**Optimale loting en matching**

**voor het VO Amsterdam**

Ko Anthonisse

12 april 2021

© Copyright J.M. Anthonisse, 2021

**SAMENVATTING**

In het kader van de overstap van basisonderwijs naar voortgezet onderwijs worden de leerlingen over de VO scholen verdeeld op basis van de lijstjes van voorkeuren voor scholen die de leerlingen hebben ingediend. In deze notitie wordt de kwaliteit van die verdeling (matching) afgemeten aan de aantallen leerlingen die op lage voorkeuren worden geplaats: hoe minder hoe beter.

In deze notitie wordt een methode onderzocht die als expliciete doelstelling heeft om die aantallen te minimaliseren. Dit gebeurt met de gegevens van het Amsterdamse VO van de laatste jaren.

De optimalisate zou in elk van de laatste jaren nagenoeg alle leerlingen in hun top 3 van voorkeuren hebben geplaatst. Dat is veel beter dan wat in de huidige praktijk wordt gerealiseerd: honderden leerlingen buiten hun top 3, in 2020 zelfs ruim 600.

Met nagenoeg alle leerlingen in hun top 3 zullen er naar verwachting niet heel veel leerlingen zijn die door het indienen van strategische valse voorkeuren zullen proberen om een betere plek te bemachtigen. Met de data 2019 wordt dit bevestigd. Als de leerlingen beredeneerd/weloverwogen manipuleren zijn het er niet heel veel, met behoorlijke kans op een betere plek en een niet te verwaarlozen kans op een slechtere plek. Dit manipuleren heeft weinig effect op de aantallen leerlingen die op $1^e$, $2^e$, $3^e$ positie worden geplaatst en een tweede ronde is niet nodig.

Als er op veel grotere schaal dan verwacht mag worden strategie zou worden toegepast, dan leidt dat tot een tweede ronde. Vervelend voor de organisatie, maar ook nadelig voor de leerlingen want in die tweede ronde zullen ze hun echte eerste voorkeur zeker niet bereiken. Leerlingen en organisatie hebben een gemeenschappelijk belang bij het voorkomen van een tweede ronde. Dat zou kunnen met een verbod om overtekende scholen als lage voorkeuren te noemen.

Overwegen om al dan niet strategie toe te passen betreft een komende edite van loting&matching. Voorkeuren van de leerlingen en de capaciteiten van de scholen zijn onbekend. Gegevens van de voorgaande editie zouden kunnen worden gebruikt ter oriëntatie op strategie. Het blijkt dat toepassing van de weloverwogen gunstige strategie van 2019 op de gegevens 2020 tot grote teleurstelling zou leiden. Het is kennelijk riskant om een strategie voor een komende editie van loting&matching te baseren op resultaten van een voorgaande editie.

**INHOUD**


| | | |
|---|---|---|
| 1 | Inleiding | 1 |
| 2 | Terminologie | 3 |
| 3 | Voorrang en plaatsingsgarantie bij het Amsterdamse VO | 4 |
| 4 | Hoe het werkt | 6 |
| 5 | Toepassing op de data 2016 .. 2020 | 9 |
| 6 | De lengte van de voorkeurslijst | 10 |
| 7 | Al dan niet manipuleren | 12 |
| 8 | Tevredenheid met plaatsing | 14 |
| 9 | Beredeneerd manipuleren 2019 | 17 |
| 10 | Unilateraal manipuleren | 21 |
| 11 | Multilateraal manipuleren | 22 |
| 12 | Meer manipuleren | 26 |
| 13 | Contra manipulatie | 29 |
| 14 | Toepassing op gegevens 2020 | 33 |
| 15 | Een virtuele discussie | 35 |
| 16 | Afsluitende opmerkingen | 37 |
| 17 | Verwijzingen | 38 |


**BIJLAGEN**


1 Welke methode wordt nu eigenlijk gebruikt?
2 Een overbodige omrekening
3 Waarderingen voor voorkeursplaatsen
4 Plaatsingen en populariteit eerste voorkeuren
5 Kleinschalig onderzoek unilateraal manipuleren
6 Toepassingen van sociaal optimaliseren

## 1. Inleiding

In Amsterdam maken jaarlijks meer dan zevenduizend leerlingen de overstap van basisonderwijs naar voortgezet onderwijs. De verdeling van de leerlingen over de scholen van voortgezet onderwijs is gebaseerd op de voorkeuren van de leerlingen. Sinds 2016 gebruikt OSVO[1] de methode **RSD** (dus niet **DA** of **DA-STB,** zie ook Bijlage 1) om de toewijzing van leerlingen aan plaatsen te berekenen,.

Tabel 1 geeft de resultaten van de afgelopen jaren[2], het persbericht van OSVO geeft iets afwijkende waarden, zie Bijlage 2.

| Jaar | Aantal leerlingen | Aantal met voorrang[3] | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|---|
| 2016 | 7543 | 1108 | 83.3 | 92.5 | 95.5 | 298 | 37 |
| 2017 | 7638 | 1058 | 84.9 | 91.9 | 94.9 | 345 | 44 |
| 2018 | 7420 | 947 | 87.1 | 93.6 | 97.0 | 218 | 6 |
| 2019 | 7580 | 881 | 83.1 | 91.4 | 95.3 | 344 | 12 |
| 2020 | 7672 | 832 | 80.2 | 88.0 | 91.4 | 616 | 41 |

Tabel 1: Gerealiseerde matchings methode **RSD**, exclusief handmatige plaatsingen wegens garantie

Elk jaar worden enkele honderden leerlingen niet op een van hun eerste drie voorkeuren geplaatst. Voor de overgrote meerderheid van de betreffende leerlingen en ouders is dit geen acceptabele uitkomst[4]. Voor een kleine minderheid van ouders en leerlingen is uitsluitend plaatsing op eerste voorkeur acceptabel[5]. Omdat de meeste pijn bij de laagst geplaatsten zit is het uit het oogpunt van *welfare* gewenst om die aantallen te verminderen.

Deze notitie gaat over een methode die de plaatsingen van de leerlingen optimaliseert. Werkend van laag naar hoog niveau van voorkeur wordt telkens het aantal leerlingen op dat niveau geminimaliseerd, waarbij de aantallen op de lagere (afgehandelde) niveaus niet mogen toenemen.

De methode heeft als naam **Bmin,** als afkorting van backward minimization.

---

[1] OSVO is het samenwerkingsverband van de scholen voor voortgezet onderwijs, Zie website verenigingosvo.nl.
[2] Voor 2016/17/18/19 zijn de gevens ontleend aan het rapport Cijfermatige/Beschrijvende/Eerste analyse van het betreffende jaar. Voor 2020 zijn de gegevens gebaseerd op de MEPS data. Van de 41 door RSD niet geplaatsten werden 32 leerlingen na afloop van RSD, op grond van de plaatsingsgarantie, alsnog handmatig geplaatst. In de Beschrijvende analyse 2020 is geen onderscheid gemaakt tussen plaatsing door RSD en handmatig.
[3] Voorrang is beperkt tot de school van eerste voorkeur.
[4] Plaatsing in top 3 is voor 62/55 % van de ouders/leerlingen acceptabel. Voor plaatsing in top 4 zijn de percentages 36/33 [Tevredenheidsonderzoek 2018]. Percentages verschillen per schooladvies.
[5] Plaatsing in top 2 is voor 83/79 % van de ouders/leerlingen acceptabel. Die minderheid is dus 17/21 %. Ook hier verschillen de percentages per schooladvies. Zie Tevredenheidsonderzoeken van de verschillende jaren.

**Verder opzet van deze notitie**

Allereerst wordt enige terminologie geïntroduceerd. Daarna wordt ingegaan op voorrang en plaatsingsgarantie bij het Amsterdamse VO. Het blijkt dat de werkwijze van OSVO niet strategyproof is. Vervolgens wordt aan de hand van twee voorbeelden de werking van de methode besproken, gevolgd door toepassing op de data 2016…2020. Het optimaliseren blijkt zeer goede resultaten te leveren. Aansluitend wordt gekeken naar de benodigde lengte van de voorkeurslijsten, vier of vijf voorkeuren zou in de genoemde jaren voldoende zijn geweest.

De rest van de notitie gaat voornamelijk over strategie.

Allereerst gaat het over de vraag op grond waarvan leerlingen zouden besluiten om al dan niet strategie toe te passen. Aansluitend wordt tevredenheid met plaatsing besproken: leerlingen die niet blij zouden zijn met plek 2 kunnen geneigd zijn om strategie toe te passen als ze mogen verwachten om op plek 2 geplaatst te worden. Vervolgens gaat het over de kans dat een leerling op een bepaalde plek geplaatst zal worden. Met de data 2019 blijkt dat die kansen samenhangen met het patroon van populariteit van de eerste drie voorkeuren. Daarna wordt op basis van unilateraal manipuleren naar bruikbare strategiën gezocht. Dit wordt voortgezet met multilateraal manipuleren en dan worden verschillende strategiën toegepast op de data 2019. Het blijkt dat, op basis van tevredenheid en de kansen op plaatsing op plek 2 en 3, er niet heel veel manipulanten zullen zijn. In een bepaalde reeks hebben de manipulanten 40% kans op hogere plaatsing en 10% kans op lagere plaatsing. Dit heeft weinig effect op de aantallen leerlingen op $1^e$, $2^e$, $3^e$ positie. Daarna worden de grenzen van de aanpak opgezocht door (veel) meer leerlingen te laten manipuleren dan verwacht mag worden. Als dat er maar genoeg zijn leidt het tot een tweede ronde, wat ook in het nadeel van de leerlingen is. Vervolgens wordt een mogelijke maatregel besproken die een tweede ronde zou verhinderen.

Dan wordt de weloverwogen strategie die in 2019 goed werkte toegepast op de data 2020. Dat geeft een voor de leerlingen teleurstellend resultaat. In een bepaalde reeks hebben de manipulanten 25% kans op hogere plaatsing en 32% kans op lagere plaatsing, de overige reeksen navenant. Het is kennelijk riskant om een strategie voor een komende editie van loting&matching te baseren op resultaten van een voorgaande editie.

De notitie wordt afgesloten met een virtuele discussie en enkele opmerkingen.

## 2. Terminologie

De beschikbare plaatsen op een school kunnen zijn onderverdeeld naar onderwijsniveau, leerweg en profiel[6]. Geplaatste leerlingen worden geteld per niveau en per reguliere en/of profielklassen. Op deze manier zijn scholen opgedeeld in *afdelingen*. De voorkeuren van de leerlingen zijn afdelingen van scholen, plaatsing is op een afdeling van een school. Technisch gesproken is een afdeling een capaciteitsgroep. In de literatuur over matching gaat het altijd over leerlingen en capaciteitsgroepen die scholen worden genoemd. In lijn hiermee is in deze notitie een *school* doorgaans een afdeling van een echte school.

Een *match* of *matching* is een verdeling van de leerlingen over de scholen (=capaciteitsgroepen).

Een voorkeur van een leerling is *onbruikbaar* als de leerling geen voorrang op die school heeft en de school volbezet zal zijn met leerlingen die daar wel voorrang hebben.

De *populariteit* van een school is de verhouding tussen het aantal vermeldingen als eerste voorkeur en het aantal beschikbare plaatsen. Een school is *populair/overtekend* als de populariteit >1. De populariteit kan ook als percentage worden genoemd.

Een *manipulant* is een leerling die niet haar echte, maar valse voorkeuren indient. Het is een poging om daardoor op een betere plaats te komen. Als die poging mislukt is die leerling nog steeds een manipulant. De kwalificatie is dus niet beperkt ot degenen die succes hebben.

Een *strategie* is het op een bepaalde manier opstellen van (valse) voorkeuren. Vaak is de beste strategie het noemen van de scholen in echte voorkeursvolgorde.

Een methode is *strategyproof* als het met die methode onmogelijk is om winst te behalen door het opgeven van valse voorkeuren. **RSD** is een voorbeeld.

Als een leerling met haar echte voorkeuren aan haar voorkeursplaats a wordt toegewezen en met een strategie aan haar echte voorkeursplaats b, dan is er sprake van een *verplaatsing* van a naar b.

Een *profiel* is een rijtje getallen dat aantallen plaatsingen op $1^e$, $2^e$, $3^e$, … voorkeur geeft. De getallen kunnen ook percentages zijn. Een profiel betreft een of meer leerlingen en bevat de resultaten van een of meer runs.

[6] In 2020 waren er 63 scholen, 35 daarvan hadden twee of meer afdelingen en 28 slechts 1 afdeling.

## 3. Voorrang en plaatsingsgarantie bij het Amsterdamse VO

Bij het Amsterdamse VO is een eenvoudige *voorrangssituatie* van toepassing. Een leerling heeft voorrang op een school of niet. En die is alleen geldig voor de school van eerste voorkeur.

Als, onder deze omstandigheden, bij een school het aantal leerlingen met voorrang hoogstens gelijk is aan het aantal plaatsen, dan kunnen die leerlingen daar worden geplaatst en verder buiten de berekeningen worden gelaten.

De laatste jaren was dit in Amsterdam steeds het geval, vanaf 2018.

Omdat de voorrang slechts geldig is op de eerste voorkeur, kan een leerling met alleen voorrang op haar tweede voorkeur, overwegen om die tweede dan maar als eerste te noemen. Want anders vervalt de voorrang en dan zou plaatsing op derde voorkeur of lager het gevolg kunnen zijn. Maar door deze maipulatie vervalt de mogelijkheid van plaatsing op eerste voorkeur (aannemende dat de school voldoende plek heeft voor alle leerlingen met voorrang).

De werkwijze 'voorrang alleen geldig op eerste voorkeur' lokt zodoende strategisch gedrag uit, wat op zich niet erg hoeft te zijn, maar in dit geval ongunstig is voor iedereeen. De betreffende leerling komt niet op haar eerste voorkeur en andere leerlingen worden van hun eerste voorkeur verdrongen. Dat is *lose-lose*.

Als voorrang op elke voorkeursplaats geldig is kan de matching bijvoorbeeld met de methode **DA-STB** worden berekend, die is daar voor ontworpen.

Als voorrang alleen geldig is op eerste voorkeur kan als volgt worden gewerkt:

1. De leerlingen worden in lotingsvolgorde gezet
2. Zoveel mogelijk van de leerlingen met voorrang worden geplaatst
3. **RSD** wordt toegepast op de rest

**RSD** (Random Serial Dictatorship) houdt in dat de leerlingen een voor een, in de lotingsvolgorde, worden geplaatst op hun hoogste voorkeur waar nog plek is. Het is wie het eerst komt die het eerst maalt. **RSD** negeert eventuele voorrang, maar die is hier niet meer aan de orde.

Dit is de werkwijze die sinds 2016 in Amsterdam wordt toegepast.

Sinds 2018 wordt door OSVO een *plaatsingsgarantie* toegepast. Als een leerling onverhoopt door RSD niet wordt geplaatst maar haar voorkeurslijst is lang genoeg, dan wordt alsnog een plaats voor haar gecreëerd.

De benodigde lengte van de voorkeurslijst is afhankelijk van het schooladvies. In 2019:

- Vmbo-b, vmbo-b/k, vmbo-k met of zonder lwoo: minimaal 4 vo-scholen op de voorkeurslijst.
- Vmbo-t met of zonder lwoo en vmbo-t/havo: minimaal 6 vo-scholen op de voorkeurslijst.
- Havo, havo/vwo en vwo: minimaal 12 vo-scholen op de voorkeurslijst.

Als scholen tellen hier alleen de echte scholen, dus meerdere afdelingen van dezelfde echte school tellen samen als één school.

Tijdens het Garantie-overleg wordt, per leerling, de voorkeurslijst vanaf eerste voorkeur doorlopen. Wanneer de vertegenwoordiger van de betreffende school van eerste voorkeur een aanbod kan doen steekt hij zijn hand op. Mocht dit niet mogelijk zijn dan wordt gekeken naar de school van

tweede voorkeur enzovoorts. De leerling wordt geplaatst op die school waarvan de vertegenwoordiger zijn hand heeft opgestoken.

Deze werkwijze, de voorkeurslijst vanaf eerste voorkeur doorlopen, lokt strategisch gedrag uit. Leerlingen kunnen in de staart van hun voorkeurslijst populaire scholen noemen in plaats van hun echte voorkeuren, met als opzet om niet geplaatst te worden als **RSD** geen plaatsing in de top van de lijst geeft. En dan hopen dat het Garantie-overleg wel een plaatsing in de top van de lijst geeft.

Dit strategische gedrag is te ontmoedigen door in het Garantie-overleg de voorkeurslijst niet in een vaste volgorde te doorlopen maar in een volgorde op basis van loting. Simpel: genummerde briefjes, schudden, en er telkens een pakken tot een vertegenwoordiger zijn hand opsteekt.

Samenvattend kan worden vastgesteld dat, hoewel **RSD** strategyproof is, de werkwijze van OSVO zeker niet strategyproof is.

Desondanks zijn suggesties voor alternatieve methoden steeds afgedaan met de zeer nadrukkelijke opmerking: maar dat lokt strategie uit, of woorden van gelijke strekking[7]. Alsof er een verbod is op het zelfs maar onderzoeken van de omvang en effecten van die strategie. Een dogmatisch standpunt, zonder onderbouwing.

Er lijkt wel iets veranderd te zijn. Op 21 augustus 2020 schreef de voorzitter van OSVO (de heer Oudkerk) een brief aan de wethouder (mevrouw Moorman). Daarin werden twee scenario's gepresenteerd die eventueel in 2021 zouden kunnen worden toegepast (het kwam er niet van). Het eerste scenario hield in dat de volgende methode zou worden gebruikt: één stap Boston en daarna **RSD**. **Boston** is het bekende in- en uitloten, en is niet strategy-proof (het kan lonen om de eerste echte voorkeur te verzwijgen), maar wordt vaak toegepast. Mogelijk is de aversie tegen strategie iets afgezwakt.

[7] Ik herinner me een vergadering (ik heb de video gezien) van de raadscommissie waarin een raadslid een andere methode suggereerde (één stap Boston, daarna RSD). Een van de aanwezige deskundigen reageerde daarop fel: Dan weet ik wel hoe ik op mijn eerste voorkeur kan komen. Daarmee was het punt afgesloten. Die reactie was pure retoriek/intimidatie. Ik denk dat het onmogelijk is om met die methode via valse voorkeuren gegarandeerd op eerste voorkeur geplaatst te worden.

## 4. Hoe het werkt

Deze sectie bevat een informele beschrijving van de werking van de methode. Een formele beschrijving wordt elders gegeven.

Het *eerste voorbeeld* is ter inleiding. Er zijn 6699 leerlingen, elk van hen heeft vier voorkeuren opgegeven, en uit voorbereidende berekeningen is al gebleken dat alle leerlingen binnen hun top 4 kunnen worden geplaatst, uiteraard ook binnen de capaciteiten van de scholen. In dit geval is een vereenvoudigde versie van de methode van toepassing[8].

In elke stap van de methode wordt een voorlopige matching berekend, die van de laatste stap is de definitieve.

Stap 1 bestaat uit het minimaliseren van het aantal plaatsingen op $4^e$ voorkeur, onder de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd
- Alle leerlingen worden geplaatst

Het resultaat is dat geen leerlingen op $4^e$ voorkeur behoeven te worden geplaatst.

Stap 2 bestaat uit het minimaliseren van het aantal plaatsingen op $3^e$ voorkeur, onder de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd
- Alle leerlingen worden geplaatst
- Geen leerlingen worden op $4^e$ voorkeur geplaatst

Het resultaat is dat 78 leerlingen op $3^e$ voorkeur moeten worden geplaatst.

Stap 3 bestaat uit het minimaliseren van het aantal plaatsingen op $2^e$ voorkeur, onder de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd
- Alle leerlingen worden geplaatst
- Geen leerlingen worden op $4^e$ voorkeur geplaatst
- 78 leerlingen worden op $3^e$ voorkeur geplaatst

Het resultaat is dat 1270 leerlingen op $2^e$ voorkeur moeten worden geplaatst.

In stap 3 wordt een matching gevonden met 5351/ 1270/ 78 leerlingen op $1^e$/ $2^e$/ $3^e$ voorkeur, daarmee zijn de berekeningen voltooid. De 5351 op $1^e$ voorkeur volgen uit de voorwaarde dat alle leerlingen moeten worden geplaatst, in combinatie met geen op $4^e$, 78 op $3^e$, en 1270 op $2^e$ voorkeur.

In elke stap wordt het aantal leerlingen berekend dat op een bepaalde voorkeur moet worden geplaatst. In de daarop volgende stappen wordt dat aantal plaatsingen als voorwaarde meegenomen.

Het gaat steeds over *aantallen* leerlingen. In stap 3 wordt de definitieve matchnig gevonden, dan blijkt *welke* leerlingen op $3^e$, $2^e$, $1^e$ voorkeur zijn geplaatst.

---

[8] Deze aanpak is door Klaren (2016) in een ruimer verband besproken. Zie ook Bijlage 6.

In het *tweede voorbeeld* zijn er weer 6699 leerlingen, elk van hen heeft minstens één en hoogstens drie voorkeuren opgegeven. Er zijn 265/ 78/ 6356 leerlingen met 1/ 2/ 3 voorkeuren. Aan de hand hiervan wordt de algemene vorm van de methode beschreven.

In elke stap van de methode worden twee voorlopige matchings berekend.

Stap 1a bestaat uit het maximaliseren van het aantal plaatsingen van leerlingen met 3 voorkeuren, onde de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd

Het blijkt dat alle 6356 leerlingen kunnen worden geplaatst, ergens in hun top 3.

Stap 1b bestaat uit het minimaliseren van het aantal plaatsingen op $3^e$ voorkeur, onder de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd
- Alle 6356 leerlingen met 3 voorkeuren worden geplaatst

Het resultaat is dat 76 leerlingen op $3^e$ voorkeur moeten worden geplaatst.

