



Risicofactoren voor transitie in en uit armoede

Samenvatting

Jan van der Laan, Joep Burger, Marijn Detiger, Naomi Schalken, Wouter van Andel,
Marion van den Brakel, Sita Tan

Juli 2021



Voorwoord

Deze samenvatting geeft een beknopte beschrijving van de inhoudelijke resultaten van het onderzoek naar risicofactoren voor transities in en uit armoede dat het Centraal Bureau voor de Statistiek (CBS) op verzoek van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties heeft uitgevoerd. Verdere toelichting bij de data, resultaten en methoden is te vinden in het bijbehorende methodologische rapport ([CBS, 2021](#)).

Inhoud

1. Inleiding
2. Data
3. Methode
4. Resultaten
5. Conclusie
6. Discussie

1. Inleiding



Inleiding

- a. Probleemstelling
- b. Onderzoeksvragen

Probleemstelling

Armoede is een relatief groot probleem: ruim 1 miljoen mensen leefden in 2019 in een huishouden onder de lage-inkomensgrens (6,2% van de Nederlandse bevolking), van wie bijna 400 duizend al ten minste vier jaar achtereen (2,5% van de bevolking) ([CBS, 2020](#)).

→ Kan vroegtijdige signalering van risicofactoren voor armoede via armoedepreventie en -bestrijding bijdragen aan effectief armoedebeleid?

Onderzoeksvragen

- Hoe goed kunnen we op basis van (combinaties van) kenmerken de kans schatten dat iemand arm blijft dan wel arm wordt?
- Voor welke aspecten kan de combinatie van populatiedekkende registerdata en machine learning tot nieuwe inzichten leiden over risicofactoren voor transitie in en uit armoede?

2. Data



2. Data

- a. Doelvariabele - armoededefinitie
- b. Kenmerken
- c. Doelpopulatie
- d. Periode

Doelvariabele - armoededefinitie

Doelvariabele: indicator voor armoede 1 jaar na het referentiejaar

Hierbij is iemand arm als

- het gestandaardiseerd besteedbaar huishoudinkomen onder de lage inkomensgrens ligt
- EN het gestandaardiseerd liquide vermogen onder de helft van de lage inkomensgrens ligt.

Kenmerken

- Ruim 500 potentiële risicofactoren
- Thema's:
 - demografie (o.a. leeftijd, geslacht, huishouden, geboorteland)
 - economie (o.a. opleiding, werk, inkomen)
 - gezondheid (o.a. zorgkosten)
 - criminaliteit (o.a. detentie)
- Persoonskenmerken en huishoudkenmerken
- Periodiciteit: tijdsonafhankelijk, jaar, maand

Doelpopulatie

- Personen van 18+ in particuliere huishoudens (exclusief institutionele huishoudens en studentenhuishoudens) waarvan inkomen en vermogen bekend zijn.
- Dit is opgesplitst in 2 subpopulaties:
 - *Arme populatie*: personen die in referentiejaar arm zijn (ongeveer 0,6 miljoen per referentiejaar)
 - *Niet-arme populatie*: personen die in referentiejaar niet arm zijn (ongeveer 12 miljoen per referentiejaar)

Periode

- Twee referentiejaren: 2013 en 2017
- We schatten met twee modellen – voor de arme populatie en voor de niet-arme populatie – de kans op armoede in 2014 en 2018 op basis van data over respectievelijk 2011–2013 en 2015–2017.

3. Methode



Methode

- [XGBoost](#) (Extreme Gradient Boosting): populaire supervised machine-learningmethode waarbij veel relatief simpele beslisbomen of classificatiebomen gecombineerd worden om de kans te schatten dat een persoon arm blijft/wordt.
- [Global surrogate model](#): Wegens de complexiteit van de gecombineerde beslisbomen in XGBoost, zijn de uitkomsten moeilijk te interpreteren. Daarom zijn eenvoudige surrogaatmodellen geschat die de geschatte kansen van de complexe modellen proberen te beschrijven. Deze surrogaatmodellen schatten vanwege hun eenvoud minder goed dan XGBoost, maar kunnen wel gebruikt worden om datagedreven deelpopulaties te definiëren die verder bestudeerd kunnen worden.
- [SHAP](#): gemiddelde bijdrage van een kenmerk aan iemands geschatte armoedekans (er wordt 'gemiddeld' over verschillende modellen met en zonder het kenmerk)

4. Resultaten



Resultaten

- a. Modelkwaliteit
- b. Wat is de kans op armoede?
- c. Wat zijn de belangrijkste risicofactoren?
- d. Welke datagedreven deelpopulaties zijn te onderscheiden?
- e. In hoeverre wijken de uitkomsten voor deelpopulaties af van landelijke uitkomsten?

