Docnr1

Centraal Planbureau

Sluit aan bij individuele situatie | *Big data biedt kansen*

CPB Policy Brief | 2018/07

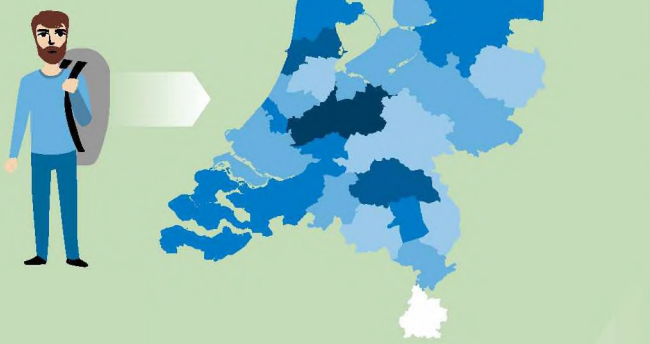
Regionale plaatsing vergunninghouders en kans op werk

Sander Gerritsen
Mark Kattenberg
Wouter Vermeulen

Slim plaatsen van vergunninghouders kan hun baankans vergroten

Baankans (%) 10 jaar na uitplaatsing, voor vergunninghouders die eind jaren '90 naar Nederland kwamen

- 55-57,5
- 52,5-55
- 50-52,5
- 47,5-50
- 45-47,5
- 42,5-45
- 40-42,5



De baankansen voor vergunninghouders zijn niet in elke regio hetzelfde: in de **meest gunstige** regio ligt de baankans bijna anderhalf keer zo hoog als in de **minst gunstige** regio.

Welke regio's gunstig zijn, hangt ook af van persoonlijke kenmerken, zoals **herkomstland, leeftijd en geslacht**. **Het maakt dus uit wie waar geplaatst wordt.**

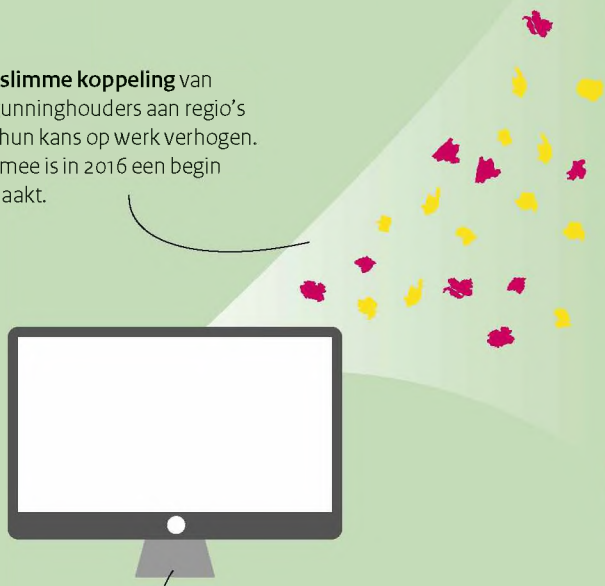
Deze regio's zijn ***relatief gunstig voor vergunninghouders uit voormalig Sovjet Unie en Joegoslavië****, en relatief ongunstig voor vergunninghouders uit overige landen.*

Deze regio's zijn relatief ongunstig voor vergunninghouders uit voormalig Sovjet Unie en Joegoslavië, en ***relatief gunstig voor vergunninghouders uit overige landen.***

Deze regio's zijn relatief ongunstig voor vrouwelijke vergunninghouders <30 jaar, en ***relatief gunstig voor mannelijke vergunninghouders <30 jaar.***

Deze regio's zijn ***relatief gunstig voor vrouwelijke vergunninghouders <30 jaar,*** *en relatief ongunstig voor mannelijke vergunninghouders <30 jaar.*

Een slimme koppeling van vergunninghouders aan regio's kan hun kans op werk verhogen. Hiermee is in 2016 een begin gemaakt.



Recent onderzoek voor de VS en Zwitserland (Bansak e.a., 2018) laat zien dat het gebruik van **big-datatechnieken** hierbij veel kan opleveren. Dit lijkt ook op te gaan voor Nederland.

Samenvatting

Voor de kans om werk te vinden maakt het uit in welke regio asielmigranten met verblijfsvergunning een woning aangeboden krijgen. Tot voor kort werd er bij de uitplaatsing van vergunninghouders echter geen rekening gehouden met hun aansluiting bij de regionale arbeidsmarkt. Aan de hand van arbeidsmarktprestaties en verhuisbewegingen van asielmigranten die eind jaren negentig naar Nederland kwamen, laten we zien dat het zinvol kan zijn om dit wel te doen.

De kans dat vergunninghouders tien jaar na uitplaatsing een baan hebben, blijkt in de meest gunstige regio bijna anderhalf keer zo hoog als in de minst gunstige. Dit vertaalt zich ook in aanzienlijke verschillen tussen regio's in het beroep op de bijstand. De regionale werkloosheid verklaart een deel van deze verschillen. Daarnaast hangt het regionale patroon samen met persoonlijke kenmerken als leeftijd, geslacht en herkomstland. Het maakt dus uit wie waar geplaatst wordt.

Voor de arbeidsmarktintegratie kan het daarom zinvol zijn om bij uitplaatsing rekening te houden met de aansluiting van kenmerken van vergunninghouders bij de regionale arbeidsmarkt. In 2016 is een start gemaakt met dit beleid, maar er zijn manieren om informatie over verschillen in de kans op werk nog meer te benutten. Internationaal onderzoek wijst bijvoorbeeld op de rol die een data-gedreven toewijzingsalgoritme kan spelen in het verbeteren van de koppeling van vergunninghouders aan regio's, en dat tegen beperkte kosten. Ook voor Nederland lijkt dit algoritme veelbelovend. Eerder beginnen met het screenen van migranten die op basis van herkomstland een grote kans hebben op het verkrijgen van een verblijfsvergunning, is een andere optie. Om de effectiviteit van dergelijk beleid te evalueren, is nader onderzoek nodig.

Het aanpassen van de huidige spreiding van aantallen vergunninghouders over het land op basis van inwonertal verhoogt hun kans op een baan waarschijnlijk aanzienlijk minder dan het slim koppelen van vergunninghouders aan regio's. Wel kan minder uitplaatsing van vergunninghouders naar dunbevolkte regio's het aantal verhuisbewegingen na uitplaatsing beperken. Hoewel de helft van alle vergunninghouders tien jaar later nog steeds in de regio woont waar ze uitgeplaatst zijn, trekken mensen vooral uit dunbevolkte gebieden weg. Aanpassing van de spreiding kan echter ook nadelen hebben, zoals meer segregatie en afbrokkeling van het draagvlak voor het opnemen van vergunninghouders.

1 Inleiding

Voor veel asielmigranten die in Nederland een verblijfsstatus krijgen, is de afstand tot de arbeidsmarkt groot. Een op de tien 18- tot 65-jarigen die in 2014 een verblijfsvergunning kregen (vergunninghouders), had tweeënhalf jaar later werk (CBS, 2018). Van de vluchtelingengolf die eind jaren negentig naar Nederland kwam, was na vijf jaar minder dan de helft aan het werk en de achterstand ten opzichte van migranten die om andere redenen naar Nederland kwamen, was na vijftien jaar nog steeds zichtbaar (Engbersen e.a., 2015).[1] Een belangrijke beleidsvraag is dan ook hoe de arbeidsmarktintegratie van vergunninghouders te bespoedigen. Dit staat ook centraal in het integratiebeleid van het huidige kabinet.[2]

Uitplaatsingsbeleid is hiervoor een van de mogelijke instrumenten. Dit beleid gaat over waar vergunninghouders terechtkomen, als zij vanuit een opvanglocatie naar een woning verhuizen. Tot voor kort werd hierbij echter nauwelijks gekeken naar gevolgen voor de arbeidsmarktintegratie. Volgens de huisvestingswet moeten vergunninghouders over het land worden verspreid op basis van inwonertal. Een gemeente met twee keer zoveel inwoners krijgt een twee maal zo grote taakstelling om vergunninghouders te huisvesten – onafhankelijk van regionale arbeidsmarktperspectieven. Hierbij werd in het verleden geen rekening gehouden met de match tussen individuele vergunninghouders en regionale arbeidsmarkten. Tegen de achtergrond van een verhoogde instroom van asielmigranten is in 2016 gestart met een vroege screening en matching op arbeid en opleiding om vergunninghouders gericht te plaatsen in arbeidsmarktregio's waar ze de meeste kansen hebben op integratie en participatie ('kansrijk koppelen').[3]

Uit de internationale literatuur blijkt dat arbeidsmarktomstandigheden in de regio waar vergunninghouders worden geplaatst, langdurig van invloed zijn op de kans op werk en op de hoogte van hun inkomen (Åslund en Rooth, 2007). Uitplaatsingsbeleid in Zweden, dat in de jaren tachtig meer vergunninghouders naar perifere gebieden stuurde, heeft bijvoorbeeld geleid tot slechtere arbeidsmarktprestaties. Bovendien trokken veel mensen na uitplaatsing uit deze gebieden weg (Edin e.a., 2004). Onderzoek voor de Verenigde Staten en Zwitserland laat zien dat matching de arbeidsmarktintegratie ook bij een gegeven taakstelling aanzienlijk kan verbeteren (Bansak e.a., 2018). De OESO beveelt dan ook aan om bij de spreiding van vergunninghouders rekening te houden met regionale arbeidsmarktperspectieven (OESO, 2016).

Deze policy brief brengt de rol van de uitplaatsingsregio voor de arbeidsmarktprestaties van vergunninghouders in Nederland in beeld. De volgende paragraaf laat zien dat de arbeidsmarktintegratie ook hier in sommige regio's beter verloopt dan in andere en dat het

[1] Maliepaard e.a. (2017) rapporteren dat 57% van deze groep na vijftien jaar een baan heeft. Voor autochtone Nederlanders in dezelfde leeftijdsgroep ligt de arbeidsmarktparticipatie rond de 80% en voor de overige niet-westerse bevolking is dit rond de 65%.
[2] Zie de brief 'Verdere Integratie op de Arbeidsmarkt: de economie heeft iedereen nodig!' van minister van Sociale Zaken en Werkgelegenheid Wouter Koolmees aan de voorzitter van de Tweede Kamer, 30 maart 2018 (link).
[3] Rijk en gemeenten hebben dit afgesproken in het Uitwerkingsakkoord Verhoogde Asielinstroom, 28 april 2016 (link).

uitmaakt wie waar geplaatst wordt. Paragraaf 3 laat zien dat de helft van alle vergunninghouders tien jaar later nog steeds in de regio woont waar ze zijn uitgeplaatst, en dat mensen vooral uit dunbevolkte regio's wegtrekken. De laatste paragraaf gaat in op de betekenis van onze bevindingen voor het huidige uitplaatsingsbeleid. Bij deze policy brief hoort een achtergronddocument met uitgebreidere analyses en verantwoording (Gerritsen e.a., 2018).

2 Arbeidsmarktprestaties naar regio

In deze paragraaf brengen we arbeidsmarktprestaties van vergunninghouders in verband met de regio waar ze na het verkrijgen van een verblijfsvergunning zijn geplaatst. Na uitplaatsing verhuizen sommige vergunninghouders naar andere regio's in Nederland of naar het buitenland. Hierop komen we in de volgende paragraaf terug. Op de plek waar vergunninghouders uiteindelijk terechtkomen, heeft het beleid echter geen directe invloed, terwijl de uitplaatsingsregio wel een directe beleidskeuze is. Daarom staat de uitplaatsingsregio in onze analyse centraal en niet de regio waar vergunninghouders in de jaren daarna gaan wonen.[4]

We richten ons hierbij op arbeidsmarkteffecten op de langere termijn, tien jaar na uitplaatsing. Vlak na uitplaatsing is toetreding tot de arbeidsmarkt voor veel vergunninghouders immers ver weg. Dit betekent dat we de vluchtelingenstroom die rond 2015 ons land binnenkwam, buiten beschouwing laten en, net als Engbersen e.a. (2015), kijken naar de vluchtelingen die tussen 1995 en 1999 naar Nederland kwamen.

In het bijzonder beschouwen we de groep vergunninghouders die tussen 1995 en 1999 werd ingeschreven in de gemeentelijke basisadministratie en die tussen 1995 en 2005 werd uitgeplaatst. Binnen deze groep kijken we naar vergunninghouders die op het moment van uitplaatsing tussen de vijftien en vijftig jaar oud zijn. Een vergunninghouder is in onze analyse uitgeplaatst als hij of zij niet meer op een locatie van het Centraal Orgaan opvang Asielzoekers (COA) woont en daar ook niet meer naar terugkeert. Het adres van deze eerste woning buiten een opvanglocatie bepaalt de uitplaatsingsregio. We gebruiken dezelfde indeling van Nederland in 35 arbeidsmarktregio's als het COA.[5]

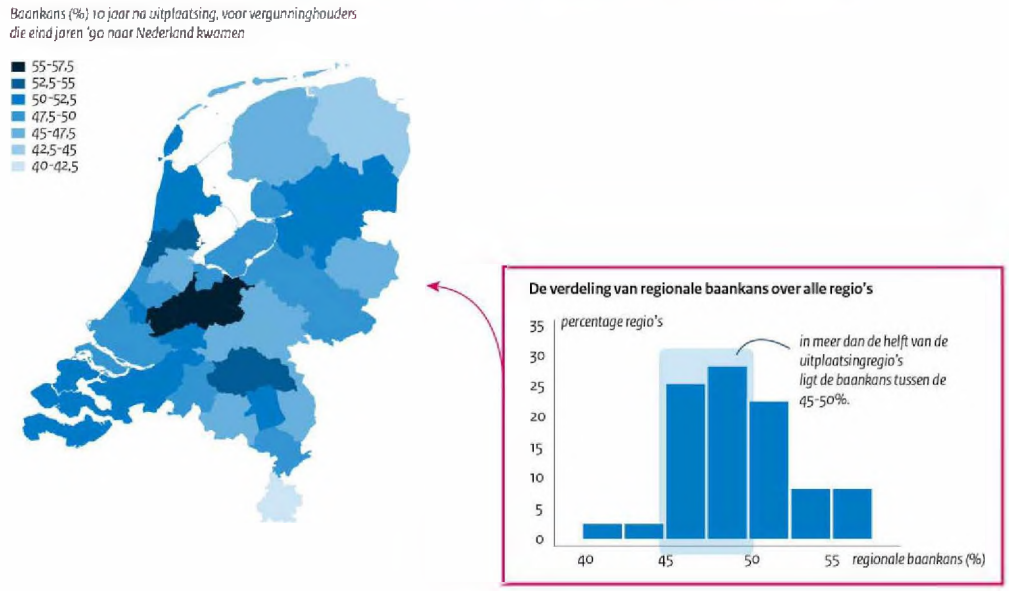
Arbeidsmarktprestaties van vergunninghouders meten we af aan het percentage dat tien jaar na uitplaatsing een baan van minstens 12 uur per week heeft (de 'baankans').[6] Daarnaast kijken we naar de kans dat vergunninghouders tien jaar na uitplaatsing een beroep doen op de bijstand. In de figuren die we in de rest van deze policy brief laten zien, zijn deze uitkomsten gecorrigeerd voor waargenomen kenmerken van vergunninghouders,

[4] Onze resultaten zijn echter vergelijkbaar als we niet kijken naar de uitplaatsingsregio, maar naar de regio waar vergunninghouders tien jaar later wonen.
[5] Deze indeling is in 2012 in overleg tussen gemeenten en het UWV tot stand gekomen (link).
[6] In Gerritsen e.a. (2018) kijken we ook nog naar een aantal andere uitkomstmaatstaven, zoals de hoogte van het inkomen van vergunninghouders die een baan hebben en de duur van uitplaatsing tot de eerste baan. De grens van 12 uur per week heeft te maken met de beschikbaarheid van historische gegevens.

zoals leeftijd, geslacht, gezinssamenstelling en land van herkomst, en voor het jaar van uitplaatsing en de conjunctuur.[7]

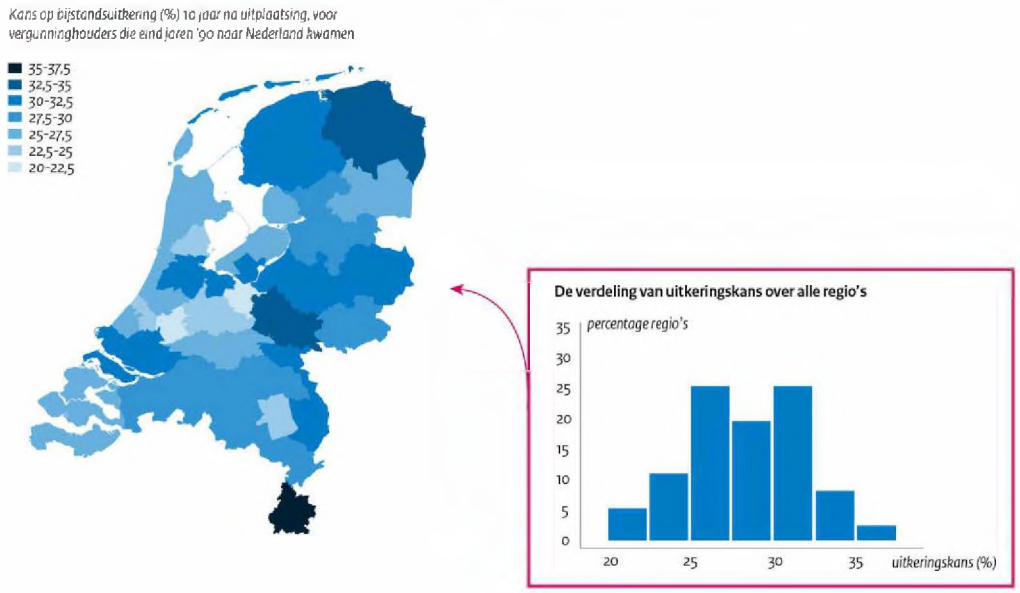
Het effect van uitplaatsingsregio op arbeidsmarktprestaties van vergunninghouders blijkt aanzienlijk. Figuur 1 zet de baankans naar uitplaatsingsregio in een kaartje. De kans dat vergunninghouders die in de regio's Amersfoort, Midden-Utrecht of Midden-Holland zijn geplaatst, tien jaar later een baan hebben (55%), is bijna anderhalf keer zo hoog als dezelfde kans voor vergunninghouders die in Zuid-Limburg terechtkwamen (41%). Belangrijk is echter niet zozeer waar de baankans hoog of laag is, want dat hoeft voor vluchtelingen die nu ons land binnenkomen niet hetzelfde te zijn als voor de vluchtelingengolf van eind jaren negentig, maar dat er aanzienlijke verschillen tussen regio's zijn. Daarom brengt een histogram naast het kaartje de spreiding van regionale baankansen op een andere manier in beeld. De hoogte van een balkje staat voor het percentage van de 35 arbeidsmarktregio's waarop de betreffende baankans van toepassing is. Zo ligt de baankans na tien jaar in ruim de helft van de uitplaatsingsregio's tussen de 45% en de 50%. Voor bijna de helft van de regio's is de baankans dus groter of kleiner.

Figuur 1 **De verdeling van regionale baankans per uitplaatsingsregio**



[7] Het opleidingsniveau van asielmigranten bij aankomst in Nederland is niet bekend. Bij de gerapporteerde regio-effecten kan een effect van selectie van vergunninghouders naar regio's op basis van niet-geobserveerde kenmerken spelen, maar gezien het uitplaatsingsproces waarin de match met de arbeidsmarktregio nauwelijks een rol speelde, verwachten we dat dit effect beperkt is.

Figuur 2 De verdeling van uitkeringskans per uitplaatsingsregio



Figuur 2 laat dezelfde plaatjes zien voor de kans dat vergunninghouders tien jaar na uitplaatsing een beroep doen op de bijstand. Deze kans varieert nog meer tussen uitplaatsingsregio's dan de kans om tien jaar later een baan te hebben. Er is wel een sterke samenhang: vergunninghouders die uitgeplaatst zijn in regio's met een hoge baankans, doen doorgaans minder beroep op de bijstand.

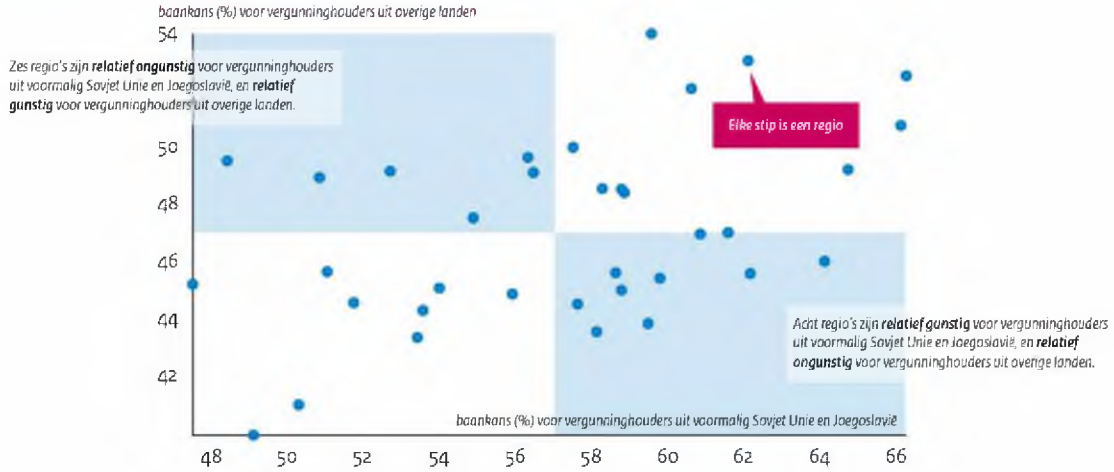
Lokale arbeidsmarktomstandigheden lijken een belangrijke determinant van het ruimtelijke patroon in figuur 1. Vergunninghouders die uitgeplaatst zijn in regio's met een hoge werkloosheid, hebben tien jaar later minder vaak werk. Dit sluit aan bij bevindingen uit de internationale literatuur (Åslund en Rooth, 2007). Toch verklaart de regionale werkloosheid niet meer dan een derde van de variatie in regionale baankans. Andere factoren spelen dus ook een belangrijke rol. Hierbij valt te denken aan aspecten zoals de aansluiting van de regionale arbeidsvraag bij de kennis en vaardigheden van vergunninghouders, maar verschillen in het lokale integratiebeleid kunnen ook van invloed zijn. We vinden geen aanwijzing voor een sterk effect van het aandeel minderheden, of de bevolkingsdichtheid in de uitplaatsingsregio.

Of een regio gunstig is, hangt ook af van wie er geplaatst wordt. Dit brengen we in beeld in figuur 3. Dit figuur toont de regionale baankans voor asielmigranten uit Europa en de voormalige Sovjet Unie, afgezet tegen de regionale baankans voor asielmigranten uit de rest van de wereld. Asielmigranten uit de eerste groep zijn doorgaans hoger opgeleid en voor hen is de culturele barrière voor (arbeidsmarkt-)integratie mogelijk ook kleiner.[8] Regio's in het kwadrant rechtsboven zijn voor beide groepen bovengemiddeld gunstig en regio's in het kwadrant linksonder zijn voor beide groepen minder gunstig dan gemiddeld. Een aanzienlijk deel van de regio's staat echter in een van de andere twee kwadranten, wat betekent dat ze voor de ene groep gunstig zijn en voor de andere groep juist niet. In Amsterdam, Rotterdam

[8] Zie de onderwijsindex van het United Nations Development Programme (link).

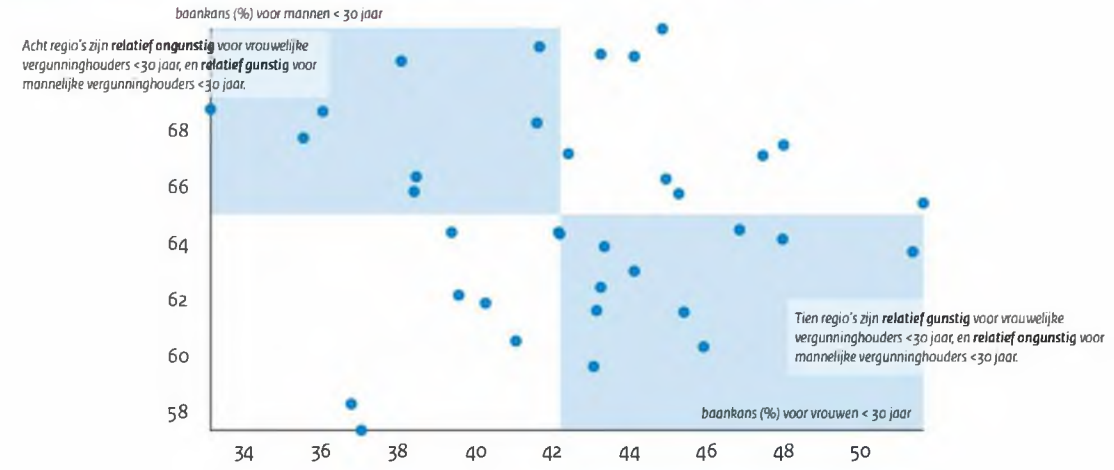
en Limburg hebben vergunninghouders uit de voormalige Sovjet Unie en Joegoslavië bijvoorbeeld relatief vaak een baan, terwijl vergunninghouders uit overige landen hier juist relatief moeilijk aan het werk komen.

Figuur 3 **Verschillen in regionale baankans naar herkomstland**



Dergelijke verschillen vinden we ook met betrekking tot andere persoonskenmerken.[9] Figuur 4 toont bijvoorbeeld een plaatje met dezelfde opzet als figuur 3, maar dan voor mannelijke dan wel vrouwelijke vergunninghouders die op het moment van uitplaatsing jonger dan dertig jaar zijn. Voor deze twee groepen verschilt het ruimtelijke patroon nog sterker dan voor herkomstland: of een regio gunstig is voor mannen onder de dertig zegt nagenoeg niets over of deze ook gunstig is voor vrouwen onder de dertig.

Figuur 4 **Verschillen in regionale baankans naar leeftijd en geslacht**



Er bestaan allerlei mogelijke verklaringen voor het resultaat dat regio's voor de ene groep wel gunstig zijn en voor de andere juist niet. De aansluiting bij de regionale arbeidsvraag kan bijvoorbeeld afhankelijk zijn van het opleidingsniveau, dat weer samenhangt met het

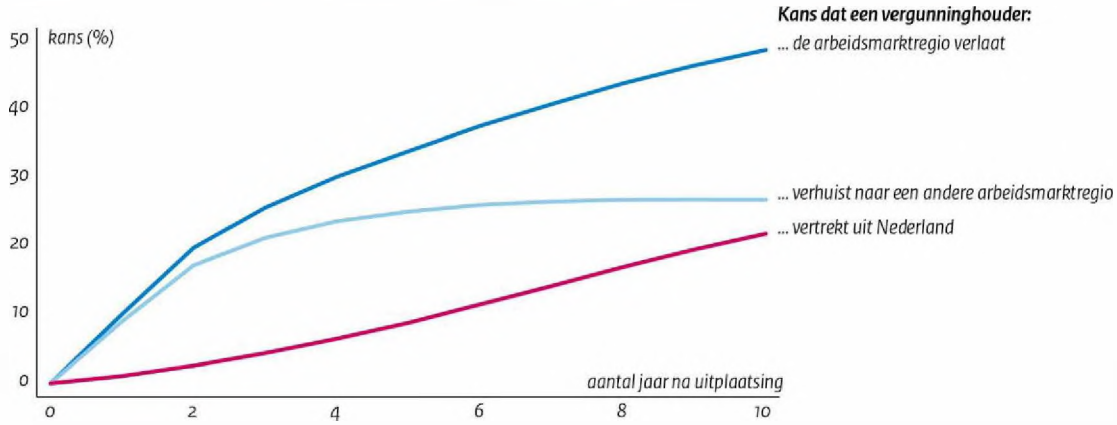
[9] Hierbij moet worden bedacht, dat verschillen ook ontstaan door ruis. Door naar deelgroepen te kijken wordt het aantal observaties, waarop de inschatting van de regionale baankans is gebaseerd, namelijk kleiner. De regionale patronen voor de deelgroepen die we hier bespreken, verschillen echter wel statistisch significant van elkaar.

herkomstland van de vergunninghouder. Bij fysiek zwaar werk kunnen leeftijd en geslacht een rol spelen. De nabijheid van migranten met een zelfde achtergrond kan zowel een positieve als een negatieve invloed hebben op de arbeidsmarktintegratie. Via lokale migrantennetwerken wordt bijvoorbeeld informatie over werk gedeeld (Beaman, 2012; Damm, 2014), maar veel omgang met mensen uit hetzelfde herkomstland vormt mogelijk ook een barrière voor de inburgering. Deze invloed lijkt ook nog eens te variëren met persoonskenmerken, zoals het opleidingsniveau (Edin e.a., 2003).[10] Welke rol deze en andere verklaringen precies spelen, kunnen we op basis van dit onderzoek niet zeggen. Echter, ook zonder precies te weten welke factoren de verschillen verklaren, kan het gegeven dat een regio voor de ene groep gunstiger is dan voor de andere, toch helpen om de koppeling van vergunninghouders aan regio's beter te maken. Hierop komen we terug in hoofdstuk 4.

