} . β^{bc} er en matrice af estimater til link-outcomes i en række estimationer med hver konsekvens som afhængig variabel. Denne matrice korrigeres derpå med et empirisk estimat på persistensen i det givne konsekvensforbrug fra år 1 til år 2, fra år 1 til år 3 og fra år 1 til år 4. Det er dog muligt at korrigere for persistens i konsekvenserne på én gang og estimere $\tilde{\beta}^{bc}$ direkte for hvert år, hvilket gør det let at implementere. Konkret sker det ved at estimere en model for hvert af alle konsekvensår 1-4³.

$$\underbrace{\beta^{bc} \cdot \beta^{pers}}_{\tilde{\beta}^{bc}} \cdot Base^{bc} \quad (2)$$

I det følgende anvendes β^{bc} blot som betegnelse for $\tilde{\beta}^{bc}$.

2.2.4 Timing

Timing vedrører det forhold, at indsatsår og kalender/budgetår ikke nødvendigvis følges ad. Det er særligt relevant, når det gælder skoleindsatser, da det kan forventes, at der sker opstart efter sommerferien. For at tage højde for, at indsatsen i sådanne tilfælde ikke har fuld virkning det i første indsatsår, skales resultaterne fra første år med en faktor t . Modellen har en tidshorisont på 4 år, hvor der specificeres t_1 til t_4 , hvor eksempelvis t_4 kommer i anvendelse det fjerde år, indsatsen pågår.

$$t_1 = (t \quad 1 \quad 1 \quad 1), t_2 = (t \quad 1 \quad 1), t_3 = (t \quad 1), t_4 = (t),$$

Defaultværdien for opstart efter sommerferien er:

$$\frac{5}{11} = 0,455$$

³ Den oprindelige tanke var at danne persistenskorrektionen β^{pers} ved at estimere én model for konsekvensår 1 og derpå estimere 3 modeller på tur; første model med konsekvens i år 2 regresseret på konsekvens i år 1 og de øvrige baggrundskarakteristika og så fremdeles.

hvilket løst svarer til 5 måneders skolegang ud af et skoleår på 11 måneder⁴.

2.2.5 Cox-transformation

Som beskrevet i Jacobsen et al. (2022) foretages der Cox-transformation af binære konsekvenser i modellen:

$$\text{Cox}(Base^{bc}, \beta^{bc}) = \frac{Base^{bc} \cdot e^{\beta^{bc} \cdot 1.65}}{1 - Base^{bc} + Base^{bc} \cdot e^{\beta^{bc} \cdot 1.65}} - Base^{bc}$$

$Base^{bc}$ er i denne sammenhæng målt i procent, og β^{bc} betegner den procentvise ændring forårsaget af indsatsen. $\text{Cox}(Base^{bc}, \beta^{bc})$ bliver for binære outcome-mål dermed ændringen i procentpoint.

Det fremgår af beskrivelserne af konsekvenserne i kapitel 5, hvilke konsekvenser der er binære. Der er tre styk i alt, og de måles alle ved "modtager foranstaltning".

2.2.6 Beregningsdele

På baggrund af ovenstående forhold og som følge af en fleksibel specificering af indsætter i beregningsmodulet optræder der en række beregningsdele i beregningsmodulet, som alle bidrager til det samlede resultat. Dette afsnit beskriver, hvordan beregningerne ser ud givet indsatsår og antal tilfælde af gentagelse. Afsnit 2.3 beskriver den samlede aggregering, hvor der tages højde for antallet af elever, der udsættes for indsatsen på et givent tidspunkt.

Alle beregningsdele er på formen $R_{x,y}$ og har dimensionerne $1 \times k$, hvor k er antallet af konsekvenser i modellen. x er indsatsåret, og y er antallet af gentagelser af indsatsen. For overblikkets skyld anvendes i flere tilfælde Hadamard-multiplikation indikeret ved "o". I modsætning til traditionel matricemultiplikation er Hadamard-multiplikation celle-til-celle-multiplikation af matricer med samme dimensioner.

Indsatsår 1

For indsatsår 1 foretages der kun én beregning, hvor afsættet i ligning (1) er udvidet med timing t_1 og diskontering for alle 4 år af konsekvenser:

⁴ For skoleåret 2021/2022 i Randers Kommune er der 90 skoledage i august-december 2021 og 200 skoledage i alt. Med det udgangspunkt vil korrektionsfaktoren til sammenligning blive $90/200 = 0,45$. <https://www.randers.dk/media/29350/feriekalender-2021-2022.pdf>.

$$\underbrace{R_{1,0}}_{1 \times k} = \underbrace{\beta^{ab}}_{1 \times 1} \cdot \underbrace{Base^{ab}}_{1 \times 1} \cdot \underbrace{t_1}_{1 \times 4} \circ \underbrace{disk_{1,4}}_{1 \times 4} \cdot \underbrace{\beta^{bc}}_{4 \times k} \circ \underbrace{Base^{bc}}_{4 \times k} \circ \underbrace{Pris}_{1 \times k}$$

Her er matricerne for baseline konsekvensforbrug, konsekvensestimater og priser givet ved:

$$\beta^{bc} = \begin{pmatrix} \beta_{a_1 k_1}^{bc} & \cdots & \beta_{a_1 k_k}^{bc} \\ \vdots & \ddots & \vdots \\ \beta_{a_4 k_1}^{bc} & \cdots & \beta_{a_4 k_k}^{bc} \end{pmatrix}$$

$$Base^{bc} = \begin{pmatrix} Base_{a_1 k_1} & \cdots & Base_{a_1 k_k} \\ \vdots & \ddots & \vdots \\ Base_{a_4 k_1} & \cdots & Base_{a_4 k_k} \end{pmatrix}$$

$$Pris = (p_{k_1} \quad \cdots \quad p_{k_k})$$

Indsatsår 2

I indsatsår 2 er der nye elever, som er blevet gamle nok til at få indsatsen. Disse elever vil bidrage med:

$$\underbrace{R_{2,0}}_{1 \times k} = \underbrace{\beta^{ab}}_{1 \times 1} \cdot \underbrace{Base^{ab}}_{1 \times 1} \cdot \underbrace{t_2}_{1 \times 3} \circ \underbrace{disk_{2,4}}_{1 \times 3} \cdot \underbrace{\beta^{bc}}_{3 \times k} \circ \underbrace{Base^{bc}}_{k \times k} \circ \underbrace{Pris}_{1 \times k}$$

Gengangere vil bidrage med:

$$\underbrace{R_{2,1}}_{1 \times k} = \underbrace{\beta^{ab}}_{1 \times 1} \cdot \underbrace{Base^{ab}}_{1 \times 1} \cdot \underbrace{d_1}_{1 \times 1} \cdot \underbrace{t_2}_{1 \times 3} \circ \underbrace{disk_{2,4}}_{1 \times 3} \cdot \underbrace{\beta^{bc}}_{3 \times k} \circ \underbrace{Base^{bc}}_{k \times k} \circ \underbrace{Pris}_{1 \times k}$$

Indsatsår 3

I indsatsår 3 er der nye elever, gengangere og 2 x gengangere. Disse elever vil henholdsvis bidrage med:

$$\underbrace{R_{3,0}}_{1 \times k} = \underbrace{\beta^{ab}}_{1 \times 1} \cdot \underbrace{Base^{ab}}_{1 \times 1} \cdot \underbrace{t_3}_{1 \times 2} \circ \underbrace{disk_{3,4}}_{2 \times 2} \cdot \underbrace{\beta^{bc}}_{2 \times k} \circ \underbrace{Base^{bc}}_{k \times k} \circ \underbrace{Pris}_{1 \times k}$$

$$\underbrace{R_{3,1}}_{1 \times k} = \underbrace{d_1}_{1 \times 1} \cdot \underbrace{R_{3,0}}_{1 \times k}$$

$$\underbrace{R_{3,2}}_{1 \times k} = \underbrace{d_2}_{1 \times 1} \cdot \underbrace{R_{3,0}}_{1 \times k}$$

Indsats år 4

Endelig vil indsatsår 4 bidrage med:

$$\underbrace{R_{4,0}}_{1 \times k} = \underbrace{\beta^{ab}}_{1 \times 1} \cdot \underbrace{Base^{ab}}_{1 \times 1} \cdot \underbrace{t_4}_{1 \times 1} \circ \underbrace{disk_{4,4}}_{1 \times 1} \cdot \underbrace{\beta^{bc}}_{1 \times k} \circ \underbrace{Base^{bc}}_{k \times k} \circ \underbrace{Pris}_{1 \times k}$$

$$\underbrace{R_{4,1}}_{1 \times k} = \underbrace{d_1}_{1 \times 1} \cdot \underbrace{R_{4,0}}_{1 \times k}$$

$$\underbrace{R_{4,2}}_{1 \times k} = \underbrace{d_2}_{1 \times 1} \cdot \underbrace{R_{4,0}}_{1 \times k}$$

$$\underbrace{R_{4,3}}_{1 \times k} = \underbrace{d_3}_{1 \times 1} \cdot \underbrace{R_{4,0}}_{1 \times k}$$

2.3 Aggregering

Beregningsmodulen skal naturligvis tage højde for antallet af elever, der udsættes for en given indsats på et givent tidspunkt. Med afsæt i et eksempel udfoldes beregningen af aggregeringen nedenfor.

I Tabel 2.2 vises et eksempel, hvor en indsats gives i 4.-7. klassetrin gennem 4 år. I første indsatsår vil ingen elever naturligvis være gengangere. R-værdierne er dog afhængige af klassetrin, hvorfor R-værdien ikke vil være den samme for alle klassetrin i indsatsåret. Klassetrinnet er indikeret i parentes. I Indsatsår 4 kommer alle værdierne $R_{4,0}() - R_{4,3}()$ i spil som følge af det varierende antal gengangere.

Tabel 2.2 Eksempel på indsats og beregningsdele

Klassetrin	Indsatsår 1	Indsatsår 2	Indsatsår 3	Indsatsår 4
4	$R_{1,0}(4)$	$R_{2,0}(4)$	$R_{3,0}(4)$	$R_{4,0}(4)$
5	$R_{1,0}(5)$	$R_{2,1}(5)$	$R_{3,1}(5)$	$R_{4,1}(5)$
6	$R_{1,0}(6)$	$R_{2,1}(6)$	$R_{3,2}(6)$	$R_{4,2}(6)$
7	$R_{1,0}(7)$	$R_{2,1}(7)$	$R_{3,2}(7)$	$R_{4,3}(7)$

Antallet af elever benævnes $e_{y,c}$ hvor y er indsatsår, og c er klassetrin⁵. Derved kan aggregeringen foretages. For at bevare overblikket foretages aggregeringen ét indsatsår ad gangen i det følgende, og de lægges sammen til sidst.

Den generelle aggregering i indsatsår 1 vil være:

⁵ | beregningsmodulen er udregningen forenklet til at lade $\tilde{e}_{y,c} = \frac{\sum_{c=\min(kl)}^{\max(kl)} e_{y,c}}{\max(kl) - \min(kl)}$, hvilket vil sige, at eleverne er ligeligt fordelt mellem klassetrin.

$$RT_1 = \sum_{c=\min(kl.)}^{\max(kl.)} R_{1,0}(c)e_{1,c}$$

hvor $kl.$ indikerer klassetrin. For de øvrige indsatsår bliver aggregeringen henholdsvis

$$RT_2 = R_{2,0}(\min(kl.))e_{2,\min(kl.)} + \sum_{c=\min(kl.)+1}^{\max(kl.)} R_{2,1}(i)e_{2,c}$$

$$RT_3 = R_{3,0}(\min(kl.))e_{3,\min(kl.)} + R_{3,1}(\min(kl.) + 1)e_{3,\min(kl.)+1} + \sum_{c=\min(kl.)+2}^{\max(kl.)} R_{3,2}(c)e_{3,c}$$

og

$$RT_4 = R_{4,0}(\min(kl.))e_{4,\min(kl.)} + R_{4,1}(\min(kl.) + 1)e_{4,\min(kl.)+1} \\ + R_{4,2}(\min(kl.) + 2)e_{4,\min(kl.)+2} + \sum_{c=\min(kl.)+3}^{\max(kl.)} R_{4,3}(c)e_{4,c}$$

Det samlede resultat bliver derved:

$$RT = RT_1 + RT_2 + RT_3 + RT_4$$

3 Indsatsen to-voksen-ordninger

BUVM ønskede prototypen udviklet med afsæt i en enkelt indsats. I dialog med VIVE faldt valget på indsatsen to-voksen-ordninger. Jf. Jacobsen et al. (2022) kan Andersen et al. (2020) anvendes som empirisk afsæt, og det blev aftalt med BUVM, at prototypen blev udviklet med dette som empirisk grundlag. Dette kapitel beskriver nærmere, hvordan studiets forskellige data og resultater er blevet omsat, og hvordan indsatsen er blevet konkretiseret til at kunne indgå som parametre i beregningerne i modellen.

Andersen et al. (2020) evaluerer en to-voksen-ordning, som blev implementeret i Danmark i 2012-2013. Indsatsen blev rullet ud som et lodtrækningsforsøg, hvor 6. klasser på 105 skoler fik en to-voksen-indsats. Der er blev benyttet to typer af indsats: Den ene indsats bestod i, at klassen havde to læreruddannede i klassen gennemsnitligt 11,6 lektioner om ugen i 85 % af skoleåret, mens den anden indsats bestod i, at klassen havde en læreruddannet og en ikke-læreruddannet i klassen gennemsnitligt 20,7 lektioner om ugen i 85 % af skoleåret. Studiet identificerer statistisk signifikante effekter på resultater af læsetests, men ingen signifikante resultater for matematiktests.

3.1 Indsatsvarianter

Som beskrevet har Andersen et al. (2020) evalueret to beslægtede, men alligevel forskellige varianter af indsatsen to-voksen-ordning. I dialog med BUVM er det derfor besluttet at præsentere brugeren for to forskellige indsatsvarianter:

- En ekstra ikke-læreruddannet
- En ekstra læreruddannet.

Årsagen hertil er, at Andersen et al. (2020) finder forskellige effekter af de to indsatsvarianter, hvilket der naturligvis skal tages højde for i beregningerne.

Det er ligeledes vigtigt at specificere indsatsvariant, hvis prototypen på et senere tidspunkt skal udvides med et omkostningsmodul, da de to indsatsvarianter ikke nødvendigvis vil koste det samme.

3.2 Helårseffekter

For at lette beregningerne i modellen foretages en omregning af resultaterne i Andersen et al. (2020), at således de korresponderer med det tilfælde, hvor indsatsen er implementeret hele året. Omregnet til helårseffekter betyder det,

at *effect size* ved fuld implementering af indsatsvarianten ”en ekstra læreruddannet” er 0,244, mens *effect size* ved fuld implementering af indsatsvarianten ”en ekstra ikke-læreruddannet” er 0,209 – under antagelse af, at indsats-effekten er lineær længden af indsatsen på et skoleår. Tabel 3.1 opsummerer denne beregning for begge indsatsvarianter.

I beregningsmodulet er det muligt at specificere indsatsens intensive margin (implementeringsgrad). Det bliver for brugeren derved muligt at regne på indsatser, der kun implementeres delvist, som det var tilfældet i Andersen et al. (2020). Det vil eksempelvis korrespondere til det tilfælde, hvor to-voksen-ordningen alene implementeres i visse fag eller på visse dage.

Tabel 3.1 Beregning af helårseffekter

Indsats	Indsatsvariant	(1) Rappor- teret effekt- stør- relse (dansk)	(2) Gns. klokke- timer pr. år i 6. klasse ³	(3) Gns. klokke- timer pr. uge ⁴	(4) Lektio- ner pr. uge med indsats ⁵	(5) Klokke- timer pr. uge med indsats ⁶	(6) Andel af år med indsats ⁵	(7) Omreg- net hel- årseffekt ⁷
To-vok- sen ord- ning	En ekstra lærer- uddannet	0,086 ¹	841	21,01	11,6	8,7	85%	0,244
To-vok- sen ord- ning	En ekstra ikke- læreruddannet	0,131 ²	841	21,01	20,7	15,5	85%	0,209

- Note: 1 Andersen et al. (2020), tabel 4, kolonne 1, række 1
2 Andersen et al. (2020), tabel 4, kolonne 1, række 2
3 Jf. uddannelsesstatistik.dk, Planlagte timer - Skoleårets længde, 2012/2013, 6. klasse, vægtet gennemsnit
4 Et skoleår består af 200 dage/40 uger
5 Jf. Andersen et al. (2020)
6 En lektion er 0,75 klokketimer
7 Helårseffekten udregnes ved: (7) = (1) · (3) / (5) / (6).

