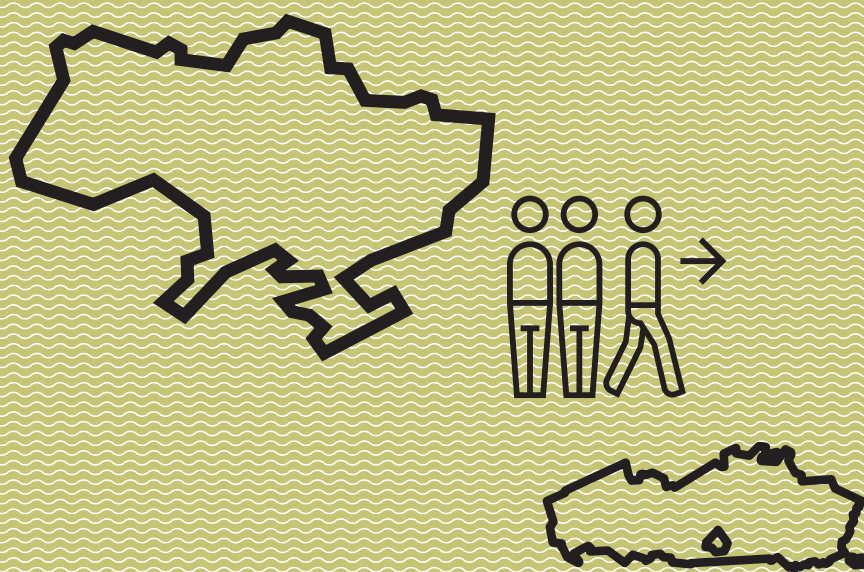




Vlaanderen
is sterk in cijfers



RAPPORT STATISTIEK
VLAANDEREN, 2025/5

STATISTIEKONTWIKKELING

GEBRUIK VAN RIJKSREGISTERGEGEVENS VOOR HET MONITOREN VAN OEKRAÏENSE ONTHEEMDEN IN VLAANDEREN

INGRID SCHOCKAERT, REMBERT DE BLANDER,
LISA VAN LANDSCHOOT & JAN PICKERY

STATISTIEK
VLAANDEREN

INHOUD

Samenvatting	3
Inleiding	5
1 Administratieve vertraging op de Rijksregistergegevens	7
2 Beschikbare gegevens	11
3 Methode van herweging (tool 1)	13
3.1 Principe	13
3.2 Bepalen van de observatiebasis en veronderstellingen	14
3.3 Berekening van de kansen	15
3.4 Berekening van het aantal bewegingen	16
4 Methode van Microsimulatie (tool 2)	19
4.1 Principe	19
4.2 Observatiebasis en startpunt van de simulatie	20
4.3 Modelleren van de immigraties	22
4.4 Modelleren van buitenlandse emigraties en verhuisbewegingen vanuit en naar elk van de gewesten	22
4.5 Het simulatiemodel	24
5 Resultaten en validatie	27
Besluit	32
Literatuurlijst	35
Bijlagen	37

SAMENVATTING

Sinds 24 februari 2022 woedt er oorlog in Oekraïne, wat een vluchtelingenstroom naar Europa en België op gang bracht. Om deze crisis aan te pakken richtte de Vlaamse overheid de Taskforce “*Vlaanderen helpt Oekraïne*” op, die de opvang en integratie van Oekraïense ontheemden coördineerde. Voor een doeltreffend beleid waren snelle, betrouwbare gegevens nodig. Het Belgisch statistiekbureau Statbel leverde daarvoor gegevens uit het Rijksregister aan de Vlaamse Statistische Autoriteit (VSA), die ze analyseerde ter ondersteuning van het beleid.

GEBRUIK VAN RIJKSREGISTERGEGEVENS VOOR HET MONITOREN VAN OEKRAÏENSE ONTHEEMDEN IN VLAANDEREN

INGRID SCHOCKAERT, REMBERT DE BLANDER,
LISA VAN LANDSCHOOT & JAN PICKERY

Het Rijksregister is een betrouwbare informatiebron over Oekraïense ontheemden, dat bovendien systematisch wordt aangevuld en gewijzigd zodra mensen aankomen in of vertrekken uit een gemeente. Vóórdat een registratie in de gemeente kenbaar wordt gemaakt aan het Rijksregister, moet er echter een administratieve procedure worden doorlopen, die door de werklast bij gemeenten tijdens de crisis soms grote vertraging oplep. Deze vertraging leidt tot een onderschatting van het actuele aantal ontheemden in recente Rijksregisterextracties.

De VSA ontwikkelde 2 monitoringtools om de onvolledigheid door administratieve vertraging te corrigeren. De methode van Herweging gebruikt het principe van terugrekening (back-calculation) om het werkelijke aantal migratiebewegingen, en dus ook het werkelijke aantal ontheemden, te schatten. Ze steunt hiervoor op de Horvitz-Thompson schatter waarbij aan elke migratiebeweging een gewicht wordt toegekend op

basis van de geschatte kans dat die beweging al gemeld is aan het Rijksregister.

De van Microsimulatie vertrekt van “levenslopen” van individuen vanaf hun eerste registratie in een Belgische gemeente en bestaande uit opeenvolgende woonplaatsen en een eventuele emigratie naar het buitenland. Aan de hand van geschatte verhuis- en emigratiekansen en uitgaand van een constante immigratiestroom naar België, wordt een dag-per-dag simulatie gemaakt van het gewest van verblijf van de ontheemden.

De methoden werden toegepast op de Rijksregisterdoorsnede van 26 mei 2023 en gevalideerd met de veel latere extractie van 14 februari 2024, die vrijwel alle bewegingen tot 26 mei 2023 bevat. Beide methoden onderschatten vaak het aantal migratiebewegingen, maar de afwijkingen waren groter bij de Microsimulatie. De methode van Herweging gaf daarom een accuratere inschatting van het aantal ontheemden dan de Microsimulatie.

INLEIDING

Sinds 24 februari 2022 is er oorlog in Oekraïne. Die bracht een stroom aan vluchtelingen richting Europa, en dus ook richting België, op gang. Om aan deze massale toestroom het hoofd te bieden, besliste Europa om de richtlijn rond tijdelijk ontheemden (richtlijn 2001/55/EG van de Raad van 20 juli 2001) te activeren voor vluchtelingen uit Oekraïne, waardoor de zware procedure gerelateerd aan een klassieke aanvraag tot internationale bescherming (asielprocedure) kon worden vermeden. Momenteel genieten meer dan 6 miljoen mensen uit Oekraïne tijdelijke bescherming in Europa (UNHCR, 16 januari 2025). In België heeft de Dienst Vreemdelingenzaken tussen het begin van de oorlog en eind 2024 meer dan 91.806 attesten tijdelijke bescherming afgegeven aan Oekraïense ontheemden (24 december 2024), waarvan meer dan 44.000 in de eerste 3 maanden van de crisis (Statbel).

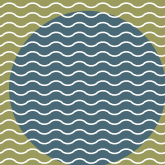
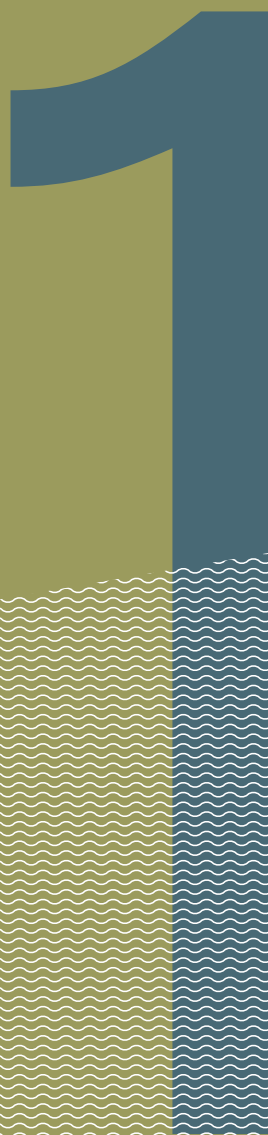
Om de vluchtelingen crisis het hoofd te bieden, werd door de Vlaamse overheid de Taskforce “*Vlaanderen helpt Oekraïne*” opgericht. Deze taskforce coördineerde in eerste instantie de opvang en later ook de integratie van Oekraïense ontheemden in Vlaanderen. Daarbij was het cruciaal om snel te beschikken over correcte en actuele gegevens over zowel het aantal ontheemden als hun opvangbehoeften. Hiervoor stelde het Belgische statistiekbureau Statbel systematisch gegevens uit het Rijksregister ter beschikking aan de Vlaamse Statistische Autoriteit (VSA). De VSA analyseerde deze gegevens om het beleid van de Taskforce “*Vlaanderen helpt Oekraïne*” te ondersteunen.

Het Rijksregister is een betrouwbare informatiebron over Oekraïense ontheemden, dat bovendien systematisch wordt aangevuld en gewijzigd zodra mensen aankomen in of vertrekken uit een gemeente. Met opeenvolgende doorsneden uit het Rijksregister kan de evolutie van de ontheemden tijdens de crisis dan ook systematisch worden opgevolgd.

Vóórdat een registratie in de gemeente kenbaar wordt gemaakt aan het Rijksregister, moet er echter een administratieve procedure worden doorlopen, waaronder een verificatie door de wijkagent van de effectieve verblijfplaats en de gezinssamenstelling.

Merk op dat deze procedure niet alleen wordt doorlopen bij aankomst uit het buitenland, maar ook bij elke herhuisvesting en vertrek naar het buitenland. Er kan dus heel wat tijd verstrijken tussen registratie in de gemeente en het moment waarop deze informatie effectief zichtbaar is in het Rijksregister. Dat maakte recente gegevens over aantallen en spreiding van ontheemden onvolledig en bemoeilijkte de planning van opvang en voorzieningen.

