

Mangelfull evaluering av inntaksmodeller

Tommy Odland

Sean Meling Murray

16. april 2020

De mest sentrale bidragene til debatten om inntaksmodeller baserer seg ikke på matematisk og samfunnsøkonomisk forskning som allerede er gjort på denne type problem.

Det pågår en politisk diskusjon om hvorvidt karakterbaserte inntak skal innføres i alle fylker, og Inntaksutvalget i Oslo arbeider med å evaluere ulike modeller. Problemstillingen har lenge vært kjent for matematikere og økonomer (Gale, 1962). I litteraturen beskrives flere modellegenskaper som er ønskelige, uavhengig av politisk ståsted. Vi vil gå gjennom disse i denne artikkelen.

Noen av de mest sentrale bidragene til debatten om inntaksmodeller er en rapport publisert av OsloMet (Serediak, 2020), og et forslag til en progresjonsbasert modell utarbeidet av forskere ved FHI (Flatø, 2019). Arbeidene er interessante, men vurderer i liten grad modellene opp mot kjente egenskaper som effektivitet, stabilitet, strategisikkerhet og insentiver. I denne artikkelen evalueres arbeidet som er gjort opp mot disse egenskapene, og to nye inntaksmodeller presenteres.

Velkjent teori

Oppgaven handler om å velge en modell som plasserer elever på skoler på beste måte. Når alle elever er tildelt skoleplass, sier vi at elevene er *fordelt*. Vi referer også til dette som modellens *fordeling*.

Fra et matematisk perspektiv kan man angripe problemet på flere måter. De vanligste er matematisk optimering og mekanismedesign. Matematisk optimering handler om å maksimere én eller flere målsettinger gitt begrensede ressurser. Mekanismedesign er studiet av hvordan å lage rettfærdige mekanismer for å oppnå en ønsket målsetting. Modeller basert på mekanismedesign antar at elevene lager en prioritert ønskeliste over skolene, og at skolene lager en rangering av elevene. To viktige begreper i denne sammenheng er effektivitet og stabilitet (Abdulkadiroğlu, 2003).

En fordeling er *effektiv* dersom man ikke kan endre fordelingen slik at én eller flere elever blir mer fornøyde med sin plassering, uten at andre blir mindre fornøyde.

En fordeling er *stabil* dersom den ikke fører til misunnelse. Anta for eksempel at Kari er plassert på skole B, men har skole A høyere på ønskelisten. Dersom skole A rangerer Kari høyere enn minst én elev som allerede har fått plass på skole A, sier man at Kari er misunnelig. Hun foretrekker skole A og skole A foretrekker henne, men hun ble plassert på skole B. Dersom det finnes elever som Kari i en fordeling, er den ustabil.

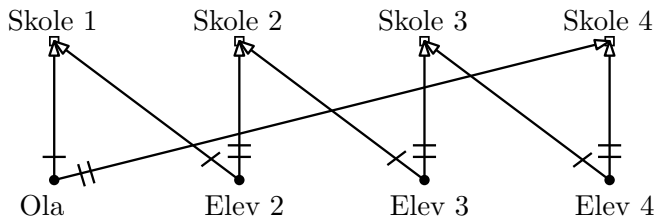
Optimeringsmodeller fordeler elever på skoler etter ett eller flere kriterier (Franklin, 1973). Et naturlig kriterie vil være å maksimere hvor fornøyde elevene er i sum med skoleplassene de er blitt tildelt. I praksis skjer dette ved at modellen får poeng for hver gode plassering den oppnår. Eksempelvis kan man tildele modellen tre poeng for å innfri et førsteønske, to poeng for et andreønske, ett poeng for et tredjeønske og null poeng for skoler som ikke er blant elevens ønsker. Modellen gir en fordeling som maksimerer summen av poengene, og kan implementeres med grafteoretiske algoritmer (Ahujia, 1992). Om ønskelig kan optimeringen være avhengig av elevenes karakterer, slik at modellen til en viss grad prioriterer elever med høyt karaktersnitt.