3.3 Målgruppe og depreciering

Prototypens evidensgrundlag består som bekendt blot af et enkelt studie. Ud fra et forsigtighedsprincip kan der på den ene side derfor argumenteres for, at modellen alene skal kunne regne på situationer, der mere eller mindre 1:1 matcher det empiriske grundlag. Det vil sige, at modellen eksempelvis kun ville tilbyde brugeren at regne på indsatser af to-voksen-ordningen implementeret i 6. klasse i samme omfang som i Andersen et al. (2020). På den anden side gør

det modellen mindre relevant i praksis, da den ikke nødvendigvis vil være ramme for den situation, brugeren kan tænkes at stå i.

Tanken bag modellen er at bløde op på forsigtighedsprincippet og tillade, at brugeren anvender modellen til at regne på økonomiske konsekvenser af indsatser, som til en vis grad afviger fra det empiriske grundlag. Det kunne eksempelvis være at regne på effekten af to-voksen-ordninger i 5. klasse implementeret i halvdelen af skoleåret. Dette vil afvige fra det empiriske grundlag, hvilket brugeren gøres bekendt med. Alle parametre skal desuden kunne overskrives af brugeren, således at der er mulighed for at foretage korrektioner baseret på en konkret vurdering fra brugerens side.

Som beskrevet i Andersen et al. (2020) var målgruppen for lodtrækningsforsøget 6. klasse. Ud fra dette empiriske grundlag og ovenstående afvejning vurderer VIVE, modellens beregninger på indsatser af to-voksen-ordninger vil være mest retvisende for elever på mellemtrinnet, dvs. 4., 5. og 6. klasse⁶. I mangel af empiri og for at imødekomme brugere, der måtte ønske at foretage beregninger på folkeskolens øvrige klassetrin, tillader modellen dog også beregninger på ind- og udskoling (henholdsvis 0.-3. og 7.-9. klasse).

En bruger kan ønske at foretage beregninger på implementering af to-voksen-ordninger over flere år. Det vil sige en situation, hvor de samme elever udsættes for indsatsen gentagne gange over flere klassetrin. I det tilfælde er det nødvendigt at forholde sig til, hvilken effekt indsatsen vil have, når den modtages gentagne gange.

Andersen et al. (2020) belyser ikke denne situation, og der er dermed ingen evidens for, hvor virkningsfuld indsatsen vil være, hvis den bliver gentaget. Der findes VIVE bekendt ikke andre studier, der giver denne evidens. Ud fra et princip om konsistens i behandlingen af effekt-estimer i modellen regner modellen i udgangspunktet med ens effektstørrelser i første og eventuelle efterfølgende indsatsår. Denne lineære ekstrapolation af effektstørrelser over år er således alene en antagelse og ikke empirisk funderet i Andersen et al. (2020).

Hvis brugeren ud fra konkret viden mener, at denne restriktion ikke er retvisende, er det som ved alle andre parametre i modellen muligt at justere deprecieringen. Det er muligt at justere i begge retninger, således en indsats kan have stigende marginaleffekt over indsatsår.

⁶ Der kunne argumenteres for, at modellen skulle kunne lade brugeren regne på indsatser for 4.-8. klasse ud fra et symmetri-argument. Begrundelsen for kun at tilbyde brugeren at regne på 4.-6. klasse er, at der er tanker om at udvide nuværende prototype med et udvælgelses- og vægtningsmodul til tilfælde med flere indsatsrelevante resultater fra forskningslitteraturen. I dette kommende modul forventes netop at benytte inddelingen med indskoling, mellemtrin og udskoling. Der tages derfor højde for denne forventede inddeling i denne version af prototypen.

4 Beregningsmodul

Beregningsmodul er helt centralt i REFUD-prototypen og har til formål at omsætte data fra Danmarks Statistiks forskermaskine, evidens fra forskningslitteraturen og brugerinput til konsekvensberegninger for en given indsats. Til det formål har VIVE udviklet et beregningsmodul i R med pakken Shiny. Afsættet for udviklingen er indsatsen to-voksen-ordninger med Andersen et al. (2020) som empirisk grundlag. Beregningsmodul er dog udviklet til at være generisk, således at nye data, ny evidens og nye indsatser relativt let kan tilføjes.

Beregningsmodul er blevet testet og virker (som minimum) i RStudio Desktop 2021.09.0+351 (RStudio Team, 2021) med R version 4.1.2 ("Bird Hippie") (R Core Team, 2021). Beregningsmodul er baseret på Shiny 1.7.1 (Chang et al., 2021). Figur 4.1 giver en komplet oversigt over R-pakker. Beregningsmodul versionstyres via Git/GitHub og distribueres som bundle-fil, således at historikken er tilgængelig.

VIVE forestår alene arbejdet med at udvikle en prototype på et beregningsmodul, mens selve brugergrænsefladen udvikles af STIL/Netcompany. Den valgte R Shiny-løsning leverer dog automatisk en webbaseret brugergrænseflade. For at understrege, at denne brugergrænseflade ikke som sådan er en del af beregningsmodul, kaldes brugergrænsefladen i prototypen et *UI-værktøj* (User Interface). Uagtet dets status har dette UI-værktøj været et nyttigt dialogredskab i forbindelse med præsentation og diskussion af beregningsmodulets funktionalitet. I afsnit 4.2 har værktøjet netop denne funktion, og skærbilleder af UI-værktøjet bruges til at beskrive brugerens interaktion med beregningsmodul og selve brugerrejsen.

4.1 Programstruktur og filbeskrivelser

Beregningsmodul består rent teknisk af en række forskellige filer. Dette afsnit beskriver programstrukturen og de enkelte filer. Beregningsmodul ligger i mappen `prototype` og er dermed også navnet på beregningsmodul. De to centrale filer er `server.R` og `ui.R`, mens henholdsvis data og parametre ligger i mapperne af samme navn. Programmet kan eksekveres fra RStudio, når enten `server.R` eller `ui.R` er åbne i RStudio.

Figur 4.1 giver et overblik over programstrukturen.

Figur 4.1 Programstruktur

```
Prototype
|   server.R
|   ui.R
|
+---data
|   beta_konsekvenser1.csv
|   beta_konsekvenser2.csv
|   konsekvensforbrug.csv
|   links1.csv
|   links2.csv
|   priser.csv
|
+---parametre
|   indsatser.json
|   indstillinger.json
|   kommuner.json
|   konsekvenser.json
|   links.json
```

Anm.: Prototypens mappestruktur og filer.

4.1.1 Server.R

`Server.R` udgør serversiden i beregningsmodulet og står for bearbejdning af alle data og brugerinputs. `Server.R` implementerer modellen som beskrevet i kapitel 2, og kodestrukturen i filen er fortløbende som brugerrejsen, som beskrives nærmere afsnit 4.2. Koden er kommenteret i selve `server.R`, hvorfor meget kodespecifikke kommentarer er overladt til selve filen. Nedenfor følger en introduktion til de mest centrale byggesten i R Shiny-koden.

I koden interageres med brugeren gennem to funktioner – `input()` og `output()`, som henholdsvis henter og sender til UI. For at referere til specifikke inputs eller output anvendes unikke id'er angivet på følgende vis:

- `input$<id>`
- `output$<id>`

Et princip i beregningsmodulet er, at alle parametre er fleksible og kan styres og ændres af brugeren. Det har den konsekvens, at alle inputs og outputs er variable størrelser. I nogle tilfælde er det kun brugeren, som kender den relevante oplysning. Det kunne eksempelvis være varigheden af en indsats. I andre tilfælde tilbydes brugeren en parameter af modellen, som brugeren kan forholde sig til og manipulere med, hvis det ønskes. Det kunne eksempelvis være den anvendte effektstørrelse.

Generelt skal modellen tage højde for, at inputs og outputs kan ændre sig. Det gøres ved funktionerne `reactive()` og `renderUI()`, hvor `reactive()` angiver, at en funktion er afhængig af brugerinputs og reagerer derpå, mens `renderUI()` genererer dynamisk output baseret på brugerens inputs. Figur 4.2 viser eksempler på, hvordan netop `reactive()` og `renderUI()` bruges i koden.

Figur 4.2 **Reactive() og renderUI()**

```
# Eksempel på reactive() med UI input
antal_klasser <- reactive({
  input$antal_klasser
})

# Eksempel på reactive() med både UI input og reactive input
indsats_beta <- reactive({
  req(indsats_grundlag())
  if (indsats_grundlag()==1) {
    indsats_beta_h()%>%as.matrix()
  } else if (indsats_grundlag()==2) {
    indsats_beta_j()%>%as.matrix()
  }
})

# Eksempel på renderUI() med reactive input
output$menu_indsats <- renderUI({
  selectInput("indsats", "Vælg indsats", c(Vælg='', indsats_liste()))
})
```

Anm.: Eksempler på centrale funktioner i `server.R`.

Som figuren viser, anvendes ligeledes funktionen `req()`, som sikrer, at en given funktion er evalueret, inden den aktuelle funktion evalueres. I eksemplet evalueres `indsats_beta()` (den anvendte effektstørrelse) ikke, hvis ikke brugeren har taget stilling til, på hvilket grundlag effektstørrelsen skal beregnes (`indsats_grundlag()`). Det sker tidligere i brugerrejsen (og i øvrigt automatisk, da brugeren som default bliver tilbudt helårseffekten).

4.1.2 ui.R

`ui.R` bestemmer udseendet af beregningsmodulets webbaserede brugergrænseflade. Sammenlignet med `server.R` er filen af beskeden omfang og består primært af *placeholders* til præsentation af servergenereret output. Indbygget i R Shiny er CSS-frameworket Bootstrap, som sikrer, at UI-værktøjet fremtræder responsivt. Desuden er det muligt at anvende en tema-tem-

plate til styling af brugergrænsefladen. Her er Bootstrap-temaet Flatly anvendt. Dertil er tilføjet en smule custom-CSS, der er med til at sikre et overskueligt layout på diverse tabeller.

Selve brugergrænsefladen er ikke central i VIVEs leverance, da det ikke er en del af opdraget. Af denne årsag vil `ui.R` ikke blive yderligere beskrevet her. Brugerens interaktion med beregningsmodulet udfoldes dog yderligere i afsnit 4.2.

4.1.3 Data

I sin enkelthed omsætter beregningsmodulet brugerinputs og data fra Danmarks Statistiks forskermaskine til økonomiske konsekvenser. Data fra forskermaskinen indlæses i beregningsmodulet fra en række csv-filer som angivet i Tabel 4.1. Der er et datasæt på priser og et datasæt på konsekvensforbrug, mens antallet af datasæt benævnt "beta_konsekvenser[nr].csv" og "links[nr].csv" afhænger af antallet af links i modellen. I øjeblikket er der dansk og matematik fra de nationale test, hvorfor det samlede antal links er 2.

For yderligere beskrivelser af de enkelte datasæt henvises til kapitel 5.

Tabel 4.1 Oversigt over beregningsmodulets datakilder

Navn på datainput	Indhold
beta_konsekvenser1.csv	Konsekvensestimater for link-outcome 1 opgjort på kommune og klasseniveau
beta_konsekvenser2.csv	Konsekvensestimater for link-outcome 2 opgjort på kommune og klasseniveau
konsekvensforbrug.csv	Baseline konsekvensforbrug for hver af modellens fire perioder, opgjort på kommune og klasseniveau
links1.csv	Baseline for link-outcome 1 opgjort på kommune og klasseniveau
links2.csv	Baseline for link-outcome 1 opgjort på kommune og klasseniveau
priser.csv	Priser for hver konsekvens opdelt på stat, region og kommune

4.1.4 Parametre

Beregningsmodulet er ønsket udviklet med størst mulig fleksibilitet. Konkret betyder det, at modellen skal kunne udvides og ændres, uden at selve koden skal genbesøges og redigeres. Til dette formål anvendes json-filer (JavaScript

Object Notation) til modellens inputs, mens data som bekendt indlæses fra cvs-filer. Som Tabel 4.2 angiver, indgår modellens indsatser, indstillinger og kommuner samt registre over konsekvenser og links som json-filer.

Tabel 4.2 Oversigt over beregningsmodulets parameterinputs

Navn på parameterinput	Indhold
indsatser.json	Modellens indsatser
indstillinger.json	Modellens standardindstillinger
kommuner.json	Kommunenavn og kommunekoder
konsekvenser.json	Register over alle konsekvenser i modellen
links.json	Register over alle link-data i modellen

Hvis Finansministeriet eksempelvis ændrer diskonteringsrenten til 4 %, vil modellen let kunne tilpasses ved blot af opdatere `indstillinger.json`. Nedenfor følger en nærmere beskrivelse af, hvordan indsatser og konsekvenser specificeres i modellen via json-filerne. Her er lidt flere parametre i spil sammenlignet med justering af indstillingerne.

Figur 4.3 viser et udsnit af `indsatser.json` med én variant af to-voksen-ordningen. Uafhængigt af indsatsvariant specificeres:

- Indsatsens nummer (`indsats_nr`): Indsatserne nummereres fortløbende, uden at nummeret bærer nogen betydning i sig selv.
- Indsatsens navn (`indsats_navn`): Indsatsens navn, der præsenteres for brugeren i UI.
- Kilde (`kilde`): Indsatsens kilde. Hvis det drejer sig om en effektstørrelse fundet ved sammenvejning af flere effektstørrelser, angives dette.
Beskrivelse (`beskrivelse`): Beskrivelse af kilden. Hvis der blot er tale om et studie, kan studiets abstract eksempelvis bruges.
Indsatstype (`indsatstype`): 1 = Skoleindsats, 2 = Dagtilbudsindsats.
- Links (`links`): Array med relevante links for indsatsen.
- Konsekvenser (`default_konsekvenser`): Array med relevante konsekvenser for indsatsen.

Dernæst specificeres indsatsens varianter. Hvis der ikke er nogle varianter, specificeres indsatsen blot som én variant. For hver variant specificeres:

- Variantens nummer (`variant_nr`): Varianterne nummereres fortløbende, uden at nummeret bærer nogen betydning i sig selv.

- Variantens navn (`variant_navn`): Variantens navn, der præsenteres for brugeren i UI.
- Rapporterede effekter (`beta_r`): Array af rapporterede effekter.
- Omregnede helårseffekter (`beta_h`): Array af omregnede helårseffekter.
- Målgruppe (`maalgruppe`): Målgruppe for indsatsen. Som det er specificeres, kommer venstresiden af ":" til at fremstå som valgmuligheden i UI (fx "4. klasse"), mens højresiden er selve værdien (fx 4). Default i modellen er 4.-6. klassetrin, men der er tilføjet en fuldt fleksibel indsatsvariant, der tillader beregning på samtlige (0.-9.) klassetrin.
- Inputs til beregning af intensiv margin (`intensiv_margin*`): Default beregning af intensiv margin af indsatsens intensitet givet nogle valgte kriterier. I eksemplet er valgt intensiteten i studiet
 - `_t`: Værdi af tæller
 - `_t_besk`: Beskrivelse af tæller i UI
 - `_n`: Værdi af nævner
 - `_n_besk`: Beskrivelse af nævner i UI
 - `_aar`: Andel af året
 - `_aar_besk`: Beskrivelse af andel af året i UI.
- Depreciering (`depreciering_aar1-depreciering_aar4`): Depreciering af indsatsen ved gentagelse. I dette tilfælde antages ens indsatseffekter ved gentagen indsats, og der anvendes 0 % depreciering efter første år.