3 Uit welke regio's trekken vergunninghouders weg?

De helft van alle vergunninghouders woont tien jaar later nog steeds in de regio waar ze zijn uitgeplaatst. De uitplaatsingsregio is dus behoorlijk bepalend voor waar mensen uiteindelijk terechtkomen. Toch vertrekt de andere helft naar een andere regio of naar het buitenland. Figuur 5 laat zien hoe deze cijfers zich vanaf het moment van uitplaatsing ontwikkelen. Twee jaar na uitplaatsing is bijna een op de vijf vergunninghouders al verhuisd naar een andere regio in Nederland. De groei in dit aandeel vlakt in de daaropvolgende jaren af. Het aandeel van vergunninghouders die in de eerste jaren na uitplaatsing al naar het buitenland vertrekken, is beperkt, maar dit aandeel neemt in de jaren daarop juist behoorlijk toe.[11]

Figuur 5 **Ontwikkeling vertrekkans naar andere regio of buitenland over de tijd**

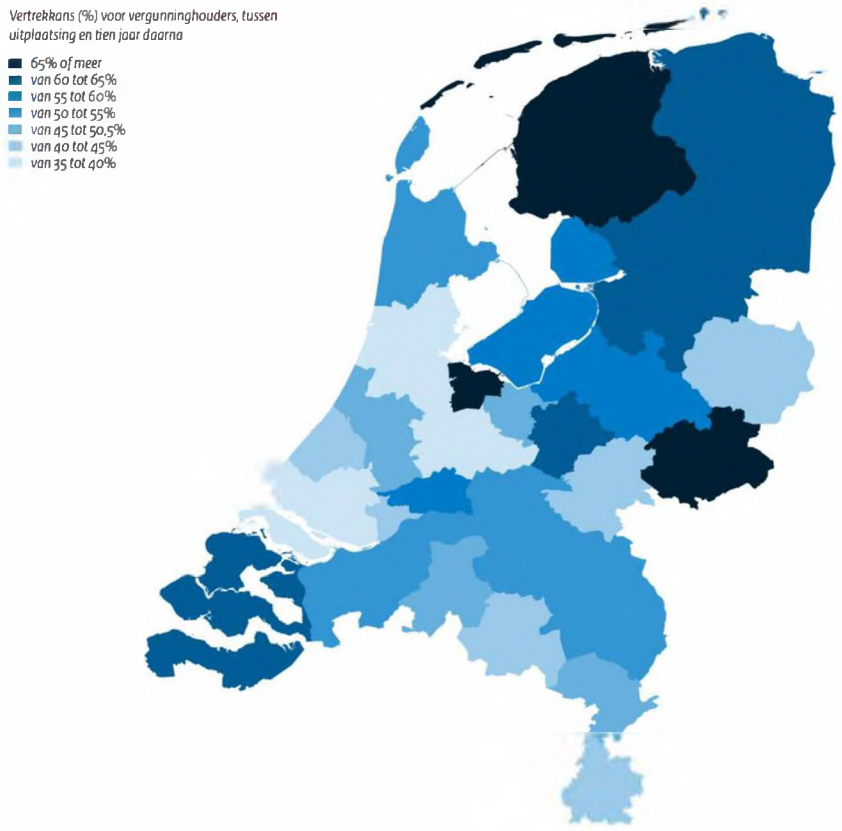


[10] Het blijkt dat vergunninghouders uit voormalig Joegoslavië relatief vaak een baan vinden in Rotterdam, Amsterdam en Limburg, waar ook relatief veel mensen uit dit herkomstland wonen. Dit suggereert in dit geval een positieve rol van lokale netwerken van mensen uit hetzelfde herkomstland.

[11] We gaan ervan uit dat een vergunninghouder naar het buitenland is vertrokken, als hij of zij niet in Nederland staat ingeschreven en ook niet is overleden. Mensen die na het verlopen van hun tijdelijke verblijfsvergunning illegaal in Nederland blijven, vallen dus ook onder deze groep.

Figuur 6 laat zien dat het aandeel van vergunninghouders die na uitplaatsing wegtrekken, wel substantieel tussen regio's verschilt. Deze vertrekkans varieert van 36% in de regio Rijnmond tot 71% in de regio Gooi en Vechtstreek. Ook uit de noordelijke regio's vertrekken relatief veel vergunninghouders. Vergunninghouders blijken vooral weg te trekken uit dunbevolkte regio's. We zien geen sterke samenhang van de totale vertrekkans met de regionale werkloosheid. Vergunninghouders in regio's met een hogere werkloosheid trekken wel vaker naar het buitenland, maar dit verklaart maar een klein deel van de regionale verschillen in emigratiekans.[12]

Figuur 6 **Vertrekkans naar andere regio of buitenland**

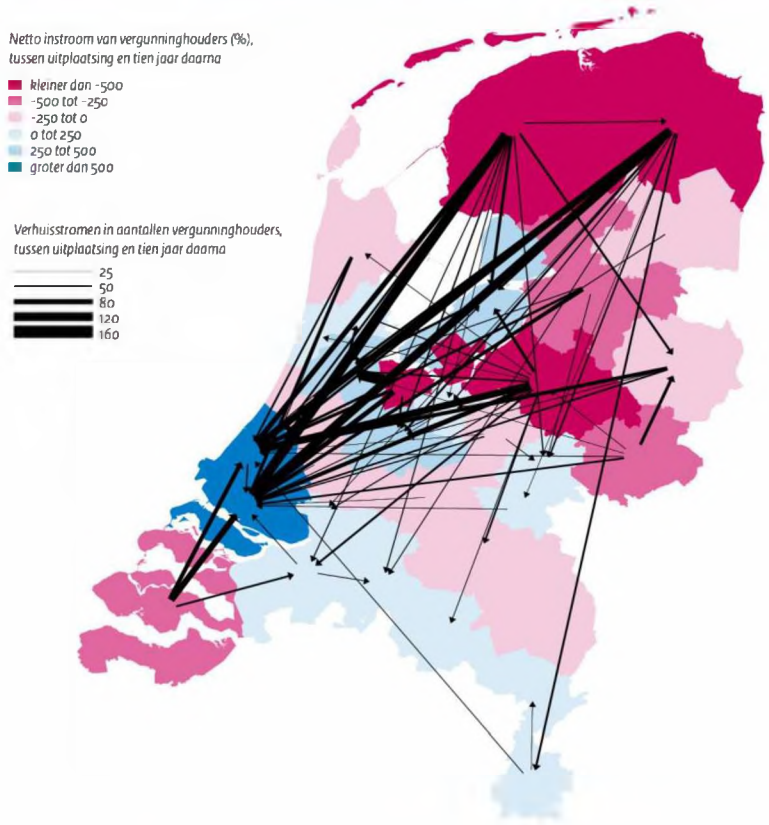


Waar verhuizen de vergunninghouders die in Nederland blijven naartoe? Figuur 7 laat voor elke regio het verschil tussen alle inkomende en vertrekkende vergunninghouders zien, gemeten vanaf de uitplaatsing tot tien jaar erna. De grootste verhuisstromen zijn er ook in weergegeven. Onder mensen die in Nederland blijven, zijn de grote steden – en dan vooral Rotterdam en Den Haag – populaire bestemmingen (zie ook Maliepaard e.a., 2015).[13] Vooral de grote verhuisstromen vanuit de noordelijke regio's in deze richting vallen op. Het belang van nabijheid van mensen uit hetzelfde herkomstland zou hierbij een rol kunnen spelen. Ook de internationale literatuur wijst hierop. Zo laat Åslund (2005) zien dat vergunninghouders

[12] Dit kan de regionale baankansen die we in de vorige paragraaf rapporteerden, enigszins vertekenen, als vergunninghouders die naar het buitenland verhuizen, qua arbeidsmarktperspectief sterk verschillen van de blijvers. De regionale baankans na twee jaar, als emigratie nog nauwelijks een rol speelt, hangt echter sterk samen met de regionale baankans na tien jaar, wat aangeeft dat dit effect in elk geval niet bepalend was.
[13] Dit patroon is nog niet zichtbaar voor het cohort asielmigranten dat in 2014 naar Nederland kwam (CBS, 2018).

in Zweden vaker verhuizen naar regio's met een hoger aandeel immigranten, en dan vooral naar regio's met veel mensen uit hetzelfde herkomstland. Damm (2009) laat zien dat dezelfde factoren voor vergunninghouders in Denemarken bepalend zijn voor of ze uit een regio wegtrekken.

Figuur 7 **Binnenlandse verhuisbewegingen na uitplaatsing**



4 Betekenis voor het uitplaatsingsbeleid

Beleid kan op allerlei manieren bijdragen aan een snellere arbeidsmarktintegratie van vergunninghouders. Eerdere studies benadrukken bijvoorbeeld het belang van een snelle start met taalles en begeleiding, een snelle asielprocedure, of het faciliteren van matching van vergunninghouders met werkgevers (Engbersen e.a., 2015; SER, 2016).[14] Bevindingen uit deze policy brief wijzen erop dat ook uitplaatsingsbeleid hierbij een rol kan spelen. Dit kan door rekening te houden met wie waar geplaatst wordt, of door taakstellingen aan te passen. In de volgende twee paragrafen gaan we hier verder op in, waarbij we getallenvoorbeelden gebruiken om een beeld te geven van de mogelijke omvang van effecten. Hieruit blijkt dat er meer te verwachten valt van het slim koppelen van vergunninghouders aan regio's, dan van het van het aanpassen van taakstellingen.

[14] Zie bijvoorbeeld OESO (2016) voor een internationaal overzicht. Voor Zwitserland tonen Hainmueller e.a. (2016) aan dat de wachttijd in asielzoekerscentra een fors negatief effect heeft op arbeidsmarktprestaties van vergunninghouders.

4.1 Kansrijk koppelen

Sinds 2016 worden vergunninghouders direct na toekenning van hun verblijfsvergunning gescreend en op basis hiervan maakt het COA een 'kansrijke koppeling' met een arbeidsmarktregio. Dit heeft alleen zin als het uitmaakt wie waar geplaatst wordt. Inderdaad laten wij zien dat een regio die gunstig is voor iemand uit voormalig Joegoslavië of de voormalige Sovjet Unie, dat niet hoeft te zijn voor iemand uit andere landen. Ook leeftijd en geslacht maken verschil. Andere kenmerken, die we niet waarnemen in onze dataset, maar die wel aan de orde komen in het screeningsgesprek, zoals het opleidingsniveau en de werkervaring van de vergunninghouder, doen er vermoedelijk ook toe. Er valt met matching dus iets te winnen, ook als de gemeentelijke taakstellingen op basis van inwonertal gehandhaafd blijven.

In de huidige opzet is de screeningsprocedure betrekkelijk licht, zodat vergunninghouders zo snel mogelijk doorgeplaatst kunnen worden naar een asielzoekerscentrum in de buurt van de regio waaraan ze gekoppeld zijn. Het gaat om een kort gesprek direct na toekenning van de verblijfsvergunning, waarin onder meer opleidingsachtergrond, werkervaring, sociaal netwerk en persoonlijke voorkeuren van de vergunninghouder aan de orde komen. Het COA vertaalt deze gegevens in een regioadvies, dat meegewogen wordt bij de uitplaatsingsbeslissing. Voor dit advies gebruikt het COA onder andere een inschatting van de regionale kans op werk, op basis van gegevens over regionale vacatures.[15]

Het is de vraag of de huidige aanpak altijd tot de koppeling met de hoogste baankans leidt. Sowieso spelen andere criteria, zoals nabijheid van familieleden en gemeentelijke taakstellingen bij de uiteindelijke uitplaatsing, ook een rol. Daarnaast zijn de verschillen in de regionale kans op werk voor veel sectoren en opleidingsachtergronden beperkt, dus er is lang niet altijd een eenduidige koppeling.[16] Bovendien blijft informatie over ervaringen van recente vergunninghouders in deze aanpak onbenut. Het regionale patroon van waar gemakkelijk werk te vinden is, kan voor deze groep echter anders zijn dan voor de rest van de beroepsbevolking.

Er zijn manieren om informatie over verschillen in de kans op werk bij de koppeling van vergunninghouders aan regio's meer te benutten. Internationaal onderzoek wijst op de rol die big-datatechnieken kunnen spelen in het verbeteren van de koppeling, tegen beperkte kosten (Bansak e.a., 2018). De essentie hiervan is om de regionale baankans van uit te plaatsen vergunninghouders zo goed mogelijk te voorspellen, op basis van hoe vergunninghouders met vergelijkbare kenmerken het hier eerder deden. De auteurs laten zien dat toepassing van hun algoritme ervoor kan zorgen dat uitgeplaatste vergunninghouders fors sneller een baan vinden. In de VS stijgt de baankans na negen

[15] Dit wordt gepubliceerd door de Samenwerkingsorganisatie Beroepsonderwijs Bedrijfsleven (link).
[16] In Gerritsen e.a. (2018) laten we zien dat vergunninghouders vooral werkzaam zijn in bedrijfstakken met een relatief zwak ruimtelijk concentratiepatroon, zoals afgemeten aan de Herfindahlindex. Vergunninghouders die een opleiding willen volgen, worden niet gekoppeld op basis van de kans op werk, maar op basis van waar voor hen geschikte opleidingen zijn. Ook dit criterium is niet altijd even onderscheidend, omdat de meeste opleidingen in meerdere regio's gevolgd kunnen worden.

maanden in hun berekening van 34% naar 48%. In Zwitserland stijgt de baankans na drie jaar van 15% naar 26%.

Een getallenvoorbeeld illustreert dat een koppeling die rekening houdt met de aansluiting van vergunninghouders bij de regionale arbeidsmarkt, ook voor Nederland veel kan opleveren. Hiertoe hebben we het algoritme van Bansak e.a. (2018) in vereenvoudigde vorm toegepast, zie het kader 'De inzet van big-datatechnieken bij de koppeling' voor meer toelichting. De gemiddelde kans dat een vergunninghouder tien jaar na uitplaatsing een baan heeft, stijgt dan fors, van 48,6 procent naar 59,3 procent. Hoewel de regionale spreiding in baankansen ook toeneemt, stijgt in bijna alle regio's de kans dat vergunninghouders een baan hebben (zie Gerritsen e.a., 2018).[17]

De inzet van big-datatechnieken bij de koppeling

In het invloedrijke tijdschrift Science beschrijven Bansak e.a. (2018) een data-gedreven toewijzingsalgoritme voor de verdeling van vergunninghouders over locaties. De eerste stap is het 'trainen' van een model op basis van arbeidsmarktprestaties van vergunninghouders die in het verleden zijn uitgeplaatst. Dit model verklaart de baankans van een vergunninghouder uit waargenomen kenmerken, zoals leeftijd, geslacht, herkomstland en opleidingsniveau (a). De wisselwerking tussen regio's en kenmerken van vergunninghouders wordt hierbij zo goed mogelijk benut. Dit levert voor elke vergunninghouder per regio een inschatting van de kans om aan het werk te gaan. In een tweede stap maakt het algoritme een vertaalslag van individuele vergunninghouders naar gezinnen. Hiermee wordt vervolgens de verdeling van gezinnen van vergunninghouders bepaald, die de gemiddelde baankans voor alle uit te plaatsen vergunninghouders zo hoog mogelijk maakt. Deze optimalisatieslag houdt rekening met beperkingen, zoals gemeentelijke taakstellingen op basis van inwonertal.

We hebben een vereenvoudigde versie van dit algoritme toegepast op onze gegevens over vergunninghouders die eind jaren negentig naar Nederland kwamen. Hiertoe hebben we regionale baankansen op basis van leeftijd, geslacht, herkomstland en jaar van uitplaatsing geschat. Vergunninghouders zijn vervolgens zo aan regio's toegewezen dat de totale baankans zo hoog mogelijk werd, terwijl het totaal aantal vergunninghouders per regio gelijk is gehouden. Hierbij is de optimale verdeling in één keer voor de hele groep bepaald, dus zonder rekening te houden met het feit dat asielzoekers een voor een een verblijfsvergunning krijgen. Daarnaast veronderstellen we dat baankansen voor individuele vergunninghouders niet afhankelijk zijn van welke andere vergunninghouders in hun regio worden geplaatst.

Het algoritme kan ook worden toegepast op asielmigranten die nu een verblijfsvergunning aanvragen. Het CBS verzamelt de gegevens over arbeidsmarktprestaties van onlangs naar Nederland gekomen vergunninghouders, die nodig zijn om het model hiervoor te trainen (CBS, 2018). Dit levert een hulpmiddel voor medewerkers van het COA, die de uiteindelijke uitplaatsingsbeslissing maken. Invoering hiervan hoeft niet veel te kosten. Bovendien vereist het geen drastische hervorming van het uitplaatsingsproces.

(a) Het opleidingsniveau is wel beschikbaar voor de VS, maar niet voor Zwitserland. Ondanks het ontbreken van deze belangrijke variabele laat het model toch indrukwekkende resultaten zien. In onze data beschikken we niet over opleidingsniveau, maar tegenwoordig wordt dit bij aankomst wel uitgevraagd.

[17] In ons rekenvoorbeeld daalt de regionale baankans voor Groningen, Groot Amsterdam, Noord-Holland Noord en Zuid-Kennemerland. Voor de meeste perifere gebieden of regio's met een hoge werkloosheid pakt de toewijzing dus gunstig uit.

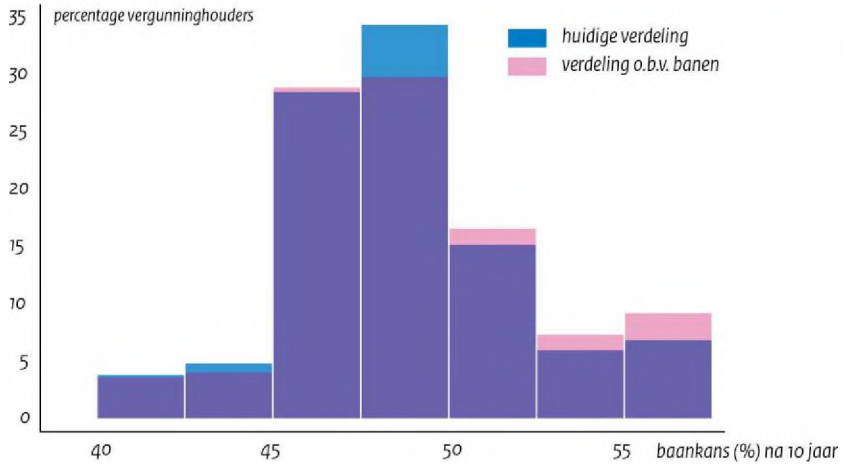
Een andere manier om meer te doen met verschillen in de kans op werk, is door op de opvanglocatie al eerder te beginnen met het screenen van migranten die op basis van herkomstland een grote kans hebben op het verkrijgen van een verblijfsvergunning.[18] Voor deze groep is er dan meer tijd om bijvoorbeeld informatie te verzamelen en te verifiëren, of om andere partijen bij het koppelproces te betrekken, zoals gemeenten of het Uitvoeringsinstituut Werknemersverzekeringen (UWV). Dit kan ook resulteren in een betere koppeling.

Over de effectiviteit van kansrijk koppelen, of een mogelijke intensivering van dit beleid, kunnen we op basis van ons onderzoek geen uitspraken doen. Een gerandomiseerd onderzoek, zoals dat ook in de medische wereld gebruikelijk is, kan hier wel uitsluitsel over geven. Dit vereist dat vergunninghouders op basis van toeval wel of niet met een bepaalde koppelmethode worden uitgeplaatst en dat hun arbeidsmarktprestaties daarna gevolgd worden. In Zwitserland start een pilot met het algoritme van Bansak e.a. (2018) volgens deze opzet.

4.2 Aanpassing van taakstellingen

Ondanks aanzienlijke verschillen in de regionale baankans, lijkt er met een realistische aanpassing van taakstellingen weinig winst te behalen. In een getallenvoorbeeld waarbij taakstellingen gebaseerd zijn op het aantal banen in een regio, en dus niet op het aantal inwoners, stijgt de gemiddelde baankans van 48,6 naar 48,9 procent.

Figuur 8 Spreiding van vergunninghouders op basis van waar werk is



[18] CBS (2018) laat zien dat de kans om een verblijfsvergunning te krijgen sterk samenhangt met het herkomstland. In het bijzonder heeft 94% van alle Syriërs en Eritreeërs die in 2014 in de asielopvang van het COA zijn ingestroomd, na tweeënhalf jaar een verblijfsvergunning.

Figuur 8 laat zien hoe dit komt. Het percentage vergunninghouders dat terechtkomt in regio's met een hoge baankans, neemt toe, maar dit gaat vooral ten koste van het percentage vergunninghouders dat in regio's met een gemiddelde baankans wordt geplaatst. Bovendien is deze verschuiving beperkt. Ook een grotere, en hierdoor minder realistische, verschuiving van taakstellingen levert weinig op.[19] Daarbij moet nog eens worden bedacht dat, naarmate vergunninghouders meer geconcentreerd worden in een klein aantal gunstige regio's, de baankans hier door de toename van het aanbod op de regionale arbeidsmarkt waarschijnlijk daalt.[20]

Het verhuispatroon van vergunninghouders na uitplaatsing kan ook reden zijn om de taakstellingen tegen het licht te houden. Twee jaar na uitplaatsing is bijna een op de vijf vergunninghouders al verhuisd naar een andere regio in Nederland. Mogelijk leidt een betere koppeling van vergunninghouders aan regio's ertoe dat minder mensen na uitplaatsing wegtrekken. We zien echter geen sterke samenhang met de regionale werkloosheid, dus meer uitplaatsing in regio's met een lage werkloosheid hoeft niet per se tot minder verhuisbewegingen te leiden.

Vergunninghouders blijken wel relatief vaak weg te trekken uit dunbevolkte regio's. Door meer vergunninghouders te plaatsen in regio's waar ze anders toch naartoe zouden trekken, zouden onnodige verhuisbewegingen – en de kosten die hiermee gepaard gaan, zoals het opnieuw opbouwen van een lokaal netwerk – vermeden kunnen worden. Meer mensen plaatsen in dunbevolkte regio's, waar de beschikbaarheid van woningen doorgaans groter is, zou juist tot meer verhuisbewegingen leiden. Bovendien ondermijnt dit verhuispatroon het argument, dat het plaatsen van vergunninghouders in dunbevolkte gebieden helpt om het draagvlak voor lokale voorzieningen in stand te houden.

Mogelijke voordelen van het aanpassen van gemeentelijke taakstellingen moeten worden afgewogen tegen een aantal potentiële nadelen. Zo kan het loslaten van de verdeling op basis van inwonertal segregatie in de hand werken, als hierdoor meer vergunninghouders in gebieden terechtkomen waar het aandeel minderheden al relatief groot is.[21] Daarnaast ervaren veel mensen deze verdeling als eerlijk (Bansak e.a., 2016). Het loslaten hiervan kan daarom ten koste gaan van het lokale draagvlak voor het opnemen van vergunninghouders, wat de integratie juist belemmert.

[19] Als we alle vergunninghouders bijvoorbeeld uitplaatsen in de helft van de regio's waar de baankans het hoogst is, dan stijgt de gemiddelde baankans naar 51,8 procent. In deze rekensom zijn vergunninghouders binnen de zeventien regio's met de hoogste baankans verdeeld op basis van bevolkingsomvang. Zie Gerritsen e.a. (2018).
[20] In combinatie met kansrijk koppelen kan het aanpassen van taakstellingen wel veel opleveren, zie Bansak e.a. (2018).
[21] Voor de arbeidsmarktintegratie van vergunninghouders hoeft segregatie niet per se nadelig te zijn (Edin e.a., 2003).

Referenties

Åslund, O, 2005, Now and forever? Initial and subsequent location choices of immigrants, *Regional Science and Urban Economics*, Vol. 35:141–165.

Åslund, O. en D.O. Rooth, 2007, Do When and Where Matter? Initial Labor Market Conditions and Immigrant Earnings, *Economic Journal*, Vol. 117:422–448.

Bansak, K., J. Hainmueller, en D. Hangartner, 2016, Aristotelian Equality and International Cooperation: Europeans Prefer a Proportional Asylum Regime, Working Paper.

Bansak, K., J. Ferwerda, J. Hainmueller, A. Dillon, D. Hangartner, D. Lawrence,, & J. Weinstein, 2018, Improving refugee integration through data-driven algorithmic assignment, *Science*, Vol. 359: 325–329.

Beaman, L. A., 2012, Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US, *The Review of Economic Studies*, Vol. 79: 128–161.

CBS, 2018, Uit de startblokken, cohortstudie naar recente asielmigratie, Den Haag: Centraal Bureau voor de Statistiek.

Damm, A. P., 2009, Determinants of recent immigrants' location choices: quasi-experimental evidence, *Journal of Population Economics*, Vol. 22: 145–174.

Damm, A. P., 2014, Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants, *Journal of Urban Economics*, Vol. 79: 139–166.

Edin, P.-A., P. Fredriksson en O. Åslund, 2003, Ethnic enclaves and the economic success of immigrants: evidence from a natural experiment, *Quarterly Journal of Economics*, Vol. 118: 329–57.

Edin, P.-A., P. Fredriksson en O. Åslund, 2004, Settlement policies and the economic success of immigrants, *Journal of Population Economics*, Vol. 17: 133–55.

Engbersen, G., J. Dagevos, R. Jennissen, L. Bakker, A. Leerkes, J. Klaver en A. Odé, 2015, Geen tijd verliezen: Van opvang naar integratie van asielmigranten, Den Haag, Wetenschappelijke Raad voor het Regeringsbeleid.

Gerritsen, S.B., M.A.C. Kattenbergen W. Vermeulen, 2018, Regionale plaatsing vergunninghouders en kans op werk, CPB Achtergronddocument, Den Haag: Centraal Planbureau.

Hainmueller, J., D. Hangartner en D. Lawrence, 2016, When lives are put on hold: Lengthy asylum processes decrease employment among refugees, *Science advances*, Vol. 2: e1600432.

Maliepaard, M., B. Witkamp en R. Jennissen, 2017, Een kwestie van tijd? De integratie van asielmigranten: een cohortonderzoek, Den Haag: WODC Cahier 2017-3.

OESO, 2016, Making Integration Work: Refugees and Others in Need of Protection, Parijs: OECD Publishing.

SER, 2016, Nieuwe wegen naar een meer succesvolle arbeidsmarktintegratie van vluchtelingen, Den Haag: Sociaal-Economische Raad.

Dit is een uitgave van:

Centraal Planbureau
Postbus 80510 | 2508 GM Den Haag
T (088) 984 60 00

Mei 2018

Stanford ETH zürich

ipl immigration policy lab

Working Paper Series

Working Paper No. 20-06

Leveraging the Power of Place: A Data-Driven Decision Helper to Improve the Location Decisions of Economic Immigrants

Jeremy Ferwerda, Nicholas Adams-Cohen, Kirk Bansak, Jennifer Fei, Duncan Lawrence, Jeremy M. Weinstein, and Jens Hainmueller

Leveraging the Power of Place: A Data-Driven Decision Helper to Improve the Location Decisions of Economic Immigrants

Jeremy Ferwerda[1,2,*], Nicholas Adams-Cohen[1,*], Kirk Bansak[1,3,*], Jennifer Fei[1], Duncan Lawrence[1], Jeremy Weinstein[1,4], and Jens Hainmueller [1,4,5,†]

[1]Immigration Policy Lab, Stanford University
[2]Department of Government, Dartmouth College
[3]Department of Political Science, University of California San Diego
[4]Department of Political Science, Stanford University
[5]Graduate School of Business, Stanford University
[*]Equal contributor
[†]Project director and corresponding author. Contact: jhain@stanford.edu.

July 2020

A growing number of countries have established programs to attract immigrants who can contribute to their economy. Research suggests that an immigrant's initial arrival location plays a key role in shaping their economic success. Yet immigrants currently lack access to personalized information that would help them identify optimal destinations. Instead, they often rely on availability heuristics, which can lead to the selection of sub-optimal landing locations, lower earnings, elevated outmigration rates, and concentration in the most well-known locations. To address this issue and counteract the effects of cognitive biases and limited information, we propose a data-driven decision helper that draws on behavioral insights, administrative data, and machine learning methods to inform immigrants' location decisions. The decision helper provides personalized location recommendations that reflect immigrants' preferences as well as data-driven predictions of the locations where they maximize their expected earnings given their profile. We illustrate the potential impact of our approach using backtests conducted with administrative data that links landing data of recent economic immigrants from Canada's Express Entry system with their earnings retrieved from tax records. Simulations across various scenarios suggest that providing location recommendations to incoming economic immigrants can increase their initial earnings and lead to a mild shift away from the most populous landing destinations. Our approach can be implemented within existing institutional structures at minimal cost, and offers governments an opportunity to harness their administrative data to improve outcomes for economic immigrants.

1 Introduction

Immigration has long been recognized as a driver of economic growth (Peri, Shih and Sparber 2015; Kerr et al. 2016). Immigrants increase the size and diversity of the workforce, fill skill shortages, start businesses, and contribute to innovation (Hunt and Gauthier-Loiselle 2010; Borjas 1995; Burchardi et al. 2020). To encourage these positive effects, many countries have complemented family-based and humanitarian admission streams with economic immigration programs, which prioritize the admission of skilled professionals. A prominent example is Canada's Express Entry system, through which approximately 100,000 immigrants are admitted each year. Applicants earn points for qualifications, such as language ability, educational degrees, and occupational experience, as well as other factors that have been shown to be associated with long-term economic success in Canada. Applicants who are above a certain threshold for that particular application round receive invitations to apply for permanent residence (Desiderio and Hooper 2016). Several other countries, such as Australia, New Zealand, and the United Kingdom, have implemented similar policies (Kerr et al. 2017).

The goal of these programs is to admit immigrants who are likely to succeed economically and contribute to the destination country. Yet despite these programs' intentions, immigrants nevertheless face a number of barriers to economic success. For instance, while a subset of individuals will have a preexisting job offer, the majority must select an initial location within the destination country in which to settle and begin their job search. However, immigrants generally lack access to personalized information on the locations that are aligned with their preferences and skill sets. As a result, the initial location decision can be affected by a variety of decision-making biases, such as availability heuristics. When economic immigrants choose suboptimal landing locations, economic admission programs cannot realize their full potential. Indeed, previous studies have demonstrated that the initial location of immigrants has a sizable impact on their short- as well as long-term economic outcomes (Åslund and Rooth 2007; Damm 2014; Bansak et al. 2018).