Modelkwaliteit

- Onderscheidend vermogen: voor twee willekeurige personen, één die arm is gebleven en één die uit de armoede is gekomen, schat het model in 83% van de gevallen een hogere armoedekans voor degene die arm is gebleven. Analooch schat het model in 95% van de gevallen een hogere armoedekans voor een persoon die arm is geworden dan voor een persoon die niet-arm is gebleven. Dit betekent dat de modellen geschikt zijn om in kaart te brengen hoe potentiële risicofactoren de armoedekans beïnvloeden.
- Classificatie: het model voor de arme populatie schat op individueel niveau beter of iemand wel of niet arm wordt dan het model voor de niet-arme populatie. Voor de arme populatie zijn de recall (terecht positieven/ waargenomen positieven) en precision (terecht positieven/ voorspelde positieven) respectievelijk 94% en 80% en voor de niet – arme populatie 44% en 41%. Dit betekent dat de modellen minder geschikt zijn om op individueel niveau personen te classificeren als wel of niet arm.



Wat is de kans op armoede?

- a. Arme populatie: gemiddeld 71% kans om arm te blijven
- b. Niet-armer populatie: gemiddeld 1,2% kans om arm te worden

Top-20 belangrijkste risicofactoren om arm te blijven

1. **Alleenstaande**
2. **Inkomensarm (referentiejaar - 1)**
3. Wel /niet werkzaam (referentiejaar)
4. **Leeftijd**
5. Fractie van maanden dat persoon bijstand had (referentiejaar)
6. Fractie van maanden dat persoon bijstand had (laatste helft van referentiejaar)
7. Aantal maanden geleden dat persoon werknemer was
8. Aantal maanden sinds wijziging in partnerschap
9. **Inkomensarm (referentiejaar - 2)**
10. **Aantal maanden sinds start van huishouden**
11. Aantal maanden sinds geboorte kind
12. **Fractie van maanden dat persoon een partner had (referentiejaar)**
13. **Eigen woning (referentiejaar)**
14. **Opleidingsniveau**
15. **Generatie migratieachtergrond**
16. **Huurwoning met huurtoeslag (referentiejaar)**
17. Bijstand als belangrijkste inkomstenbron
18. Aantal kinderen in huishouden
19. **Loon als voornaamste inkomensbron van huishouden (referentiejaar)**
20. **Aantal gewerkte uren (december van referentiejaar)**



Dikgedrukte risicofactoren komen ook voor in de top-20 voor de niet-arme populatie.

Top-20 belangrijkste risicofactoren om arm te worden

1. **Eigen woning (referentiejaar)**
2. Economisch onzelfstandig (referentiejaar)
3. **Leeftijd**
4. **Huurwoning met huurtoeslag (referentiejaar)**
5. **Aantal gewerkte uren (december van referentiejaar)**
6. **Aantal maanden sinds start van huishouden**
7. Eigen woning (referentiejaar - 2)
8. **Generatie migratieachtergrond**
9. **Opleidingsniveau**
10. Eigen woning (referentiejaar - 1)
11. **Alleenstaande**
12. **Inkomensarm (referentiejaar - 1)**
13. Aantal maanden geleden dat persoon zelfstandig was
14. Thuiswonend kind
15. **Loon als voornaamste inkomensbron van huishouden (referentiejaar)**
16. Fractie van maanden dat persoon voltijd werkzaam was (referentiejaar)
17. Aantal gewerkte uren (november van referentiejaar)
18. **Inkomensarm (referentiejaar - 2)**
19. **Fractie van maanden dat persoon een partner had (referentiejaar)**
20. Fractie van maanden dat persoon een pensioen had (laatste helft van referentiejaar)



Welke deelpopulaties kunnen in de arme populatie worden onderscheiden?