Stap 2a bestaat uit het maximaliseren van het aantal plaatsingen van leerlingen met 2 voorkeuren, onder de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd
- Alle 6356 leerlingen met 3 voorkeuren worden geplaatst
- 76 leerlingen worden op $3^e$ voorkeur geplaatst

Het blijkt dat van de 78 er 74 leerlingen kunnen worden geplaatst, ergens in hun top 2.

Stap 2b bestaat uit het minimaliseren van het aantal plaatsingen op $2^e$ voorkeur, onder de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd
- Alle 6356 leerlingen met 3 voorkeuren worden geplaatst
- 76 leerlingen worden op $3^e$ voorkeur geplaatst
- 74 leerlingen met 2 voorkeuren worden geplaatst

Het resultaat is dat 1238 leerlingen op $2^e$ voorkeur moeten worden geplaatst.

Stap 3a bestaat uit het maximaliseren van het aantal plaatsingen van leerlingen met 1 voorkeur, onder de voorwaarden dat

- De capaciteiten van de scholen worden gerespecteerd
- Alle 6356 leerlingen met 3 voorkeuren worden geplaatst
- 76 leerlingen worden op $3^e$ voorkeur geplaatst
- 74 leerlingen met 2 voorkeuren worden geplaatst
- 1238 leerlingen worden op $2^e$ voorkeur geplaatst

Het blijkt dat van de 265 er 244 leerlingen kunnen worden geplaatst.

In stap 3a wordt een matching gevonden met 5360/ 1238/ 76 leerlingen op $1^e$/ $2^e$/ $3^e$ voorkeur, en 25 ongeplaatste leerlingen, stap 3b is overbodig.

De volgende tabel geeft een overzicht van de berekeningen.

| stap | voorkeur | aantal leerlingen | maximum aantal daarvan dat plaatsbaar is op voorkeur of beter | minimum aantal op voorkeur | geplaatst |
|---|---|---|---|---|---|
| 1 | 3 | 6356 | 6356 | 76 | 76 |
| 2 | 2 | 78 | 74 | 1238 | 1238 |
| 3 | 1 | 265 | 244 | | 5360 |

Tabel 2 : Verloop van de berekeningen, voorbeeld met drie voorkeuren

De methode begint, in stap 1a, met de leerlingen die het grootste aantal voorkeuren hebben ingevuld. Er wordt berekend hoeveel van hen kunnen worden geplaatst. Dat geldt ook als alle leerlingen evenveel voorkeuren hebben opgegeven. In dat geval zijn de sub-stappen 2a, 3a, … leeg omdat er geen leerlingen met de bijbehorende aantallen voorkeuren zijn.

Leerlingen die veel voorkeuren noemen komen eerder in aanmerking voor plaatsing dan leerlingen die weinig voorkeuren noemen. Dit is een ontmoediging voor leerlingen om slechts enkele voorkeuren te noemen. Als voldoende leerlingen weinig voorkeuren noemen kan het gebeuren dat veel leerlingen niet worden geplaatst. Dan moet een tweede ronde worden georganiseerd, wat niet aantrekkelijk is voor de organisatie. Bovendien is dit niet aantrekkelijk voor die leerlingen, want hun hoogste voorkeuren zullen dan niet meer beschikbaar zijn.

Zoals gezegd gaan de berekeningen steeds over *aantallen* leerlingen, na de laatste stap blijkt dan wel *welke* leerlingen op $1^e$/ $2^e$/ $3^e$ voorkeur zijn geplaatst. Dat is afhankelijk van de volgorde waarin de leerlingen staan. Daarom worden de leerlingen, voorafgaand aan stap 1a, in een lotingsvolgorde gezet.

De aantallen zijn niet afhankelijk van de lotingsvolgorde. Dus als in de proefmatching de capaciteiten zo zijn aangepast dat er weinig leerlingen op lage voorkeuren komen, dan is dat in de definitieve matching, met alleen een andere lotingsvolgorde, ook zo.

Met **RSD** zijn de aantallen wel afhankelijk van de lotingsvolgorde. Dus in de definitieve matching kunnen de aantallen afwijken van wat in de proefmatching was gezien.

## 5. Toepassing op de data 2016 .. 2020

Voor elk van de vijf jaren is **Bmin** toegepast, met de voorkeuren en capaciteiten zoals in de MEPS data was vastgelegd.

Tabel 3 geeft de resultaten.

| Jaar | Aantal leerlingen | Aantal met voorrang | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|---|
| 2016 | 7453 | 1108 | 83.6 | 97.7 | 99.3 | 10 | 41 |
| 2017 | 7638 | 1058 | 84.8 | 97.4 | 99.5 | 1 | 38 |
| 2018 | 7420 | 947 | 87.0 | 99.5 | 99.9 | 0 | 10 |
| 2019 | 7580 | 881 | 82.4 | 98.7 | 99.7 | 0 | 23 |
| 2020 | 7672 | 832 | 79.5 | 93.4 | 98.5 | 92 | 25 |

Tabel 3: matchings met methode **Bmin**

Uit vergelijking met Tabel 1 blijkt dat de percentages op eerste voorkeur geplaatste leerlingen niet veel verschillen van die met **RSD**, de grootste verschillen zijn 0.7 procentpunt in in 2019 en 2020.

De percentages in top 3 geplaatste leerlingen verschillen wel veel van die met **RSD**, ten gunste van **Bmin**. Het kleinste verschil is 2.9 procentpunt in 2018, het grootste is 6.6 procentpunt in 2020.

Deze verschillen komen ook naar voren in de aantallen op vierde of lagere voorkeur geplaatste leerlingen. Met **RSD** zijn het er honderden in elk van de jaren 16/17/18/19, met **Bmin** zijn het er slecht weinige. In 2020 gaf **RSD** 616 plaatsingen op vierde of lagere voorkeur, met **Bmin** zijn dat er 92, en alle op 4$^{e}$ voorkeur.

*Het effect van het vervangen van **RSD** door **Bmin** is voor de leerlingen dat de kans op plaatsing buiten de top 3 sterk daalt en dat de kans op plaatsing op plek 1 marginaal kleiner wordt.*

In 2016 en 2018 geeft **Bmin** iets meer ongeplaatste leerlingen dan **RSD**, in 2017 iets minder, in 2019 duidelijk meer en in 2020 duidelijk minder.

Van de door **Bmin** niet geplaatste leerlingen waren er in de jaren 2016/17/18/19/20 slechts 7/ 1/ 0/ 2/ 1 met drie voorkeuren, de overige niet geplaatsten hadden minder dan drie voorkeuren.

Om alle leerlingen geplaatst te krijgen zullen ze voldoende aantallen voorkeuren moeten opgeven.

## 6. De lengte van de voorkeurslijst

Het is noch voor de leerlingen noch voor de organisatie prettig als er (veel) ongeplaatste leerlingen zijn. De organisatie moet een tweede ronde beginnen en de leerlingen weten dat de beste plaatsen niet meer beschikbaar zijn.

De methode **Bmin** nodigt de leerlingen uit om veel scholen in hun voorkeurslijst te noemen. Want leerlingen met meer voorkeuren hebben grotere kans op plaatsing dan leerlingen met minder voorkeuren.

De vraag is nu hoe lang de lijstjes minimaal zouden moeten zijn om alle leerlingen geplaatst te krijgen.

Allereerst is onderzocht of, in elk van de jaren 2016,…,2020, drie voorkeuren voldoende zou zijn geweest.

In 2016/17/18/19/20 waren er 661/674/523/480/430 leerlingen met slechts één of twee voorkeuren. Voor deze leerlingen zijn op een plausibele manier aanvullende scholen gekozen.

Voor elk jaar zijn 10 runs uitgevoerd, met in elke run een andere aanvulling van de voorkeuren tot drie.

Tabel 4 geeft de resultaten.

| Jaar | Aantal leerlingen | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal op aanvulling geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|---|
| 2016 | 7453 | 83.8 | 97.9 | 99.7 | 13.3 | 11.7 | 0.0 |
| 2017 | 7638 | 84.7 | 97.5 | 99.8 | 3.2 | 11.8 | 0.0 |
| 2018 | 7420 | 87.0 | 99.6 | 99.9 | 0.0 | 4.9 | 0.0 |
| 2019 | 7580 | 82.3 | 98.9 | 99.9 | 0.0 | 6.8 | 0.0 |
| 2020 | 7672 | 80.8 | 93.9 | 98.7 | 0.0 | 7.6 | 94.5 |

Tabel 4: matchings methode **Bmin**, gemiddeld over 10 runs, voorkeuren aangevuld tot drie

In de eerste drie jaren zijn drie voorkeuren voldoende, in 2020 is dat niet het geval.

In 2020 waren er 595 leerlingen met hoogstens drie voorkeuren. In 2020 waren vier voorkeuren wel voldoende, zie Tabel 5.

| Jaar | Aantal leerlingen | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal op aanvulling geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|---|
| 2020 | 7672 | 79.3 | 93.4 | 98.7 | 94.1 | 8.4 | 0.0 |

Tabel 5: 2020, matchings methode **Bmin**, gemiddeld over 10 runs, voorkeuren aangevuld tot vier

Het valt op (Tabellen 4 en 5) dat slechts weinig leerlingen op een aanvullende voorkeur zijn geplaatst. Zonder de aanvullingen werden tientallen leerlingen niet geplaatst, zie Tabel 3. Met de aanvullingen worden ze wel geplaatst, en dat dan meestal op een oorspronkelijke voorkeur. Dank zij de aanvullingen komen ze in de optimalisatie eerder aan de beurt en worden dan meestal op een hoge voorkeur geplaatst. Kennelijk is het minder belangrijk welke scholen als aanvulling worden toegepast, het is wel belangrijk dat het lijstje lang genoeg is.

In 2016..19/2020 zou het noemen dan meer dan drie/vier voorkeuren geen effect hebben op het profiel van de plaatsing van de leerlingen, dat is een eigenschap van de methode .

*De **conclusie** van deze berekeningen is dat het voor de organisatie verstandig is als de leerlingen minstens vier (of veiligheidshalve vijf) voorkeuren noemen[9]. De organisatie kan dit als advies geven of als eis stellen.*

*Leerlingen die in hun top 4/5 voorkeuren noemen die mogelijk volbezet raken met voorrangsleerlingen doen er goed aan om extra voorkeuren te noemen ter compensatie van het mogelijk onbruikbaar zijn van hun eerste voorkeuren.*

Als alle leerlingen dan hun echte voorkeuren noemen mag verwacht worden dat ze nagenoeg allen in hun top 3 worden geplaatst en in overgrote meerderheid zelfs in hun top 2 van voorkeuren.

De vraag is nu of verwacht mag worden dat de leerlingen hun echte voorkeuren zullen noemen, en wat de effecten zijn als ze dat niet doen.

[9] Dit zou afhankelijk kunnen zijn van het schooladvies.

## 7. Al dan niet manipuleren

Als een methode niet strategyproof is zullen leerlingen zich afvragen of het voordeel zou opleveren om, op een bepaalde manier, valse voorkeuren in te leveren. De vraag is op welke argumenten zij een besluit zouden baseren.

Elke leerling zou per voorkeursplaats met een getal (nut, utility) kunnen aangeven wat haar waardering voor die plaats is, het getal bij de tweede voorkeur is niet hoger dan dat bij de eerste, enzovoort.

Deze getallen kunnen worden gebruikt om in simulaties te bepalen of de leerling voordeel zou hebben bij een bepaalde strategie: bereken in veel runs de gemiddelde waarde zonder strategie en de gemiddelde waarde met strategie. Het verschil is basis voor de keus.

Tabel 6 bevat voor een bepaalde leerling de resultaten van 300 runs. Met de echte voorkeuren zijn er 238/59/3 plaatsingen op $1^e/2^e/3^e$ voorkeur. Met strategie zijn het er 288/5/7.

| van | naar voorkeur | | | totaal |
|---|---|---|---|---|
| voorkeur | 1 | 2 | 3+ | |
| 1 | 236 | | 2 | 238 |
| 2 | 49 | 5 | 5 | 59 |
| 3 | 3 | | | 3 |
| totaal | 288 | 5 | 7 | 300 |

Tabel 6: resultaten 300 runs

Als deze leerling de volgende waarderingen heeft:

| | Plaats van voorkeur | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4+ |
| waarde | 100 | 70 | 0 | 0 |

Tabel 7: Voorbeeld van waarderingen

dan is de gemiddelde waarde van de plaatsingen met strategie (28800+350)/300 = 97.17, en is de gemiddelde waarde met echte voorkeuren (23800+4130)/300 = 93.10, dus kan de strategie worden aanbevolen, met een te verwachten winst van 4.07.

Deze winstverwachting is het gemiddelde van de kansverdeling die in Tabel 8 is opgenomen.

| waarde | -100 | -70 | 0 | +30 | +100 |
|---|---|---|---|---|---|
| frequentie | 2 | 5 | 241 | 49 | 3 |

Tabel 8: kansverdeling

De spreiding van deze kansverdeling is 19.46, dus de voorspellende waarde van de gemiddelde winst 4.07 is gering, de variatiecoëfficiënt = 4.8.

Ook als de spreiding kleiner zou zijn blijft het zo dat de leerling niet met een reeks van toepassingen van de methode te maken heeft, maar het precies één toepassing. Het is *once in a lifetime*.

Als de linker-onderhoek van een tabel als Tabel 6 veel meer verplaatsingen bevat dan de rechter-bovenhoek dan is te verwachten dat de strategie aanbeveling verdient, ongeacht de waarderingen voor de plaatsen. En als het omgekeerde het geval is, dan is de strategie naar verwachting af te wijzen, ook weer ongeacht de waarderingen voor de plaatsen.

In het voorbeeld is er in 52 van de 300 runs een gunstig  effect van de strategie, in 7 runs een ongunstig effect. Voor een willekeurig run is er een kans 17.3% op gunstig effect en 2.3% op ongunstig effect van de strategie. Op grond hiervan kan voor het toepassen van de strategie worden gekozen, de kans op gunstig is beduidend groter dan de kans op ongunstig.

Als de aantallen gunstige en ongunstige runs niet veel verschillen is er geen betrouwbare basis om voor de strategie te kiezen.

Met waarde eerste voorkeur = 100 en waarde 3$^{e}$ en lagere voorkeuren = 0 zijn winstverwachting en spreiding alleen afhankelijk van de waarde van de tweede voorkeur. Die waarde is individueel bepaald, het is lastig om een bepaalde keuze te onderbouwen.

Maar gelukkig blijkt uit de simulaties dat de resultaten vaak zo zijn dat het teken van de gemiddelde winst onafhankelijk is van de waarderingen voor de voorkeuren.

Als het aantal gunstige runs groter is dan het aantal ongunstige, dan is de gemiddelde winst in het algemeen positief. Analoog als het aantal ongunstige runs groter is dan het aantal gunstige.

Dit geldt voor de data waarmee de simulaties zijn uitgevoerd.

Voor een leerling die moet beslissen om al dan niet te manipuleren is het gegeven: met deze strategie gemiddeld positief resultaat, minder informatief dan: met deze strategie kans X op gunstig effect en kans Y op ongunstig effect.

Al met al redenen om de waarderingen geen centrale plaats te geven. Zie ook Bijlage 3.

## 8. Tevredenheid met plaatsing

In het kader van het jaarlijkse tevredenheidsonderzoek is aan de ouders gevraagd: Welke school op de voorkeurslijst was nog acceptabel geweest? Aan de leerlingen is gevraagd: Met welke scholen op jouw voorkeurslijst was je nog blij geweest? Tabel 9 bevat een deel van de antwoorden[10].

| | vmbo-b, b/k, k (inclusief lwoo) | | vmbo-t, t/havo (inclusief lwoo) | | havo, havo/vwo | | vwo | |
|---|---|---|---|---|---|---|---|---|
| | ouders | leerl | ouders | leerl | ouders | leerl | ouders | leerl |
| eerste keuze | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| top 2 | 56 | 57 | 77 | 71 | 88 | 85 | 92 | 93 |
| top 3 | 21 | 17 | 45 | 35 | 71 | 62 | 83 | 78 |
| top 4 | 11 | 8 | 15 | 12 | 41 | 36 | 62 | 54 |
| steekproef | 330 | 298 | 519 | 461 | 709 | 607 | 851 | 701 |

Tabel 9: Per top van voorkeuren en schooladvies het percentage ouders dat plaatsing in die top acceptabel noemt en het percentage leerlingen dat blij zou zijn met die plaatsing, Dialogic 2019

Deze vraagstelling gaat dus niet over de gerealiseerde plaatsing.

Voor veel ouders is plaatsing op tweede voorkeur acceptabel, dat loopt van 56% bij vmbo-b tot 92% bij vwo. Onder de ouders is geen breed gedeelde obsessie met plaatsing op eerste voorkeur.

De Stichting VSA[11] is wel sterk gericht op de eerste voorkeur:

> Ieder jaar opnieuw staat een groep kinderen buiten spel omdat zij worden uitgeloot voor de school of het schooltype van hun keuze.[12]

Met betrekking tot het opmerkelijke verschil tussen vmbo-b en vwo staat In het Dialogic rapport over 2018:

> *Een mogelijke verklaring is dat ouders met een vmbo-advies, gezien de lengte van de lijst van vier scholen, een plek op 3 of 4 negatiever ervaren dan andere ouders (waarbij een plaatsingsaanbod binnen de top 3 in verhouding tot de langere lengte van de voorkeurslijst positiever is).*

De percentages leerlingen die niet blij zijn met plaatsing op tweede en derde voorkeur staan in Tabel 10. Dit betreft 2019 en (2018)

| plaatsing op voorkeur | vmbo-b, b/k, k (inclusief lwoo) | vmbo-t, t/havo (inclusief lwoo) | havo, havo/vwo | vwo |
|---|---|---|---|---|
| 2 | 43 (49) | 29 (31) | 15 (16) | 7 ( 9) |
| 3 | 83 (82) | 65 (67) | 38 (41) | 22 (22) |

Tabel 10: Percentages leerlingen niet blij met plaatsing op tweede/derde voorkeur, Dialogic 2019 (2018)

Wat hier opvalt is het grote verschil in percentages bij de verschillende voorkeuren en de verschillen tussen de schooladviezen. Er zijn geen grote verschillen tussen de twee jaran.

[10] Overgenomen uit Tabel 16 en Tabel 17 in het rapport Dialogic Evaluatie 2019.
[11] Stichting Vrije Schoolkeuze Amsterdam
[12] website stichtingvsa.nl

Tabel 11 bevat de resultaten van **Bmin**, uitgeplitst naar de schooladviezen, het betreft de leerlingen die in de optimalisatie werden betrokken.

| plaatsing op voorkeur | vmbo-b, b/k, k (inclusief lwoo) | vmbo-t, t/havo (inclusief lwoo) | havo, havo/vwo | vwo | totaal |
|---|---|---|---|---|---|
| 1 | 1263.76 | 1303.91 | 1226.50 | 1482.83 | 5277 |
| 2 | 132.24 | 231.09 | 460.50 | 443.17 | 1267 |
| 3 | | | 78.00 | | 78 |
| totaal | 1396 | 1535 | 1765 | 1926 | 6622 |

Tabel 11: 1000 runs, gemiddelde aantallen leerlingen

Gemiddeld over de 1000 runs zouden 253.62 leerlingen niet blij zijn, dat is 3.83% van deze 6622 leerlingen.

De volgende Tabel 12 gaat wel over de gerealiseerde plaatsing, over de mate van tervredenheid daarmee.

| niveau | % dat (zeer) tevreden is |
|---|---|
| vmbo-b, vmbo-b/k, vmbo-k (+lwoo) | 83% |
| vmbo-t (+lwoo), vmbo-t/havo | 80% |
| havo, havo/vwo | 77% |
| vwo | 88% |

Tabel 12: Percentage ouders dat (zeer) tevreden is met de plaatsing van hun kind uitgesplitst naar school-advies (n=2.309), Dialogic 2020

Om dit enigszins in perspectief te plaatsen volgt nu een overzicht van de percentages leerlingen die op eerste voorkeur werden geplaatst.

| niveau | % op eerste voorkeur geplaatst |
|---|---|
| vmbo-b, vmbo-b/k, vmbo-k (+lwoo) | 96% |
| vmbo-t (+lwoo), vmbo-t/havo | 82% |
| havo, havo/vwo | 69% |
| vwo | 81% |

Tabel 13: Percentages plaatsing op eerste voorkeur, data 2020

Het is merkwaardig dat bij vmbo-b 96% op eerste voorkeur is geplaatst en dat toch maar 83% (zeer) tevreden is. Analoog bij vmbo-t. Bij havo en vwo is er ook tevredenheid met plaatsing op tweede of lagere voorkeur.Het Dialogic rapport over 2020 bevat een overzicht[13] van tevredenheid met plaatsingsaanbod, gesplitst naar niveau van voorkeur.

| geplaatst op plek | % ontevreden | % zeer ontevreden |
|---|---|---|
| 2-3 | 6 | 3 |
| 4-5 | 24 | 11 |
| 6+ | 23 | 42 |

Tabel 14: Ontevredenheid ouders met plaatsingsaanbod, Dialogic 2020

Van de leerlingen die op plek 2-3 zijn geplaatst is 91% daarmee niet ontevreden. Dat is meer dan op grond van Tabel 9 te verwachten was, maar dat was 2019.

[13] Figuur 23 in het rapport Dialogic Evaluatie 2020.

Het is duidelijk dat plaatsingen buiten de top 3 ongewenst zijn.

De neiging om te manipuleren (=valse voorkeuren inleveren) kan zeker voorkomen bij leerlingen die vrezen dat ze met hun echte voorkeuren op een plaats zullen komen waarmee ze ontevreden of zeer ontevreden zouden zijn. Dit geldt uiteraard niet als de methode van matching het onmogelijk maakt om met valse voorkeuren winst te behalen. Met **RSD** is dat zo, met **Bmin** niet.