Det er en viktig prinsipiell forskjell mellom mekanismedesign og optimering. Førstnevnte fordeler én etter én elev, uten nødvendigvis å ta hensyn til alle elevenes ønsker sett under ett. Med andre ord kan modeller basert på mekanismedesign føre til at plasseringen av én elev "ødelegger" for mange andre. Optimeringsmodeller fordeler elever basert på hva som er best for flest mulig. Uformelt kan man si at optimeringsmodeller evner å "se hele bildet". De produserer effektive fordelinger der mange elever blir fornøyde. Prisen man betaler er at optimeringsmodeller stort sett ikke er stabile. Det er viktig å være bevisst på disse egenskapene, fordi de ofte er i konflikt med hverandre (Roth, 1982).

Hva er rettfærdig?

Et viktig grunnprinsipp er at en inntaksmodell skal være mest mulig rettfærdig. Man kan teste sin egen oppfattelse av rettfærdighet i figur 1. Figuren illustrerer et problem der hver skole (representert

ved firkantene øverst) har én ledig plass. Pilene ut fra hver elev (representert ved sirklene nederst) indikerer ønskene til eleven. Linjene som går på tvers av pilene indikerer prioriteten til ønskene til hver elev. Én strek angir førsteønsket og to streker angir andreønsket. Ola har oppgitt Skole 1 som førsteønske og Skole 4 som andreønske. Anta at Ola har karaktersnitt 5,0 og at de andre elevene har karaktersnitt 4,9. Hver skole har kun én ledig plass. Hvilken fordeling er mest rettferdig?



Figur 1: Figuren illustrerer et problem der hver skole (representert ved firkantene øverst) har én ledig plass. Pilene ut fra hver elev (representert ved sirklene nederst) indikerer ønskene til eleven. Linjene som går på tvers av pilene indikerer prioriteten til ønskene til hver elev. Én strek er førsteønsket, to streker er andreønsket, og så videre.

En karakterbasert modell basert på mekanisme-design vil gi Ola førsteønsket, men på bekostning av resten av elevene, som ikke vil få innfridd sine førsteønsker. En optimeringsmodell evner å se helheten. Den kan gi Ola sitt andreønske slik at alle andre elever får innfridd sine førsteønsker.

Noen vil foretrekke å alltid innfri Olas førstevalg, fordi han har høyest karaktersnitt. Andre vil maksimere antallet elever som blir fornøyde, og derfor se helt bort fra karakterer. Et kompromiss er også mulig om man tildeler optimeringsmodellen flere poeng for å innfri ønskene til elever med høyere karaktersnitt. Ola vil da prioriteres foran de andre dersom han har høyt karaktersnitt *i forhold til* de andre. Men hvor gode karakterer må Ola ha for at modellen skal prioritere ham fremfor tre andre elever? Dette er et politisk spørsmål, men optimeringsmodellen er åpen for hele spekteret av meninger.

Hva ønsker man fra en modell?

Uavhengig av politisk ståsted, bør alle som jobber med innstaksmodeller være enige om noen grunnprinsipper.

Elevene bør, i den grad det er mulig, spares for å måtte tenke strategisk når de søker skoleplass. En modell er *strategisikker* dersom det ikke er fordelaktig for elever å oppgi usanne ønsker (Abdulkadiroğlu, 2003). Ingen av dagens karakterbaserte modeller, og ingen av modellene evaluert av Inntaksutvalget,

er strategisikre. Optimeringsmodeller er heller ikke strategisikre, men i praksis er de vanskelige å manipulere.

I hvilken grad karakterer skal påvirke inntaket er et politisk spørsmål. Likevel bør perverse incentiver som aktivt belønner dårlig innsats på skolen unngås. Senere i artikkelen viser vi at kvotering ofte fører til slike incentiver.

På et overordnet nivå er det fordelaktig om det er enkelt å forstå hva modellens definisjon av rettferdighet er. Dette betyr ikke at detaljene i beregningene trenger å være enkle. Et ønske om å unngå komplisert matematikk er ikke en unnskyldning for å implementere en modell med negative egenskaper som ellers kunne vært unngått. En modell bør være matematisk forankret; man må kunne bevise at modellen innehar egenskapene man ønsker.

