

TOR 2005/55



# Technisch verslag: Respons en wegingscoëfficiënten SONAR-cohorten

Ilse Laurijssen

Onderzoeksgroep TOR  
Vakgroep Sociologie  
Vrije Universiteit Brussel





## **Inhoudstafel**

|   |           |
|---|-----------|
| <b>1. Inleiding.....</b>  | <b>1</b>  |
| <b>2. Welke indicaties zijn er over de selectiviteit van de respons?.....</b> | <b>3</b>  |
| 2.1. Wat vertellen de globale responscijfers ons?.....                        | 5         |
| 2.2. Socio-demografische kenmerken van niet-deelnemers.....                   | 7         |
| 2.3. Gegevens over de woning.....   | 9         |
| <b>3. Representativiteit en weging van de cohorten.....</b>                   | <b>11</b> |
| 3.1. De steekproeven volgens socio-demografische kenmerken.....               | 12        |
| 3.2. Afleiding van de wegingsfactoren.....                                    | 14        |
| 3.3. Verdere toetsing van de veralgemeenbaarheid over de cohortes.....        | 17        |
| 3.3.1. Onderwijsniveau cohorte 1980 op 21 jaar.....                           | 17        |
| 3.3.2. Minimum kwalificaties vergeleken tussen de cohorten.....               | 18        |
| <b>4. Conclusie.....</b>  | <b>20</b> |
| <b>5. Addendum : wegingscoëfficiënten op 26 jaar.....</b>                     | <b>21</b> |
| 5.1. Toevoeging: wegingscoëfficiënten cohorte 76(26).....                     | 22        |
| <b>Bibliografie.....</b>  | <b>25</b> |
| <b>Bijlagen.....</b>  | <b>26</b> |
| Vergelijkbaarheid onderzoeksgegevens met bevolkingsgegevens.....              | 27        |
| Wegingscoëfficiënten cohorte 76(23).....                                      | 29        |
| Wegingscoëfficiënten cohorte 78(23).....                                      | 30        |
| Wegingscoëfficiënten cohorte 80(23).....                                      | 31        |
| Wegingscoëfficiënten cohorte 76(26).....                                      | 32        |



## 1. Inleiding

In dit verslag worden de wegingscoëfficiënten beschreven voor verschillende gerealiseerde steekproeven in het kader van het SONAR-project.

De ‘Studiegroep van ONderwijs naar ARbeidsmarkt’ - kortweg SONAR - is een interuniversitaire onderzoeksgroep (Universiteit Gent, Universiteit Antwerpen, Vrije Universiteit Brussel, Hoger Instituut Voor de Arbeid) met een interdisciplinaire samenstelling (sociologen, economen, pedagogen, psychologen). SONAR werd opgericht naar aanleiding van onderzoeksinitiatieven van de Afdeling Beleidscoördinatie van het Departement Onderwijs van de Vlaamse Gemeenschap om onderzoek uit te voeren over de manier waarop jongeren de overgang van school naar werk maken<sup>1</sup>.

### Overzicht van de SONAR dataverzameling:

|      | Cohorte 1976 | Cohorte 1978 | Cohorte 1980 |
|------|--------------|--------------|--------------|
| ...  |              |              |              |
| 1999 | op 23 jaar   |              |              |
| 2000 |              |              |              |
| 2001 |              | op 23 jaar   |              |
| 2002 | op 26 jaar   |              |              |
| 2003 |              |              | op 23 jaar   |
| 2004 |              | op 26 jaar   |              |
| 2005 | op 29 jaar   |              |              |

Het onderzoeksprogramma van SONAR omvat een geregelde bevraging van leeftijdscohortes en hun herbevraging op latere tijdstippen (zie bovenstaande tabel). Drie geboortecohorten werden bevraged op 23-jarige leeftijd: jongeren geboren in 1976 in het najaar van 1999 (N=3015), geboortecohorte 1978 in 2001 (N=3002), en jongeren met geboortjaar 1980 in 2003 (N=2993). Het zijn deze drie groepen respondenten, hierna ook aangeduid als c76, c78 en c80, die in dit verslag aan bod komen. Deze 23-jarigen worden door SONAR eveneens op latere leeftijd herbevraged, namelijk op 26 jaar (voor c76 in 2002, voor c78 in 2004), en voor de eerst opgestarte geboortecohorte is eveneens een laatste enquête op 29-jarige leeftijd voorzien.

In dit document staat de representativiteit van de basissteekproeven op 23 jaar centraal. De representativiteit van de steekproef wordt bepaald door de onder- of oververtegenwoordiging van bepaalde bevolkingscategorieën. In een representatieve steekproef komen de verschillende categorieën in dezelfde verhouding voor als in de bevolking, en geeft de steekproef een correct beeld van de volledige populatie. Afwijkingen op de evenredige vertegenwoordiging brengen de veralgemeenbaarheid van vaststellingen in gevaar. Daarom wordt hier onderzocht in welke mate de gerealiseerde steekproeven representatief zijn voor de populatie van Vlaamse 23-jarigen in de respectieve geboortecohorten.

Hiervoor wordt gestart met een overzicht van de responscijfers. Hoe lager de globale respons, hoe meer ruimte voor selectieve uitval. Op zich is uitval geen probleem, wel als de uitval structureel is, dus wanneer ze sterker geconcentreerd is in bepaalde bevolkingscategorieën dan andere. In de mate dat de (non) respons niet toevallig is, komt de representativiteit en dus de veralgemeenbaarheid van de resultaten van het onderzoek naar de bevolking in het gedrang.

<sup>1</sup> Het onderzoeksprogramma van SONAR werd gefinancierd door het Ministerie van de Vlaamse Gemeenschap in het kader van het Programma Beleidsgericht Onderzoek (contractnummers PBO97/15/42 – PBO98/56/76 – PBO99B/4/13), nadien via het Steunpunt voor Beleidsrelevant Onderzoek “Loopbanen van leerlingen en studenten in het onderwijs en de overgang van onderwijs naar arbeidsmarkt”.

## Inleiding

Een vergelijking van de respons-rates tussen verschillende groepen, geeft een eerste indicatie van de kwaliteit van de steekproefsamenstelling en de kwaliteit van de onderzoeksgegevens. Vervolgens bekijken we de selectiviteit in respons aan de hand van extra gegevens verkregen via de contactbladen. Hierop registreerden de enquêteurs registreerden bij de contactpogingen een aantal gegevens over respondenten die weigerden deel te nemen. Door de kenmerken van weigeraars en deelnemers aan het onderzoek met elkaar te vergelijken, onderzoeken we verder of er een logica zit in de uitval.

Vervolgens toetsen we de representativiteit van de gerealiseerde steekproeven rechtstreeks door de onderzoeksgegevens te vergelijken met de samenstelling van de volledige populatie op een aantal socio-demografische kenmerken (verstedelijkingsgraad, geslacht, en onderwijsniveau). We stellen een lichte vertekening vast die aansluit bij resultaten van ander survey-onderzoek, namelijk dat sociaal-economisch zwakkere categorieën minder goed vertegenwoordigd zijn. Gelukkig is de afwijking niet te groot, en kan ze worden geredieerd door het gewicht van de betrokken categorieën in de steekproef aan te passen. Op basis van de bevolkingsgegevens worden tot slot de wegingscoëfficiënten bepaald en verder getoetst. Die uiteindelijke wegingscoëfficiënten kunnen dan bij de latere data-analyses gebruikt worden om veralgemeenbare uitspraken te doen.