In this study, we propose a data-driven decision helper that leverages administrative data and machine learning methods to improve the initial location decisions of economic immigrants. Drawing on behavioral insights that have been used to improve decisions in other policy domains (Thaler and Sunstein 2009; OECD 2017), our approach seeks to enhance the choice architecture that shapes immigrants' decisions by providing them with systematic personalized information, delivered in the form of informational nudges, about which locations in the destination country would likely be beneficial to them. Building upon the outcome-based matching algorithm developed in Bansak et al. (2018), the decision helper provides newly invited immigrants with location recommendations that align with their preferences and maximize their expected economic outcomes. These recommendations draw on machine learning models applied to administrative data, which predict how immigrants with similar profiles have fared across possible landing locations, in combination with elicited preferences. The recommendations from the decision helper are not meant to be binding, but provide additional information that

assists newly invited economic immigrants to make more informed location decisions.

To illustrate the potential of our approach, we evaluate administrative tax and landing data on recent cohorts from the economic immigration programs within Canada's Express Entry system. We find that landing locations are highly concentrated, and many economic immigrants settle in destinations that are sub-optimal from the perspective of predicted earnings. Using backtests and simulations, we find that providing data-driven location recommendations could significantly increase the annual income of economic immigrants and more widely disperse the benefits of economic migration across Canada. These gains would be realized at limited marginal cost, since the Canadian government already collects the administrative data used to train the models and communicates regularly with economic immigrants throughout the application process. Although the decision helper should be tested prospectively via a randomized controlled trial to evaluate its full impact on a variety of outcomes, our results suggest that nudging incoming economic immigrants with personalized information could improve their outcomes and create an opportunity for governments to leverage their existing data to offer an innovative resource at scale.

2 A Data-Driven Decision Helper

2.1 Motivation

Our approach is motivated by a growing body of research that demonstrates the importance of the initial landing location in shaping immigrants' outcomes. For example, studies have used quasi-experimental designs to demonstrate that an immigrant's initial landing location has an impact on short- as well as long-term economic success (Åslund and Rooth 2007; Damm 2014; Bansak et al. 2018). Similarly, when examining outcomes among non-immigrants, experiments have shown that families who were randomly offered housing vouchers to move to lower-poverty neighborhoods had improved long-term outcomes in terms of earnings and educational attainment for their children (Chetty, Hendren and Katz 2016; Ludwig et al. 2013).

The initial destination choice is also consequential given that many immigrants tend to remain in their landing location (Kaida, Hou and Stick 2020; Mossad et al. 2020). For example, in Canada, more than 80% of recent economic class immigrants remained in their arrival cities ten years later (Kaida, Hou and Stick 2020). In addition, landing locations are often highly concentrated. For example, in Australia, more than half of all recent immigrants settled in Greater Sydney and Greater Melbourne, and only 14% settled outside the major capital cities (Tuli 2019). If initial settlement patterns concentrate immigrants in a few prominent landing regions, many areas of the country may not experience the economic growth associated with immigration. Moreover, undue concentration may impose costs in the form of congestion in local services, housing, and labor markets. To address the uneven distribution of immigrants, governments including Canada and Australia have implemented policy reforms to regionalize immigration

and encourage settlement outside of well-known major cities (Taylor et al. 2014; Fotros 2018; Hugo 2008; Brezzi et al. 2010).

Although the evidence suggests that the initial landing location shapes immigrants' outcomes, choosing an optimal destination from the large set of potential options is a formidable task. Research suggests that immigrants consider the location of family or friends, the perceived availability of employment opportunities, or preferences regarding climate, city size, and cultural diversity (Chiswick and Miller 2004; Hyndman, Schuurman and Fiedler 2006; Akbari and Harrington 2007; Massey 2008; Brezzi et al. 2010; Tonkin and Tonkin 1993; Damm 2009; Mossad et al. 2020). While some of these considerations may lead immigrants to correctly identify an optimal location, research also suggests that location decisions are impaired by common cognitive biases. One such bias, which has been well documented as a powerful influence across many choice settings with incomplete information (Tversky and Kahneman 1974; Thaler and Sunstein 2009), is an availability heuristic. This heuristic suggests that immigrants prioritize places they have heard about, and that they overlook less prominent locations even though these locations may actually align with their preferences and skills.

For example, in the Canadian context, studies indicate that many location decisions are linked to the international prominence of destinations (Di Biase and Bauder 2005). As Bégin-Gillis (2010) argues, "many immigrants choose Toronto simply because that is all they know of Canada" (also see McDonald (2004)). Similarly, Teo (2003) concludes that "unfamiliarity means that decisions regarding their initial destination are often reliant on secondary information sourced from earlier migrants, immigration companies, the Internet or other sources" (also see Fotros (2018)). Recognizing that perceptions of places can skew location decisions toward prominent cities, some provinces have attempted to influence perceptions by providing prospective immigrants with information about less prominent locations. Evaluating these programs, Bégin-Gillis (2010) notes that when "prospective immigrants are provided with more information and provided with more choices, they often choose differently."

Interventions that use behavioral insights to counteract the effects of cognitive biases have been used in a wide variety of policy domains (Thaler and Sunstein 2009; OECD 2017). Our approach builds on these behavioral insights, and seeks to enhance the choice architecture for newly invited economic immigrants by systematically providing personalized recommendations as they make a decision about where to settle in the destination country. The recommendations from our decision helper reflect individuals' location preferences as well as data-driven predictions about the locations where they are likely to attain the highest earnings given their profile. The recommendations thus act as informational nudges that counteract the effects of limited information and cognitive biases, assisting immigrants in making more informed location decisions.

The primary anticipated users of the tool are economic immigrants who have been invited to apply for permanent residence via an economic admissions stream and are in the process of selecting a landing location within the destination country. Since governments communicate regularly with immigrants during this transition period to provide

information about the immigration process, the decision helper could be offered online via a user interface at minimal cost. We expect that immigrants' likelihood of using the decision helper tool will depend on their prior level of certainty in their destination decision. As a result, the decision helper provides a complementary source of information with minimal disruption to existing sets of resources that guide decision making. Our decision helper could—with appropriate adjustments—also be useful for other immigrants or even Canadian-born residents as an informational tool in deciding if and where to move. Note that even Canadian-born residents typically do not have access to granular administrative data which would allow them to discern how workers with similar skill-sets and backgrounds fare in various locations.

2.2 Design

The decision helper approach combines three main stages: modeling and prediction, preference constraints, and recommendations. Figure 1 is a flowchart of these different steps.

2.2.1 Modeling/Prediction

Our approach leverages individual-level administrative data from prior immigrant arrivals. Governments routinely gather information on applicants to economic immigration programs, ranging from individuals' skills, education, and prior job experience to their age, gender, and national origin. Using unique identifiers, these background characteristics can be merged with applicants' initial landing locations and economic outcomes, such as earnings or employment. Although governments with economic admission streams collect these data as part of normal program administration, to our knowledge they have not been systematically leveraged to predict how immigrants with different profiles fare across various landing locations.

Our approach leverages these historical data to fit a set of location-specific, supervised machine learning models that serve as the basis for recommended landing locations for future immigrants. These models learn how immigrants' background characteristics and skill sets are related to taxable earnings within each potential landing location, while also accounting for local trends over time. The models can then be used to predict an economic immigrant's expected earnings at any of the possible landing locations. To minimize bias in the models' earnings predictions, a broad set of characteristics related to immigrants' backgrounds, qualifications, and skills are included as predictors. Furthermore, to reduce the possibility that observed patterns are driven by location-specific self-selection bias, the data used to train the model could exclude prior immigrants with standing job offers, family ties to specific locations, or other special situations. See the Supplementary Material (SM) appendix for a formalization and decomposition of the possible selection bias in the models, along with discussion on how such bias can be limited.

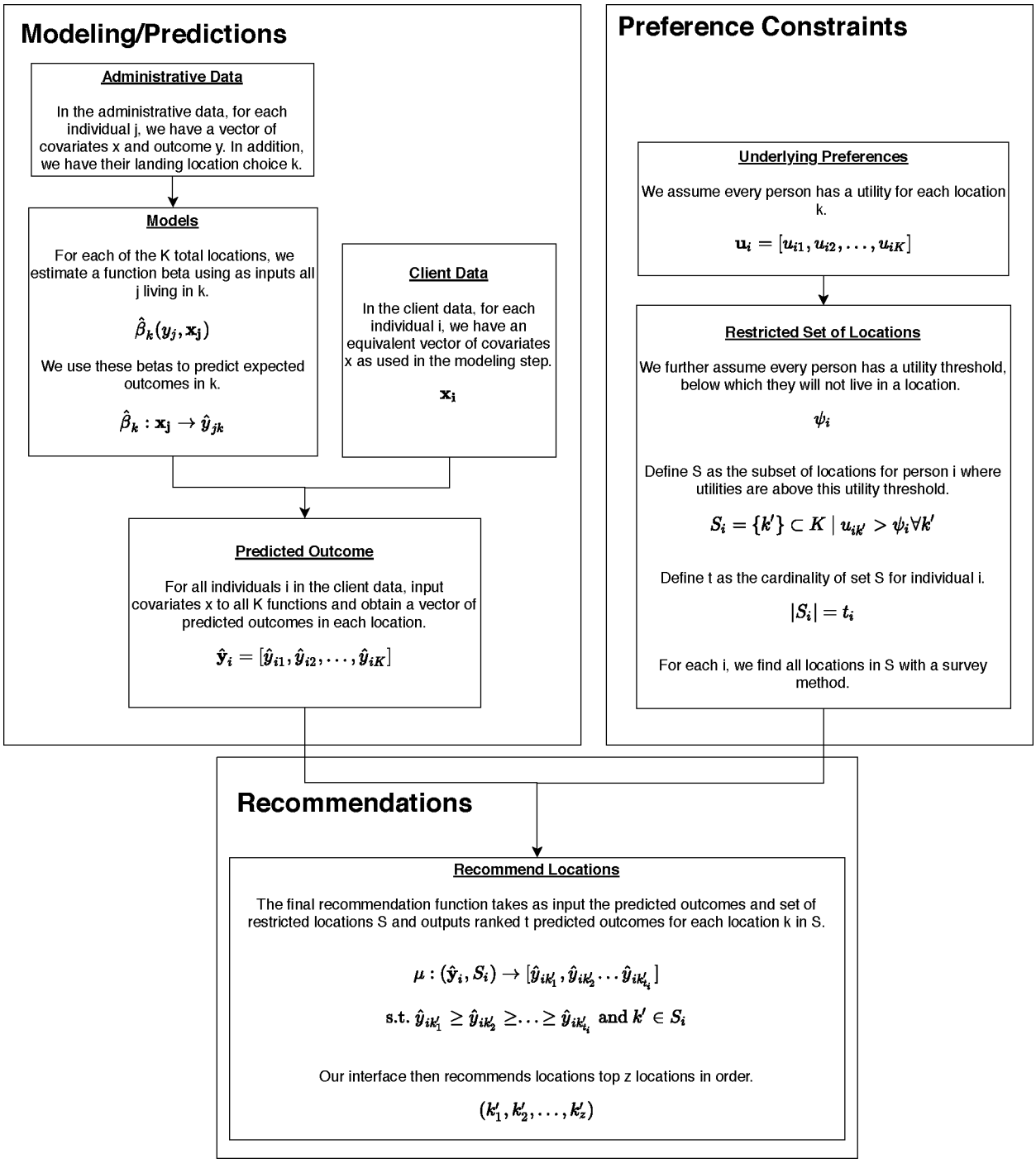


Figure 1: Decision Helper

2.2.2 Preference Constraints

Next, the decision helper elicits and incorporates individuals' location preferences. Even if a particular location is predicted to be the best for an immigrant in terms of expected earnings, if the immigrant strongly prefers not to live there, then recommending such a location will result in limited uptake and downstream dissatisfaction. To accommodate individual preferences, we identify the list of locations each user would consider unacceptable, and then limit the choice set to the remaining locations prior to optimizing on expected earnings.

Our approach is agnostic to the specific method used to rule out unacceptable locations prior to making recommendations. For instance, users can be directly queried to indicate regions of the country where they are unwilling to settle. Alternately, users can be asked to provide preferences regarding specific location characteristics via direct questioning or conjoint survey tasks, spanning identifiable features such as urban density, climate, and the relative availability of amenities. These can be mapped onto observable characteristics of each landing location in order to identify acceptable locations according to the expressed preferences.

2.2.3 Recommendations

After taking individual preferences into account to constrain the choice set, the remaining locations are then ranked with respect to the individual's predicted earnings in each location. The decision helper delivers these recommendations to the user in the form of an informational nudge via an online interface. Users can be given either a reduced number of the top ranked locations (e.g. the top 3 locations) or a full ranked list of the locations, along with accompanying information. Users choose whether to use the tool and follow the recommendations. This approach is non-coercive, seeking only to inform those who can benefit while not interfering with immigrants who may already have solid plans or private information guiding their selection of a particular location.

3 Empirical Analysis

We illustrate the potential of a decision helper to recommend initial landing locations for newly invited immigrants by evaluating data from Canada's flagship Express Entry system. Express Entry is a system that manages applications to Canada's high skilled economic immigration programs, which select skilled workers for admission through a points-based system.[1] The Express Entry application process involves several stages and is designed to select applicants who are most likely to succeed in Canada (Immigration, Refugees and Citizenship Canada 2019). First, eligible candidates create and submit a

[1] These economic immigration programs are the Federal Skilled Worker Program, the Federal Skilled Trades Program, the Canadian Experience Class, and a portion of the Provincial Nominee Program.

profile to indicate their interest in moving to Canada. If candidates meet the minimum requirements for one of the programs managed by Express Entry, they are entered into the Express Entry pool, awarded points based on information in their profile and ranked according to the Comprehensive Ranking System (CRS). The CRS awards points based on human capital characteristics including language skills, education, work experience, age, employment and other aspects that previously have been shown to be associated with long-term economic success in Canada. Factors in the CRS are generally grouped under two categories: core points and additional/bonus points. Candidates with the highest rankings in the pool are invited to apply online for permanent residence following regular invitation rounds. If candidates' CRS scores are above a specified threshold for an invitation round, they receive an invitation to apply for permanent residence, to be submitted within 90 days of receiving an invitation (Immigration, Refugees and Citizenship Canada 2019). If an application is approved by an IRCC officer, permanent resident visas are issued so that the applicant and his or her accompanying family members can be admitted to Canada. Processing times for Express Entry profiles vary depending on the program of admission, but the majority of applications are processed within six months.

Since the system initially launched in 2015, it has steadily expanded. In 2018, 280,000 Express Entry profiles were submitted, and 92,331 people were admitted to Canada through the Express Entry system (Immigration, Refugees and Citizenship Canada 2019). The growing importance of this system is mirrored by similar developments within other advanced economies. For instance, Australia and New Zealand also use a similar expression of interest process to determine applicants' eligibility and offer invitations to apply for permanent residence. Countries such as Austria, Japan, South Korea, and the United Kingdom also have aspects of points-based admissions systems built into their economic immigration programs.

3.1 Data

We draw on data from the Longitudinal Immigration Database (IMDB)—the integrated administrative database that Immigration Refugees and Citizenship Canada reports on the outcomes of immigrants. IMDB was initially a basic linkage between tax files and the Permanent Residents Database. The IMDB (2019 release) includes more than 12 million immigrants who landed in Canada between 1952 and 2018 and income tax records from 1982 to 2017, as well as all temporary residents, Express Entry Comprehensive Ranking System scores, citizenship uptake, and service usage for settlement programs. We subset the data to include principal applicants who arrived between 2012 and 2017 under the following programs: the Federal Skilled Worker program, the Federal Skilled Trades program, and the Canadian Experience Class. Applications for these admission streams are managed by the Express Entry system. We further subset the data to exclude individuals who were minors at the time of arrival ($\leq$18), as well as individuals who did not file a tax return while living in Canada. Given that the Express Entry system does not apply to Quebec, we also exclude all immigrants who first landed in Quebec or entered

on an immigration program run by Quebec. The final sample size consists of 203,290 unique principal applicants.

3.2 Measures

The outcome measure is immigrants' individual annual employment income, measured at the close of the first full calendar year after arrival. We model this outcome as a function of a variety of predictors that are either prior to or contemporaneous with an immigrant's arrival. Predictors used in the modeling stage include age at arrival, citizenship, continent of birth, education, family status, gender, intended occupation, skill level, English ability, French ability, having a prior temporary residence permit for study in Canada, having a prior temporary residence permit for work in Canada, having previously filed taxes in Canada, arrival month, arrival year, immigration category, and Express Entry indicator. See Table S2 in the SM appendix for more information and summary statistics on these measures.

We map immigrants' landing locations to a specific Economic Region (ER) within Canada, using census subdivision codes. As regional predictors, we also include the population and the unemployment rate within the ER in the quarter of immigrants' arrival. An ER is a Canadian census designation that groups neighboring census divisions to proxy regional economies. We use the ER as the primary unit throughout the analysis. There are 76 ERs in total, but in our analysis there are 52 after excluding Quebec and merging the smallest ERs using standard census practices.

3.3 Models

The modeling approach is based on the methodology developed in Bansak et al. (2018). We first merge historical data on immigrants' background characteristics, economic outcomes, and geographic locations. Using supervised machine learning methods, we fit separate models across each ER estimating an immigrant's annual employment income as a function of the predictors described above. These models serve as the basis for the decision helper tool's recommendations, as they allow for the generation of annual employment income predictions across each ER that are personalized to each immigrant's background characteristics.

As our modeling technique, we use stochastic gradient boosted trees. We use 10-fold cross-validation within the training data to select tuning parameter values, including the interaction depth, bag fraction, learning rate, and number of boosting iterations.[2] More

[2] The cross-validated R^2 for our primary set of models (where the units of analysis are principal applicants) is 0.54. Within the context of incomes – which are highly skewed and difficult to predict – this is relatively high. This represents a substantial improvement over the R^2 for an analogous linear regression model using cross-validation (0.34). See the SM appendix for more details, including a breakdown of the relative importance of the predictors in the boosted trees models.

details are provided in the SM appendix.

3.4 Simulations

To estimate how our proposed decision helper would affect income and influence location decisions, we perform a series of backtests using historic Express Entry cohorts. Specifically, we implement a series of simulations in which the decision helper provides recommendations to individual immigrants. We then simulate uptake of these recommendations, and for individuals who follow the recommendation, we compare the expected income at that location and at the location where they actually landed.

After training the models, we input the background characteristics of 2015 and 2016 Express Entry principal applicants ($n = 17,640$) to obtain predicted income across ERs. These predictions serve as the basis for the simulated recommendations. The degree to which immigrants would follow such recommendations is unknown. To model these dynamics, the simulations vary two parameters that reflect different assumptions concerning the influence of the recommendation on immigrants' location decisions.

The first parameter is the compliance rate, denoted by π, which is defined as the probability that individuals will follow the recommendations. We assume that the probability of following a recommendation decreases linearly across income quantiles. We apply an upper bound π_{max} to the individuals with the lowest actual income. We then linearly interpolate to a value of $\pi = 0$ across the income distribution. Each individual within the prediction set thus receives an individual compliance parameter, π_i. Functionally, the average compliance rate across the distribution is $\pi_{max}/2$. For example, in the simulations with $\pi_{max} = .30$, on average 15 percent of immigrants are expected to follow the recommendation, and the probability varies from a high of 30 percent for the lowest-income immigrants to close to zero percent for the highest-income immigrants. We chose to vary π_i as a function of income under the assumption that wealthier individuals are more likely to have self-selected into location-specific employment opportunities. Simulations that impose a uniform compliance parameter instead also suggest substantial gains (see SM appendix).

The second parameter is the number of acceptable locations, denoted by ϕ. Each applicant is assumed to have a set of idiosyncratic preferences regarding locations, which results in a ranked preference order of ERs, ranging from the most attractive (1) to the least attractive (52). The ϕ parameter determines how many top-preference-ranked ERs are included within the optimization. Location preferences are unmeasured within the administrative data, and must be inferred. Using the landing ER as the dependent variable, we fit a multinomial logit model and proxy immigrants' preferences using their predicted probability of landing in each ER. After obtaining predictions for each Express Entry case, we rank order locations by each individual's predicted probability of landing, randomly breaking ties. For each individual, the resulting preference ranks are then used in conjunction with the parameter ϕ to define their set of acceptable locations, which serves as the initial set of locations considered when selecting the locations with the highest expected incomes. To guarantee that gains are not entirely driven by subsets

of locations with certain characteristics, we run simulations entirely removing certain ERs from consideration, and find no major deviation from our core results (see SM appendix).

We conduct a simulation for various combinations of parameters. For each immigrant we consider only the top ϕ preference-ranked locations, and return the three locations within this subset that are expected to yield the highest employment income for the immigrant. We assume that individuals who follow the recommendations have an equal probability of selecting each of these three locations. However, with probability (1-π_i), individuals will select their original location rather than any of the recommendations. For each case, we draw from a uniform distribution bounded by 0 and 1 to determine whether the case will take the recommendation or not. If (draw) > π, cases are assumed not to have followed the recommendation, and their location is recorded as their actual location. Their income is recorded as their predicted income within their actual location; for these immigrants there is no gain in income from using the tool. For immigrants where (draw) < π, we perform a second draw to determine which of the three locations they will select. For these immigrants, the expected gain in income is computed as the difference between the expected income they would earn in the recommended location and the location they would have chosen without the tool. After performing these random draws for each of the 2015 and 2016 Express Entry users, we obtain the total expected difference in location counts and income.

4 Results from Empirical Analysis

4.1 Current Settlement Patterns

As shown in Figure 2, economic immigrants who arrived in the 2015 and 2016 arrival cohorts through the Express Entry system are highly concentrated within a few regions of Canada. About 78% settled in one of the four largest ERs as their initial destination, and 31% of immigrants selected Toronto. In stark contrast, only about 44% of the overall population is concentrated in those four ERs.[3]

To what extent do these concentrated settlement patterns support the goal of maximizing incomes? We evaluate this by estimating each economic immigrant's expected income in every potential landing region as a function of their background characteristics and qualifications (see SM for details). Among economic immigrants who selected one of the four most selected locations, Figure 3 displays how the selected ER would rank in terms of expected earnings relative to all other potential ERs. For example, a rank of 1 for a given immigrant in the top left panel indicates that, at the time of arrival, the models estimate that Toronto ranked first (i.e. best) out of 52 possible landing locations in terms of the expected employment income for that immigrant.

Although the models suggest that a small subset of individuals selected an initial

[3] We exclude Quebec from this computation to have an accurate comparison.

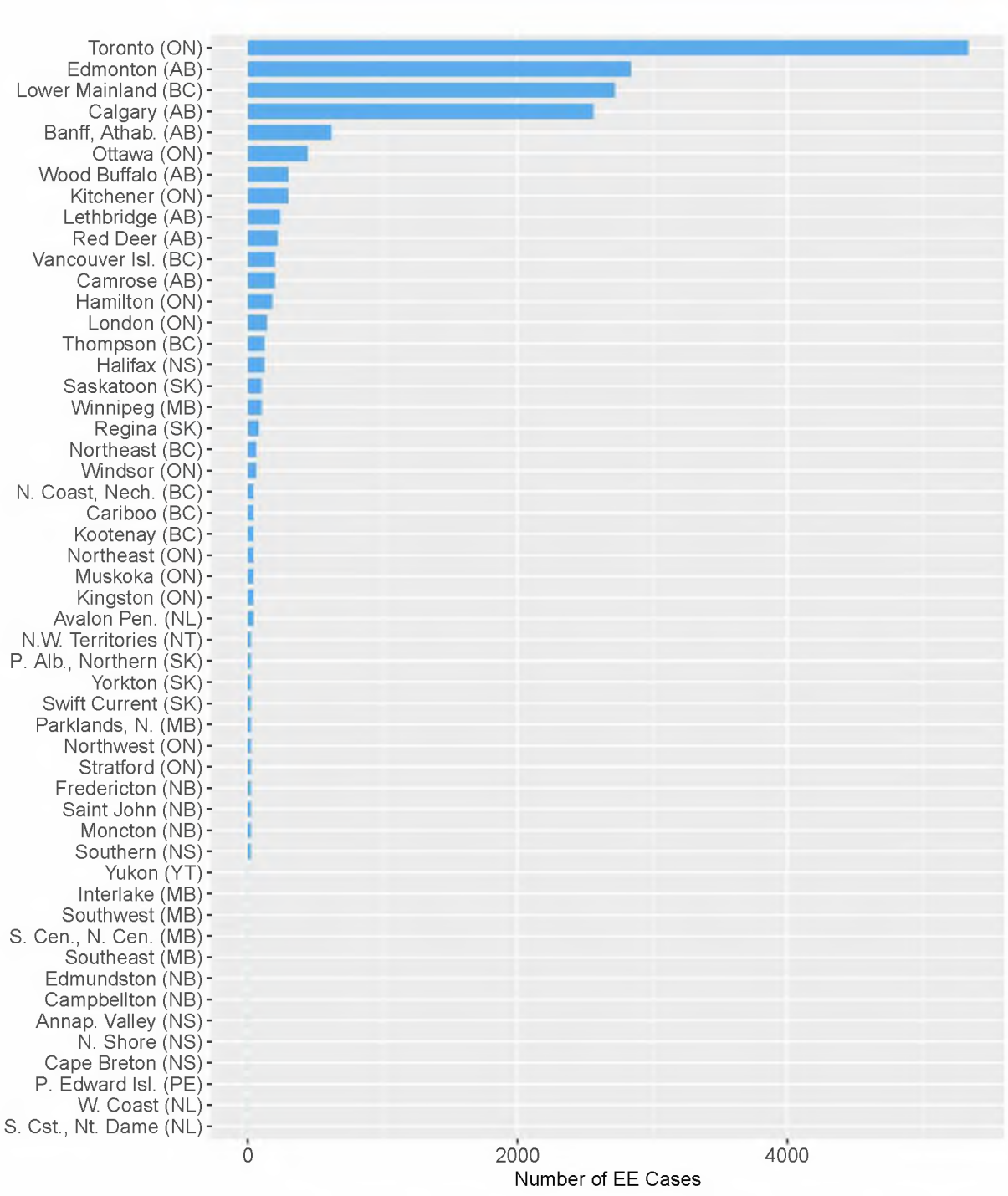


Figure 2: Landing Economic Regions for Express Entry Cohorts Arriving 2015-2016. N=17,640.

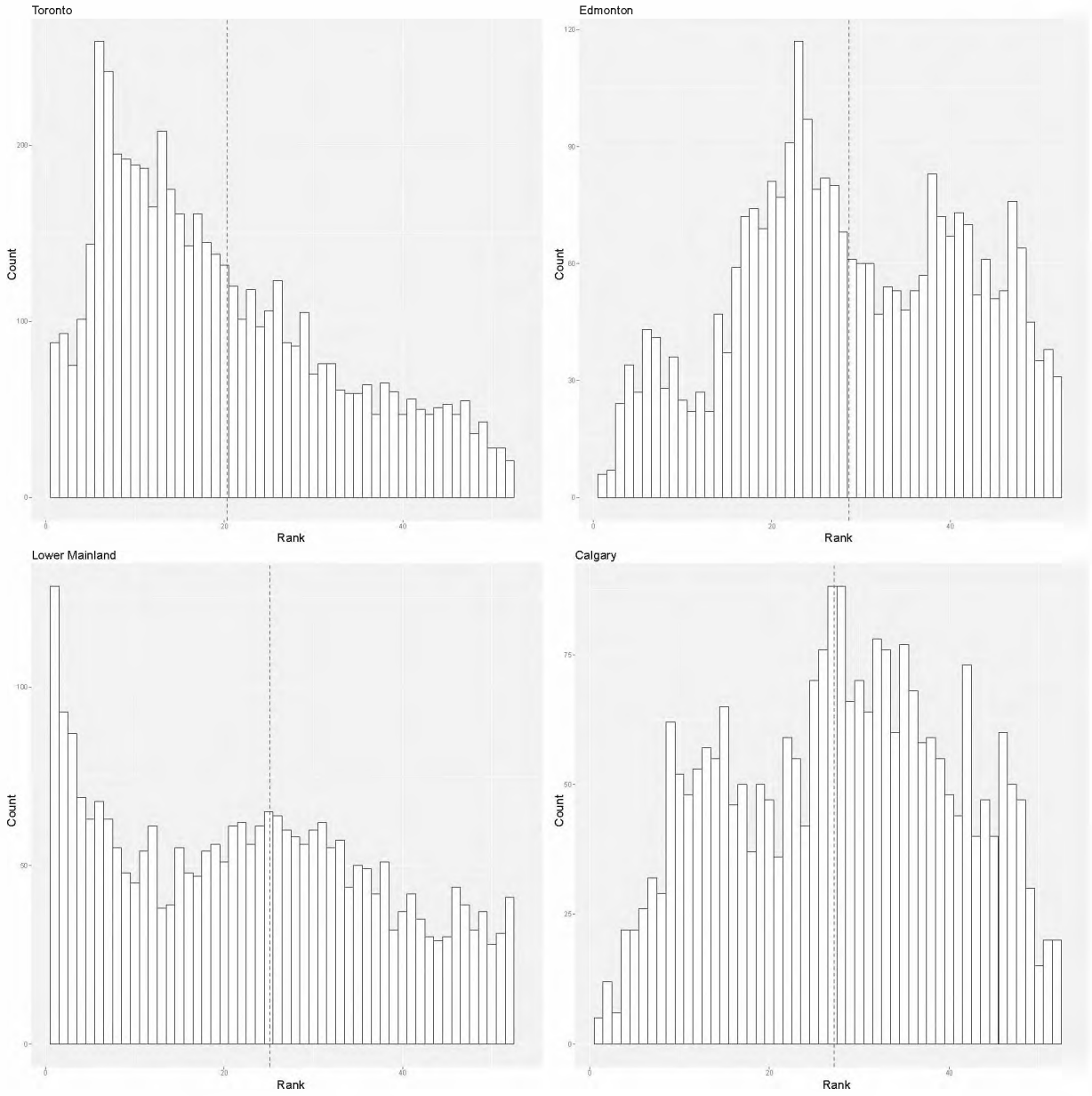


Figure 3: Estimated Annual Income: Rank of Landing Economic Region vs. All Economic Regions. N=17,640

location that maximizes their expected income, we find that for many economic immigrants the chosen location is far from optimal in terms of expected income. For instance, among economic immigrants who chose to settle in Toronto, that landing location only ranked approximately 20th on average out of the 52 ERs in terms of maximizing expected income in the year after arrival. In other words, the data suggest that for the average economic immigrant who settled in Toronto, there were 19 other ERs where that immigrant had a higher expected income than in Toronto. The situation is similar for other prominent locations, including Edmonton, Lower Mainland, and Calgary, where the average ranks are 28, 24, and 26, respectively. Across Canada as a whole, the average rank is a mere 26.5. This suggests that many immigrants do not select locations where individuals with similar background characteristics tend to achieve the best economic outcomes, and there is potential to improve immigrants' landing choices.