Figur 4.3 Uddrag fra indsatser.json

```
"indsats_nr": 1,
"indsats_navn": "To-voksen ordning",
"kilde": "Andersen et al. (2018)",
"beskrivelse": "[udeladt her]",
"indsatstype": 1,
"links": [1, 2],
"default_konsekvsr":
,2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 37],
"variant":
  [
    {
      "variant_nr": "1",
      "variant_navn": "En ekstra læreruddannet",
      "beta_r": [0.086, 0],
      "beta_h": [0.244, 0],
      "maalgruppe":
        [
          {
            "4. klasse": 4,
            "5. klasse": 5,
            "6. klasse": 6
          }
        ],
      "inputs":
        [
          {
            "intensiv margin t": 8.7,
            "intensiv_margin_t_besk": "Antal timer med to-voksenordning pr.
",
            "intensiv margin n": 21.01,
            "intensiv_margin_n_besk": "Gns. antal timer i skoleåret pr. uge",
            "intensiv margin aar": 85,
            "intensiv_margin_aar_besk": "Andel af skoleåret med to-
ksenordning"
          }
        ],
      "depreciering":
        [
          {
            "depreciering_aar1": 0,
            "depreciering_aar2": 0,
            "depreciering_aar3": 0,
            "depreciering_aar4": 0
          }
        ]
    }
  ]
...

```

Anm.: Uddrag af filen indsatser.json.

Figur 4.4 viser et udsnit af konsekvenser.json. For hver konsekvens specificeres fire parametre:

- Konsekvensnummer (*konsekvens_nr*): Konsekvenser nummereres fortløbende, uden at nummeret bærer nogen betydning i sig selv.
- Konsekvensdimension (*konsekvens_dimension*): Den overordnede type af konsekvens. Anvendes p.t. ikke, men er til rådighed til det kommende UI.
- Konsekvensnavn (*konsekvens_navn*): Navnet på konsekvensen, der præsenteres for brugeren.
- Enhed (*enhed*): Enheden, som konsekvensen er mål i. Anvendes p.t. ikke, men er til rådighed til det kommende UI.

Figur 4.4 Uddrag fra konsekvenser.json

```
[{
  "konsekvenser":
    [
      {
        "konsekvens_nr": 1,
        "konsekvens_dimension": "Grundskole",
        "konsekvens_navn": "Specialskole",
        "Enhed": "Kr. pr. uge"
      },
      {
        "konsekvens_nr": 2,
        "konsekvens_dimension": "Grundskole",
        "konsekvens_navn": "Grundskole 0.-9. klasse, almen kommunal",
        "enhed": "Kr. pr. uge"
      },
      {
        "konsekvens_nr": 3,
        "konsekvens_dimension": "Grundskole",
        "konsekvens_navn": "Grundskole 10. klasse, almen kommunal",
        "Enhed": "Kr. pr. uge"
      },
      ...
    ]
}]
```

Anm.: Uddrag af filen konsekvenser.json.

4.2 UI-værktøj og brugerrejse

Dette afsnit præsenterer UI-værktøjet og brugerens interaktion med beregningsmodul gennem brugerrejsen.

UI-værktøjet er dynamisk, da brugerens valg påvirker, hvilke øvrige valg brugeren efterfølgende kan foretage. Når en beregning er gennemført, er det muligt at gå tilbage og rette forskellige valg og værdier. I den situation er det dog vigtigt at genberegne resultaterne, da de ellers kan være misvisende. Da UI-værktøjet ikke som sådan er en del af leverancen af beregningsmodul, er der ikke indbygget check af formularindtastninger mv., da værktøjet kun er til internt brug, hvor vi selv har ansvaret for at indtaste meningsfyldte oplysninger.

På alle sider er det muligt at navigere frem og tilbage, ligesom det er muligt at nulstille alt. Øverst er det muligt at navigere rundt i hele beregningsmodul. Dog vil flere sider fremstå tomme, da de kræver, at der er foretaget valg.

4.2.1 Forside

Det første, brugeren møder, er forsiden vist i Figur 4.5. Her skal brugeren i menuen til venstre vælge

- Indsatstype (dagtilbudsindsats eller skoleindsats)
- Kommune (en specifik kommune eller en gennemsnitlig kommune).



4.2.2 Indsats

Det næste, brugeren skal forholde sig til, er selv indsatsen – dels hvilken indsats og indsatsvariant der skal regnes på, dels i hvilket omfang indsatsen implementeres. Specifikt skal brugeren forholde sig til:

1. Indsats.
2. Linkdata (dansk eller matematik).
3. Indsatsvariant.
4. Klassestrin (givet mulige klassestrin, jf. specifikationen af indsatsen).
5. Antal klasser pr. år med indsats.
6. Tidspunkt for opstart af indsats (efter nytår eller efter sommerferien).
7. Varigheden af indsatsen (1 år, 2 år, 3 år eller 4 år/permanent).

Elever udsat for indsatsen for hvert år givet den valgte varighed. Brugeren hjælpes på vej her, da der i udgangspunktet regnes med, at den gennemsnitlige klassestørrelse er 22, hvorfor feltet er udfyldt på forhånd, når først antallet af klasser er udfyldt. Klassestørrelsesangivelsen har udelukkende til formål at beregne, hvor mange elever der får indsatsen, til aggregering af gevinster på individniveau og kan ikke benyttes til at justere indsats effekten for den enkelte elev. Indsats effekten pr. elev af en to-voksen-ordning er i modelberegningerne således den samme ved en klassestørrelse på 2 som ved en klassestørrelse på 28.

Derudover kan brugeren rette beregningsgrundlaget, når det gælder henholdsvis effekt og depreciering. Enten kan de givne værdier bruges, eller også kan brugeren justere henholdsvis effekt og depreciering.

4.2.3 Links

I næste skridt får brugeren oplyst det lokale baseline af link-outcome for hvert klassetrin, givet valgt kommune og valgt link-outcome. Med lokal baseline forstås lokal standardafvigelse på link-outcome. Da modellen regner med effect size som inputparameter for en indsats' effekt, er enheden for beregningerne netop standardafvigelser. En effektstørrelse på eksempelvis 0,20 betyder, at eleven klarer sig 0,20 standardafvigelser bedre end forventet. I gennemsnit vil standardafvigelsen være én, men lokalt kan en standardafvigelse være henholdsvis større eller mindre end én. Det skal der korrigeres for i beregningerne.

Hvis brugeren ønsker at justere det anvendte lokale baseline-niveau, er det muligt, hvorved det justerede niveau anvendes som beregningsgrundlag. I praksis forventes det dog at være sjældent, at en bruger ønsker at korrigere, men situationen kan opstå, hvis der eksempelvis regnes på en indsats på en skole, som systematisk afviger fra kommunens andre skoler, når det gælder standardafvigelser på link-outcome.

4.2.4 Konsekvenser

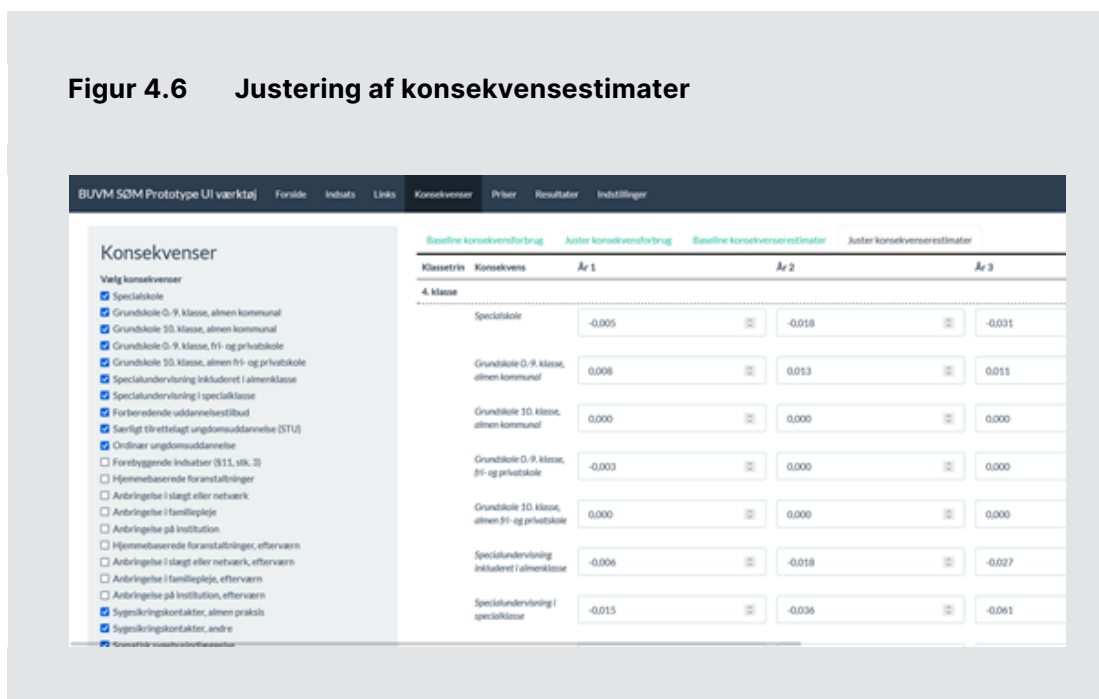
Brugeren bliver dernæst præsenteret for en lang række konsekvenser under siden "Konsekvenser". Konsekvenserne er oplyst i menuen til venstre som vist i Figur 4.6. Til hver konsekvens er en checkboks. Hvis en given konsekvens er blandt defaultkonsekvenserne, er der sat flueben i checkboksen på forhånd. Kriterierne for defaultkonsekvenser diskuteres indgående i kapitel 5.

I højre side kan brugeren klikke sig rundt blandt fire forskellige faner, som viser tabeller over henholdsvis:

- Baseline konsekvensforbrug
- Juster konsekvensforbrug
- Baseline konsekvensestimater
- Juster konsekvensestimater.

Der vises kun tal for de konsekvenser, som er valgt i menuen til venstre. Igen kan alt justeres, og Figur 4.6 illustrerer netop, hvordan brugeren eksempelvis selv kan justere konsekvensestimaterne i modellen. Hvis det justerede konsekvensforbrug og de justerede konsekvensestimater skal anvendes som beregningsgrundlag, skal det angives nederst i menuen til venstre.

Figur 4.6 Justering af konsekvensestimater



Anm.: Figuren illustrerer, hvordan brugeren eksempelvis selv kan justere konsekvensestimaterne i modellen.

4.2.5 Priser

Næste skridt er modellens priser. Brugeren præsenteres udelukkende for priser for de konsekvenser, som brugeren valgte i forrige skridt. Hvis brugeren ønsker at anvende andre priser, er det som altid muligt at justere.

4.2.6 Resultater

Sluttelig når brugeren frem til siden "Resultater", hvor brugeren selv skal initiere den endelige beregning ved tryk på en knap. Årsagen til, at beregningen ikke sker automatisk, er for at undgå for lange load-tider.

Når resultaterne er beregnet, kan brugeren se:

- Samlet resultat (aggregeret for hvert klassetrin)
- Opdelt resultat (stat, region og kommune)
- Beregningskomponenter.

Beregningskomponenterne viser matricer med alle beregningsmodulets delresultater. Det forventes ikke, at dette skal vise den endelige slutbruger i det kommende UI. Til gengæld gør beregningselementerne det nemmere at følge udregningerne i prototypen.

Hvis der mangler oplysninger til at foretage en beregning for en given konsekvens i et givent klassetrin, vil resultatet i det tilfælde blot vise 0. En sådan situation kunne eksempelvis ske ved, at estimater er insignifikante eller ved datadiskretionering som følge af for få observationer.

4.2.7 Indstillinger

Under siden "Indstillinger" er det muligt at justere:

- Diskonteringsrente
- Justeringsfaktor for opstart efter sommerferien.

Standardindstillingerne er automatisk indlæst fra `indstillinger.json` med værdier som beskrevet i afsnit 2.2.

4.3 Nye indsatser, nye data og ny evidens

Beregningsmodulet er kodet generisk, så det er muligt let at udvide med nye data, nye indsatser og ny evidens. Dette afsnit giver en beskrivelse af, hvordan det gøres gennem en række eksempler. Eksemplerne illustrerer hver enkelt tilføjelse, men tilføjelserne kan fint ske i flere dimensioner på samme tid.

Eksempel: Ny indsats eller ny indsatsvariant

En ny indsats eller indsatsvariant tilføjes blot `indsatser.json` for alle relevante parametre beskrevet i afsnit 4.1.4.

Eksempel: Nyt link-outcome

Et nyt link-outcome tilføjes ved:

1. at opdatere registeret `links.json`
2. at tilføje ny datasæt med baseline niveau `links[nr.].csv`
3. at tilføje link-outcomet samt relevante estimater til én eller flere indsatser `indsatser.json`.

Eksempel: Nye konsekvenser

Ny konsekvenser tilføjes ved:

1. at opdatere registeret `konsekvenser.json`
2. at opdatere `priser.csv` med priserne på de nye konsekvenser
3. at opdatere `konsekvensforbrug.csv` med baseline konsekvensforbrug
4. at opdatere `beta_konsekvenser[1-antal links].csv` med nye β^{bc} -estimater for de nye konsekvenser
5. at opdatere eksisterende indsatser i `indsatser.json`, for hvilke de nye konsekvenser skal være default.

Eksempel: Ny evidens til eksisterende indsats

Prototypen er bygget med afsæt i Andersen et al. (2020), hvor der hentes evidens på effekten af to-voksen-ordning på de nationale test i dansk og matematik. Derpå findes en statistisk beregnet sammenhæng til modellens konsekvenser. Som beskrevet i afsnit 2.1 er det nødvendigt, at modellen også kan håndtere direkte evidens og indirekte evidens, da sådan evidens kan opstå i litteraturen i fremtiden for indsatsen to-voksen-ordning eller allerede findes for andre indsatser. Beregningsmodulet har mulighed for at beregne sådanne tilfælde – omend en smule håndholdt i den nuværende version⁷.

Indirekte evidens (β^{bc} fra litteraturen)

Indirekte evidens kan i beregningsmodulet implementeres ved manuelt at rette β^{bc} under konsekvensestimaterne for de relevante konsekvenser.

Direkte evidens (β^{ac} fra litteraturen)

Direkte evidens på en eller flere konsekvenser kan i beregningsmodulet implementeres ved manuelt at skalere β^{bc} med α_{k_i} således at:

⁷ Det er på et senere tidspunkt på tegnebrættet at udvide beregningsmodulet med en prioriteringsmekanisme som beskrevet i Jacobsen et al. (2022). I et sådant tilfælde vil håndteringen af direkte og indirekte evidens kunne kodes til at kunne beregnes mere dynamisk og fleksibelt.

$$\beta^{ac} = \alpha_k \beta^{bc}.$$

hvor α_k varierer over konsekvenserne med direkte evidens. I det tilfælde, hvor β^{ac} ikke yder estimer over hele tidshorizonten af interesse (1-4 år), er det nødvendigt at korrigere med persistens i konsekvenserne.

5 Beregninger på Danmarks Statistiks forskermaskine

Leverancen af inputs til REFUD-prototypen baseres på de detaljerede registre, der er tilgængelige på Danmarks Statistiks forskermaskine. Med udgangspunkt i den i afdækningsrapporten opstillede model (opsummeret i afsnit 2.1) omhandler beregningerne på forskermaskinen tre konkrete output-dimensioner:

1. Lokale baseline standardafvigelser på link-outcomes ($Base^{ab}$)
2. Betingede korrelationsmatricer for sammenhængen mellem link-outcome og konsekvenser rensset for individspecifik baggrundskarakteristika i en op til 4-årig horisont (β^{bc})
3. Lokale baseline standardafvigelser på modellens konsekvenser ($Base^{bc}$)

Herudover er enkelte priser (P) beregnet ved hjælp af registerdata.

At baseline standardafvigelserne på både link-outcome og endelige konsekvenser måles lokalt på kommunalt niveau, giver således brugeren et mere retvisende billede af de forventede nettogevinster ved modellens indsatser for den enkelte kommune. Herudover er der i modellen tilføjet mulighed for beregning af indsatsers nettogevinster baseret på et landsgennemsnit. Dette er alene tænkt som en rettesnor for brugere, der ikke har én konkret kommune i tankerne, eller som er i tvivl om generaliserbarheden af modellens estimater, hvis de kommunespecifikke karakteristika ændres. Indsatseffekterne på landsplan skal således fortolkes med forsigtighed, da modellen ikke tager højde for generelle ligevægtseffekter, fx som følge øget personaleefterspørgsel som konsekvens af en landsdækkende to-lærer-ordning, der eventuelt kan føre til lønpres for lærere og dermed øgede offentlige udgifter.