Mangelfull evaluering

OsloMet har på oppdrag av Inntaksutvalget i Oslo kommune analysert flere inntaksmodeller (Serediak, 2020). Fokuset er på statistiske forskjeller når modellene anvendes på søkermassen i 2018. Resultatene er til en viss grad interessante, men det er uheldig at modellenes egenskaper ikke er viet mer plass i rapporten. Statistisk analyse er lite relevant for modeller med dårlige egenskaper. I det følgende går vi kort gjennom modellene evaluert av OsloMet, med fokus på stabilitet, effektivitet, strategisikkerhet og incentiver.

Karakterbasert inntak

Elevene rangeres etter karaktersnitt og blir plassert i denne rekkefølgen. Modellen er effektiv og stabil. Den incentiverer høyt karaktersnitt. Enhver elev blir plassert uten hensyn til hvordan det påvirker elevene som plasseres senere. Mange grupper av elever vil ha identisk karaktersnitt, og fordelingen blir urettferdig om disse elevene rangeres tilfeldig.

Loddtrekning

Elevene plasseres én etter én på skoler i rekkefølge gitt av en loddtrekning. Modellen er effektiv, og incentiverer verken til god eller dårlig innsats på skolen. Det rapporteres at loddtrekning fører til lav segregering i Oslo-skolen, men dette er utelukkende et resultat av tilfeldigheten i loddtrekningen. Ulike loddtrekninger vil gi ulike resultater, som i varierende grad reduserer segregering og innfrir elevenes ønsker (se figur 3). Loddtrekning vil i unødvendig stor grad ofre elevenes ønsker for å oppnå lav segregering.

Kvotering av elever med lavt snitt

De 20% av elevene med lavest karaktersnitt plasseres først, deretter plasseres de resterende 80% rangert etter karaktersnitt. Under denne modellen har 20% av elevene insentiver til å prestere dårlig for å forbli i kvoten. Elever som ligger på grensen kan bevisst prestere dårlig for å komme inn i kvoten. Modellen kan oppfattes som urettferdig av de resterende 80%.

Prioritetsbasert inntak

Elevenes førsteønsker blir fordelt først, deretter andreønsker, osv. Modellen er aldri strategisikker. En strategisk elev vil spekulere i andre elevers ønsker og utfallet av modellen, fordi det lønner seg å kun oppgi skoler der man forventer at man vil komme inn. Fordi modellen fører til at elever ikke oppgir samme ønsker, er den verken stabil eller effektiv med hensyn til elevenes reelle ønsker. Modellen har vært brukt i Boston i USA, men ble vraket fordi den oppfordrer sterkt til strategisk søking (Abdulkadiroğlu, 2003).

Ingen av modellene ovenfor er strategisikre når elevene kun får oppgi et begrenset antall ønsker. Jo flere ønsker som er mulige å oppgi, dess mindre strategisk tenkning kreves. Dette kan forklares med utgangspunkt i en karakterbasert modell. Noen elever har ikke høyt nok karaktersnitt til å komme inn på populære skoler, men ønsker å gå der. Elever som oppgir reelle ønsker risikerer å ikke komme inn og dermed bli tilfeldig plassert på en skole. Elever som nedjusterer forventningene sine ved å oppgi usanne ønsker får større kontroll over hvor de blir plassert.

Svak progresjonsmodell

I utredningen til ekspertutvalget om kjønnsforskjeller i skoleprestasjoner (NOU 2019:3) skisseres en modell basert på elevers faglige progresjon. Modellen beskrives mer detaljert i (Flatø, 2019). I korte trekk deles elevene inn i mestringsnivåer basert på resultater fra 8. trinn, og konkurrerer om inntak med karakterene fra 10. trinn innen sitt tildelte mestringsnivå. Hvert mestringsnivå gis en kvote av inntaket på hver skole. Modellen kan virke forlokkende fordi den fremstår som mer kompleks enn de ovennevnte, men den har flere svært uheldige egenskaper.