4.2 Changes in Expected Income and Arrival Locations

Figure 4 displays the results from backtests that examine how a data-driven decision helper tool may influence the expected incomes and location decisions of economic immigrants who enter through the Express Entry system (see SM for details). The top panel shows the estimated effects on the average expected income one year after arrival for economic immigrants across the entire backtest cohort of 2015 and 2016 arrivals. We simulate effects using a varying set of parameters, including the share of immigrants who are assumed to follow the recommendations (horizontal axis) and the number of locations that are considered for each recommendation, based on the immigrants' modeled location preferences (colors and symbols). These results report the average gain from 100 simulation runs.

The simulations suggest gains in expected annual incomes, even under scenarios in which compliance is low and/or location preferences are highly restrictive. For example, using the assumption that on average only 10% of immigrants settle in one of the recommended locations, and that individuals' location preferences will rule out 42 of the 52 possible locations as unacceptable (the scenario labeled "Top 10"), the simulation yields an average gain in expected annual employment income one year after arrival of $1,100, averaged across the full cohort. This amounts to a cumulative gain of $55 million in total income for every 50,000 cases that enter Canada via the Express Entry system. Note that these gains are entirely driven by the 10% of immigrants who follow the recommendations, since we assume zero gains for the rest of the cohort. Immigrants who do follow the recommendation increase their expected annual employment income one year after arrival by $10,600 on average, relative to the estimated income at the location they would have selected without using the decision helper. These gains are large relative to the observed average first-year income within the prediction sample ($49,900).

Across the full cohort, the total expected gains from implementing the decision helper would be larger if immigrants had less restrictive location preferences and/or more immigrants followed the recommendations. For example, under the assumption that 15% of immigrants follow the recommendations, and the recommended locations

are chosen from a set of 25 acceptable locations, the average expected annual income one year after arrival across the cohort increases by $3,400. The SM demonstrates that these results are similar across various robustness checks, including replicating the analysis with cost of living adjusted income (Figure S3), recommending locations to maximize the joint income of principal applicants and their spouses (Figure S4), or removing smaller, larger, or growing ERs from consideration (Figure S7).

The lower panel in Figure 4 displays the anticipated impact on the distribution of economic immigrants across initial landing regions under the 15% compliance and Top 25 location simulation scenario. The results suggests that we would see a mild shift from the most populous destinations toward mid-sized landing regions. For instance, about 15% of immigrants who chose one of the four largest locations would have chosen an alternate location in Canada if they had followed the recommendation. Although there is not a marked redistribution of arrivals toward the smallest locations, the estimates for smaller locations are likely conservative given that the preferences for our simulation were derived from data on the existing residential patterns of immigrants across Canada. Figure S5 in the SM shows the expected distribution if no location preferences were taken into account. While these scenarios find that the majority of outflows continue to be associated with the four largest ERs, we find more movement into a subset of the smaller locations when we do not restrict locations based on the inferred preferences.

4.3 Changes in Expected Income for Subgroups

We also assess the distribution of potential gains across subgroups to understand the differential effects our approach could have for economic immigrants of various backgrounds. Figure 5 shows the estimates of the change in average expected incomes one year after arrival across a variety of different subgroups, again using the assumption that 15% of immigrants would follow the recommendations and that recommended locations are chosen from a set of 25 acceptable locations. While expected gains vary as a function of individuals' characteristics, the overall increase in income does not appear to be the result of disproportionate benefit on the part of any particular demographic or socioeconomic groups. Instead, we find comparable average gains across a range of subgroups of economic immigrants, including groups stratified by gender, education level, case size, landing year, and immigration category.[4]

5 Potential Limitations

The impact assessment is limited to backtests applied to historical data. Such backtests are commonly used to examine the potential impact of new approaches, but they cannot

[4] Results for case size subgroups are only shown for case sizes of 1 and 2 due to an insufficient number of cases of size greater than 2.

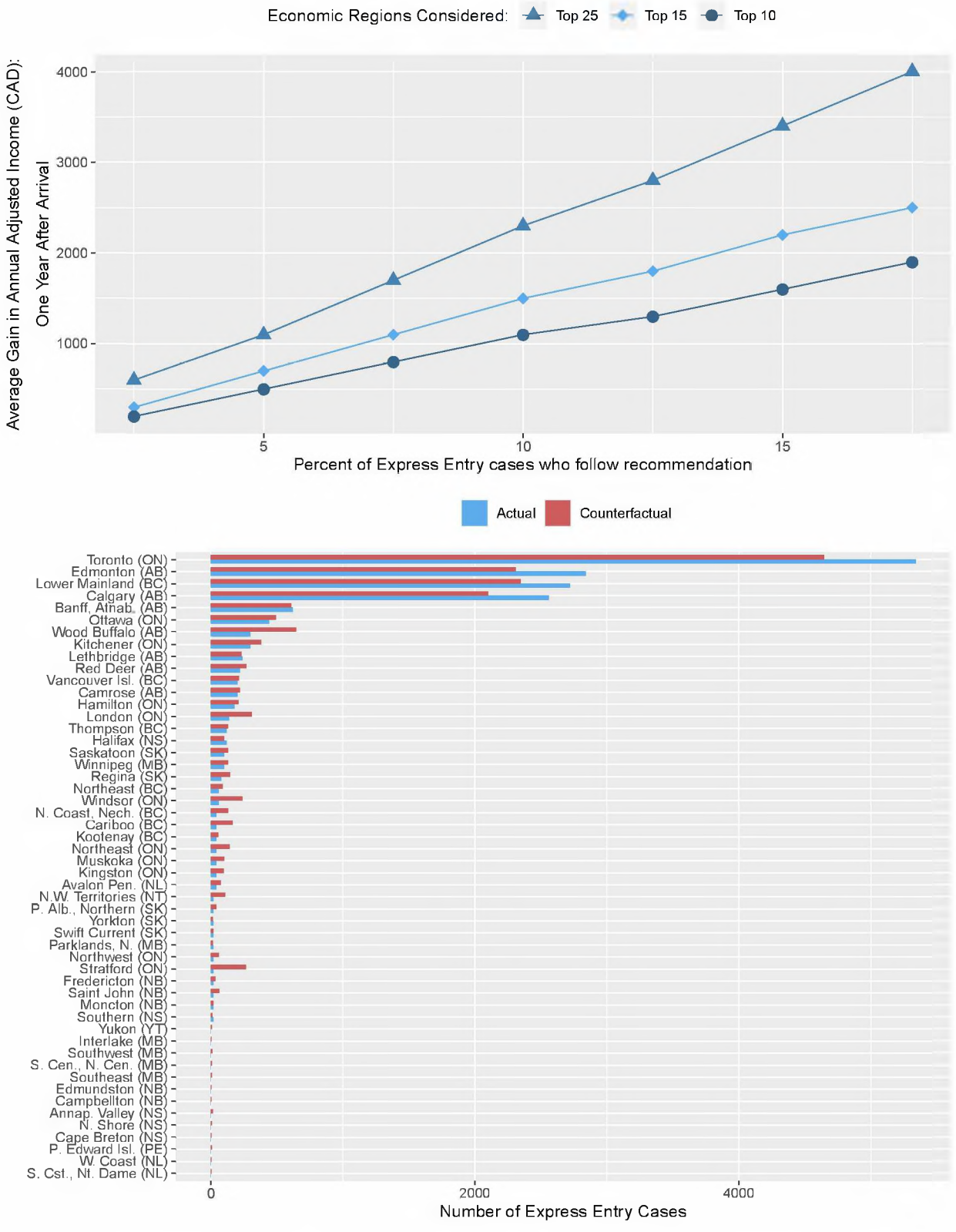


Figure 4: Estimated Average Income Gains and Shifts in Arrival Locations. N=17,640

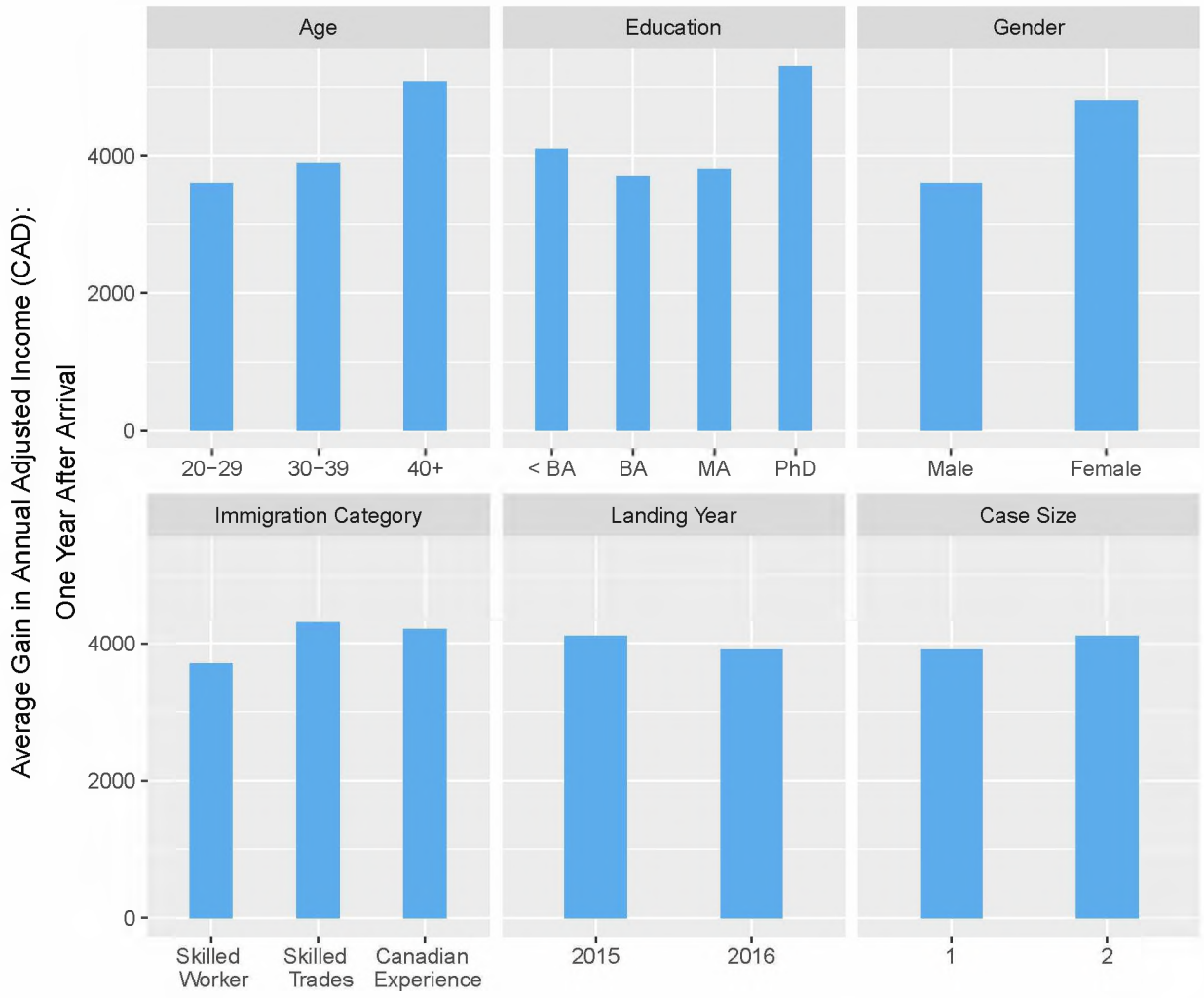


Figure 5: Estimated Average Income Gains for Subgroups. N=17,640

fully capture all the factors that may affect the potential impact of our decision helper in a prospective implementation.

For instance, economic outcomes could be influenced by compositional effects if many immigrants with a similar profile receive the same location recommendation. Modeling such effects in a backtest is challenging because their direction and magnitude is theoretically ambiguous. An increased concentration could lower expected incomes due to saturation, or alternatively, increase incomes due to local agglomeration economies, which are common for high-skilled migration (Kerr et al. 2017). Similarly, economic immigrants account for a relatively small fraction of the local labor market, implying that the direct impact of the tool on labor market saturation is difficult to determine a priori.

Although we do not model compositional effects directly in the backtests, in a prospective application the decision helper can learn potential compositional effects over time as the models are continually updated based on observed data from new arrivals, as well as local economic conditions at the time of arrival. Therefore, should a particular location cease to be a good match for a particular immigrant profile due to increased concentration, the models would adjust to this pattern over time and no longer recommend that location. In addition, the approach can be adjusted to incorporate location-specific quotas if desired by governments.

The results in a prospective implementation could also differ from the backtests if the expected incomes predicted by the machine learning models over- or under- estimate the actual incomes that newly invited immigrants attain if they were to follow the recommendations. Such prediction errors could occur for a variety of reasons, including immigrants selecting into locations based on unobserved characteristics. Such unobserved characteristics can be separated into two broad categories. The first category includes unobserved characteristics that are unrelated to any particular location, and hence could be associated with higher (or lower) earnings potential in all possible locations. Examples would include an individual's unmeasured abilities or motivation. The second category includes unobserved characteristics that are unique to specific locations, and hence represent an earnings advantage for a particular individual only in a select location (or select locations). Examples would include an individual's unknown job offer or social network (e.g. family members) in a specific location. See the SM appendix for a more comprehensive discussion, formalization, and decomposition of the possible selection bias.

These concerns about selection bias driving the results in our backtests are partially addressed by the fact that we train flexible models on a rich set of covariates derived from application data to economic admissions programs. That is, conditional on the broad set of background and skills-based characteristics we observe, it is not likely that individuals have self-selected into locations as a function of unobserved non-location-specific characteristics that account for the full magnitude of the estimated gains we observe in the backtests. For instance, for individuals who are identical on the observed characteristics (e.g. same age, education, profession, skills, etc.), their variation in un-

measured variables such as motivation would need to both be integral in their location choices and significantly affect their earnings potential (see SM for details).

Finally, while the backtest results suggest the possibility of gains if the decision helper were implemented, it is important to obtain reliable estimates of impact through a prospective randomized-controlled trial (RCT). An RCT would randomly assign new arrivals to receive recommendations from the decision helper, allowing for a rigorous evaluation of the tool's impact on a variety of outcomes, including incomes, satisfaction, and location patterns.

6 Potential Risks

To examine potential risks, it is important to consider how introducing a decision-helper tool would influence the status quo. The decision-helper provides additional information to incoming economic immigrants so that they can make more informed location decisions, from among the set of possible locations that match their preferences. In doing so, it does not limit immigrants' agency to choose their final settlement location. Given that the recommendations are based on historical data, the predictions may be subject to error. As a result, the decision helper should be viewed as a complement, rather than a replacement, of the existing information streams and processes that governments use to inform immigrants about potential destinations.

Care needs to be taken to transparently communicate the locational recommendations to users. In particular, users should be made aware that the recommendations are based upon the goal of maximizing the particular metric of near-term income (or whichever specific metric has been applied) and that the predictions reflect the outcomes that recent immigrants with similar profiles have attained in the past. This does not guarantee that the user's realized income will be optimal at the recommended location, as expected income cannot take into account all possible factors that are unique to an individual. Nor does it imply that the recommended location would necessarily be optimal in terms of other possible life goals or long-term earnings. While we do not have evidence that selecting locations based on near-term earnings will have a negative impact on these longer-term outcomes, this is a theoretically possible risk that would need to be monitored over the course of a prospective implementation.

7 Discussion

A growing number of countries have implemented economic immigration programs to attract global talent and generate growth. However, economic immigrants lack access to personalized information that would help them identify their most beneficial initial settlement locations. In this study, we propose a data-driven decision helper that delivers informational nudges to counteract the effects of limited information and cognitive biases. The decision helper harnesses insights from administrative data to recommend the

locations that would maximize their expected incomes and align with their preferences. We illustrate its potential by conducting backtests on historical data from the Canadian Express Entry system. The results suggest that economic immigrants currently select sub-optimal locations, and that there could be gains in expected incomes from providing data-driven, personalized location recommendations. While the results from our backtests suggest potential gains in the Canadian context, it is important to assess the impact in the context of a pilot initiative with a randomized controlled trial.

The decision helper outlined in this study is adaptable and could be rapidly implemented within existing institutional structures in various countries. First, as we demonstrate via application to Canadian data, many governments are already collecting administrative data that can be used to generate recommendations. Moreover, governments administering economic admission programs engage in regular communication with applicants, implying that the decision helper can easily be made accessible to a large group of users. In light of the gains observed in the backtests, these limited costs suggest a positive return on investment, even in scenarios where only a small share of immigrants follow the recommendations. Second, the approach is flexible in terms of implementation and can be adjusted to the specific priorities identified by the destination government. For example, the decision helper could be used to improve other measurable integration outcomes (for example, longer-term measures of income), incorporate location-specific quotas, and allow for a wide variety of approaches to elicit preferences and display recommendations. Third, the approach is designed as a learning system such that the models for the predictions are continually updated using observed data from new arrivals and changing local economic conditions. The decision helper therefore learns synergies between personal characteristics and landing locations as they evolve over time and adjusts the recommendations accordingly. Finally, the approach supports individual agency. The decision helper provides immigrants with personalized recommendations that help them make a more informed decision, but immigrants decide whether to use it, and they can decline the recommendations. The decision helper thus complements rather than replaces existing information streams and processes.

In sum, a data-driven decision helper holds the potential to assist incoming economic immigrants in overcoming informational barriers and choosing better landing locations. In addition, the approach we outline offers governments the ability to leverage administrative data to increase economic returns within the structures of their existing admission process. Together, we expect these factors to improve the well-being of economic immigrants and the communities in which they settle.

8 Acknowledgments

This study was completed as part of a Data Partnership Arrangement between Immigration, Refugees and Citizenship Canada (IRCC) and the Immigration Policy Lab. The analysis, conclusions, opinions, and statements expressed in the material are those of the authors, and not necessarily those of the IRCC. This research received generous sup-

port from Eric and Wendy Schmidt by recommendation of Schmidt Futures. We also acknowledge funding from the Charles Koch Foundation. These funders had no role in the data collection, analysis, decision to publish, or preparation of the manuscript.

References

Akbari, Ather H and Jennifer S Harrington. 2007. Initial location choice of new immigrants to Canada. Technical report. Working Paper 05-2007, Atlantic Metropolis Centre.

Åslund, Olof and Dan-Olof Rooth. 2007. "Do when and where matter? Initial labour market conditions and immigrant earnings." *The Economic Journal* 117(518):422–448.

Bansak, Kirk, Jeremy Ferwerda, Jens Hainmueller, Andrea Dillon, Dominik Hangartner, Duncan Lawrence and Jeremy Weinstein. 2018. "Improving refugee integration through data-driven algorithmic assignment." *Science* 359(6373):325–329.

Bégin-Gillis, Margot. 2010. "Immigrant settlement in rural Nova Scotia: Impacting the location decisions of newcomers." *Papers in Canadian Economic Development* 12:1–18.

Borjas, George J. 1995. "The economic benefits from immigration." *Journal of Economic Perspectives* 9(2):3–22.

Brezzi, Monica, Jean-Christophe Dumont, Mario Piacentini and Cécile Thoreau. 2010. Determinants of localization of recent immigrants across OECD regions. Technical report. Paper for OECD Workshop "Migration and Regional Development", June 7, 2010, Paris.

Burchardi, Konrad, Thomas Chaney, Tarek Alexander Hassan, Stephen Terry and Lisa Tarquinio. 2020. Immigration, Innovation, and Growth. Technical report. NBER Working Paper No. 27075.

Chetty, Raj, Nathaniel Hendren and Lawrence F Katz. 2016. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." *American Economic Review* 106(4):855–902.

Chiswick, Barry R and Paul W Miller. 2004. "Where immigrants settle in the United States." *Journal of Comparative Policy Analysis: Research and Practice* 6(2):185–197.

Damm, Anna Piil. 2009. "Determinants of recent immigrants' location choices: Quasi-experimental evidence." *Journal of Population Economics* 22(1):145–174.

Damm, Anna Piil. 2014. "Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants." *Journal of Urban Economics* 79:139–166.

Desiderio, Maria Vincenza and Kate Hooper. 2016. The Canadian expression of interest system: A model to manage skilled migration to the European Union? Technical report. Migration Policy Institute Europe.

Di Biase, Sonia and Harald Bauder. 2005. "Immigrant settlement in Ontario: Location and local labour markets." *Canadian Ethnic Studies* 37(3):114.

Fotros, Homayoon. 2018. Destination matters: Policy options to balance the distribution of Iranian immigrants in Canada. Technical report. PhD Thesis, Simon Frazer University.

Hugo, Graeme. 2008. "Immigrant settlement outside of Australia's capital cities." *Population, Space and Place* 14(6):553–571.

Hunt, Jennifer and Marjolaine Gauthier-Loiselle. 2010. "How much does immigration boost innovation?" *American Economic Journal: Macroeconomics* 2(2):31–56.

Hyndman, Jennifer, Nadine Schuurman and Rob Fiedler. 2006. "Size matters: Attracting new immigrants to Canadian cities." *Journal of International Migration and Integration* 7(1):1.

Immigration, Refugees and Citizenship Canada. 2019. Express Entry year-end report 2018. Technical report.

Kaida, Lisa, Feng Hou and Max Stick. 2020. "Are refugees more likely to leave initial destinations than economic immigrants? Recent evidence from Canadian longitudinal administrative data." *Population, Space and Place* p. e2316.

Kerr, Sari Pekkala, William Kerr, Çağlar Özden and Christopher Parsons. 2016. "Global talent flows." *Journal of Economic Perspectives* 30(4):83–106.

Kerr, Sari Pekkala, William Kerr, Çağlar Özden and Christopher Parsons. 2017. "High-skilled migration and agglomeration." *Annual Review of Economics* 9:201–234.

Ludwig, Jens, Greg J Duncan, Lisa A Gennetian, Lawrence F Katz, Ronald C Kessler, Jeffrey R Kling and Lisa Sanbonmatsu. 2013. "Long-term neighborhood effects on low-income families: Evidence from moving to opportunity." *American Economic Review* 103(3):226–31.

Massey, Douglas S. 2008. *New Faces in New Places: The Changing Geography of American Immigration*. Russell Sage Foundation.

McDonald, James Ted. 2004. "Toronto and Vancouver bound: The location choice of new Canadian immigrants." *Canadian Journal of Urban Research* pp. 85–101.

Mossad, Nadwa, Jeremy Ferwerda, Duncan Lawrence, Jeremy M Weinstein and Jens Hainmueller. 2020. "In search of opportunity and community: Internal migration of refugees in the United States." *Science Advances* (Forthcoming).

OECD. 2017. Behavioural insights and public policy: Lessons from around the world. Technical report.

Peri, Giovanni, Kevin Shih and Chad Sparber. 2015. "STEM workers, H-1B visas, and productivity in US cities." *Journal of Labor Economics* 33(S1):S225–S255.

Taylor, Andrew J, Lauren Bell, Rolf Gerritsen et al. 2014. "Benefits of skilled migration programs for regional Australia: Perspectives from the Northern Territory." *Journal of Economic & Social Policy* 16(1):35.

Teo, Sin Yih. 2003. Imagining Canada: Tracing the cultural logics of migration amongst PRC immigrants in Vancouver. Technical report. PhD Thesis, University of British Columbia.

Thaler, Richard H and Cass R Sunstein. 2009. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Penguin.

Tonkin, Sue and Sue Tonkin. 1993. Initial location decisions of immigrants: Results from the longitudinal survey of immigrants to Australia (LSIA) pilot. Technical report. Australian Government Pub. Service.

Tuli, Sajeda. 2019. "Migrants want to live in the big cities, just like the rest of us." *The Conversation* March 31, 2019.

Tversky, Amos and Daniel Kahneman. 1974. "Judgment under uncertainty: Heuristics and biases." *Science* 185(4157):1124–1131.

Supplementary Material

S1 Decision Helper Workflow

In this section, we provide additional details and formalize our decision helper approach. This workflow consists of three stages, modeling/prediction, preference constraints, and recommendations. We repeat the visualization of the workflow with additional formal notation in Figure 1, and describe each step of the process in detail.

S1.1 Modeling/prediction

In the first stage, we use training data to build a series of models that predict expected outcomes in a particular location. This process begins by gathering a set of **Administrative Data**, an individual-level dataset containing information about prior immigrant arrivals. This dataset must consist of individuals that are similar to the eventual users of the decision helper tool.

For each individual in the administrative data, we need three pieces of information: a collection of covariates, a measurable outcome, and a choice of landing location. Because our goal is to determine unique synergies between individual-level profiles and outcomes in a particular landing spot, we estimate models for each location separately.

For each individual in the administrative data $j = 1,\ldots,m$, let the outcome of interest be denoted y_j and the landing decision denoted $w_j \in \{1,\ldots,K\}$. Let $\vec{x}_j$ represent a p-dimensional vector of relevant covariates for individual j, and x_{ir} represent the r-th feature in the p-dimensional vector. Our goal in the model training portion of the workflow is to predict the outcome based on the relevant covariates and specified landing location; that is, to estimate function β mapping $\vec{x}_j$ to y_j. As we want to find separate functional forms for each location $k = 1,\ldots,K$, we estimate K total functions $\beta_k(\vec{x}_j|w_j = k)$. We find an approximation $\hat{\beta}_k$ to β_k by minimizing the expected value of a specified loss function $L(y_j,\beta_k(\vec{x}_j)$ over the joint distribution of $(y,\vec{x})$:

$$\hat{\beta}_k = \underset{\beta_k}{\operatorname{argmin}}\, \mathbb{E}_{(y,\vec{x})}\, L(y,\beta_k(\vec{x}))$$

After estimating $\hat{\beta}_k$ for all K landing locations in the administrative data, we then apply these models to the **Client Data**, the potential users of the decision helper tool. For each client $i = 1,\ldots,n$, we have the same set of covariates used to train the models $\vec{x}_i$. By inputting $\vec{x}_i$ to each of the $\hat{\beta}_k$ location functions, we obtain individual predicted outcomes in each possible location. The following delineates each step in this process:

1. Denote the administrative data by matrix $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} y_1 & w_1 & x_{11} & \dots & x_{1r} & \dots & x_{1p} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_j & w_j & x_{j1} & \dots & x_{jr} & \dots & x_{jp} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_m & w_m & x_{m1} & \dots & x_{mr} & \dots & x_{mp} \end{bmatrix}$$

2. Train a set of K models,

$$\mathbf{L} = \{\hat{\beta_1}(\vec{x}_j), \dots, \hat{\beta_k}(\vec{x}_j), \dots, \hat{\beta_K}(\vec{x}_j)\}$$

as follows.

For $k = 1, \dots, K$:

a) Subset $\mathbf{A}$ to individuals for whom $w_j = k$ and call this $\mathbf{A}_k$

$$\mathbf{A}_k = \begin{bmatrix} y_1 & x_{11} & \dots & x_{1r} & \dots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_j & x_{j1} & \dots & x_{jr} & \dots & x_{jp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{m_l} & x_{m_k 1} & \dots & x_{m_k r} & \dots & x_{m_k p} \end{bmatrix}_{w=k} = \begin{bmatrix} y_1 & \vec{x}_1 \\ \vdots & \vdots \\ y_j & \vec{x}_j \\ \vdots & \vdots \\ y_m & \vec{x}_{m_k} \end{bmatrix}_{w=k}$$

Where m_k denotes the number of individuals in the administrative data for whom $w_j = k$.

b) Using the data in $\mathbf{A}_k$, model and estimate $\hat{\beta}_k$.

Note that while there are many ways to potentially model $\hat{\beta}_k$, we have found that using supervised machine learning methods provides the best flexible solution to capture complex non-linearities, interactions between covariates, and automatically engage in feature selection. To avoid overfitting on the training set, wherin $\hat{\beta}_k$ has very high predictive power in the training set but low out-of-sample predictive power, it is necessary to use cross-validation in the training process.

3. Denote the client data by matrix $\mathbf{C}$.

$$\mathbf{C} = \begin{bmatrix} x_{11} & \dots & x_{1r} & \dots & x_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \dots & x_{jr} & \dots & x_{jp} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nr} & \dots & x_{np} \end{bmatrix} = \begin{bmatrix} \vec{x}_1 \\ \vdots \\ \vec{x}_j \\ \vdots \\ \vec{x}_n \end{bmatrix}$$

4. For all clients in **C** and all k locations, estimate $\beta_k : \vec{x}_i \rightarrow \hat{y}_i$ as follows:

   For $i = 1,\ldots,n$

   For d = 1, . . . , k

   Estimate $\beta_k(\vec{x}_i)$ by applying the k-th model in **L** to $\vec{x}_i$, where $\hat{\beta}_k(\vec{x}_i) = \hat{y}_{ik}$

   Arrange $\hat{y}_{ik}$ into a vector $\vec{\hat{y}}_i = [\hat{y}_{i1}, \hat{y}_{i2}, \ldots, \hat{y}_{iK}]$

5. Produce a matrix of predicted outcomes **M**, with rows corresponding to clients and columns responding to potential landing locations as follows.

$$\mathbf{M} = \begin{bmatrix} \vec{\hat{y}}_1 \\ \vdots \\ \vec{\hat{y}}_i \\ \vdots \\ \vec{\hat{y}}_n \end{bmatrix} = \begin{bmatrix} \hat{} & & \hat{} & & \hat{} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \hat{y}_{i1} & \cdots & \hat{y}_{ik} & \cdots & \hat{y}_{iK} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \hat{y}_{n1} & \cdots & \hat{y}_{nk} & \cdots & \hat{y}_{nK} \end{bmatrix}$$

This represents the final ouptut of the modeling/prediction phase of the workflow

S1.2 Preference Constraint

The next stage of our approach involves eliciting clients' underlying preferences and ruling out locations that are inconsistent with these preferences. Specifically, we assume that for every client $i = 1,\ldots,n$, preferences for each location $k \in \{1,\ldots,K\}$ can be expressed by a utility value $u_{ik} \in \mathbb{R}$. The set of utility values over each K location is arranged in a vector $\vec{u}_i = [u_{i1}, u_{i2}, \ldots, u_{iK}]$.