Til beregning af punkt 1.-3. er der på forskermaskinen skabt et kode-univers. Her benyttes SAS 9.4 til konstruktion og rensning af rådata fra de administrative danske registre, mens output beregnes og printes til Excel via Stata 17⁸. I både SAS- og Stata-delen af universet vil master-filer guide igennem de forskellige trin i etableringen af modelstrukturen, indhentning og bearbejdning af rådata til konstruktion af konsekvenser, baggrundskarakteristika og link-outcomes samt estimation af output. Universet tager udgangspunkt i kodestrukturen, der blev skabt i forbindelse med udviklingen af Socialstyrelsens SØM. For i videst muligt omfang at sikre en smidig og effektiv databehandling i RE-

⁸ Input i beregningsmodulet er csv-filer, men da det ikke er tilladt at udsende csv-filer fra forskermaskinen, konverteres Excel-filerne efter udsendelse til csv.

FUD sker bearbejdelsen af registre fra DST's grunddatabank med afsæt i Soci- alstyrelsens oversættelse og videreudvikling af VIVEs oprindelige Stata-base- rede SØM-model i SAS. Således er universet kodet op ved hjælp af makroer, der hurtigt og enkelt tillader opdatering af modellen med nye datakohorter, konsekvensmål og link-outcomes.

5.1 Modelstrukturen og beregninger på Danmarks Statistiks forskermaskine

Vi anvender registerdata til at konstruere en overordnet populationsstruktur på forskermaskinen: elevrammen. Alle registerbaserede beregninger på for- skermaskinen foregår ved hjælp af denne ramme. Specifikt anvendes special- undervisningsregisteret (UDSP) koblet med den seneste version af instituti- onsregisteret (INST2020) til at udtrække alle elever i folkeskolens normalklas- ser og indplacere disse på 0.-9. klassetrin på baggrund af uddannelses-delko- den UDEL. Specialundervisningsregisteret, og dermed også elevrammen, dækker perioden 2011/2012-2020/2021, men ud fra ønsket om at bruge de mest aktuelle data og på grund af det markante regimeskifte i forbindelse med folkeskolereformen i 2014 begrænses elevrammen til skoleåret 2014/2015 og frem.⁹ Denne afgrænsning tillader dog stadig minimum to dataårgange i hele modellens 4-årige horisont for stort set alle outcomes (undtaget er outcomes baseret på landspatientregistrene).

Elevrammen kobles herefter med modellens øvrige datainput: link-outcomes, konsekvenser og baggrundsvariable, hvorefter beregningsrammens input kan konstrueres. Elevens kommunale tilhørsforhold baserer sig alene på admini- strationskommunen for den skole, hvor elever er indskrevet ved skoleårets af- slutning. For yderligere detaljer om strukturen for kode-universet på forsker- maskinen henvises til master-filen til databearbejdelse af grunddata i Bilag 3 og master-filen til beregning af output i Bilag 4.

5.1.1 Lokale standardafvigelser på link-outcomes

REFUD-prototypen tager udgangspunkt i de effektstørrelser, der er tilveje- bragt gennem Afdækningsrapportens litteraturstudie Jacobsen et al. (2022). Specifikt tages udgangspunkt i to-voksen-indsatsen beskrevet i Andersen et al. (2020). Som tilfældet er med dette studie, vil effektstørrelserne for indsats- effekter på børn- og ungeområdet ofte være målt direkte på læring (som app- roksimeret ved testscores/karakterer) eller forskellige mål for trivsel og mental

⁹ Afhængig af link-outcome kan der dog forekomme yderligere periodemæssige begrænsninger, se fx afsnit 5.2 om De Nationale Test som link-outcomes.

sundhed. Sådanne direkte effekter kan ikke omregnes til umiddelbare økonomiske gevinster og betegnes dermed i modellens regi som link-outcomes, idet modellens beregnede økonomiske indsatseffekter på forbrug af offentlige ydelser og aktiviteter estimeres gennem disse 'links'.

I det første af beregningsrammens input omsættes litteraturens effektstørrelser til lokalt (kommunalt) niveau ved at multiplicere effektstørrelsen med den kommunespecifikke standardafvigelse på (et tilsvarende) link-outcome.¹⁰ Konkret beregnes de lokale standardafvigelser af link-outcomes ($Base^{ab}$) som standardafvigelsen af hver enkelt link-outcome for hver kommune baseret på skolernes administrationskommune på hvert klassetrin.

Idet det ofte vil have interesse at kunne beregne indsatseffekter for klassetrin, der ikke umiddelbart dækkes af tilgængelige link-outcomes, benyttes Stata's *ipolate*-kommando til lineært at interpolere mellemliggende standardafvigelser separat for hver kommune, mens lokale standardafvigelser på klassetrin *før* første link-outcome måling erstattes med den tidligste måling for hver kommune. Fx benyttes den lokale standardafvigelse på 2. klasse test scores i Dansk, læsning også for både 0. og 1. klasse, der ikke bliver testet obligatorisk i Dansk, læsning. Tilsvarende erstattes lokale standardafvigelser på klassetrin *efter* den senest tilgængelige link-outcome med denne (fx, 8. classes testscoren i matematik benyttes også i 9. klasse, der ikke har obligatoriske matematiktest – se også afsnit 5.2). Således afgrænses modellens estimationsflade alene baseret på en vurdering af effektstudiernes generaliserbarhed i beregningsmodulet. For to-voksenordningen gælder det, at nettogevinsterne af indsatsen kan beregnes på mellemtrinnet (4.-6. klasse, se afsnit 3.3).

5.1.2 Betingede korrelationer mellem link-outcome og konsekvenser

Efter at have skaleret litteraturens effektstørrelse af en given indsats til lokalt niveau i afsnit 5.1.1 ovenfor, estimerer vi sammenhængen mellem niveauet for det givne link-outcome og forbrug af modellens konsekvenser. Dette gøres ved hjælp af en statistisk regressionsmodel, hvor der samtidig tages højde for en række baggrundsfaktorer om barnet samt barnets initiale forbrug af flere af ydelserne.

Nedenfor følger en nærmere beskrivelse af estimationsmetoden.

¹⁰ Idet effektstørrelser angiver standardiserede effekter, vil fortolkningen af model-estimerne være gyldig, så længe modellens link-outcome måler nogenlunde samme dimension og dermed rangerer eleverne på samme vis i forhold til hinanden. Dermed kræver REFUD-prototypen ikke nøjagtigt sammenfald mellem litteraturens udfaldsmål og det inkluderede link-outcome, hvilket muliggør benyttelse af andre test scores, end effektstudierne baserer sig på (fx test scores fra læsetesten i De Nationale Test på et andet klassetrin eller test scores fra et andet testregime) til skalering af effektstørrelsen til lokalt niveau.

Estimationsmetode

Vi estimerer sammenhængen mellem modellens link-outcome (θ) i året og det standardiserede forbrug af en offentlige ydelse/aktivitet i en fireårig horisont ($Y_{k,t}, t = 1, \dots, 4$). Konsekvensmålet i år 1 til 4 standardiseres til middelværdi 0 og standardafvigelse 1 for hvert klassetrin. Standardiseringen har alene til formål at sammenhængen mellem link-outcome og konsekvensforbrug efterfølgende skaleres til det lokale kommunale niveau ved at multiplicere med den lokale standardafvigelse på forbruget.

For hver konsekvens ($Y_k, k = 1, \dots, K$) estimeres således følgende regressionsmodel for hvert af indsatsens relevante klassetrin med tilgængelig link-outcome (i det følgende undertrykkes subskription for konsekvensnummer og klassetrin) separat for $t = 1, \dots, 4$:

$$Y_{it} = \delta_t^0 + \beta_t^{bc} \theta_{it} + \delta_t^1 Y_{it=0} + \delta_t^2 X_i + \gamma_t^m + \epsilon_{it},$$

Hvor $Y_{it=0}$ er en vektor af elevens (i) initiale forbrug af offentlige ydelser målt i tidspunkt 0, dvs. i kalenderåret inden link-outcomet måles. Tilsvarende er X_i en vektor af baggrundskarakteristika for eleven, γ_t^m angiver kommune-fixed effects, og ϵ_t er individ-specifikke fejllid. Tabel 5.1 giver et fuldstændigt overblik over de inkluderede kontrolvariable. $\beta_t^{bc}, t = 1, \dots, 4$ er koefficienterne af interesse for hver af de K konsekvenser på hvert tilgængeligt klassetrin og angiver samvariationen mellem link-outcome og konsekvensforbrug henholdsvis i året, et år, to år og tre kalenderår efter måling af link-outcome, når elevens initialforbrug og baggrundskarakteristika holdes konstante. Standardfejl klyngekorrigeres på kommuneniveau for at tage højde for forventet korrelation af fejllidene inden for kommunerne.¹¹

Ikke alle datasammenhænge mellem link-outcome og konsekvensforbrug medtages i modelberegningen. Hvis en β^{bc} -koefficient er for usikkert estimeret, erstattes den med 0, og det tilhørende konsekvensmål udgår af beregningen af indsatsens nettoresultat. Selvom der i de senere år har været en tendens til at skelne mindre til mål for statistik signifikans i fortolkningen af økonomiske sammenhænge, har der traditionelt været en konsensus om at benytte et signifikansniveau på 0,05 for en tilstrækkelig overbevisende statistisk sammenhæng. Imidlertid stiger sandsynligheden for at inkludere falske statistiske sammenhænge (type I-fejl) for et givent signifikansniveau med antallet af analyser, der foretages, hvilket man bør holde sig for øje i anvendelsen af

¹¹ Fordi den budgetterede for den enkelte skole påvirkes af det forventede budget for kommunens øvrige skoler samt kommunens øvrige økonomi generelt og fx dagtilbud, praksis for afdækning af støttebehov, specialtilbud mv. varierer på kommunalt niveau, forventes en betydelig korrelation i modellens fejllid såvel som i selve formationen af link-outcome (elevens faglighed) inden for kommunerne. VIVE vurderer derfor, at både estimatet på sammenhængen mellem link-outcome og konsekvens samt inferensen herfor bør korrigeres for klynger ved hjælp af fixed effects og klynge-robuste varians-estimator på kommunalt niveau.

modeller, der benytter sig af mange samtidige hypotesetest. REFUD-prototypen anvender således et relativt konservativt signifikansniveau på 0,01 for tilstrækkelig statistisk signifikans på den enkelte sammenhæng.

Idet det kun vil være muligt at opnå β^{bc} -koefficienter på klassetrin med tilgængelige link-outcomes, erstattes manglende regressionskoefficienter på øvrige klassetrin vha. en metode tilsvarende afsnit 5.1.1 på de lokale link-outcome standardafvigelser, dvs. regressionskoefficienter på mellemliggende klassetrin interpoleres lineært, mens manglende regressionskoefficienter før (efter) første (sidste) måling erstattes af estimatet.

Tabel 5.1 Baggrundsvariable indeholdt i regressionsmodeller til bestemmelse af forholdet mellem link-outcome og konsekvenser

Variabel (måleenhed)	Kilde
Barnets baggrundsfaktorer	
Alder (målt i år, $t = 0$), kvadreret alder, indikator for køn, indikator for ikke-vestlig oprindelse, placering i mors fødselsrækkefølge, indikator for flerfoldsfødsel	BEF, IEPE
<i>Fødselsinformation</i> : indikatorer for lav fødselsvægt (1.500-2.500 g) og meget lav fødselsvægt (< 1.500 g), indikatorer for apgarscore under 7 og apgarscore 7-9, gestationsalder ved fødsel (dage), mors BMI ved fødsel og indikator for, om mor røg under graviditet	MFR
Barnets initialforbrug af konsekvenser ¹	
<i>Specialundervisning</i> : indikatorer for henholdsvis specialundervisning i almenklasse, specialklasse og specialskole samt antal uger modtaget specialundervisning i almenklasse	UDSP, INST
<i>Sociale foranstaltninger</i> : indikatorer for henholdsvis forebyggende indsats, hjemmebaseret indsats, anbringelse i netværkspleje, anbringelse i plejefamilie, anbringelse på institution	BUFO, BUAF
<i>Sundhedsydelse under sygesikringen</i> : kontakter med praktiserende læge, kontakter med øvrige behandlere	SSSY
<i>Somatisk hospitalskontakt</i> : indlæggelsesdage, ambulante besøg, skadestuebesøg	LPR
<i>Psykiatrisk hospitalskontakt</i> : indlæggelsesdage, ambulante besøg, skadestuebesøg	PSYK_ADM
Forældres baggrundskarakteristika ²	
<i>Mors diagnoser</i> : indikatorer for henholdsvis tidligere eller i året diagnosticeret med opmærksomhedsforstyrrelse samt tidligere eller i året diagnosticeret med angst, depression eller anden affektiv lidelse	LPR, PSYK
<i>Mors arbejdsmarkedstilknytning på 5 kategorier</i> : lønmodtager (ref.), arbejdsløs, dagpengemodtager, modtager af overførselsindkomst, modtager andre sociale ydelser	AKM
<i>Mors uddannelsesniveau på 3 kategorier</i> : gymnasial uddannelse eller lavere (ref.), videregående uddannelse, erhvervsuddannelse	UDDA
<i>Mors indkomst</i> : disponibel indkomst (ln), lønindkomst (ln, smal), arbejdsdage i året	BEF, IND, BFL
Indikator for om mor er gift/samlevende, mors alder ved fødsel, mors kvadrerede alder ved fødsel	BEF
<i>Fars diagnoser</i> : indikatorer for henholdsvis tidligere eller i året diagnosticeret med opmærksomhedsforstyrrelse samt tidligere eller i året diagnosticeret med angst, depression eller anden affektiv lidelse	LPR, PSYK
<i>Fars arbejdsmarkedstilknytning på 5 kategorier</i> : lønmodtager (ref.), arbejdsløs, dagpengemodtager, modtager af overførselsindkomst, modtager andre sociale ydelser	AKM
<i>Fars uddannelsesniveau på 3 kategorier</i> : gymnasial uddannelse eller lavere (ref.), videregående uddannelse, erhvervsuddannelse	UDDA
<i>Fars indkomst</i> : disponibel indkomst (ln), lønindkomst (ln, smal), arbejdsdage i året	BEF, IND, BFL
Indikator for om far er gift/samlevende, fars alder ved fødsel, fars kvadrerede alder ved fødsel	BEF

Anm.: Tabellen opsummerer de inkluderede kontrolvariable i regressionsmodellen til estimation af samvariationen mellem link-outcome og konsekvensforbrug for eleven. Ud over regressorerne oplyst i tabellen inkluderes også indikatorer for årstal.

Note: ¹ Målt i kalenderåret før analyse ($t=0$). Se Tabel 5.3 for detaljer omkring konsekvensmålene.

² Forældres baggrundskarakteristika er målt i kalenderåret før analyse ($t=0$), medmindre andet er angivet.

5.1.1 Lokale baseline standardafvigelser

Herefter beregnes standardafvigelsen for hvert konsekvensmål for hver kommune til skalering af samvariationen mellem link-outcome og konsekvensmål til lokalt (kommunalt) niveau.¹² Modellen benytter separate baseline standardafvigelser på konsekvensmålene for hvert år i modellens 4-årige tidshorisont ($Base_t^{bc}, t = 1, \dots, 4$), og således justeres den estimerede samvariation mellem link-outcome og konsekvensforbrug i år 4 med standardafvigelsen på eleverne i kommunes konsekvensforbrug i år 4 ($Base_4^{bc}$) baseret på skolernes administrationskommune.

5.2 Link-outcomes

REFUD-prototypen indeholder baseret på Andersen et al. (2020) to link-outcomes til estimation af indsatseffekter på modellens konsekvenser (dvs. forbrug af offentlige ydelser og aktiviteter): testscores fra De Nationale Tests læse- og matematiktest, se Tabel 5.2). Disse danner dermed grundlag for estimation af lokale standardafvigelser på link-outcomes ($Base^{ab}$) i afsnit 5.1.1 og den betingende sammenhæng mellem link-outcomes og konsekvensforbrug (β^{bc}) i afsnit 5.1.2.