Dit rapport bespreekt 2 monitoringstools die de VSA ontwikkelde om de onvolledigheid van doorsneden uit het Rijksregister te remediëren. Het rapport is opgebouwd als volgt. In deel 2 wordt ingegaan op het effect van de administratieve vertraging op de gegevens van het Rijksregister en worden de 2 monitoringstools geïntroduceerd. In deel 3 wordt dieper ingegaan op de beschikbare gegevens. De methode van Herweging (tool 1) wordt uiteengezet onder deel 4. De methode van Microsimulatie (tool 2) wordt uiteengezet onder deel 5. De resultaten van de schattingen van het aantal bewegingen en het aantal ontheemden alsook de validatie van beide methoden, worden uiteengezet in deel 6 van het rapport. Ten slotte wordt nog kort ingegaan op de voor- en nadelen van de methoden en op hun mogelijke verdere toepassing en ontwikkeling.

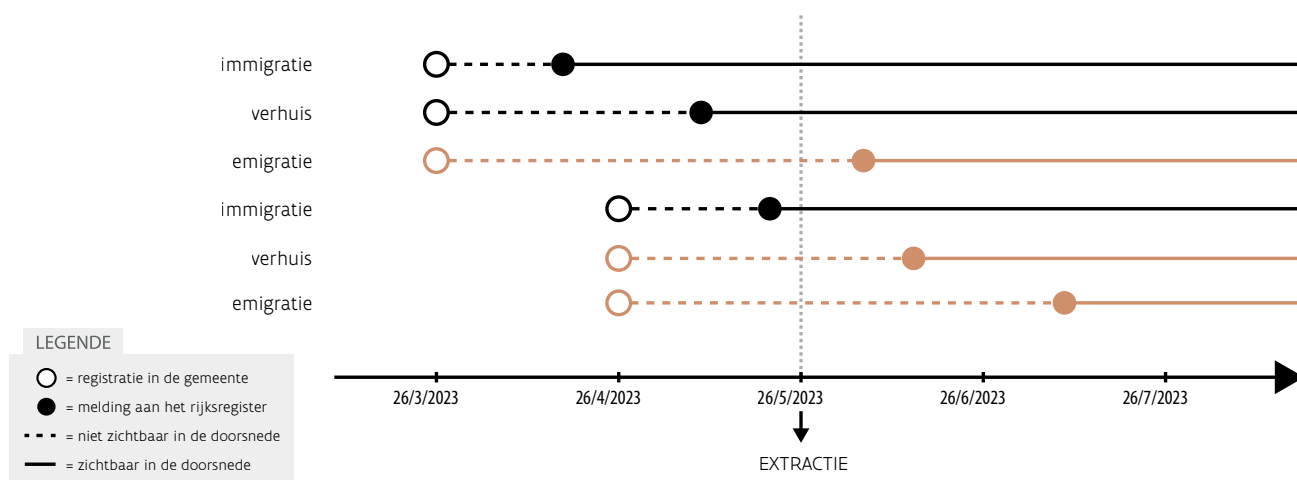


ADMINISTRATIEVE VERTRAGING OP DE RIJKSREGISTERGEGEVENS

Om de evolutie van het aantal ontheemden en hun spreiding over het Belgische grondgebied te monitoren, stelde Statbel systematisch doorsnedes uit het Rijksregister ter beschikking aan de VSA. Door de hoge instroom en frequente herhuisvesting tijdens de crisis, liep de werklast bij gemeenten op en liepen de administratieve procedures soms grote vertraging op. Daardoor waren er meer registraties in de gemeenten dan er in de doorsnedes van het Rijksregister zichtbaar waren. Hoe recenter de extractie, hoe groter de kans dat ze onvolledig is.

Dit wordt geïllustreerd in Figuur 1. Het fictieve voorbeeld betreft een doorsnede van het Rijksregister op 26 mei 2023. Er vonden 6 beweging plaats. Bewegingen kunnen immigraties uit het buitenland zijn, verhuisbewegingen binnen België of emigraties naar het buitenland. De open cirkels stellen de datum van registratie in de gemeente voor, de volle cirkels de datum van hun melding aan het Rijksregister. Vanaf dan zijn bewegingen “zichtbaar” in de doorsnede van 26 mei 2023. Voor 3 bewegingen vallen de meldingen vóór het moment van doorsnede, 3 bewegingen werden erna gemeld (oranje) en komen er dus niet in voor. Voor deze laatste bewegingen is geen enkele informatie beschikbaar.

Figuur 1. Administratieve vertraging op meldingen aan het Rijksregister



Dit rapport bespreekt 2 monitoringstools die de VSA ontwikkelde om de onvolledigheid van doorsnedes uit het Rijksregister te remediëren. Bij iedere nieuwe doorsnede van Statbel schatten de tools:

- het werkelijke aantal tijdelijk ontheemden in België,
- hun verdeling over de 3 gewesten.

De doelstelling betreft een zogenaamde *nowcasting*: het voorspellen van de huidige situatie die nog onbekend is omdat de benodigde gegevens nog niet beschikbaar zijn (Banbura e.a., 2013; Wang e.a., 2017). Op basis van het aantal *waargenomen* bewegingen – immigraties, emigraties en verhuisbewegingen – sinds het begin van de crisis tot aan de meest recente doorsnede, werd een schatting gemaakt van het *werkelijke* aantal.

Een eerste tool is gebaseerd op het principe van terugrekening (*back-calculation*) (Brookmeyer & Gail, 1988). De term verwijst naar de reconstructie van niet-geobserveerde gebeurtenissen uit patronen van gebeurtenissen die wel al geobserveerd zijn. Dit principe wordt vaak toegepast in epidemiologische studies. Besmettingen zelf zijn immers niet zichtbaar, maar kunnen wel worden afgeleid uit de (daarop volgende) ziekte-uitbraken (Heisterkamp e.a., 1989; Harris, 1990; Donker e.a., 2011, Höhle & der Heiden e.a., 2014, van de Kastele e.a., 2019, Salmon e.a., 2015, Bastos e.a., 2019 en Günther e.a., 2021). Andere contexten waarin observatievertragingen voorkomen betreffen onderhouds- en verzekeringscontracten (Crevecoeur e.a. 2019), waar mensen pas een tijd na een panne of een ongeluk een dossier aan de verzekeraar overmaken. In dit rapport wordt het aantal registraties van immigraties, emigraties en verhuisbewegingen van de ontheemden die nog niet zichtbaar zijn door de ‘administratieve vertraging’ afgeleid uit de patronen in registraties die wel al zichtbaar zijn in het Rijksregister. De methode in dit rapport bouwt verder op de theorie van herweging.

Een tweede tool situeert zich in de traditie van dynamische microsimulatie. Dynamische microsimulatie is een techniek waarmee individuele gedragingen worden gesimuleerd in de toekomst, uitgaande van geobserveerde gedragingen uit het verleden (O’Donoghue, 2001). De levenslopen van mensen worden als het ware stap voor stap geconstrueerd om te zien hoe hun situatie in de loop van de tijd verandert. Deze methode wordt onder andere gebruikt om inkomensverschillen, bevolkingsgroei, verkeersstromen en gezondheidsontwikkelingen in kaart te brengen en te simuleren in de toekomst. Een overzicht van benaderingswijzen kan worden teruggevonden in Spielauer (2006), Willekens (2009) en Li & O’Donoghue (2013). In dit rapport wordt deze techniek toegepast om de immigraties, verhuisbewegingen en emigraties van ontheemden die in het verleden geobserveerd werden, te simuleren tijdens de periode van de administratieve vertraging.

In dit rapport wordt verwezen naar de 2 tools als “de methode van Herweging” en “de methode van Microsimulatie”. Beide tools werden ontwikkeld om telkens opnieuw gebruikt te kunnen worden bij elke nieuwe beschikbare doorsnede van het Rijksregister. De eerste methode werd

ontwikkeld aan het begin van de crisis, toen er enkel behoefte was aan een (snelle) accurate telling van het actuele aantal ontheemden. Naarmate de situatie evolueerde, kwam echter de noodzaak aan betrouwbare vooruitzichten naar voren, om zo de planning rond huisvesting van nieuwkomers te ondersteunen. Daarom werd overgestapt op de ontwikkeling van een microsimulatiemodel, dat niet alleen geschikt is voor nowcasting, maar ook projecties op langere termijn mogelijk maakt.

Beide methoden worden in dit rapport gedemonstreerd op basis van de doorsnede van gegevens uit het Rijksregister van 26 mei 2023. Er werd gekozen voor deze doorsnede omdat hieraan enerzijds voldoende extracties voorafgaan waarop de schattingen gebaseerd kunnen worden (vanaf 14 juli 2022), en anderzijds nog voldoende extracties erop volgen om een validatiebestand te creëren met informatie uit recentere extracties (tot 14 februari 2024). Recente extracties bevatten immers met heel grote waarschijnlijkheid *alle* bewegingen van *alle* ontheemden op 26 mei 2023, ook deze die pas later werden gemeld aan het Rijksregister. Met dit volledige bestand als vergelijkingsbasis wordt de accuraatheid van beide methoden als monitoringstool voor het opvolgen van de Oekraïense ontheemden gevalideerd.



BESCHIKBARE GEGEVENS

Sinds 15 juli 2022 ontvangt de Vlaamse Statistische Autoriteit (VSA) van Statbel periodiek een doorsnede van Rijksregistergegevens over de Oekraïense ontheemden. Tot september 2022 waren de extracties ongeveer wekelijks beschikbaar, daarna werd de frequentie teruggebracht tot tweewekelijks en sinds februari 2023 naar maandelijks. De laatste ontvangen extractie dateert van 14 februari 2024.

De extracties bevatten gepseudonomiseerde gegevens over leeftijd, geslacht en huishoudsamenstelling van elk individu bij de eerste inschrijving in het Rijksregister. Daarnaast bevatten ze ook informatie van elk van de “bewegingen” van de individuen sinds aankomst in België. Er wordt een onderscheid gemaakt tussen 3 types bewegingen: immigraties vanuit het buitenland, verhuisbewegingen binnen België en emigraties naar het buitenland¹. Voor elke beweging wordt de datum ervan bij registratie in de gemeente weergegeven.