We further assume that every individual has a utility threshold below which they will find a location unacceptable to live in. We denote this utility threshold ψ_i. We denote the subset of acceptable locations for each individual i as $S_i = \{k'\} \subset K$, and define acceptable locations in $\{k'\} \subset K$ as those locations where an individual's utility value is above their utility threshold value ψ_i.

$$S_i = \{k'\} \subset K | u_{ik'} > \psi_i \forall k'$$

Given every individual has their own utility vectors $\vec{u}_i$ and utility threshold value ψ_i, the cardinality of set S_i is different for each i. Define $t_i = |S_i|$, the number of acceptable locations for person i. We assume $t_i \geq 1$; that is that S_i is non-empty and at least one location is above the utility threshold.

We find set S_i acceptable locations for each i with a survey method. As expressed in the main paper, we are agnostic as to which survey device is used, as long as the method allows us to restrict the set of locations to those consistent with i's underlying preferences.

S1.3 Recommendations

The final stage of our workflow uses as input the predicted outcome vector $\vec{\hat{y}}_i$ and set of feasible locations S_i to produce a final set of recommendations for individual i. This process is formalized as follows:

Define μ as a function with predicted outcome vector $\vec{\hat{y}}_i$ and set of feasible locations S_i as inputs. The function μ then outputs vector $\hat{y}_{iR}$ which ranks expected outcomes for all t_i locations within feasible set S_i.

$$\mu : (\vec{\hat{y}}_i, S_i) \rightarrow \vec{\hat{y}}_{iR} = [\hat{y}_{ik'_1}, \hat{y}_{ik'_2} ... \hat{y}_{ik'_{t_i}}]$$

$$\text{s.t. } \hat{y}_{ik'_1} \geq \hat{y}_{ik'_2} \geq ... \geq \hat{y}_{ik'_{t_i}} \text{ and } k' \in S_i$$

Define z as the maximum number of recommendations to present to the user, and z'_i as the minimum between z and t_i.

$$z'_i = min(z, t_i)$$

The final interface will recommend the top z'_i locations to user i order. Various formats could be used to present the recommendations.

$$(k'_1, k'_2, ..., k'_{z'_i})$$

S2 Properties of the Modeling/Prediction Stage

For individuals denoted by i, let Y_i denote observed outcomes, A_i denote their chosen locations, and X_i denote their observed characteristics (which can denote a vector of covariates or a single fully stratifying variable). Further, let $Y_i(a)$ denote the potential outcome for individual i in any location $a \in S_A$, where S_A denotes the set of possible locations. In other words, $Y_i(a)$ represents the outcome that individual i would achieve if that individual had chosen location a, and $Y_i = Y_i(A_i)$.[5] In the modeling/prediction stage of the decision helper, the goal is to determine the optimal location for each individual as a function of their observed characteristics. In other words, for each stratum $X_i = x$ and at each location $a \in S_A$, the goal is to determine the following quantity of interest:

$$\theta_a(x) \equiv E[Y_i(a) | X_i = x]$$

where the expectation (and all expectations presented hereafter) is defined over the distribution of the population of interest (i.e. the population for whom the decision helper

[5] Note that the definition of the potential outcomes implies the stable unit treatment value assumption (SUTVA).

is targeted).

The goal is then to use this quantity for all $a \in S_A$ to determine each individual's optimal location(s)—that is, the location(s) for which the quantity is highest, perhaps subject to additional constraints—and then deliver informational nudges to encourage individuals to land in these locations.

However, a key impediment to using $\theta_a(x)$ in this ideal manner is that $\theta_a(x)$ is not necessarily identifiable with observed data. Instead, what is identified is the following:

$$\theta'_a(x) \equiv E[Y_i(a)|X_i = x, A_i = a] = E[Y_i|X_i = x, A_i = a]$$

That is, it is based upon $\theta'_a(x)$ (or estimates thereof) that optimal locations will be inferred for each individual, and these inferences may not perfectly match the true optimal locations as defined by $\theta_a(x)$.

Therefore, it is useful to characterize the potential bias of $\theta'_a(x)$ with respect to $\theta_a(x)$ in order to (a) understand the extent to which that bias may result in suboptimal informational nudges and (b) identify concrete actions that can be taken to limit or eliminate the bias. To do so, the following additional quantities are first defined:

$$\theta''_a(x) \equiv E[Y_i(a)|X_i = x, A_i \neq a]$$

$$p_a(x) \equiv P(A_i = a|X_i = x)$$

In addition, assume that $0 < p_a(x)$ for all a and x, and note that $\theta_a(x) = \theta'_a(x)p_a(x) + \theta''_a(x)(1 - p_a(x))$. Hence, the bias of $\theta'_a(x)$ with respect to $\theta_a(x)$ is bounded by the following:

$$B_a(x) \equiv \lim_{p_a(x)\to 0} \left(\theta'_a(x) - \theta_a(x)\right) = \theta'_a(x) - \theta''_a(x)$$

$$= E[Y_i(a)|X_i = x, A_i = a] - E[Y_i(a)|X_i = x, A_i \neq a]$$

This term is a form of selection bias that represents, within a stratum of x, the extent to which the mean potential outcome in location a is higher for individuals who actually choose a versus individuals who do not choose a.

If it can be assumed that $Y_i(a) \perp\!\!\!\perp A_i|X_i$, then the selection bias is eliminated (i.e. selection on observables). However, it could be that the potential outcomes are also related to unobserved characteristics of an individual that may also be correlated with location choices. Such unobserved characteristics can be separated into two broad categories. The first, denoted by U_i, are unobserved characteristics that are unrelated to any particular location. Examples would include an individual's unmeasured abilities or motivation. The second, denoted by V_{ai}, are unobserved characteristics that are unique to a particular location a in question. Examples would include an individual's unknown job offer or social network (e.g. family members) in location a.

Taking these unobserved characteristics into account, for any location a let the potential outcome $Y_i(a)$ be modeled as an arbitrary (and arbitrarily complex) function of

X_i, U_i, and V_{ai} as well as an exogenous error term:

$$Y_i(a) = g_a(X_i, U_i, V_{ai}) + \epsilon_i$$

where $E[\epsilon_i | X_i, A_i] = 0$. By extension, we have the following:

$$B_a(x) = E[Y_i(a)|X_i = x, A_i = a] - E[Y_i(a)|X_i = x, A_i \neq a]$$

$$= E[g_a(X_i, U_i, V_{ai}) + \epsilon_i | X_i = x, A_i = a] - E[g_a(X_i, U_i, V_{ai}) + \epsilon_i | X_i = x, A_i \neq a]$$

$$= \int g_a(x,u,v_a) dF_{U_i,V_{ai}|X_i=x,A_i=a}(u,v_a) - \int g_a(x,u,v_a) dF_{U_i,V_{ai}|X_i=x,A_i\neq a}(u,v_a)$$

$$= \int \int g_a(x,u,v_a) dF_{V_{ai}|U_i=u,X_i=x,A_i=a}(v_a) dF_{U_i|X_i=x,A_i=a}(u)$$

$$- \int \int g_a(x,u,v_a) dF_{V_{ai}|U_i=u,X_i=x,A_i\neq a}(v_a) dF_{U_i|X_i=x,A_i\neq a}(u)$$

where $F_{U,V_a|X,A}$ denotes the joint conditional distribution function of U and V_a given X and A; $F_{V_a|U,X,A}$ denotes the conditional distribution function of V_a given U, X, A; and $F_{U|X,A}$ denotes the conditional distribution function of U given X and A, all in the population of interest.

These results help to highlight what assumptions are required, and what corresponding design decisions could be made, to limit or eliminate this bias. For instance, provided a sufficiently rich set of covariates are observed in X_i, the following assumption may hold:

$$U_i \perp\!\!\!\perp A_i | X_i$$

In words, this assumption states that within strata of X, individuals are not self-selecting into locations as a function of unobserved non-location-specific characteristics. For instance, for individuals who are identical on X_i (e.g. same age, education, profession, skills, etc.), their variation in unmeasured variables U_i such as motivation is unrelated to their location choices A_i. Under the assumption that $U_i \perp\!\!\!\perp A_i | X_i$, $F_{U|X,A} = F_{U|X}$ and hence the bias term simplifies to:

$$B_a(x) = \int \int g_a(x,u,v_a) \{ dF_{V_{ai}|U_i=u,X_i=x,A_i=a}(v_a) - dF_{V_{ai}|U_i=u,X_i=x,A_i\neq a}(v_a) \} dF_{U_i|X_i=x}(u)$$

In other words, under this assumption, the bias is driven by the difference in the distribution of V_{ai} for individuals who choose a versus do not choose a, by joint strata of X and U. If we make this assumption in the context of the simulated backtests applied to the Canada Express Entry applicants, then for the resulting estimated gains to be driven purely by bias, this would mean that the average estimated gains among compliers can be accounted for by bias attributed solely to location-specific links or advantages that the individuals who chose any particular Economic Region had over otherwise identical individuals who did not choose that Economic Region. In other words, the individuals who select into a particular location have an average annual employment income ad-

vantage at that location of between $11,000 and $22,700 (depending on the simulation scenario) due to pre-determined job offers or family/social network ties compared to otherwise identical individuals who chose different locations.

Another assumption that could be made is that V_{ai} is constant in the population of interest, which could be ensured by design by redefining the population of interest and excluding individuals accordingly, e.g. excluding all individuals likely to have pre-determined job offers.[6] Under the assumption that $V_{ai} = \tilde{v}_a$ for all individuals in the population of interest, the bias term simplifies to the following:

$$B_a(x) = \int g_a(x, u, \tilde{v}_a) dF_{U_i|X_i=x, A_i=a}(u) - \int g_a(x, u, \tilde{v}_a) dF_{U_i|X_i=x, A_i \neq a}(u)$$

If the previous assumption that $U_i \perp\!\!\!\perp A_i | X_i$ is added back in, the bias is completely eliminated:

$$B_a(x) = \int g_a(x, u, \tilde{v}_a) \{ dF_{U_i|X_i=x}(u) - dF_{U_i|X_i=x}(u) \} = 0$$

S3 Application of Decision Helper Workflow: Canada Express Entry

The empirical application of our proposed decision helper workflow analyzed data from Canada's Express Entry system. While we provide an overview of this method in the body of our paper, here we provide additional methodological details on how we implement the **Model/Predictions** and **Preference Constraint** portion of our workflow.

S3.1 Data Sources

We merged three datasets at the Federal Research Data Center in Ottawa to conduct our analysis:

- IMDB Integrated Permanent and Non-permanent Resident File (PNRF) 1980-2018 (2019 release)
- IMDB Tax Year Files (t1ff) 2013-2017 (2019 release)
- Express Entry File (2018 release, case level data)

[6] Note that excluding those with job offers from the training data set would have meant excluding a significant proportion of immigrants who came through Express Entry in the first 2 years (2015-2016). However, this limitation becomes less salient as the share of admissions with job offers has declined considerably in recent years with the reduction in number of points for arranged employment in the CRS. For example, in 2017-2019 only about 10% of invited candidates had a job offer or arranged employment.

In addition, we leveraged several additional datasets to provide supplementary information on population levels, unemployment rates, and price indices by geographic region:

- Canada Labour Force Survey (LFS): A monthly survey providing data on the labour market, including estimates of employment and demographics of the working population. Estimates are available at different levels of geographic aggregation, including Economic Region.
- Canadian Rental Housing Index: Public index compiled by the BC Non-Profit Housing Association, based on the 2016 Census. The index reports the average rental price for a single-family apartment, by Census Subdivision (CSD).

A full list of considered variables is found in Table S1. Note that all analyses was conducted in the Federal Research Data Centre in Ottawa and all data output presented here was approved for release.

S3.1.1 Administrative Data

Although the population of interest consists of immigrants entering through Canada's Express Entry system, limited data on this relatively new initiative required us to expand our training data to similar economic immigrants entering Canada. Specifically, in addition to including all Express Entry clients entering between 2015 and 2016 who filed a tax return, we expand our training set to include Non-Express Entry clients who entered between 2012 and 2016 and filed a tax return under four admission categories that would be managed by the Express Entry system: Federal Skilled Workers (A1111), Skilled Trades (A1120), Canadian Experience (A1130), and Provincial Nominees (A1300).

We restrict our training set by removing:

- Individuals who were selected by Quebec or landed in Quebec, given that Express Entry does not apply to this province.
- Individuals whose yearly income in the year after arrival exceeded the 99th percentile, to avoid overfitting to outliers.
- Accompanying children (LANDING_AGE > 18)
- Individuals who died in the year of arrival or in the following year

This set of training data corresponds to matrix $\mathbf{A}$ in the decision helper helper workflow.

S3.1.2 Client Data

The client dataset consists of the Express Entry cohort who entered prior to 2017 (the final year available in the outcome data), along with their associated characteristics from the PNRF file. This represents the group of economic immigrants we consider in all our simulation results. A set of descriptive statistics for this subgroup is found in Table S2.

This set of prediction data corresponds to **C** in the general decision helper helper workflow.

S3.2 Modeling Decisions

Our workflow allows for a wide variety of potential models to be used in predicting outcomes. In this section, we describe the particular modeling decisions we made in the context of analyzing Canadian immigration data.

S3.2.1 Models

We used a supervised machine learning framework to fit and train our models. We use this class of models due to their ability to both flexibly fit the training data while retaining high out-of-sample accuracy with proper model tuning. While any number of supervised machine learning methods might be applicable, we chose to use gradient boosting machines due to their ability to automatically engage in feature selection and discover complex interactions between covariates.

We implement the modeling stage on a location-by-location basis. Specifically, for each economic region, we first subset the training data to those individuals who originally landed in that location, and fit the supervised learning model using individuals' background characteristics to predict their employment earnings. We model synergies using stochastic gradient boosted trees, which we run with a customized script using the gradient boosting machine (gbm) package within R.

In our implementation of gradient boosted trees, we used 10-fold cross-validation within the training data to select tuning parameter values, including the interaction depth (the maximum nodes per tree), bag fraction (the proportion of the training set considered at each tree expansion), learning rate (the size of each incremental step in the algorithm), and number of boosting iterations (number of trees considered).

To determine the best model, we first fix an interaction depth, bag fraction, and learning rate. For this fixed set of parameters, we then fit models over a sequence of boosting iterations (normally, 1 to 1,000 trees). For each model, we calculate the cross-validation root mean square error (RMSE), and choose the model which minimizes this error. To avoid potentially choosing a local minimum, if the best model is within 100 trees of the maximum number of trees we consider, we re-run this process by increasing the maxi-

mum considered trees by 500. We repeat this process as many times as necessary, and record the final tree count and RMSE for the fixed interaction depth, bag fraction, and learning rate.

We repeat the above process tuning over different values of interaction depth, bag fraction, and learning rate. Finally, we pick the model with the set of parameters with the lowest cross-validation RMSE for each separate location model. The set of parameters we consider are:

- Interaction Depth: 5-7
- Learning Rate: .1 and .01
- Bag Fraction: .5-.8 by .15

The set of final models, one for each location, correspond to the set of **L** models in the general workflow.

In order to investigate which covariates are the most predictive of income, we calculate a variable importance measure for each predictor in every separate tuned location model (see summary.gbm in the gbm package for details on how this statistic is calculated). We present these variable importance measures in Figure S1.

This figure demonstrates one of the advantages of fitting each location model separately – in each model, the importance measures of each covariate differs, demonstrating how the set of characteristics that lead to better or worse economic outcomes vary between ERs. Some overall trends emerge, with occupation and citizenship in the top most influential covariates in every model. Whether or not a client had a previous temporary residence permit is also a highly influential predictor in certain Economic regions, especially in Toronto. We further note certain variables have little influence on predicted income across each model, such as the language (French and English) indicators and landing year.

S3.2.2 Predicted Outcomes

We then apply each fitted model to the prediction set to estimate the income for new Express Entry clients should they select the economic region in question. For the prediction set, we remove any Express Entry client with the provincial nomination (A1300) category, as these clients do not have flexibility in choosing the initial landing location. This process is performed separately and independently for each location, which yields a vector of predicted income across possible economic regions for each individual within the prediction set. The final result is a matrix of predicted annual income with rows representing individual Express Entry clients and columns representing economic regions,

corresponding to **M** in the general workflow.

In order to asses model fit, we compare predicted income within the principal applicants' actual location to their observed income in that location in the top panel of Figure S2. The bottom two panels show the histogram of predicted incomes and actual incomes respectively. Overall, predictions are well calibrated, albeit slightly more conservative than observed income at the tails of the distribution.

In our implementation, the cross-validated R^2 for the tuned PA model is 0.54. Within the context of incomes – which are highly skewed and difficult to predict – this is relatively high. This represents a substantial improvement over the R^2 for an analogous linear regression model using cross-validation (0.34). The RMSE for the tuned model is 29,486, as compared to an observed mean income of 58,000 and a standard deviation of 41,600.

S3.3 Approximating Locational Preferences

While a prospective use case would use a survey restrict locations to a set that align with an individual's underlying preferences, we were unable to engage in this exercise in our backtests. Therefore, we use the administrative data and existing migration patterns to estimate preferences. We explain the details of this methodology in this section.

S3.3.1 Underlying Preferences

Upon entering Canada, clients plausibly have a series of idiosyncratic locational preferences related to geographic location, climate, local demographics, labor markets, and other potential factors. We approximated individual locational preferences by examining how Express Entry clients with different background characteristics varied in terms of their original landing locations. Using the landing Economic Region as the dependent variable, we fit a multinomial logit model on Express Entry clients and a random subset of 20% of the non-Express Entry economic immigrants. Given that these preferences are proxies, we use a coarse set of covariates including education, birth region, age, immigration category, case size, and indicators for work and study permits. These predictions approximate the vector of utilities **u** in the workflow.

S3.3.2 Restricted Set of Locations

After obtaining predictions for each Express Entry case, we rank order locations by predicted preferences, randomly breaking ties. The resulting ranks are used in conjunction with the parameter ϕ (the number of acceptable locations we consider) to determine the initial set of locations considered when selecting the locations with the top optimal income. This set of ϕ acceptable locations represents subset $\mathbf{S_i}$ in the workflow.

S4 Robustness Checks

Below we discuss several robustness checks to our study. In particular, we outline the impact of our backtest when considering 1) cost-of-living adjustments, 2) maximizing principal applicant plus spouse income, and 3) alternative simulation specifications.

S4.1 Cost-of-living Adjustments

An important consideration influencing relative quality of life across landing locations relates to living costs. To take this factor into account during the recommendation process, we ran alternate simulations where we define outcomes as total income less estimated yearly rental costs of a two bedroom apartment. To our knowledge, rental prices are the most granular cost index currently available across small geographic regions. The estimates we pull are from the Canadian Rental Housing Index, a public index compiled by the BC Non-Profit Housing Association and based on the 2016 Census. The index reports the average rental price for a single-family apartment, by Census Subdivision (CSD).

In Figure S3, we replicate the results in our main paper, demonstrating average gains in employment under various simulation parameters in the top panel and visualizing expected movement patterns in the bottom panel.

S4.2 Principal Applicant and Spouse Model

While our main paper reports our findings for principal applicants only, we additionally fit a set of models that consider both principal applicants and their accompanying partners. As a simplifying assumption, we assume that both individuals in a case have similar (joint) locational preferences, and derive these preferences from a PA-only model. The income predictions, however, take into account the joint income of the PA and partner divided by the number of adults in the family unit (average family income). We then estimate the models assuming a family unit will move jointly. In Figure S4, we replicate the results in the main paper using this approach, with similar results.

S4.3 Alternative Simulation Specifications

Our simulations vary two parameters: the number of acceptable locations considered (ϕ) and the compliance rate (π). In this section, we consider the impact of further varying these parameters on our core results.

S4.3.1 No Locational Preferences

In our main analysis, we infer regional preferences by analyzing existing residential patterns and then using these estimated preferences to restrict the choice set in our simu-

lations. However, expected income gains are maximized when no locational preferences are taken into account. In Figure S5, we show simulated movement patterns with no locational preferences under different compliance rates. Relative to the case presented in the main paper, the results similarly suggest that the majority of outflow is from the four largest locations, but display increased recommendations to smaller locations.

S4.3.2 Constant Compliance Rate

In the body of the paper, we present results where we vary the compliance rate π as a function of income. Specifically, we specify an upper bound π_{max} to the individuals with the lowest actual income before linearly interpolating to the value $\pi = 0$ across the entire income distribution. Thus, each individual receives a heterogeneous compliance parameter π_i, and the average compliance rate in a particular simulation run is reported as $\pi_{max}/2$.

To ensure that results are not driven by this modeling decision, we repeat our simulations with a constant compliance rate in an individual simulation run. In each of these tests, we set a single π that represents each individuals' likelihood of complying, which is constant across income. We present these results in Figure S6, which reveals very similar potential average gains in income across each simulation.

S4.3.3 Removing Economic Regions

In order to evaluate whether the gains we report in the main analysis of our paper are being driven by a specific subset of ERs, we rerun the simulations exactly as described in the body of the paper but remove from consideration certain subsets of landing locations. That is, if an individual 'complies' with probability π_i, we limit the set of locations they can potentially move to in the simulation.

We begin by specifying three alternative models, in each case excluding a subset of ERs that could potentially drive our results. In the first alternative model, we do not allow individuals to move to the largest ERs, those with a population greater than 1,500,000 according to the 2016 census. In the second, we extend this to include large and growing ERs, defined as all ERs with a population greater than 1,000,000 and a growth rate in population from the 2011 to 2016 census above 5%. In the third, we remove the smallest ERs – those with a population less than 100,000 in the 2016 census. A full list of removed ERs in each specification is listed in the Table S3 below.

The results of these alternative model runs are found in Figure S7. Overall, these plots demonstrate the gains we find in our main analysis are not driven by one of the ER subsets we define above. In the top-left panel, we presents results from a simulation considering "All ERs," effectively replicating our main results in the body of our paper. In each of the three alternative specifications, we see that average gains across each compliance and preference parameter do not substantially differ from this baseline.

Another way we check against the impact of a single ER on our core results is by running a "leave-one-out" robustness check. In this test, we run a series of 52 simulations, in each case dropping one of the 52 total ERs from consideration. Other than removing this single ER from the choice set, the simulations are run exactly as described in the body of the paper. For simulation, we calculate the mean gain in annual income. We present the average of these gains and the 95% confidence interval across the 52 simulations in Figure S8. We again see little change to our core results, demonstrating no single ER drives the average gain in income in our simulations.

S5 Figures

Figure S1: Variable Importance Statistics for Tuned Location Models (Principal Applicants)

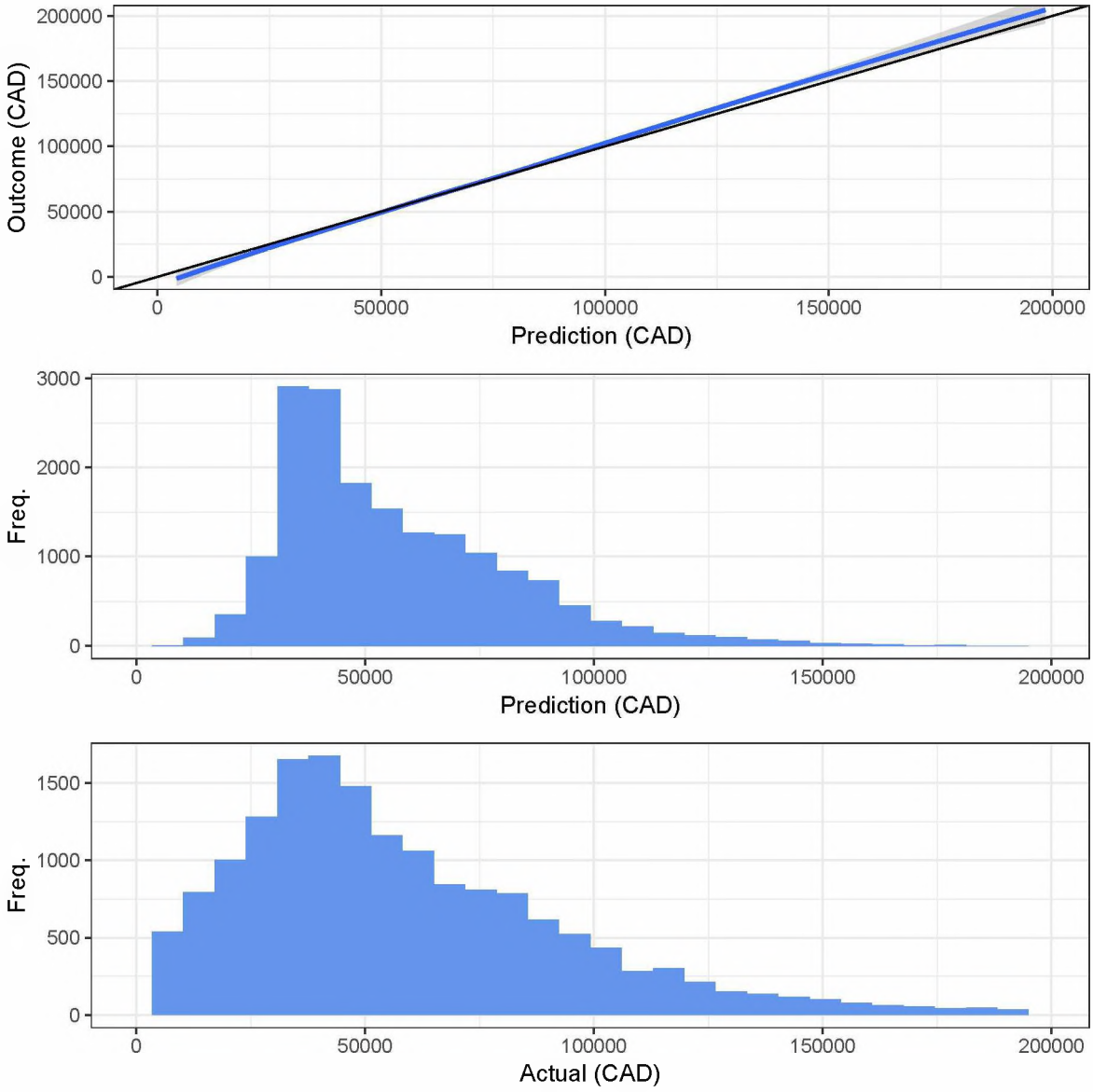


Figure S2: Calibration Plot: Model Fit for Express Entry Clients

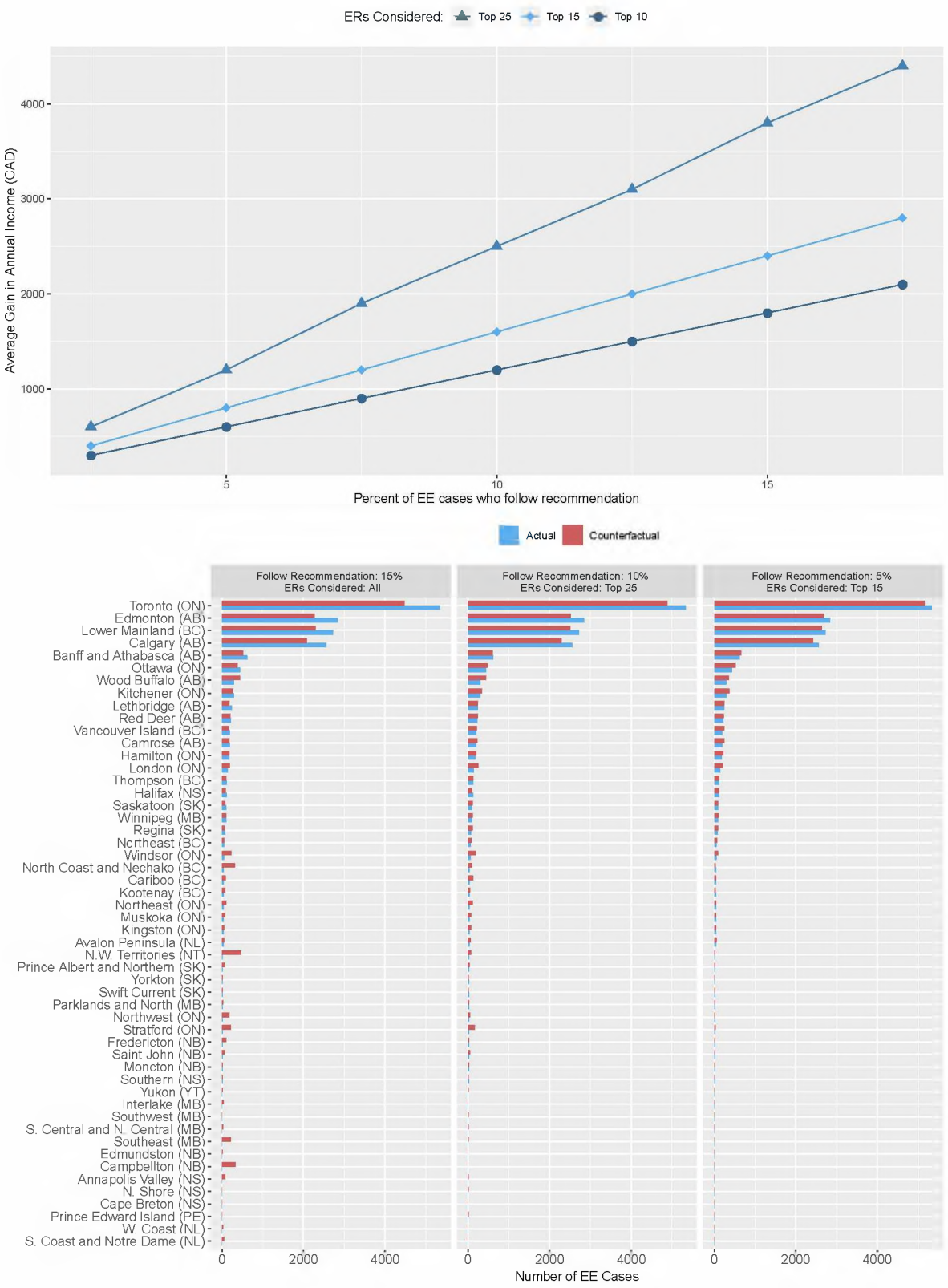


Figure S3: Estimated Average Income Gains and Shifts in Arrival Locations with CPI Adjustments