## 2. Welke indicaties zijn er over de selectiviteit van de respons?

Voor de enquêtering van de drie cohorten werd telkens een gelijkaardige twee-trapste steekproeftrekking gevolgd om tot een bevraging van 3000 respondenten te komen<sup>2</sup>. In een eerste stap werden 100 regionaal omschreven steekproefpunten geselecteerd, a rato van het aantal jongeren tussen 20 en 24 jaar in de Vlaamse gemeenten. Per steekproefpunt werden vervolgens personen uit het betrokken geboortjaar geselecteerd via het rijksregister, telkens 120 personen, die werden ingedeeld in 30 reeksen van eerste adressen en drie vervangadressen.

Via deze selectieprocedure werd het veldwerk telkens gestart met een a-selecte of toevallige steekproef, waarvan representativiteit verwacht kan worden. Vervolgens werden diverse mechanismen ingebouwd om een zo hoog mogelijke respons te bereiken, en selectiviteit in de respons te beperken.

### 1. Wijze van bevraging

Bij elke enquête verliep de dataverzameling via een face-to-face interview. Er werd voor deze methode geopteerd omdat ze een complexer en langer interview toelaat, maar ze staat er ook om bekend dat ze de beste garantie biedt om personen zo goed mogelijk te bereiken. Bij andere methoden, als men al kan vertrekken van een aselekt staal van de bevolking, ligt het probleem vooral in de lagere bereidheid tot deelname omdat men zich niet rechtstreeks aangesproken voelt.

### 2. Vertrouwen en motivatie respondenten

De respondenten werden vooraf per brief, en nadien ook door de enquêteur, op de hoogte gebracht van het onderzoek en gemotiveerd om mee te werken. De respondenten konden eveneens bij de coördinator of de onderzoekers van het project informatie inwinnen.

### 3. Opvang non-respons door vervanging

De potentiële respondenten werden gegroepeerd in aan elkaar gekoppelde adressen om weigeringen en afwezigheden op te kunnen vangen. Via deze methode kon doorgaans gemakkelijk het beoogde quotum van 30 respondenten per steekproefpunt gehaald worden. Hieruit volgt logischerwijze ook een respons gelijk verdeeld over de geselecteerde regio's.

De adressen werden bovendien niet louter toevallig gegroepeerd, maar binnen de gemeenten aan elkaar gekoppeld op basis van het geslacht, de postcode, en soms zelfs de straat. Deze koppeling van 'gelijkaardige' personen aan elkaar, hoewel niet steeds tot in het laatste detail doorgevoerd, geeft de bijkomende garantie dat een persoon die niet geënquêteerd kon worden, eerder 'vervangen' zou worden door een persoon met gelijkaardige kenmerken.

### 4. Gecontroleerde contactprocedure

Tot slot werd de enquêteurs een zeer strikte contact- en vervangprocedure opgelegd. Enkel wanneer bepaalde voorwaarden voldaan waren, maar nog geen interview werd afgenomen, mochten de enquêteurs doorgaan naar het volgende adres. Om te beginnen diende men de respondent persoonlijk te bereiken. Weigeringen door anderen, bijvoorbeeld huisgenoten, mochten niet als definitieve weigering beschouwd worden. Pas na ofwel een rechtstreekse weigering tot deelname, ofwel drie contactpogingen, op verschillende tijdstippen van de dag en de week, mocht naar het volgende adres worden overgegaan. Bij personen die tijdelijk belemmerd waren, die bijvoorbeeld stelden weinig tijd te hebben, diende een afspraak gemaakt te worden.

Deze procedures dienden de nodige garanties in te bouwen om telkens een zo representatief mogelijk samengestelde steekproef van respondenten te realiseren. Maar elk sociaal-wetenschappelijk survey-

<sup>2</sup> Voor de details van het opzet en het verloop van elk van de enquêtes op 23 jaar, verwijzen we naar de technische rapporteringen van de betrokken cohorten: c76(23): Coppieters et. al. (2002); c78(23): Belet et. al. (2003); c80(23): Belet et. al. (2004b).

Welke indicaties zijn er over de selectiviteit van de respons?

onderzoek wordt geconfronteerd met de vraag naar de veralgemeenbaarheid van de verzamelde gegevens. In de volgende paragraaf bekijken we daarom verder welke indicaties we hebben over de verhouding tussen en de verschillen tussen wie wel en wie niet deelnam aan het onderzoek.



## 2.1. Wat vertellen de globale responscijfers ons?

In dit deel geven we een overzicht van de respons van de drie uitgevoerde onderzoeken. De verhouding tussen deelnemers en niet-deelnemers, geeft de sterkte van de uitval in de onderzoeken weer.

Voor elk van de drie onderzochte cohorten vermeldt onderstaande tabel het bruto- en het netto-respons percentage<sup>3</sup>. Het bruto-respons cijfer (het aandeel geslaagde enquêtes op het geheel van gecontacteerde personen) blijkt het hoogst in de laatste cohorte, en het laagst in de middelste cohorte. De variaties naar geslacht zijn relatief beperkt, naar provincie zijn de verschillen groter, waar West-Vlaanderen en Limburg telkens gekenmerkt worden door een hogere respons, en Vlaams-Brabant door een relatief lagere respons. De belangrijkste categorie van non-respons betreft telkens respondenten die expliciet weigerden deel te nemen aan het onderzoek. Bepaalde non-respons categorieën wijzen echter eerder op een onmogelijkheid om aan het onderzoek deel te nemen, dan een 'echte' weigering, namelijk personen die verhuisd, overleden, zwaar ziek of langdurig afwezig waren, en andere non-respons categorieën geven aan dat de respondent niet tot het doelpubliek van het onderzoek hoort, namelijk personen die nooit onderwijs in België volgden of het Nederlands onvoldoende beheersen. Wanneer we deze personen buiten beschouwing laten dan bekomen we het netto-responscijfer dat in elke cohorte ongeveer 10 procentpunten hoger ligt dan het bruto-cijfer.

### Responscijfers, naar geslacht en provincie

|                  | C76(23) |               |               | C78(23) |               |               | C80(23) |               |               |
|------------------|---------|---------------|---------------|---------|---------------|---------------|---------|---------------|---------------|
|                  | %       | Bruto respons | Netto respons | %       | Bruto respons | Netto respons | %       | Bruto respons | Netto respons |
| <b>Geslacht</b>  |         |               |               |         |               |               |         |               |               |
| man              | 51,2    | 58,8          | 67,2          | 50,3    | 54,1          | 62,8          | 50,9    | 62,3          | 72,0          |
| vrouw            | 48,7    | 56,1          | 65,4          | 49,7    | 54,5          | 63,6          | 49,1    | 63,8          | 74,3          |
| <b>Provincie</b> |         |               |               |         |               |               |         |               |               |
| Antwerpen        | 27,2    | 51,4          | 62,4          | 27,0    | 51,0          | 60,3          | 31,0    | 63,5          | 73,1          |
| Vlaams Brabant   | 16,4    | 52,7          | 63,0          | 16,8    | 45,6          | 53,7          | 15,0    | 57,8          | 72,1          |
| West-Vlaanderen  | 18,5    | 66,0          | 70,2          | 19,8    | 60,4          | 69,1          | 16,0    | 69,2          | 75,7          |
| Oost-Vlaanderen  | 23,7    | 59,3          | 67,5          | 23,8    | 57,4          | 66,7          | 21,0    | 61,9          | 72,7          |
| Limburg          | 14,2    | 64,9          | 72,0          | 12,7    | 62,5          | 70,3          | 17,0    | 63,5          | 72,3          |
| <b>Totaal</b>    | 100,0   | 57,5          | 66,4          | 100,0   | 54,3          | 63,2          | 100,0   | 63,0          | 73,1          |

Uiteraard geven de netto-cijfers nog steeds een globaal beeld dat niet te reduceren is tot louter de motivatie van de respondent. Ook de moeite en overredingskracht die de enquêteur aan de dag legt kan van belang zijn, naast de rol van technische elementen en de beoordeling door de onderzoekers van de enquêtes in het mislukken of weerhouden van een eventueel begonnen enquête.