Tabel 5.2 Tilgængelige link-outcomes i REFUD-prototypen

Link-outcome	Kilde	Tilgængelige klassetrin	Registerbehandling	Linknummer
Testscore i Dansk, læsning	De Nationale Test	2., 4., 6. og 8. klasse	Den standardiserede testscore (middelværdi 0, standardafvigelse 1 inden for testår og klassetrin) er konstrueret som det standardiserede gennemsnit af de standardiserede testscores for hvert af de tre profilområder.	1
Testscore i matematik	De Nationale Test	3., 6. og 8. ¹ klasse	Den standardiserede testscore (middelværdi 0, standardafvigelse 1 inden for testår og klassetrin) er konstrueret som det standardiserede gennemsnit af de standardiserede testscores for hvert af de tre profilområder.	2

Anm.: Tabellen angiver REFUD-prototypens tilgængelige link-outcomes, hvordan de er konstrueret, samt for hvilke klassetrin disse er relevante. Testscores i dansk, læsning og matematik er indhentet fra De Nationale Test-data fra STIL.

De Nationale Test blev suspenderet i skoleåret 2019/2020 og er derfor kun tilgængelige til og med 2018/2019.

Note: ¹ Matematiktesten i 8. klasse blev først tilføjet De Nationale Test i skoleåret 2017/2018.

¹² Både af hensyn til gyldigheden af modelresultaterne samt Danmarks Statistiks retningslinjer vedrørende diskretionering beregnes lokale baseline standardafvigelser ikke for klassetrin hvor færre end ti individer har/ikke har et positivt forbrug af ydelser. Desuden beregnes standardafvigelser for visse konsekvenser ikke, hvis de falder uden for de angivne aldersbegrænsninger (se Tabel 5.3).

De Nationale Test måler elevdygtighed på tre forskellige domæner inden for hvert fag, kaldet profilområder. For at opnå en samlet testscore i et fag, standardiseres testscores til middelværdi 0 og standardafvigelse 1 i hver test først inden for hvert profilområde, klassetrin og testår. Det er nødvendigt at standardisere baseret på testår, da testene løbende er blevet udviklet og opdateret, og testforudsætningerne samt item-banken, fra hvilken testspørgsmålene trækkes, har ændret sig fra år til år. Herefter beregnes for hver test et gennemsnit af den standardiserede score for profilområderne, før dette gennemsnit igen standardiseres på baggrund af klassetrin og testår. Denne standardisering tillader, at testscores på tværs af klassetrin og år er umiddelbart sammenlignelige inden for samme fag. De standardiserede, samlede testscores kan nu tilføjes elevrammen baseret på elev-id og skoleår, hvorefter standardafvigelserne for hvert link-outcome (testfag) kan beregnes for hvert klassetrin og kommune. Testscores for test, der er foretaget uden for påkrævede klassetrin, blændes i beregningen af lokale standardafvigelser og betingede korrelationer.

De Nationale Test blev suspenderet i skoleåret 2019/2020. I koblingen med elevrammen, der er afgrænset af skolereformen i 2014/2015, kan testdata fra De Nationale Test dermed foreløbig benyttes som link-outcome i modelberegninger i perioden 2014/2015-2018/2019.

5.3 Konsekvenser

Prototypen benytter en fast liste af konsekvenser, der påvirker de offentlige budgetter, til hvilke effekt-estimer fra litteraturen omsættes enten via link-outcomet eller direkte ved sammenfald mellem modellem konsekvens og effektstudiets outcome. En stor del af disse tager udgangspunkt i Socialstyrelsens SØM v. 2.5, som fx antal uger i året, eleven modtager specialundervisning, og kontakter med praktiserende læge i året under sygesikringen, mens enkelte er omdefinert for i højere grad at have relevans for populationen af normalklasse-elever i folkeskolen, fx opdeling af grundskole forbrug i kommunale folkeskoler samt fri- og privatskoler, og andre er tilføjet som helt nye. De nye konsekvenser, der er udarbejdet i forbindelse med prototypen, er beskrevet nederst i Tabel 5.3. Særligt bør nævnes, at de nye konsekvenser omhandler elevens forældre, hvor der måles på eventuelle indsats effekter på disses arbejdsmarkedsudfald samt sygefravær blandt elevens lærere. Den fuldstændige liste over de i modellen tilgængelige konsekvenser er angivet i Tabel 5.3. For detaljeret information om de øvrige konsekvenser henvises læseren til Jacobsen et al. (2018), Jacobsen & Beuchert (2018) og Jacobsen & Ibsen (2021).

Data til konstruktion af konsekvenser er ved modeludvikling generelt tilgængelig til og med året 2020. Undtagelser herfra tæller sundhedsydelseskonsekvenser, der er baseret på Landspatientregisteret, og er tilgængelige til og med 2018, samt sociale foranstaltninger og personalefravær, der er tilgængelige til og med 2019.

5.3.1 Default konsekvenser

Idet det fulde omfang af konsekvensmål præsenteret i Tabel 5.3 vurderes som omfattende i forhold til, hvad man i mange tilfælde vil forvente, at en skolebaseret indsats vil kunne påvirke, og for dermed at undgå, at beregnede indsats-effekter baseres på tilfældige sammenhænge med statistisk signifikans, inkluderer beregningsrammen i udgangspunktet ikke samtlige konsekvenser til beregning af indsatseffekter. Det er dog altid muligt for brugeren selv at bestemme, hvilke konsekvenser indsatsens nettogevinst skal beregnes over.

Til bestemmelse af modellens default benyttes de konsekvenser, der forventes at være umiddelbart påvirket af en generel skolebaseret indsats gennem elevernes faglighed, trivsel og lærernes arbejdsmiljø. Således er konsekvenser omhandlende elevens videre skolegang, beskæftigelse og risikofyldt adfærd samt forældres beskæftigelse og lærernes sygefravær medtaget som default konsekvenser. Elevens skolegang og ungdomsuddannelse er umiddelbart relateret til elevens faglige kompetencer og vil dermed naturligt påvirkes af en skolebaseret indsats som to-voksen-ordningen. Ligeledes medtages elevens forbrug af sundhedsydelser og kontakt med politi, retsvæsen og kriminalforsorg, da disse betragtes som registerbaserede mål for risikofyldt adfærd hos eleven, der relaterer sig til (mis)trivsel.

Sluttelig inkluderes beskæftigelsesinformation for modellens aktører: eleven, dennes forældre og lærere. Som default indgår eleven og forældrene med deres lønindkomster. For eleven måler denne dermed både mængden af fritidsarbejde, men også eventuelt fuldtidsarbejde efter endt skolegang. Forældrenes lønindkomst skal her måle forældres beskæftigelsesrespons til deres barns forbedrede skole-outcomes på både den intensive og ekstensive margin. Da førsteordenseffekter af barnets skole-outcomes primært forventes at påvirke den intensive margin, udelades konsekvenser omhandlende forældrenes indkomstoverførsler og modtagne beskæftigelsesindsatser fra defaultlisten. Lærerenes sygefravær antages ligeledes at være påvirket af deres arbejdsbetingelser, herunder direkte gennem en skolebaseret indsats og indirekte gennem indsatsens påvirkning på elevernes sociale og faglige kompetencer. I de tilfælde, hvor effektstudierne fra litteraturen ikke præsenterer estimater af indsatseffekten direkte på lærerfravær, vil REFUD-prototypens beregning således basere sig alene på sidstnævnte indirekte mekanisme.

Tabel 5.3 Modellens konsekvenser opdelt på hoveddimension og aktivitet

Hoveddimension	Aktivitet	Enhed	Kilde(r)	Default konsekvens	Nr.
Grundskole og specialundervisning (elev)	Specialskole ^{1,2}	Antal uger indskrevet	Elevregister (KOTRE) og Institutionsregister (INST)	Ja	1
	Kommunal grundskole, 0.-9. klasse almen ¹	Antal uger indskrevet	Elevregister (KOTRE) og Institutionsregister (INST)	Ja	2
	Kommunal grundskole, 10. klasse almen ^{1,8}	Antal uger indskrevet	Elevregister (KOTRE) og Institutionsregister (INST)	Ja	3
	Fri- og privat grundskole, 0.-9. klasse almen ¹	Antal uger indskrevet	Elevregister (KOTRE) og Institutionsregister (INST)	Ja	4
	Fri- og privat grundskole, 10. klasse almen ^{1,8}	Antal uger indskrevet	Elevregister (KOTRE) og Institutionsregister (INST)	Ja	5
	Specialundervisning i almen klasse ^{1,2}	Antal uger indskrevet	Elevregister (KOTRE) og Specialundervisning (UDSP)	Ja	6
	Specialklasse ^{1,2}	Antal uger indskrevet	Elevregister (KOTRE) og Specialundervisning (UDSP)	Ja	7
Ungdomsuddannelse (elev) ^{2, 8}	Forberedende uddannelsesstilbud (ekskl. STU)	Antal uger indskrevet	Elevregister (KOTRE)	Ja	8
	Særligt tilrettelagt uddannelse (STU)	Antal uger indskrevet	Elevregister (KOTRE)	Ja	9
	Ordinær ungdomsuddannelse	Antal uger indskrevet	Elevregister (KOTRE)	Ja	10
Sociale foranstaltninger til børn og unge under 18 år ²	Forebyggende indsatser (§ 11, stk. 3)	Modtager foranstaltning	Børn og Unge Forebyggende Foranstaltninger (BUFO)		11
	Hjemmebaserede foranstaltninger	Modtager foranstaltning	Børn og Unge Forebyggende Foranstaltninger (BUFO) og Børn og Unge Anbragte Forløbsregister (BUAF)		12
	Anbringelse i slægts- eller netværkspleje	Antal dage	Børn og Unge Anbragte Forløbsregister (BUAF)		13
	Anbringelse i familiepleje	Antal dage	Børn og Unge Anbragte Forløbsregister (BUAF)		14
	Anbringelse på institution	Antal dage	Børn og Unge Anbragte Forløbsregister (BUAF)		15
Efterværn (til 18-22-årige) ^{2, 7}	Efterværn med hjemmebaserede foranstaltninger	Modtager foranstaltning	Børn og Unge Forebyggende Foranstaltninger (BUFO) og Børn og Unge Anbragte Forløbsregister (BUAF)		16
	Efterværn med anbringelse i slægt eller netværk	Antal dage	Børn og Unge Anbragte Forløbsregister (BUAF)		17
	Efterværn med anbringelse i familiepleje	Antal dage	Børn og Unge Anbragte Forløbsregister (BUAF)		18
	Efterværn med anbringelse på institution	Antal dage	Børn og Unge Anbragte Forløbsregister (BUAF)		19

Hoveddimension	Aktivitet	Enhed	Kilde(r)	Default kon- sekvens	Nr.
Sundhedsydelse (elev) ³	Praktiserende læge	Antal sygesikrings- kontakter	Sygesikringsregisteret (SSSY)	Ja	20
	Øvrige behandlere under sygesikringen	Antal sygesikrings- kontakter	Sygesikringsregisteret (SSSY)	Ja	21
	Somatisk indlæggelse	Antal sengedage	Landspatientregisteret (LPR)	Ja	22
	Somatisk ambulantbesøg	Antal besøg	Landspatientregisteret (LPR)	Ja	23
	Somatisk skadestuebesøg	Antal besøg	Landspatientregisteret (LPR)	Ja	24
	Psykiatrisk indlæggelse	Antal sengedage	Landspatientregisteret Psykiatri (PSYK)	Ja	25
	Psykiatrisk ambulantbesøg	Antal besøg	Landspatientregisteret Psykiatri (PSYK)	Ja	26
	Psykiatrisk skadestuebesøg	Antal besøg	Landspatientregisteret Psykiatri (PSYK)	Ja	27
Politi, retsvæsen og kriminalforsorg (elev) ^{3,6}	Fængselsdage	Antal dage	Kriminalstatistiske Afgørelser (KRAF) og Kriminalstatistiske Indsættelser	Ja	28
	Domme for vold og sædelighedsforbrydelser	Antal domme	Kriminalstatistiske Afgørelser (KRAF)	Ja	29
	Domme for indbrud, tyveri og hærværk	Antal domme	Kriminalstatistiske Afgørelser (KRAF)	Ja	30
	Domme for overtrædelse af færdselsloven og øvrige særlove	Antal domme	Kriminalstatistiske Afgørelser (KRAF)	Ja	31
Indkomstoverførsler og beskæftigelsesindsats (elev) ^{3,7}	Kontanthjælp	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		32
	Førtidspension	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		33
	Uddannelseshjælp	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		34
	Vejledning og opkvalificering	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		35
	Øvrig overførsel (ekskl. SU)	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		36
Beskæftigelse (elev) ⁴	Lønindkomst	Kr. pr. år	Detaljeret lønmodtagerdata fra e-Indkomst (BFL)	Ja	37
Sociale serviceydelse (elev) ^{3,7}	Herberg og forsorgshjem (§ 110)	Antal dage	Boformerregisteret (BO-FORM_F)		38
	Stofmisbrugsbehandling, dag	Antal dage	Register over Stofmisbrugere i Behandling (SIB)		39
	Stofmisbrugsbehandling, døgn	Antal dage	Register over Stofmisbrugere i Behandling (SIB)		40
	Alkoholmisbrugsbehandling	Antal dage	Det National Alkoholbehandlingsregister (NAB)		41
	Kontanthjælp	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		42

Hoveddimension	Aktivitet	Enhed	Kilde(r)	Default konsekvens	Nr.
Indkomstoverførsler og beskæftigelsesindsats (mødre) ¹	Førtidspension	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		43
	Uddannelseshjælp	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		44
	Vejledning og opkvalificering	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		45
	Øvrige overførsler (ekskl. SU)	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		46
Beskæftigelse (mødre) ⁴	Lønindkomst	Kr. pr. år	Detaljeret lønmodtagerdata fra e-Indkomst (BFL)	Ja	47
Indkomstoverførsler og beskæftigelsesindsats (fædre) ¹	Kontanthjælp	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		48
	Førtidspension	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		49
	Uddannelseshjælp	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		50
	Vejledning og opkvalificering	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		51
	Øvrige overførsler (ekskl. SU)	Antal uger	Beskæftigelsesministeriets forløbsdatabase (DREAM)		52
Beskæftigelse (fædre) ⁴	Lønindkomst	Kr. pr. år	Detaljeret lønmodtagerdata fra e-Indkomst (BFL)	Ja	53
Personalefravær ⁵	Lærernes sygefravær	Antal fuldtidsdage	Fraværsperioder (FRPE) og Læreres Kompetencedata fra STIL	Ja	54

Anm.: Tabellen angiver, hvilke konsekvenser og ydelser der estimeres indsatseffekter for i prototypen, herunder måleenhed og registerkilde, samt om de i udgangspunktet er inkluderet i måling af indsatsers nettogevinst, eller om de skal tilvælges af brugeren.

Note: ¹ Grundskolen (0.-9. klasse) og grundskolen (10. klasse) er ved hjælp af elevregisteret opgjort som antal uger indskrevet i det pågældende år. Disse er hver især opdelt på institutionstype via Institutionsregisteret, således at INST3 = 1012 angiver antal uger indskrevet i kommunal grundskole og INST3 = 1013 angiver antal uger indskrevet i fri- eller privatskole. Denne opdeling af konsekvensmålet for grundskoleaktivitet er nyudviklet til REFUD-prototypen. Konsekvensmål for 0.-9. klasse måles til og med 9. klassetrin.

² Konsekvensmålet er indhentet fra Socialstyrelsens SØM v. 2.5 (se Jacobsen & Beuchert (2018) for detaljeret beskrivelse).

³ Konsekvensmålet er indhentet fra Socialstyrelsens SØM v. 2.5 (se Jacobsen et al. (2018) for detaljeret beskrivelse).

⁴ Beskæftigelse måles som noget nyt ved hjælp af AJO_SMALT_LOENBELOEB i den detaljerede lønmodtagerdata fra e-Indkomst. Dette tillader i højere grad justeringer af beskæftigelse på den intensive margin i forhold til det tilsvarende konsekvensmål baseret på DREAM (se Jacobsen et al. (2018)), hvilket vurderes relevant for både elevpopulationen, der er karakteriseret ved fritidsarbejde, og for forældrepopulationens respons til en skolebaseret indsats. Beskæftigelse på elevniveau måles tidligst fra det klassetrin, hvor den øverste alderspercentil er 13 år.

