



**BØRNE- OG
UNDERVISNINGS-
MINISTERIET**
STYRELSEN
FOR IT OG LÆRING

Styrelsen for It og Læring
Kontor for Data på
Ungdomsuddannelser og
VEU

Lyseng Allé 1
8270 Højbjerg
Tlf. nr.: 89 37 66 66
E-mail: stil@stil.dk
www.stil.dk
CVR-nr.: 13223459

Metodenotat om beregningsgrundlaget bag de estimerede konsekvensberegninger i forbindelse med udmeldelsen af de beregningstekniske kapaciteter

10. oktober 2022
Sagsnr.: 22/19308

I forbindelse med forberedelserne til forhandlinger om en ny elevfordelingsmodel på gymnasieområdet har Børne- og Undervisningsministeriet udviklet en konsekvensberegningsmodel, som estimerer de forventede konsekvenser af en ny elevfordelingsmodel. Den samme konsekvensberegningsmodel er anvendt til at estimere de forventede konsekvenser af en elevfordelingsmodel med de beregningstekniske kapaciteter. Dette notat gennemgår metoden og beregningsgrundlaget bag konsekvensberegningsmodellen.

Det skal bemærkes, at konsekvensberegningsmodellen er en forsimplet og stiliseret udgave af den aftalte model – Den Koordinerede Tilmelding til Gymnasiale Ungdomsuddannelser. Konsekvensberegningsmodellen adskiller sig således på en række punkter fra den algoritme, som vil blive anvendt i implementeringen af aftalen. Det betyder, at konsekvensberegningerne er omfattet af en lang række beregningsusikkerheder, jf. nedenstående afsnit om samme.

Konsekvensberegningsmodellen

Konsekvensberegningsmodellen baserer sig på et analytisk fordelingsværktøj (algoritme), der simulerer en forenklet udgave af ny fordeling med nye regler af historiske ansøgere på individniveau. Dernæst foretages der en sammenligning til fordelingen af ansøgere i dag ved de gældende regler.

Selve fordelingsværktøjet er baseret på en fordelingsalgoritme, som tager afsæt i en såkaldt *controlled school choice model* (Abdulkadirouglu and Sönmez, 2003)¹, som er karakteriseret ved, at der kan sikres en mere blandet

¹ Abdulkadirouglu and Sönmez (2003): School Choice: A Mechanism Design Approach

elevsammensætning via kvoter for forskellige elevgrupper. Fordelingsværktøjet bygger på, at institutioner i bestemte geografiske områder (fordelingszoner) skal afspejle området ift. forældreindkomst, samt at fordelingen sker efter fordelingsmekanismen *priority matching* på 1-3. prioriteten og efter *deferred acceptance* på efterfølgende prioriteter.

Beregningsusikkerheder

Konsekvensberegningerne beror på en række beregningstekniske antagelser og simplificeringer, som medfører en betydelig beregningsusikkerhed. De væsentligste årsager til beregningsusikkerheden er:

- *Forudsat uændret søgeadfærd*, som følge af anvendelsen af historiske søgedata, hvor gældende regler kan have betydning for søgeadfærden. Den nye fordelingsmodel kan afstedkomme ændrede søgemønstre og vil i fordelingszoner som hovedregel sikre, at ansøgere angiver flere prioriteter end i dag. Begge forhold forventes at påvirke resultaterne
- *Særgrupper med forrang* er ikke indarbejdet i modellen. Dette gælder bl.a. profilelever, pre-IB elever, Team Danmark elever og ASF-elever.
- *Krydsøgning til ikke-gymnasiale uddannelser samt private gymnasiale uddannelser* er ikke indarbejdet i modellen
- *Forskel mellem tilgangspopulationen af 1. g'ere i 2021*, som præsenterer situationen i dag, og *ansøgerpopulationen i 2021*, som ligger til grund for den skønnede fordeling af ansøgere givet den nye fordelingsmodel. Disse to elevgrupper kan ikke sammenlignes 1:1, og sammenligninger mellem fordelingen i dag (tilgangen) og fordelingen ved en ny fordelingsmodel (ansøgere) er derfor også behæftet med stor usikkerhed. I de historiske data kan der fx i perioden fra de prioriterede ansøgninger til gymnasial uddannelse er afsendt til start på uddannelsen ske omvalg af uddannelse og/eller institution. Sådanne omvalg indgår ikke i beregning af prioritetsopfyldelsen.
- *Ufordelte ansøgere* er ikke medtaget i beregningerne, da det ikke vides, hvor regionen placerer disse ansøgere.
- *Udvælgelse på baggrund af lodtrækning* ved identisk rejsetid medfører alt andet lige, at fordelingsresultaterne indebærer et vist element af tilfældighed.