Voor immigraties uit het buitenland is de gemeente, en dus ook het gewest, van aankomst gegeven, voor emigraties de gemeente (gewest) van vertrek. Voor elke verhuisbeweging zijn de gemeenten (gewesten) van vertrek en aankomst gegeven. Er werden ook 43 overlijdens geregistreerd, maar omwille van het kleine aantal werd hiermee in de verdere analyses geen rekening gehouden. Geboorten werden niet in de extracties opgenomen.

Merk op dat de datum waarop een beweging wordt gemeld aan het Rijksregister ontbreekt in de gegevens. De datums van opeenvolgend extracties worden daarom gebruikt om een *meldingsinterval* te construeren. Dit interval wordt gedefinieerd als de periode tussen de datum van de laatste extractie waarin een beweging nog niet voorkwam en de datum van de extractie waarin een beweging voor het eerst voorkwam. De melding aan het Rijksregister ligt tussen beide datums in.

¹ De immigraties betreffen eerste registratie, herregistratie na verblijf in Oekraïne of in een ander land, en na een ambtshalve schrapping. De emigraties zijn bewegingen richting Oekraïne of een ander land en de ambtshalve schrappingen. Een verhuis betreft elke verandering van adres binnen België.



METHODE VAN HERWEGING (TOOL 1)

3.1 PRINCIPE

Het aantal tijdelijk ontheemden in de Belgische gewesten wordt berekend door het aantal internationale en intergewestelijke immigraties sinds het begin van de crisis tot aan de datum van de extractie (26 mei 2023) op te tellen, en daar het aantal internationale en intergewestelijke emigraties van af te trekken. Er moet dus een inschatting gemaakt worden van het aantal migratiebewegingen.

Door administratieve vertraging tussen het moment van een beweging geregistreerd door de gemeente en de melding ervan aan het Rijksregister, ligt het aantal bewegingen in de Rijksregisterextractie echter lager dan het werkelijke aantal. De bewegingen die al aan het Rijksregister zijn gemeld, kunnen daarom beschouwd worden als een soort van *steekproef* van alle bewegingen die tot op de datum van extractie hebben plaatsgevonden.

Vanuit deze optiek kan het “werkelijke” aantal bewegingen worden geschat op basis van het geobserveerde aantal gemelde bewegingen via de Horvitz-Thompson schatter (1952), gekend uit de literatuur over steekproefbewegingen. Aan elke al gemelde beweging i wordt een gewicht w_i toegekend. Dat gewicht is het omgekeerde van de kans p_i dat beweging i al gemeld werd aan het Rijksregister op de datum van extractie. Het totaal aantal bewegingen N is dan gelijk aan de som van deze gewichten w_i over alle gemelde bewegingen:

$$N = \sum_i w_i = \sum_i 1/p_i$$

In Figuur 1 is bijvoorbeeld de helft van de bewegingen al gemeld op de datum van extractie en de andere helft niet. De kans dat een beweging al gemeld is, is dus gelijk aan 0,5. Het totaal aantal bewegingen kan daardoor berekend worden als

$$\frac{1}{0.5} + \frac{1}{0.5} + \frac{1}{0.5} = 6$$

In werkelijkheid zijn het aantal nog te melden bewegingen en dus ook de meldingskans niet gekend. Deze kans wordt daarom geschat op basis van informatie over de bewegingen die wel al aan het Rijksregister gemeld werden.

De kans dat een beweging is opgenomen in het Rijksregister hangt op de eerste plaats af van de dag waarop de beweging plaatsvond. Hoe recenter de beweging plaatsvond, hoe kleiner de kans dat de beweging al gemeld werd aan het Rijksregister omwille van de administratieve vertraging. Bewegingen

die al langer geleden plaatsvonden, hebben daarentegen een heel grote kans om al in het Rijksregister zichtbaar te zijn. Vanaf een bepaalde tijdsspanne is die kans hoogstwaarschijnlijk zelfs 100%.

Laat D de Datum zijn van een beweging geregistreerd door de gemeente, en laat M de datum zijn waarop een beweging gemeld werd aan het Rijksregister. Vanaf M is een beweging dus zichtbaar. Laat e de datum voorstellen van de Extractie waarop we het aantal ontheemden in een gewest willen kennen. Laat verder X bijkomende kenmerken van de bewegingen voorstellen (bijvoorbeeld kenmerken van de gemeenten waarin of waartussen de beweging plaatsvond). De kans dat een beweging i gemeld is aan het Rijksregister op een bepaalde datum e (de dag van extractie) is gelijk aan de kans dat M kleiner is dan e , gegeven de datum van de beweging d en de achtergrondkenmerken x :

$$p_i = P(M \leq e | D = d, X = x) = f(e, d, x)$$

Volgens basisprincipes van kansrekening kan deze gelijkheid worden geschreven als:

$$p_i = 1 - P(M > e | D = d, X = x) = 1 - P(M - d > e - d | D = d, X = x)$$

De kansenfunctie neemt hierbij de vorm aan van een survival-functie $S(t | D, X) = P(T > t | D, X)$, waarbij T de tijd voorstelt vanaf de registratie bij de gemeente tot de melding van de beweging aan het Rijksregister. Merk op dat voor de berekening van kans p_i deze survival-functie geëvalueerd moet worden op basis van de datum van extractie en de datum waarop de beweging plaatsvond, $t = e - d$.

3.2 BEPALEN VAN DE OBSERVATIEBASIS EN VERONDERSTELLINGEN

Het probleem is dat de survival-functie $S(t | D, X)$ niet zomaar kan worden geschat op de geobserveerde data. De geobserveerde data bevatten immers enkel de bewegingen al gemeld aan het Rijksregister en bewegingen met een korte meldingstijd zijn daarin per definitie oververtegenwoordigd. Dit leidt tot een onderschatting van de survival-functie en een overschatting van de gewichten.

Om dit probleem te vermijden, gebruiken we niet de datum van extractie als het eindpunt van observatie, maar wel een moment in de tijd waarop we zeker zijn dat alle registraties in de gemeenten ook al gemeld werden aan het Rijksregister. Met andere woorden, het eindpunt moet worden teruggezet met een periode die langer is dan de langste geobserveerde observatievertraging (Casper & Cook, 2012). Zo goed als geen enkele geobserveerde vertraging was langer dan 150 dagen. Het eindpunt van de observatieperiode werd daarom gezet op 26 december 2022.

Om de administratieve vertraging van bewegingen na 26 december 2022 te voorspellen moet worden aangenomen dat de vertraging bij deze bewegingen dezelfde trends vertonen als de vertraging bij de bewegingen vóór 26 december 2022. Met andere woorden, er wordt aangenomen dat de survival-functie niet afhangt van de datum van de registratie van de beweging: $S(t|D,X) = S(t|X)$.

Om deze aanname te toetsen, werd in het schattingsmodel de datum van registratie van de bewegingen expliciet opgenomen. Meer specifiek wordt gebruik gemaakt van dummy-variabelen voor elke maand van registratie (zie tabellen 1 tot 3 in bijlage 1). Voor geen van de onderzochte bewegingstypes – immigratie, emigratie of verhuisbewegingen – werd echter een duidelijk verband gevonden tussen de maand van registratie en de duur van administratieve vertraging. Hoewel deze vaststelling strikt genomen alleen geldt voor de periode waarop de schatting betrekking heeft, ondersteunt deze vaststelling wel de veronderstelling dat we de kansverdeling van de duur van administratieve vertraging kunnen extrapoleren naar de periode na 26 december 2022. Bovendien is het aannemelijk dat, naarmate de crisis stabiliseerde, er ook een stabielere en uniformere administratief verwerkingsritme werd bereikt.

De veronderstellingen dat (1) de administratieve vertraging maximaal 150 dagen is en (2) een constante verdeling volgt doorheen de tijd, laten toe om de duur tussen de registratie en de melding aan het Rijksregister te modelleren op alle gemelde bewegingen die werden geregistreerd in een gemeente tussen 15 juni 2022 (datum van eerste extractie) en 26 december 2022 (150 dagen voor de datum van extractie), en dit model te gebruiken om de administratieve vertraging te schatten voor bewegingen na 26 december 2022.

3.3 BEREKENING VAN DE KANSEN

Om de kansen p_i te berekenen werden survival-functies $S(t|X)$ geschat voor elk type beweging afzonderlijk: immigraties uit het buitenland, verhuisbewegingen tussen gewesten, en emigraties naar het buitenland. Er werd gebruik gemaakt van een parametrische survival-functie, op basis van een Weibull distributie (Therneau, 2020). De Weibull-distributie werd gekozen vanwege haar vermogen om zowel toenemende als afnemende kansverdelingen in verschillende vormen te modelleren. Als monitoringtool moest het model snel inzetbaar zijn op elk moment tijdens de crisis, zonder dat vooraf telkens de validiteit van de veronderstelde distributie geverifieerd hoefde te worden.

De schattingen houden rekening met de volgende covariaten. De duur tot de melding van een immigratie uit het buitenland wordt geschat conditioneel op gewest van aankomst, of het een immigratie naar een centrumstad betreft en de maand van immigratie. De duur tot de melding van een

intergewestelijke verhuisbeweging wordt geschat conditioneel op gewest van vertrek en op gewest van aankomst, of het een beweging van en naar een centrumstad betreft en de maand van de binnenlandse migratie. De duur tot de melding van een emigratie naar het buitenland tenslotte, wordt geschat volgens gewest van vertrek, of het een emigratie vanuit een centrumstad betreft en de maand van vertrek. De geografische covariaten weerspiegelen de variatie in de administratieve werking tussen de 3 gewesten en tussen grote steden en kleinere gemeenten. De registratiemaand verifieert de aanname van een constante survival-functie in de tijd.