Progresjonsmodellen skiller skarpt mellom en elevs faglige utvikling før og etter 8. klasse. Modellen antar at faglige evner i 8. klasse er et resultat av flaks i livslotteriet (foreldrenes utdanningsnivå, elevens evner, osv.). Etter 8. klasse er faglig progresjon et resultat av elevens ambisjon. Basert

på disse antakelsene hevdes det at modellen er ressursnøytral og ambisjonssensitiv. Det skarpe skillet i årsaksbestemmelsen til faglige prestasjoner er en forenkling av en kompleks problemstilling. Vi stiller oss undrende til påstanden om at modellen vil forene ideologiske motsetninger; modellen tar et standpunkt om at ambisjon og innsats er viktigere enn evner og ressurser. I det følgende ser vi på noen av de negative egenskapene til den progresjonsbaserte modellen.

Anta at Ola er i mestringsnivå 1 og har skole A som førsteønske. Kari er i mestringsnivå 2 og har skole B som førsteønske. Dersom skole A har fylt kvoten for mestringsnivå 1 og skole B har fylt kvoten for mestringsnivå 2, vil både Ola og Kari havne på skoler de ikke ønsker. For Ola og Kari er det ingen trøst at modellen er ressursnøytral og ambisjonssensitiv; de kunne byttet skole og begge ville vært mer fornøyde. Sagt med andre ord: Modellen er ikke effektiv.

Modellen gir alle elever i 8. klasse insentiv til å aktivt prestere dårlig. Dette anerkjenner forfatterne. De foreslår derfor at et datasystem skal avdekke underprestasjon, og delegerer i praksis problemet videre til en ny, uspesifisert modell. Det fremstår som urealistisk at et datasystem skal klare å avdekke strategisk underprestering utover rimelig tvil. Dette reiser også et viktig prinsipielt spørsmål: Kan man straffe elever fordi de underpresterer?

Oppsummert er progresjonsmodellen ikke effektiv, den har sterke negative insentiver og den krever et komplekst og hittil ukjent datasystem som kan avdekke juks. I praksis straffer den også elever som tilfeldigvis havner nederst i et gitt mestringsnivå, noe forfatterne selv poengterer.

To nye modeller som reduserer segregering

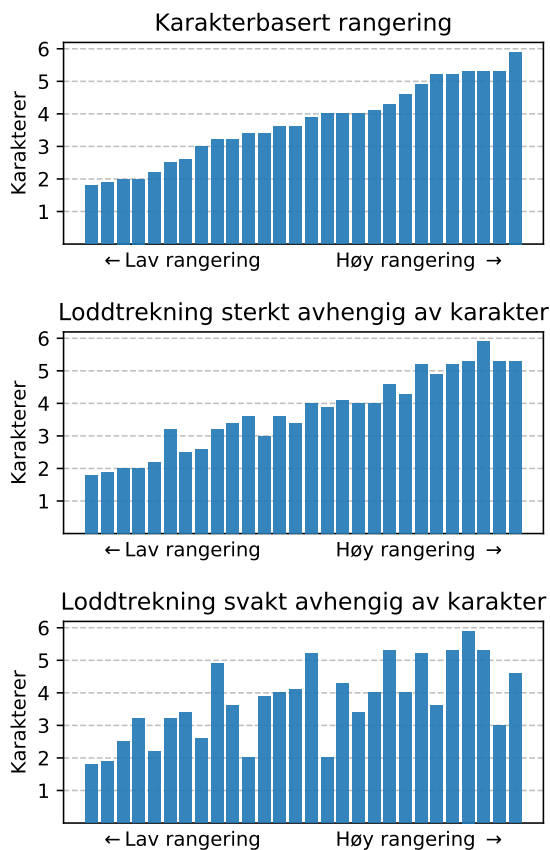
Et av målene ved å utrede alternative inntaksmodeller er å redusere segregering i Oslo-skolen (Serediak, 2020). Her er det flere utfordringer. Segregering er ikke entydig definert i rapporten, de foreslåtte modellene har betenkelige insentiver og ingen av modellene optimerer direkte for lavere segregering.