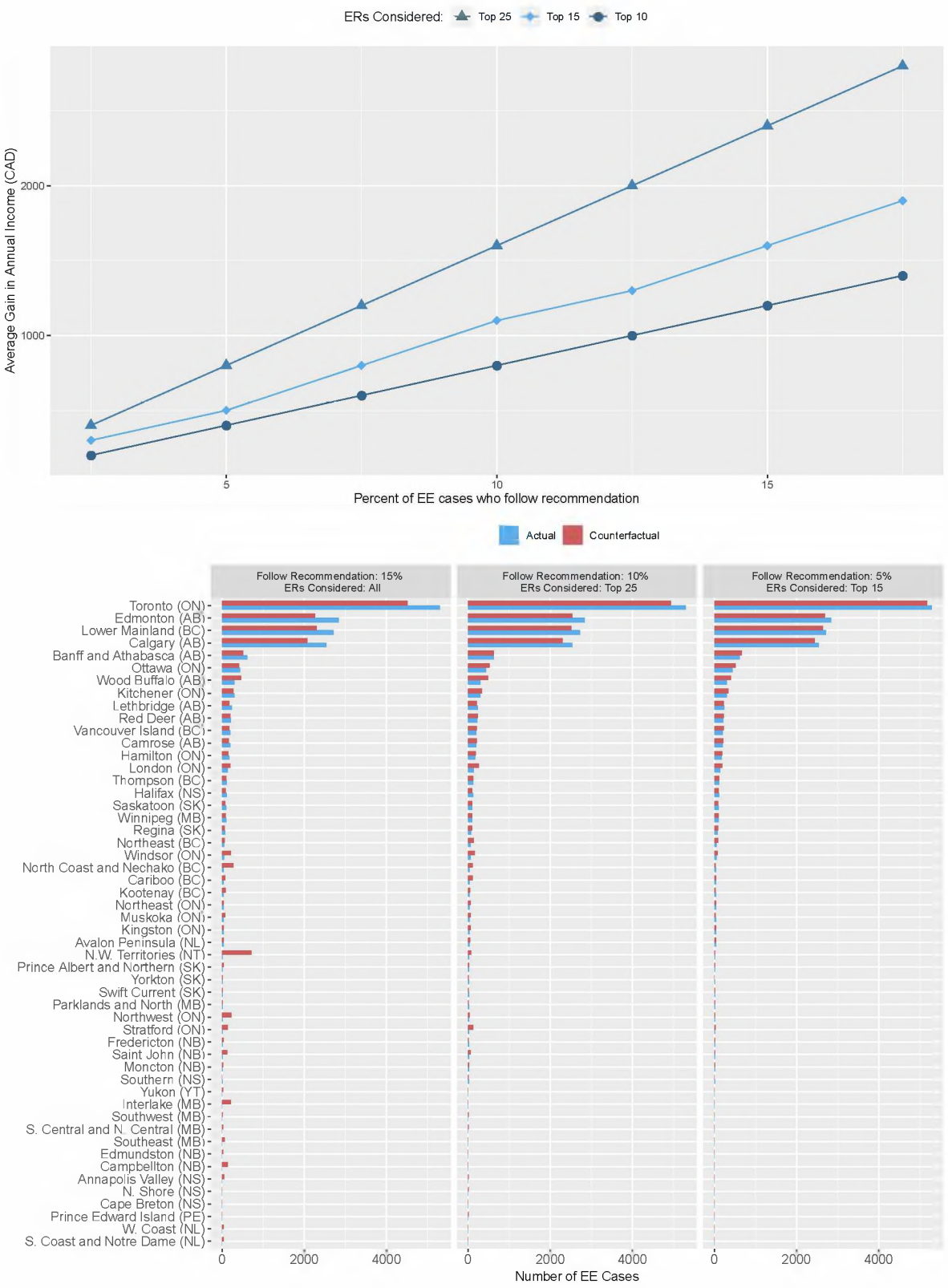


Figure S4: Estimated Average Income Gains and Shifts in Arrival Locations for Principal Applicant and Spouse Model

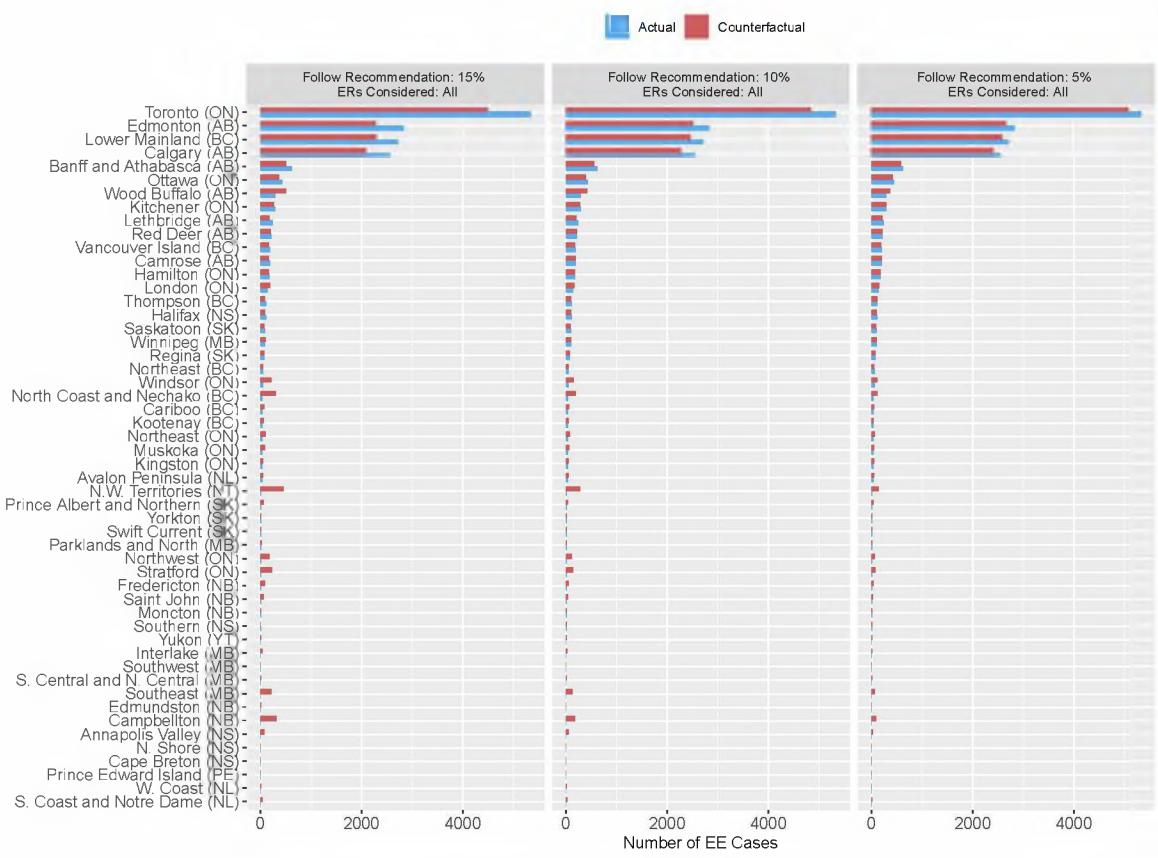


Figure S5: Movement Under Various Simulation Parameters

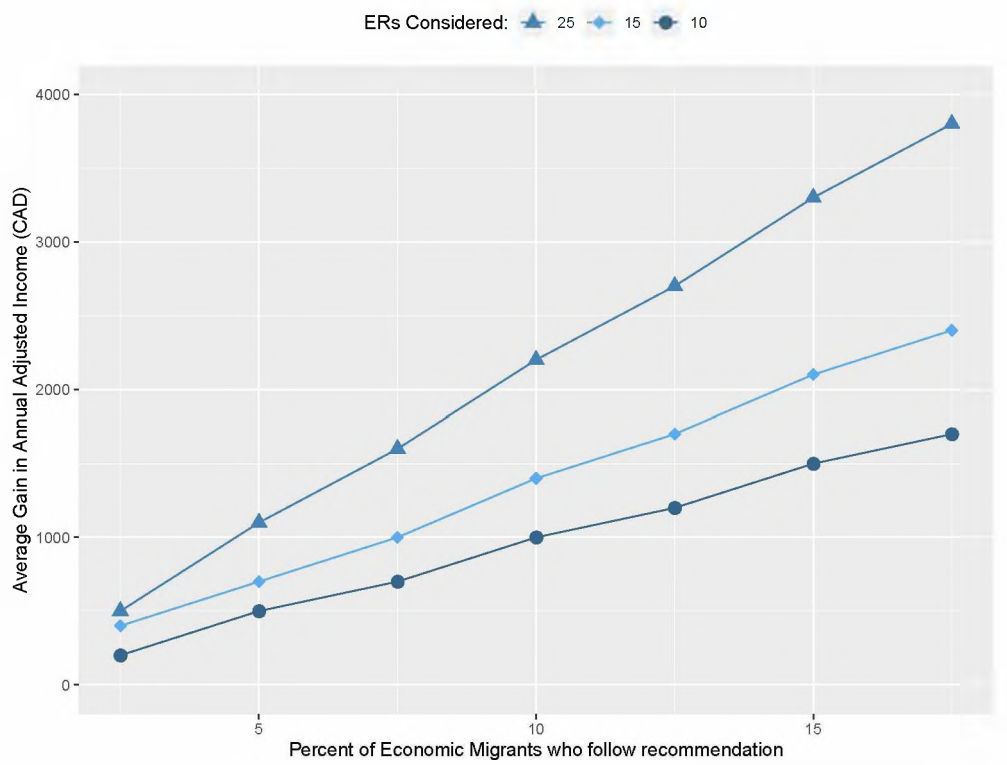


Figure S6: Constant Compliance Rate: Estimated Average Income Gains. N=17,640

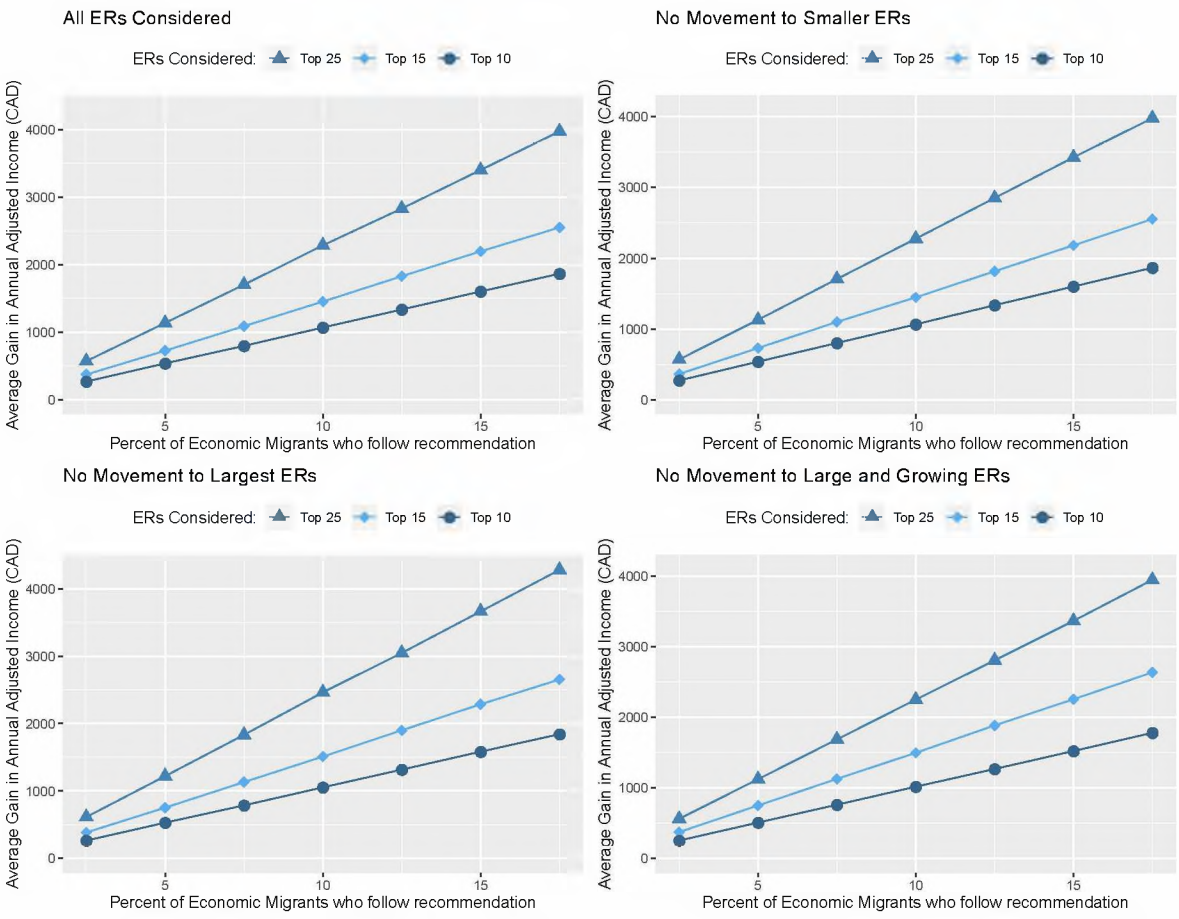


Figure S7: Removing Subsets of ERs. N=17,640

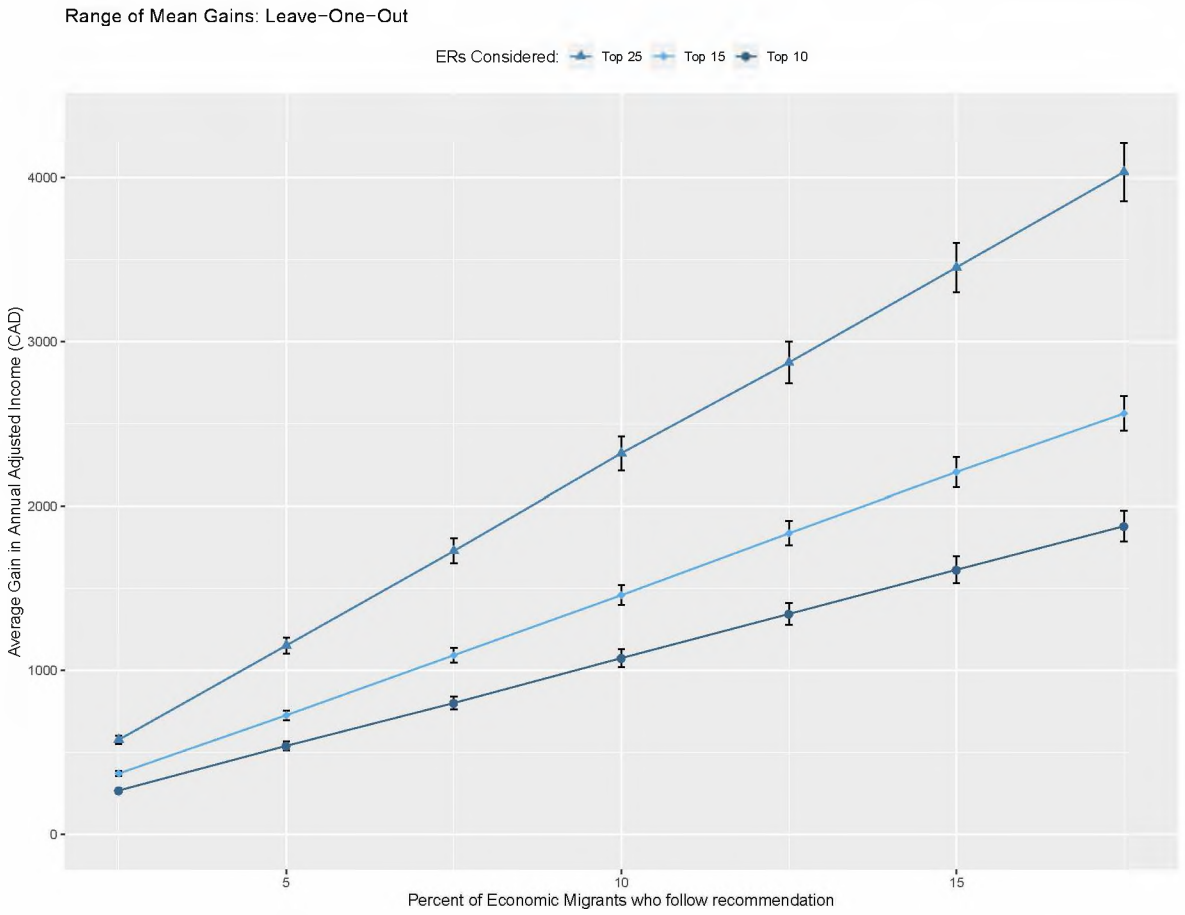


Figure S8: Average Gains Across 52 Leave-One-Out Simulations. N=17,640

S6 Tables

Original Variable Name	Source	Description
COUNTRY_BIRTH[1,4]	PNRF	
COUNTRY_CITIZENSHIP[1,4]	PNRF	
CSQ_IND[2]	PNRF	Quebec program indicator
DEATH_INDICATOR[2]	PNRF	
DESTINATION_ER[1,2,3]	PNRF	Intended Economic Region (ER) of landing
EDUCATION_QUALIFICATION[1]	PNRF	
EXPRESS_ENTRY_IND[1,2]	PNRF	Express Entry flag
FAMILY_STATUS[1,2]	PNRF	
GENDER[1]	PNRF	
IMMIGRATION_CATEGORY_CENSUS[1,2]	PNRF	Admission category
LANDING_AGE[1]	PNRF	
LANDING_MONTH[1]	PNRF	
LANDING_YEAR[1,2]	PNRF	
LEVEL_OF_EDUCATION[1]	PNRF	
NOC3_CD11[1]	PNRF	3-digit expected occupation code
NUMBER_STUDY_PERMITS[1]	PNRF	
NUMBER_WORK_PERMITS[1]	PNRF	
OFFICIAL_LANGUAGE[1]	PNRF	
SKILL_LEVEL_CD11[1]	PNRF	
EI___I[1]	Tax	Individual employment income (excluding self-employment)
PREFILER_IND[1]	Tax	Whether an individual filed a return on a TR permit
POPULATION_ER[1]	LFS	Quarterly population
PRICE_INDEX[3]	External	Average yearly rental price
UNEMPLOYMENT_ER[1]	LFS	Quarterly unemployment

[1] = Variable used to train machine learning models
[2] = Variable used to subset the data
[3] = Variable used to adjust final predictions

[4] = Variable aggregated to the continent level for modeling

Table S1: Variable Names

	mean	sd	min	max		mean	sd	min	max
Annual Income per Head (CAD)	49900.00	33600.00	0	246600	English: No	0.02	0.14	0	1
Age	33.15	5.99	22	65	English: Yes	0.98	0.14	0	1
Birth Region: The Americas	0.09	0.29	0	1	French: No	0.97	0.16	0	1
Birth Region: Europe	0.24	0.43	0	1	French: Yes	0.03	0.16	0	1
Birth Region: Africa	0.06	0.23	0	1	Prefiler: No	0.13	0.34	0	1
Birth Region: Asia	0.59	0.49	0	1	Prefiler: Yes	0.87	0.34	0	1
Birth Region: Oceania	0.02	0.14	0	1	TR: No TR	0.12	0.32	0	1
Citizenship: United States	0.03	0.17	0	1	TR: Study	0.01	0.09	0	1
Citizenship: Mexico	0.02	0.13	0	1	TR: Study+Work	0.31	0.46	0	1
Citizenship: Jamaica	0.01	0.11	0	1	TR: Work	0.57	0.5	0	1
Citizenship: Brazil	0.01	0.11	0	1					
Citizenship: France	0.03	0.16	0	1					
Citizenship: Germany	0.01	0.11	0	1					
Citizenship: Poland	0.01	0.10	0	1					
Citizenship: Russia	0.01	0.10	0	1					
Citizenship: Ukraine	0.01	0.12	0	1					
Citizenship: Ireland	0.05	0.21	0	1					
Citizenship: United Kingdom	0.06	0.24	0	1					
Citizenship: Nigeria	0.02	0.15	0	1					
Citizenship: South Africa	0.01	0.10	0	1					
Citizenship: Iran	0.01	0.09	0	1					
Citizenship: China	0.04	0.20	0	1					
Citizenship: South Korea	0.03	0.16	0	1					
Citizenship: Philippines	0.14	0.35	0	1					
Citizenship: India	0.28	0.45	0	1					
Citizenship: Pakistan	0.02	0.12	0	1					
Citizenship: Australia	0.02	0.14	0	1					
Citizenship: Other	0.18	0.38	0	1					
Education: Less than BA	0.41	0.49	0	1					
Education: BA	0.27	0.44	0	1					
Education: MA	0.28	0.45	0	1					
Education: PhD	0.03	0.18	0	1					
Male	0.67	0.47	0	1					
Female	0.33	0.47	0	1					
Unit Size	1.43	0.50	1	3					
Landing Year: 2015	0.29	0.45	0	1					
Landing Year: 2016	0.71	0.45	0	1					
Landing Month (1-12)	7.40	3.40	1	12					
Category: Skilled Worker program	0.46	0.50	0	1					
Category: Skilled Trades program	0.09	0.29	0	1					
Category: Canadian Experience Class	0.45	0.50	0	1					
Industry: ArtCultureSport	0.04	0.20	0	1					
Industry: Computer	0.20	0.40	0	1					
Industry: Education_Law_Govt	0.07	0.25	0	1					
Industry: Extraction	0.00	0.03	0	1					
Industry: Finance	0.09	0.29	0	1					
Industry: FoodTourism	0.12	0.33	0	1					
Industry: Health	0.04	0.21	0	1					
Industry: Management_Misc	0.03	0.16	0	1					
Industry: ManualTrades	0.10	0.31	0	1					
Industry: Manufacturing	0.01	0.08	0	1					
Industry: NatResources_9080	0.01	0.10	0	1					
Industry: No_Info	0.01	0.08	0	1					
Industry: Sales	0.05	0.22	0	1					
Industry: Services	0.18	0.38	0	1					
Industry: SocialServices	0.01	0.12	0	1					
Industry: Technical	0.03	0.18	0	1					
Skill Level: Managerial	0.10	0.30	0	1					
Skill Level: Professionals	0.35	0.48	0	1					
Skill Level: Skilled and Technical	0.55	0.50	0	1					

Table S2: Descriptive Statistics: Express Entry Principal Applicants

ER_CODE	Name	Large	Rapidly Growing	Small
3530	Toronto	X	X	
5920	Lower Mainland	X	X	
4830	Calgary		X	
4860	Edmonton		X	
3540	Kitchener		X	
4660	Interlake			X
4680	Parklands and North			X
4840	Banff and Athabasca			X
4740	Yorkton			X
1350	Edmundston			X
5980	N.E. (B.C.)			X
4620	S. Central and N. Central			X
5960	North Coast and Nechako			X
4640	S. Central and N. Central			X
6110	N.W. Territories			X
4670	Parklands and North			X
5970	North Coast and Nechako			X
4760	Prince Albert and Northern			X
1020	S. Coast and Notre Dame			X
6010	Yukon			X

Table S3: Removing Economic Regions Robustness Check

Analyse mogelijke risico's inputdata Kansrijke Koppeling

1. Inleiding

Het Centraal Orgaan opvang Asielzoekers (COA) heeft het CBS gevraagd te helpen met het identificeren van mogelijke risico's van het toepassen van GeoMatch in Nederland. GeoMatch is een aanbevelingsinstrument voor de optimale verdeling van statushouders over arbeidsmarktregio's, waarbij baankansen worden geschat met machine learning. Het CBS richt zich hierbij vooral op de kwaliteit van de inputdata, die deels door het CBS zijn geleverd, terwijl Deloitte zich zal richten op het model en de juridische/ethische aspecten. In deze nota beschrijven we eerst kort de opzet van het model (GeoMatch) en trainingsdata. Daarna beschrijven we onze onderzoeksmethodiek en geven een aantal aandachtspunten aan wat betreft datakwaliteit.

2. GeoMatch en trainingsdata

GeoMatch is een aanbevelingsinstrument voor de optimale verdeling van statushouders over arbeidsmarktregio's (*labor market regions*, LMR's). De optimale allocatie maximaliseert de gemiddelde kans dat een statushouder werk vindt, gegeven een maximum aantal statushouders per regio. Om dit optimalisatieprobleem op te lossen worden baankansen per regio geschat met een supervised machine learning algoritme. Per regio is een gradient boosting algoritme getraind op historische data om de relaties te leren tussen kenmerken van statushouders en het vinden van werk in het eerste jaar dat statushouders woonachtig zijn in de betreffende regio. GeoMatch is ontwikkeld door het Immigration Policy Lab (IPL) aan o.a. Stanford University en ETH Zürich (Bansak et al. 2018). Het is uitdrukkelijk bedoeld als een hulpmiddel voor medewerkers die uiteindelijk beslissen aan welke arbeidsmarktregio statushouders worden toegewezen. In Zwitserland wordt een pilot uitgevoerd met de tool. In Nederland is de tool nog niet getest in het veld.

Voor het COA is GeoMatch getraind en geoptimaliseerd op Nederlandse data. Hierbij zijn twee databronnen gebruikt: COA's register met achtergrondkenmerken, procedurele informatie en locatiegegevens van alle Nederlandse statushouders (vanaf 2014), en het Asielcohortenbestand van het CBS met informatie over (onder andere) werk en opleiding van de statushouders in Nederland vanaf 2014.

Voor elk van de 35 arbeidsmarktregio's die COA hanteert is een algoritme getraind om de kans te voorspellen op betaald werk in ten minste één maand in het jaar na verhuizing. Elk van deze modellen werkt met eigen trainingsdata: namelijk de statushouders die door COA aan die regio zijn toegewezen.

Het gebruik van algoritmes kan een waardevolle aanvulling zijn om evidence-based beslissingen te maken. Omdat het algoritme gebruikt wordt om beslissingen over individuen te ondersteunen, is het extra belangrijk om de data en het algoritme te begrijpen. Daarom is aan het CBS gevraagd om mee te kijken met de samenstelling van de gebruikte dataset. Daarbij beoordelen wij niet of de data van voldoende kwaliteit is, maar wijzen wij enkel aandachtspunten aan om rekening mee te houden bij eventuele toepassing van GeoMatch in de praktijk.

3. Onderzoeksmethodiek

Voor het onderzoek zijn verschillende documenten bestudeerd:

- Beschikbare documentatie over het model, met name het onderzoeksrapport opgesteld door IPL (Bansak et al, 2021).
- R-scripts van de voorbewerking van de data: controleren, bewerken en samenvoegen van de bronbestanden tot het uiteindelijke inputbestand voor het model. Deze scripts zijn aangeleverd door IPL.
- Het gebruikte inputbestand van het model, aangeleverd door IPL.
- Ten slotte hebben we van IPL schriftelijke antwoorden ontvangen op onze aanvullende vragen over de data.

Daarnaast woonden we een informatiesessie bij, georganiseerd door COA, waarin het huidige toewijzingsproces werd besproken. Hierin kwam ook de dataverzameling voor het model en de inpassing van het model in het toewijzingsproces aan bod.

Tijdens het bestuderen van bovenstaande informatie hebben we het Toetsingskader algoritmes van de Algemene Rekenkamer gebruikt om richting te bepalen. Aangezien wij niet beoordelen, maar alleen aandachtspunten constateren, hebben we het kader niet als een checklist gebruikt en zijn de resultaten ook zo niet gepresenteerd. Wel presenteren we de aandachtspunten en risico's die we zijn tegengekomen. Deze zijn opgenomen in het volgende hoofdstuk.

Het CBS heeft geen analyses op de data uitgevoerd. Wanneer wij een analyse relevant vinden, hebben we deze als aanbeveling opgenomen in het volgende hoofdstuk.

4. Datakwaliteit

Bij het beoordelen van datakwaliteit worden mogelijke risico's voor datakwaliteit hieronder gegroepeerd onder de thema's Representatie en Meting. Dit zijn terugkerende thema's in generieke raamwerken (Groves and Lyberg 2010 voor enquêtes; Zhang 2012 en De Waal et al. 2019 voor statistieken op basis van meerdere bronnen). Risico's voor representatie ontstaan als niet alle eenheden en hun kenmerken in de doelpopulatie worden waargenomen (of doordat eenheden

worden waargenomen die niet tot de doelpopulatie behoren). Risico's voor meting ontstaan als gegevens over een eenheid onjuist of onbetrouwbaar zijn.

4.1 Representatie

Representatie gaat over vertegenwoordiging van de totale populatie in de beschikbare data. Als er bepaalde groepen ondervertegenwoordigd zijn in de data, of helemaal missen, maakt het model geen betrouwbare voorspellingen voor deze groep. Het is daarom belangrijk om bewust te zijn van de representativiteit van de beschikbare data. In principe wordt in het COA-model met integrale data gewerkt. De IBIS-data is een dataset met daarin alle asielzoekers die door COA werden opgevangen; de asielcohortendata bevat alle statushouders die sinds 2014 in Nederland een verblijfsvergunning asiel toegekend kregen. In de basis zouden we daarom geen problemen verwachten voor de representativiteit. De modellen worden echter getraind per regio, en in elke regio zal de samenstelling afwijken van de samenstelling van de gehele populatie statushouders. In afzonderlijke regio's kunnen er daarom toch risico's voor de representativiteit ontstaan. Hieronder beschrijven we enkele aandachtspunten over de representativiteit.