⁵ Lærer kobles til elevrammen ved hjælp af klassebetegnelsesvariablen i lærernes kompetencedata fra STIL. Fraværsmålet er baseret på fraværstimer (FRAVTIMER) omregnet til fuldtidsdage (7,4 timer pr. dag) for fraværsårsag = 1100 (egen sygdom) i fraværspærederegisteret FRPE. Fraværstimer opsummeres i to halvår: januar-juli og august-december, således at sygefraværet kan tilskrives de skoleår, hvor læreren har undervist en given klasse. Læreren fravær tildeles en klasse baseret på lærerens relative undervisningsbelastning i klassen i Lærernes Kompetencedata, hvoraf den enkelte elev tilskrives lærerens vægtede fravær divideret med klassestørrelsen i fra specialundervisningsregisteret UDSP. Dette mål forudsætter således, at lærerfraværet er lineært med antallet af elever i indsatsskolen.

⁶ Konsekvensmål måles tidligst fra det klassetrin, hvor den øverste alderspercentil er 15 år.

⁷ Konsekvensmål måles tidligst fra det klassetrin, hvor den øverste alderspercentil er 18 år.

⁸ Konsekvensmål måles tidligst fra 9. klassetrin.

5.4 Priser

For at kvantificere konsekvensestimaterne i sammenlignelige og økonomisk håndterbare størrelser, skaleres effekt-estimerne på hver konsekvens med en enhedspris. Med undtagelse af beskæftigelse for elever og forældre opgøres konsekvenserne som ændringer i forbrug af ydelser og aktiviteter, og værdien af den enkelte aktivitet beregnes herefter som konsekvensforbruget multipliceret med enhedsprisen.

Dette stiller to centrale krav til modellens enhedspriser (se også Jacobsen & Beuchert, 2018):

1. Prisen skal lige præcis dække omkostningen ved, at en borger forbruger en ydelse eller offentlig service, og må således ikke være overlappende med priser for øvrige ydelser.
2. Så vidt muligt skal alle omkostninger ved den pågældende indsats være inkluderet, dvs. prisen skal også indeholde administrations-, ledelses-, sagsbehandlings- og andre overheadomkostninger.

Prototypen indeholder standardforslag til enhedsprisen for hver konsekvens baseret på landsgennemsnittet; det er dog muligt at justere standardindstillingen for brugere med særlig indsigt i omkostningsniveauet for en given kommune. Tabel 5.4 giver et fuldstændigt overblik over prototypens indeholdte standardpriser.

Ved sammenfald af konsekvenser mellem REFUD-prototypen og Socialstyrelsens SØM er benyttet priserne fra SØM version 2.5.¹³ Enhedspriser på konsekvenser inden for BUVM's ressortområde er dog opdateret i samarbejde med BUVM uanset konsekvenssammenfald med øvrige SØM-modeller.

Enkelte priser beregnes direkte ved hjælp af de administrative registre på forskermaskinen. Det drejer sig om de gennemsnitlige marginalskattesatser for elever og forældre til beregning af "prisen" (statens og kommunens indtægt) ved indsatseffekter på elever og forældres arbejdsudbud, samt enhedsomkostninger på sundhedsydelserne baseret på sygesikringskontakter og kontakt med somatisk sygehus.

I det følgende gennemgås udregningen af enhedspriserne på de ydelser, der ikke er sammenfaldende med priserne oplyst i Socialstyrelsens SØM version 2.5 (se også Jacobsen et al., 2018; Jacobsen & Beuchert, 2018; Jacobsen & Ibsen, 2021).

¹³ Den seneste version af den [Socialøkonomiske Investeringsmodel, SØM](#), kan hentes her.

Tabel 5.4 Modellens enhedspriser og fordeling mellem stat, region og kommune (2021-priser)

Hoveddimension	Aktivitet	Enhed	Enhedspris i alt	Stat	Region	Kommune
Grundskole og specialundervisning (elev) ¹	Kommunal grundskole, 0.-9. klasse almen	Kr. pr. uge	1.466	0	0	1.466
	Kommunal grundskole, 10. klasse almen	Kr. pr. uge	1.466	0	0	1.466
	Fri- og privat grundskole, 0.-9. klasse almen	Kr. pr. uge	*	*	*	*
	Fri- og privat grundskole, 10. klasse almen	Kr. pr. uge	*	*	*	*
	Specialskole	Kr. pr. uge	7.821	0	0	7.821
	Specialklasse	Kr. pr. uge	4.558	0	0	4.558
	Specialundervisning i almen klasse	Kr. pr. uge	5.105	0	0	5.105
Ungdomsuddannelse (elev) ¹	Forberedende uddannelsstilbud (ekskl. STU)	Kr. pr. uge	2.761	966	0	1.794
	Særligt tilrettelagt uddannelse (STU)	Kr. pr. uge	4.572	0	0	4.572
	Ordinær ungdomsuddannelse	Kr. pr. uge	1.793	1.793	0	0
Sociale foranstaltninger til børn og unge under 18 år ²	Forebyggende indsatser (§11, stk. 3)	Kr. pr. modtager	61.604	3.640	0	57.965
	Hjemmebaserede foranstaltninger	Kr. pr. modtager	172.264	15.933	0	156.330
	Anbringelse i slægts- eller netværkspleje	Kr. pr. dag	587	24	0	563
	Anbringelse i familiepleje	Kr. pr. dag	1.539	22	0	1.517
	Anbringelse på institution	Kr. pr. dag	3.458	323	0	3.135
Efterværn (til 18-22-årige) ²	Efterværn med hjemmebaserede foranstaltninger	Kr. pr. modtager	172.264	15.933	0	156.330
	Efterværn med anbringelse i slægt eller netværk	Kr. pr. dag	587	24	0	563
	Efterværn med anbringelse i familiepleje	Kr. pr. dag	1.539	22	0	1.517
	Efterværn med anbringelse på institution	Kr. pr. dag	3.458	323	0	3.135
Sundhedsydelse (elev) ^{2,3}	Praktiserende læge	Kr. pr. kontakt	156	0	132	24
	Øvrige behandlere under sygesikringen	Kr. pr. kontakt	440	0	335	105

Hoveddimension	Aktivitet	Enhed	Enhedspris i alt	Stat	Region	Kommune
	Somatisk indlæggelse	Kr. pr. senge- dag	4.781	0	3.452	1.329
	Somatisk ambulantbesøg	Kr. pr. besøg	2.291	0	1.678	613
	Somatisk skadestuebesøg	Kr. pr. besøg	1.504	0	1.113	390
	Psykiatrisk indlæggelse	Kr. pr. senge- dag	3.885	0	1.554	2.331
	Psykiatrisk ambulantbesøg	Kr. pr. besøg	1.944	0	1.361	583
	Psykiatrisk skadestuebesøg	Kr. pr. besøg	1.944	0	1.361	583
Politi, retsvæsen og kriminalforsorg (elev) ⁴	Fængselsdage	Kr. pr. dag	1.696	1.696	0	0
	Domme for vold og sædelighedsforbrydelser	Kr. pr. dom	59.974	59.974	0	0
	Domme for indbrud, tyveri og hærværk	Kr. pr. dom	27.437	27.437	0	0
	Domme for overtrædelse af Færdselsloven og øvrige særlove	Kr. pr. dom	15.119	15.119	0	0
Indkomstoverførsler (elev) ⁵	Kontanthjælp	Kr. pr. uge	2.806	561	0	2.245
	Førtidspension	Kr. pr. uge	3.621	724	0	2.897
	Uddannelseshjælp	Kr. pr. uge	1.682	336	0	1.345
	Øvrig overførsler (ekskl. SU)	Kr. pr. uge	1.828	914	0	914
Beskæftigelsesindsats (elev) ⁵	Vejledning og opkvalificering	Kr. pr. uge	3.543	709	0	2.834
Beskæftigelse (elev) ³	Lønindkomst pr. år	Marginalskat (%)	0.104	0.090	0	0.014
Sociale serviceydelser (elev) ⁴	Herberg og forsorgshjem (§ 110)	Kr. pr. dag	1.087	544	0	544
	Stofmisbrugsbehandling, dag	Kr. pr. dag	204	0	0	204
	Stofmisbrugsbehandling, døgn	Kr. pr. dag	2.376	0	0	2.376
	Alkoholmisbrugsbehandling	Kr. pr. dag	89	0	0	89
Indkomstoverførsler (mødre) ⁵	Kontanthjælp	Kr. pr. uge	3.621	724	0	2.897
	Førtidspension	Kr. pr. uge	2.806	561	0	2.245

Hoveddimension	Aktivitet	Enhed	Enhedspris i alt	Stat	Region	Kommune
	Uddannelseshjælp	Kr. pr. uge	1.682	336	0	1.345
	Øvrig overførsler (ekskl. SU)	Kr. pr. uge	3.543	709	0	2.834
Beskæftigelsesindsats (mødre) ⁵	Vejledning og opkvalificering	Kr. pr. uge	1.828	914	0	914
Beskæftigelse (mødre) ³	Lønindkomst pr. år	Marginalskat (%)	0.405	0,203	0	0.202
Indkomstoverførsler (fædre) ⁵	Kontanthjælp	Kr. pr. uge	3.621	724	0	2.897
	Førtidspension	Kr. pr. uge	2.806	561	0	2.245
	Uddannelseshjælp	Kr. pr. uge	1.682	336	0	1.345
	Øvrig overførsler (ekskl. SU)	Kr. pr. uge	3.543	709	0	2.834
Beskæftigelsesindsats (fædre) ⁵	Vejledning og opkvalificering	Kr. pr. uge	1.828	914	0	914
Beskæftigelse (fædre) ³	Lønindkomst pr. år	Marginalskat (%)	0.442	0.229	0	0.213
Personalets fravær ¹	Lærerfravær	Kr. pr. dag	*	*	*	*

Anm.: Tabellen angiver de priser, der er tilgængelig som standardindstilling i REFUD. De sidste tre kolonner angiver omkostningerne fordelt på henholdsvis stat, region og kommune. Alle priser er 2021-priser.

Note: ¹ Priser og fordeling af omkostninger på grundskole- og ungdomsuddannelsesområdet samt tilsvarende for personalefravær i grundskolen er udarbejdet i samarbejde med BUVM. Se rapporttekst for yderligere detaljer.

² Priser og fordeling af omkostninger er indhentet fra Socialstyrelsens SØM v. 2.5. Disse er dokumenteret nærmere i Jacobsen & Beuchert (2018).

³ Priser og fordeling af omkostninger på ydelser under sygesikringen og somatisk hospital samt marginalsattesatser er baseret på egne beregninger på registerdata. Priserne på sundhedsområdet er beregnet ved hjælp af beregningsmetoden beskrevet i Jacobsen & Beuchert (2018). Se rapporttekst for yderligere detaljer.

⁴ Priser og fordeling af omkostninger er indhentet fra Socialstyrelsens SØM, v. 2.5. Disse er dokumenteret nærmere i Jacobsen & Ibsen (2021).

⁵ Priser og fordeling af omkostninger er indhentet fra Socialstyrelsens SØM v. 2.5. Disse er dokumenteret nærmere i Jacobsen et al. (2018).

* Leveres af BUVM efterfølgende.

Kilde: Egne beregninger og SØM v. 2.5. For dokumentation, se Jacobsen et al. (2018), Jacobsen & Beuchert (2018) og Jacobsen & Ibsen (2021). Priser er fremskrevet til 2021-priser ved hjælp af Statistikbankens Tabel PRIS9 og Økonomisk Redegørelse 2020 (Finansministeriet, 2020).

5.4.1 Priser: Grundskole og specialundervisning

Leveres af BUVM.

5.4.2 Priser: Ungdomsuddannelse

Leveres af BUVM.

5.4.3 Priser: Behandling under sygesikringen og somatisk hospitalskontakt

Enhedspriser på behandlinger under sygesikringen og somatisk hospitalskontakt er baseret på egne registerberegninger på Danmarks Statistiks forskermaskine. Priserne er opgjort på samme måde som for børnemålgrupperne i Socialstyrelsens SØM v. 2.5 (se Jacobsen et al. (2018) og Jacobsen & Beuchert (2018)), men er her baseret på gruppen af 0-17-årige i 2019 (det senest tilgængelige dataår).

Priserne for *sygesikringskontakter* opgøres på baggrund af variabelen BRUHON fra sygesikringsregisteret (SSSY2019) som gennemsnitligt bruttohonorar til læge mv. pr. kontakt for henholdsvis kontakter med praktiserende læge (SPEC2 = 70, 80) og kontakter med øvrige behandlere (SPEC2 = 01-68, 81-93). Fordelingen af omkostningerne på kommune og region tager udgangspunkt i Bekendtgørelse om den kommunale medfinansiering på sundhedsområdet (BEK 1881 af 01/12/2020), hvor kommunerne finansierer 14 % af udgifterne for 0-2-årige for kontakter med almen læge, fysioterapi, tandlæge, foterapi, psykolog og kiropraktor, og 7 % af de tilsvarende udgifter for de 3-64-årige. For øvrige specialer tilskrives 45 % kommunen (dog maksimalt 2.690 kr. i 2020-priser) for de 0-2-årige og 20 % (dog maksimalt 1.614 kr. i 2020-priser) for de 3-64-årige.

Priserne for *somatisk hospitalskontakt* opgøres på baggrund af variabelen TOTALPRIS_DRG fra registeret over DRG-grupperede LPR-kontakter (DRGSOMA_KONTAKT2018) som den gennemsnitlige pris pr. liggedag for indlæggelsesdage (PATIENTTYPE = 0), skadestuebesøg (PATIENTTYPE = 2 og INDMAADE = 1) og ambulantsbesøg (PATIENTTYPE = 2 og INDMAADE = 2). Som ovenfor følger fordelingen af omkostningerne på kommune og region Bekendtgørelse om den kommunale medfinansiering på sundhedsområdet (BEK 1881 af 01/12/2020), hvor kommunen tilskrives 45 % (dog maksimalt 26.902 kr. i 2020-priser) for de 0-2-årige og 20 % (dog maksimalt 16.140 kr. i 2020-priser) for de 3-64-årige.

5.4.4 Priser: Beskæftigelse

Skatteindtægten af elevers og forældres *beskæftigelse* (målt som lønindkomst) beregnes som den gennemsnitlige marginalsattesats til kommune og stat i henholdsvis populationen af unge i alderen 13-18 år, populationen af mødre og populationen af fædre. For at få et retvisende billede af indsatselvernes indtjeningsmuligheder inkluderer prisberegningen kun elever, der er 13 år eller ældre, og da elever i indsatsåret i udgangspunktet vil fylde 16 i året, hvor de forlader grundskolen, antages marginalsattesatsen i indsatsåret at skulle beregnes for de 13-15-årige, men da modellen har en 4-årig tidshorisont, vil disse elever være 16-18 år ved modellens udløb. For at tage højde for

de meget forskellige indtjeningsmuligheder, der er for elever i og i årene efter grundskolen, indarbejdes modellens 4-årige horisont således ved at vægte de 14- og 17-årige dobbelt og de 15-16-årige trefold i gennemsnitsberegningen af marginalskat.

Skatteberegningen følger ligesom Socialstyrelsens SØM v. 2.5 principperne beskrevet på Skatteministeriets hjemmeside og benytter gældende skattesatser i 2021¹⁴.

5.4.5 Priser: Lærerfravær

Leveres af BUVM.

Enhedsprisen på *lærerfravær* opgøres som den gennemsnitlige undervisningsbelastning pr. dags fuldtidsarbejde skaleret med den gennemsnitlige vikardækning af undervisningstimer, hvor læreren er fraværende, multipliceret med timeprisen for en gennemsnitlig vikartype.

¹⁴ <https://www.skm.dk/skattetal/satser/skatte-og-afgiftsberegning/skatteberegningseksempel-for-en-ugift-skatteyder-i-2021/>

Litteratur

Andersen, S.C., Beuchert, L., Nielsen, H.S. & Thomsen, M.K. (2020). The Effect of Teacher's Aides in the Classroom: Evidence from a Randomized Trial. *Journal of the European Economic Association*, 18(1), 469-505.

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A. & Borges, B. (2021). *Shiny: Web Application Framework for R*: <https://shiny.rstudio.com>

Finansministeriet (2021). *Dokumentationsnotat – den samfundsøkonomiske diskonteringsrente*. København: Finansministeriet.