Deelpopulatie	Aandeel	Gemiddelde armoedekans
Hele jaar bijstand, alleenstaand, 42-63	14%	95%
Hele jaar bijstand, niet alleenstaand, kind geboren minder dan 3 jaar geleden	3%	90%
Hele jaar bijstand, alleenstaand, 18-41	6%	87%
Niet hele jaar bijstand, partnerschap niet gewijzigd, niet werkend, alleenstaand, 18-63	10%	86%
Hele jaar bijstand, niet alleenstaand, kind geboren ten minste 3 jaar geleden	17%	76%
*Inwoner gemeente Heerlen	1%	74%
Niet hele jaar bijstand, partnerschap niet gewijzigd, niet werkend, niet alleenstaand, vorig jaar inkomensarm	12%	73%
*Inwoner gemeente Amsterdam	10%	73%
*Ouder in eenouderhuishouden	12%	68%
*Laag opleidingsniveau, niet thuiswonend, 18-26, volgt geen opleiding	3%	68%
Hele jaar bijstand, alleenstaand, 64+	3%	66%
Niet hele jaar bijstand, partnerschap niet gewijzigd, niet werkend, alleenstaand, 64+	2%	60%
Niet hele jaar bijstand, partnerschap niet gewijzigd, werkend, vorig jaar inkomensarm	10%	59%
Niet hele jaar bijstand, partnerschap niet gewijzigd, niet werkend, niet alleenstaand, vorig jaar niet inkomensarm	7%	55%
Niet hele jaar bijstand, partnerschap gewijzigd, alleenstaand	1%	55%
*Werknemer, groot aandeelhouder, zelfstandig ondernemer, overig zelfstandige of meewerkend gezinslid, 18-74	21%	48%
Niet hele jaar bijstand, partnerschap niet gewijzigd, werkend, vorig jaar niet inkomensarm	9%	43%
Niet hele jaar bijstand, partnerschap gewijzigd, niet alleenstaand	5%	23%

* Deelpopulatie op basis van inhoudelijke gronden.
Overige deelpopulaties volgen uit de data.



Welke deelpopulaties kunnen in de niet-arme populatie worden onderscheiden?

Beschrijving	Aandeel	Gemiddelde armoedekans
Geen eigen woning, economisch onzelfstandig, huurwoning met huurtoeslag	2,9%	17,7%
*Laag opleidingsniveau, niet thuiswonend, 18-26, volgt geen opleiding	0,4%	12,3%
*Ouder in eenouderhuishouden	3,7%	5,3%
Geen eigen woning, economisch onzelfstandig, geen huurwoning met huurtoeslag	5,3%	3,3%
Geen eigen woning, economisch zelfstandig, huurwoning met huurtoeslag	5,4%	2,9%
*Inwoner gemeente Amsterdam	4,4%	2,5%
Geen eigen woning, economisch zelfstandig, geen huurwoning met huurtoeslag, zelfstandig ondernemer of zelfstandig ondernemer geweest in de afgelopen 1,5 jaar	1,4%	2,4%
*Inwoner gemeente Heerlen	0,5%	1,8%
Eigen woning, geen zelfstandig ondernemer of langer dan 3 jaar geleden zelfstandig ondernemer, economisch onzelfstandig, woont in een huishouden dat korter dan 79 maanden bestaat	9%	1,4%
Eigen woning, zelfstandig ondernemer of zelfstandig ondernemer geweest in de afgelopen 3 jaar	7,5%	1,1%
*Werknemer, groot aandeelhouder, zelfstandig ondernemer, overig zelfstandige of meewerkend gezinslid, 18-74	58,4%	0,7%
Geen eigen woning, economisch zelfstandig, geen huurwoning met huurtoeslag, geen zelfstandig ondernemer of ergens tussen de 1,5 en 3 jaar geleden zelfstandig ondernemer geweest	15,7%	0,6%
Eigen woning, geen zelfstandig ondernemer of langer dan 3 jaar geleden zelfstandig ondernemer, economisch onzelfstandig, woont in een huishouden dat langer dan 79 maanden bestaat	5,4%	0,3%
Eigen woning, geen zelfstandig ondernemer of langer dan 3 jaar geleden zelfstandig ondernemer, economisch zelfstandig, woont in een huishouden dat minder dan 82 maanden bestaat, niet hele jaar voltijd baan	16,6%	0,2%
Eigen woning, geen zelfstandig ondernemer of langer dan 3 jaar geleden zelfstandig ondernemer, economisch zelfstandig, woont in een huishouden dat minder dan 82 maanden bestaat, voltijd baan in het afgelopen jaar	12,6%	0,1%
Eigen woning, geen zelfstandig ondernemer of langer dan 3 jaar geleden zelfstandig ondernemer, economisch zelfstandig, woont in huishouden dat langer dan 82 maanden bestaat	18,2%	0,1%