Deze responscijfers kunnen wel een indicatie geven van moeilijker of gemakkelijker bereikbare groepen, maar door het gebruik van gekoppelde adressen kan dit weinig invloed hebben gehad op de uiteindelijke verdeling van de steekproef over geslacht en provincie. Zo blijken de verdelingen over die kenmerken van de cohorten ondanks de variaties in de responscijfers, redelijk stabiel. Elk adres waar geen enquête gerealiseerd kon worden, diende immers in een vooraf vastgelegde volgorde vervangen te worden door een ander adres in dezelfde provincie, dat zo veel mogelijk iemand betrof

<sup>3</sup> De cijfers zijn gebaseerd op de technische rapporteringen van de drie cohorten; de netto-respons werd telkens op dezelfde wijze afgeleid zoals dat beschreven staat voor c80(23), mogelijkheid van afrondingsfouten.

## Welke indicaties zijn er over de selectiviteit van de respons?

met hetzelfde geslacht in dezelfde gemeente of zelfs straat. Op die manier zijn variaties in spreiding over de provincie en geslacht, niet louter het gevolg van selectieve respons vanuit de respondenten of enquêteurs, maar ook – en vooral – een indicatie van verschillen in samenstelling van de oorspronkelijke steekproef.

Uiteraard kan de non-respons nog steeds selectief zijn op kenmerken waarvoor het onderzoeksdesign geen speciale maatregelen nam.

Sociaal-wetenschappelijke onderzoeken worden typisch gekenmerkt door een lagere deelname van personen uit sociaal zwakkere sociale groepen dan van personen uit hogere sociale strata. De vergelijking van de kenmerken van deelnemers en niet-deelnemers zegt ons onrechtstreeks iets over de onder- of oververtegenwoordiging van bepaalde groepen in de onderzoeksgegevens. Vaak stelt zich echter het probleem dat onderzoekers wel gegevens hebben over de respondenten die meewerkten aan het onderzoek, maar niet specifiek over de respondenten die weigerden, aangezien bij deze groep geen enquête werd afgenomen.

Voor twee van de drie SONAR cohortes beschikken we echter over bijkomende informatie over de personen die niet deelnamen aan het onderzoek, die we kunnen vergelijken met de kenmerken van de respondenten zelf. Voor de cohorten 76(23) en 80(23) zijn een aantal gegevens gevraagd aan mensen die niet wilden of konden deelnemen aan het onderzoek. Bij cohorte c80(23) dienden de enquêteurs bovendien de woning van alle personen die ze bezochten (dus ook de respondenten) te beoordelen. Beide soorten gegevens komen hierna aan bod, om de mogelijke vertekening door de uitval beter in kaart te brengen.

## 2.2. Socio-demografische kenmerken van niet-deelnemers

Bij twee van de drie onderzoeken, werden respondenten die niet wilden of konden deelnemen, verzocht om toch een beperkt aantal algemene vragen te beantwoorden. Het ging daarbij vooral om basiskenmerken als huidige werksituatie en het behaalde onderwijsniveau. Op die manier hoopten we iets te weten te komen over de verschillen of gelijkenissen in de profielen van deelnemers en niet-deelnemers. Een groot succes vormt die werkwijze niet, met slechts een zeer beperkt aantal personen dat niet deelnam aan het onderzoek, maar wel bereid was om de bijkomende gegevens te verstrekken, of waarvan die gegevens op zijn minst bekend zijn: bij c76(23) gaat het in totaal om 190 personen, bij c80(23) om 181. Een ruime basis om te vergelijken met de gegevens van de respondenten die uiteindelijk wel in het bestand zijn opgenomen, vormen deze data dus niet. Daarom worden de vergelijkingen hier gedaan over ruime categorieën.

### Kenmerken van niet-deelnemers vergeleken met de respondenten

|                                     | C76(23)         |              | C80(23)         |              |
|-------------------------------------|-----------------|--------------|-----------------|--------------|
|                                     | Niet-deelnemers | Respondenten | Niet-deelnemers | Respondenten |
| <b>Huidige situatie<sup>4</sup></b> | (N=185)         | (N=3010)     | (N=181)         | (N=2993)     |
| Werkend                             | 78,9%           | 73,5%        | 76,2%           | 65,9%        |
| Student                             | 16,2%           | 15,4%        | 12,7%           | 15,6%        |
| <b>Onderwijsniveau<sup>5</sup></b>  | (N=177)         | (N=3008)     | (N=181)         | (N=2993)     |
| Geen diploma secundair onderwijs    | 10,2%           | 8,4%         | 5,0%            | 9,2%         |
| Beroepssecundair onderwijs          | 19,2%           | 22,5%        | 17,7%           | 21,9%        |
| Technisch secundair onderwijs       | 19,2%           | 17,2%        | 28,7%           | 18,0%        |
| Algemeen secundair onderwijs        | 19,2%           | 15,8%        | 8,8%            | 12,4%        |
| Hoger onderwijs één cyclus          | 20,3%           | 24,0%        | 17,7%           | 24,8%        |
| Universitair onderwijs              | 11,9%           | 12,1%        | 14,9%           | 13,8%        |
| Andere                              |                 |              | 7,2%            |              |

Hoewel de vraagstelling naar huidige situatie ruimte laat voor interpretatie en de vergelijking niet sluitend maakt, lijkt het voor beide cohorten dat bij de niet-deelnemers meer werkenden zijn dan bij de deelnemers, vooral bij c80 is het verschil groter. Het aandeel studenten verschilt eveneens bij de deelnemers en degenen die weigerden deel te nemen, maar niet zo duidelijk want in tegengestelde richting in beide cohorten. Verschillen tussen de cohorten zouden te wijten kunnen zijn aan veranderde arbeidsmarktomstandigheden, maar hoe deze de verschillen tussen deelnemers en niet-deelnemers zouden beïnvloeden is minder duidelijk.

Het onderwijsniveau is ook niet helemaal vergelijkbaar, maar de indeling in categorieën is zo vergelijkbaar mogelijk gemaakt. Een moeilijkheid is dat bij de niet-deelnemers van c80 een ruime groep “andere” geantwoord heeft, waarvan het grootste deel wellicht te rekenen is bij de laagst

4 Het gaat om de hoofdsituatie; de niet-deelnemers van c76(23) werd eveneens apart gevraagd of men nog bezig was met hoger onderwijs, waar het percentage duidelijk hoger ligt (21,0% - N=181), te wijten aan de combinatie van studeren met werken; voor de respondenten werd voorrang gegeven aan werken (bij c80 overlap tussen beide bij 1,4%; bij c76 onderschatting studenten, want in combinatie met nooit gewerkt).

5 BSO is inclusief deeltijds onderwijs en post-secundair niet-hoger onderwijs; TSO is inclusief KSO; universitair onderwijs is inclusief het hoger onderwijs buiten de universiteit van 2 cycli; bij niet-deelnemers van cohorte c80(23) werd het deeltijds onderwijs niet expliciet gevraagd; een gedeelte daarvan zit wellicht in de categorie “andere”.