Jacobsen, R.H., Andreasen, A.G., Keilow, M., Krassel, K.F., Lesner, R.V., Mortensen, N.P., Rangvid, B.S. & Verner, M. (2022). *Afdækning af muligheder for udbredelse af SØM til indsatser på børne- og undervisningsområdet*. København: VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.

Jacobsen, R.H., Arendt, J.N., Verner, M., Kollin, M.S., Halling, C. & Kolstrup, K. (2018). *Socialøkonomisk investeringsmodel (SØM) – Dokumentation*. København: VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.

Jacobsen, R.H. & Beuchert, L. (2018). *Den Socialøkonomiske Investeringsmodel: Udvidelse med udsatte børn og unge – Dokumentation*. København: VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.

Jacobsen, R.H. & Ibsen, J.T. (2021). *Opdatering af priser til vidensdatabasen i SØM – Politi, retsvæsen og kriminalforsorg samt sociale serviceydelser*. København: VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Wien: R Foundation for Statistical Computing.

RStudio Team (2021). *RStudio: Integrated Development for R*: <https://www.rstudio.com/>

Bilag 1 Ordforklaringer

β^{ab}	Se indirekte evidens.
β^{ac}	Se direkte evidens.
β^{bc}	Se indirekte evidens.
Baggrundvariabel, baggrundskarakteristika	Oplysninger, der kontrolleres for ved estimation af betingede korrelationer mellem link-outcomes og konsekvenser.
Baseline	Tidspunkt for måling af variable. Enten $t = 0$ eller $t = 1$. Se herunder for hvert klassetrin og hver kommune.
Baseline konsekvensforbrug (individet)	Individets baseline konsekvensforbrug benyttes som regressor i estimationen af den betingede korrelation mellem link-outcome og konsekvens og måles til $t = 0$ (før første indsatsår).
Baseline standardafvigelse	Baseline standardafvigelser angiver (lokale) kommunale standardafvigelser på henholdsvis link-outcomes og konsekvenser på hvert af analysepopulationens klassetrin og dermed til tid $t = 1$.
Betinget korrelation	Lineær statistisk sammenhæng mellem to variable, hvor der kontrolleres for en eller flere baggrundsvARIABLE. Måles for hvert klassetrin.
Default konsekvenser	Konsekvenser, som med rimelighed kan forventes at påvirkes af en generel skolebaseret indsats og derfor er valgt på forhånd for brugeren i beregningsmodulet.
Depreciering	Et mål, der beskriver, hvor meget en indsats' effektivitet aftager ved gentagelse. Antages i udgangspunktet at være 0 i mangel på empiri.
Direkte evidens	Det tilfælde, hvor virkningen af en indsats er målt direkte på en eller flere konsekvenser.
Indirekte evidens	Det tilfælde, hvor virkningen af en indsats er målt via et link-outcome. Sammenhængen mellem link-outcome og konsekvens kan hentes i litteraturen eller estimeres baseret på de administrative registre hos Danmarks Statistik.
Indsats	En handling, der ændrer status quo og giver anledning til at forvente økonomiske konsekvenser over en 4-årig periode.
Indsatsvariant	Forskellige, men dog sammenlignelige udformninger af en given indsats.
Ekstensiv margin	Indsatsens omfang med hensyn til varighed, klassetrin og antal elever.
Helårseffekt	Effekt fra litteraturen omregnet (lineært) til effekt ved fuld implementering i et år.
Intensiv margin	Intensitet, med hvilken en indsats implementeres i løbet af skoleåret.
Konsekvens, konsekvensforbrug, konsekvensmål	Et udfald, til hvilken der kan knyttes økonomi, og som er relevant i statslig, regional eller kommunal budgetmæssig sammenhæng. Dækker forbrug af offentlige ydelser, services og aktiviteter.
Konsekvens-estimat	Estimatet på den betingede korrelation mellem link-outcome og konsekvens.
Link-outcome	Et udfald, der påvirkes af indsatsen og benyttes i litteraturen, men til hvilken der ikke kan knyttes økonomi. Bruges som bindeled mellem indsats og konsekvens.
Lokalt niveau	De parametre, der indgår i beregningen, måles på kommunalt niveau.
Outcome	Se konsekvens.
Persistens	Et mål for, hvor vedholdende effekten af en indsats kan forventes at være over tid. Bestemmes empirisk som korrelationen mellem link-outcome i år $t = 1$ og konsekvens i år $t = 2, 3, 4$.
Pris	Enhedsprisen, der knytter sig til konsekvenserne. Benyttes til at knytte de estimerede indsats effekter på forbrug af konsekvenser til økonomi.
Timing	Tidsmæssig skævhed mellem indsatsperiode og budgetperiode korrigeres med en korrektionsfaktor. Default i modellen er opstart af indsats efter sommerferien frem for ved budgetårets start, da indsats på skoleområdet typisk følger skoleåret. Korrektionsfaktoren er 0,455.

Bilag 2 Versioner af R-pakker

```
sessionInfo()
R version 4.1.2 (2021-11-01)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 18363)

Matrix products: default

locale:
 [1] LC_COLLATE=Danish_Denmark.1252  LC_CTYPE=Danish_Denmark.1252    LC_MONETARY=Danish_Denmark.1252 LC_NUMERIC=C                    LC_TIME=Danish_Denmark.1252

attached base packages:
 [1] stats      graphics  grDevices  utils      datasets  methods   base

other attached packages:
 [1] shinybusy_0.2.2 shinyjs_2.0.0 scales_1.1.1 forcats_0.5.1 stringr_1.4.0 dplyr_1.0.7 purrr_0.3.4 readr_2.0.2 tidyr_1.1.4 tibble_3.1.5 ggplot2_3.3.5
 [12] tidyverse_1.3.1 shiny_1.7.1

loaded via a namespace (and not attached):
 [1] tidymodels_1.1.1 haven_2.4.3 colorspace_2.0-2 vctrs_0.3.8 generics_0.1.1 htmltools_0.5.2 utf8_1.2.2 rlang_0.4.12 later_1.3.0
 [10] pillar_1.6.4 glue_1.4.2 withr_2.4.2 DBI_1.1.1 dbplyr_2.1.1 modelr_0.1.8 readxl_1.3.1 lifecycle_1.0.1 munsell_0.5.0
 [19] gtable_0.3.0 cellranger_1.1.0 rvest_1.0.2 htmlwidgets_1.5.4 tzdb_0.2.0 fastmap_1.1.0 httpuv_1.6.3 fansi_0.5.0 broom_0.7.10
 [28] Rcpp_1.0.7 xtable_1.8-4 promises_1.2.0.1 backports_1.3.0 jsonlite_1.7.2 mime_0.12 fs_1.5.0 hms_1.1.1 digest_0.6.28
 [37] stringi_1.7.5 grid_4.1.2 cli_3.1.0 tools_4.1.2 magrittr_2.0.1 crayon_1.4.2 pkgconfig_2.0.3 ellipsis_0.3.2 xml2_1.3.2
 [46] reprex_2.0.1 lubridate_1.8.0 rstudioapi_0.13 assertthat_0.2.1 httr_1.4.2 R6_2.5.1 compiler_4.1.2
```

Anm.: REFUD-prototypen kan som minimum afvikles i et R-miljø bestående af disse pakker.

Bilag 3 Kodestruktur til databearbejdelse på Danmarks Statistiks forskermaskine

```

/*****/
/* PROGRAMNAVN      : Master-fil.sas
/* PROJEKT         : REFUD
/* FORMÅL          : Master file, der guider bearbejdning af grunddata og elevramme og kobling med link-outcomes
og konsekvenser/forbrug af ydelser/aktiviteter. Derudover også beregning af enhedspriser for 0-17-årige.
Denne fil beskriver strukturen og rækkefølgen af programmerne på SAS-serveren, og kan desuden
køre programmerne direkte.
/* OPGØRELSESÅR    : Fastsættes ved makro_aar (se PREAMBLE)
/*****/
/* Programmet baserer sig på et program oprettet af VLHA fra Socialstyrelsen i forbindelse
/* med oversættelsen af Socialstyrelsens SØM til SAS fra VIVES oprindelige Stata-baserede model.
/* Programmet er udvidet og tilpasset REFUD-universet.
/*****/

/* Program oprettet af      : ANNA
/* Oprettelsesdato        : 29-11-2021
/* Opdatering             :
/*****/

/*****
OBS! for at mindske datasættens størrelse, sættes variablene så vidt muligt til at have length 3.
Dette er dog IKKE muligt for variable, der opgøres med decimaler, eller som har værdier, der overstiger
ca. 8000. Hvis opgørelsen af variabelen ændres, kan det derfor være nødvendigt at ændre dette!
/*****

/*****
ROAD MAP:

1) PREAMBLE
Indlæser data- og workdatastier og formater samt makro, der fastsætter modellens dataår.
*** OBS: DENNE SKAL KØRES FØR PROGRAMMERNE KAN KØRES SÆRSKILT, HVIS MAN IKKE ØNSKER AT KØRE ALT
DIREKTE GENNEM MASTER-FILEN ***

2) GRUNDDATA:
Danner et balanceret datasæt af alle personer i opgørelsesperioden (GRUND_SORT).
Her dannes et balanceret datasæt af alle personer i opgørelsesperioden (GRUND_SORT). Hver person
optræder med én observation i hvert kalenderår fra og med første gang, vedkommende findes i BEF i
perioden. I datasættet er der koblet baggrundsvariable som fx alder, socioøkonomisk status og
forældre-id.

3) BEARBEJDNING AF WORKDATA:
- Data for link-outcomes og konsekvensforbrug - datasæt navngivet med suffiks "_SORT": Datasæt med
forbrug af ydelser/aktiviteter i hvert kalenderår for hver person, der er registreret med et
forbrug. I enkelte tilfælde dannes desuden kontrolvariable.
- Hjælpedata:
  Datasæt til yderligere kontrolvariable.

4) LINK OUTCOMES: Kobling af grunddatasæt og forbrug af ydelser/aktiviteter
- Forbruget af ydelser/aktiviteter leades. Dvs. der dannes en variabel med forbruget af den
pågældende ydelse/aktivitet for hvert kalenderår efter året for indsats/link-outcome-måling.
Variablene er markeret med et suffiks, som angiver hvornår variablene er målt relativt til
hinanden. År "_0" anvendes altid som indsatsåret. Det er dette år, hvor alle de tidsafhængige
baggrundsvariable er målt. Året "_1" er typisk året, hvor succes/ikke-succes opgøres, og årene
fra og med "_2" og frem anvendes derfor typisk som outcome i regressionerne.

5) POPULATIONSRAMME: Kobling af data fra 4) indeholdende baggrundsdata, link-outcomes og konsekvensforbrug
med indsatspopulation.
Data eksporteres herefter til Stata, hvor beregningsmodulets inputs konstrueres.
Kodestrukturen tillader således samtidigt arbejde med flere forskellige rammer for
indsatspopulationer. Fx skoleelever og børn i dagtilbud særskilt.

6) PRISER: Beregning af enhedspriser (sygesikringskontakter, LPR-kontakter og marginale skattesatser).
Beregnes for 0-17-årige. Marginale skattesatser beregnes vægtet for 13-18-årige samt for populationen
af forældre i elevrammen.
Disse udsendes blot fra forskermaskinen og indsættes direkte i beregningsrammens pris-.JSON-fil.

/*****/

```

```

/*****
1. PREAMBLE: OPRETTER STIER OG INDLESER FORMATER (SKAL KØRES FØR PROGRAMMRNE ENKELTVIS KAN KØRES)
*****/

* Opretter stier til data;
libname DST "D:\Data\rawdata\708135\Grunddata"; /* Rawdata fra DST */
libname ekstern "D:\Data\rawdata\708135\Eksterne data"; /* Rawdata fra SDS, STIL */
libname workdata "D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\Workdata\"; /*
Workdata */

* Indlæs formatbibliotek;
libname fmt "\\srvfsenas1\data\formater\SAS formater i Danmarks Statistik\FORMATKATALOG' access=readonly;

libname udfmt "\\srvfsenas1\data\formater\SAS formater i Danmarks Statistik\SAS_dataaet\Disced";
libname vstfmt "\\srvfsenas1\data\formater\SAS formater i Danmarks Statistik\SAS_dataaet\Statistikbank";

* Opretter pseudo-makroer til stier;
%let CodePathMacro = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\1. Macros;
*Makroer der anvendes flere gange i programmer + opgørelsesår og specielle navne på datasæt;
%let CodePathProg2 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\2. Grunddata;
%let CodePathProg3 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\3. Workdata bearbejdning;
%let CodePathProg4 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\4. Flet link-outcomes og
outcomes;
%let CodePathProg5 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\5. Populationsramme;
%let CodePathProg6 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\6. Priser;

%let CodePathOutput= D:\Data\workdata\708135\BUVM_SOM\Output\;

%include "&CodePathMacro.\Macros_aar.sas"; *Opretter makroer for dataår og -navne.
***OBS: RET HER for tilgængeligt data;

/*****
2. DANNER GRUNDDATASÆT (GRUND_SORT)
*****/
%include "&CodePathProg2.\GRUND.sas"; ***OBS: SKAL OPDATERES MED BUVMs ADGANG TIL REGISTRE I GRUNDVIVE (BEF,
DOD m.fl.);

/*****
3. DANNER DATASÆT MED LINK-OUTCOMES, FORBRUG AF YDELSER/AKTIVITETER (*_SORT) SAMT DIVERSE HJELPEDATASÆT
*****/
*Link-outcomes; ***OBS: INDSÆT HER for dannelse af yderligere link-outcomes***;
%include "&CodePathProg3.\DNT.sas"; *DNT_SORT;

*Konsekvenser;
%include "&CodePathProg3.\DREAM.sas"; *DREAM_OVER_SORT;
%include "&CodePathProg3.\SSSY.sas"; *SSSY, SSSY_SORT;
%include "&CodePathProg3.\LPRPSYK.sas"; *LPR_ADM, LPR_DIAG, LPR_SORT, PSYK_SORT, ANGSTDEP_FIRST,
OPMFORSTYR_FIRST;
%include "&CodePathProg3.\KRAF.sas"; *KRAF1_SORT;
%include "&CodePathProg3.\KRIN.sas"; *KRAF2_SORT;
%include "&CodePathProg3.\NAB.sas"; *ALKOHOL_SORT;
%include "&CodePathProg3.\SIB.sas"; *STOFMISBRUG_SORT;
%include "&CodePathProg3.\BOFORM.sas"; *HERBERG_SORT;
%include "&CodePathProg3.\BUFOBUAF.sas"; *BUFOBUAF_SORT;
%include "&CodePathProg3.\UDSP.sas"; *UDSP_SORT ;
%include "&CodePathProg3.\KOTRE.sas"; *KOTRE_SORT;
%include "&CodePathProg3.\BFL.sas"; *BFL_SORT;
%include "&CodePathProg3.\FRPE.sas"; *FRPE_SORT, laerere_sygefravaer_sort;

*Kontrolvariable;
%include "&CodePathProg3.\MFR.sas"; *MFR;

/*****
4. KOBLER LINK-OUTCOMES OG LEADED FORBRUG AF YDELSER/AKTIVITETER PÅ FULD POPULATION (0-17-årige)
*****/
%include "&CodePathProg4.\LINK_OUTCOMES.sas";

/*****
5. DANNER DATASÆT MED INDSATSPOPULATION OG TILFØJER BAGGRUNDSDATA, LINK-OUTCOMES OG KONSEKVENSDATA
*****/
%include "&CodePathProg5.\SKOLEINDSATS_RAMME.sas";
*INDSÆT HER, hvis modellen udvides ud over 0.-9. klasse;

/*****
6. BEREGNER ENHEDSPRISER
*****/
%include "&CodePathProg6.\PRISER_SSSY.sas"; *Priser på kontakter til sygesikring;
%include "&CodePathProg6.\PRISER_LPR_DRGSOMA.sas"; *Priser på kontakt til somatisk hospital;
%include "&CodePathProg6.\PRISER_MARGSKATINDKOMST.sas"; *Priser på marginalskat af indkomst;

```

Anm.: Master-filen er skrevet i SAS 9.4.