Uit de tabellen 4 en 5 blijkt dat er met **Bmin** maar weinig leerlingen op derde en lagere voorkeuren worden geplaatst, gerekend over alle leerlingen is er dus maar een kleine kans op een dergelijke mogelijk teleurstellende gebeurtenis. Straks zal blijken dat de kansen voor verschillende leerlingen heel verschillend zijn.

## 9. Beredeneerd manipuleren 2019

Met de data 2019 zijn de voorkeuren met één loting aangevuld tot drie. Vervolgens zijn 1000 runs uitgevoerd waarbij alleen de eerste drie voorkeuren van de leerlingen werden gebruikt. In elke run werd een andere lotingsvolgorde van de leerlingen gebruikt. In de presentatie van de resultaten wordt gebruik gemaakt van een indeling van de scholen naar populariteit;

De volgende tabel geeft een verdeling van de populariteit van scholen in drie intervallen en per interval een letter ter aanduiding van dat interval.

| interval | letter | omschrijving | aantal |
|---|---|---|---|
| populariteit ≤ 1 | n | school niet overtekend | 102 |
| 1 < populariteit ≤ 2 | P | school enkel overtekend | 22 |
| 2 < populariteit | Q | school dubbel overtekend | 6 |

Tabel 15: Intervallen populariteit

Tabel 16 bevat per voorkeur de uitsplising van de plaatsingen naar populariteit van de betreffende scholen.

| populariteit | voorkeur | | | Totaal |
|---|---|---|---|---|
| | 1e | 2e | 3e | |
| n | 3735.95 | 1110.05 | 78.00 | 4924 |
| P | 1395.71 | 147.29 | | 1543 |
| Q | 145.33 | 9.67 | | 155 |
| totaal | 5277 | 1267 | 78 | 6622 |

Tabel 16 : 1000 runs, gemiddeld aantal plaatsingen

Per leerling bestaat het patroon van populariteit van de voorkeuren uit de opeenvolging van de letters die bij de voorkeuren horen.

Tabel 17 betreft de leerlingen die in de optimalisatie werden betrokken, met als eerste voorkeur een niet overtekende school.

| patroon | aantal plaatsing | | | totaal | % plaatsing | | |
|---|---|---|---|---|---|---|---|
| populariteit | 1e | 2e | 3e | | 1e | 2e | 3e |
| nnn | 1623.97 | 89.03 | | 1713 | 95 | 5 | |
| nnP | 599.91 | 69.09 | | 669 | 90 | 10 | |
| nnQ | 154.30 | 13.70 | | 168 | 92 | 8 | |
| nPn | 726.70 | 2.10 | 1.20 | 730 | 100 | 0 | 0 |
| nPP | 287.00 | 2.00 | | 289 | 99 | 1 | |
| nPQ | 103.00 | | | 103 | 100 | | |
| nQn | 144.06 | | 0.94 | 145 | 99 | | 1 |
| nQP | 81.00 | | | 81 | 100 | | |
| nQQ | 16.00 | | | 16 | 100 | | |
| totaal | 3735.95 | 175.91 | 2.14 | 3914 | 95 | 4 | 0 |

Tabel 17: 1000 runs, leerlingen met 1e voorkeur niet overtekend, gemiddeld aantal leerlingen en percentages per plaats, per populariteitspatroon

Van de 3914 leerlingen met een niet overtekende school als eerste voorkeur wordt per run gemiddeld 95.5% op die eerste voorkeur geplaatst. Het valt op dat vanaf nPn plaatsing op 1e voorkeur nagenoeg zeker is.

In de loop van de jaren is al enkele malen de suggestie gedaan[14] om alle leerlingen met een niet overtekende eerste voorkeur *a priori* op hun eerste voorkeur te plaatsen en het loten te beperken tot de overige leerlingen. De met **Bmin** berekende matching komt daar vanzelf bij in de buurt.

De gemiddeld 4.5% die niet op eerste voorkeur komt (maar wel op 2e of 3e) dragen er toe bij om de 356 leerlingen die door **RSD** op 4e of lager of niet werden geplaatst, binnen de top 3 te halen.

De weinige plaatsingen op 3e zijn op een niet overtekende school, dat geldt ook voor bijna alle plaatsingen op 2e voorkeur. Als de 2e voorkeur wel overtekend is, dan is de plaatsing bijna altijd op 1e voorkeur.

Leerlingen in deze groep met niet overtekende eerste voorkeur die (zeer) ontevreden zouden zijn met plaatsing op hun derde voorkeur kunnen zeker geneigd zijn om te manipuleren als hun voorkeuren nQn of nPn zijn, de kans op plaatsing op 3e is overigens erg klein.

Leerlingen met nQn zouden hun 3e voorkeur kunnen vervangen door een P- of een Q-school, dan lijkt plaatsing op 1e voorkeur verzekerd. Dat zal mogelijk zo zijn als enkele leerlingen dat doen, maar als het grootschalig wordt gedaan is dat zeker niet zeker.

Analoog voor leerlingen met nPn.

Leerlingen in deze groep voor wie plaatsing op 2e of 3e voorkeur niet acceptabel zou zijn kunnen geneigd zijn om te manipuleren als hun voorkeuren nnP, nnQ, nnn, nPP, nPn of nQn zijn, de kans op plaatsing op 2e of 3e is overigens hoogstens 10% (bij nnP).

Deze leerlingen zouden dan zowel hun 2e als hun 3e voorkeur door een P- of Q-school kunnen vervangen. Succes is niet verzekerd.

Tabel 18 betreft de leerlingen die in de optimalisatie werden betrokken, met als eerste voorkeur een enkel overtekende school.

| patroon populariteit | aantal plaatsing | | | totaal | % plaatsing | | |
|---|---|---|---|---|---|---|---|
| | 1e | 2e | 3e | | 1e | 2e | 3e |
| Pnn | 383.26 | 427.74 | | 811 | 47 | 53 | |
| PnP | 124.14 | 251.86 | | 376 | 33 | 67 | |
| PnQ | 29.36 | 61.64 | | 91 | 32 | 68 | |
| PPn | 440.60 | 36.57 | 1.83 | 479 | 92 | 8 | 0 |
| PPP | 220.98 | 13.02 | | 234 | 94 | 6 | |
| PPQ | 28.82 | 2.18 | | 31 | 93 | 7 | |
| PQn | 99.55 | | 5.45 | 105 | 95 | | 5 |
| PQP | 41.00 | | | 41 | 100 | | |
| PQQ | 28.00 | | | 28 | 100 | | |
| totaal | 1395.71 | 793.01 | 7.28 | 2196 | 64 | 36 | 0 |

Tabel 18: 1000 runs, leerlingen met 1e voorkeur enkel overtekend, gemiddeld aantal leerlingen en percentages per plaats, per populariteitspatroon

[14] Niet door mij.

Van de 2196 leerlingen met een enkel overtekende school als eerste voorkeur wordt per run gemiddeld 63.6% op 1e voorkeur geplaatst en bijna alle anderen op 2e voorkeur. Het valt op dat vanaf PPn kans op plaatsing op 1e voorkeur > 90% is.

Leerlingen met Pnn zouden er geen baat bij hebben om alleen hun 3e voorkeur door een P- of Q-school te vervangen. Vervanging van alleen de 2e voorkeur door een P- of Q-school zou de kans op 1e voorkeur verhogen, met risico op 3e voorkeur. Vervanging van zowel 2e als derde voorkeur door P of Q zou mogelijk wel voordelig kunnen zijn. Mogelijk niet als veel leerlingen dat doen.

Tabel 19 betreft de leerlingen die in de optimalisatie werden betrokken, met als eerste voorkeur een dubbel overtekende school.

| patroon | aantal plaatsing | | | totaal | % plaatsing | | |
|---|---|---|---|---|---|---|---|
| populariteit | 1e | 2e | 3e | | 1e | 2e | 3e |
| Qnn | 3.88 | 114.12 | | 118 | 3 | 97 | |
| QnP | 1.12 | 71.88 | | 73 | 2 | 98 | |
| QnQ | | 11.00 | | 11 | | 100 | |
| QPn | 38.22 | 51.19 | 27.59 | 117 | 33 | 44 | 24 |
| QPP | 28.24 | 32.76 | | 61 | 46 | 54 | |
| QPQ | 2.53 | 7.47 | | 10 | 25 | 75 | |
| QQn | 20.10 | 1.91 | 40.99 | 63 | 32 | 3 | 65 |
| QQP | 51.24 | 7.76 | | 59 | 87 | 13 | |
| totaal | 145.33 | 298.09 | 68.58 | 512 | 28 | 58 | 13 |

Tabel 19: 1000 runs, leerlingen met 1e voorkeur dubbel overtekend, gemiddeld aantal leerlingen en percentages per plaats, per populariteitspatroon

Van de 512 leerlingen met een dubbel overtekende school als eerste voorkeur wordt per run gemiddeld 28.4% op 1e voorkeur geplaatst en 58.2% op 2e voorkeur.

De plaatsingen op 3e voorkeur zijn alle op een niet overtekende school, dat geldt ook voor de plaatsingen op 2e voorkeur, voor zover er een niet overtekende school als 2e voorkeur is genoemd.

De leerlingen met QQn en QPn hebben de grootste kansen om op 3e voorkeur te worden geplaatst. Zij zouden dat mogelijk kunnen ontlopen door als 3e voorkeur een P-school of een Q-school te noemen.

Leerlingen met QPQ zouden mogelijk winst kunnen behalen door als 3e voorkeur een P-school te noemen. Dat lijkt a-typisch. Maar het zijn slechts 10 leerlingen en de Q-scholen in QPP zijn mogelijk andere dan de eerste voorkeuren in QPQ.

In het bovenstaande zijn de groepen leerlingen met hetzelfde patroon van populariteit van hun eerste voorkeuren behandeld alsof het homogene groepen zijn. In bijlage 4 wordt onderzocht in hoeverre dat het geval is. De resultaten zijn geen aanleiding geweest om voor de simulaties deze groepen verder uit te splitsen.

De 180 leerlingen met voorkeuren QPn en QQn hebben aanzienlijke kans om op 3e voorkeur geplaatst te worden, zij vormen de groep “vaak 3”.

De 1551 leerlingen met voorkeuren Pnn, PnP, PnQ , Qnn, QnP, QnQ, QPP en QPQ vormen de groep “vaak 2”, want ze hebben aanzienlijke kans op 2e voorkeur geplaatst te worden.

Deze twee groepen kunnen met reden manipuleren als zij die 3e of 2e voorkeursplaats niet acceptabel vinden. Of zij die plaats al dan niet acceptabel vinden hangt af van het schooladvies, zoals in Tabel 9 is aangegeven. Hieronder wordt die tabel gecombineerd met de aantallen leerlingen in de twee groepen.

| | vmbo-b, b/k, k (inclusief lwoo) | vmbo-t, t/havo (inclusief lwoo) | havo, havo/vwo | vwo | totaal |
|---|---|---|---|---|---|
| percentage niet blij met voorkeur 2 | 43 | 29 | 15 | 7 | |
| aantal in vaak 2 | 306 | 334 | 360 | 551 | 1551 |
| manipulanten voorkeur 2 | 131.58 | 96.86 | 54.00 | 38.57 | 321.01 |
| percentage niet blij met voorkeur 3 | 83 | 65 | 38 | 22 | |
| aantal in vaak 3 | 4 | 46 | 117 | 13 | 180 |
| manipulanten voorkeur 3 | 3.32 | 29.90 | 44.46 | 2,86 | 80.54 |
| manipulanten totaal | 134.90 | 126.76 | 98.46 | 41.43 | 401.55 |
| | | | | | |
| leerlingen | 1535 | 1765 | 1396 | 1926 | 6622 |
| % manipulanten | 8.79 | 7.18 | 7.05 | 2.15 | 6.06 |

Tabel 20 : Schatting aantallen manipulanten met reden

Er zijn 1551+180=1731 leerlingen die een reden kunnen hebben om te manipuleren, dat is 26.14%. Met toepassing van de percentages “niet blij” is de verwachting dat maximaal 401.55 daartoe over gaan, dat is 6.06%

Het is niet gezegd dat alle leerlingen die niet blij zouden zijn ook daadwerkelijk zouden manipuleren, niet blij van de leerlingen is niet hetzelfde als niet acceptabel voor de ouders.

Desondanks zijn in de eerste reeks van simulaties de percentages manipulanten gelijk aan de percentages niet blij zoals ze in deze tabel staan.

*De populariteit van hun eerste drie voorkeursscholen kan de leerlingen een schatting leveren van de kansen op plaatsing op 1e, 2e, 3e voorkeur. Als de leerling plaatsing op haar 3e voorkeur of op haar 2e (en dan dus ook 3e) ongewenst vindt kan zij overwegen om te proberen die plaatsing te verhinderen door niet haar echte voorkeuren in te leveren.*

Deze aanpak is er op gebaseerd dat de populariteit van de scholen bekend is, met name of scholen al dan niet overtekend zijn. Voor de komende editie van loting & matching is dat uiteraard niet exact bekend. De cijfers van de vorige editie[15] kunnen enig houvast geven. Wat betreft de aantallen plaatsen moet daarbij bedacht worden dat ze in het kader van de proefmatching kunnen worden gewijzigd, dat tijdens de matching nog 4% kan worden toegevoegd, en dat de handmatige plaatsing op basis van plaatsingsgarantie ook extra plaatsen oplevert.

[15] Staan in het betreffende rapport Beschrijvende/Cijfermatige Analyse.

## 10. Unilateraal manipuleren

Unilateraal manipuleren houdt in dat één leerling valse voorkeuren opgeeft, alle andere leerlingen geven hun echte voorkeuren. Het effect van unilateraal manipuleren kan voor veel leerlingen, stuk voor stuk, worden onderzocht. Dit leidt niet tot een advies aan een individuele leerling, want het is niet realistisch om te veronderstellen dat die leerling als enige zal manipuleren. Integendeel, als een leerling overweegt om te manipuleren dan is het verstandig om te veronderstellen dat anderen dat ook zullen doen. En als veel anderen manipuleren is het niet altijd verstandig om daarin mee te gaan.

Het is zinvol om unilateraal manipuleren te onderzoeken omdat het inzicht kan geven in de effecten van verschillende strategieën. Zie Bijlage 5 voor een kleinschalig onderzoek.

De leerlingen met QPn of QQn zouden strategie 213 kunnen overwegen, waarbij eerste en tweede echte voorkeur van plaats wisselen. De strategiën 12X, 1X2 en 1XY zijn ook te overwegen als het unilateraal is. Daarbij zijn X en Y overtekende scholen die uit de lagere voorkeuren worden gekozen. 12X en 1XY zijn al in de sectie over beredeneerd manipuleren aan de orde gekomen.

## 11. Multilateraal manipuleren

Multilateraal manipuleren houdt in dat meerdere leerlingen gelijktijdig valse voorkeuren opgeven.Verondersteld wordt dat ze dit onafhankelijk van elkaar doen. Zo niet, dan is er sprake van coalities en dat is een heel ander onderwerp.

Met de gegevens 2019 is, met één bepaalde aanvulling tot drie voorkeuren, onderzocht wat de effecten van verschillende strategieën voor het opgeven van valse voorkeuren zouden zijn geweest. Leerlingen die wegens voorrang zeker op 1e voorkeur komen passen geen strategie toe. Per leerling wordt met een bepaalde kans geloot of deze leerling zal manipuleren. Leerlingen voor wie de strategie niet toepasbaar is kunnen uiteraard niet manipuleren.

Er worden drie groepen leerlingen onderscheiden. De eerste groep (Vaak 3) bestaat uit de 180 leerlingen met voorkeuren QPn of QQn, die volgens Tabel 19 grote kans op plaatsing op 3e voorkeur hebben. De tweede groep (Vaak 2) bestaat uit de 1551 leerlingen met voorkeuren Pnn, PnP, PnQ , Qnn, QnP, QnQ, QPP en QPQ, die volgens Tabel 18 en Tabel 19 grote kans op plaatsing op 2e voorkeur hebben. De derde groep (Overig) bestaat uit de overige 4891 leerlingen met een kleine kans op plaatsing op 2e of 3e voorkeur.

In alle simulaties is gewerkt met drie voorkeuren, verkregen door in de data 2019 waar nodig aan te vullen.

**Simulaties A, B, C en D**

Er zijn in eerste instantie vier simulaties uitgevoerd, deze worden in Tabel 21 gespecificeerd.

| Simulatie | Groep leerlngen | | |
|---|---|---|---|
| | Vaak 3 | Vaak 2 | Overig |
| A | 1XY | 1X2 | |
| B | 1X2 | 1XY | |
| C | 1XY | 12X | |
| D | 213 | 1X2 | |

Tabel 21: Per simulatie de toegepaste strategie

Per simulatie zijn de manipulanten door loting uit de groepen geselecteerd, op grond van de percentages “niet blij” die in Tabel 20 zijn vermeld, dus ook afhankelijk van het schooladvies. Met deze gegevens zijn telkens 100 runs van **Bmin** uitgevoerd.

Qua manipulanten zijn dit situaties die op grond van de beschikbare gegevens redelijkerwijs te verwachten zijn.

De volgende Tabel geeft de resultaten van de simulaties A, B, C en D. Ter vergelijking ook de resultaten zonder manipuleren, daarbij zijn de aanvullingen op voorkeuren als echte voorkeuren gezien.

| simulatie | Aantal manipulanten | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|
| | 0 | 82.3 | 99.0 | 100.0 | 0 | 0 |
| A | 374 | 82.0 | 97.8 | 99.5 | 35.9 | 0 |
| B | 379 | 81.1 | 96.9 | 98.5 | 115.3 | 0 |
| C | 344 | 82.1 | 98.8 | 99.8 | 15.3 | 0 |
| D | 413 | 81.1 | 98.1 | 99.8 | 15.0 | 0 |

Tabel 22: Resultaten simulaties, gemiddelden over 100 runs

In elke simulatie werden 413 leerlingen geloot die zouden manipuleren. Het daadwerkelijke aantal is lager bij de eerste drie omdat bij sommige leerlingen geen vervangende school X of Y van voldoende populariteit beschikbaar was.

In elk van deze simulaties worden alle leerlingen geplaatst, een tweede ronde is niet aan de orde. Nagenoeg alle leerlingen worden in hun top 3 geplaatst. De plaatsingen op 4$^{e}$ of lagere voorkeuren zijn uitsluitend manipulanten. De tabel is gebaseerd op de echte voorkeuren van de leerlingen, met de nominale voorkeuren zouden ze allen in hun top 3 zijn geteld.

De volgende Tabel bevat de effecten van het manipuleren op de plaatsing van de leerlingen.

| simulatie | manipulanten | | niet manipulanten | |
|---|---|---|---|---|
| | hoger | lager | hoger | lager |
| A | 157.99 | 43.98 | 92.75 | 275.92 |
| B | 131.72 | 115.26 | 86.21 | 285.09 |
| C | 38.88 | 36.78 | 114.58 | 139.37 |
| D | 147.17 | 52.47 | 97.39 | 253.51 |

Tabel 23 : Verplaatsingen wegens manipuleren, aantallen leerlingen, gemiddelden over 100 runs

In simulatie A zijn er per run gemiddeld 172.03 (46%) manipulanten zonder effect op de plaatsing. De kansen op hogere respectievelijk lagere plaatsing zijn 42.24% en 11.76%. Dat is niet direct onaantrekkelijk voor wie overweegt strategie te gaan toepassen. Hierbij moet bedacht worden dat het betrekking heeft op de 374 manipulanten die allen deze strategiën toepassen.

Bij simulaties B en C blijkt dat met die strategiën de aantallen op hogere en lagere plaatsing niet veel verschillen. Dat is niet aantrekkelijk. Simulatie D is vergelijkbaar met A.

De verplaatsingen van de niet-manipulanten moeten gezien worden in de context van hun aantal: het zijn er rond 6200, dus de kans om als niet-manipulant effect van het manipuleren te ondervinden is klein.

De volgende tabel bevat de aantallen als percentage van de aantallen manipulanten en niet manipulanten.

| simulatie | manipulanten | | niet manipulanten | |
|---|---|---|---|---|
| | hoger | lager | hoger | lager |
| A | 42.24 | 11.76 | 1.48 | 4.42 |
| B | 34.75 | 30.41 | 1.38 | 4.57 |
| C | 11.30 | 10.69 | 1.83 | 2.22 |
| D | 35.63 | 12.70 | 1.57 | 4.08 |

Tabel 24 : Verplaatsingen wegens manipuleren, percentages, gemiddelden over 100 runs

Voor de leerlingen is simulatie A, of beter: de strategiën van simulatie A, het meest aantrekkelijk. De kans op hogere plaatsing is duidelijk groter dan die op lagere plaatsing, hoewel die laatste niet verwaarloosbaar is.