Nedenfor presenterer vi to nye modeller som reduserer segregering. Disse modellene unngår perverse insentiver og kan tilpasses ulike politiske målsettinger.

Vektet loddtrekning gir gode insentiver

I stedet for en ren loddtrekning kan man foreta en vektet loddtrekning, der vinnere sjansene er avhengig av karaktersnitt. En parameter angir i hvilken grad vinnere sjansene påvirkes av karakterer (se figur 2). Selve parametervalget er en politisk beslutning.

Vektet loddtrekning insentiverer alltid innsats, fordi gode karakterer hjelper enhver elev. Samtidig vil elever med lave karakterer i snitt bli løftet opp på den nye rangeringen. Segregering vil reduseres som følge av tilfeldigheten i den vektete loddtrekningen.



Figur 2: Rangeringen av elevene endres under vektet loddtrekning. Den øverste figuren viser rangering basert på karakterer. Den midterste figuren viser en rangering fra en vektning loddtrekning som er sterkt avhengig av karakterer. Nederst ser man en rangering som i mindre grad er avhengig av karakterer.

Optimering reduserer segregering

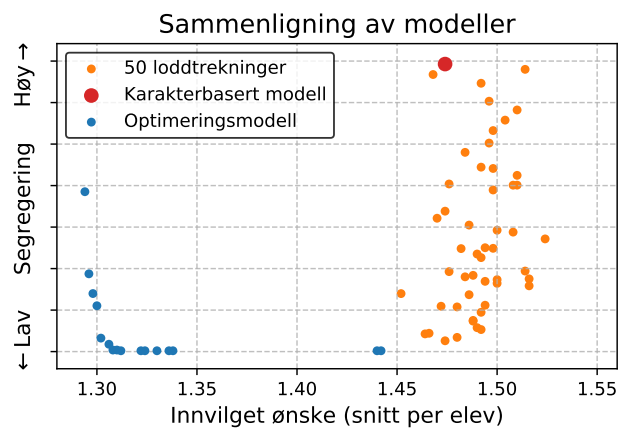
Ved å presist kvantifisere segregering kan man formalisere problemet som et matematisk optimeringsproblem. Algoritmer for multi-objektiv optimering kan maksimere innfridde ønsker samtidig som de minimerer segregering. Segregering kan eksempelvis defineres som variasjonen i gjennomsnittskarakterene i skolene. Er variansen høy, kan det tolkes som høy grad av segregering i skolene. Andre mål på segregering er mulig, men denne definisjonen vil eliminere såkalte A- og B-skoler uten at det innfører perverse insentiver.

Fordelingen som minimerer segregering er ikke nødvendigvis den som gjør elevene mest fornøyd, og målsettingene vil være i konflikt med hverandre. Dette vil gjøre seg særlig gjeldende om man

tilegner mer vekt til elever med høyere karaktersnitt. Optimeringsmodellen gir derfor flere løsningsforslag til hvordan å fordele elever. Hvert løsningsforslag representerer den minste mengden av elevenes ønsker som må ofres for en gitt reduksjon av segregering (se figur 3). Dette er i kontrast til modeller som bruker kvoter og tilfeldighet i håp om at det skal redusere segregering, men som ikke optimerer for lav segregering direkte. Figur 3 viser at en sterk reduksjon i segregering er mulig med begrenset inngrep i innfrielse av ønsker.

I optimeringsmodellen må brukeren bestemme i hvilken grad elever med høyt karaktersnitt skal vektas når ønsker innfris. Modellen returnerer flere løsningsforslag, og man velger løsning basert på i hvilken grad målsettingen om lav segregering skal gå ut over elevenes egne ønsker. Dette er politiske beslutninger som modellen er fleksibel nok til å kunne ta høyde for. Basert på parametervalgene kan modellen reduseres til

1. en ren karakterbasert modell,
2. en modell som ser bort fra karakterer og maksimerer hvor fornøyd elevene er i sum,
3. en modell som effektivt fjerner segregering på minst mulig bekostning av elevenes ønsker, eller
4. en hvilken som helst kombinasjon av disse målsettingene.