Bilag 4 Kodestruktur til beregning af model-input på Danmarks Statistiks forskermaskine

```
*Generel start-up
clear all
set more off, perm
cap log close
set rmsg on

/*****
*****
STIER
*****
*****/
global dopath "D:\Data\workdata\708135\BUVM_SOM\2. Estimationsprogrammer_Stata\" //Programmer
global datapath "$dopath\Estimationsdata\" //Estimationsdata eksporteret fra SAS
global resultspath "D:\Data\workdata\708135\BUVM_SOM\Output\" //Output til hjemsendelse
cd "$datapath\Temp\" //Midlertidig data

/*****
*****
DATASET: populationsramme koblet med baggrundsdata, link-outcomes og forbrug af
ydelse/aktiviteter.
Navn på datasæt, der indlæses, skal alle slutte på "_linkoutcomes".
I den nedenstående makro medtages kun den resterende del af navnet på datasættet.
*****/
global datalist ///
skoleramme

*Sæt startår
global startaar 2014 //Startår for modellens dataoutput,
//Beholder i udgangspunktet kun år efter skolereformen
/*****
*****/
MODELKARAKTERISTIKA
/*****
*****/
*Sæt link-outcomes til estimation (se links-liste i makro.do)
global links_nr "1 2"

*Sæt konsekvenser til estimation (se fuld konsekvensliste i makro.do)
global konsekvens_nr "1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54"

*Sæt modellens klassetrin
global klassetrin "0 1 2 3 4 5 6 7 8 9"

/*****
*****
SIGNIFIKANSNIVEAU
*****
*****/
*Sæt signifikansniveau
global signifikans = 0.01 //Betingede korrelationer sættes til 0, hvis estimatets
//P-værdi overstiger signifikansniveauet

/*****
*****
PROGRAMMER
*****
*****/
*Makroer for samtlige outcomes samt tilhørende navne og beskrivelser
do "$dopath\Makroer.do" //Makro fastsætter konsekvensliste og linksliste samt klassetrinsafgrænsning for links

*Baseline standardafvigelser på link-outcomes på kommune-niveau
do "$dopath\Linkoutcomes_lokale_sd.do"

*Baseline standardafvigelser på konsekvensforbrug på kommune-niveau
do "$dopath\Konsekvens_lokale_sd.do"

*Betingede korrelationer mellem link-outcomes og konsekvensforbrug i samme år, år 2, år 3 og år 4.
do "$dopath\Korrelation_konsekvens_link_outcome.do"
```

Anm.: Masterfilen er skrevet i Stata 17.

Bilag 5 Versioner af R-pakker

```
sessionInfo()
R version 4.1.2 (2021-11-01)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 18363)

Matrix products: default

locale:
 [1] LC_COLLATE=Danish_Denmark.1252  LC_CTYPE=Danish_Denmark.1252    LC_MONETARY=Danish_Denmark.1252 LC_NUMERIC=C                    LC_TIME=Danish_Denmark.1252

attached base packages:
[1] stats      graphics  grDevices  utils      datasets  methods   base

other attached packages:
 [1] shinybusy_0.2.2 shinyjs_2.0.0  scales_1.1.1  forcats_0.5.1  stringr_1.4.0  dplyr_1.0.7  purrr_0.3.4   readr_2.0.2    tidyr_1.1.4    tibble_3.1.5  ggplot2_3.3.5
[12] tidyverse_1.3.1 shiny_1.7.1

loaded via a namespace (and not attached):
 [1] tidymodels_1.1.1  haven_2.4.3      colorspace_2.0-2  vctrs_0.3.8      generics_0.1.1   htmltools_0.5.2  utf8_1.2.2      rlang_0.4.12     later_1.3.0
[10] pillar_1.6.4      glue_1.4.2       withr_2.4.2       DBI_1.1.1         dbplyr_2.1.1     modelr_0.1.8     readxl_1.3.1    lifecycle_1.0.1  munsell_0.5.0
[19] gtable_0.3.0      cellranger_1.1.0  rvest_1.0.2       htmlwidgets_1.5.4  tzdb_0.2.0       fastmap_1.1.0    httpuv_1.6.3    fansi_0.5.0      broom_0.7.10
[28] Rcpp_1.0.7        xtable_1.8-4     promises_1.2.0.1  backports_1.3.0   jsonlite_1.7.2   mime_0.12        fs_1.5.0        hms_1.1.1        digest_0.6.28
[37] stringi_1.7.5     grid_4.1.2       cli_3.1.0         tools_4.1.2       magrittr_2.0.1   crayon_1.4.2     pkgconfig_2.0.3  ellipsis_0.3.2   xml2_1.3.2
[46] reprex_2.0.1      lubridate_1.8.0  rstudioapi_0.13  assertthat_0.2.1  httr_1.4.2       R6_2.5.1         compiler_4.1.2
```

Anm.: REFUD-prototypen kan som minimum afvikles i et R-miljø bestående af disse pakker.

Bilag 6 Kodestruktur til databearbejdelse på Danmarks Statistiks forskermaskine

```

/*****
/* PROGRAMNAVN      : Master-fil.sas
/* PROJEKT          : REFUD
/* FORMÅL           : Master file, der guider bearbejdning af grunddata og elevramme og kobling med link-outcomes
og konsekvenser/forbrug af ydelser/aktiviteter. Derudover også beregning af enhedspriser for 0-17-årige.
Denne fil beskriver strukturen og rækkefølgen af programmerne på SAS-serveren, og kan desuden
køre programmerne direkte.
/* OPGØRELSESÅR    : Fastsættes ved makro_aar (se PREAMBLE)
/*****
/* Programmet baserer sig på et program oprettet af VLHA fra Socialstyrelsen i forbindelse
/* med oversættelsen af Socialstyrelsens SØM til SAS fra VIVES oprindelige Stata-baserede model.
/* Programmet er udvidet og tilpasset REFUD-universet.
/*****

/* Program oprettet af      : ANNA
/* Oprettelsesdato         : 29-11-2021
/* Opdatering              :
/*****

/*****
OBS! for at mindske datasættens størrelse, sættes variablene så vidt muligt til at have length 3.
Dette er dog IKKE muligt for variable, der opgøres med decimaler, eller som har værdier, der overstiger
ca. 8000. Hvis opgørelsen af variabelen ændres, kan det derfor være nødvendigt at ændre dette!
/*****

/*****
ROAD MAP:

1) PREAMBLE
   Indlæser data- og workdatastier og formater samt makro, der fastsætter modellens dataår.
   *** OBS: DENNE SKAL KØRES FØR PROGRAMMERNE KAN KØRES SÆRSKILT, HVIS MAN IKKE ØNSKER AT KØRE ALT
   DIREKTE GENNEM MASTER-FILEN ***
2) GRUNDDATA:
   Danner et balanceret datasæt af alle personer i opgørelsesperioden (GRUND_SORT).
   Her dannes et balanceret datasæt af alle personer i opgørelsesperioden (GRUND_SORT). Hver person
   optræder med én observation i hvert kalenderår fra og med første gang, vedkommende findes i BEF i
   perioden. I datasættet er der koblet baggrundsvariable som fx alder, socioøkonomisk status og
   forældre-id.
3) BEARBEJDNING AF WORKDATA:
   - Data for link-outcomes og konsekvensforbrug - datasæt navngivet med suffiks " SORT": Datasæt med
   forbrug af ydelser/aktiviteter i hvert kalenderår for hver person, der er registreret med et
   forbrug. I enkelte tilfælde dannes desuden kontrolvariable.
   - Hjælpedata:
     Datasæt til yderligere kontrolvariable.
4) LINK OUTCOMES: Kobling af grunddatasæt og forbrug af ydelser/aktiviteter
   - Forbruget af ydelser/aktiviteter ledes. Dvs. der dannes en variabel med forbruget af den
   pågældende ydelse/aktivitet for hvert kalenderår efter året for indsats/link-outcome-måling.
   Variablene er markeret med et suffiks, som angiver hvornår variablene er målt relativt til
   hinanden. År "_0" anvendes altid som indsatsåret. Det er dette år, hvor alle de tidsafhængige
   baggrundsvariable er målt. Året "_1" er typisk året, hvor succes/ikke-succes opgøres, og årene
   fra og med "_2" og frem anvendes derfor typisk som outcome i regressionerne.
5) POPULATIONSRAMME: Kobling af data fra 4) indeholdende baggrundsdata, link-outcomes og konsekvensforbrug
   med indsatspopulation.
   Data eksporteres herefter til Stata, hvor beregningsmodulets inputs konstrueres.
   Kodestrukturen tillader således samtidigt arbejde med flere forskellige rammer for
   indsatspopulationer. Fx skoleelever og børn i dagtilbud særskilt.
6) PRISER: Beregning af enhedspriser (sygesikringskontakter, LPR-kontakter og marginale skattesatser).
   Beregnes for 0-17-årige. Marginale skattesatser beregnes vægtet for 13-18-årige samt for populationen
   af forældre i elevrammen.
   Disse udsendes blot fra forskermaskinen og indsættes direkte i beregningsrammens pris-.JSON-fil.

/*****

```

```

/*****
1. PREAMBLE: OPRETTER STIER OG INDLÆSER FORMATER (SKAL KØRES FØR PROGRAMMRNE ENKELTVIS KAN KØRES)
*****/

* Opretter stier til data;
libname DST "D:\Data\rawdata\708135\Grunddata"; /* Rawdata fra DST */
libname ekstern "D:\Data\rawdata\708135\Eksterne data"; /* Rawdata fra SDS, STIL */
libname workdata "D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\Workdata\"; /*
Workdata */

* Indlæs formatbibliotek;
libname fmt "\\srvfsenas1\data\formater\SAS formater i Danmarks Statistik\FORMATKATALOG' access=readonly;

libname uddfmt "\\srvfsenas1\data\formater\SAS formater i Danmarks Statistik\SAS_dataaet\Disced";
libname vstfmt "\\srvfsenas1\data\formater\SAS formater i Danmarks Statistik\SAS_dataaet\Statistikbank";

* Opretter pseudo-makroer til stier;
%let CodePathMacro = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\1. Macros;
*Makroer der anvendes flere gange i programmer + opgørelsesår og specielle navne på datasæt;
%let CodePathProg2 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\2. Grunddata;
%let CodePathProg3 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\3. Workdata bearbejdning;
%let CodePathProg4 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\4. Flet link-outcomes og
outcomes;
%let CodePathProg5 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\5. Populationsramme;
%let CodePathProg6 = D:\Data\workdata\708135\BUVM_SOM\1. Databearbejdning_SAS\6. Priser;

%let CodePathOutput= D:\Data\workdata\708135\BUVM_SOM\Output\;

%include "&CodePathMacro.\Macros_aar.sas"; *Opretter makroer for dataår og -navne.
***OBS: RET HER for tilgængeligt data;

/*****
2. DANNER GRUNDDATASÆT (GRUND_SORT)
*****/
%include "&CodePathProg2.\GRUND.sas"; ***OBS: SKAL OPDATERES MED BUVMS ADGANG TIL REGISTRE I GRUNDTVIVE (BEF,
DOD m.fl.);

/*****
3. DANNER DATASÆT MED LINK-OUTCOMES, FORBRUG AF YDELSER/AKTIVITETER (*_SORT) SAMT DIVERSE HJÆLPEDATASÆT
*****/
*Link-outcomes; ***OBS: INDSÆT HER for dannelse af yderligere link-outcomes***;
%include "&CodePathProg3.\DNT.sas"; *DNT_SORT;

*Konsekvenser;
%include "&CodePathProg3.\DREAM.sas"; *DREAM_OVER_SORT;
%include "&CodePathProg3.\SSSY.sas"; *SSSY, SSSY_SORT;
%include "&CodePathProg3.\LPRPSYK.sas"; *LPR_ADM, LPR_DIAG, LPR_SORT, PSYK_SORT, ANGSTDEP_FIRST,
OPMFØRSTYR_FIRST;
%include "&CodePathProg3.\KRAF.sas"; *KRAF1_SORT;
%include "&CodePathProg3.\KRIN.sas"; *KRAF2_SORT;
%include "&CodePathProg3.\NAB.sas"; *ALKOHOL_SORT;
%include "&CodePathProg3.\SIB.sas"; *STOFMISBRUG_SORT;
%include "&CodePathProg3.\BOFORM.sas"; *HERBERG_SORT;
%include "&CodePathProg3.\BUFOBUAF.sas"; *BUFOBUAF_SORT;
%include "&CodePathProg3.\UDSP.sas"; *UDSP_SORT;
%include "&CodePathProg3.\KOTRE.sas"; *KOTRE_SORT;
%include "&CodePathProg3.\BFL.sas"; *BFL_SORT;
%include "&CodePathProg3.\FRPE.sas"; *FRPE_SORT, laerere_sygefravaer_sort;

*Kontrolvariable;
%include "&CodePathProg3.\MFR.sas"; *MFR;

/*****
4. KOBLER LINK-OUTCOMES OG LEADED FORBRUG AF YDELSER/AKTIVITETER PÅ FULD POPULATION (0-17-årige)
*****/
%include "&CodePathProg4.\LINK_OUTCOMES.sas";

/*****
5. DANNER DATASÆT MED INDSATSPOPULATION OG TILFØJER BAGGRUNDSDATA, LINK-OUTCOMES OG KONSEKVENSDATA
*****/
%include "&CodePathProg5.\SKOLEINDSATS_RAMME.sas";
*INDSÆT HER, hvis modellen udvides ud over 0.-9. klasse;

/*****
6. BEREGNER ENHEDSPRISER
*****/
%include "&CodePathProg6.\PRISER_SSSY.sas"; *Priser på kontakter til sygesikring;
%include "&CodePathProg6.\PRISER_LPR_DRGSOMA.sas"; *Priser på kontakt til somatisk hospital;
%include "&CodePathProg6.\PRISER_MARGSKATINDKOMST.sas"; *Priser på marginalsikat af indkomst;

```

Anm.: Master-filen er skrevet i SAS 9.4.

Bilag 7 Kodestruktur til beregning af model-input på Danmarks Statistiks forskermaskine

```
*General start-up
clear all
set more off, perm
cap log close
set rmsg on

/*****
*****
STIER
*****
*****/
global dopath "D:\Data\workdata\708135\BUVM_SOM\2. Estimationsprogrammer_Stata" //Programmer
global datapath "$dopath\Estimationsdata" //Estimationsdata eksporteret fra SAS
global resultspath "D:\Data\workdata\708135\BUVM_SOM\Output\" //Output til hjemsendelse
cd "$datapath\Temp\" //Midlertidig data

/*****
*****
DATASÆT: populationsramme koblet med baggrundsdata, link-outcomes og forbrug af
ydelse/aktiviteter.
Navn på datasæt, der indlæses, skal alle slutte på "_linkoutcomes".
I den nedenstående makro medtages kun den resterende del af navnet på datasættet.
*****
*****/
global datalist ///
skoleramme

*Sæt startår
global startaar 2014 //Startår for modellens dataoutput,
//Beholder i udgangspunktet kun år efter skolereformen
/*****
*****/
MODELKARAKTERISTIKA
*****
*****/
*Sæt link-outcomes til estimation (se links-liste i makro.do)
global links_nr "1 2"

*Sæt konsekvenser til estimation (se fuld konsekvensliste i makro.do)
global konsekvens_nr "1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54"

*Sæt modellens klassetrin
global klassetrin "0 1 2 3 4 5 6 7 8 9"

/*****
*****
SIGNIFIKANSNIVEAU
*****
*****/
*Sæt signifikansniveau
global signifikans = 0.01 //Betingede korrelationer sættes til 0, hvis estimatets
//P-værdi overstiger signifikansniveauet

/*****
*****
PROGRAMMER
*****
*****/
*Makroer for samtlige outcomes samt tilhørende navne og beskrivelser
do "$dopath\Makroer.do" //Makro fastsætter konsekvensliste og linksliste samt klassetrinsafgrænsning for links

*Baseline standardafvigelser på link-outcomes på kommune-niveau
do "$dopath\Linkoutcomes_lokale_sd.do"

*Baseline standardafvigelser på konsekvensforbrug på kommune-niveau
do "$dopath\Konsekvens_lokale_sd.do"

*Betingede korrelationer mellem link-outcomes og konsekvensforbrug i samme år, år 2, år 3 og år 4.
do "$dopath\Korrelation_konsekvens_link_outcome.do"
```

Anm.: Masterfilen er skrevet i Stata 17.

VIVÉ