De verplaatsingen van de manipulanten in simulaties A en B staan in de volgende tabellen.

| van voorkeur | naar voorkeur | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4+ | totaal |
| 1 | 142.82 | | 2.22 | 11.49 | 156.53 |
| 2 | 151.69 | 29.21 | 5.81 | 21.28 | 207.99 |
| 3 | 6.30 | | | 3.18 | 9.48 |
| totaal | 300.81 | 29.21 | 8.03 | 35.95 | 374 |

Tabel 25: Simulatie A, verplaatsingen manipulanten, gemiddeld over 100 runs

| van voorkeur | naar voorkeur | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4+ | totaal |
| 1 | 132.02 | | | 27.62 | 159.64 |
| 2 | 114.34 | | | 81.03 | 195.37 |
| 3 | 17.38 | | | 6.61 | 23.99 |
| totaal | 263.74 | | | 115.26 | 379 |

Tabel 26: Simulatie B, verplaatsingen manipulanten, gemiddeld over 100 runs

In de nu volgende twee tabellen zijn de verplaatsingen uitgesplitst naar populariteit van de eerste drie scholen van de manipulanten.

| patroon populariteit | leerlingen | | | | percentages | | |
|---|---|---|---|---|---|---|---|
| | hoger | gelijk | lager | totaal | hoger | gelijk | lager |
| Pnn | 76.99 | 108.65 | 2.36 | 188 | 41 | 58 | 1 |
| PnP | 30.96 | 16.00 | 1.04 | 48 | 64 | 33 | 2 |
| PnQ | 16.05 | 8.95 | | 25 | 64 | 36 | |
| Qnn | 8.75 | 14.13 | 8.12 | 31 | 28 | 46 | 26 |
| QnP | 4.15 | 9.09 | 3.76 | 17 | 24 | 53 | 22 |
| QnQ | 0.93 | 2.84 | 0.23 | 4 | 23 | 71 | 6 |
| QPn | 15.41 | 8.63 | 21.96 | 46 | 34 | 19 | 48 |
| QPP | 2.25 | 3.24 | 6.51 | 12 | 19 | 27 | 54 |
| QPQ | 2.50 | 0.50 | | 3 | 83 | 17 | |
| totaal | 157.99 | 172.03 | 43.98 | 374 | 42 | 46 | 12 |

Tabel 27: simulatie A, verplaatsingen, aantallen manipulanten, gemiddeld over 100 runs

In simulatie A zijn de uitkomsten voor de leerlingen met Pnn, PnP, PnQ, QnQ en QPQ gunstig in de zin dat de kans op lager klein is. Voor Qnn en Qnp zijn de kansen op hoger en op lager nagenoeg gelijk. Voor QPn en QPP is de kans op lager duidelijk groter dan de kans op hoger. Deze uitspraken tegen de achtergrond dat alle 374 leerlingen manipuleren.

| patroon populariteit | leerlingen | | | | percentages | | |
|---|---|---|---|---|---|---|---|
| | hoger | gelijk | lager | totaal | hoger | gelijk | lager |
| Pnn | 59.23 | 97.65 | 31.12 | 188 | 32 | 52 | 17 |
| PnP | 25.49 | 11.84 | 10.67 | 48 | 53 | 25 | 22 |
| PnQ | 5.53 | 0.47 | | 6 | 92 | 8 | |
| Qnn | 8.22 | 0.85 | 21.93 | 31 | 27 | 3 | 71 |
| QnP | 2.25 | | 12.75 | 15 | 15 | | 85 |
| QPn | 20.58 | 10.88 | 28.54 | 60 | 34 | 18 | 48 |
| QPP | 0.03 | 1.90 | 4.07 | 6 | | 32 | 68 |
| QQn | 10.39 | 8.43 | 6.18 | 25 | 42 | 34 | 25 |
| totaal | 131.72 | 132.02 | 115.26 | 379 | 35 | 35 | 30 |

Tabel 28: simulatie B, verplaatsingen, aantallen manipulanten, gemiddeld over 100 runs

In simulatie B is alleen bij PnQ de kans op lager een kleine kans. Bij Pnn, PnP, PnQ en Qnn is de kans op hogere plaatsing groter dan de kans op lagere plaatsing. Bij Qnn, QnP, QPn en QPP ligt dat andersom. Deze uitspraken tegen de achtergrond dat alle 374 leerlingen manipuleren.

Simulatie D was toegevoegd om het effect van strategie 213 voor de 180 leerlingen QPn en QQn te onderzoeken. Onder de 413 leerlingen waren er 60 QPn en 25 QQn. Voor QPn waren de percentages hoger/gelijk/lager : 5/ 56/ 39, voor QQn 17/ 49/ 35. Dus onder deze omstandigheden ongunstig voor deze leerlingen.

Hierboven werden steeds gemiddelden gerapporteerd, dat is opgebouwd uit verplaatsingen van individuele leerlingen. Nu volgen gegevens over individuele leerlingen.

| simulatie | aantal manipulanten | beter in | | gelijk in | | slechter in | |
|---|---|---|---|---|---|---|---|
| | | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs |
| A | 374 | 142 | 92 (67) | 185 | 137 | 39 | 33 |
| B | 379 | 121 | 71 (51) | 141 | 90 | 108 | 90 |
| C | 344 | 16 | 13 ( 7) | 311 | 228 | 17 | 16 |
| D | 413 | 132 | 76 (42) | 221 | 165 | 56 | 35 |

Tabel 29: Simulaties A, B, C en D aantallen manipulanten, 100 runs

In simulatie A zijn er 92 van de 374 leerlingen (24.60%) die in minstens 75 van de 100 runs een hogere plaats verwerven. Tot die 92 behoren 67 leerlingen die in elke run een betere plaats krijgen. En 33 leerlingen (8.82%) die in minstens 75 runs een lagere plaatsing krijgen.

In simulatie C zijn er overigens 228 leerlingen (66.28%) waar vaak geen effect optreedt.

Op individueel niveau is maar voor weinig leerlingen een enigszins betrouwbare voorspelling van het effect van manipuleren te geven. Het effect wordt grotendeels door de lotingsvolgorde bepaald.

## 12. Meer manipuleren

Er zijn twee simulaties uitgevoerd, A2 en A3, waarbij de percentages leerlingen die niet blij zouden zijn met plaatsing op 2e of 3e voorkeur, zijn verdubbeld respectievelijk verdrievoudigd ten opzichte van de percentages in Tabel 20.

De volgende Tabel geeft de resultaten van de simulaties A, A2 en A3.

| simulatie | Aantal manipulanten | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|
| | 0 | 82.3 | 99.0 | 100.0 | 0 | 0 |
| A | 374 | 82.0 | 97.8 | 99.5 | 35.9 | 0 |
| A2 | 709 | 82.5 | 96.9 | 99.4 | 43.7 | 0 |
| A3 | 966 | 82.8 | 96.4 | 99.4 | 48.3 | 0 |

Tabel 30: Resultaten simulaties, gemiddelden over 100 runs

In beide simulaties worden alle leerlingen geplaatst, een tweede ronde is niet aan de orde. Nagenoeg alle leerlingen worden in hun top 3 geplaatst. De plaatsingen op 4e of lagere voorkeuren zijn uitsluitend manipulanten. De tabel is gebaseerd op de echte voorkeuren van de leerlingen, met de nominale voorkeuren zouden ze allen in hun top 3 zijn geteld.

De volgende Tabel bevat de effecten van het manipuleren op de plaatsing van de leerlingen, als percentages van de aantallen manipulanten respectievelijk niet manipulanten.

| simulatie | manipulanten | | niet manipulanten | |
|---|---|---|---|---|
| | hoger | lager | hoger | lager |
| A | 42.24 | 11.76 | 1.48 | 4.42 |
| A2 | 34.09 | 9.66 | 1.53 | 5.28 |
| A3 | 29.58 | 10.11 | 1.57 | 5.46 |

Tabel 31: Percentages verplaatsingen wegens manipuleren, aantallen leerlingen, gemiddelden over 100 runs

Naarmate er meer manipulanten zijn daalt de kans op succes, maar die is wel duidelijk groter dan die op een averechts resultaat.

Nu volgen gegevens over individuele leerlingen.

| simulatie | aantal manipulanten | beter in | | gelijk in | | slechter in | |
|---|---|---|---|---|---|---|---|
| | | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs |
| A | 374 | 142 | 92 | 185 | 137 | 39 | 33 |
| A2 | 709 | 192 | 124 | 449 | 295 | 53 | 36 |
| A3 | 966 | 193 | 138 | 696 | 387 | 58 | 39 |

Tabel 32: Simulaties A2 en A3 aantallen manipulanten, 100 runs

Al met al zijn de resultaten van deze simulaties niet alarmerend. De aantallen leerlingen met een bepaald effect in ≥ 75 runs zijn niet hoog: de effecten worden voornamelijk door de lotingsvolgorde bepaald.

De tot nu toe besproken simulaties zijn in feite gebaseerd op verondersteld zeer rationeel gedrag van de leerlingen. De leerling kijkt naar de populariteit van haar eerste drie plaatsen, maakt een schatting van de kansen op plaatsing op 2e of 3e voorkeur (info in deze notitie), en beslist dan om al dan niet een bepaalde strategie te volgen. De percentages van Dialogic beschrijven hoeveel leerlingen dat zouden kunnen doen op basis van “niet blij”.

Het is misschien irrationeel om uit te gaan van dat rationele gedrag, waar ook nog wat werk aan vastzit.

In de nu volgende simulaties zijn het willekeurige leerlingen die manipuleren, dus niet beperkt tot degenen die met echte voorkeuren vaak op 2e of 3e voorkeur komen. Dit is niet gebaseerd op de veronderstelling dat de leerlingen dat werkelijk zo zouden doen, het doel is alleen om te zien bij welke aantallen manipulanten het fout loopt, dat wil zeggen wanneer er veel leerlingen niet geplaatst worden.

De volgende Tabel geeft de resultaten van de simulaties in serie E.

| simulatie | Aantal manipulanten | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|
| | 0 | 82.3 | 99.0 | 100.0 | 0 | 0 |
| E15 | 894 | 82.2 | 97.8 | 99.5 | 37.6 | 0 |
| E30 | 1908 | 81.9 | 96.4 | 99.0 | 78.8 | 0 |
| E60 | 3651 | 80.6 | 93.3 | 97.2 | 208.6 | 6 |
| E75 | 4643 | 81.2 | 92.5 | 95.9 | 238.6 | 75 |
| E90 | 5543 | 83.0 | 92.6 | 94.9 | 210.8 | 179 |

Tabel 33: Resultaten simulaties, serie E, gemiddelden over 100 runs

De percentages manipulerende leerlingen zijn 15/30/60/75/90. De manipulanten loten 50/50 om strategie 1X2 danwel 1XY toe te passen.

De aantallen manipulanten zijn lager dan overeenkomen met de percentages van de 6622 leerlingen die hierin betrokken zijn omdat voor sommige leerlingen geen vervangende school X of Y beschikbaar was.

Bij 60% manipulanten zijn er 6 niet geplaatste leerlingen, dat loopt op tot 75 bij 75% en 179 bij 90% manipulanten. Zolang het percentage manipulanten niet veel hoger is dan 60 is er geen omvangrijke tweede ronde te verwachten.

De volgende Tabel bevat de effecten van het manipuleren op de plaatsing van de leerlingen, als percentages van de aantallen manipulanten respectievelijk niet manipulanten:

| simulatie | manipulanten | | niet manipulanten | |
|---|---|---|---|---|
| | hoger | lager[16] | hoger | lager |
| E15 | 16.88 | 4.72 | 1.85 | 4.62 |
| E30 | 16.24 | 4.60 | 1.38 | 8.83 |
| E60 | 10.50 | 6.61 | 0.66 | 16.85 |
| E75 | 9.76 | 7.81 | 0.46 | 21.64 |
| E90 | 9.24 | 8.40 | 0.50 | 22.17 |

Tabel 34: Verplaatsingen wegens manipuleren, percentages leerlingen, gemiddelden over 100 runs

Het valt op dat het manipuleren bij rond 80% van de manipulanten geen effect heeft.

Het percentage manipulanten met hogere plaatsing daalt naarmate er meer manipulanten zijn. Het percentage manipulanten met lagere plaatsing stijgt naarmate er meer manipulanten zijn. Het valt te betwijfelen of leerlingen geneigd zijn om te manipuleren als de kans op succes 10% is, tegen kans 8% op slechter resultaat. Dat doet zich voor vanaf het punt dat het manipuleren ook voor de organisatie onaantrekkelijk is wegens niet te verwaarlozen tweede ronde.

Leerlingen en organisatie hebben een gedeeld belang bij het tegengaan van hoge percentages manipulanten.

De percentages met hogere plaatsing zijn duidelijk lager dan die in Tabel 32. Met beredeneerd manipuleren is de kans op succes groter dan met willekeurig manipuleren.

Nu volgen gegevens over individuele leerlingen.

| simulatie | aantal manipulanten | beter in | | gelijk in | | slechter in | |
|---|---|---|---|---|---|---|---|
| | | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs |
| E15 | 894 | 143 | 113 | 707 | 676 | 45 | 36 |
| E30 | 1908 | 297 | 225 | 1530 | 1441 | 81 | 63 |
| E60 | 3651 | 335 | 222 | 3066 | 2924 | 224 | 190 |
| E75 | 4643 | 397 | 254 | 3897 | 3695 | 296 | 250 |
| E90 | 5543 | 456 | 249 | 4661 | 4352 | 338 | 222 |

Tabel 35: Simulaties serie E aantallen manipulanten, 100 runs

Er zijn relatief weinig manipulanten die in minstens 75 runs beter af zijn. Er zijn relatief veel leerlingen waarbij in minstens de helft van de runs geen verplaatsing optreedt. Dat ligt ook voor de hand: het zijn willekeurige leerlingen die manipuleren, met de echte voorkeuren wordt ruim 80% op eerste voorkeur geplaatst, dus daar kan geen verbetering zijn.

[16] Inclusief de niet geplaatsten. In een tweede ronde komen ze niet op echte eerste voorkeur, hooguit een enkeling op echte tweede voorkeur.

## 13. Contra manipulatie

Manipuleren is gericht op het verkrijgen van een betere positie dan met de echte voorkeuren te verwachten is. De gedachte is om zodanige voorkeuren op te geven dat plaatsing op lage voorkeuren onwaarschijnlijk is.

Een drastische manier om dat te bereiken zou kunnen zijn om helemaal geen lage voorkeuren te noemen, maar alleen de echte eerste of de eerste twee.

Maar met **Bmin** zijn leerlingen met weinig voorkeuren in het nadeel ten opzichte van leerlingen met meer voorkeuren. Zekerheid van plaatsing daalt met het aantal voorkeuren. Het advies is dat leerlingen minimaal drie/vier/vijf voorkeuren noemen. Het verzwijgen van lagere voorkeuren is onverstandig. Het minimum aantal zou ook als eis kunnen worden gesteld, want als veel leerlingen slechts enkele voorkeuren noemen zal er een tweede ronde komen, en dat is ook niet in het belang van de organisatie.

Een andere manier om plaatsing op lagere voorkeuren tegen te gaan is om een of meer echte lagere voorkeuren te vervangen door scholen waarvan verwacht wordt dat ze volbezet zullen worden door andere leerlingen, bijvoorbeeld de strategie 1XY. Als dat zo is zijn er twee mogelijkheden: de leerling komt op 1$^{e}$ voorkeur of de leerling wordt niet geplaatst. In het laatste geval komt de leerling in de tweede ronde en dan is de oorspronkelijke eerste voorkeur inmiddels volbezet en heeft het manipuleren gefaald. Dat is ook het geval als de leerling op een van de vervangende scholen wordt geplaatst.

In 2016 en 2017 waren er enkele scholen met minder capaciteit dan het aantal leerlingen met voorrang. Die scholen worden dus volledig bezet met leerlingen die daar voorrang hebben. Om dat te bewerkstelligen worden die scholen als onbruikbaar uit de voorkeuren van de andere leerlingen geschrapt. Dat heeft allereerst een technische reden ter vereenvoudiging van het optimaliseringsmodel.

Met een andere invulling van het model zouden die scholen in de voorkeuren mogen blijven staan. Maar dan zijn het de ideale vervangende scholen om mee te manipuleren. Want het is zeker dat de leerling zonder voorrang daar niet geplaatst zal worden. Het tegengaan van manipulatie is een tweede reden om die onbruikbare voorkeuren te schrappen.

Ter ontmoediging van manipuleren zou kunnen worden bepaald dat de lagere voorkeuren geen overtekende scholen mogen zijn.

Het verbod op overtekende scholen is een bruikbare optie voor derde en lagere voorkeuren.

Er zijn twee manieren om dat te implementeren.

De eerste mogelijkheid is om, voorafgaand aan de optimalisatie, de overtekende scholen uit de voorkeuren te schrappen. Dat verhoogt de stress bij het opstellen van de voorkleuren, want de lijstjes moeten niet te kort worden.

De tweede mogelijkheid is om alleen de scholen te verbieden die in de vorige editie overtekend waren. Dat lijstje kan gepubliceerd worden voor men de voorkeuren gaat opstellen.

De bepaling heeft effect op de voorkeuren die worden opgegeven. In de echte voorkeuren 2019 zijn er 2416 leerlingen met een overtekende school als derde voorkeur. Maar in de 1000 runs is geen van die leerlingen ooit op derde voorkeur geplaatst. De 2416 leerlingen zouden gedwongen zijn om hun

overtekende derde en lagere voorkeuren te schrappen. Dan is hun nieuwe derde voorkeur een niet overtekende school. Als de leerling daarop geplaatst wordt is dat op de echte 4e of lagere voorkeur.

In 2019 werden 2341 leerlingen met een overtekende school als derde voorkeur in de optimalisatie betrokken. Het multilateraal vervangen van die 3e voorkeur door de eerste niet overtekende school uit de 4e en lagere voorkeuren (en zonodig uit de bij het schooladvies passende scholen) had als effect dat rond 43 van deze leerlingen op 4e of lagere echte voorkeur werden geplaatst. Dat zijn de slachtoffers van het verbod.

Het verbod heeft, zoals bedoeld, invloed op de mogelijkheden om te manipuleren.

De leerlingen met voorkeuren QPn en QQn kunnen strategie 12X niet toepassen. Ze kunnen wel de 3e voorkeur vervangen door de meest populaire niet overtekende school uit de 4e en lagere voorkeuren (en zonodig die keus uit de bij het schooladvies passende scholen), dat is strategie 12x, waarbij x een hogere populariteit heeft dan de oorspronkelijke 3e voorkeur. Zij kunnen strategie 1X2 niet toepassen, want hun echte tweede voorkeur is overtekend.

Bij strategie 1Xy wordt voor y de meest populaire niet overtekende school uit de 4e en lagere voorkeuren gekozen (en zonodig die keus uit de bij het schooladvies passende scholen) , waarbij y een hogere populariteit heeft dan de oorspronkelijke 3e voorkeur.

Het vervangen van de overtekende 3e voorkeur door een niet overtekende school heeft als gevolg dat de 59 leerlingen die in Tabel 20 in de regel QQP staan, dan tot QQn worden gerekend. En analoog voor de overige groepen leerlingen.

Het verbod is geen optie voor de tweede voorkeur. Er zijn 2675 leerlingen met een overtekende school als tweede voorkeur. 418 van hen worden in een of meer van de 1000 runs op die voorkeur geplaatst. Het verbod zou deze leerlingen sterk kunnen benadelen.

De reeks simulaties F is analoog aan de reeks E, het verschil is dat in F geen overtekende scholen als 3e of lagere voorkeur mogen worden genoemd.

De volgende Tabel geeft de resultaten van de simulaties in serie F.

| simulatie | Aantal manipulanten | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|
| | 0 | 82.7 | 99.0 | 99.4 | 43.4 | 0 |
| F15 | 784 | 83.2 | 97.8 | 98.9 | 82.5 | 0 |
| F30 | 1646 | 83.2 | 96.3 | 98.3 | 132.4 | 0 |
| F60 | 3207 | 83.5 | 93.8 | 97.1 | 218.6 | 0 |
| F75 | 4048 | 83.9 | 92.2 | 96.2 | 287.7 | 0 |
| F90 | 4834 | 83.6 | 90.5 | 95.4 | 351.2 | 0 |

Tabel 36: Resultaten simulaties, serie F, gemiddelden over 100 runs

De percentages manipulerende leerlingen zijn 15/30/60/75/90. De manipulanten loten 50/50 om strategie 1X2 danwel 1XY toe te passen, waarbij op de 3e voorkeur geen overtekende school mag staan.

De aantallen manipulanten zijn lager dan overeenkomen met de percentages van de 6622 leerlingen die hierin betrokken zijn omdat voor sommige leerlingen geen vervangende school X of Y beschikbaar was. Het zijn ook minder manipulanten dan in de serie E.

Het uitsluiten van overtekende scholen op 3e voorkeur treft doel[17], er zijn geen ongeplaatste leerlingen. Naarmate er meer manipulanten zijn daalt het percentage in top 3 geplaatste leerlingen.

De volgende Tabel bevat de effecten van het manipuleren op de plaatsing van de leerlingen als percentages van de aantallen manipulanten respectievelijk niet manipulanten:

| simulatie | manipulanten | | niet manipulanten | |
|---|---|---|---|---|
| | hoger | lager | hoger | lager |
| F15 | 20.26 | 3.51 | 2.24 | 4.41 |
| F30 | 19.72 | 4.59 | 2.02 | 7.85 |
| F60 | 15.99 | 5.92 | 1.31 | 13.18 |
| F75 | 13.94 | 6.78 | 1.53 | 16.87 |
| F90 | 12.14 | 8.22 | 0.79 | 21.16 |

Tabel 37: Serie F, verplaatsingen wegens manipuleren, percentages leerlingen, gemiddelden over 100 runs

Het valt op dat het manipuleren bij rond 80% van de manipulanten geen effect heeft.

Het percentage manipulanten met hogere plaatsing daalt naarmate er meer manipulanten zijn. Het percentage manipulanten met lagere plaatsing stijgt naarmate er meer manipulanten zijn.

Het is opvallend dat in serie F de percentages manipulanten met hogere plaatsing groter zijn dan in serie E. Het zijn in serie F uiteraard andere leerlingen dan in serie E, die succes hebben.