Figur 3: Sammenligning av loddtrekning, karakterbasert modell og optimeringsmodellen på simulerte data, der noen skoler er mer populære enn andre. Hver prikk representerer en fordeling av 500 elever på fem skoler. Den horisontale akse viser hvilket ønske elevene i snitt fikk innvilget. Den vertikale akse viser segregering målt som standardavviket av gjennomsnittskarakterene i skolene. Optimeringsmodellen vektas her alle elevene likt, uavhengig av karakter. Den returnerer en mengde løsninger: man velger løsning basert på hvor villig man er til å ofre elevenes ønsker til fordel for lavere segregering.

Oppsummering

Å velge en god inntaksmodell er et utfordrende problem, men det finnes modellegenskaper som kan evalueres uansett politisk ståsted: stabilitet, effektivitet, strategisikkerhet og insentiver. En god modell bør ha en tydelig oppfattelse av rettferdighet, og den må være matematisk forankret.

Ekspertutvalget og OsloMet gjør viktig arbeid, men vi er kritiske til modellering som er uten sterk matematisk forankring. I modellarbeid bør skillet mellom subjektive politiske ønsker og objektive modellegenskaper være tydelig. Simuleringer på den tidligere søkermassen er av liten verdi om modellene har uønskede egenskaper eller vil endre søkemønsteret.

Som et alternativ presenterte vi to nye modeller: vektet loddtrekning og en optimeringsmodell. Disse modellene reduserer segregering uten uheldige insentiver. De kan også parametriseres for å tilrettelegge for ulike politiske beslutninger, noe de andre modellene ikke er i stand til. Våre forslag er imidlertid ikke perfekte. Alle modeller er bundet av uunngåelige avveininger mellom matematiske egenskaper. Det beste man kan gjøre er å konstruere en modell som beviselig innehar de egenskapene man anser som viktigst.

Våre modeller er blitt implementert, og programvaren er offentliggjort [1]. Detaljerte beskrivelser av modellene tilgjengelige på internett [2][3]. Beskrivelsen er også sendt til Inntaksutvalget. I skrivende stund har vi ikke mottatt svar.

Fotnoter

1. <https://github.com/tommyod/inntaksmodeller>

2. <https://tommyodland.com/articles/2020/fritt-skolevalg-del-2-karakterer/>
3. <https://tommyodland.com/articles/2020/fritt-skolevalg-del-3-segregering/>

Referanser

- Abdulkadiroğlu A. og Sönmez T. (2003). School Choice: A Mechanism Design Approach. *American Economic Review*, Vol 93, No. 3.
- Ahujia R. K., Magnanti T. L. og Orlin J. B. (1993). *Network flows: Theory, algorithms and applications*. New Jersey: Prentice-Hall.
- Franklin A. og Koenigsberg E. (1973). Computed School Assignments in a Large District. *Operations Research*, Vol. 21, No. 2.
- Flatø M., Solli I. F., Stoltenberg C. og Torvik F. A. (2019). Progresjonsbasert opptak - et mer rettferdig inntakssystem. *Bedre Skole*, nr. 4-2019.
- Gale D. og Shapley L. S. (1962). College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, Vol. 69, No. 1.
- NOU, Norges offentlige utredninger (2019:3). Nye sjanser - bedre læring. Kjønnforskjeller i skoleprestasjoner og utdanningsløp. Oslo: Utdannings og forskningsdepartementet.
- Roth A. E. (1982). The Economics of Matching: Stability and Incentives. *Mathematics of Operations Research*, Vol. 7, No. 4.
- Serediak O. og Helland H. (2020). "Inntak til Oslos videregående skoler. Analyse av simulerte inntaksmodeller", OsloMet Skriftserie 2020 nr 1.