## Welke indicaties zijn er over de selectiviteit van de respons?

geschoolden. Bij beide cohorten lijkt de laagst geschoolde groep eerder ondervertegenwoordigd, net zoals de respondenten met een diploma technisch secundair onderwijs. De hoogst geschoolden, en dan vooral degenen met een diploma hoger onderwijs van één cyclus, lijken oververtegenwoordigd in de gerealiseerde steekproef, net zoals de respondenten met een diploma beroepssecundair onderwijs. Wat de extremen betreft, komen de vaststellingen overeen met de verwachting van hogere deelname door sociaal sterkere groepen.

Naast de gegevens vermeld in de tabel, werden niet-deelnemers van c80(23) ook gevraagd naar de nationaliteit: slechts twee respondenten antwoordden een andere dan Belgische nationaliteit te hebben (1,1%), terwijl 2,5% van de gerealiseerde steekproef een buitenlandse nationaliteit opgeeft. Deze cijfers geven geen aanwijzing voor een ondervertegenwoordiging van allochtonen in het onderzoek. De meest voor de hand liggende conclusie hier lijkt echter dat de gegevens over de niet-deelnemers wellicht onvoldoende betrouwbaar zijn om te veralgemenen naar alle niet-deelnemers, en bijgevolg geen goede basis vormen voor de inschatting van selectieve respons. Het aantal niet-deelnemers waarvoor we over informatie beschikken is niet alleen relatief klein, maar bovendien neemt de deze groep vermoedelijk een selectieve positie in binnen de totale groep van niet-respondenten.<sup>6</sup>

---

6 Verdere analyse van de non-respons lijkt aan te geven dat die groep niet-deelnemers waarvan die informatie verzameld kon worden, niet noodzakelijk een goede afspiegeling vormt van alle niet-deelnemers. Om te beginnen betreft het uiteraard het meest bereikbare segment uit de niet-deelnemers. Zo werd de meerderheid van deze personen wel persoonlijk gecontacteerd (79% tegenover 40%). Net zoals globaal vormen weigeringen de hoofdreden voor niet-deelname, maar in vergelijking met de volledige non-respons is deze reden veel uitgesprokener (72% tegenover 33% in c76), de tweede reden voor respons, problemen met de adressen is daarentegen veel minder aanwezig (5% tegenover 18%).

### 2.3. Gegevens over de woning

Op basis van de responscijfers, weten we wel hoe groot de uitval is, maar blijft de vraag naar de selectiviteit in de uitval open. De vergelijking van kenmerken van deelnemers en niet-deelnemers aan een onderzoek, geeft onrechtstreeks een indicatie over de sociale achtergrond van de (non)participanten. Hierboven rapporteerden we het profiel van niet-deelnemers waarvan toch nog de werksituatie en het onderwijsniveau bekend was. Dit blijkt echter onvoldoende betrouwbaar als vergelijkingsbasis.

Bij c80(23) registreerden de enquêteurs ook een aantal gegevens over de woning van alle personen, ongeacht hun deelname, op het contactblad. Deze gegevens worden courant beschouwd als een indicator van de sociaal-economische status van de respondent, en kunnen hier gebruikt worden om een beter beeld te verkrijgen van de sociale positie van de weigeraars.

Van 4822 bezochte personen noteerden de enquêteurs welk type van woning het betrof, en of het gebouw beschikte over een parlofoon<sup>7</sup>. We vergelijken de antwoorden voor de groep respondenten met de antwoorden van degenen die niet deelnamen aan het onderzoek.

**Welke van de volgende omschrijvingen past het best voor de woning van de respondent ?  
Beschikt de woning / het gebouw waarin de woning is gelegen over een parlofoon ?**

|  | <b>niet-<br/>deelnemer<br/>s</b> | <b>deelnemer<br/>s</b> | <b>totaal</b> | <b>kans op<br/>deelname</b> |
|--|----------------------------------|------------------------|---------------|-----------------------------|
| <b>Type woning</b>                           |                                  |                        |               |                             |
| Eengezinswoning                              | (64,3%)                          | (79,3%)                | (73,7%)       | (67,9%)                     |
| • alleenstaand                               | 30,2%                            | 41,8%                  | 37,7%         | 70,4%                       |
| • half-open bebouwing                        | 12,2%                            | 14,0%                  | 13,2%         | 66,3%                       |
| • rijhuis of gesloten bebouwing              | 21,9%                            | 23,5%                  | 22,8%         | 64,8%                       |
| Een boerderij of hoeve (nog in bedrijf)      | 0,6%                             | 1,3%                   | 1,0%          | 79,2%                       |
| Appartement, woning in meergezinswoning      | (27,9%)                          | (15,0%)                | (19,8%)       | (48,1%)                     |
| • met in totaal 2 woningen                   | 2,7%                             | 2,1%                   | 2,3%          | 57,3%                       |
| • met in totaal 3 tot 9 woningen             | 17,8%                            | 9,9%                   | 12,8%         | 48,9%                       |
| • met in totaal meer dan 9 woningen          | 7,4%                             | 3,0%                   | 4,6%          | 41,2%                       |
| Enige woning in gebouw met andere bestemming | 1,1%                             | 1,5%                   | 1,4%          | 69,7%                       |
| Andere                                       | 6,1%                             | 2,9%                   | 4,1%          | 45,4%                       |
| <b>Parlofoon?</b>                            |                                  |                        |               |                             |
| ja   | 33,7%                            | 23,1%                  | 27,0%         | 54,1%                       |
| neen   | 66,3%                            | 76,9%                  | 73,0%         | 66,6%                       |
| <b>Totaal aantal</b>                         | 1746                             | 3001                   | 4747          | 4747                        |

Gegeven de veel grotere aantallen, geven deze cijfers een betere indicatie van de mogelijke selectiviteit van de respons. Bovendien speelt hier niet zozeer de respondent een rol (die niet de toestemming moet geven), maar wel de enquêteur die niet geneigd zal zijn naar het sociaal-wenselijke te antwoorden (de gegevens zijn ook voor quasi alle gecontacteerden beschikbaar; enkel

<sup>7</sup> Deze 4822 personen hebben naast de effectieve respondenten (3001) en de niet-respondenten (1746 ; bepaald op basis van de contactgegevens), ook betrekking op 75 andere personen. We veronderstellen dat die bijkomende 75 personen misschien wel bezocht werden door de enquêteurs, maar dan louter anticiperend op een mogelijke contactname, en rekenen hen hier dan ook niet bij de niet-deelnemers.

## Welke indicaties zijn er over de selectiviteit van de respons?

voor 13 gecontacteerde personen ontbreken de gegevens). En zelfs van respondenten die verhuisd bleken, krijgen we via deze methode een zicht op de socio-economische status.

Er blijkt een vrij duidelijke lijn te zitten in de deelnamekansen van de bewoners van bepaalde woningtypes. Degenen die in een appartement wonen hebben een beduidend lagere kans op deelname aan het onderzoek dan wie in een eengezinswoning woont. Dit gaat ook op voor die respondenten waarvan de woning als “andere” geklasseerd werd, maar het feit dat dit woningtype beduidend vaker is aangestipt bij niet-deelnemers doet vermoeden dat enquêteurs de woning van niet-deelnemers minder vaak effectief beoordeeld hebben. De relatief hoge deelname cijfers voor mensen wonend in een boerderij of een gebouw met een andere bestemming, betreffen geen statistisch significant hogere kans op deelname in vergelijking met een (alleenstaande) eengezinswoning. Degenen die wonen in een boerderij hebben wel vaker deelgenomen dan degenen die in een appartement wonen.

Binnen de globale woningtypes zijn er nog duidelijke (statistisch significante) verschillen: naarmate de eengezinswoning meer in gesloten bebouwing ligt, daalt de kans op deelname, wie in een appartement woont, heeft minder deelgenomen aan de enquête naarmate het appartementsgebouw meer woningen omvat.