Het percentage met hogere plaatsing is kleiner dan met beredeneerd manipuleren (Tabe 32).l

Nu volgen gegevens over individuele leerlingen.

| simulatie | aantal manipulanten | beter in | | gelijk in | | slechter in | |
|---|---|---|---|---|---|---|---|
| | | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs |
| F15 | 784 | 160 | 121 | 597 | 556 | 24 | 14 |
| F30 | 1646 | 312 | 231 | 1256 | 1158 | 67 | 41 |
| F60 | 3207 | 468 | 349 | 2553 | 2326 | 178 | 123 |
| F75 | 4048 | 518 | 372 | 3272 | 2973 | 228 | 163 |
| F90 | 4834 | 530 | 359 | 3909 | 3588 | 325 | 256 |

Tabel 38: Simulaties serie F aantallen manipulanten, 100 runs

Het verbod op het noemen van overtekende scholen op 3e en lagere voorkeur was ingesteld om te voorkomen dat een tweede ronde nodig zou zijn. De vraag is nu of dat verbod ook enig effect zou hebben als er alleen met reden wordt gemanipuleerd.

[17] Het zou mis kunnen lopen als slechts enkele niet overtekende scholen door veel leelingen worden genoemd.

Simulaties AA en BB zijn analoog aan simulaties A en B, het verschil is het verbod met betrekking tot de 3e voorkeur. De eerste resultaten staan in Tabel 39.

| simulatie | Aantal manipulanten | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|
| AA | 461 | 82.7 | 97.1 | 98.3 | 126.3 | 0 |
| BB | 313 | 82.9 | 97.4 | 98.7 | 100.2 | 0 |

Tabel 39: Resultaten simulaties AA en BB, gemiddelden over 100 runs

In simulatie AA zijn er meer leerlingen op voorkeur 4 of lager dan in simulatie A (Tabel 22). In geen van de runs was een tweede ronde nodig.

De volgende Tabel bevat de effecten van het manipuleren op de plaatsing van de leerlingen, als percentages van de aantallen manipulanten respectievelijk niet manipulanten:

| simulatie | manipulanten | | niet manipulanten | |
|---|---|---|---|---|
| | hoger | lager | hoger | lager |
| AA | 39.41 | 21.57 | 1.47 | 4.02 |
| BB | 46.63 | 16.02 | 1.81 | 4.03 |

Tabel 40: Simulaties AA en BB, verplaatsingen wegens manipuleren, percentages leerlingen, gemiddelden over 100 runs

Uit vergelijking met Tabel 24 blijkt dat met simulatie AA het percentage manipulanten met lagere plaatsing bijna verdubbeld is ten opzichte van simulatie A. In simulatie BB is dat percentage ongeveer gehalveerd ten opzichte van simulatie B.

Nu volgen gegevens over individuele leerlingen.

| simulatie | aantal manipulanten | beter in | | gelijk in | | slechter in | |
|---|---|---|---|---|---|---|---|
| | | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs |
| AA | 461 | 162 | 111 | 186 | 86 | 84 | 61 |
| BB | 313 | 131 | 95 | 132 | 56 | 40 | 19 |

Tabel 41: Simulaties AA en BB, aantallen manipulanten, 100 runs

Uit vergelijking met Tabel 29 blijkt dat er iets meer manipulanten in minstens 75 runs beter af zijn dan in simulaties A en B.

## 14. Toepassing op gegevens 2020

In de voorgaande secties is aan de hand van de gegevens 2019 onderzocht wat de effecten van verschillende vormen van strategisch gedrag in 2019 zouden zijn geweest. Dat is allemaal wijsheid[18] achteraf.

De strategiën zijn gebaseerd op de data 2019, ze moeten aan andere data worden getoetst.

Nu gaat het er om of strategie die in 2019 gunstig effect had, dat ook in 2020 zou hebben gehad. De vraag is of de lessen van het ene jaar toepasbaar zijn in het volgende jaar.

Om dat te ondezoeken worden de strategiën zonder meer op de gegevens 2020 toegepast. Dit houdt onder meer in dat de definitie van de groepen vaak 3 en vaak 2 gelijk is aan die in 2019.

De eerste resultaten zijn:

| simulatie | Aantal manipulanten | Percentage op eerste voorkeur geplaatst | Percentage In top 2 geplaatst | Percentage in top 3 geplaatst | Aantal op voorkeur 4 of lager geplaatst | Aantal niet geplaatst |
|---|---|---|---|---|---|---|
| | 0 | 79.5 | 93.4 | 98.5 | 92.0 | 25.0 |
| A | 264 | 79.0 | 93.4 | 98.1 | 99.1 | 45.0 |
| A2 | 541 | 79.8 | 93.1 | 97.6 | 84.9 | 96.0 |
| A3 | 768 | 80.4 | 92.9 | 97.5 | 62.6 | 131.0 |

Tabel 42: Data 2020, simulaties A, A2, A3, gemiddelden over 100 runs

De percentages verschillen niet heel veel van die zonder manipuleren. Het aantal niet geplaatste leerlingen stijgt met het aantal manipulanten.

De volgende Tabel bevat de effecten van het manipuleren op de plaatsing van de leerlingen, als percentages van de aantallen manipulanten respectievelijk niet manipulanten:

| simulatie | manipulanten | | niet manipulanten | |
|---|---|---|---|---|
| | hoger | lager[19] | hoger | lager |
| A | 24.81 | 31.62 | 2.61 | 3.04 |
| A2 | 17.09 | 27.89 | 3.20 | 3.03 |
| A3 | 14.17 | 24.05 | 3.70 | 2.81 |

Tabel 43: Data 2020, simulaties A, A2, A3, verplaatsingen wegens manipuleren, percentages leerlingen, gemiddelden over 100 runs

In vergelijking met dezelfde strategiën in 2019, zie Tabel 32, halveren de percentages manipulanten met hogere plaatsing en is er bijna een verdrievoudiging van de percentages met lagere plaatsing. Het percentage met lagere plaatsing is duidelijk groter dan dat met hogere plaatsing. Dit is onaantrekkelijk. De strategiën van 2019 leiden in 2020 tot veel teleurstelling.

[18] If any

[19] Inclusief de niet geplaatsten.

Nu volgen gegevens over individuele leerlingen.

| simulatie | aantal manipulanten | beter in | | gelijk in | | slechter in | |
|---|---|---|---|---|---|---|---|
| | | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs | ≥ 50 runs | ≥ 75 runs |
| A | 264 | 74 | 54 | 110 | 100 | 81 | 81 |
| A2 | 541 | 74 | 26 | 308 | 232 | 148 | 140 |
| A3 | 768 | 84 | 43 | 509 | 412 | 175 | 173 |

Tabel 44: Data 2020, simulaties A, A2, A3, aantallen manipulanten, 100 runs

In simulatie A is 20.5% van de manipulanten in minstens 75 runs beter af, in 2019 (Tabel 33) was dat 24.6%. In simulaties A2 en A3 zijn de betreffende percentages 4.8 en 5.6, in 2019 waren dat 17.5 en 14.3.

De verplaatsingen van de manipulanten in simulatie A staan in de volgende tabel.

| van voorkeur | naar voorkeur | | | | N | totaal |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4+ | | |
| 1 | 56.73 | 0.09 | 2.17 | 1.33 | 4.72 | 65.04 |
| 2 | 59.51 | 57.29 | 9.16 | 17.27 | 7.45 | 150.68 |
| 3 | 4.00 | 2.00 | 1.00 | 8.45 | 26.83 | 42.28 |
| totaal | 120.24 | 59.38 | 12.33 | 27.05 | 45.00 | 264.00 |

Tabel 45: Data 2020, simulatie A, verplaatsingen manipulanten, gemiddeld over 100 runs

In de nu volgend tabel zijn de verplaatsingen uitgesplitst naar populariteit van de eerste drie scholen van de manipulanten.

| patroon | leerlingen | | | | | percentages | | | |
|---|---|---|---|---|---|---|---|---|---|
| populariteit | hoger | gelijk | lager | niet | totaal | hoger | gelijk | lager | niet |
| Pnn | 37.70 | 37.30 | 3.00 | | 78.00 | 48 | 48 | 4 | |
| PnP | 11.14 | 31.76 | 4.10 | | 47.00 | 24 | 68 | 9 | |
| PnQ | 6.01 | 6.99 | | 1.00 | 14.00 | 43 | 50 | | 7 |
| Qnn | 3.65 | 12.19 | 8.16 | 1.00 | 25.00 | 15 | 49 | 33 | 4 |
| QnP | 1.20 | 7.44 | 7.36 | 1.00 | 17.00 | 7 | 44 | 43 | 6 |
| QnQ | | 14.00 | | 5.00 | 19.00 | | 74 | | 26 |
| QPn | 5.81 | 0.34 | 9.85 | 12.00 | 28.00 | 21 | 1 | 35 | 43 |
| QPP | | 2.00 | 6.00 | 2.00 | 10.00 | | 20 | 60 | 20 |
| QPQ | | 3.00 | | 9.00 | 12.00 | | 25 | | 75 |
| QQn | | | | 14.00 | 14.00 | | | | 100 |
| totaal | 65.51 | 115.02 | 38.47 | 45.00 | 264.00 | | | | |

Tabel 46: Data 2020, simulatie A, verplaatsingen, aantallen manipulanten, gemiddeld over 100 runs

## 15. Een virtuele discussie

In de loop van de jaren heb ik, naar aanleiding van notities over andere methoden, allerlei opmerkingen gekregen. Mijn reacties op een aantal daarvan volgen nu, qua methode geactualiseerd naar **Bmin**.

Je beweert dat je met **Bmin** "optimale" matchings berekent. Ik heb het even nagekeken, maar **Bmin** zet minder leerlingen op hun eerste voorkeur dan **RSD** doet. Dus **RSD** is beter en de meeste ouders zijn dat met mij eens.

> Als je het aantal plaatsingen op eerste voorkeur wilt maximaliseren zou je **RSD** moeten afwijzen en **Boston** moeten propageren, of liever nog **Fmax** (zie Bijlage 6). Met het focus geheel op 1e voorkeur negeer je de pijn van de plaatsingen op lage voorkeuren.
>
> In 2020 plaatste **RSD** 616 leerlingen (8.0%) buiten hun top 3, met **Bmin** zouden dat er 92 zijn (1.2%, allen op voorkeur 4), dus winst voor 616 - 92 = 524 leerlingen die met **Bmin** wel in hun top 3 komen. Daar staat dan tegenover dat het percentage op plek 1 daalt van 80.2 met **RSD** naar 79.5 met **Bmin**. Dat zijn rond 50 leerlingen die plek 1 verliezen.
>
> Uitgaande van de verdeling door **Bmin** kun je dus 50 leerlingen extra op plek 1 zetten door 524 leerlingen uit hun top 3 te verdringen. Dat is wat mij betreft niet te rechtvaardigen.
>
> Focus geheel op 1e voorkeur is ook verbonden aan de suggestie dat ouders en leerlingen alleen die plaats acceptabel zouden vinden. Maar plaatsing op plek 2 wordt niet door iedereen als een ramp ervaren. Volgens Dialogic is plek 2 voor rond 80% van de ouders acceptabel, de percentages zijn wel verschillend voor de verschillende schooladviezen.
>
> Misschien kun je ouders eens vragen of ze liever een methode hebben waarmee ze met kans 8.0% buiten hun top 3 komen dan een methode waarmee dat kans 1.2% is. In het eerste geval is de kans op plaatsing op eerste voorkeur 80.2%, in het tweede 79.5%. Dus een marginaal kleinere kans op plek 1 tegenover een relevant kleinere kans op een plek buiten top 3.

Die methode **Bmin** is wel erg ingewikkeld, ik zou dat aan de ouders niet kunnen uitleggen.

> In dat geval kun je het uitleggen maar beter aan anderen overlaten. De vraag is natuurlijk of ouders een methode tot in alle details zouden willen of moeten begrijpen. De eerste uitleg is het bekijken van de Tabellen 1 en 3. Het is net als met navigatie-software. Weinigen zullen weten hoe dat werkt, en zeker niet tot in de details. Dat is geen reden om het niet te gebruiken, in het gebruik blijkt dat het goed werkt. De tweede uitleg is het rustig doornemen van sectie 4. Daarna wordt het wiskunde.

Optimalisatie is geschikt voor navigatie en planning van produktie en dergelijke, besef je wel dat het hier over kinderen gaat en een belangrijke stap in hun leven?

> Ja, dat besef ik. En daarom is het van belang om een methode te gebruiken die kwalitatief goede matchings levert. 600 leerlingen buiten hun top 3 is bedroevende kwaliteit en voor al die kinderen een demotiverende slechte start.

**Bmin** is niet strategyproof en zal daarom aan strategie ten onder gaan, begin er dus maar niet aan.

> Het eerste is waar, maar het tweede is een vooroordeel. **Boston** is ook niet strategyproof en wordt toch op grote schaal gebruikt. "Aan strategie ten onder gaan" betekent wat mij betreft dat strategie tot een omvangrijke tweede ronde leidt, hetgeen voor de organisatie zeer

ongewenst is. Uit de simulaties 2019 blijkt dat er geen tweede ronde is als de strategie is beperkt tot leerlingen die een reden hebben om strategie toe te passen. Pas als veel leerlingen (meer dan te verwachten is) zonder reden strategie toepassen is een tweede ronde aan de orde. Maar in die tweede ronde worden geen leerlingen op hun echte eerste voorkeur geplaatst. Het is dus ook in het belang van de leerlingen dat er niet op grote schaal strategie wordt toegepast. Uit de simulaties 2019 blijkt ook dat met een verbod om overtekende scholen als derde of lagere voorkeur te noemen, geen tweede ronde nodig is. Het is een paardenmiddel, maar het werkt wel.

Ja, wij zijn voor het gebruik van **Bmin**, want dan gaan wij goed geld verdienen met het adviseren van ouders/leerlingen over het opstellen van hun voorkeurslijst.

Ik gun iedereen omzet en winst, maar ik heb een hekel aan bedrog.

Een (te) gemakkelijke commerciële aanpak zou zijn om systematisch de echte voorkeuren te adviseren, met wellicht een cosmetische aanpassing op de laagste voorkeuren. Dan zal 80% van de leerlingen op hun 1e voorkeur komen, net als met de volledige echte voorkeuren. Dit dit is bedrog en zal snel worden doorzien.

Een eerlijke aanpak zou zijn om op basis van *no cure, no pay* te werken. Omdat na afloop van de matching de (opgegeven) voorkeuren van de leerlingen en ook de gebruikte lotingnummers beschikbaar zijn (MEPS data) kan achteraf per leerling worden gecontroleerd of de geadviseerde voorkeuren inderdaad een betere plaatsing gaven dan de echte voorkeuren. Het ligt dan voor de hand dat de ouders slechts betalen als dat het geval is. Het ligt ook voor de hand dat de ouders compensatie krijgen als de gerealiseerde plaatsing slechter is dan met de echte voorkeuren zou zijn behaald.

Veel geld verdienen is gemakkelijker gezegd dan gedaan. Zolang er geen onderbouwing is beschouw ik die bewering als bluf, *wishful thinking* en pure retoriek om **Bmin** in diskrediet te brengen.

Overigens, de MEPS data mogen niet commercieel worden gebruikt. Daar heb ik voor getekend en ik neem aan dat dit voor iedereen geldt die deze data ter beschikking heeft gekregen.

Ik heb bij verschillende punten zo mijn twijfels of de manier waarop je **Bmin** hebt onderzocht methodologisch wel helemaal verantwoord is.

Van mijn kant zou ik graag weten welke twijfels bij welke punten dat zijn. Ik denk dat mijn aanpak zeker verdedigbaar is, ik heb zorgvuldig gewerkt en hopelijk bijvoorbeeld geen programmeerfouten gemaakt. Maar opmerkingen zijn altijd welkom.

Omdat je kennelijk hoge prijs stelt op methodologische zorgvuldigheid zul je het met me eens zijn dat, als ik een steekje zou hebben laten vallen, je geen oordeel over **Bmin** kunt uitspreken maar dat moet opschorten tot de kwestie is opgehelderd.

## 16. Afsluitende opmerkingen

Met de expliciete doelstelling “weinig leerlingen op lage voorkeuren” levert de methode **Bmin** beduidend veel betere matchings dan nu met **RSD** worden gevonden. Dit is ook het geval als er op beperkte schaal weloverwogen strategie zou worden toegepast. Als op grote schaal (meer dan te verwachten is) strategie zou worden toegepast, dan zou een ongewenste tweede ronde nodig kunnen zijn. De verwachtingen over de neiging om strategie toe te passen zijn gebaserd op enkele tabellen in de Tevredenheidsonderzoeken van Dialogic. Het kan nuttig zijn om op dit punt nader onderzoek te doen als voorbereiding op een keuze om eventueel **Bmin** te gaan toepassen. Overigens zou in 2019 een verbod op het noemen van overtekende scholen op derde en lagere voorkeuren ook afdoende zijn geweest om een tweede ronde te voorkomen.

## 17. Verwijzingen

[1] J.M. Anthonisse, J.K.Lenstra: Lineaire Programmering, Notities t.b.v. een cursus Operations Research in de Cultuurtechniek, 1973, rapport BC 7/73, Mathematisch Centrum (now CWI), Amsterdam.

[2] Jeanine Daems: Over schoolkeuze en toewijzing, Volkskrant, 9 maart 2013.

[3] Dan Gusfield and Robert W. Irving: The Stable Marriage Problem. The MIT Press, 1989.

[4] L. M. Klaren: Weighted bipartite matching and the assignment of students to highschools in Amsterdam, Technical University, Eindhoven, 2016.

[5] John Rawls: A Theory of Justice. Oxford University Press, revised edition1999 (original edition 1971).

[6] John Vande Vate. Linear programming brings marital bliss, Operations Research Letters 8 (1989) 147-153.

http://cgm.cs.mcgill.ca/~avis/courses/251/2012/ktlaerid/LP-bliss.pdf

**Een opmerking**

De column van Jeanine Daems citeerde een moeder uit Utrecht die het maar raar vond dat ze de eerste schoolvoorkeur van haar kind zou moeten verzwijgen om kans op een acceptabele plaats te hebben. Met de methode **Boston** kan dat voorkomen. Jeanine Daems riep de wethouders van Nederland op om iets beters te gaan gebruiken. Als de wethouders zich deskundig zouden hebben laten adviseren dan had de keus op **RSD** (Random Serial Dictatorship) of een variant van **DA** (Deferred Acceptance) of **TTC** (Top Trading Cycles) kunnen vallen. Bij die methoden loont het niet om een voorkeur te verzwijgen of anderszins valse voorkeuren in te dienen.

**Bijlage 1: Welke methode wordt nu eigenlijk gebruikt?**

Bovenstaande vraag heb ik in mijn notitie van 19 april 2019, over een andere methode, ook gesteld en beantwoordt. Omdat er nieuwe feiten zijn bespreek ik de kwestie opnieuw.

Er is een serieuze discrepantie tussen de Keuzegids (voor de leerlingen) en de Kernprocedure Overstap PO-VO ( Afspraken van schoolbesturen en gemeente Amsterdam over de overstap van primair onderwijs naar voortgezet onderwijs).

In de Keuzegids wordt de methode **RSD** uitgelegd, in de Kernprocedure staat dat **DA-STB** wordt gebruikt, en dat laatste is ten onrechte. **RSD** en **DA-STB** zijn geen synoniemen, het zijn twee totaal verschillende methoden. Wel leveren de twee methoden onder de Amsterdamse omstandigheden dezelfde verdeling van de leerlingen over de scholen. Dus qua resultaat zijn ze equivalent. Maar dat kan geen excuus zijn om het verkeerde antwoord te geven op de vraag hoe het resultaat wordt berekend.

Het ietwat saaie relaas hieronder leidt tot de volgende conclusies:

1. In 2016 heeft de firma Skrepr, onder begeleiding van de heer Van der Klaauw, verbonden aan de Vrije Universiteit te Amsterdam, een ict-applicatie voor loting & matching gebouwd die uitgaat van de rekenmethode *Random Serial Dictatorship* (RSD).
2. Deze applicatie wordt sinds 2016 voor OSVO gebruikt.
3. Zeker in 2020 was het voor OSVO duidelijk dat RSD werd gebruikt. In de Kernprocedure 2020 staat dat Skrepr en de heer Van der Klaauw een verklaring met die inhoud ondertekenen.
4. Toch staat in de Kernprocedures van 2020 en 2021 dat **DA-STB** wordt gebruikt.
5. In de Beschrijvende Analyse die jaarlijks na afloop van de Loting & Matching door de heer Van der Klaauw en anderen wordt opgesteld staat sinds 2016 ten onrechte dat **DA-STB** is gebruikt.

Er wordt sinds 2016 systematisch een verkeerde voorstelling van zaken gegeven. Dat leidt tot misverstanden en zou gecorrigeerd moeten worden.

Nu het relaas.
Het gaat eerst over Keuzegids en Kernprocedure, daarna over een brief en tenslotte over de jaarlijkse Beschrijvende Analyse.

**Keuzegids en Kernprocedure**
In de Keuzegids[20] wordt uitgelegd hoe de loting & matching verlopen:

1. Elke leerling krijgt een lotnummer toegewezen door de computer. Hierbij is een notaris aanwezig.
2. In volgorde van het lotnummer worden alle leerlingen geplaatst. Dus als jouw lotnummer aan de beurt is, wordt er gekeken of de VO-school die boven aan je voorkeurslijst staat plek heeft.

[20] Identieke tekst in 2020 en 2021.

3. Is die VO-school al vol, dan wordt gekeken naar je tweede voorkeur, en als die vol is naar je derde voorkeur, etc.
4. Als jij een plek hebt, dan gaat het centrale systeem door met de leerling met het volgende lotnummer.
5. Op deze manier worden alle leerlingen in volgorde van hun lotnummer geplaatst op de VO-school die het hoogst op hun voorkeurslijst staat en nog plek heeft op het moment dat zij aan de beurt zijn.

Er staat niet bij dat dit gebeurt nadat de leerlingen met voorrang zijn geplaatst, maar dat is geen bezwaar. Er staat ook niet bij dat deze methode een naam heeft. Die naam is RSD (Random Serial Dictatorship). De leerlingen komen in lotingsvolgorde aan de beurt en dan is het: wie het eerst komt die het eerst maalt. Wie met Google zoekt op Random Serial Dictatorship vindt allerlei sites die ditzelfde verhaal vertellen.