Ontbrekende doelvariabelen:
Er zijn statushouders waarvan het label (de doelvariabele wel/geen baan) ontbreekt en die dus ontbreken in de trainingsdata. Dit is een mogelijke bron van vertekening als het ontbreken van een label correleert met een van de verklarende variabelen van het model, en mensen met bepaalde kenmerken daardoor niet meegenomen worden in de trainingsset.
Voorgestelde analyse: Vergelijk de samenstelling van de doelpopulaties met en zonder doelvariabele.

Ontbrekende kenmerken:
Het gebruikte algoritme (gradient boosting) is gebaseerd op beslisbomen. Dit soort algoritmen zijn robuust tegen ontbrekende waardes op kenmerken. Eenheden met een ontbrekende waarde op een kenmerk waarop gesplitst wordt, worden toegewezen aan het blad dat het beste past bij het doel (minimaliseren van de voorspelfout). Er is echter een set kenmerken die voor alle personen in de trainingsdata ontbreken. Deze kenmerken zijn verkregen uit gesprekken met statushouders, maar zijn niet opgenomen in de dataset, bijvoorbeeld het netwerk en de ambities van de statushouder. COA-medewerkers nemen deze mee in de afweging over de best passende arbeidsmarktregio als aanvulling op de andere kenmerken die wel in de dataset zijn meegenomen, zoals opleiding of werkgeschiedenis. Ten opzichte van de COA-medewerker, mist de inputdata voor het model dus informatie om de baankansen te voorspellen.
Stel dat COA-medewerkers mensen met kenmerk X (bijvoorbeeld ambitie) toewijzen aan regio A, omdat daar goede baankansen zijn voor mensen met kenmerk X. Als kenmerk X niet in de dataset is opgenomen, en het algoritme daarom kenmerk X niet meeneemt, voorspelt het algoritme ten onrechte dat iedereen goede/slechte baankansen heeft in regio A.
Aanbeveling: Draag zorg dat alle werknemers die met het model werken zich bewust zijn dat de voorspelde kansen zijn gemaakt op basis van alleen de kenmerken die in de trainingsdata zijn opgenomen, en dat ze hun toewijzing baseren op de combinatie van voorspelde modelkansen en de overige kenmerken (bijvoorbeeld netwerk en ambitie).

Regio-afhankelijke samenstelling:
De samenstelling van de trainingsdata per arbeidsregio kan afwijken van de samenstelling van de gehele populatie statushouders (door de selectieve allocatie) en van de samenstelling van toekomstige cohorten (denk aan vluchtelingen uit andere landen als bijvoorbeeld nieuwe conflicten

ontstaan). Elk regionaal model zal zwaarder leunen op de relatief veel voorkomende groepen in die regio (een regio heeft mogelijk vooral statushouders met een bepaalde nationaliteit of werksector). Dit leidt wellicht tot vertekening van de voorspellingen van relatief zeldzame groepen in deze regio. Daarnaast kan het zo zijn dat bepaalde groepen helemaal niet voorkomen in bepaalde regio's. Het model kan dan geen inschatting maken van het succes van mensen met kenmerk X in regio A. Dit maakt de vergelijkbaarheid tussen regio's, en de afweging van het model over in welke regio's de statushouder de hoogste baankans heeft, moeilijk.
Voorstel: dit representativiteitsrisico is inherent aan de regionale opzet van het model. Dit risico kan gemitigeerd worden door het creëren van bewustzijn bij medewerkers, of er rekening mee te houden in de modelopzet. Om de disbalans in de regionale trainingsdata te balanceren zou het model bijvoorbeeld getraind kunnen worden met grotere gewichten voor relatief zeldzame groepen en lagere gewichten voor relatief veel voorkomende groepen. Daarvoor is wel inzicht nodig in hoeverre de samenstelling per regio afwijkt van de gehele populatie statushouders (alle regio's samen), en hoe de samenstelling in de regio's verschilt ten opzichte van elkaar. Een procentuele vergelijking van welke subgroepen in welke regio's wonen geeft al een eerste beeld van de mate van balans/disbalans. Ook worden zo subgroepen in beeld gebracht die in bepaalde regio's niet of amper voorkomen en waarvoor de voorspellingen van het model minder betrouwbaar zijn. Om inzicht te krijgen in de mate van afwijking tussen regio's, en het risico dat de gemiddeld baankans voor die regio vertekend is, kan gebruikt worden gemaakt van de variatiecoëfficiënt (CV) .[1] Bij de COA-medewerkers kan bewustzijn gecreëerd worden door de wijze waarop zij de output van het model te zien krijgen. Als zij, naast de aanbevolen regio's, ook de geschatte baankans en een betrouwbaarheidsscore per regio te zien krijgen, geeft hen dit waardevolle informatie over de robuustheid van de aanbeveling. Informatie over de robuustheid helpt de COA-medewerker bij de afweging om de aanbeveling van het model te volgen, of om additionele informatie uit gesprekken met de statushouder zwaarder te laten wegen.

4.2 Meting

Risico's in de meting ontstaan bij meetfouten. Als data niet volledig of onjuist zijn, zal het model dat op basis daarvan gebouwd wordt de werkelijkheid ook niet volledig weerspiegelen. Daarom is het van belang om bewust te zijn van wat de data wel en niet correct reflecteert, en hier rekening mee te houden bij interpretatie van resultaten.

Meting doelvariabele, het effect van verhuizingen van statushouders:
Het is mogelijk dat statushouders verhuizen in het eerste jaar na toewijzing aan de arbeidsmarktregio. Dit gegeven wordt in de inputdata van het model niet meegenomen. Daarin wordt namelijk geen onderscheid gemaakt tussen statushouders die het volledige eerste jaar in de arbeidsmarktregio blijven wonen en statushouders die in het eerste jaar al naar een andere arbeidsmarktregio verhuizen. Het zou zo kunnen zijn dat statushouders verhuizen en in de nieuwe arbeidsmarktregio een baan vinden. In de inputdata wordt dit geregistreerd alsof een baan in de

[1] Bij enquêtes wordt de variatiecoëfficiënt (CV) van de responskansen gebruikt als maat voor het risico op vertekening van populatieschattingen (Van Berkel et al. 2020). Bij een enquête zouden mensen van de ene achtergrond vaker kunnen reageren dan mensen met een andere achtergrond (dus een grote variatie in de responskans per subgroep). De enquêteresultaten zijn dan niet representatief voor mensen van alle achtergronden. Hoe kleiner de variatie in de responskansen en hoe groter de gemiddelde responskans, hoe kleiner de vertekening. Individuele responskansen worden geschat door de relatie te modelleren tussen achtergrondvariabelen en de responsindicator. Op vergelijkbare wijze zouden allocatiekansen gemodelleerd kunnen worden uit achtergrondkenmerken en een indicator die aangeeft of een statushouder wel of niet is toegewezen aan de betreffende arbeidsmarktregio. De CV van de allocatiekansen zou gebruikt kunnen worden als maat voor het verschil in samenstelling tussen een arbeidsmarktregio en de gehele populatie statushouders, en als maat voor het risico op vertekening van de gemiddelde baankans in een arbeidsmarktregio.

toegewezen arbeidsmarktregio wordt gevonden. Deze keuze is gemaakt om bias te verminderen: het model wordt niet getraind op factoren die je nog niet kan weten op het moment van toewijzing aan een arbeidsmarktregio (zoals of mensen overlijden of verhuizen uit de toegewezen arbeidsmarktregio). Tegelijkertijd zou deze keuze juist een mogelijke bron van bias kunnen zijn als er arbeidsmarktregio's zijn waar een groot deel van de statushouders snel weer verhuist. De voorspellingen van het model voor deze arbeidsmarktregio kloppen dan niet. Wij raden aan om een extra analyse uit te voeren om te bepalen in hoeverre dit risico van toepassing is.
Aanbevolen analyse: model trainen op statushouders die nog steeds woonachtig zijn in de arbeidsmarktregio *waaraan ze zijn toegewezen en kijken of dit de resultaten verandert, of controleer hoe vaak het voorkomt dat personen naar een andere arbeidsmarktregio verhuizen).*

Meting doelvariabele, het effect van de tijdsperiode:
Augustus 2020 is geen representatief meetmoment voor het meten van baankansen onder statushouders (of alle inwoners van Nederland), vanwege de uitbraak van het coronavirus in Nederland en de bijbehorende maatregelen in maart 2020 en later. Statushouders werken relatief vaak op tijdelijke contracten in sectoren als de horeca (CBS, 2021). Het is aannemelijk dat deze groep relatief hard geraakt is door de maatregelen en op dat moment noodgedwongen minder is gaan werken of werkloos is geworden. Door baankansen in augustus 2020 te meten, worden de baankansen van statushouders dus wellicht onderschat. Ook zou deze onderschatting kunnen verschillen tussen de verschillende arbeidsmarktregio's.
Voorgestelde analyse: verkrijg inzicht in de verschillen tussen augustus 2020 en de periode vóór coronamaatregelen (in hoeverre werken statushouders minder uren, of in andere sectoren). Om aan de hand van deze resultaten te bepalen hoe representatief de resultaten zijn voor een typisch jaar. Indien mogelijk: herhaal de analyse op een periode zonder coronamaatregelen (bijvoorbeeld januari 2019-januari 2020).

Referenties

Algemene Rekenkamer (2020). Toetsingskader algoritmes. Online gepubliceerd: https://www.rekenkamer.nl/onderwerpen/algoritmes/algoritmes-toetsingskader

Bansak, K. J. Ferwerda, J. Hainmueller, A. Dillon, D. Hangartner, D. Lawrence, J. Weinstein (2018). Improving refugee integration through data-driven algorithmic assignment. *Science*, 359, 325–329, doi: 10.1126/science.aao4408.

Bansak, K., J. Fei, N. Adams-Cohen, J. Ferwerda, D. Hangartner and J. Hainmueller (2021). GeoMatch Netherlands. Retrospective impact evaluation and feasibility study for the Central Agency for the Reception of Asylum Seekers (COA). Internal report to COA.

Van Berkel, K, S. van der Doef and B. Schouten (2020). Implementing adaptive survey design with an application to the Dutch Health Survey. *Journal of Official Statistics*, 36(3), 609–629, doi: 10.2478/jos-2020-0031.

CBS (2021). Asiel en integratie 2021. Cohortstudie asielzoekers en statushouders. Den Haag: Centraal Bureau voor de Statistiek.

De Waal, T., A. van Delden and S. Scholtus (2019). Multi-source statistics: Basic situations and methods. *International Statistical Review*, 88(1), 203–228, doi: 10.1111/insr.12352.

Groves, R.M. and L. Lyberg (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, 74 (5), 849–879, doi: 10.1093/poq/nfq065.

Zhang, L.-C. (2012). Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica*, 66(1), 41–63, doi:10.1111/j.1467-9574.2011.00508.x.

Docnr4

GeoMatch™
Connecting people to places

ipl immigration policy lab

IPL <> COA Partnership: GDPR Sensitive Data Limitations Update

May 2022

Centraal Orgaan opvang asielzoekers (COA)

Immigration Policy Lab (IPL)

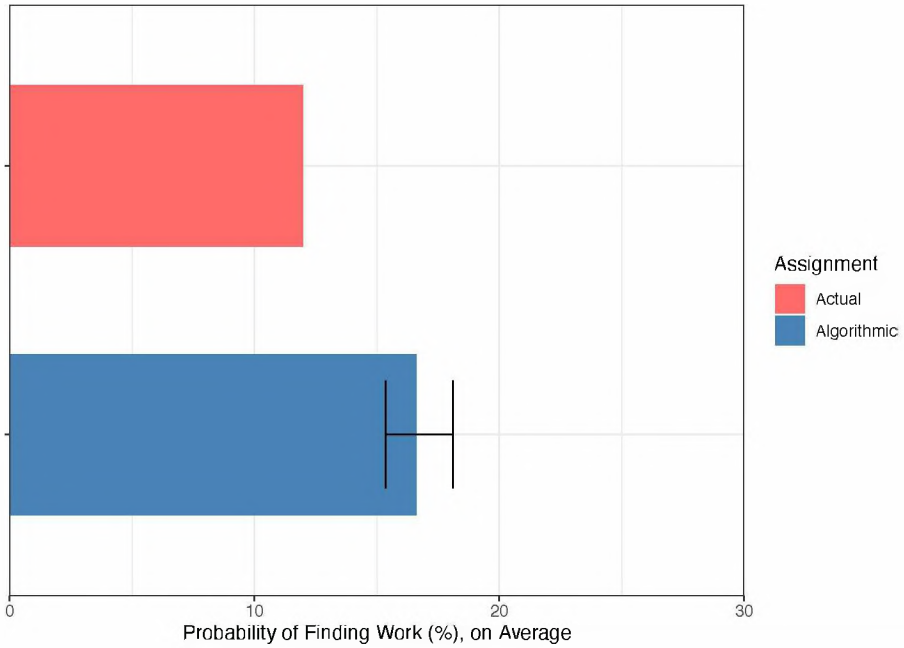
Agenda

Our discussion today

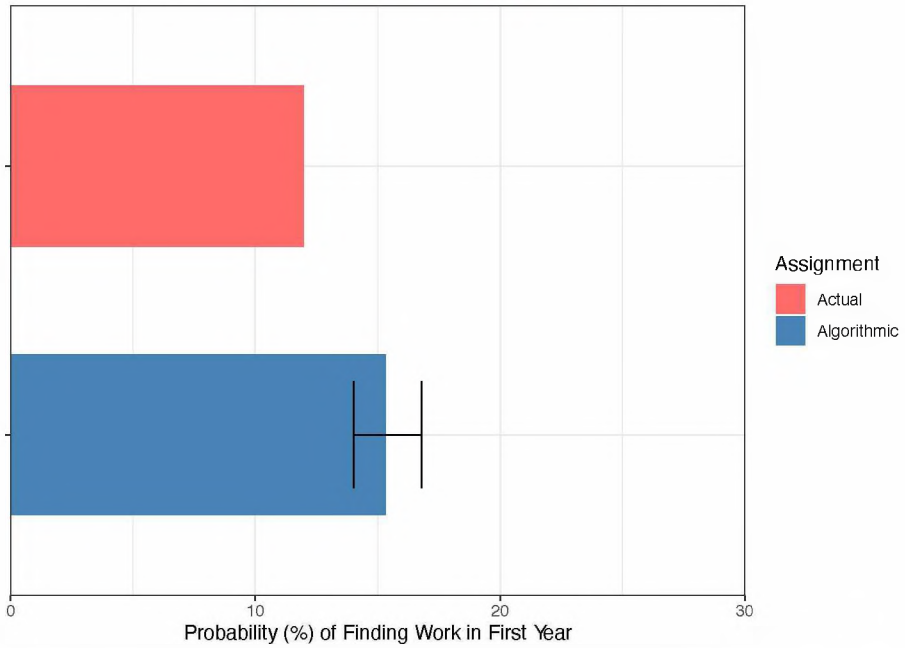
1. **Impact on overall gains**
2. Impact on subgroups
3. Impact on subgroups (nationalities)
4. Implications and discussion

Impact on overall gains

With sensitive variables:
16.61% employment rate predicted (compared to the actual rate of 11.97%)



Without sensitive variables:
15.33% employment rate predicted. A 3.3 percentage points increase compared to the initial 4.67 = **loss of 30% of the gains**



Estimated impact on access to employment*: compared to the status quo, we would now estimate that **200 more status holders** would obtain a job in their first year, compared to **280 more** when using all the variables.

*Based on initial backtest's assumption of an annual cohort size of 6,000 adult status holders

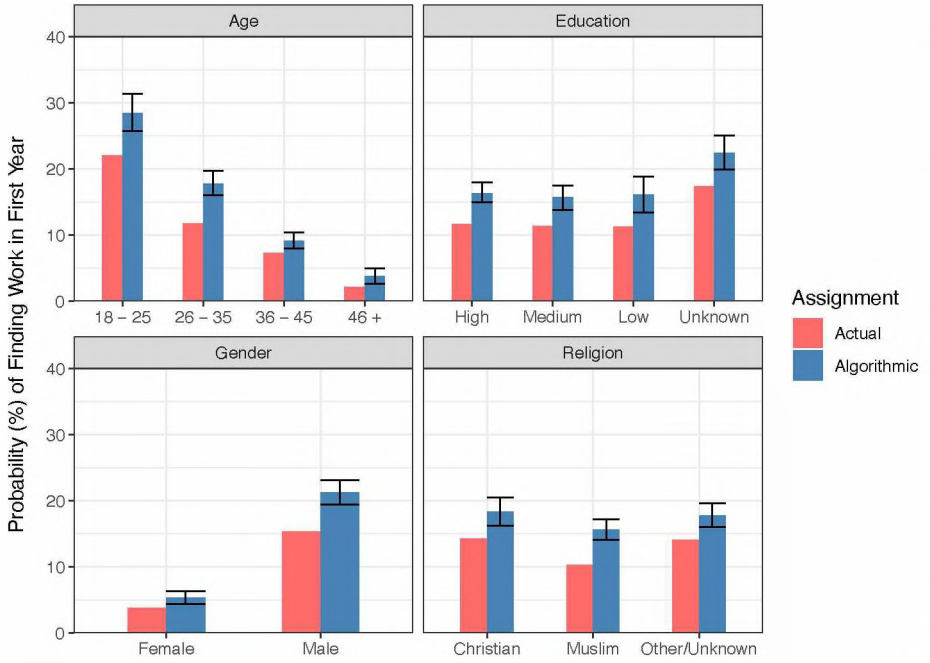
Agenda

Our discussion today

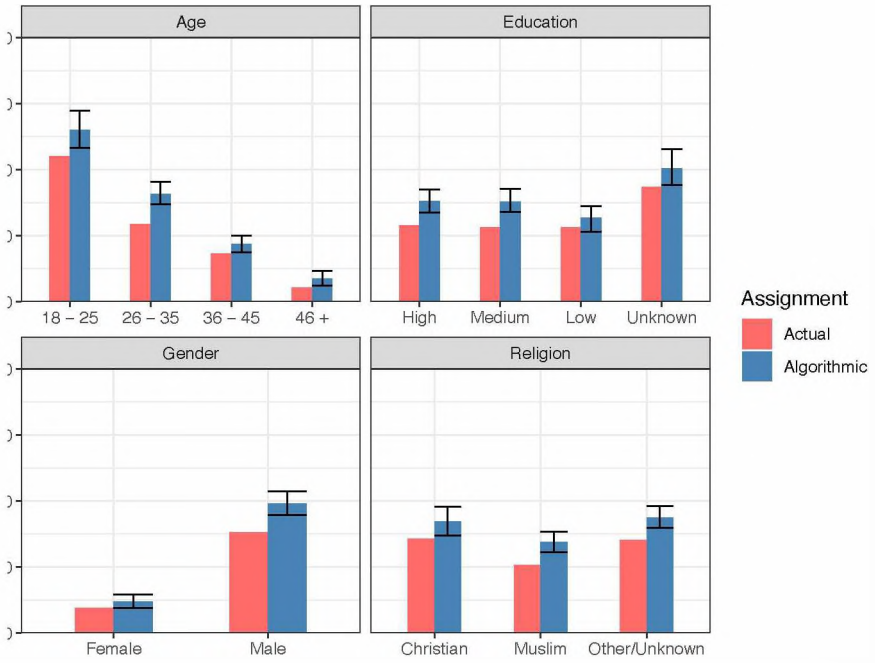
1. Impact on overall gains
2. **Impact on subgroups**
3. Impact on subgroups (nationalities)
4. Implications and discussion

Impact on subgroups

With sensitive variables



Without sensitive variables



Among the subgroups shown here, the losses appear to be equally distributed

ipl immigration policy lab

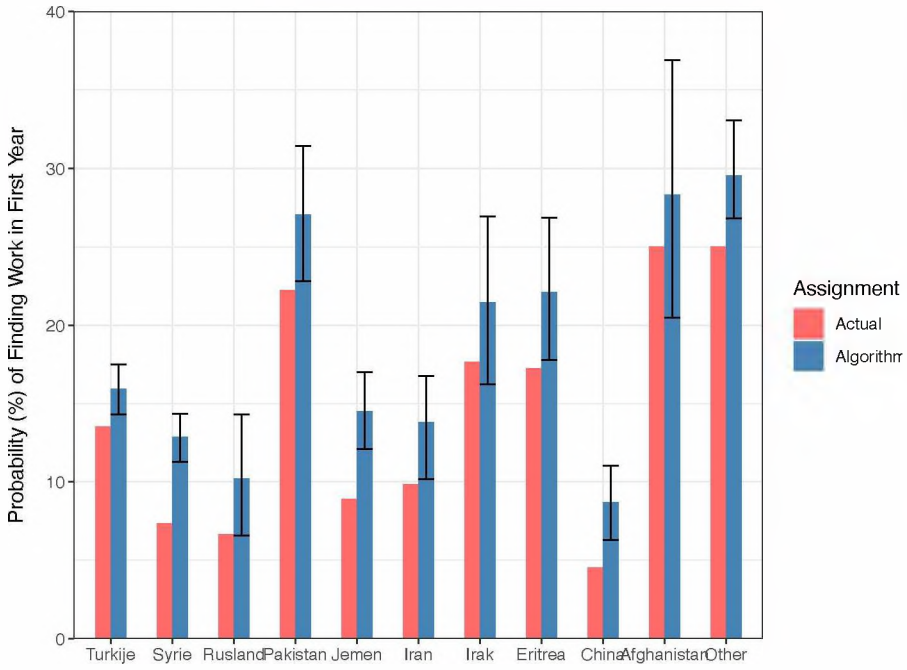
Agenda

Our discussion today

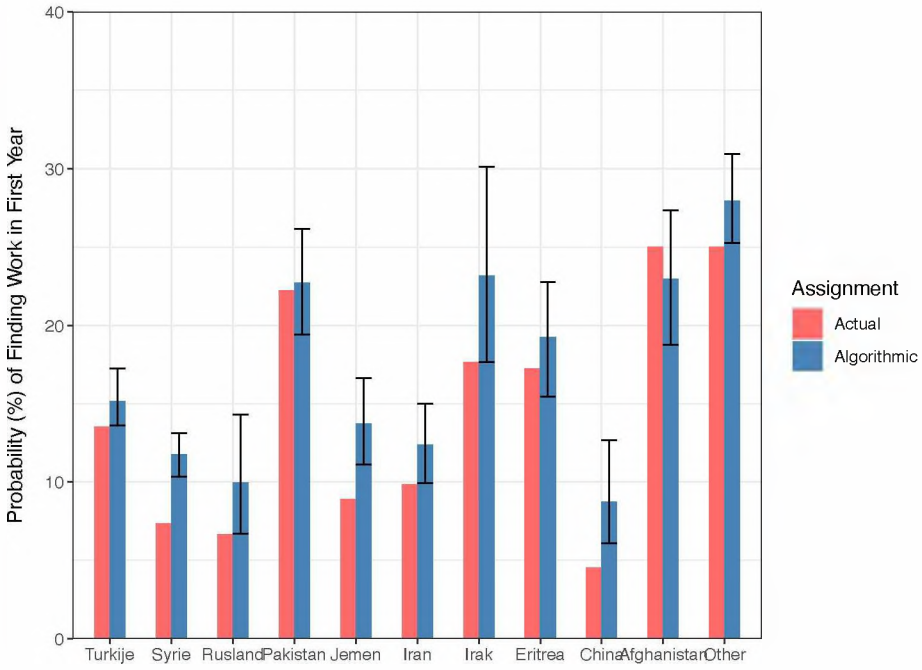
1. Impact on overall gains
2. Impact on subgroups
3. **Impact on subgroups (nationalities)**
4. Implications and discussion

Impact on subgroups (nationalities)

With sensitive variables



Without sensitive variables



Impact potentially negative on some nationalities (but small sample sizes)

ipl immigration policy lab

Agenda

Our discussion today

1. Impact on overall gains
2. Impact on subgroups
3. Impact on subgroups (nationalities)
4. **Implications and recommendations**

Implications and risks

Data science and fairness implications/risks of losing access to sensitive variables

- **A 30% loss in expected overall gains**
  - 3.3 percentage point expected increase in employment (compared to 4.64 with more variables) → we would lose 30% of the gains
  - An estimated 200 status holders would find employment in their first year from GeoMatch (compared to 280 with more variables)
- **A potential (but hard to estimate) negative impact on some subgroups**
  - It is possible that some groups might lose compared to the status quo if the algorithm can't use sensitive data
  - These groups are too small in the backtest cohort to be certain
- **Fairness concerns**
  - We could only measure the impact on sensitive subgroups during monitoring (if we have enough non-missing data on these variables)
  - If we do find that some groups end up worse-off, we would be unable to make any correction in the algorithm
  - We are currently developing new methods to ensure fairness across groups, but cannot implement these methods without access to the variables

Recommendations

Recommendations for proceeding without sensitive variables

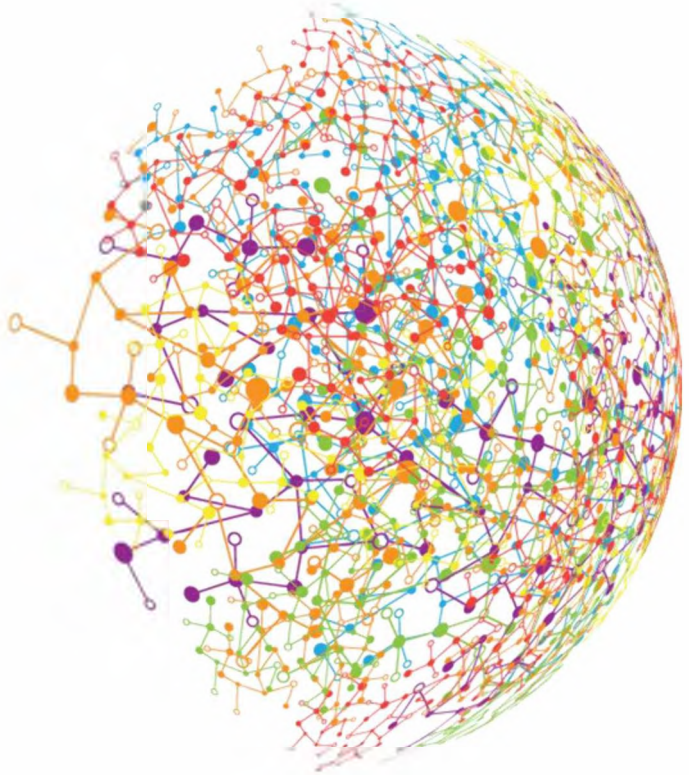
- **Strong preference for the inclusion of sensitive variables**
  - Larger gains
  - More evenly distributed gains across key subgroups
  - Greater ability to help the refugee population and key subgroups
  - Recourse if fairness metrics are violated

- **Can still proceed without sensitive variables**
  - Overall gains still present
  - Would need to think through our monitoring & response plan, and ethical considerations, surrounding fairness

Docnr5

Deloitte.

COA Centraal Orgaan opvang asielzoekers



Algorithm analyse

Tussentijdse rapportage

3 Juli 2023
KD/avl/23-0957

Achtergrond en activiteiten

Deloitte heeft in de zomer van 2022 werkzaamheden uitgevoerd als onderdeel van de onderzoeksfase van het project “kansrijke koppeling met behulp van AI”. Die werkzaamheden hebben geresulteerd in een rapport met bevindingen, bestaande uit observaties en aanbevelingen. Op basis van deze observaties en aanbevelingen hebben COA en IPL het algoritme aangepast in voorbereiding op de pilotfase. In deze tweede analyse zal Deloitte, zoals overeengekomen in de opdrachtbrief met kenmerk KD/js/23-0546, het aangepaste algoritme in twee stappen analyseren. Deze tussentijdse rapportage bevat de resultaten van stap 1 van deze tweede analyse.

Stap 1.

Tijdens de eerste stap analyseert Deloitte de perspectieven 'Model', 'Data’ en ‘Ethiek’ op basis van de technische algoritme documentatie en het monitoringsplan. De nadruk zal liggen op de opvolging van de volgende aanbevelingen uit de eerste rapportage van Deloitte*; 1, 2, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, CBS-1, CBS-3, CBS-4 en CBS-5.

De resultaten van stap 1 zijn opgenomen in de hierna volgende tabellen. De observaties en aanbevelingen (gezamenlijk: “bevindingen”) uit de eerste rapportage van Deloitte, zijn één op één overgenomen uit het document met referentie KD/lr/22-1275. Op basis van de stap 1 analyse is per bevinding een status opgenomen (open, conditioneel gesloten, gesloten) in de 'Status' kolom, met bijbehorende toelichting in de 'Uitleg' kolom.

Stap 2.

In de tweede stap analyseert Deloitte de overgebleven onderdelen die niet zijn geanalyseerd in stap 1. De nadruk zal liggen op de andere, door COA opgestelde, documenten die niet zijn geanalyseerd in stap 1. Minimaal zal Deloitte de opvolging van de volgende aanbevelingen analyseren*: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 17, 18, CBS-2 en CBS-4.

Deze tussentijdse rapportage heeft enkel betrekking op stap 1. De resultaten van stap 2 volgen in een later stadium.

Deloitte voert de analyse uit op basis van documenten die zijn aangeleverd door het COA als onderdeel van de algoritmedocumentatie. De hoofddocumenten van deze analyse zijn:

- NLD_GeoMatch_TechnicalDocumentation_20230418.pdf (hierna: “Technische Documentatie)”;
- Annex_GeoMatchNetherlands_PilotMonitoringPlan_20230416.pdf (hierna: “Monitoringsplan”);
- 230419 Audit 2.1 Deloitte project Kansrijke Koppeling m.b.v. AI.xlsx (hierna: “Bevindingen Status Excel”);
- Annex GeoMatchNetherlands_UpdatedBacktests_060222 (hierna: “Backtest rapport”).