Merk ook op dat in de modellering rekening moeten worden gehouden met interval censoring omdat de exacte dag van een melding niet gekend is (Sun, 2006). Enkel het tijdsinterval van de melding tussen de datum van de laatste extractie waarin de beweging niet zichtbaar was en de eerste extractie waarin ze wel zichtbaar was, is gekend. De schattingsresultaten zijn terug te vinden in tabellen 1 tot 3 van Bijlage 1.

3.4 BEREKENING VAN HET AANTAL BEWEGINGEN

De survival-functie laat toe om voor elke beweging in de Rijksregisterdoorsnede de kans te schatten dat deze beweging al gemeld is, voor elke dag sinds het begin van de crisis, gegeven de datum van registratie in de gemeente en zijn karakteristieken. Eens deze kans bekend is, kan deze gebruikt worden als weging om het totaal aantal bewegingen te schatten op het moment van de extractie.

Standaardfouten op het aantal herwogen bewegingen worden geschat aan de hand van een Monte-Carlo simulatie van 200 trekkingen uit de verdelingen van de geschatte parameters (Robert & Casella, 2000).



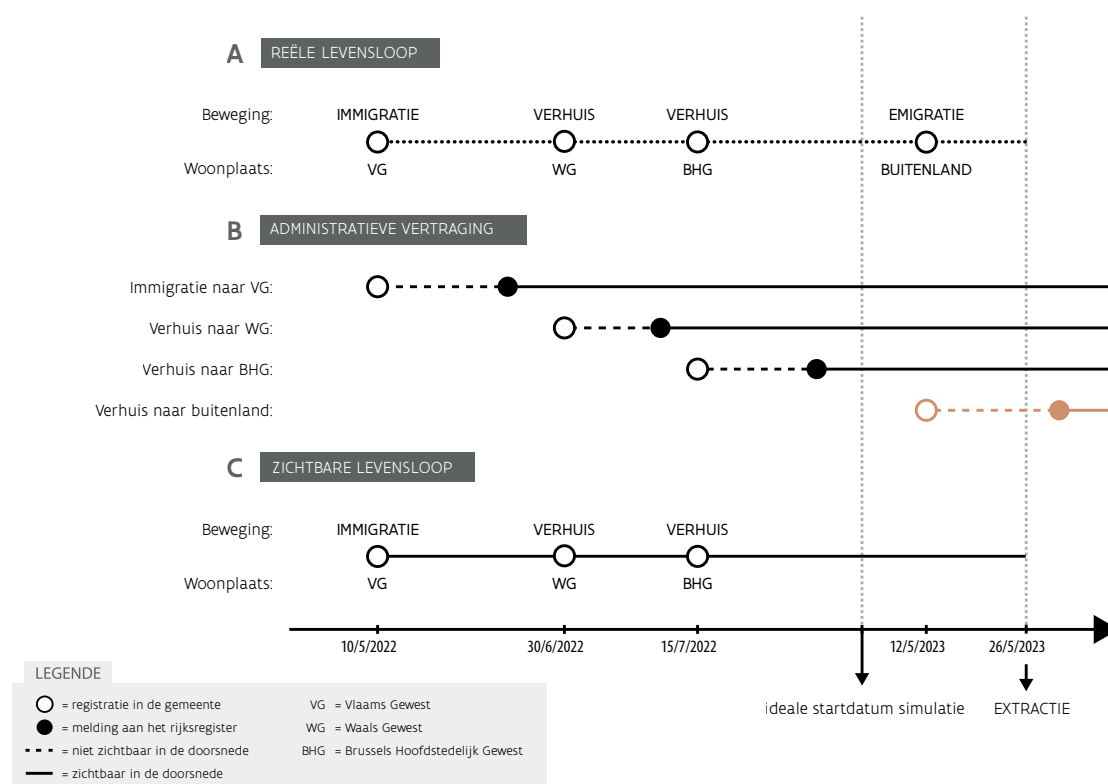
METHODE VAN MICROSIMULATIE (TOOL 2)

4.1 PRINCIPE

Zoals bij de methode van Herweging, wordt bij de methode van Microsimulatie het aantal tijdelijk ontheemden in de Belgische gewesten berekend door het aantal internationale en intergewestelijke immigraties sinds het begin van de crisis tot aan de datum van de extractie (26 mei 2023) op te tellen en daar het aantal internationale en intergewestelijke emigraties van af te trekken. Er moet dus een inschatting gemaakt worden van het aantal bewegingen.

De methode van microsimulatie gaat uit van de “levenslopen” van ontheemden, opgebouwd uit hun opeenvolgende bewegingen tussen verblijfplaatsen. Figuur 2 stelt zulk een fictieve levensloop voor (deel A in de figuur). De datum van immigratie naar België (10 mei 2022) en de eerste geregistreerde verblijfplaats in België vormen het beginpunt van de levensloop. Het vertrek naar het buitenland (12 mei 2023) markeert het eindpunt. Elke verhuisbeweging tussen gewesten (op 30 juni 2022 en op 15 juli 2022) betekent het begin van een nieuwe verblijfsperiode in een (ander) gewest.

Figuur 2. Fictief voorbeeld van een reële en geobserveerde “levensloop van bewegingen” van ontheemden



Op basis van een doorsnede van het Rijksregister (26 mei 2023), waarin elke beweging met hun datum, oorsprong en bestemming sinds het begin van de crisis wordt opgenomen, is het in principe eenvoudig om deze levenslopen te construeren. Echter, bewegingen kortbij de datum van extractie zijn vaak niet zichtbaar omwille van de administratieve vertraging. Deel B in figuur 2 illustreert het effect van deze vertraging op de geobserveerde levensloop. De immigratie naar België en de 2 verhuisbewegingen werden gemeld aan het Rijksregister vóór de datum van de doorsnede op 26 mei 2023, het vertrek naar het buitenland (12 mei 2023) echter nog niet. Bijgevolg heeft de zichtbare levensloop slechts 3 toestanden (VG, WG en BHG) en lijkt de duur van verblijf in het Brusselse Gewest te lopen tot het moment van de doorsnede (Figuur 2C). Kortom, de administratieve vertraging zorgt ervoor dat op het moment van de deeldoorsnede noch de emigratie, noch de resulterende woonplaats (buitenland) en evenmin de duur in de vorige woonplaats correct worden geobserveerd.

Het doel van de microsimulatie is om, op basis van het deel van de 'levensloop' van bewegingen dat geobserveerd werd, een simulatie te maken van het verdere verloop tijdens de periode dat nog niet alle bewegingen gemeld werden aan het Rijksregister. Deze procedure impliceert verschillende stappen.

Eerst moet worden vastgesteld tot welk tijdstip de bewegingen als volledig gemeld kunnen worden beschouwd en vanaf wanneer simulatie nodig is om ontbrekende bewegingen aan te vullen. Dat moment is namelijk niet waarneembaar. Vervolgens wordt, vanaf deze startdatum voor simulatie, een inschatting gemaakt van het dagelijkse aantal immigraties vanuit het buitenland om zo simulatiegewijs de pool van vluchtelingen aan te kunnen vullen met de immigranten die nog niet zichtbaar zijn in het Rijksregister. Ten derde worden bewegingen die geregistreerd werden vóór de startdatum gebruikt om schattingen te maken van internationale emigratie- en binnenlandse verhuiskansen vanuit en naar elk gewest alsook van de duur van verblijf in elk gewest om voor alle vluchtelingen (zichtbare en gesimuleerde immigranten) de levensloop verder te kunnen simuleren.

4.2 OBSERVATIEBASIS EN STARTPUNT VAN DE SIMULATIE

De eerste stap in de simulatieprocedure is het bepalen van de startdatum van de simulatie. Hoe verder deze startdatum teruggezet wordt in de tijd, hoe korter de observatieperiode voor de schattingen en hoe langer de periode die moet overbrugd worden door simulatie. Een startdatum kort bij de datum van extractie beperkt de simulatieperiode, maar kan daarentegen leiden tot een vertekende observatiebasis en kan ervoor zorgen dat niet alle niet-geobserveerde bewegingen gesimuleerd worden.

In het fictieve voorbeeld van figuur 2 ligt de “ideale startdatum” voor simulatie vlak voor het vertrek naar het buitenland, de enige beweging die niet werd geobserveerd in de doorsnede en dus nog moet worden gesimuleerd om de levensloop te vervolledigen. Een latere startdatum impliceert dat de emigratiebeweging niet of te laat zou worden gesimuleerd. Een te vroege startdatum zou maken dat ook de 2de verhuisbeweging zou moeten worden gesimuleerd (en uitgesloten wordt van de observatiebasis voor de schattingen), terwijl die in de extractie eigenlijk geobserveerd werd.

Het is dus belangrijk om de *recentst* mogelijke startdatum te bepalen vóór dewelke alle of toch zo veel mogelijk bewegingen geobserveerd werden. Indien we de startdatum terugzetten in de tijd met een periode gelijk aan de langste administratieve vertraging, kunnen we vrij zeker zijn dat alle bewegingen vóór die startdatum effectief geobserveerd werden - in de veronderstelling dat de observatievertraging constant blijft (of afneemt) in de tijd (Casper & Cook, 2012). In de praktijk geeft deze regel echter aanleiding tot een periode van 150 dagen die zou moeten overbrugd worden door simulatie (zoals al uitgelegd in deel 3.2).

Om deze periode te beperken, werd voor een andere strategie gekozen. Er werd gebruik gemaakt van de schattingen van observatievertragingen bij immigratie, emigratie en verhuisbewegingen, zoals beschreven in deel 4. De startdatum voor simulatie werd teruggezet in de tijd met een periode die overeenkomt met 95% van de geschatte vertragingen van de beweging met de grootste gemiddelde observatievertraging. In het geval van de huidige oefening, is dat immigratie. Voor immigratie bleek dat 95% van de geschatte vertragingen korter was dan 100 dagen.