In de Kernprocedure wordt jaarlijks, sinds 2016, *ten onrechte* beweerd dat de methode **DA-STB** wordt gebruikt.

Dit is een van de twee varianten van DA (Deferred Acceptance). In beide varianten wordt in meerdere ronden gewerkt. In elke ronde melden leerlingen zich aan bij scholen[21], zij worden daar niet geplaatst maar komen in de wacht te staan voor zover de school capaciteit heeft. In de tweede en latere ronde kunnen leerlingen die al in de wacht stonden alsnog worden geweigerd. Ze worden dan verdrongen door nieuwe aanmeldingen. Dit is duidelijk een heel wat ingewikkelder proces dan **RSD**. Pas na de laatste ronde worden de leerlingen die dan in de wacht staan definitief geplaatst. Vandaag de naam: Deferred Acceptance = Uitgestelde plaatsing.

Het verschil met **RSD** is evident, bij **RSD** is geen sprake van enig uitstel in plaatsing van leerlingen.

In **Kernprocedure 2017** (pagina 4) staat te lezen:

> Centrale loting & matching
> De leerlingen worden bij de centrale loting & matching zo veel mogelijk gelijktijdig geplaatst door middel van het algoritme Deferred Acceptance-Single Tie Break (DA-STB). Hier krijgen de leerlingen onder toezicht van de notaris van de computer één lotnummer. Leerlingen worden op volgorde van hun lotnummer geplaatst op de VO-school die zo hoog mogelijk op hun voorkeurslijst staat.

Onder de naam **DA-STB** wordt hier **RSD** beschreven. Fout en verwarrend. Als je **RSD** de naam **DA-STB** geeft, welke naam ga je **DA-STB** dan geven?

In de **Kernprocedure 2018** (pagina 29) staat:

> Het OSVO heeft voor de Matching gekozen voor een algoritme dat gebaseerd is op het zogenaamde Deferred Acceptance (DA) algoritme. Gelijk aan de overstap in 2016 en 2017 zal ook in 2018 de variant Single Tiebreak (DA-STB) worden gebruikt. Bij deze wijze van centraal loten en matchen krijgt elke leerling een (1) lotnummer. In volgorde van de lotnummers worden de leerlingen geplaatst op de hoogst mogelijke plek op hun voorkeurslijst.

Wederom fout en verwarrend.

[21] Dit is beeldend taalgebruik, alles wordt door de computer gedaan.

In de **Kernprocedure Overstap PO-VO 2019** staat op pagina 35:

> Vereniging OSVO heeft voor de Centrale Loting & Matching gekozen voor een algoritme dat gebaseerd is op het beginsel van Deferred Acceptance (DA). Voor het schooljaar 2019/2020 betreft dit de variant Single Tie- break (STB). Deze variant van het algoritme wordt dus ook wel aangeduid als DA-STB.

Op pagina 36 staat, onder het kopje **Uitvoering**:

> De definitieve Centrale Loting & Matching verloopt eveneens volgens het DA-STB algoritme, dat in grote lijnen als volgt functioneert:
>
> 1. Elke leerling (in feite elk leerlingnummer) die meedoet aan de Centrale Loting & Matching krijgt een lotnummer toegekend (de regels rondom voorrang worden toegepast bij het toekennen van de lotnummers) door de computer in aanwezigheid van een notaris.
> 2. De leerlingen worden vervolgens in volgorde van hun lotnummer geplaatst op een zo hoog mogelijke plek op de voorkeurslijst waar nog plek is op het moment dat de leerling aan de beurt is. Dus eerst wordt gekeken of er nog plek is op VO-school 1, zo ja, dan wordt de leerling geplaatst. Als de VO-school op nummer 1 geen plek meer heeft, dan wordt gekeken naar de VO- school op nummer 2, enz.
> 3. Als een leerling onverhoopt niet geplaatst kan worden op een VO-school van zijn lijst …

Die grote lijnen zijn helemaal niet hoe **DA-STB** functioneert. Het is wel exact hoe de methode **RSD** functioneert.

In de **Kernprocedure 2020** wordt een ander verhaal verteld.

Op pagina 30 staat:

> Er wordt gebruik gemaakt van het algoritme DA-STB. Bij deze vorm van Centrale Loting & Matching komen veel leerlingen op hun VO-school van eerste voorkeur.

Dit is de enige plaats in het document waar **DA-STB** wordt genoemd.

Op pagina 38 staat:

> Onder begeleiding van de heer Van der Klaauw, verbonden aan de Vrije Universiteit te Amsterdam, heeft OSVO, namens de VO-besturen, uitgaande van een uitkomst die voldoet aan *Deferred Acceptance - Single Tie-Breaking Mechanism* en het *Random Serial Dictatorship* een matchingsprocedure en bijbehorend Algoritme laten ontwikkelen. OSVO heeft, namens de VO-besturen, aan Skrepr gevraagd een ict-applicatie (LMS-applicatie) te bouwen die uitgaat van deze matchingsprocedure, de rekenmethode *Random Serial Dictatorship* en het Algoritme. De ELK-applicatie draagt de benodigde gegevens over aan de LMS-applicatie.

Hier wordt Skrepr als leverancier van de software genoemd. Interessant is dat wordt uitgegaan van een uitkomst die

> voldoet aan *Deferred Acceptance - Single Tie-Breaking Mechanism* en het *Random Serial Dictatorship*

Het is namelijk zo dat, als er geen voorrang (meer) in het spel is (de leerlingen met voorrang zijn al geplaatst) en als voor beide methoden dezelfde lotingsvolgorde wordt gebruikt, **DA-STB** en **RSD** dezelfde matching leveren. Het ligt dan voor de hand om het eenvoudige **RSD** te implementeren[22].

Op pagina 44/45 staat:

> Skrepr zal voorafgaand aan de Definitieve Matching 1 een verklaring ondertekenen inhoudende dat:
> a. de LMS-applicatie is ontwikkeld op basis van de door OSVO benoemde uitgangspunten;
> b. de kwaliteit van LMS-applicatie zodanig is dat de Definitieve Matching behoort overeen te komen met de ingevoerde leerlinggegevens, de ingevoerde schoolgegevens 2020-2021 en de Benoemde Uitkomst;
> c. de LMS-applicatie bij de Definitieve Matching 1 lotnummers toekent aan leerlingen;
> d. de LMS-applicatie bij de Definitieve Matching 1 uitgaat van deze matchingsprocedure, de rekenmethode *Random Serial Dictatorship* en het Algoritme en is afgestemd met de heer Van der Klaauw van de Vrije Universiteit te Amsterdam.
> Skrepr zal deze verklaring overhandigen aan OSVO en OSVO zal de notaris verzoeken deze verklaring in bewaring te nemen.
>
> De heer Van der Klauw zal voorafgaand aan de Definitieve Matching 1 een verklaring ondertekenen inhoudende dat aan Skrepr ten behoeve van de door hen ontwikkelde LMS-applicatie een algoritme is aangeleverd dat uitgaat van deze matchingsprocedure, de rekenmethode *Random Serial Dictatorship* en het Algoritme. De heer Van der Klaauw zal deze verklaring overhandigen aan OSVO en OSVO zal de notaris verzoeken deze verklaring in bewaring te nemen.

**Zowel Skrepr als de heer van der Klaauw tekenen dus een verklaring die inhoudt dat RSD wordt gebruikt.** En **RSD** wordt al sinds 2016 gebruikt.

Tenslotte nog de **Kernprocedure 2021**, daarin wordt helemaal geen verhaal verteld.

Op pagina 28 staat:

> Er wordt gebruikgemaakt van het algoritme DA-STB.

Dit is de enige plaats in het document waar **DA-STB** wordt genoemd. Hoe het algoritme werkt wordt niet verteld. Er is geen enkele verwijzing naar **RSD**.

Het is theoretisch mogelijk dat voor 2021 inderdaad **DA-STB** wordt geprogrammeerd en toegepast. Dat lijkt me onwaarschijnlijk want verspilling van tijd en geld. Het is voor mij wel zeker dat die mededeling op pagina 28 falikant fout is. Het begint op bewust bedrog te lijken.

In het kader van de nagestreefde transparantie zou OSVO dit moeten corrigeren en gewoon zeggen hoe het zit.

Zoals gezegd leveren **DA-STB** en **RSD** onder de omstandigheden van VO Amsterdam hetzelfde resultaat. De equivalentie qua resultaat is geen excuus om het verkeerde antwoord te geven op de vraag hoe het resultaat wordt berekend.

---

[22] Op 11 mei 2018 bracht ik een bezoek aan Skrepr (te Urk) en werd mij verzekerd dat inderdaad RSD werd toegepast, sinds 2016.

Het is alsof iemand sinds 2016 alsmaar vertelt dat zijn route van Amsterdam naar Leeuwarden *gebaseerd* is op rijden door Noord-Holland en over de Afsluitdijk. Gevraagd naar details blijkt dat hij altijd langs Lelystad en Lemmer komt. En als dan wordt gezegd: “Je komt dus niet over de Afsluitdijk” is zijn antwoord: “Maar ik kom wel in Leeuwarden”, het is toch equivalent! En hij blijft elk jaar weer zeggen dat hij door Noord-Holland gaat. Dat is toch wel een beetje belachelijk.

**Een brief**

Op 21 augustus schreef de voorzitter van OSVO (de heer Oudkerk) een brief aan de wethouder (mevrouw Moorman). Daarin werden twee scenario’s gepresenteerd die eventueel in 2021 zouden kunnen worden toegepast (het kwam er niet van).

> *Scenario 1: Eerst Boston, daarna Deferred Acceptance (DA) (met single tie-breaking)*
> Leerlingen melden zich bij één of meerdere scholen aan. Scholen met meer plaatsen dan aanmeldingen, nemen iedereen die zich heeft aangemeld aan. Scholen met minder plaatsen dan aanmeldingen, nemen leerlingen aan op basis van loting. De beschikbare plekken worden willekeurig onder de aangemelde leerlingen verdeeld. Dit heet in matchingssysteem-termen “boston”.
>
> Leerlingen die niet kunnen worden geplaatst op de school waar ze zich hebben aangemeld gaan naar een tweede ronde.
>
> In deze ronde worden via een matchingssysteem, leerlingen alsnog geplaatst op een school waar nog plek is. Dit gebeurt door elke leerling een willekeurig lotnummer te geven en op volgorde van lotnummer de leerlingen te plaatsen op een school van voorkeur waar nog plek is (DA met single tie-breaking).

Wederom dezelfde fout, de tweede ronde is **RSD** maar wordt **DA-STB** genoemd.

> *Scenario 2: Deferred Acceptance (DA) (met single tie-breaking) met voorrang op basis van woonplaats.*
> Via een centraal systeem melden leerlingen zich aan voor een aantal scholen in Amsterdam. Op basis van de postcode van de basisschool worden leerlingen in een aantal (bijv. 3) groepen verdeeld. De postcode van de basisschool wordt dan als richtlijn gebruik voor de woonplaats van de leerling.
>
> Leerlingen worden op volgorde van voorrang gematcht met een school van voorkeur.
>
> Binnen dit scenario wordt gedacht aan drie hoofdgroepen op basis van woonplaats:
> 1. Leerlingen woonachtig in Amsterdam
> 2. Leerlingen woonachtig buiten Amsterdam met geen of onvolledig aanbod voortgezet onderwijs
> 3. Leerlingen woonachtig buiten Amsterdam met volledig aanbod voortgezet onderwijs
>
> Daarbij wordt in eerste instantie uitgegaan van:
> ● Leerlingen behorende bij groep 1 voorrang hebben op leerlingen die behoren tot groep 2 en dat de leerlingen uit groep 2 voorrang hebben op de leerlingen uit groep 3.
> ● Bij voorrang op basis van onderwijsconcept geldt hetzelfde; leerlingen uit groep 1 voorrang op leerlingen uit groep 2 en leerlingen uit groep 2 voorrang op leerlingen uit groep 3.

> Verdere uitwerking moet aantonen of dit haalbaar en tot de gewenste uitkomsten leidt of dat er nog een nieuwe subgroep dan wel volgorde van voorrang moet komen.

Dit is wel correct. Nu is **RSD** niet toepasbaar want er is voorrang mogelijk op alle voorkeursplaatsen. Voor dit soort situaties is DA ontworpen, en ligt STB voor de hand.

In Scenario 2 kom je niet ver met "**DA-STB**" zoals die in Scenario 1, tweede ronde, is omschreven.

De argeloze lezer denkt bij scenario 2: aha, **DA-STB** toepassen, ik heb in scenario 1, tweede ronde, gezien hoe dat werkt, simpel. Maar dan komt die lezer bedrogen uit, want wat in scenario 1 **DA-STB** werd genoemd negeert alle voorrang en is dus niet toepasbaar.

**Beschrijvende Analyse**

Jaarlijks wordt, na afloop van de Loting & Matching, een Beschrijvende Analyse[23] opgesteld waarin allerlei cijfermateriaal over voorkeuren en plaatsingen wordt gepresenteerd. In 2020 waren de opstellers de hoogleraren Bas van der Klaauw (VU) en Hessel Oosterbeek (UvA). In de jaren daarvoor waren er naast hen nog twee andere personen bij betrokken.

In 2015 stond in de inleiding van het rapport dat DA-MTB was gebruikt, en dat klopte. In de jaren 2016, 2017, 2018 en 2019 werd in de inleiding steevast gemeld dat **DA-STB** werd gebruikt, wat dus niet het geval was. In de inleiding van het rapport over 2020 staat:

> Sinds 2015 wordt daarbij gebruik gemaakt van het zogenoemde student-proposing Deferred Acceptance (DA) systeem.

Dat klopt voor 2015, maar niet voor de jaren daarna, sinds 2016 wordt **RSD** toegepast, en minstens één van de auteurs weet dat.

De heer Van der Klaauw was begeleider bij de ontwikkeling van de software door Skrepr. In maart 2020 ondertekende hij een verklaring dat **RSD** werd gebruikt, in mei van datzelfde jaar was het opeens DA dat werd toegepast, volgens de inleiding van de Beschrijvende Analyse 2020. Sinds 2016 wordt die onwaarheid gedebiteerd.

**Tenslotte**

Er was eens een man die bedreven was in het bouwen van bescheiden meubilair. Apetrots liet hij een nieuw geslaagd werkstuk aan vrienden en bekenden zien en riep dan: heb ik dat niet even mooi in elkaar geschroefd? Toen merkte iemand op dat hij alles gespijkerd had, er was geen schroef aan te pas gekomen. Op deze kritiek had hij als repliek: het is toch een mooi ding, schroeven of spijkers is toch equivalent! En de volgende keer was het weer: heb ik dat niet even mooi in elkaar geschroefd?
De man had een buurman die ook wat wilde gaan knutselen en een tijdje bij hem in de leer kwam. Op een dag ging de buurman naar de ijzerwinkel om schroeven te kopen. De medewerker van de winkel gaf hem schroeven van de gevraagde maat, waarop de buurman zei: die wil ik niet met die rare ribbels, ik moet gladde hebben.

[23] In de loop van de jaren zijn ook andere titels gebruikt.

Onlangs hoorde ik over een medewerker (M) van een adviesbureau dat hij een gesprek had met een functionaris (F) van een scholenorganisatie.
F: we willen een andere methode voor loting & matching gaan gebruiken, het wordt **DA-STB**.
M: is uit voorraad leverbaar, maar er is bij jullie geen voorrang aan de orde, dus waarom niet **RSD**, die is ook nog heel wat gemakkelijke aan ouders en leerlingen uit te leggen.
F: die willen we echt niet, we hebben gekozen voor een methode waarmee de leerlingen in lotingsvolgorde naar hun hoogste nog beschikbare voorkeur gaan.
M: dat is precies wat **RSD** doet.
F: ik moet eens even gaan terugkoppelen.

Zulke bizarre situaties krijg je als de dingen niet bij hun echte naam worden genoemd.

**Bijlage 2: Een overbodige omrekening**

Op 2 april 2020 maakte OSVO in een persbericht de resultaten van de Centrale loting en matching bekend. In Tabel 2.1 worden de resultaten van de eerste ronde vergeleken met het resultaat van een telling van de plaatsingen zoals die in het MEPS bestand zijn vastgelegd. In dat bestand staan de plaatsingen zoals op basis van de voorkeuren door **RSD** zijn berekend, plus de 32 handmatige plaatsingen op grond van plaatsinggarantie, samen dus ook het resultaat van de eerste ronde.

| Voorkeur | Persbericht | | MEPS bestand | |
|---|---|---|---|---|
| | Aantal | Cumulatief % | Aantal | Cumulatief % |
| 1 | 6208 | 80.92 | 6183 | 80.59 |
| 2 | 629 | 89.12 | 600 | 88.41 |
| 3 | 252 | 92.42 | 264 | 91.85 |
| 4 | 163 | 94.54 | 171 | 94.08 |
| 5 | 142 | 96.39 | 142 | 95.93 |
| 6+ | 268 | 99.89 | 303 | 99.89 |
| ongeplaatst | 9 | 100.00 | 9 | 100.00 |

Tabel 2.1: Vergelijking persbericht met MEPS bestand

Het persbericht geeft een iets geflatteerd beeld ten opzichte van het bestand. De verschillen zijn niet groot en nauwelijks relevant voor het algemene oordeel over de resultaten.

De verklaring van de verschillen gaat aan de hand van een voorbeeld.

Tabel 2.2 bevat de voorkeuren van een leerling volgens MEPS en volgens het persbericht.

| Voorkeur in MEPS | School | Afdeling | Voorkeur in persbericht |
|---|---|---|---|
| 1 | Metis Montessori Lyceum | Technasium - v.a. havo | 1 |
| 2 | Metis Montessori Lyceum | v.a. havo | 1 |
| 3 | St. Nicolaaslyceum | v.a. havo | 2 |
| 4 | Spinoza Lyceum | Muziekplus - v.a. havo | 3 |
| 5 | Spinoza Lyceum | v.a. havo | 3 |
| 6 | Fons Vitae Lyceum | v.a. havo | 4 |
| 7 | Gerrit van der Veen College | v.a. havo | 5 |
| 8 | Montessori Lyceum Amsterdam | v.a. havo | 6 |
| 9 | Cartesius 2 | v.a. havo | 7 |

Tabel 2.2: Voorkeur volgens MEPS en volgens persbericht

De leerling is op haar 9e voorkeur (Cartesius 2) geplaatst en dat wordt in het persbericht als plaatsing op 7e voorkeur gerapporteerd. Cartesius 2 is inderdaad de zevende unieke school in dit lijstje van voorkeuren. Ten behoeve van het persbericht worden de voorkeursplaatsen omgerekend naar voorkeurscholen.

In het persbericht staat:

80.92% kreeg een aanbod van de voortgezet onderwijs school van hun eerste voorkeur.

Letterlijk genoemen is dat waar, maar voor 25 leerlingen is dat geen aanbod voor hun eerste voorkeur.

In het persbericht staat ook:

96.39% kreeg een aanbod van een voortgezet onderwijs school in de top 5 van hun keuze.

Die top 5 van hun keuze bestaat uit de eerste 5 verschillende scholen in hun lijst van voorkeuren, het is niet de top 5 van hun voorkeuren. Die 96.39% bevat leerlingen die lager dan hun 5$^{e}$ voorkeur zijn geplaatst.

Deze omrekening van plaats naar school flatteert de resultaten enigszins en draagt niet bij aan de transparantie van het geheel. Het omrekenen kan wat mij betreft beter worden nagelaten, minder werk en meer duidelijkheid.

**Bijlage 3: Waarderingen voor voorkeursplaatsen**

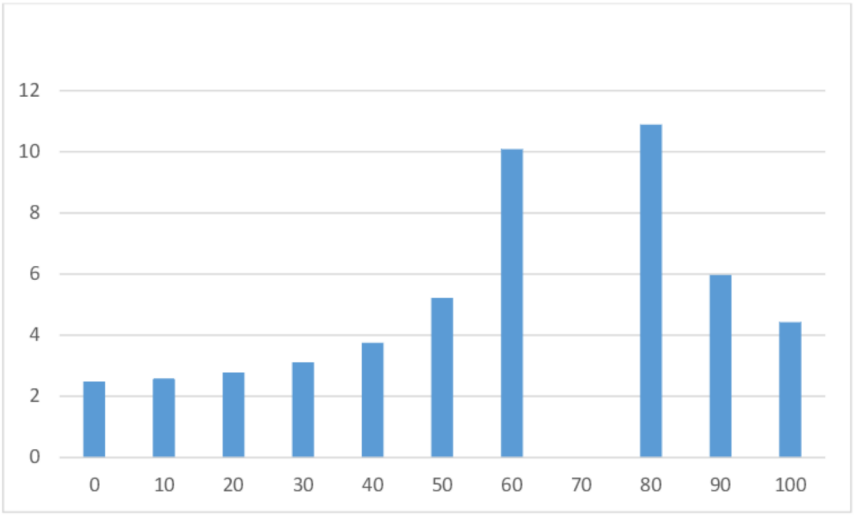
De waarderingen voor de verschillende plaatsen kunnen zonder bezwaar zo worden genormeerd dat $W_1 = 100$ en ze vormen uiteraard een niet-stijgende rij. Omdat plaatsing op $4^e$ en lagere voorkeuren ongewenst is kan $W_4 = 0$ worden gesteld. Gemakshalve wordt hierna ook $W_3 = 0$ gebruikt.

