



13. juni 2022

Metodenotat om beregningsgrundlaget bag aftalen om en ny elevfordelingsmodel på gymnasieområdet – opdateret i forbindelse med tillægsaftalen

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. Dette notat gennemgår metoden og beregningsgrundlaget bag konsekvensberegningssmodellen.

Det skal bemærkes, at konsekvensberegningssmodellen er en forsimplet og stiliseret udgave af den aftalte model – Den Koordinerede Tilmelding til Gymnasiale Ungdomsuddannelser. Konsekvensberegningssmodellen 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.

Konsekvensberegningssmodellen

Konsekvensberegningssmodellen 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 elevsammensætning via kvoter for forskellige elevgrupper. Fordelingsværktøjet bygger på, at institutioner i bestemte geografiske områder (for-

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

delingszoner) 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
- *Beregningsteknisk eksempel på en stram kapacitetsfastsættelse*, som bygger på overordnede og mekanisk fastsatte skøn for at reducere den eksisterende overkapacitet og sikre en minimumskapacitet.
- *Forskel mellem tilgangspopulationen af 1. g'ere i 2020*, som præsenterer situationen i dag, og *ansøgerpopulationen i 2020*, 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
- *Ufordelte ansøgere* er ikke medtaget i beregningerne, da det ikke vides, hvor regionen placerer disse ansøgere. Antallet af ufordelte ansøgere er meget tæt forbundet med kapacitetsfastsættelse, som i det beregningstekniske grundlag er fastsat mekanisk. I den kommende kapacitetsfastsættelse for 2023, der tager udgangspunkt i nærmere analyse, forventes antallet af ufordelte at blive reduceret ift. aftalt model.
- *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.
<ul style="list-style-type: none"> • 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. • Populationen beriges med elever, som tilgår en af uddannelserne i august uden om ”optagelse.dk”. Denne gruppe udgør i høj grad de såkaldte ”eftertilmeldere”. • Der ses bort fra ansøgere, som tilgår et privat gymnasium. • Der ses bort fra ansøgninger til private gymnasier eller til ikke-gymnasiale uddannelser (fx eud og 10. klasse). • Der ses bort fra ansøgere med uoplyste bopælsoplysninger. • Der ses bort fra ansøgninger med manglende rejsetid, <i>jf. uddybning i skema 2.</i> 	<ul style="list-style-type: none"> • Den første registrerede elevtilgang på de gymnasiale institutioner. • 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. • Baseret på samme uddannelsesvalg som ved ansøgerpopulationen. 	<ul style="list-style-type: none"> • Alle offentlige institutioner, som udbyder uddannelserne stx, hf, hhx, htx og pre-IB indgår i beregningerne. Pre-IB omkodes til stx/htx. • Der er foretaget korrektioner af data på et mindre antal institutioner for at kunne opgøre data på afdelingsniveau. • Beregningerne tager udgangspunkt i institutionernes struktur fra 2019.

Skema 2. Definitioner og datavalg

Prioritetsopfyldelse	<p>Ansøgerpopulationen (algoritmeberegningerne):</p> <ul style="list-style-type: none"> • 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. <p>Tilgangspopulationen (beregningerne for ”i dag”):</p> <ul style="list-style-type: none"> • Hvis eleven har ansøgt den tilgæet 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”.
-----------------------------	--

	<ul style="list-style-type: none"> • Hvis eleven ikke har søgt via ”optagelse.dk”, typisk eftertilmelder, antages det, at den tilgæet institution er elevens 1. prioritet. • Elevens prioritetsofyldelse bliver opdelt efter den uddannelse, som eleven er tilgæet.
Rejsetid	<p>Individernes rejsetid mellem deres bopæl og en given institution opgøres således:</p> <ul style="list-style-type: none"> • 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. <p>Opgørelse af beregningerne for individernes rejsetid:</p> <ul style="list-style-type: none"> • Der ses bort fra ansøgere/elever, hvor rejsetiden til den tildelte/tilgæet institution mangler og hvor det ikke var muligt at approksimere rejsetiden. • 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.
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	Beregningstekniske eksempler på en zoneinddeling, hvor landet var inddelt i 28 zoner. Der henvises i øvrigt til særskilt notat om beregningstekniske zoner.