* Deelpopulatie op basis van inhoudelijke gronden.
Overige deelpopulaties volgen uit de data.



4e. Resultaten deelpopulaties

In hoeverre wijken uitkomsten voor deelpopulaties af van landelijke uitkomsten?

- In hoeverre wijken demografische **kenmerken** af van de verdeling in Nederland?
- Hoe verhoudt de **armoedekans** zich tot de armoedekans in Nederland?
- Zijn bepaalde **risicofactoren** belangrijker dan voor heel Nederland?

De uitkomsten voor de arme deelpopulaties vergelijken we met de uitkomsten voor de gehele arme populatie in Nederland. Analoog vergelijken we de uitkomsten voor de niet-arme deelpopulaties met de uitkomsten voor de gehele niet-arme populatie in Nederland.



Onderzochte deelpopulaties

Voor zowel de arme als niet-arme populatie:

- i. Deelpopulatie met kleinste/ grootste armoedekans
- ii. Amsterdam
- iii. Heerlen
- iv. Laagopgeleide jongeren
- v. Werkenden
- vi. Alleenstaande ouders

4e. Resultaten deelpopulaties – arme populatie

Arme deelpopulatie met grootste kans om uit de armoede te komen

- Arm
 - Niet hele jaar bijstand; voornamelijk het hele jaar geen bijstand
 - Partnerschap gewijzigd in het afgelopen jaar; voornamelijk een partner gekregen
 - Niet alleenstaand
-
- Kenmerken: met name relatief veel jongeren, personen met een Nederlandse migratieachtergrond, vrouwen, ongehuwde partners, relatief minder gehuwde partners met kinderen.
 - Armoedekans: 23% tegenover 71% landelijk.
 - Risicofactoren: de top-20 belangrijkste kenmerken komt grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo is met name wijziging in partnerschap belangrijker dan voor heel Nederland, doordat wijziging in partnerschap in deze deelpopulatie – per definitie – relatief vaak voorkomt. Verder is de kans om uit de armoede te komen – net als voor de hele arme populatie – groter, naarmate de wijziging langer geleden is binnen het afgelopen jaar. Daarbij is dit effect sterker voor deze deelpopulatie.

Arme deelpopulatie Amsterdam

- Arm
- Inwoner van Amsterdam

- Kenmerken: met name relatief meer personen met een eerste generatie migratieachtergrond en meer alleenstaanden, relatief minder personen met een Nederlandse achtergrond.
- Armoedekans: 73% tegenover 71% landelijk.
- Risicofactoren: de uitkomsten zijn grotendeels vergelijkbaar met de landelijke uitkomsten.

Arme deelpopulatie Heerlen

- Arm
- Inwoner van Heerlen

- Kenmerken: met name relatief meer personen met een Nederlandse achtergrond, met een tweede generatie migratieachtergrond, alleenstaanden, relatief minder gehuwde partners met kinderen.
- Armoedekans: 74% tegenover 71% landelijk.
- Risicofactoren: de uitkomsten zijn grotendeels vergelijkbaar met landelijke uitkomsten. Wel is het kenmerk alleenstaand wat belangrijker voor deze deelpopulatie. Net als landelijk gezien hebben alleenstaanden een relatief grote kans om arm te blijven en niet-alleenstaanden een relatief kleine kans. Dit effect is voor deze deelpopulatie sterker dan landelijk gezien. Daarbij bevat deze deelpopulatie relatief veel alleenstaanden.