Bij het terugzetten in de tijd van de startdatum voor simulatie met 100 dagen wordt een kans getolereerd van 5% dat een immigratie niet wordt geobserveerd en ook niet (of te laat) wordt gesimuleerd. Voor emigratie en verhuis ligt deze kans net iets lager. Hiertegenover staat echter dat, in vergelijking met het terugzetten van de startdatum met een tijdsperiode gelijk aan de langst geobserveerde administratieve vertraging, 51 dagen minder moeten worden gesimuleerd en de geobserveerde bewegingen uit deze periode in de observatiebasis voor de schattingen worden opgenomen. De simulatie wordt gestart op 100 dagen voor de extractie van 26 mei 2023: dat is vanaf 15 februari 2023. De observatiebasis voor de schatting van immigraties, verhuisbewegingen en emigraties betreft de bewegingen die geregistreerd werden vóór 15 februari 2023.

4.3 MODELLERING VAN DE IMMIGRATIES

De datum van immigratie en de eerste woonplaats initieert de levensloop van de ontheemden. Tot 15 februari 2023 zijn de immigraties, hun datum en gewest van bestemming verondersteld gekend. Na deze datum moeten ze worden gesimuleerd.

De stromen van ontheemden naar België worden in grote mate bepaald door het verloop van de oorlog in Oekraïne, het verloop van de opvang in naburige landen en Europese overeenkomsten over de spreiding van ontheemden. Deze tendensen zijn moeilijk voorspelbaar en kwantificeerbaar. In afwezigheid van elke redelijke basis voor het bouwen van een hypothese rond immigratie, wordt uitgegaan van een constante instroom en wordt het aantal dagelijkse immigraties vanaf 15 februari 2023 gelijkgesteld aan het daggemiddelde van een bepaalde observatieperiode voorafgaand aan 15 februari 2023. Omdat we een monitoringtool ontwikkelen, hebben we gekozen voor een observatieperiode van 120 dagen om toevallige schommelingen door plotselinge oorlogsgebeurtenissen of feestdagen als kerstmis en nieuwjaar te beperken. Tijdens deze 120 dagen tonen de gegevens een gemiddelde van 61 immigraties per dag.

4.4 MODELLERING VAN BUITENLANDSE EMIGRATIES EN VERHUISBEWEGINGEN VANUIT EN NAAR ELK VAN DE GEWESTEN

Om de verhuisbewegingen en emigraties naar het buitenland te simuleren vanaf 15 februari 2023 wordt een schatting gemaakt van de dagelijkse kansen op een verhuis naar één van de 3 gewesten of naar het buitenland, en dit vanuit elk van de gewesten en gegeven de eigenschappen van elk individu. De gegevensbasis voor deze berekening betreft alle bewegingen geregistreerd tussen 1 juli 2022 en 15 februari 2023. De eerste 4 maanden van de crisis werden niet gebruikt omdat de bewegingen toen vaak het gevolg waren van door het beleid vastgelegde (her)-huisvestingsprocedures en daarom een andere logica volgden.

De schatting van de kansen gebeurt in 2 stappen. In een eerste stap wordt een schatting gemaakt van de duur tot verhuis of emigratie naar het buitenland, gegeven de eigenschappen van individuen en het gewest van verblijf. Daarna wordt in een 2de stap voor elke binnenlandse verhuis het gewest van bestemming bepaald.

In **stap 1** wordt gebruik gemaakt van een parametrisch competing risk model met Weibull-distributie voor de duur tot binnenlandse verhuis of emigratie naar het buitenland. De Weibull-distributie werd ook hier gekozen vanwege haar flexibiliteit en haar vermogen verschillende kansverdelingen te modelleren. Een competing risk model wordt gebruikt wanneer er sprake is van 2 gebeurtenissen - binnenlandse verhuis en emigratie naar het

buitenland - die elkaar wederzijds uitsluiten, maar waarbij de hypothese van onafhankelijkheid tussen gebeurtenissen onrealistisch of onverifieerbaar is (Tsiatis, 1975; Haller e.a., 2013). De kans dat iemand verhuist hangt samen met de kans dat deze persoon naar het buitenland vertrekt, en omgekeerd. Een verhuis van de eerste noodopvang naar een meer permanente huisvesting bijvoorbeeld wijst op een verdere integratie en hangt samen met de kans om langere tijd in België te verblijven.

Dit rapport volgt de benadering van Mahani & Sharabiani (2019), geïmplementeerd in het R-package CFC. Deze benadering schat verhuis en emigratie als afzonderlijke (parametrische) survival-functies, waarbij de ene beweging als censurerende gebeurtenis voor de andere wordt beschouwd. In tweede instantie worden de schattingsresultaten geïntegreerd tot dagelijkse oorzaaks-specifieke cumulatieve kansen op interne verhuis en emigratie naar het buitenland, gegeven de eigenschappen van de individuen.

Naast de traditionele demografische covariaten zoals leeftijd, geslacht, huishoudcompositie (bij eerste registratie), werden ook een dummy voor eerste woonplaats in een centrumstad, het aantal voorbije verhuisbewegingen, het gewest van waaruit de migratie vertrekt en de maand van immigratie in België in het model opgenomen. De schattingsresultaten zijn opgenomen in tabellen 4 en 5 van Bijlage 2.

De covariaat *maand van immigratie* werd opgenomen om de veronderstelling van temporele stabiliteit in de duur tot internationale emigratie of verhuis te toetsen. Immers, de schattingen op basis van de gegevens van vóór 15 februari 2023 dienen om de bewegingen te simuleren in de periode na deze datum tot het moment van de doorsnede van 26 mei 2023. Een centrale veronderstelling in het model is daarom dat de duur tot internationale emigratie en tot binnenlandse verhuisbewegingen constant blijven in de tijd. De maand van immigratie situeert de start van de levensloop - en dus ook de daarop volgende verblijfsperiodes - in de tijd.

Tabellen 4 en 5 van Bijlage 2 tonen aan dat de maand van immigratie positief geassocieerd is met de kans tot zowel emigratie naar het buitenland als binnenlandse verhuis. Vooral de relatie met binnenlandse verhuisbewegingen is uitgesproken, met een geschatte coëfficiënt van 0,32². De positieve associatie impliceert dat in de simulatie de duur tot emigratie en vooral de duur tot binnenlandse verhuis zal worden overschat en het aantal bewegingen onderschat.

2 Deze bevinding is in overeenstemming met het verloop van de opvangcrisis. In de beginfase werden ontheemden voornamelijk toegewezen aan gemeenten die hun beschikbare opvangcapaciteit doorgaven aan Fedasil. Deze eerste generaties ontheemden hadden vaak een beperkt sociaal netwerk en hun mogelijkheden om zelfstandig alternatieve huisvesting te vinden, was relatief gering. In de daaropvolgende maanden veranderde dit: de volgende generaties ontheemden vonden vaak zonder tussenkomst van de overheid onderdak bij familie of vrienden. Deze sociale netwerken speelden een belangrijke rol in het faciliteren van de doorstroom naar de private (huur)markt.

In *stap 2* voor de berekening van de kansen, wordt de bestemming bepaald van binnenlandse verhuisbewegingen: het Vlaamse, Waalse of Brusselse Hoofdstedelijk Gewest. Het beperkte aantal intergewestelijke verhuisbewegingen laat echter geen tijdsafhankelijke modellering toe. Daarom wordt de bestemming van deze interne migratiebewegingen vastgesteld op basis van de verhuiskansen tussen de gewesten over de hele periode tussen 1 juli 2022 en 15 februari 2023 - gegeven de woonplaats vóór verhuis en de individuele kenmerken leeftijd, geslacht, huishoudcompositie en eerste woonplaats in een centrumstad. In Tabel 6 van bijlage 3 worden de gemiddelde verhuiskansen per gewest van herkomst weergegeven. Door de dagelijkse kansen op interne verhuis te vermenigvuldigen met deze transitiekansen tussen gewesten, zijn ook de kansen op verhuis binnen en tussen gewesten gekend.

De volledige multinomiale kansenverdeling op emigratie naar het buitenland, naar het Vlaamse Gewest, het Waalse Gewest of het Brusselse Gewest is nu gekend voor elk van de individuen, gegeven de eigenschappen van individuen, de woonplaats vóór verhuis of emigratie *en* de duur sinds vorige beweging.

4.5 HET SIMULATIEMODEL

De startpopulatie voor de simulatie betreft de bevolking van Oekraïense ontheemden geregistreerd in een Belgische gemeente op 15 februari 2023, en hun eigenschappen: leeftijd, geslacht, huishoudcompositie bij eerste registratie, datum van eerste registratie, gewest, een dummy voor eerste woonplaats in een centrumstad, aantal voorbije verhuisbewegingen en datum van laatste beweging.

In eerste instantie worden elke dag volgend op 15 februari 2023 tot 26 mei 2023 61 immigranten aan deze startbevolking toegevoegd. Hun individuele eigenschappen alsook het gewest van bestemming worden willekeurig toegekend proportioneel met de verdeling in de observaties.

In tweede instantie worden de binnenlandse verhuisbewegingen en emigraties naar het buitenland gesimuleerd. Dit gebeurt op basis van de geschatte multinomiale kansverdeling: emigratie naar het buitenland of verhuizing naar het Vlaamse, Brusselse of Waalse Gewest, gegeven de individuele eigenschappen. De beslissing welk individu welke beweging maakt gebeurt door voor elk individu een willekeurige kans te genereren, die vervolgens wordt toegewezen aan de bijbehorende categorie van de geschatte verdeling.

Deze procedure maakt het mogelijk om, voor zowel de startpopulatie van 15 februari 2023 als voor dagelijks gesimuleerde internationale immigranten, per dag te bepalen of iemand binnen of tussen gewesten verhuist, of emigreert naar het buitenland. Het geheel van deze individuele trajecten

levert het totaal aantal dagelijkse bewegingen op en de opeenvolgende verblijfplaatsen van de hele bevolking.