Bij een bepaalde leerling had een manipulatie, waarmee 100 runs werden gedaan, het volgende resultaat.

| | 1 | 2 | 3 | 4 | totaal |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | 14 | 80 | | 6 | 100 |
| 3 | | | | | |
| 4 | | | | | |

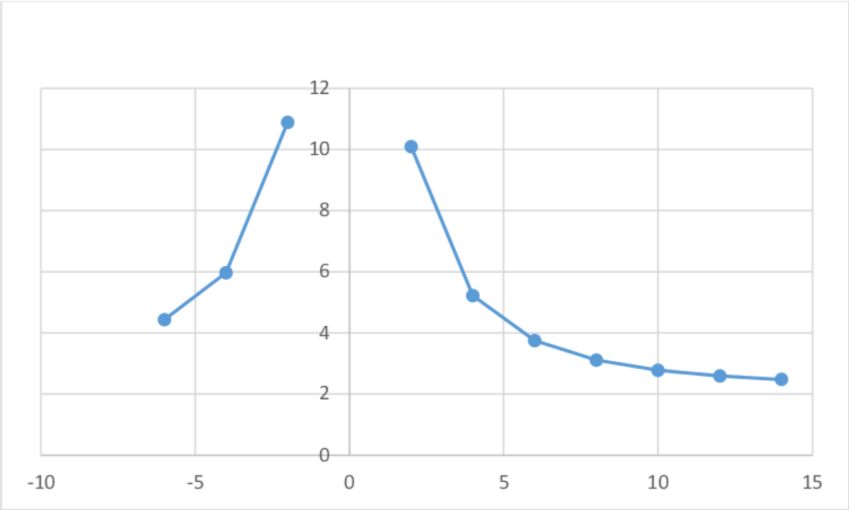
Tabel 3-1: Voorbeeld individuele verplaatsingstabel

Met de echte voorkeuren was het altijd plaatsing op $2^e$ voorkeur, met de strategie 14 maal op $1^e$ en 6 maal op $4^e$ voorkeur. De hogere plaatsing is gemiddelde winst van $14*( 100 – W_2 )$, de lagere plaatsing is gemiddeld verlies van $6* W_2$. Samen is dit $1400 – 20 * W_2$. De winstverwachting is $14 – 0.20 * W_2$. Met $W_2 = 70$ is er winst noch verlies. Als $W_2 > 70$ is er gemiddeld verlies, als $W_2 < 70$ is er gemiddeld winst. In het eerste geval zou de manipulatie niet moeten worden toegepast. Onderstaande figuur toont de variatiecoëfficiënt van de winst als funnctie van $W_2$




Figuur 3-1 : Voorbeeld variatiecoëfficiënt van de winst als funnctie van $W_2$

De volgende figuur toont de variatiecoëfficiënt van de winstverwachting als funnctie van de winstverwachting.




Figuur 3-2: Voorbeeld variatiecoëfficiënt van de winstverwachting als funnctie van de winstverwachting.

De voorspellende waarde van de winstverwachting daalt naarmate die verwachting dichter bij 0 komt, althans in dit voorbeeld.

Voor wat betreft de berekening is de verplaatsing van 2$^{e}$ naar 4$^{e}$ voorkeur equivalent met verplaatsing van 2$^{e}$ naar 3$^{e}$, gevolgd door verplaatsing van 3$^{e}$ naar 4$^{e}$. De laatste verplaatsing is neutraal qua winst en verlies.

De verplaatsingstabel is voor de berekeningen te gebruken:

| | 1 | 2 | 3 | 4 | totaal |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | 14 | 80 | 6 | | |
| 3 | | | | (6) | |
| 4 | | | | | |

Tabel 3-2 : voorbeeld individuele berekening, eerste stap

De 6 in cel [2,3] wordt verplaatst naar cel [3,2] en verandert van teken.

| | 1 | 2 | 3 | 4 | totaal |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | 14 | 80 | | | |
| 3 | | -6 | | | |
| 4 | | | | | |

Tabel 3-3 : voorbeeld individuele berekening, tweede stap

Nu bevat de nevendiagonaal de coefficienten om de winst te berekenen.

Deze berekeningen kunnen analoog op elke verplaatsingstabel worden uitgevoerd. Als de nevendiagonaal geen negatieve getallen bevat is er gemiddeld geen verlies, ongeacht de waarderingen voor de plaatsen, zijn er geen positieve getallen dan is er gemiddeld geen winst. Zijn er zowel positieve als negatieve, dan is de gemiddelde winst afhankelijk van de waarderingen.

Als de getallen in absolute waarde klein zijn, dan rijst de vraag of er, statistich gesproken, wel een bruikbaar resultaat is.

Uit de individuele verplaatsingstabellen in de simulaties A, B, C en D blijkt het volgende met betrekking tot winst en waarderingen voor plaatsen.

| simulatie | nul | pos | mix | neg | totaal |
|---|---|---|---|---|---|
| A | 30 | 260 | 53 | 31 | 374 |
| B | 17 | 198 | 54 | 110 | 379 |
| C | 118 | 187 | 6 | 33 | 344 |
| D | 60 | 260 | 39 | 54 | 413 |

Tabel 3-4: Per simulatie aantallen leerlingen met bepaald effect van het manipuleren, 100 runs

De kolom nul bevat de aantallen leerlingen waarbij de manipulatie gemiddeld winst = 0 oplevert, ongeacht de waarderingen voor de voorkeursplaatsen.

De kolom pos bevat de aantallen leerlingen waarbij de manipulatie gemiddeld geen verlies oplevert, ongeacht de waarderingen voor de voorkeursplaatsen.

De kolom mix bevat de aantallen leerlingen waarbij de gemiddelde winst afhankelijk is van de waarderingen voor de voorkeursplaatsen.

De kolom neg bevat de aantallen leerlingen waarbij de manipulatie gemiddeld geen winst oplevert, ongeacht de waarderingen voor de voorkeursplaatsen.

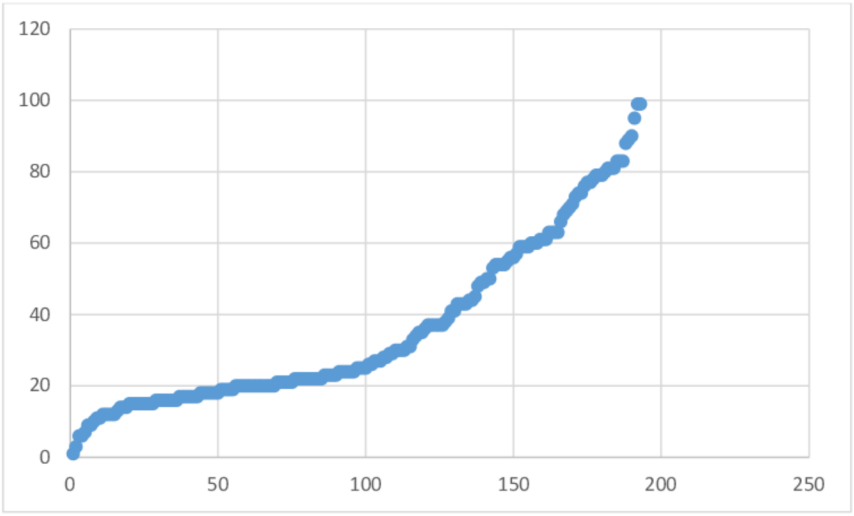
Hierbij is waarde = 100 voor $1^e$ voorkeur en waarde = 0 voor $3^e$ en lagere voorkeuren.

Enkele opmerkingen bij simulatie A, daarna ook nog resultaten van andere simulaties.

Onder de 30 leerlingen in kolom nul zijn er 3 waarbij wel verplaatsingen optreden, dat was in elke run van $3^e$ naar een lagere voorkeur, die $3^e$ en lagere voorkeuren hebben alle waarde = 0, zodoende altijd winst = 0.

Tot de 260 leerlingen in kolom pos behoren 67 leerlingen die in elke run een betere plaats kregen, de overige 193 komen zo aan de orde.

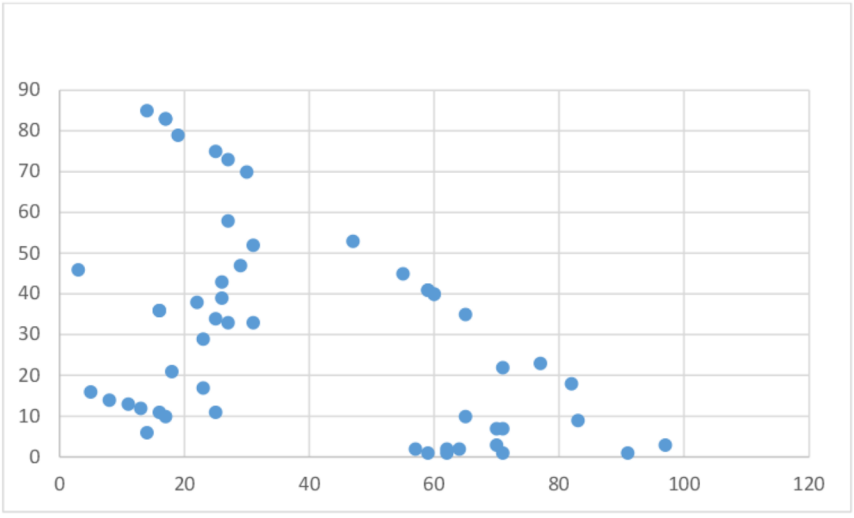
Onder de 31 leerlingen in kolom neg zijn er 13 die zowel verplaatsingen naar hogere als verplaatsingen naar lagere voorkeuren hebben. Er was bijvoorbeeld een leerling die in 6 runs van $1^e$ naar $4^e$ of lager werd verplaatst, eenmaal van $2^e$ naar $1^e$ en in 93 runs niet vanaf $1^e$ werd verplaatst.




Figuur 4-1: Simulatie A, 193 leerlingen, aantallen runs met hogere plaatsing.

De Figuur 4-1 heeft betrekking op de overige 193 leerlingen. Deze werden in geen enkele run lager geplaatst. De grafiek toont de aantallen runs met hogere plaatsing. De eerste waarneming is een leerling die in 99 runs niet verplaatst werd (bleef op 1$^{e}$ voorkeur) en in één run van 2$^{e}$ naar 1$^{e}$ voorkeur ging. Dat kan een toevalstreffer zijn.

Bij 53 leerlingen (in kolom mix van Tabel 3-4 ) hangt de gemiddelde winst wel af van de waarderingen voor de plaatsen. Deze leerlingen zijn in Figuur 4-2 weergegeven.




Figuur 4-2 : Simulatie A, 53 leerlingen, horizontaal aantal runs met hogere plaatsing, verticaal aantal runs met lagere plaatsing

De leerlingen links-onder in de grafiek hebben veel runs waarin ze niet werden verplaatst. De lijn van links-boven naar rechts-onder bevat de leerlingen die in (bijna) elke run werden verplaatst.

Het eerste punt op deze lijn betreft een leerling met in 14 runs een hogere plaatsing en in 85 runs een lagere. Haar winstverwachting is 13.00 – 0.98 $W_2$, dus positief als $W_2 < 13.27$.

Het laatste punt op deze lijn betreft een leerling met in 97 runs een hogere plaatsing en in 3 runs een lagere. Haar winstverwachting is $97.00 - 1.00\ W_2$, dus positief als $W_2 < 97.00$.

Tenslotte een punt in de linker-onderhoek. Dat betreft een leerling met in 13 runs een hogere plaatsing en in 12 runs een lagere. Haar winstverwachting is $13.00 - 0.25\ W_2$, dus positief als $W_2 < 52.00$.

De volgende Tabel toont, voor simulatie A, het verband tussen het aantal gunstige runs *minus* het aantal ongunstige runs en de afhankelijkheid van de winstverwachting van de waarderingen.

| gunstig minus ongunstig | nul | pos | mix | neg |
|---|---|---|---|---|
| -100 | 3 | | | 9 |
| -99..-90 | | | | 12 |
| -89..-80 | | | | |
| -79..-70 | | | 1 | 3 |
| -69..-60 | | | 3 | 2 |
| -59..-50 | | | 1 | |
| -49..-40 | | | 3 | |
| -39..-30 | | | 1 | |
| -29..-20 | | | 3 | |
| -19..-10 | | | 5 | |
| -9..-1 | | | 8 | 5 |
| 0 | 27 | | | |
| 1..9 | | 7 | 5 | |
| 10..19 | | 48 | 4 | |
| 20..29 | | 54 | 2 | |
| 30..39 | | 19 | 1 | |
| 40..49 | | 12 | 1 | |
| 50..59 | | 15 | 4 | |
| 60..69 | | 13 | 7 | |
| 70..79 | | 12 | 2 | |
| 80..89 | | 9 | | |
| 90..99 | | 4 | 2 | |
| 100 | | 67 | | |
| totaal | 30 | 260 | 53 | 31 |

Tabel 3-5: simulatie A, aantallen leerlingen

In kolom pos is het aantal gunstige runs groter dan het aantal ongunstige. Omgekeerd: als het aantal gunstige runs groter is dan het aantal ongunstige, dan staat de overgrote meerderheid van de gevallen in kolom pos.

De volgende Tabel toont, voor simulatie C, het verband tussen het aantal gunstige runs *minus* het aantal ongunstige runs en de afhankelijkheid van de winstverwachting van de waarderingen.

| gunstig minus ongunstig | nul | pos | mix | neg |
|---|---|---|---|---|
| -100 | 1 | | | 1 |
| -99..-90 | | 2 | 1 | 1 |
| -89..-80 | | | | |
| -79..-70 | | | | 7 |
| -69..-60 | | | 1 | 2 |
| -59..-50 | | | 1 | |
| -49..-40 | | | | 2 |
| -39..-30 | | | | |
| -29..-20 | | | | 3 |
| -19..-10 | | | | 4 |
| -9..-1 | 5 | 58 | 1 | 13 |
| 0 | 112 | 13 | | |
| 1..9 | | 87 | | |
| 10..19 | | 6 | | |
| 20..29 | | | | |
| 30..39 | | 4 | | |
| 40..49 | | 3 | | |
| 50..59 | | 2 | | |
| 60..69 | | 1 | | |
| 70..79 | | 2 | 1 | |
| 80..89 | | 2 | | |
| 90..99 | | | 1 | |
| 100 | | 7 | | |
| totaal | 118 | 187 | 6 | 33 |

Tabel 3-6 : simulatie C, aantallen leerlingen

Er zijn twee leerlingen met veel meer ongunstige dan gunstige runs in de kolom pos. Als een leerlingen in bijna alle runs verplaatst wordt van 3$^{e}$ naar 4$^{e}$ of lagere voorkeur, dan hebben deze verplaatsingen bijdrage 0 in de winstverwachting. Met enkele verplaatsingen van 3$^{e}$ naar hogere voorkeur is dan de winstverwachting niet negatief.

**Bijlage 4: Plaatsingen en populariteit eerste voorkeuren**

De eerste regel QPn in Tabel 4-1 bevat de percentages plaatsing van de betreffen de 117 leerlingen zoals in Tabel XX zijn genoemd. Elke leerling heeft een schooladvies, zoals havo/vwo. De volgende regel QPn betreft de 23 leerlingen met dat advies, en daarbij worden de percentages plaatsing genoemd.

| patroon populariteit | schooladvies | % plaatsing | | | afwijking | leerlingen |
|---|---|---|---|---|---|---|
| | | 1e | 2e | 3e | | |
| QPn | alles | 33 | 44 | 24 | | 117 |
| QPn | havo/vwo | 17 | 28 | 55 | 32 | 23 |

Tabel 4-1: Voorbeeld

Het getal afwijking geeft aan in hoeverre de percentages bij het advies afwijken van de algemene percentages (de regel alles).

De berekening voor havo/vwo verloopt als volgt.

De cumulatieve percentages bij alles zijn 33 77 100.

Die van havo zijn 17 45 100.

De afwijking is het grootste absolute verschil van deze twee over de plaatsen. In dit geval dus

77 - 45 = 32. Deze berekeningen zijn voor alle patronen van populariteit en alle schooladviezen uitgevoerd.

De volgende tabellen bevatten de afwijkingen per patroon van populariteit, met uitsplitsing naar schooladvies en is beperkt to de regels waar de afwijking ≥ 10 is.

| patroon populariteit | schooladvies | % plaatsing | | | afwijking | leerlingen |
|---|---|---|---|---|---|---|
| | | 1e | 2e | 3e | | |
| nnP | alles | 90 | 10 | | | 669 |
| nnP | havo/vwo | 70 | 30 | | 20 | 59 |
| nnP | havo | 75 | 25 | | 15 | 77 |
| nnP | vmbo-b, met lwoo | 100 | | | 10 | 20 |
| nnP | vmbo-b/k, met lwoo | 100 | | | 10 | 16 |
| nnP | vmbo-b/k | 100 | | | 10 | 17 |
| nnP | vmbo-b | 100 | | | 10 | 8 |
| nnP | vmbo-k, met lwoo | 100 | | | 10 | 8 |
| nnP | vmbo-k | 100 | | | 10 | 51 |
| nnP | vmbo-t, met lwoo | 100 | | | 10 | 3 |
| | | | | | | |
| nnQ | alles | 92 | 8 | | | 168 |
| nnQ | havo/vwo | 82 | 18 | | 10 | 37 |
| nnQ | vwo | 67 | 33 | | 25 | 3 |

Tabel 4-2: eerste voorkeur niet overtekend, per schooladvies afwijking percentages van algemene percentages

Er zijn 3914 leerlingen met een niet overtekende eerste voorkeur, daarvan hebben er 299 een schooladvies met afwijking ≥ 10, dat is 7.6%.

| patroon populariteit | schooladvies | % plaatsing | | | afwijking | leerlingen |
|---|---|---|---|---|---|---|
| | | 1e | 2e | 3e | | |
| Pnn | alles | 47 | 53 | | | 811 |
| Pnn | havo/vwo | 17 | 83 | | 30 | 57 |
| Pnn | havo | 22 | 78 | | 25 | 78 |
| Pnn | vmbo-b, met lwoo | 79 | 21 | | 32 | 52 |
| Pnn | vmbo-b/k, met lwoo | 81 | 19 | | 34 | 29 |
| Pnn | vmbo-k, met lwoo | 78 | 22 | | 31 | 1 |
| Pnn | vmbo-k | 59 | 41 | | 12 | 83 |
| Pnn | vmbo-t, met lwoo | 86 | 14 | | 39 | 1 |
| Pnn | vmbo-t | 62 | 38 | | 15 | 104 |
| Pnn | vwo | 37 | 63 | | 10 | 270 |
| | | | | | | |
| PnP | alles | 33 | 67 | | | 376 |
| PnP | havo/vwo | 14 | 86 | | 19 | 26 |
| PnP | havo | 14 | 86 | | 19 | 23 |
| PnP | vmbo-k | 47 | 53 | | 14 | 18 |
| PnP | vmbo-t, met lwoo | 86 | 14 | | 53 | 1 |
| PnP | vmbo-t/havo | 52 | 48 | | 19 | 18 |
| | | | | | | |
| PnQ | alles | 32 | 68 | | | 91 |
| PnQ | havo/vwo | 9 | 91 | | 23 | 21 |
| PnQ | vmbo-k | 42 | 58 | | 10 | 2 |
| PnQ | vmbo-t/havo | 53 | 47 | | 21 | 25 |
| PnQ | vmbo-t | 42 | 58 | | 10 | 23 |
| PnQ | vwo | | 100 | | 32 | 4 |
| | | | | | | |
| PPn | alles | 92 | 8 | 0 | | 479 |
| PPn | havo/vwo | 72 | 27 | 1 | 20 | 33 |
| PPn | havo | 74 | 22 | 4 | 18 | 36 |
| | | | | | | |
| PPQ | alles | 93 | 7 | | | 31 |
| PPQ | havo | 78 | 22 | | 15 | 5 |

Tabel 4-3: eerste voorkeur enkel overtekend, per schooladvies afwijking percentages van algemene percentages

Er zijn 2196 leerlingen met een enkel overtekende eerste voorkeur, daarvan hebben er 910 een schooladvies met afwijking ≥ 10, dat is 41.4%.

| patroon populariteit | schooladvies | % plaatsing | | | afwijking | leerlingen |
|---|---|---|---|---|---|---|
| | | 1e | 2e | 3e | | |
| Qnn | alles | 3 | 97 | | | 118 |
| Qnn | vmbo-b/k | 20 | 80 | | 17 | 2 |
| Qnn | vmbo-b | 17 | 83 | | 14 | 2 |
| Qnn | vmbo-k | 19 | 81 | | 16 | 12 |
| | | | | | | |
| QnP | alles | 2 | 98 | | | 73 |
| QnP | vmbo-k | 16 | 84 | | 14 | 6 |
| | | | | | | |
| QPn | alles | 33 | 44 | 24 | | 117 |
| QPn | havo/vwo | 17 | 28 | 55 | 32 | 23 |
| QPn | havo | 22 | 30 | 48 | 25 | 31 |
| QPn | vmbo-b | 100 | | | 67 | 1 |
| QPn | vmbo-k | 100 | | | 67 | 3 |
| QPn | vmbo-t/havo | 35 | 65 | | 23 | 29 |
| QPn | vmbo-t | 32 | 68 | | 23 | 17 |
| QPn | vwo | 62 | 38 | | 29 | 13 |
| | | | | | | |
| QPP | alles | 46 | 54 | | | 61 |
| QPP | vmbo-t/havo | 31 | 69 | | 15 | 11 |
| QPP | vwo | 66 | 34 | | 20 | 12 |

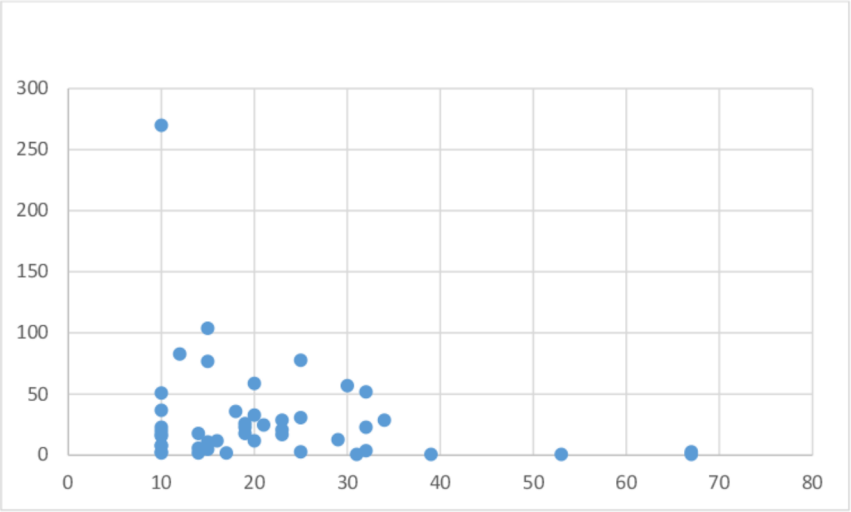
Tabel 4-4: eerste voorkeur dubbel overtekend, per schooladvies afwijking percentages van algemene percentages

Er zijn 512 leerlingen met een dubbel overtekende eerste voorkeur, daarvan hebben er 162 een schooladvies met afwijking ≥ 10, dat is 31.6%.