Arme laag opgeleide jongeren

- Arm
- Laag opleidingsniveau
- 18- 26 jaar
- Niet schoolgaand of studerend
- Geen thuiswonend kind.

- Kenmerken: met name relatief veel personen met Nederlandse achtergrond of 2e generatie migratieachtergrond, alleenstaanden, ongehuwde partners, relatief minder gehuwde partners.
- Armoedekans: 68% tegenover 71% landelijk.
- Risicofactoren: de top-20 belangrijkste kenmerken komt grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo zijn de kenmerken wijziging in partnerschap en kind geboren belangrijker, met name doordat deze risicogroepen relatief vaak voorkomen. Daarbij geldt net als landelijk dat personen met een recente wijziging in partnerschap een grotere kans hebben om uit armoede te komen, terwijl een wijziging langer geleden gepaard gaat met een grotere kans om arm te blijven. Daarnaast geldt net als landelijk gezien, dat de kans om arm te blijven toeneemt naarmate het korter geleden is dat er een kind is geboren.

Arme werkenden

- Arm
 - Werknemer, directeur-grotaandeelhouder, zelfstandig ondernemer, overig zelfstandige of meewerkend gezinslid
-
- Kenmerken: met name relatief veel personen van ongeveer 20-50 jaar, personen met een Nederlandse achtergrond, mannen, partners met kinderen, ongehuwde partner zonder kinderen, en relatief minder personen met een eerste generatie migratieachtergrond en minder alleenstaanden.
 - Armoedekans: 48% tegenover 71% landelijk.
 - Risicofactoren: de top-20 belangrijkste kenmerken komen grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo is recente wijziging in partnerschap belangrijker dan landelijk. Dit komt zowel door de samenstelling van de populatie als doordat personen met een recent gewijzigd partnerschap in deze deelpopulatie een kleinere kans hebben om arm te blijven dan personen met een recent gewijzigd partnerschap in Nederland.



Arme alleenstaande ouders

- Arm
- Ouder in eenouderhuishouden

- Kenmerken: met name relatief meer personen van ongeveer 25-55 jaar, vrouwen en wat meer personen met een tweede generatie migratieachtergrond.
- Armoedekans: 68% tegenover 71% landelijk.
- Risicofactoren: de top-20 belangrijkste kenmerken komen grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo zijn de kenmerken aantal kinderen in het huishouden en tijd sinds geboorte van een kind met name belangrijker, doordat het effect van deze kenmerken in deze deelpopulatie sterker is dan landelijk. Zo neemt de kans om arm te blijven onder arme alleenstaande ouders meer dan gemiddeld toe, naarmate het korter geleden is dat er een kind geboren is. Daarnaast hebben arme alleenstaande ouders met één kind– net als landelijk– een kleinere kans om arm te blijven, en met drie of meer kinderen juist een hogere kans. Dit effect is voor arme alleenstaande ouders sterker dan landelijk.

4e. Resultaten deelpopulaties – niet-arme populatie

Niet-arme deelpopulatie met grootste kans om arm te worden

- Niet arm
- Woont in een huurwoning met huurtoeslag
- Economisch onzelfstandig

- Kenmerken: met name relatief veel jongeren, eerste en tweede generatie migranten, vrouwen, alleenstaanden, eenoudergezinnen, thuiswonende kinderen, relatief minder gehuwde partners.
- Armoedekans: 18% tegenover 1,2% landelijk.
- Risicofactoren: de top-20 belangrijkste kenmerken komt grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo hebben alleenstaanden - net als landelijk gezien - een grotere kans om arm te worden en niet-alleenstaanden een kleinere. Dit effect is in de deelpopulatie sterker dan landelijk gezien. Ook zijn er relatief veel personen die minder dan drie jaar geleden een bijstandsuitkering gehad hebben, waardoor dit kenmerk belangrijker is dan landelijk. Verder valt onder meer op dat er relatief veel personen zijn die in de twee jaren voorafgaand aan het referentiejaar inkomensarm waren. Met name hierdoor is inkomensarmoede voor deze deelpopulatie belangrijker.