Door deze hele procedure 200 keer te herhalen, wordt het gemiddelde aantal bewegingen verkregen alsook de variantie ervan over de Monte-Carlo-simulaties. Door bovendien in plaats van de puntschatting van de parameters uit de duuranalyses, een willekeurige trekking uit de parameterdistributie te nemen, wordt in de variantie van het aantal bewegingen ook rekening gehouden met de schattingsfouten op de modelparameters (Robert & Casella, 2000).



De in deel 3 en 4 voorgestelde methoden resulteren in een schatting van het aantal internationale en intergewestelijke migraties vanaf het begin van de crisis tot op de datum van de Rijksregisterdoorsnede van 26 mei 2023. Door de geschatte internationale en intergewestelijke immigraties naar de gewesten te verminderen met het aantal internationale en intergewestelijke emigraties wordt een inschatting gemaakt van het aantal ontheemden.

Om de 2 voorgestelde methoden voor now-casting te kunnen gebruiken als een monitoringtool, is het echter noodzakelijk een idee te hebben in hoeverre ze het *werkelijke* aantal bewegingen en ontheemden kunnen benaderen. Een validatie van beide methoden dringt zich daarom op.

Op het moment van de extractie van 26 mei 2023 bevatten de gegevens enkel de bewegingen die al gemeld waren aan het Rijksregister. Echter, de nog niet gemelde bewegingen werden beetje bij beetje zichtbaar in de daaropvolgende doorsneden. Op basis van informatie uit de meest recente, beschikbare extractie van 14 februari 2024, kunnen daarom de bewegingen in de extractie van 26 mei 2023 vervolledigd worden en een bestand worden opgesteld van *alle* bewegingen en ontheemden. Door de aantallen in dit volledige bestand te vergelijken met de resultaten van de schattingen wordt een beeld verkregen van de vertekening van beide methodes, in absolute en in relatieve termen.

Tabel 7 en 8 tonen de resultaten van de schattingen van het aantal bewegingen en ontheemden per gewest, en de validatie ervan, respectievelijk voor de methode van Herweging en de methode van Microsimulatie.

Figuur 3. Bewegingen en ontheemden

Aantal per gewest, 26 mei 2023, schatting en validatie van de herweging

	Schatting		Validatie	
	Aantal	Std.fout	Absolute afwijking	Relatieve afwijking
Vlaams Gewest				
Immigraties	37.096,96	1,840	-94,04*	-0,003
WG>VG	766,62	0,058	-3,38*	-0,004
BHG>VG	490,76	0,062	-8,24*	-0,017
Emigraties	4.989,68	0,762	-184,32*	-0,036
VG>WG	251,26	0,034	0,26*	0,001
VG>BHG	564,96	0,110	-26,04*	-0,044
Ontheemden	32.548,44	2,022	104,44*	0,003
Waals Gewest				
Immigraties	13.770,22	0,595	-60,78*	-0,004
VG>WG	251,26	0,034	0,26*	0,001
BHG>WG	130,44	0,021	-2,56*	-0,019
Emigraties	2.307,66	0,419	-43,34*	-0,018
WG>VG	766,62	0,058	-3,38*	-0,004
WG>BHG	362,83	0,065	4,83*	0,013
Ontheemden	10.714,82	0,734	-21,19*	-0,002
Brussels Hoofdstedelijk Gewest				
Immigraties	12.623,91	2,340	139,91*	0,011
VG>BHG	564,96	0,110	-26,04*	-0,044
WG>BHG	362,83	0,070	4,83*	0,013
Emigraties	709,37	0,270	-25,63*	-0,035
BHG>VG	490,76	0,060	-8,24*	-0,017
BHG>WG	130,44	0,020	-2,56*	-0,019
Ontheemden	12.221,13	2,360	155,12*	0,013

* Afwijking significant tot $p < 0.05$

VG: Vlaams Gewest, BHG: Brussels Hoofdstedelijk Gewest, WG: Waals Gewest.

Er is een lichte maar statistisch significante overschatting van het aantal ontheemden in het Vlaamse Gewest (0,3%) en het Brusselse Hoofdstedelijke Gewest (1,3%). In het Waalse Gewest wordt het aantal ontheemden licht onderschat (0,2%).

Het aantal migratiebewegingen wordt onderschat, met uitzondering van het aantal internationale immigraties naar het Brusselse Hoofdstedelijke Gewest en de interne stromen van het Vlaamse naar het Waalse en van het Waalse naar het Brussels Hoofdstedelijke Gewest. De afwijkingen tussen het geschatte en werkelijke aantal bewegingen schommelen tussen -3,6% en +1,3%. In absolute termen zijn vooral de afwijking tussen het geschatte en het werkelijke aantal internationale migratiestromen aanzienlijk (tussen -184,32 en +139,91).

Figuur 4. Bewegingen en ontheemden

Aantal per gewest, 26 mei 2023, schatting en validatie van de simulatie

	Schatting		Validatie	
	Aantal	Std.fout	Absolute afwijking	Relatieve afwijking
Vlaams Gewest				
Immigraties	38.893	0,250	1702,31*	0,046
WG>VG	720	0,063	-48,34*	-0,063
BHG>VG	383	0,040	-124,41*	-0,249
Emigraties	5.119	0,446	-54,82*	-0,011
VG>WG	214	0,024	-41,81*	-0,167
VG>BHG	452	0,046	-142,38*	-0,241
Ontheemden	34.213	0,506	1768,58*	0,055
Waals Gewest				
Immigraties	13.646	0,190	208,21*	0,015
VG>WG	214	0,024	-41,81*	-0,167
BHG>WG	109	0,015	-24,09*	-0,181
Emigraties	2.471	0,221	123,32*	0,052
WG>VG	720	0,063	-48,34*	-0,063
WG>BHG	325	0,041	-30,46*	-0,085
Ontheemden	10.453	0,268	97,80*	0,009
Brussels Hoofdstedelijk Gewest				
Immigraties	12.229,84	0,190	447,48*	0,036
VG>BHG	452,22	0,050	-41,81*	-0,167
WG>BHG	325,13	0,040	-30,46	-0,085
Emigraties	691,11	0,140	-37,05	-0,050
BHG>VG	383,13	0,040	-124,41	-0,249
BHG>WG	108,72	0,010	-24,09	-0,181
Ontheemden	11.824,22	0,250	460,18	0,038

* Afwijking significant tot $p < 0.05$

VG: Vlaams Gewest, BHG: Brussels Hoofdstedelijk Gewest, WG: Waals Gewest.

De simulatie-oefening resulteert in een aanzienlijke overschatting van het aantal ontheemden in het Vlaamse Gewest (5,5%) en het Brusselse Hoofdstedelijke Gewest (3,8%). In het Waalse Gewest blijft de overschatting beperkt tot 0,9%.

Het aantal migratiebewegingen wordt onderschat, met uitzondering van het aantal internationale immigraties naar de 3 gewesten, evenals de internationale emigraties uit het Waals Gewest. De afwijkingen tussen het geschatte en werkelijke aantal bewegingen schommelen tussen -24,9% en +4,6%.

In absolute termen zijn vooral de afwijking tussen het geschatte en het werkelijke aantal internationale immigraties aanzienlijk. In het Vlaamse Gewest loopt dit op tot 1.702,31 immigraties, in het Brusselse Hoofdstedelijke Gewest tot 477,48 en in het Waalse Gewest tot 208,21.

BESLUIT

Voor het opvolgen van de tijdelijk ontheemden uit Oekraïne steunde het beleid van het begin van de crisis op informatie uit het Rijksregister. Er is echter een vertraging tussen de registratie van migratiebewegingen van ontheemden in de gemeenten en de melding daarvan aan het Rijksregister. Hierdoor is het aantal migratiebewegingen en het aantal ontheemden dat beleidsmakers observeren in recente doorsneden van het Rijksregister onvolledig, wat leidt tot een incompleet beeld van deze bevolkingsgroep.

In dit rapport werd de werkwijze uiteengezet van 2 monitoringstools die de VSA ontwikkelde in het kader van haar opdracht voor de Task-Force *“Vlaanderen helpt Oekraïne”* om op basis van observaties uit doorsneden uit het Rijksregister die Statbel ter beschikking stelde, een meer correcte inschatting te maken van het aantal ontheemden in België en hun verdeling over de 3 gewesten.

Er werd vertrokken van de observaties van de internationale migraties en binnenlandse verhuisbewegingen. Het totaal aantal binnenlandse en internationale immigraties naar een gewest sinds het begin van de crisis in maart 2022, verminderd met het totale aantal binnenlandse en internationale emigraties uit een gewest vanaf het begin van de crisis, resulteert immers in het aantal ontheemden op het moment van de doorsnede.

Er werden 2 modellen ontwikkeld om het aantal bewegingen te schatten: een model op basis van Herweging en een model op basis van Microsimulatie. Deze methoden voorgesteld in dit rapport werden ontwikkeld als respons op beleidsvragen in de loop van de crisis. De methodologische keuzes werden dan ook mede gestuurd door de beleidsprioriteiten op dat moment en de doelstelling om een monitoringstool te bouwen om geautomatiseerde berekeningen op een later moment mogelijk te maken. De methode van Herweging werd ontwikkeld aan het begin van de crisis, wanneer het in kaart brengen van het aantal ontheemden en hun geografische spreiding over de gewesten centraal stond. Toen er naast nowcasting ook behoefte kwam aan projecties over mogelijke toekomstige ontwikkelingen, werd de overstap gemaakt naar de Microsimulatiemethode.

De methoden verschillen fundamenteel qua werkwijze, onderliggende assumpties, gegevensbasis en controlevariabelen. De methode van Herweging focust op de schatting van de administratieve vertraging zelf: de tijd die verstrijkt tussen het moment van registratie van een beweging en de melding ervan aan het Rijksregister. Ze steunt daarvoor op de Horvitz-Thompson-schatter, waarbij aan elke gemelde beweging een gewicht wordt toegekend op basis van de geschatte kans dat de beweging al gemeld is. Deze kans wordt, voor internationale immigraties, emigraties en interne verhuisbewegingen afzonderlijk en uitgaand van een constante administratieve vertraging in de tijd, berekend via een survival-analyse, rekening houdend met geografische kenmerken van de beweging.