Van de 6622 leerlingen hebben er 1366 een schooladvies met afwijking ≥ 10, dat is 20.6%.

De maximale afwijking is bij QPn, vmbo-b en vmbo-k, dat betreft 4 leerlingen.

De volgende grafiek toont de aantallen leerlingen per grootte van de afwijking.




Grafiek 4-1: Aantallen leerlingen bij grootte van afwijking

Het hoogste punt is de 270 leerlingen met afwijking 10 bij Pnn, vwo.

De volgende Tabel geeft de aantallen leerlingen per interval van de grootte van afwijking.

| Interval afwijking | | leerlingen |
|---|---|---|
| 0 | 10 | 5711 |
| 11 | 20 | 527 |
| 21 | 30 | 274 |
| 31 | 40 | 110 |
| 41 | 50 | 0 |
| 51 | 60 | 1 |
| 61 | 70 | 4 |

Tabel 4-5: aantallen leerlingen per interval van de grootte van afwijking.

**Bijlage 5: Kleinschalig onderzoek unilateraal manipuleren**

De gegevens 2019 zijn, waar nodig, aangevuld tot drie voorkeuren.

De echte voorkeuren van een leerling, zoals ze in de aangevulde MEPS data staan, worden aangegeven als 1,2,3,4,… .

In één run is onderzocht wat de effecten van verschillende strategieën voor het unilateraal opgeven van valse voorkeuren zouden zijn geweest.

De run is telkens per leerling tweemaal toegepast, eenmaal met de echte voorkeuren 1, 2, 3 en eenmaal met valse voorkeuren waarbij bijvoorbeeld 3 door een andere school was vervangen. Er werd dus steeds met drie voorkeuren gewerkt.

Het effect van een strategie is een aantal verplaatsingen naar hogere *versus* een aantal verplaatsingen naar lagere voorkeuren. Voor leerlingen die hun 3e voorkeur willen ontlopen kan worden verondersteld dat ze 1e en 2e voorkeur ongeveer gelijk waarderen. Dan is het aantal verplaatsingen van 3e naar hoger *versus* het aantal van hoger naar 3e of lager een indicatie van de waarde van het effect. Voor leerlingen die hun 2e en 3e voorkeur willen ontlopen kan worden verondersteld dat ze hun 2e en 3e voorkeur ongeveer gelijk waarderen. Dan is het aantal verplaatsingen van lager naar 1e *versus* het aantal van 1e naar lager een indicatie van de waarde van het effect.

**Plaatsing op 3e voorkeur vermijden, permutaties**

Allereerst zijn de 78 leerlingen bekeken die met hun echte voorkeuren op hun 3e voorkeur werden geplaatst. Voor een steekproef van 31 van deze leerlingen is voor elke permutatie van de drie voorkeuren onderzocht wat het effect daarvan was. Er werden 18 leerlingen gevonden die met strategie 2, 1, 3 op hun valse eerste = echte tweede voorkeur werden geplaatst. Bij de overige leerlingen en de overige permutaties was er geen effect.

Naar aanleiding hiervan is voor een steekproef van 101 leerlingen met voorkeuren QPn of QQn, het effect van deze strategie bekeken. Tabel 5-1 bevat de resultaten.

| voorkeuren | beter | gelijk | slechter | totaal |
|---|---|---|---|---|
| QPn | 12 | 35 | 21 | 68 |
| QQn | 11 | 11 | 11 | 33 |
| totaal | 23 | 46 | 32 | 101 |

Tabel 5-1: 101 leerlingen, resultaten strategie 2,1,3

De 23 gevallen met succesvol manipuleren waren alle verplaatsingen van 3e naar 2e voorkeur. Van de 32 gevallen met slechter resultaat waren er 30 verplaatsing van 1e naar 2e voorkeur, de andere van 1e naar 3e.

Met unilateraal manipuleren valt deze strategie dus te overwegen door leerlingen die tot QPn of QQn behoren en 1e en 2e voorkeur als gelijkwaardig beschouwen en 3e voorkeur willen ontlopen.

**Plaatsing op 3e voorkeur vermijden, substitutie**

Voor een steekproef van 16 uit de 78 leerlingen (met echte voorkeuren op 3e geplaatst) zijn de strategieën 1, 2, A onderzocht waarbij A = achtereenvolgens de 4e en lagere voorkeuren van de leerling. Tabel 5-2 bevat de resultaten.

| school A | plaatsing met strategie | | totaal |
|---|---|---|---|
| | beter | slechter | |
| niet overtekend | 51 | 66 | 117 |
| wel overtekend | 35 | 2 | 37 |
| totaal | 86 | 68 | 154 |

Tabel 5-2: Aantallen plaatsingen 16 leerlingen

Alle verbeteringen zijn van 3e naar hoger, alle verslechteringen zijn van 3e naar de vervangende school A, dus 4e of lagere voorkeur.

Naar aanleiding hiervan is eerst voor de 180 leerlingen met voorkeuren QPn of QQn, het effect van de strategie 1, 2, X bekeken. Hierbij is X de meest populaire school in de voorkeuren 4 en lager, X moet hogere populariteit hebben dan de 3e voorkeur. Als X niet beschikbaar is in de voorkeuren van de leerling, dan wordt X geselecteerd uit de scholen die bij het schooladvies van de leerling passen. Tabel 5-3 bevat de resultaten.

| voorkeuren | beter | gelijk | slechter | totaal |
|---|---|---|---|---|
| QPn | 36 | 76 | 5 | 117 |
| QQn | 43 | 18 | 2 | 63 |
| totaal | 79 | 94 | 7 | 180 |

Tabel 5-3: 180 leerlingen, resultaten strategie 1,2,X, met X = meest populair

In de 79 gevallen met succesvol manipuleren waren er 46 verplaatsingen van 3e naar 1e , 23 met verplaatsing van 3e naar 2e voorkeur en 10 van 2e naar 1e voorkeur. De verplaatsingen naar lagere voorkeuren waren 5 van 1e naar 2e en 2 van 1e of 2e naar de vervangende school X (4e of lager).

Al met al dus 69 verplaatsingen van 3e naar hoger en 2 van hoger naar lager dan 3e .

Strategie 1,2,X geeft betere resultaten dan strategie 2,1,3.

Vervolgens is voor dezelfde 180 leerlingen ook de strategie 1,X,2 onderzocht. De resultaten staan in Tabel 5-4.

| voorkeuren | beter | gelijk | slechter | totaal |
|---|---|---|---|---|
| QPn | 55 | 17 | 45 | 117 |
| QQn | 41 | 18 | 4 | 63 |
| totaal | 96 | 35 | 49 | 180 |

Tabel 5-4: 180 leerlingen, resultaten strategie 1,X,2 met X = meest populair

Van de 96 verbeteringen zijn er 34 (bij QPn) verplaatsingen van 2e naar 1e voorkeur, de overige zijn van 3e naar 1e. Van de 49 verslechteringen zijn er 18/24/7 van 1e /2e /3e voorkeur naar de vervangende school X.

Hier zijn 62 verplaatsingen van 3e naar hoger en 42 van hoger naar 3e of lager.

Deze strategie 1,X,2 geeft geen betere resulaten dan 1, 2, X.

**Plaatsing op 2e of 3e voorkeur vermijden, permutaties**

Voor een steekproef van 61 leerlingen uit degenen die in de run op 2e of 3e voorkeur werden geplaatst zijn de effecten van alle permutaties van de eerste drie voorkeuren onderzocht.

In 34 gevallen gaf dit verplaatsing naar een hogere voorkeur, in 144 gevallen was er geen effect en in 127 gevallen was er een verplaatsing naar een lagere voorkeur. Tabel 5-5 geeft de resultaten per permutatie.

| permutatie | beter | gelijk | slechter | totaal |
|---|---|---|---|---|
| 132 | 27 | 2 | 32 | 61 |
| 213 | 1 | 60 | | 61 |
| 231 | | 54 | 7 | 61 |
| 312 | 6 | 3 | 52 | 61 |
| 321 | | 25 | 36 | 61 |
| totaal | 34 | 144 | 127 | 305 |

Tabel 5-5: 61 leerlingen, op 2e of 3e geplaatst, effecten alle permutaties eerste drie voorkeuren

Van de 34 verbeteringen waren er 33 verplaatsingen van 2e naar 1e en 1 van 3e naar 2e. De verslechteringen waren alle van 2e naar 3e.

Naar aanleiding hiervan is de strategie 132 onderzocht voor een steekproef van 200 leerlingen uit degenen die met hun populariteitspatroon met kans minstens 50% op 2e of 3e voorkeur werden geplaatst. Tabel 5-6 bevat de resultaten.

| patroon populariteit | beter | gelijk | slechter | totaal |
|---|---|---|---|---|
| Pnn | 9 | 29 | 49 (9) | 87 |
| PnP | 34 | 18 | 2 (0) | 54 |
| PnQ | 5 | 2 | | 7 |
| Qnn | 1 | 1 | 18 (2) | 20 |
| QnP | 3 | | 3 (0) | 6 |
| QnQ | | 1 | | 1 |
| QPn | | 4 | 5 (3) | 9 |
| QPP | 3 | 3 | 3 (1) | 9 |
| QPQ | 1 | | | 1 |
| QQn | | 3 | 3 (3) | 6 |
| totaal | 56 | 61 | 83 (18) | 200 |

Tabel 5-6: 200 leerlingen, strategie 132

De 56 verbeteringen zijn verplaatsingen van 2e naar 1e voorkeur. De 83 verslechteringen zijn 18 verplaatsingen van 1e naar 3e en 65 verplaatsingen van 2e naar 3e voorkeur.

In de kolom ‘slechter’ staan tussen ( ) de aantallen verplaatsingen van 1e naar lagere voorkeur.

Voor de leerlingen met PnP kan het unilateraal toepassen van de strategie interessant zijn. Dit geldt ook voor enkele andere groepen, onder de veronderstelling dat die leerlingen hun 2e en 3e voorkeur ongeveer gelijk waarderen.

**Plaatsing op 2e of 3e voorkeur vermijden, substituties**

Een andere aanpak om te proberen plaatsing op 2e of 3e voorkeur te vermijden is het toepassen van de strategie 1,X,Y of de strategie 1,X,2. Hierbij zijn X en Y overtekende scholen, bijvoorbeeld de meest populaire scholen uit de lagere voorkeuren. Als X of Y niet beschikbaar is in de lagere voorkeuren, dan wordt een andere, bij het schooladvies passende, overtekende school gebruikt.

Tabel 5-7 geeft de resultaten van unilateraal strategie 1XY met een steekproef van 200 leerlingen, weer uit degenen met kans ≥ 50% op plaatsing op 2e of 3e voorkeur. Bij het samenstellen van de steekproef werden 24 leerlingen verworpen omdat X en/of Y niet beschikbaar was.

| patroon populariteit | beter | gelijk | slechter | totaal |
|---|---|---|---|---|
| Pnn | 47 | 40 | 11 (2) | 98 |
| PnP | 37 | 17 | 7 (1) | 61 |
| PnQ | 1 | | 2 (0) | 3 |
| Qnn | 8 | 3 | 10 (0) | 21 |
| QnP | 2 | | 3 (0) | 5 |
| QPn | 5 | 1 | 1 (1) | 7 |
| QPP | 1 | | 4 (2) | 5 |
| totaal | 101 | 61 | 38 (6) | 200 |

Tabel 5-7: 200 leerlingen, strategie 1,X,Y

98 van de verbeteringen zijn verplaatsingen van 2e naar 1e voorkeur, 3 zijn van 3e naar 1e. Van de 38 verslechteringen zijn 6 verplaatsingen van 1e naar 4e of lager en 32 verplaatsingen van 2e naar 4e of lager. In de kolom 'slechter' staan tussen ( ) de aantallen verplaatsingen van 1e naar lagere voorkeur.

Onder de veronderstelling dat de leerlingen hun 2e en 3e voorkeur ongeveer gelijk waarderen is het aantrekkelijk om deze strategie, mits unilateraal, toe te passen.

Tabel 5-8 geeft de resultaten van unilateraal strategei 1,X,2 toepassen op dezelfde steekproef van 200 leerlingen.

| patroon populariteit | beter | gelijk | slechter | totaal |
|---|---|---|---|---|
| Pnn | 48 | 39 | | 87 |
| PnP | 36 | 18 | | 54 |
| PnQ | 5 | 2 | | 7 |
| Qnn | 10 | 7 | 3 (0) | 20 |
| QnP | 4 | | 2 (0) | 6 |
| QnQ | | 1 | | 1 |
| QPn | 3 | 1 | 5 (2) | 9 |
| QPP | 4 | 3 | 2 (1) | 9 |
| QPQ | 1 | | | 1 |
| QQn | 3 | 3 | | 6 |
| totaal | 114 | 74 | 12 (3) | 200 |

Tabel 5-8: 200 leerlingen, strategie 1,X,2

108 van de verbeteringen zijn verplaatsingen van 2e naar 1e voorkeur, 6 zijn van 3e naar 1e. Van de 12 verslechteringen zijn er 3/8/1 van 1e /2e /3e voorkeur naar de vervangende school X. In de kolom 'slechter' staan tussen ( ) de aantallen verplaatsingen van 1e naar lagere voorkeur.

Onder de veronderstelling dat de leerlingen hun 2e en 3e voorkeur ongeveer gelijk waarderen is het aantrekkelijk om deze strategie, mits unilateraal, toe te passen.

Strategie 1,X,2 geeft, bij unilaterale toepassing, meer verbeteringen en minder verslechteringen dan strategie 1,X,Y.

**Conclusies unilateraal manipuleren**

Leerlingen met voorkeuren QPn en QQn kunnen unilateraal strategie 1,2,X gebruiken als ze plaatsing op 3$^{e}$ voorkeur willen vermijden. In de testrun was dat succesvol voor 79 van de 180 leerlingen, voor 94 had het geen effect en voor 7 was er een nadelig effect.

Leerlingen die, met echte voorkeuren, kans ≥ 50% hebben om op 2$^{e}$ of 3$^{e}$ voorkeur geplaatst te worden kunnen strategie 1,X,2 toepassen om dat te vermijden. In de testrun was dat succesvol voor 114 van de 200 leerlingen, voor 74 had het geen effect en voor 12 was er een nadelig effect.

Als verplaatsingen van 2$^{e}$ naar lagere voorkeur niet als een echte verslechtering tellen, dan levert strategie 1,X,Y iets minder kwaliteit dan 1, X, 2.

**Bijlage 6: Toepassingen van sociaal optimaliseren**

Er volgen drie consultatieprojecten, uitgevoerd door de Stichting Mathematisch Centrum (sinds 2001 Centrum Wiskunde & Informatica).

In alle drie gaat het over het verdelen van schaarste. Het sociale aspect is dat de aandacht speciaal uitgaat naar degenen die in de verdeling het slechtst af zijn.

In 1973 was ik betrokken bij een consultatieproject met betrekking tot een probleem in de agrarische sector. In het kader van een ruilverkaveling is de vraag hoe de beschikbare kavels aan de betrokken bedrijven toe te wijzen. De te verdelen grond bestaat uit een aantal vakken. De verdeling over de bedrijven wordt beschreven door variablelen $x_{vb}$ = de fractie van vak v toegewezen aan bedrijf b. Op grond van deze variabelen kan de opbrengst voor bedrijf b worden berekend, in het model is dat de variabele $y_b$. De aanpak is nu om de minimale opbrengst te maximaliseren. Dat wordt gedaan met behulp van een te maximaliseren variabele z en de voorwaarden $z \leq y_b$ voor b = 1, 2, 3, … . Stel $z_m$ is de gevonden maximale waarde van z en $Y_b$ is de gevonden waarde van $y_b$ . Dan zal voor een of meer bedrijven gelden dat $z_m = Y_b$ , een of meer bedrijven verhinderen verdere verhoging van z. In de volgende stap van de berekeningen geldt nu voor deze bedrijven de voorwaarde $y_b = Y_b$ ( $= z_m$ ) en voor de overige bedrijven blijven de voorwaarden $z \leq y_b$ van toepassing en wederom wordt z gemaximaliseerd. Dit proces wordt herhaald tot voor elk bedrijf de opbrengst is gevonden.

Het eindresultaat is goed verdedigbaar. Het is onmogelijk om de opbrengst van bedrijf A te verhogen zonder die van een bedrijf B te verlagen, terwijl die van B toch al niet hoger is dan die van A.

Dit is een gestyleerde beschrijving van de aanpak die in [1] is gerapporteerd. Daarin is een complicatie die zou kunnen optreden niet besproken, en dit is niet de plaats om daar wel op in te gaan.

In 1982 is door B.J.Lageweg en mijzelf een consultatieproject uitgevoerd met betrekking tot het toewijzen van studenten aan keuzevakken tandheelkunde. Per keuzevak was een aantal plaatsen beschikbaar, elke student had een lijst ingeleverd van een tiental vakken, in volgorde van voorkeur. De vraag was om de methode **Boston** (het bekende in- en uitloten) door de computer te laten uitvoeren. Dat hebben we gedaan, elk student kreeg een plaats.
Het viel ons wel op dat met **Boston** het aantal plaatsingen in de tweede ronde (dus op tweede voorkeur) afhankelijk was van welke leerlingen in de eerste ronde op hun eerste voorkeur waren geplaatst. Dat leek ons ongewenst.

Dit leidde tot een optimaliseringsmodel. In het model zijn er variabelen die de plaatsing van de leerlingen beschrijven. Er zijn voorwaarden die rzorgen dat elke leerling wordt geplaatst en voorwaarden die er voor zorgen dat de plaatsingen binnen de capaciteiten van de keuzevakken vallen. En er zijn variabelen $y_p$ = het aantal op voorkeur p geplaatste leerlingen.
In de eerste stap van de berekeningen wordt $y_1$ gemaximaliseerd, $Y_1$ is het gevonden maximum.
In de tweede stap van de berekeningen wordt $y_2$ gemaximaliseerd, met als extra voorwaarde $y_1 = Y_1$ .
En zo gaat dit verder. In stap p wordt bepaald hoeveel studenten op voorkeur p worden geplaatst, onder de extra voorwaarden dat de in de voorgaande stappen berekende aantallen op de hogere voorkeuren 1, 2, … , p-1 worden gehandhaafd. In elke stap wordt een nieuwe verdeling van de studenten over de vakken gevonden, de verdeling van de laatste stap is de definitieve.
Deze aanpak is een voorbeeld van forward maximization (**Fmax**).

Met **Boston** en **Fmax** worden de aantallen studenten op hoge voorkeuren gemaximaliseerd. De ervaring leert dat er dan ook veel leerlingen op lage voorkeuren zijn. En dat is uit sociaal oogpunt ongewenst.

Dit leidde tot een tweede optimaliseringsmodel, **Bmin**, waarbij, werkend van laag naar hoog, de aantallen plaatsingen worden geminimaliseerd. In elke stap p onder de extra voorwaarden dat de in de voorgaande stappen berekende aantallen op de lagere voorkeuren p+1, p+2, … worden gehandhaafd. Dit is in sectie 4 beschreven.

In het jaarverslag 1982 van het MC werd hierover een rapport aangekondid, dat is nooit verschenen. Over beide methoden is door ons niets gepubliceerd, tot deze notitie. De aanpak **Bmin** wordt door Klaren (2016) besproken, **Fmax** door Gusfield & Irving (1989), zie ook Vande Vate (1989) .

Dezelfde aanpak bleek in 1984 toepasbaar op een probleem bij de PCGD.
Het betreft de overschakeling op nieuwe acceptgirokaarten in de periode medio 1985 tot eind 1986.

Bij de PCGD worden de nieuwe en de oude werkwijze naast elkaar gebruikt, de capaciteit oud daalt en de capaciteit nieuw groeit. Elke incassant (een organisatie die acceptgirokaarten aan leden/abonnees stuurt) moeten in één keer op de nieuwe kaart overstappen. Bij de incassanten zijn er seizoenseffecten, zoals kwartaalbetalingen. De momenten van overstap moeten dus zo worden gekozen dat er tot eind 1986 geen tekort aan verwerkingscapaciteit optreedt.

Elke incassant heeft een voorkeursmaand voor de overstap, de laatheid van een incassant is de wachttijd vanaf de voorkeursmaand.

De oplossing is gebaseerd op het minimaliseren van de maximale laatheid. Dat wordt herhaald toegepast. In de oplossing kan het aantal incassanten met een gegeven laatheid niet worden verkleind zonder het aantal dat een hogere laatheid heeft te vergroten.

Dit is in overeenstemming met de gedachten van Rawls [5]. Bijvoorbeeld:

> Once the difference principle is accepted, however , the minimum is to be set at that point which, … , maximizes the expectations of the least advantaged group.

Mijn belangstelling voor het school match probleem herleefde dank zij een column van Jeanine Daems in de Volkskrant (2013).