Datagrundlag

Datagrundlaget tager udgangspunkt i tre populationer, som er beskrevet i skema 1. Datagrundlaget er derudover baseret på en række definitioner og datavalg, som er udfoldet i skema 2.

Skema 1. Populationer i datagrundlaget

<i>Ansøgerpopulation</i>	<i>Tilgangspopulation</i>	<i>Institutionspopulation</i>
Populationen ligger til grund for konsekvensberegninger af en ny fordeling af ansøgere (algoritmeberegninger).	Populationen ligger til grund for fordelingen ”i dag” under gældende regler (baseline).	Populationen af de institutioner, som ansøgerne fordeles til.
• Rettidige ansøgere fra ”optagelse.dk” med gymnasial uddannelse som 1. prioritet (stx, hf, hhx, htx og pre-IB). Ansøgere til pre-IB omkodes til stx/htx. • Der ses bort fra ansøgninger til private gymnasier eller til ikke-gymnasiale uddannelser (fx eud og 10. klasse). • Ansøgere med uoplyst bopælsoplysninger antages at have samme kommuneoplysninger som institutionen på deres første prioritet. • Der er foretaget korrektioner af data på et mindre antal institutioner for at følge institutionernes struktur for 2021, når der har været uoverensstemmelser i data.	• Den første registrerede elevtilgang i 1.g/1.hf på de gymnasiale institutioner fra skolestart til og med 31. oktober. • Der ses dermed bort fra ansøgere, som ender med at falde fra/søger om orlov frem mod skolestart eller skifter til en ikke-gymnasial uddannelse. • Elever, der tilgår en gymnasial uddannelse uden at søge via ”optagelse.dk”, indgår i populationen, dog ikke på hf. • Baseret på samme uddannelsesvalg som ved ansøgerpopulationen. • Der er foretaget korrektioner af data på et mindre antal institutioner, for at følge institutionernes struktur for 2021, når der har været uoverensstemmelser i data.	• Alle offentlige institutioner², som udbyder uddannelserne stx, hf, hhx, htx og pre-IB indgår i beregningerne. Pre-IB omkodes til stx/htx. • Beregningerne tager udgangspunkt i institutionernes struktur fra 2021.

² Der skal for så vidt angår Sankt Annæ Gymnasium ikke tages hensyn til elevsammensætningen, jf. § 11 i. Baggrunden for det er jf. lovbemærkningerne, at Sankt Annæ optager alle elever som profilelever, dvs. at alle optagne ansøgere har fået forrang til at blive optaget på institutionen, fordi de opfylder institutionens kriterier for at være profilelev. Sankt Annæ Gymnasium indgår derfor ikke i konsekvensberegningerne.