Bij de methode van Microsimulatie vormen de registraties in de gemeenten van de instroom van ontheemden naar België, de opeenvolgende verblijfplaatsen en de emigraties naar het buitenland ("levenslopen") de basis voor het modelleren van migratiegedrag. Op basis van verhuis- en emigratiekansen en uitgaand van een constante immigratiestroom naar België wordt een dag-per-dag simulatie gemaakt van de verblijfplaatsen en bewegingen van ontheemden, waarmee het totale aantal actuele verblijvers in elk gewest wordt ingeschat. De verhuis- en emigratiekansen worden geschat via een survival-analyse op basis van gegevens ver genoeg terug in de tijd zodat ze kunnen geacht worden niet vertekend te zijn door administratieve vertraging in meldingen aan het Rijksregister. De schattingen houden rekening met individuele kenmerken en gaan uit van constante kansen voor en na de startdatum van simulatie.

Voor beide methoden werd in dit rapport een voorbeeld uitgewerkt op basis van de Rijksregisterdoorsnede van 26 mei 2023. Aan deze doorsnede gingen er voldoende extracties vooraf om de methoden te kunnen implementeren en waren er voldoende daaropvolgende om ze te valideren. De validatiedataset werd opgebouwd uit de meest recente doorsnede beschikbaar uit het Rijksregister (14 februari 2024). Deze bevat immers met heel grote waarschijnlijkheid *alle* bewegingen van *alle* ontheemden tot 26 mei 2023, ook deze die pas later werden gemeld aan het Rijksregister. De resultaten van de 2 methoden werden gevalideerd door ze te vergelijken met het aantal migratiebewegingen en het aantal ontheemden geobserveerd in de validatiedataset.

Een tweede toepassing van de methoden op een andere doorsnede zou de validatie zonder twijfel sterker maken. Alleen is er geen doorsnede beschikbaar in een tijdsvenster dat genoeg verschilt van deze van de gekozen doorsnede van 26 mei 2023 met voldoende observaties, zowel ervoor als erna, om de methode goed te kunnen schatten én te toetsen.

Algemeen kan gesteld worden dat de methode van Herweging een accuratere inschatting geeft van het aantal ontheemden dan de methode van Microsimulatie. Beide methoden onderschatten vaak het aantal bewegingen, maar de afwijkingen zijn groter bij de methode van Microsimulatie. Deze laatste methode berust op de vooronderstelling van een constant aantal dagelijkse immigraties en er wordt uitgegaan van temporele stabiliteit in de duur tot internationale emigratie of verhuis. Met andere woorden, de methode van Microsimulatie veronderstelt dat het migratiegedrag van opeenvolgende cohorten niet verandert, wat in tijden van (oorlogs-)crisis een betwistbare aanname is. Ook de methode van Herweging gaat uit van temporele stabiliteit, maar hier betreft het de duur van het administratieve proces tussen de registratie van migraties in de gemeente en de melding aan het Rijksregister. Deze hypothese werd ondersteund door de gegevens en het is bovendien aannemelijk dat, na verloop van tijd in de crisis, opnieuw een stabiel administratief verwerkingsritme wordt bereikt.

Timely data worden steeds belangrijker nu crisissen naar verwachting sneller op elkaar zullen volgen. Daarom is het zinvol te onderzoeken of beide tools, oorspronkelijk ontwikkeld voor het monitoren van het aantal ontheemden in België, ook kunnen bijdragen aan de correctie van tellingen in andere onvolledige administratieve databronnen. Vooral de methode van Herweging is in dit opzicht relevant, omdat zij een eenvoudige en snelle correctie van administratieve data mogelijk maakt. Daarnaast kunnen de wegingscoëfficiënten die aan deze methode ten grondslag liggen, ook mogelijk rechtstreeks worden ingezet in statistische analyses op onvolledige administratieve data.

Tegelijk is het waardevol om de mogelijkheden van andere modellen te verkennen – zoals die van Bastos e.a. (2019) en Günther e.a. (2021) – die een gezamenlijke modellering van gebeurtenissen en meldingsvertragingen toelaten. Deze modellen maken het mogelijk de vertraging per gebeurtenis te differentiëren en variatie in de tijd te modelleren, waardoor ze ook accuratere en robuustere resultaten kunnen genereren dan de hier voorgestelde herweging.

De kracht van de microsimulatiemethode ligt in het kunnen maken van vooruitzichten voorbij de periode van administratieve vertraging. Tegelijk bleek deze methode minder efficiënt in het corrigeren van de tellingen. Het kan daarom nuttig zijn verder uit te zoeken of de resultaten van de herwegingsmethode zouden kunnen gebruikt worden om vooruitzichten te kalibreren.

Ten slotte pleiten de bevindingen uit dit rapport voor het systematisch opnemen van niet alleen de datum van events zelf, maar ook de datum van opname ervan in administratieve databases zoals het Rijksregister voor statistische analyses gebaseerd op administratieve data. Bij gebrek hieraan moet de administratieve vertraging worden afgeleid uit de opeenvolgende datums van extractie, zoals bij de oefening in dit rapport. Dit kan de resultaten van de schattingen aanzienlijk beïnvloeden, zeker bij een grote tijdsspanne tussen extracties.

LITERATUURLIJST

Banbura, M., Giannone, D., Modugno, M. & Reichlin, L. (2013). Now-casting and the real-time data flow. In: Elliott, G., Timmermann, A. (red.), *Handbook of Economic Forecasting*. Elsevier: 195–237.

Bastos, L.S., Economou, T., Gomes, M.F.C., Villela, D.A.M., Coelho, F.C., Cruz, O.G., Stoner, O., Bailey, T. & Codeço, C.T. (2019). A modelling approach for correcting reporting delays in disease surveillance data. In: *Statistics in Medicine*, 38, 4363–4377.

Brookmeyer, R. & Gail, M. H. (1988). A method for obtaining short-term projections and lower bounds on the size of the aids epidemic. In: *Journal of the American Statistical Association*, 83(402), 301–308.

Casper, T.C. & Cook, T.D. (2012). Estimation of the Mean Frequency Function for Recurrent Events when Ascertainment of Events Is Delayed. In: *The International Journal of Biostatistics*, 8(1), 1–20.

Crevecoeur, J., Antonio, K. & Verbelen, R. (2019). Modeling the number of hidden events subject to observation delay. In: *European Journal of Operational Research*, 277, 930–944.

Donker, T., van Boven, M., van Ballegooijen, W.M., van't Klooster, T.M., Wielders, C.C. & Wallinga, J. (2011). Nowcasting pandemic influenza A/H1N1 2009 hospitalizations in the Netherlands. In: *European Journal of Epidemiology*, 26, 195–201.

Günther, F., Bender, A., Katz, K., Küchenhoff, H. & Höhle, M. (2021). Nowcasting the COVID-19 pandemic in Bavaria. In: *Biometrical Journal*, 63, 490–502.

Haller, B., Schmidt, G. & Ulm, K. (2013). Applying competing risks regression models: an overview. In: *Lifetime Data Analysis*, 19, 33–58.

Harris, J.E. (1990). Reporting delays and the incidence of AIDS. In: *Journal of the American Statistical Association*, 85, 915–924.

Heisterkamp, S.H., Jager, J.C., Ruitenberg, E.J., Van Druten, J.A.M. & Downs, A.M. (1989). Correcting reported AIDS incidence: A statistical approach. In: *Statistics in Medicine*, 8, 963–976.

Höhle, M. & der Heiden, M. (2014). Bayesian nowcasting during the STEC O104:H4 outbreak in Germany, 2011. In: *Biometrics*, 70, 993–1002.

Horvitz, D.G. & Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. In: *Journal of the American Statistical Association*, 47, 663–685.

<https://statbel.fgov.be/nl/visuals/oekraiense-ontheemden>, geraadpleegd op 24 december 2024

<http://data2.unhcr.org/en/situations/ukraine>, geraadpleegd op 16 januari 2025.

Li, J. & O'Donoghue, C. (2013), A survey of dynamic microsimulation models: Uses, model structure and methodology, In: *International Journal of Microsimulation*, 6(2); 3-55.

Mahani, A.S. & Sharabiani, M.T.A. (2019). Bayesian, and non-Bayesian, cause-specific competing risk analysis for parametric and nonparametric survival functions: The R package CFC. In: *Journal of Statistical Software*, 89, 1–29.

O'Donoghue, C. (2001). Dynamic Microsimulation: A Methodological Survey. In: *Brazilian Electronic Journal of Economics*, 4.

Robert, C. & Casella, G. (2000). Monte Carlo Statistical Method. In: *Technometrics*, 42.

Salmon, M., Schumacher, D., Stark, K. & Höhle, M. (2015). Bayesian outbreak detection in the presence of reporting delays. In: *Biometrical Journal*, 57, 1051–1067.

Spielauer, M. (2006). 'The "life-course" model, a competing risk cohort microsimulation model: source code and basic concepts of the generic microsimulation programming language ModGen'. *Working Paper WP 2006-046*, Max Planck Institute for Demographic Research, Rostock.

Sun, J. (2006). *The Statistical Analysis of Interval-censored Failure Time Data*. Springer Science & Business Media.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. In: *Proceedings of the National Academy of Sciences*, 72, 20–22.

Therneau, T. M. (2020). *A Package for Survival Analysis in R*. <https://CRAN.R-project.org/package=survival>

van de Kastele, J., Eilers, P.H.C. & Wallinga, J. (2019). Nowcasting the number of new symptomatic cases during infectious disease outbreaks using constrained p-spline smoothing. In: *Epidemiology*, 30, 737–745.