Het document Bevindingen Status Excel is geen onderdeel van de algoritmedocumentatie van het COA, hierin is enkel opgenomen hoe COA opvolging heeft gegeven aan de gemaakte observaties en aanbevelingen.

Deloitte heeft niet gecontroleerd of de aangeleverde set van informatie compleet is.

* De details van de eerste analyse inclusief de observaties en aanbeveling zijn te vinden in eerste rapport met referentie KD/lr/22-1275.

Tussentijdse rapportage (1/8)

	Observatie	Aanbeveling	Status	Uitleg
1	De vaststelling van rollen en verantwoordelijkheden in de gehele algoritmelevenscyclus is beperkt uitgewerkt in de algoritmedocumentatie voor interne en externe belanghebbenden die betrokken zijn bij de ontwikkeling van het algoritme.	Verrijk de algoritmedocumentatie met de rollen en verantwoordelijkheden van interne en externe actoren die (eind)verantwoordelijk is (zijn) voor bepaalde werkzaamheden.	Gesloten	• COA heeft het document *Annex DARE chart.png* toegevoegd aan de algoritmedocumentatie. Dit bestand omschrijft de rollen, COA stakeholders die betrokken zullen zijn voorafgaand en/of tijdens de pilot, en de stakeholder impact op het project. Tijdens stap 1 van deze analyse heeft COA hier stakeholders aan toegevoegd, namelijk CBS als data provider en IPL als ontwikkelaar en beheerder van het algoritme.
2	De externe belanghebbenden worden pas tijdens de pre-pilot via de Landelijke Regietafel Migratie & Integratie ("LRT") betrokken. Hiermee ontstaat het risico dat fundamentele veranderingen later in het project worden doorgevoerd en/of hun belangen niet volledig worden gewaarborgd, wat tot een risico op vooringenomenheid van het algoritme kan leiden.	Betrek alle belanghebbenden zo vroeg mogelijk in de pre-pilot bij de besluitvorming door te vragen naar de waarden, uitgangspunten en belangen die voor hen belangrijk zijn bij de toepassing van het algoritme.	Conditioneel gesloten	• Het document *Annex Stakeholdersession.docx* bevat de notulen van een twee uur durende stakeholdersessie op 7 december 2022, waarin externe stakeholders zijn geïnformeerd over het project en gevraagd om hun ideeën, waardes en principes in de context van het algoritme. Verder is COA voornemens voorafgaand aan de pilot nog een stakeholder sessie te organiseren waarin meer details worden gedeeld. • De externe stakeholder Vereniging van Nederlandse Gemeenten (hierna: "VNG") was niet aanwezig bij bovengenoemde stakeholdersessie. COA heeft aangegeven dat dezelfde sessie met VNG is gehouden op 19 december 2022, en dat hieruit geen additionele punten naar voren zijn gekomen. • Deze bevinding kan worden gesloten indien COA in de algoritmedocumentatie opneemt dat bovenstaande sessie met VNG heeft plaatsgevonden en wat de uitkomsten daarvan waren.
5	Het COA is afhankelijk van externe partijen voor het gebruik en beheersing van het algoritme. Er zijn geen uitgewerkte overeenkomsten wat de diensten en verantwoordelijkheden van het IPL en het CBS definieert, en het COA beschikt zelf nog niet over de nodige expertise.	Verrijk de algoritmedocumentatie met de rollen en verantwoordelijkheden van interne en externe actoren die (eind)verantwoordelijk is (zijn) voor bepaalde werkzaamheden.	Open	• COA heeft aangegeven een Service Level Agreement tijdens stap 2 te delen. • Deze bevinding blijft hiermee open, en zal tijdens stap 2 opnieuw geanalyseerd worden op basis van de Service Level Agreement.
6	Er is gebrek aan beschikbaarheid van analyses en gemaakte (ontwerp)keuzes binnen het algoritme.	Verrijk de algoritmedocumentatie met de uitgevoerde analyses, waar mogelijk de uitkomsten van de analyses, de gemaakte ontwerpkeuzes en de mechanismen om gedefinieerde werkbegrippen te waarborgen.	Open	• In sectie 6c van de Technische Documentatie is beschreven welke datakwaliteitchecks en analyses zijn uitgevoerd (zie bevinding 11). Verder heeft COA meerdere analyses toegevoegd in de Technische Documentatie rondom uitlegbaarheid (zie bevinding 13 en 14). Daarnaast beschrijft COA in sectie 2, 3 en 4 dat de methodologie ontwerpkeuzes bevat, zoals "Mapping" functies en optimalisatie criteria (zie bevinding 16). • De uitkomsten en conclusies van checks en analyses worden niet beschreven. Daarnaast worden ontwerpkeuzes aangegeven, maar de uitkomst en argumentatie van deze ontwerpkeuzes niet beschreven. • De bevinding blijft hiermee open, en zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie.

Tussentijdse rapportage (2/8)

	Observatie	Aanbeveling	Status	Uitleg
7	Er is geen monitoringsplan opgesteld, waarin afspraken over het monitoren en beheer gedurende de algoritmelevenscyclus van het algoritme worden gedefinieerd; evenals de afspraken met het IPL, het CBS. Dit brengt als risico dat de monitoringsactiviteiten niet of te weinig uitgevoerd zullen worden.	Verrijk de algoritmedocumentatie met een monitoringsplan, waarin bijvoorbeeld activiteiten, toegestane marges, governance en een stopknop gedefinieerd zijn inclusief de frequentie van de activiteiten. Deze documentatie en afspraken dienen op orde te zijn voorafgaand aan het gebruik van het algoritme in de pilot. In de pilot wilt het COA het algoritme onderzoeken door het algoritme te gebruiken onder een representatief deel van de vergunninghouders.	Open	• Het Monitoringsplan beschrijft verschillende 'monitoring levels' ofwel monitoringsactiviteiten. Per monitoring level wordt het doel, benodigde data en de monitoring frequentie beschreven. Daarnaast worden details van de uit te voeren analyses gegeven, met bijbehorende drempelwaarden en vervolgacties. Er is geen invulling gegeven aan de hoogte van de drempelwaarden voor monitoringslevel 1, 4 en 5 (nu beschreven als X% en Y%) en binnen welke termijn de vervolgacties dienen te worden opgevolgd. • Deze bevinding kan worden gesloten indien COA de drempelwaarden vaststelt en opneemt in het Monitoringsplan, en opneemt binnen welke termijn er opvolging gegeven dient te worden aan vervolgacties. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van het Monitoringplan.
8	De doelstellingen worden onvoldoende concreet uitgewerkt in lange termijn prestatie-indicatoren. Verder zijn er geen acceptabele toleranties en (fout)marges gedefinieerd op zowel korte- als lange termijndoelstellingen.	Verrijk de algoritmedocumentatie met meetbare (kwantitatieve) kwaliteitseisen van het algoritme, voorafgaand aan de pre-pilot.	Open	• COA heeft invulling gegeven aan deze bevinding door middel van het Monitoringsplan zoals beschreven onder bevinding 7. COA is voornemens drempelwaarden vast te stellen en op te nemen in de finale versie van het Monitoringsplan. • Er is niet beschreven hoe het algoritme aan het eind van de pilot geëvalueerd wordt en wanneer COA de pilot als geslaagd ziet. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van het Monitoringplan en/of algoritmedocumentatie.
9	Er zijn geen processen en/of mechanismen ingesteld om algemene feedback op te halen ten aanzien van het algoritme. Verder komt het doel van het gebruikersoverleg niet overeen met het doel van het COA om interne en externe feedback te ontvangen.	Stel een aanspreekpunt aan waar derden en algoritmegebruikers terecht kunnen met zowel hun vragen/zorgen, als algemene feedback. Wijs een eindverantwoordelijke binnen het COA aan die zorgt voor afstemming met het IPL over (mogelijke) aanpassingen aan de hand van ontvangen feedback.	Open	• COA beschrijft in de Bevindingen Status Excel de rol van centrale coördinator voor de verschillende focus groepen, en dat de project 5.1.2.e eindverantwoordelijk is voor het verwerken van feedback en het afstemmen met IPL over (mogelijke) aanpassingen naar aanleiding van ontvangen feedback. • COA geeft aan dat in een algoritmeregister zal worden opgenomen wie het aanspreekpunt is waar derden en algoritmegebruikers terecht kunnen. • Deze bevinding kan worden gesloten indien COA de beschrijvingen zoals opgenomen in de Bevindingen Status Excel opneemt in de algoritmedocumentatie en/of het algoritmeregister. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie en/of het algoritmeregister.
11	Er is onvoldoende documentatie van de analyses die uitgevoerd zijn door het IPL en/of het COA om de kwaliteit en integriteit van de data te waarborgen. Verder zijn er geen data beschrijvende statistieken beschreven in de modeldocumentatie.	Documenteer de analyses, die zijn uitgevoerd, inclusief de resultaten en conclusies, die uitgevoerd zijn om datakwaliteit te analyseren. Verder is het aanbevolen om datakwaliteit te monitoren wanneer het algoritme opnieuw getraint gaat worden.	Open	• In sectie 6c van de Technische Documentatie is beschreven welke datakwaliteitchecks zijn uitgevoerd. De uitkomst en de conclusie van deze checks is niet opgenomen in de algoritmedocumentatie. COA heeft aangegeven additionele details op te nemen in de algoritmedocumentatie. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie.

Tussentijdse rapportage (3/8)

	Observatie	Aanbeveling	Status	Uitleg
12	De definitie van eerlijkheid is niet volledig voor wat betreft de context van het algoritme. Het is niet gedefinieerd welke groepen er met elkaar vergeleken worden, welke marges er zijn en hoe eerlijkheid gemeten wordt.	Werk de definitie van eerlijkheid uniform en volledig uit inclusief procedures om eerlijkheid te meten op de testset en tijdens het monitoren van het algoritme. Toets de definitie van eerlijkheid met een diverse groep (minimaal) bestaande uit het COA, de LRT, de vergunningshouders en/of belangenbehartigers maar ook een onafhankelijke afvaardiging vanuit de maatschappij in Nederland.	Open	• Sectie 4.c.ii in de Technische Documentatie beschrijft COA's definitie van eerlijkheid vanuit (1) het perspectief van de statushouders, en (2) het perspectief van de arbeidsmarktregio's. De analyse en conclusie vanuit het perspectief van de arbeidsmarktregio's is niet opgenomen in de Technische Documentatie. COA heeft aangegeven dit voor stap 2 op te zullen nemen. • In het Monitoringsplan is opgenomen onder 'monitoring level 3', dat jaarlijks wordt uitgevoerd, of de verdeling van specifieke subgroepen (o.a., leeftijd ≥40, vrouwen, herkomstland Syrië) over arbeidsmarktregio's statistisch niet meer afwijkt dan de historisch geobserveerde maximum afwijking. COA geeft in het Monitoringsplan aan dat de lijst van subgroepen nog niet finaal is en voor stap 2 wordt gefinaliseerd. • COA geeft aan voor stap 2 de opgestelde definitie van eerlijkheid en monitoring hiervan te toetsen met stakeholders. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de Technische Documentatie, het Monitoringplan en het stakeholder proposal.
13	Het algoritme gebruikt alle attributen uit IBIS die potentieel relevant zijn voor het te behalen doel. Er wordt geen inputselectie (feature selection) toegepast om te analyseren of attributen relevant zijn. Het algoritme kan hierdoor onnodig complex zijn en onnodig veel data gebruiken. Het gebruiken van onnodig veel data is in strijd met het dataminimalisatie-principe uit de AVG wanneer de data bestaat uit persoonsgegevens.	Pas input selectie toe om te onderbouwen waarom attributen nodig zijn in het algoritme.	Open	• COA heeft input selectie toegepast door middel van een kwantitatieve en kwalitatieve analyse. Het COA geeft aan dat voor de kwalitatieve analyse interviews met COA-medewerkers en proces experts binnen Nederland zijn gehouden. Voor de kwantitatieve analyse is gekeken naar *variable importance plots* (opgenomen in de Technische Documentatie), en naar resultaten van het model van IPL in andere landen, en academische literatuur. • In de algoritmedocumentatie is niet opgenomen hoe de input van experts is meegenomen en of dit geleid heeft tot aanpassingen in de gebruikte attributen. Daarnaast zijn de *variable importance plots* in de Technische Documentatie verouderd, en er is geen onderbouwde conclusie opgenomen in de Technische Documentatie die COA hieraan verbindt. • COA heeft verder aangegeven dat er vanuit (interne en externe) stakeholder sessies geen directe feedback is geweest op de gebruikte attributen. Het is onduidelijk of expliciet is getoetst met stakeholders dat het wenselijk is dat persoonskenmerken zoals geslacht, leeftijd en land van herkomst gebruikt worden voor de plaatsing van statushouders. • De bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie.
14	De relatie van de attributen ten opzichte van de uitkomst is alleen geanalyseerd op het belang van de attributen. Echter, de directe relatie tot de uitkomst (bijvoorbeeld met het SHAP raamwerk, Partial Dependence Plots en/of Individual Expectations Plots) is niet gedocumenteerd. Hierdoor is het algoritme niet technisch uitlegbaar en kan het algoritme leren van verkeerde relaties (spurious correlations) in de data.	Documenteer de werking van het algoritme aan de hand van globale uitlegbaarheidstechnieken.	Open	• Zoals aangegeven bij bevinding 13 zijn de *variable importance plots* in de Technische Documentatie verouderd, en er is geen onderbouwde conclusie opgenomen in de Technische Documentatie die COA hieraan verbindt. • Verder blijkt uit de *variabel importance plots* dat leeftijd een sterk verklarende variabele is in veel arbeidsmarktregio's. Er is niet geanalyseerd in hoeverre dit erin resulteert dat het algoritme bepaalde leeftijdsgroepen altijd in bepaalde arbeidsmarktregio's plaatst en of dit wenselijk is. • De bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie.

Tussentijdse rapportage (4/8)

	Observatie	Aanbeveling	Status	Uitleg
15	Tijdens de ontwikkeling van het algoritme zijn bepaalde groepen, zoals mensen met harde criteria, vergunningshouders onder het kinderpardon van 2019 en onder de EU-Turkije deal, buiten scope gelaten. Het algoritme is hierdoor niet getraind op deze groepen en werkt mogelijk niet optimaal voor deze groepen. Wanneer een dergelijke groep in de toekomst in-scope is, moet het algoritme opnieuw getraind worden. Deze beperking is niet beschreven in de algoritmedocumentatie.	Verrijk de algoritmedocumentatie met de beperking dat het algoritme alleen toepasbaar is op de huidige scope en kalibreer het algoritme opnieuw als de scope verandert.	Open	• In sectie 6.b van de Technische Documentatie wordt beschreven dat de in-scope populatie wordt vastgesteld door het filteren van minderjarigen, de vaststelling door de programma begeleider dat geen sprake is van harde criteria, en het feit dat de statushouder in aanmerking komt voor werk. Er wordt aangegeven dat dit specifiek opgenomen zal worden in de instructies voor werknemers en informatiebladen voor statushouders welke COA voor aanvang van de Pilot zal delen. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de werkinstructie.
16	Essentiele onderdelen van het algoritme, zoals het combineren van de uitkomsten van meerdere vergunningshouders in één plaatsingseenheid, het stochastic online matching component, en de onderbouwing voor ontwerpkeuzes van deze onderdelen zijn niet opgenomen in de algoritmedocumentatie.	Verrijk de algoritmedocumentatie met alle onderdelen van het algoritme inclusief de ontwerpcomponenten van het algoritme zodat het algoritme gereproduceerd kan worden door een onafhankelijke derde partij.	Open	• In secties 2, 3 en 4 van de Technische Documentatie worden de drie onderdelen van het algoritme beschreven. Voor aanvullende details verwijst COA naar wetenschappelijke publicaties en aanvullende documenten op deze publicaties. • Verder beschrijft COA in de Technische Documentatie dat er verschillende implementatieopties, zoals verschillende “mapping” functies en optimalisatie criteria, mogelijk zijn. Er is niet beschreven welke optimalisatie criteria COA heeft gekozen (bijvoorbeeld of “load balancing” wordt toegepast), waarom COA een specifieke “mapping” functie en optimalisatie criteria kiest en welke (ethische) gevolgen deze keuzes hebben op het algoritme. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie.
17	Tijdens het backtesten komt onvoldoende naar voren welke impact de inzet van het algoritme heeft op subgroepen die niet worden geplaatst met ondersteuning van het algoritme. Verder ontbreekt een analyse naar subgroepen die op basis van de leeftijdscriteria in-scope zijn maar geen (of beperkt) deel zullen uitmaken van de arbeidsmarkt, zoals vergunningshouders met bijzondere behoeften, arbeidsongeschikten en senioren.	Analyseer de impact van het algoritme op subgroepen die niet worden geplaatst met het algoritme en op subgroepen die door het algoritme worden geplaatst maar geen onderdeel zijn van de arbeidsmarkt.	Open	• COA heeft aangegeven dat voor subgroepen die niet worden geplaatst door het algoritme niet altijd data beschikbaar is, en dat het niet mogelijk is een analyse uit te voeren op subgroepen waarvoor geen data beschikbaar is. Het COA geeft aan in de werkinstructie op te nemen over de beperkingen van het algoritme binnen bepaalde target groepen. • Merk op dat een deel van de out-of-scope populatie geplaatst wordt met voorrang in bepaalde arbeidsmarktregio's (e.g. harde criteria), waardoor de verwachting is dat het algoritme slechts beperkt invloed heeft op de arbeidsmarktregio’s waar deze individuen geplaatst zullen worden. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de werkinstructie.
18	De beschrijving en de accuraatheid van het impact op loon-analyse is niet transparant. Dit heeft effect op de betrouwbaarheid van de impact op loon analyse en van de financiële voordelen analyse.	Wees transparant over de accuraatheid, en daarmee de betrouwbaarheid, van de impact op loon-analyse. Verder is het aanbevolen om niet actief te sturen op de uitkomsten van de impact op loon en financiële voordelen analyse voordat dit duidelijk is.	Conditioneel gesloten	• COA heeft in sectie 6.g.ii van de Technische Documentatie de analyse waaruit de financiële voordelen van het algoritme blijken verduidelijkt, waarbij op de aannames wordt gewezen en een betrouwbaarheidsinterval wordt gegeven. COA is voornemens dit op te nemen in een stakeholder voorstel. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van het stakeholder voorstel.

Tussentijdse rapportage (5/8)

	Observatie	Aanbeveling	Status	Uitleg
19	Het COA heeft geen inzage in de door IPL uitgevoerde analyses naar de integratiewinst van het gebruik van de bijzondere persoonsgegevens.	Verrijk de algoritmedocumentatie met een onderbouwing van de noodzakelijkheid en evenredigheid van de afzonderlijke bijzondere persoonsgegevens (niet als een geheel). Hiervoor is inzage in de analyses van het IPL essentieel. Na afloop van de analyse heeft het COA aangegeven dat de volgende versie van het algoritme geen gebruik zal maken van bijzondere persoonsgegevens.	Gesloten	• In het Backtest Rapport geeft COA inzichten in de prestaties van het algoritme met het gebruik van bijzondere persoonsgegevens en zonder bijzondere persoonsgegevens. • COA heeft aangegeven dat het algoritme niet de kenmerken *religie* en *etniciteit* zal gebruiken. Het algoritme zal wel kenmerken als *land van herkomst* en *moedertaal* gaan gebruiken. De combinatie van deze kenmerken kan leiden tot het gebruik van een bijzonder persoonsgegeven. Dit wordt verder geanalyseerd onder bevinding 20.
20	Het COA heeft geen standpunt gevormd over het gebruik van de bijzondere persoonsgegevens.	Werk het standpunt uit of het COA bijzondere persoonsgegevens mag en wil gebruiken. Toets het standpunt van het COA met relevante belanghebbenden. Na afloop van de analyse heeft het COA aangegeven dat de volgende versie van het algoritme geen gebruik zal maken van bijzondere persoonsgegevens.	Open	• In de nieuwste versie van het algoritme worden de bijzondere persoonsgegevens *religie* en *etniciteit* niet langer gebruikt. Kenmerken als *land van herkomst* en *moedertaal* worden wel gebruikt in het algoritme. • De Algemene verordening gegevensbescherming (“AVG”) beschouwt persoonsgegevens waaruit ras of etnische afkomst blijkt als bijzondere persoonsgegevens. Het risico dat het gezamenlijke gebruik van kenmerken als *land van herkomst* en *moedertaal* leidt tot het gebruik van een bijzonder persoonsgegeven is onvoldoende onderzocht en gedocumenteerd door COA. • In stap 2 zal de privacy impact assessment (“PIA”) worden aangeleverd, waarin onderbouwing, proportionaliteit, evenredigheid en mogelijke wettelijke grondslag van de gebruikte attributen zal worden gegeven. • De bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de PIA.

Tussentijdse rapportage (6/8)

	Observatie	Aanbeveling	Status	Uitleg
CBS 1	Er zijn vergunningshouders waarvan het label (de doelvariabele wel/geen baan) ontbreekt en die dus ontbreken in de trainingsdata.	Vergelijk de samenstelling van de doelpopulaties met en zonder doelvariabele.	Gesloten	• COA beschrijft in sectie 6.a van de Technische Documentatie hoe de doelpopulatie wordt vastgesteld, en heeft bevestigd dat voor de volledige populatie de doelvariabele beschikbaar is. Dit is ook op te maken uit de tabel in Appendix 2.c van de Technische Documentatie.
CBS 3	Augustus 2020 is geen representatief meetmoment voor het meten van baankansen onder vergunningshouders (of alle inwoners van Nederland), vanwege de uitbraak van het coronavirus in Nederland en de bijbehorende maatregelen in maart 2020 en later. Vergunningshouders werken relatief vaak op tijdelijke contracten in sectoren als de horeca. Het is aannemelijk dat deze groep relatief hard geraakt is door de maatregelen en op dat moment noodgedwongen minder is gaan werken of werkloos is geworden.	Verkrijg inzicht in de verschillen tussen augustus 2020 en de periode vóór coronamaatregelen (in hoeverre werken vergunningshouders minder uren, of in andere sectoren). Om aan de hand van deze resultaten te bepalen hoe representatief de resultaten zijn voor een typisch jaar. Indien mogelijk: herhaal de analyse op een periode zonder coronamaatregelen.	Open	• COA beschrijft in de Bevindingen Status Excel dat er momenteel nog geen post-COVID data beschikbaar is, en dat het daarmee nog niet haalbaar is een analyse op deze periode uit te voeren. • Er is geen analyse uitgevoerd op de verschillen tussen de pre-COVID periode en de COVID periode, omdat COA van mening is dat de pre-COVID periode niet representatief is voor de post-COVID periode. • Het COA is voornemens voorafgaand aan de start van de pilot het algoritme opnieuw te trainen op basis van de meest recente data. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie.

Tussentijdse rapportage (7/8)

	Observatie	Aanbeveling	Status	Uitleg
CBS 4	De samenstelling van de trainingsdata per arbeidsregio kan afwijken van de samenstelling van de gehele populatie vergunningshouders (door de selectieve allocatie) en van de samenstelling van toekomstige. Elk regionaal model zal zwaarder leunen op de relatief veel voorkomende groepen in die regio (een regio heeft mogelijk vooral vergunningshouders met een bepaalde nationaliteit of werksector). Dit leidt wellicht tot vertekening van de voorspellingen van relatief zeldzame groepen in deze regio. Daarnaast kan het zo zijn dat bepaalde groepen helemaal niet voorkomen in bepaalde regio's. Het model kan dan geen inschatting maken van het succes van mensen met kenmerk X in regio A. Dit maakt de vergelijkbaarheid tussen regio's, en de afweging van het model over in welke regio's de vergunningshouders de hoogste baankans heeft, moeilijk.	Dit representativiteitsrisico is inherent aan de regionale opzet van het model. Dit risico kan gemitigeerd worden door het creëren van bewustzijn bij medewerkers, of er rekening mee te houden in de modelopzet. Om de disbalans in de regionale trainingsdata te balanceren zou het model bijvoorbeeld getraind kunnen worden met grotere gewichten voor relatief zeldzame groepen en lagere gewichten voor relatief veel voorkomende groepen. Een procentuele vergelijking van welke subgroepen in welke regio's wonen geeft al een eerste beeld van de mate van balans/disbalans. Ook worden zo subgroepen in beeld gebracht die in bepaalde regio's niet of amper voorkomen en waarvoor de voorspellingen van het model minder betrouwbaar zijn. Om inzicht te krijgen in de mate van afwijking tussen regio's, en het risico dat de gemiddeld baankans voor die regio vertekend is, kan gebruikt worden gemaakt van de variatiecoëfficiënt (CV). Bij de COA-medewerkers kan bewustzijn gecreëerd worden door de wijze waarop zij de output van het model te zien krijgen. Als zij, naast de aanbevolen regio's, ook de geschatte baankans en een betrouwbaarheidsscore per regio te zien krijgen, geeft hen dit waardevolle informatie over de robuustheid van de aanbeveling.	Open	• COA beschrijft in sectie 6.b van de Technische Documentatie het representativiteitsrisico zoals beschreven in de observatie. COA beschrijft dat wanneer bepaalde groepen missen in arbeidsmarktregio's, het model dan geen inschatting kan maken voor deze groepen in de specifieke regio. Verder beschrijft COA dit niet te zijn tegengekomen in de data. • COA heeft aangegeven dit verder te monitoring als onderdeel van "monitoring level 3". Echter wordt deze monitoringsactiviteit niet omschreven in monitoring level 3 in het Monitoringsplan. COA heeft aangegeven de monitoringsactiviteit in het Monitoringsplan te beschrijven. • Daarnaast beschrijft de aanbeveling dat er bewustzijn gecreëerd kan worden bij de COA-medewerkers door de wijze waarop zij de output van het model te zien krijgen. In stap 2 zal de werkinstructie en voorlichting naar COA-medewerkers worden aangeleverd door COA. • Deze bevinding zal in stap 2 opnieuw geanalyseerd worden op basis van het Monitoringsplan en de werkinstructie.

Tussentijdse rapportage (8/8)

	Observatie	Aanbeveling	Status	Uitleg
CBS 5	Het is mogelijk dat vergunningshouders verhuizen in het eerste jaar na toewijzing aan de arbeidsmarktregio. Dit gegeven wordt in de inputdata van het model niet meegenomen. Het zou zo kunnen zijn dat vergunningshouders verhuizen en in de nieuwe arbeidsmarktregio een baan vinden. In de inputdata wordt dit geregistreerd alsof een baan in de toegewezen arbeidsmarktregio wordt gevonden. Deze keuze is gemaakt om bias te verminderen: het model wordt niet getraind op factoren die je nog niet kan weten op het moment van toewijzing aan een arbeidsmarktregio. Tegelijkertijd zou deze keuze juist een mogelijke bron van bias kunnen zijn als er arbeidsmarktregio's zijn waar een groot deel van de vergunningshouders snel weer verhuist. De voorspellingen van het model voor deze arbeidsmarktregio kloppen dan niet. Wij raden aan om een extra analyse uit te voeren om te bepalen in hoeverre dit risico van toepassing is.	Model trainen op vergunningshouders die nog steeds woonachtig zijn in de arbeidsmarktregio waaraan ze zijn toegewezen en kijken of dit de resultaten verandert, of controleer hoe vaak het voorkomt dat personen naar een andere arbeidsmarktregio verhuizen.	Open	• COA ligt in sectie 6.d van de Technische Documentatie toe waarom ze van mening is dat statushouders die verhuizen en een nieuwe baan in een andere arbeidsmarktregio vinden ook een positieve doelvariabele zouden moeten krijgen. • Indien kan worden aangetoond dat de grote van deze verhuizende populatie beperkt is in de verschillende arbeidsmarktregio's, en daarmee de impact van deze verhuizende populatie ook beperkt zal zijn, kan deze bevinding gesloten worden. • COA heeft aangegeven om voor stap 2 te onderzoeken of het mogelijk is om te analyseren of de impact van de verhuizende populatie beperkt is. • De bevinding blijft hiermee open, en zal in stap 2 opnieuw geanalyseerd worden op basis van de algoritmedocumentatie.

Disclaimers

De algoritme analyse is niet uitgevoerd in het kader van een assurance-opdracht zoals gedefinieerd in het International Framework for Assurance Engagements van de International Federation of Accountants (“IFAC”). Het is de verantwoordelijkheid van het COA om te beoordelen of de resultaten van de algoritme analyse in het perspectief van het geheel van de hen ter beschikking staande informatie en hun risicoperceptie aan de door hen te stellen eisen voldoen.

Uitsluitend het COA is verantwoordelijk voor onder meer: (a) het nemen van alle managementbeslissingen, het uitoefenen van alle managementfuncties en het dragen van alle managementverantwoordelijkheden, (b) het aanwijzen van een competente vertegenwoordiger namens het management die de analyse overziet, (c) het evalueren van de toereikendheid en de resultaten van de analyse, (d) het accepteren van de verantwoordelijkheid voor het gebruik van de resultaten van de analyse, en (e) het inrichten en onderhouden van een intern beheersingssysteem, inclusief het toezicht op de lopende activiteiten.

Deloitte wil benadrukken dat de opgeleverde deliverables (waaronder deze tussentijdse rapportage) slechts bedoeld zijn voor intern gebruik door het COA ten behoeve van het omschreven doel in de opdrachtbrief met referentie KD/js/23-0546. Verder mag deze tussentijdse rapportage ook gedeeld worden met de algoritme ontwikkelaars (in dit geval het IPL) en het CBS.

Zonder uitdrukkelijke en voorafgaande schriftelijke toestemming van Deloitte is het niet toegestaan deliverables, dan wel delen daaruit, te gebruiken voor andere doeleinden dan overeengekomen, aan derden te verspreiden of openbaar te maken, aan deliverables te refereren of uit deliverables te citeren.

Op dit rapport zijn de Algemene Voorwaarden Dienstverlening Deloitte Nederland, januari 2020, van toepassing waarvan u bijgaand een exemplaar aantreft.

Deloitte.