Skema 2. Definitioner og datavalg

Prioritetsopfyldelse	Ansøgerpopulationen (algoritmeberegningerne): • Baseret på ansøgernes faktiske prioriteter fra "optagelse.dk". • I de tilfælde, at der ses bort fra udvalgte ansøgninger, fx fordi det er ansøgninger til eud eller til et privat gymnasium, da rykkes evt. efterfølgende prioriteter op således, at ansøgernes prioriteringsrækkefølge er kontinuer: 1, 2, 3, 4 osv. • I forlængelse af ansøgernes faktiske prioriteter, tildeles ansøgerne yderligere prioriteter. Disse ekstra prioriteter udgøres af institutioner med samme uddannelse, som ansøgerens 1. prioritet, og tildeles efter korteste rejsetid op til maksimalt 45 min. • I beregningerne ses der bort fra ansøgere, der i fordelingssimuleringen ender ufordelt. • Ansøgerens prioritetsopfyldelse bliver opdelt efter den uddannelse, som ansøgeren tildeles plads på. Tilgangspopulationen (beregningerne for "i dag"): • Hvis eleven har ansøgt den tilgæede institution via "optagelse.dk" benyttes prioriteten herfra. • Hvis eleven har søgt via "optagelse.dk", men den tilgæede institution ikke fremgår af elevens prioriteringsliste anmærkes elevens prioritetsopfyldelse med "ej opfyldt". • Hvis eleven ikke har søgt via "optagelse.dk", antages det, at den tilgæet institution er elevens 1. prioritet. Dette gælder dog ikke på hf, som er fjernet fra <i>tilgangspopulationen</i>, jf. skema 1. • Elevens prioritetsopfyldelse bliver opdelt efter den uddannelse, som eleven er tilgæet.
Rejsetid	Individernes rejsetid mellem deres bopæl og en given institution opgøres således: • Rejsetiden bliver opgjort som den korteste rejsetid mellem rejsetider beregnet via offentlig transport og via cykel. • I beregningen via offentlig transport benyttes gang til, mellem og fra offentlige transportmidler. • Rejsetiden opgøres i hele minutter. • For et mindre antal individer kan rejsetiden ikke beregnes. Det kan bl.a. skyldes, at det ikke er muligt at gennemføre rejsen uden supplerede transportmidler inden kl. 8 om morgenen. For en andel af disse er det muligt at approksimere en rejsetid ud fra deres fugleflugtsafstand sammenholdt med andres individers rejsetid og fugleflugtsafstand, som er bosat i samme postnummer/kommune. Opgørelse af beregningerne for individernes rejsetid: • For ansøgere/elever, hvor rejsetiden til den tildelte/tilgæet institution mangler og hvor det ikke var muligt at approksimere rejsetiden, sættes den til 999 minutter, så ansøgerene forsat kan fordeles. • For ansøgerpopulationen (algoritmeberegninger) ses der også bort fra ufordelte i beregningerne.
Ufordelte	I algoritmeberegningerne vil der være tilfælde, hvor ikke alle kan tildeles en plads. Disse benævnes som ufordelte. Ansøgerne ender ufordelt i algoritmen, fordi der ikke er ledige pladser på enten prioriterede institutioner eller på institutioner af samme uddannelse som ansøgerens 1. prioritet inden for 45 minutters rejsetid.

Forældreindkomst	Der henvises til særskilt notat om opgørelse af forældreindkomst: https://www.uvm.dk/-/media/filer/uvm/i-fokus/tema/elevfordeling/210610-notat-om-foraeldreindkomst.pdf .
Ikke-vestlig herkomst	Ikke-vestlig herkomst følger definitioner fra Danmarks Statistik og indeholder indvandrere og efterkommere med et ikke-vestligt oprindelsesland. Hvis oplysningen er uoplyst, indgår individet <i>ikke</i> i kategorien for ikke-vestlig herkomst.
Individernes bopæl	Individernes bopælsoplysninger bliver opgjort pr. 1. marts i ansøgningsåret.
Zoneinddeling	Zoneinddeling hvor landet er inddelt i 28 zoner. Der henvises i øvrigt til særskilt notat om zoner: https://www.uvm.dk/-/media/filer/uvm/i-fokus/tema/elevfordeling/211027-notat-om-beregningstekniske-zoner.pdf .