Niet-arme deelpopulatie Amsterdam

- Niet arm
- Inwoner van Amsterdam

- Kenmerken: met name relatief veel personen van ongeveer 25-45 jaar, personen met een eerste en tweede migratieachtergrond, alleenstaanden, ongehuwde partners zonder kinderen en relatief weinig gehuwde partners.
- Armoedekans: 2,5% tegenover 1,2% landelijk.
- Risicofactoren: de top-20 belangrijkste kenmerken komt grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo is het kenmerk alleenstaand belangrijker, deels doordat er meer alleenstaanden zijn en deels doordat alleenstaanden een grotere kans hebben om arm te worden, dan wel niet-alleenstaanden een kleinere kans hebben. Verder zijn economische onzelfstandigheid en het aantal gewerkte uren met name belangrijker, doordat het effect van deze kenmerken groter is dan landelijk. Daarnaast is migratieachtergrond (eerste en tweede generatie) belangrijker dan landelijk, met name omdat deze relatief veel voorkomen in deze deelpopulatie.

Niet-arme deelpopulatie Heerlen

- Niet arm
- Inwoner van Heerlen

- Kenmerken: met name relatief meer ouderen, personen met een tweede generatie migratieachtergrond, alleenstaanden en wat minder gehuwde partners met kind.
- Armoedekans: 1,8% tegenover 1,2% landelijk.
- Risicofactoren: de uitkomsten zijn grotendeels vergelijkbaar met landelijke uitkomsten. Wel is het kenmerk huurwoning met huurtoeslag belangrijker omdat dit relatief vaak voorkomt.

Niet-arme laag opgeleide jongeren

- Niet arm
- Laag opleidingsniveau
- 18- 26 jaar
- Niet schoolgaand of studerend
- Geen thuiswonend kind

- Kenmerken: met name relatief meer personen met 1e of 2e generatie migratieachtergrond, alleenstaanden, ongehuwde partners, relatief minder gehuwde partners.
- Armoedekans: 12% tegenover 1,2% landelijk.
- Risicofactoren: de top-20 belangrijkste kenmerken komt grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo zijn de kenmerken economische onzelfstandigheid en alleenstaand belangrijker dan voor heel Nederland. Net als landelijk hebben alleenstaanden en economisch zelfstandigen een grotere kans om arm te worden dan niet-alleenstaanden en niet-economisch zelfstandigen. Met name doordat dit effect sterker is dan landelijk, zijn deze kenmerken belangrijker. Ook is het kenmerk huurwoning met huurtoeslag belangrijker dan landelijk, met name doordat dit relatief vaak voorkomt.



Niet-arme werkenden

- Niet arm
 - Werknemer, directeur-grotaandehouder, zelfstandig ondernemer, overig zelfstandige of meewerkend gezinslid
-
- Kenmerken: met name relatief meer 20-60-jarigen, mannen, gehuwde partners met kinderen en relatief minder gehuwde partners zonder kinderen.
 - Armoedekans: 0,7% tegenover 1,2% landelijk.
 - Risicofactoren: de top-20 belangrijkste kenmerken komen overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo komen in deze deelpopulatie –per definitie- relatief meer voltijd werk en meer baanuren voor, waardoor het belang hiervan groter is dan voor heel Nederland.

Niet-arme alleenstaande ouders

- Niet arm
- Ouder in eenouderhuishouden

- Kenmerken: met name relatief meer 35-60-jarigen, personen met een eerste generatie migratieachtergrond, vrouwen.
- Armoedekans: 5% tegenover 1,2% landelijk.
- Risicofactoren: de top-20 belangrijkste kenmerken komen grotendeels overeen met de landelijke top-20. Wel zijn er verschillen in effecten te zien. Zo is woning met huurtoeslag belangrijker dan voor heel Nederland, met name doordat dit relatief veel voorkomt. Verder hebben bijvoorbeeld economisch onzelfstandigen een grotere kans om arm te worden dan economisch zelfstandigen. Dit effect is in deze deelpopulatie sterker dan landelijk gezien.

5. Conclusie



Conclusie

Beantwoording onderzoeksvragen:

- a. Hoe goed kunnen we op basis van (combinaties van) kenmerken de kans schatten dat iemand arm blijft dan wel arm wordt?
- b. Voor welke aspecten kan de combinatie van populatiedekkende registerdata en machine learning tot nieuwe inzichten over risicofactoren voor transitie in en uit armoede leiden?