De aanwezigheid van een parlofoon gaat gepaard met een beduidend lagere kans op deelname aan het onderzoek. Het negatief effect van een parlofoon wordt wel quasi volledig verklaard door het feit dat de appartementen, en dan vooral die in grote gebouwen, vaker over een parlofoon beschikken.

Wanneer we veronderstellen dat hogere sociale strata eerder een alleenstaand huis bewonen, dan bevestigen deze resultaten dat hogere sociale strata meer participeren aan survey-onderzoek. Uit de voorgaande analyses kan worden besloten dat de non-respons eenduidig samenhangt met het type woning: bewoners van een vrijstaande woning, een eengezinswoning in half-open bebouwing of een boerderij participeren meer dan de bewoners van een rijhuis. De meergezinswoningen kennen de grootste kans op niet-deelname. De anonimiteit van dergelijke woningtypes is vermoedelijk groter dan bij alleenstaande woningen in een plattelandsomgeving, wat weigeren of niet open doen vergemakkelijkt. Dezelfde anonimiteitsverklaring gaat ook op voor de lagere kans op deelname wanneer er een parlofoon aan de woning aanwezig is. De aanwezigheid van dergelijke technologie bemoeilijkt duidelijk het rechtstreeks contact van de enquêteur met de respondent.

### 3. Representativiteit en weging van de cohorten

In het voorgaande hebben we indicaties gevonden voor een mogelijke ondervertegenwoordiging van mensen uit lagere strata in de steekproeven van de drie onderzoeksc cohorten. Deze conclusie trekken we op basis van de samenhang tussen de kans op deelname en het type woning, gegevens waarover we enkel beschikken voor de laatste cohorte van 23-jarigen. De vraag die zich hier dan stelt is in welke mate deze mensen ondervertegenwoordigd zijn, in elk van de drie cohortes, en of we dit dan kunnen opvangen door wegingscoëfficiënten toe te passen. De beste manier om uitsluitel te krijgen over de representativiteit, is het vergelijken van de gekende kenmerken van de onderzoeksbevolking met de bevolkingsgegevens (Billiet, 1996:225). We zochten daarom naar gegevens over de bevolking (ic Vlaamse jongeren uit het overeenkomstige geboortjaar) waarmee we de verdeling op socio-demografische kenmerken in de respectievelijke steekproeven konden vergelijken. Aan de hand daarvan zullen vervolgens wegingscoëfficiënten bepaald worden.

Bevolkingsgegevens kennen we in Vlaanderen enkel voor een beperkt aantal socio-demografische kenmerken. Op basis van het Rijksregister is bijvoorbeeld wel een selectie op de juiste leeftijd of het juiste geboortjaar mogelijk, en het geslacht, de nationaliteit en de burgerlijke staat bekend, maar om werkelijk iets te weten over de sociaal-economische achtergrond of status zijn andere databronnen nodig. Grote periodieke dataverzamelingen, zoals de Europese arbeidskrachtentelling (EAK) bleken niet de geschikte vergelijkingsbasis te kunnen leveren, doordat het aantal 23-jarigen in dergelijke steekproeven relatief beperkt is (zelfs heel wat kleiner dan in de SONAR-steekproeven). Zowel qua selectiviteit van de respons als betrouwbaarheid van die enquêtécijfers, stelt zich dan de vraag of de resultaten wel correcter, representatiever voor de bevolking zijn, dan de resultaten van het eigen onderzoeksmateriaal.

Als referentiemateriaal werden gegevens van de Volkstelling 2001 voor de geboortecohorte 1978 genomen<sup>8</sup>. Deze databron heeft het voordeel van de volledige bevolking in kaart te brengen. De selectie van jongeren geboren in 1978 vormt een goed vergelijkingspunt, aangezien deze cohorte bij SONAR in 2001, hetzelfde jaar, werd bevraagd. Dit betekent volledige vergelijkbaarheid met één van de SONAR-steekproeven: dezelfde geboortecohorte, dezelfde leeftijd (tijdstip enquête<sup>9</sup>). Bovendien ligt deze geboortecohorte juist midden de twee andere geboortecohorten, bij veralgemening van deze referentiecijfers naar de beide andere steekproeven dient dan enkel twee jaar overbrugd te worden.

We beschrijven hier achtereenvolgens de verschillen die we vaststellen, hoe de steekproeven afwijken van de bevolking, op welke manier de wegingscoëfficiënten werden bepaald, en een toetsing van de wegingen.

---

8 De verdelingen voor de geboortecohorte 1978 in de Volkstelling 2001 werden bekomen via P. Deboosere en D. Willaert, Steunpunt Demografie, Vrije Universiteit Brussel.

9 Er is enkel de afwijking dat bij de Volkstelling het referentiemoment vast ligt op 1 oktober 2001, terwijl de situatie van de respondenten van de SONAR-steekproeven die op het tijdstip van enquête zelf is, waarbij de spreiding van het veldwerk een factor kan worden.

### **3.1. De steekproeven volgens socio-demografische kenmerken**

Drie indicatoren werden gebruikt om de gegevens op te kunnen ijken: geslacht, regio, en onderwijsniveau. Deze indicatoren werden gekozen omwille van hun beschikbaarheid en vergelijkbaarheid, en om hun relevantie (zowel qua aantallen als qua verwachte vertekening of selectiviteit). We hebben ook overwogen om op basis van andere kenmerken te vergelijken en nadien eventueel te wegen: nationaliteit (maar kleine aantallen), beroepsactiviteit (maar problematische afbakeningen), woon- en gezinssituatie (maar probleem vergelijkbaarheid).

Geslacht en regio worden als vergelijkingsindicatoren gebruikt, niet zozeer om selectieve respons te ondervangen, maar veeleer om ook te bekijken in welke mate afwijkingen geïnduceerd kunnen zijn door de samenstelling van de steekproeven (met de selectie van steekproefpunten en koppeling van startadressen aan vervangadressen, cfr. supra). Het onderwijsniveau wordt hier beschouwd als het belangrijkste kenmerk om de selectiviteit van de uitval te evalueren. Gegeven voorgaande indicaties dat deelnamegraden aan survey-onderzoek en het SONAR-onderzoek positief samenhangen met de sociaal-economische status, wilden we dit kenmerk expliciet toetsen. Voor deze groepen van 23-jarigen, die misschien nog geen definitieve of duidelijke status hebben op de arbeidsmarkt, werd geoordeeld dat het bereikte onderwijsniveau op 23 jaar de beste indicator van hun (huidige, maar wellicht ook toekomstige) sociale positie was. De indelingen op de SONAR-data en de Volkstellingcijfers werden zo goed als mogelijk vergelijkbaar gemaakt (vooral voor onderwijsniveau belangrijk), wat meer in detail toegelicht wordt in de bijlagen.

Wat stellen we nu vast wanneer we de bevolkingsgegevens van 23-jarigen vergelijken met die van de drie SONAR-steekproeven (zie onderstaande tabel)? We veronderstellen hier dat de bevolkingscijfers die gebaseerd zijn op jongeren geboren in 1978 ook een geldige vergelijkingsbasis vormen voor de cohorten 1976 en 1980 (verder wordt de veralgemeenbaarheid naar de andere cohorten ook nog getoetst).

Op de verdeling naar geslacht zit een lichte afwijking, al blijkt de lichte oververtegenwoordiging van mannen in de steekproeven, in feite ook terug te komen in de bevolking.