Wang, Y., De Coning, E., Jacobs, W., Joe, P., Nikitina, L., Roberts, R., Wang, J. & Wilson, J. (2017). *Guidelines for Nowcasting Techniques*, WMO 1198, World Meteorological Organization, Geneva.

Willekens, F. J. C. (2009). Continuous-time microsimulation in longitudinal analysis. In: Zaidi, A., Harding, A & Williamson, P. (red.), *New frontiers in microsimulation modelling*, Ashgate Publishing Ltd, 413-436.

BIJLAGEN

Bijlage 1: Parameterschattingen voor administratieve vertraging op de melding aan het Rijksregister op internationale immigraties en emigraties en op verhuisbewegingen tussen gewesten

Tabel 1. Administratieve vertraging op internationale immigratie

	Coef.	Std. Fout	z	p
(Intercept)	3,3144	0,0240	137,9967	0,0000
Waals Gewest	0,0498	0,0220	2,2663	0,0234
Brussels Gewest	0,7881	0,0204	38,5496	0,0000
Centrumstad	0,1131	0,0178	6,3658	0,0000
Maand registratie aug. 2022	-0,0314	0,0268	-1,1706	0,2418
Maand registratie sept. 2022	-0,0092	0,0283	-0,3248	0,7453
Maand registratie okt. 2022	0,0651	0,0290	2,2451	0,0248
Maand registratie nov. 2022	0,0476	0,0292	1,6305	0,1030
Maand registratie dec. 2022	0,0398	0,0314	1,2681	0,2048
Log(scale)	-0,2277	0,0076	-29,8713	0,0000

Referentiecategorieën: Vlaams Gewest, eerste verblijfplaats geen centrumstad, maand registratie juli 2022.

Tabel 2. Administratieve vertraging op internationale emigratie

	Coef.	Std. Fout	z	p
(Intercept)	2,4343	0,1054	23,0930	0,0000
Waals Gewest	0,1811	0,0538	3,3639	0,0008
Brussels Gewest	0,4640	0,0892	5,1989	0,0002
Centrumstad	0,2690	0,0653	4,1180	0,0000
Maand registratie aug. 2022	-0,2156	0,1138	-1,8943	0,0582
Maand registratie sept. 2022	-0,3873	0,1119	-3,4620	0,0005
Maand registratie okt. 2022	-0,4211	0,1146	-3,6736	0,0002
Maand registratie nov. 2022	0,0629	0,1157	0,5439	0,5865
Maand registratie dec. 2022	-0,2074	0,1163	-1,7836	0,0745
Log(scale)	0,3385	0,0153	22,1251	0,0000

Referentiecategorieën: Vlaams Gewest, eerste verblijfplaats geen centrumstad, maand registratie juli 2022.

Tabel 3. Administratieve vertraging op interregionale verhuisbewegingen

	Coef.	Std. Fout	z	p
(Intercept)	3,0029	0,0238	126,0787	0,0000
Van Waals Gewest	0,0705	0,0349	2,0192	0,0435
Van Brussels Gewest	0,1263	0,0352	3,5875	0,0003
Van centrumstad	0,0464	0,0234	1,9825	0,0474
Naar Waals Gewest	0,0398	0,0360	1,1045	0,2694
Naar Brussels Gewest	0,4613	0,0344	13,4057	0,0000
Naar centrumstad	0,4901	0,0223	21,9820	0,0000
Maand registratie aug. 2022	0,0470	0,0266	1,7639	0,0777
Maand registratie sept. 2022	-0,0523	0,0268	-1,9536	0,0507
Maand registratie okt. 2022	0,0280	0,0278	1,0063	0,3143
Maand registratie nov. 2022	0,1100	0,0284	3,8806	0,0001
Maand registratie dec. 2022	-0,0310	0,0293	-1,0592	0,2895
Log(scale)	-0,2480	0,0068	-36,6224	0,0000

Bijlage 2. Parameterschattingen voor de duur tot internationale emigratie en de duur tot interne verhuisbeweging

Tabel 4. Duur tot internationale emigratie

	Coef.	Std. Fout	z	p
(Intercept)	6,4520	0,0390	163,3720	0,0000
Waals Gewest	-0,1110	0,0190	-6,0050	0,0000
Brussels Gewest	0,7840	0,0340	22,8800	0,0000
Alleenstaande ouder zonder grootouder(s)	-0,1810	0,0310	-5,8110	0,0000
Andere samenstelling met kinderen	0,0040	0,0350	0,1050	0,9160
Andere samenstelling zonder kinderen	-0,0880	0,0410	-2,1710	0,0300
Eenpersoonshuishouden man	0,1290	0,0860	1,5060	0,1320
Eenpersoonshuishouden vrouw	0,1570	0,0490	3,1930	0,0010
Paar met kinderen en grootouder(s)	0,0810	0,0460	1,7790	0,0750
Paar met kinderen zonder grootouder(s)	0,0190	0,0340	0,5550	0,5790
Paar zonder kinderen, 1 generatie	0,0630	0,0460	1,3780	0,1680
Paar zonder kinderen, 2 generaties	-0,1160	0,0460	-2,5000	0,0120
Geslacht	-0,0970	0,0190	-5,2320	0,0000
Eerste verblijfplaats centrumstad	0,4830	0,0240	19,9540	0,0000
Aantal (voorafgaande) bewegingen	0,1750	0,0260	6,8080	0,0000
Maand van immigratie	0,0470	0,0060	7,2420	0,0000
Log(scale)	-0,5390	0,0130	-42,4710	0,0000

Referentiecategorieën: Vlaams Gewest, alleenstaande ouder met grootouders, eerste verblijfplaats geen

Tabel 5. Duur tot interne verhuisbeweging

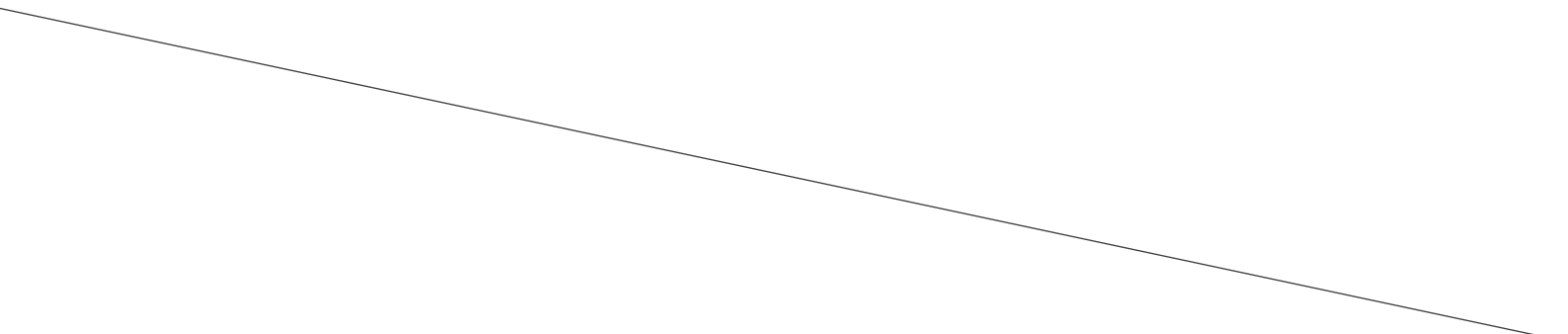
	Coef.	Std. Fout	z	p
(Intercept)	5,3470	0,0570	94,6260	0,0000
Waals Gewest	-0,1540	0,0310	-5,0510	0,0000
Brussels Gewest	0,8920	0,0410	21,9860	0,0000
Alleenstaande ouder zonder grootouder(s)	-0,0550	0,0470	-1,1580	0,2470
Andere samenstelling met kinderen	0,1250	0,0520	2,4030	0,0160
Andere samenstelling zonder kinderen	0,6050	0,0680	8,8530	0,0000
Eenpersoonshuishouden man	0,3310	0,1130	2,9240	0,0030
Eenpersoonshuishouden vrouw	0,3450	0,0710	4,8870	0,0000
Paar met kinderen en grootouder(s)	0,0900	0,0660	1,3540	0,1760
Paar met kinderen zonder grootouder(s)	0,0530	0,0500	1,0650	0,2870
Paar zonder kinderen, 1 generatie	0,2980	0,0680	4,3840	0,0000
Paar zonder kinderen, 2 generaties	0,0850	0,0730	1,1640	0,2450
Geslacht	-0,1070	0,0280	-3,8150	0,0000
Eerste verblijfplaats centrumstad	0,2130	0,0310	6,8140	0,0000
Aantal (voorafgaande) bewegingen	1,0510	0,0420	24,8970	0,0000
Maand van immigratie	0,3260	0,0080	38,8450	0,0000
Log(scale)	0,6840	0,0060	109,8760	0,0000

Referentiecategorieën: Vlaams Gewest, alleenstaande ouder met grootouders, eerste verblijfplaats geen

Bijlage 3: Gemiddelde verhuiskansen tussen gewesten

Tabel 6. Matrix van gemiddelde verhuiskansen tussen gewesten
Belgische gewesten, 1 juli 2022 – 15 februari 2023

		Naar		
		Vlaams Gewest	Waals Gewest	Brussels Gewest
Van	Vlaams Gewest	94,36%	1,45%	4,19%
	Waals Gewest	7,91%	88,21%	3,88%
	Brussels Gewest	9,14%	2,48%	88,37%



COLOFON

Verantwoordelijke uitgever

Statistiek Vlaanderen

Bezoekadres: Simon Bolivarlaan 17, 1000 Brussel

Auteurs

Ingrid Schockaert | ingrid.schockaert@vlaanderen.be

Rembert De Blander

Lisa Van Landschoot | lisa.vanlandschoot@vlaanderen.be

Jan Pickery | jan.pickery@vlaanderen.be

Grafische vormgeving

Guy De Smet

Depotnummer

D/2025/3241/469