Hoe goed kunnen we op basis van (combinaties van) kenmerken de kans schatten dat iemand arm blijft dan wel arm wordt?

- Voor zowel de arme als niet-arme populatie kan het model arme en niet-arme personen vrij goed van elkaar onderscheiden. Hierdoor zijn de modellen geschikt om in kaart te brengen hoe potentiële risicofactoren de armoedekans beïnvloeden.
- De modellen zijn minder geschikt om op individueel niveau personen te classificeren als wel of niet arm. Daarbij is beter te schatten dat iemand arm blijft dan dat iemand arm wordt.

Voor welke aspecten kan de combinatie van populatiedekkende registerdata en machine learning tot nieuwe inzichten over risicofactoren voor transitie in en uit armoede leiden?

- *Life events*: de modellen bieden naast inzicht in statische kenmerken ook inzicht in life events.
- *Deelpopulaties*: de modellen bieden inzicht in risicofactoren voor specifieke deelpopulaties. Daarbij is het mogelijk om in te zoomen op zowel data- als expertgedreven deelpopulaties. Ook is het mogelijk om een verschil tussen deelpopulatie en gehele populatie te ontleden in een verschil in samenstelling van de (deel)populatie en een verschil in effect van een risicofactor.
- *Effecten*: gedetailleerde niet-lineaire verbanden tussen risicofactor en armoedekans kunnen met de modellen in kaart worden gebracht.
- *Ranking*: de ruim 500 potentiële risicofactoren in de modellen kunnen gerangschikt worden.
- *Overgangen*: de modellen bieden inzicht in risicofactoren die samenhangen met overgangen in en uit armoede.

6. Discussie



Discussie

- a. Mogelijke toepassingen
- b. Aanknopingspunten voor beleid
- c. Aandachtspunten

Mogelijke toepassingen

- Gemeentes kunnen de uitkomsten van dit onderzoek gebruiken om specifieke doelgroepen voor beleid te bepalen. Daarbij kunnen ze onderscheid maken tussen de arme en niet-arme populatie.
- Op basis van ranking van risicofactoren kunnen gemeentes focus in beleid aanbrengen.

Aanknopingspunten voor beleid

Via de modellen is veel informatie beschikbaar. In samenwerking met gemeentes kan verkend worden wat er nodig is om ervoor te zorgen dat dit kan bijdragen aan gemeentelijk armoedebeleid. Bijvoorbeeld:

- Verder inzoomen door:
 - kenmerken verder te specificeren, bijvoorbeeld in- en uitstroom van bijstand, wijziging in partnerschap;
 - (niet-lineaire) verbanden tussen kenmerken in kaart te brengen, bijvoorbeeld leeftijd en wel/niet thuis wonen, leeftijd en huishoudensamenstelling, bijstand en geboorte;
 - in te zoomen op andere deelpopulaties, bijvoorbeeld deelpopulaties die uit de modellen naar voren komen met relatief hoge of lage armoedekansen.
- Modeluitkomsten confronteren met gemeentelijke regelingen. Bijvoorbeeld: in hoeverre komt het bereik van inkomensgerelateerde gemeentelijke regelingen overeen met de hoogrisicogroepen uit de modellen. Dit kan in kaart gebracht worden door data over gemeentelijke regelingen te koppelen aan hoogrisicokenmerken uit de modellen.



Aandachtspunten

- De modellen zijn bedoeld voor uitspraken over groepen, niet voor het classificeren van individuen.
- Relaties tussen kenmerken en armoede zijn correlaties en niet noodzakelijkerwijs causale effecten. Zo is huurwoning met huurtoeslag een goede voorspeller van armoede, maar waarschijnlijk is dit niet de oorzaak.
- De modellen focussen op voorspelkwaliteit, niet op kenmerken die beïnvloedbaar zijn door armoedebeleid.
- Deze eerste versie van de modellen is een structuurmodel, dat wegens beschikbaarheid van inkomensgegevens van de belastingdienst getraind is op data over 2011-2018. Dit betekent dat de modellen inzicht geven in algemene risicofactoren voor armoede en niet in specifieke risicofactoren gerelateerd aan de coronacrisis.