Naar regio zijn de verschillen groter. Over de cohortes gaan de afwijkingen ten opzichte van de bevolking in verschillende richtingen. Er blijkt dus geen duidelijk consistente oververtegenwoordiging van bij voorbeeld landelijke gemeenten en ondervertegenwoordiging van stedelijke gemeenten. Wellicht juist omdat de verdeling over de regio meer het gevolg is van hoe de steekproeven telkens zijn opgezet, eerder dan selectieve uitval door respondenten die weigeren. Van de drie cohorten lijkt c78 wel het sterkst afwijkende profiel naar regio te hebben, met een zeer sterke oververtegenwoordiging van de kleine stad, en lichte oververtegenwoordiging van respondenten in niet-stedelijke gemeenten, en juist een ondervertegenwoordiging van respondenten die in een grote of regionale stad wonen. Globaal komt dit patroon van ondervertegenwoordiging van verstedelijkte gebieden ook terug bij c76, maar wel minder sterk. Bij c80 lijkt de verdeling het sterkst op die van de bevolking, en als er al een systematische afwijking is, dan lijkt ze juist in tegengestelde richting te gaan, met een lichte oververtegenwoordiging van de steden.

Op onderwijsniveau lijken de afwijkingen iets systematischer. In elke steekproef ligt het aandeel laaggeschoolden zonder diploma secundair onderwijs telkens beduidend lager dan in de bevolking (8 à 9% tegenover 13%). Aan de andere kant is er een sterkere vertegenwoordiging van de personen met een diploma hoger onderwijs op 23 jaar, en nog het meest opvallend voor de groepen met een diploma van het hoger onderwijs van één cyclus (24 à 25% tegenover 22%). Verder zijn er lichtere afwijkingen: een iets hoger aandeel met een diploma van beroeps- en technisch secundair onderwijs, en een te lage vertegenwoordiging van personen met een diploma van algemeen secundair onderwijs.



## Wegingscoëfficiënten SONAR cohortes

### *Globale verdeling naar geslacht, regio en onderwijsniveau, voor weging*

|                                     | <b>Bevolking</b> |       | <b>C76(23)<sup>10</sup></b> |       | <b>C78(23)<sup>11</sup></b> |       | <b>C80(23)</b> |       |
|-------------------------------------|------------------|-------|-----------------------------|-------|-----------------------------|-------|----------------|-------|
| <b>Geslacht</b>                     |                  |       |                             |       |                             |       |                |       |
| Mannen                              | 37325            | 50,6% | 1544                        | 51,2% | 1509                        | 50,3% | 1528           | 51,1% |
| Vrouwen                             | 36482            | 49,4% | 1471                        | 48,8% | 1493                        | 49,7% | 1465           | 48,9% |
| <b>Regio</b>                        |                  |       |                             |       |                             |       |                |       |
| Grote stad                          | 8717             | 11,8% | 332                         | 11,0% | 239                         | 8,0%  | 358            | 12,0% |
| Regionale stad                      | 10595            | 14,4% | 414                         | 13,7% | 329                         | 11,0% | 479            | 16,0% |
| Kleine stad                         | 15635            | 21,2% | 575                         | 19,1% | 825                         | 27,5% | 630            | 21,0% |
| Niet-stedelijke gemeente            | 38860            | 52,7% | 1694                        | 56,2% | 1605                        | 53,5% | 1526           | 51,0% |
| <b>Onderwijsniveau<sup>12</sup></b> |                  |       |                             |       |                             |       |                |       |
| Geen diploma secundair onderwijs    | 9348             | 13,3% | 251                         | 8,4%  | 247                         | 8,2%  | 275            | 9,2%  |
| Beroepssecundair onderwijs          | 14184            | 20,1% | 677                         | 22,5% | 651                         | 21,7% | 655            | 21,9% |
| Technisch secundair onderwijs       | 11465            | 16,3% | 517                         | 17,2% | 508                         | 16,9% | 538            | 18,0% |
| Algemeen secundair onderwijs        | 11097            | 15,7% | 474                         | 15,8% | 413                         | 13,8% | 372            | 12,4% |
| Hoger onderwijs één cyclus          | 15476            | 21,9% | 723                         | 24,0% | 732                         | 24,4% | 741            | 24,8% |
| Universitair onderwijs              | 8953             | 12,7% | 365                         | 12,1% | 451                         | 15,0% | 412            | 13,8% |

Op basis van deze vergelijkingen blijkt inderdaad dat de uitval het grootst is bij de laagste sociale strata, in dit geval de laaggeschoolden zonder diploma. Globaal echter blijken de afwijkingen niet bijster groot. Niet alleen wijken de verdelingen in de steekproeven niet zeer sterk af van die in de totale bevolking, maar ook gelijken de verdelingen van de drie steekproeven opvallend meer op elkaar dan dat ze verschillen. We nemen dan ook aan dat mits weging van de steekproeven de vertekeningen door de vastgestelde over- en ondervertegenwoordigingen opgevangen kunnen worden.

10 8 missings op onderwijsniveau door van ontbreken kalender.

11 4 missings op regio door fout in respondentnummer.

12 afbakening categorieën onderwijsniveau:

- diploma secundair onderwijs: minstens een zesde leerjaar geslaagd zijn;
- bij diploma beroepssecundair onderwijs is ook de vierde graad inbegrepen, evenals vizo-opleidingen, deeltijds en buitengewoon secundair onderwijs (minstens een vijfde leerjaar);
- respondenten met een diploma van een uniformopleiding zijn gerekend bij het hoger onderwijs van één cyclus;
- het universitair onderwijs omvat zowel opleidingen gevolgd aan een universiteit als opleidingen hoger onderwijs van twee cycli;
- een aanvullend diploma hoger onderwijs is bij het universitair diploma gerekend (enkel bij NIS; in SONAR niet meegenomen)

### 3.2. Afleiding van de wegingsfactoren

We beschrijven hier de methode waarmee de wegingsfactoren afgeleid werden. De cijfers worden hier enkel getoond voor de cohorte 1978, aangezien deze cohorte het meest vergelijkbaar is met de bevolkingscijfers. De methode van afleiding voor de andere cohorten is analoog. De wegingscoëfficiënten die het resultaat zijn van de procedure zijn voor elke cohorte terug te vinden in de bijlagen, evenals de globale verdelingen op de drie indicatoren geslacht, regio, en onderwijsniveau (zowel de oorspronkelijke cijfers, als de cijfers na toepassing van de wegingscoëfficiënten voor elk van de onderzochte cohorten).

Het uitgangspunt is de vergelijking van een aantal kenmerken in de data van c78(23) en overeenkomstige gegevens uit de Volkstelling 2001 voor alle personen geboren in 1978. De wegingscoëfficiënten waarmee de verdelingen op de SONAR-data omgezet kunnen worden naar de verdelingen op basis van de bevolkingscijfers, werden in twee stappen bepaald. In een eerste stap wordt een correctie van de aantallen in de combinatie geslacht en regio uitgevoerd, in een tweede stap worden de (door de eerste stap aangepaste) aantallen naar onderwijsniveau, apart voor mannen en vrouwen, verder gecorrigeerd. De uiteindelijke wegingscoëfficiënten zijn dan te bepalen door het product van de factoren nodig voor beide correcties apart. Deze stapsgewijze procedure werd verkozen boven een rechtstreekse bepaling van de gewichten op basis van de vergelijking van de verdelingen op de drie kenmerken tegelijk. Met de gecombineerde verdelingen werden we immers geconfronteerd met relatief kleine aantallen in bepaalde cellen (zie ook verder).

#### Stap 1: correctie naar geslacht x regio (c78(23))

|                          | <i>Bevolking</i> |       | <i>Origineel</i> |       | <i>Correctie factor 1</i> |
|--------------------------|------------------|-------|------------------|-------|---------------------------|
| <b>Mannen</b>            |                  |       |                  |       |                           |
| Grote stad               | 4242             | 5,7%  | 114              | 3,8%  | 1,51                      |
| Regionale stad           | 5298             | 7,2%  | 160              | 5,3%  | 1,35                      |
| Kleine stad              | 7912             | 10,7% | 417              | 13,9% | 0,77                      |
| Niet-stedelijke gemeente | 19873            | 26,9% | 816              | 27,2% | 0,99                      |
| <b>Vrouwen</b>           |                  |       |                  |       |                           |
| Grote stad               | 4475             | 6,1%  | 125              | 4,2%  | 1,45                      |
| Regionale stad           | 5297             | 7,2%  | 169              | 5,6%  | 1,27                      |
| Kleine stad              | 7723             | 10,5% | 408              | 13,6% | 0,77                      |
| Niet-stedelijke gemeente | 18987            | 25,7% | 789              | 26,3% | 0,98                      |

In bovenstaande tabel staan de verdelingen van geslacht in combinatie met regio weergegeven voor de bevolking en de steekproef c78(23). Het gaat hier steeds om de proporties binnen het geheel van alle personen. De correctiefactor is berekend als het quotiënt van de proportie van de betrokken groep in de bevolking gedeeld door de proportie in de steekproef. Als het percentage in de steekproef vermenigvuldigd wordt met die factor, moet het percentage in de bevolking immers gereproduceerd worden. Uit die cijfers blijkt duidelijk dat de sterkste ondervertegenwoordiging zich voordoet in de steden – wat consistent is met de bevindingen op basis van de analyse van de woninggegevens – en iets meer bij de mannen dan de vrouwen.

Met deze correctiefactor is in een tweede stap een raming gemaakt van de verdeling over onderwijsniveau in de steekproef, door elke respondent mee te tellen met een gewicht gelijk aan de

## Wegingscoëfficiënten SONAR cohortes

eerste correctiefactor (in plaats van dat elke respondent voor exact 1 telt) gegeven de kenmerken geslacht en regio voor de betrokken respondent. In onderstaande tabel wordt deze gecorrigeerde verdeling over onderwijs vergeleken met de verdeling in de bevolking, ditmaal apart voor mannen en vrouwen. Op basis van de afwijkingen wordt in deze tweede stap een tweede correctiefactor afgeleid, op dezelfde wijze berekend als uitgelegd voor de eerste stap.

De tweede correctiefactor laat zien welke groepen na correctie voor de vertegenwoordiging naar geslacht en regio, nog over- en ondervertegenwoordigd zijn op het vlak van onderwijsniveau. Bij mannen en vrouwen vinden we dezelfde tendensen terug: de laagst geschoolde groep blijkt het sterkst ondervertegenwoordigd in de steekproef, de hoogst geschoolden zijn duidelijk oververtegenwoordigd.

### **Stap 2: correctie naar onderwijsniveau per geslacht (c78(23))**

|                                  | <i>Bevolking</i> |       | <i>Origineel</i> |       | <i>Correctie factor 2</i> |
|----------------------------------|------------------|-------|------------------|-------|---------------------------|
| <b>Mannen</b>                    |                  |       |                  |       |                           |
| Geen diploma secundair onderwijs | 5736             | 16,1% | 167              | 11,0% | 1,47                      |
| Beroepssecundair onderwijs       | 7264             | 20,4% | 345              | 22,7% | 0,90                      |
| Technisch secundair onderwijs    | 6808             | 19,1% | 303              | 20,0% | 0,96                      |
| Algemeen secundair onderwijs     | 6014             | 16,9% | 234              | 15,4% | 1,10                      |
| Hoger onderwijs één cyclus       | 5442             | 15,3% | 246              | 16,2% | 0,94                      |
| Universitair onderwijs           | 4288             | 12,1% | 223              | 14,7% | 0,82                      |
| <b>Vrouwen</b>                   |                  |       |                  |       |                           |
| Geen diploma secundair onderwijs | 3612             | 10,3% | 95               | 6,4%  | 1,61                      |
| Beroepssecundair onderwijs       | 6920             | 19,8% | 309              | 20,8% | 0,95                      |
| Technisch secundair onderwijs    | 4657             | 13,3% | 200              | 13,5% | 0,99                      |
| Algemeen secundair onderwijs     | 5082             | 14,5% | 189              | 12,7% | 1,14                      |
| Hoger onderwijs één cyclus       | 10034            | 28,7% | 473              | 31,9% | 0,90                      |
| Universitair onderwijs           | 4665             | 13,3% | 218              | 14,7% | 0,91                      |

Om tot slot de finale wegingscoëfficiënten te verkrijgen, werden de correctiefactoren uit beide stappen met elkaar vermenigvuldigd, voor elke combinatie van de kenmerken geslacht, regio en onderwijsniveau. Deze wegingen zijn voor elk van de cohortes opgenomen in de bijlagen, samen met de verdelingen in de steekproeven na toepassing van de wegingen. Zoals daaruit blijkt, worden door de weging de verdelingen van geslacht en onderwijsniveau in de bevolking exact gereproduceerd in de steekproeven. De uiteindelijke verdelingen op regio wijken wel lichtjes af van de verdeling in de bevolking, wat mogelijk is omdat het criterium regio niet gebruikt werd in de laatste stap om de wegingscoëfficiënten af te leiden. De afwijkingen zijn dan een gevolg van een zekere samenhang in het onderwijsniveau en de regio.

De hier beschreven tweestap procedure werd verkozen boven de gemakkelijkere werkwijze om de verdelingen op de drie indicatoren in één stap, dus gecombineerd, uit te voeren. Die werkwijze zou bovendien niet alleen de verdeling over geslacht en onderwijsniveau gereproduceerd hebben, maar ook over regio. Het basisprobleem is dat de verdelingen over onderwijsniveau per geslacht dan ook nog eens opgesplitst moesten worden per regio. Met vier vrij ongelijk gespreide regio's betekende dit dat de aantallen vrij klein zouden worden voor de kleinste regio's. Bovendien zou met deze werkwijze de samenhang tussen regio en onderwijsniveau gereproduceerd worden, ook al zou het

## Representativiteit en weging van de cohorten

misschien om een relatief betekenisloze samenhang gaan. De oefening werd wel gedaan, maar dan met regio ingedeeld in twee groepen, om het probleem van kleine aantallen te beperken. Wanneer dan gekeken werd naar de gewogen verdeling over regio, bleek de geïntegreerde methode zeker niet beter.

In de bijlagen zijn de concrete cijfers opgenomen per cohorte, en in onderstaande tabel worden de statistische kenmerken van de afgeleide wegingscoëfficiënten samengevat. De hoogste wegingscoëfficiënt bedraagt 2,35 in de cohorte 78(23), in beide andere cohorten zijn al de wegingsfactoren zelfs beduidend kleiner dan twee. De laagste coëfficiënten zijn eveneens terug te vinden in c78(23). Hiermee samenhangend is de spreiding van de wegingscoëfficiënten het grootst in de middelste cohorte.

**Wegingscoëfficiënten van de drie cohorten - statistieken**

|                       | <b>Aantal</b> | <b>Minimum</b> | <b>Maximum</b> | <b>st.afw.</b> |
|-----------------------|---------------|----------------|----------------|----------------|
| <b>Weging c76(23)</b> | 3015          | 0,79           | 1,82           | 0,21           |
| <b>Weging c78(23)</b> | 3002          | 0,63           | 2,35           | 0,30           |
| <b>Weging c80(23)</b> | 2993          | 0,75           | 1,66           | 0,20           |

De verschillen tussen de cohorten in de wegingscoëfficiënten zijn niet onverwacht. Op basis van zowel de duur van het veldwerk als de globale bruto- en netto-responscijfers, kon verwacht worden dat de meest representatieve steekproef die zou zijn van c80, gevolgd door c76, en als laatste c78. Die volgorde lijkt overeen te stemmen met wat blijkt uit de statistieken van de afgeleide wegingscoëfficiënten.

Deze statistieken geven alleszins aan dat de gewichten vrij beperkt blijven qua grootte, en dus relatief veilig gebruikt kunnen worden in analyses. In wat volgt evalueren we verder de zo afgeleide wegingscoëfficiënten voor de andere cohorten; we bekijken in welke mate de veralgemening van de methode c78(23), ic. de bevolkingsgegevens, naar de andere cohorten verantwoord lijkt.

### 3.3. Verdere toetsing van de veralgemeenbaarheid over de cohortes

Op basis van de vergelijking van de gegevens van de verschillende geboortecohorten met die van de Volkstelling 2001 bij jongeren geboren in 1978, werden de wegingscoëfficiënten bepaald. In feite zijn de gegevens van de Volkstelling enkel direct vergelijkbaar met de gegevens van c78(23), omdat de vergelijking daar betrekking heeft op dezelfde geboortecohorte, op de leeftijd van 23 jaar. Door diezelfde gegevens als referentie te gebruiken voor de andere geboortecohorten veronderstellen we dat die, toen zij 23 waren, gekenmerkt worden door dezelfde verdelingen, met andere woorden dat er zich geen veranderingen hebben voorgedaan over de tijd in de verdeling van jongeren op een bepaalde leeftijd over geslacht, regio en onderwijsniveau. Voor geslacht en regio lijkt het inderdaad weinig waarschijnlijk dat zich duidelijke evoluties hebben voorgedaan. Voor het bereikte onderwijsniveau ligt die veronderstelling misschien moeilijker, gegeven de onderwijsexpansie – de lange termijn trend van steeds meer jongeren die deelnemen aan het (hoger) onderwijs, met de daaruit volgende toename in het gemiddeld onderwijsniveau.

#### 3.3.1. Onderwijsniveau cohorte 1980 op 21 jaar

Daarom is het wenselijk te toetsen in welke mate de veralgemening van het onderwijsniveau van jongeren geboren in 1978 naar de andere cohorten problematisch zou kunnen zijn. Voor beide andere cohortes zijn geen bevolkingscijfers beschikbaar over het onderwijsniveau op 23 jaar. Voor de cohorte c80(23) is het echter wel mogelijk met de gegevens van de Volkstelling 2001 te vergelijken: de bevolkingsgegevens hebben dan betrekking op het onderwijsniveau op 21-jarige leeftijd, en doordat bij c80(23), net zoals bij de andere cohortes trouwens, de volledige onderwijsloopbaan per schooljaar geregistreerd werd, is het mogelijk het hoogst behaalde onderwijsniveau op de leeftijd van 21 af te leiden. Voor de respondenten uit c76(23) kon een gelijkaardige oefening niet gemaakt worden. De enquête vond immers plaats in 1999, op een tijdstip voor de Volkstelling 2001 dus, waardoor geen vergelijking mogelijk is. Teruggrijpen naar de voorgaande Volkstelling was evenmin een optie, omdat de respondenten dan nog te jong waren (in 1991).

**Onderwijsniveau bij c80(23) op 21 jaar: het effect van de weging getoetst<sup>13</sup>**

| <b>Onderwijsniveau</b>           | <b>Bevolking</b> |              | <b>Origineel</b> |              | <b>Na weging</b> |              |
|----------------------------------|------------------|--------------|------------------|--------------|------------------|--------------|
|                                  | <b>Man</b>       | <b>Vrouw</b> | <b>Man</b>       | <b>Vrouw</b> | <b>Man</b>       | <b>Vrouw</b> |
| Geen diploma secundair onderwijs | 17,4             | 10,6         | 11,8             | 6,9          | 16,2             | 10,7         |
| Beroepssecundair onderwijs       | 20,9             | 20,5         | 23,2             | 21,0         | 20,4             | 20,1         |
| Technisch secundair onderwijs    | 24,1             | 18,7         | 28,0             | 21,6         | 25,4             | 19,1         |
| Algemeen secundair onderwijs     | 31,1             | 35,6         | 31,0             | 36,3         | 32,8             | 37,2         |
| Hoger onderwijs één cyclus       | 5,6              | 13,7         | 6,0              | 14,1         | 5,1              | 12,7         |
| Universitair onderwijs           | 0,8              | 1,0          | 0,1              | 0,1          | 0,1              | 0,1          |

In bovenstaande tabel worden de bevolkingscijfers naast de ongewogen en de gewogen cijfers van c80(23) weergegeven, apart voor mannen en vrouwen. Er werd ook bekeken hoe de verdelingen lagen in de steekproef als enkel de correctie op de verdeling van geslacht en regio werd toegepast (eerste stap in de afleiding van de wegingscoëfficiënten). Deze bleken niet te verschillen van de originele verdeling zonder correctie (tenzij hier en daar een verschuiving van 0,1 in de percentages).

13 De gebruikte werkwijze om het onderwijsniveau op 21 jaar te bepalen van de respondenten, onderschat eventueel de aantallen die op basis van een kwalificatie in het deeltijds onderwijs bij de gediplomeerden BSO geteld kunnen worden (maar het effect kan slechts marginaal zijn). De verdelingen voor de geboortecohorte 1980 (op 21 jaar) in de Volkstelling 2001 werden, evenals de cijfers voor cohorte 1978 op 23, bekomen via P. Deboosere en D. Willaert, Steunpunt Demografie, Vrije Universiteit Brussel.

## Representativiteit en weging van de cohorten

De vergelijking van de verdelingen over onderwijsniveau in bovenstaande tabel laat duidelijk zien dat de gewogen verdeling van de steekproef veel beter aansluit bij de verdeling van het onderwijs in de bevolking). Er zijn ook geen indicaties voor een toename van het onderwijsniveau over twee jaar: de gewogen onderwijsniveaus (op basis van de steekproef c80 gecorrigeerd naar bevolkingsgegevens voor c78) tonen geen grotere groep laaggeschoolden dan de bevolkingsgegevens voor c80, maar eerder het omgekeerde.

Uiteindelijk is het wel zo dat, indien er toch een verschuiving zou zijn in het onderwijsniveau, deze slechts zeer klein kan zijn. En aangezien we veralgemenen van de middelste cohorte naar telkens een cohorte die slechts twee jaar verschilt met de referentiecohorte, zou de vertekening sowieso beperkt blijven.

### 3.3.2. Minimum kwalificaties vergeleken tussen de cohorten

Een mogelijk nadeel van het gebruik van dezelfde referentiegegevens voor de weging van de drie verschillende cohorten, is dat eventuele evoluties in kenmerken die samenhangen met de criteria waarop gewogen wordt, de facto worden verdoezeld of zelfs uitgewist.

#### *Gekwalificeerd of niet? Het effect van de wegingen getoetst<sup>14</sup>*

|                      | C76  |       | C78  |       | C80  |       |
|----------------------|------|-------|------|-------|------|-------|
|                      | Man  | Vrouw | Man  | Vrouw | Man  | Vrouw |
| <b>Ongewogen</b>     |      |       |      |       |      |       |
| Ongekwalificeerd     | 10,0 | 6,9   | 10,1 | 5,9   | 11,0 | 6,3   |
| Laag gekwalificeerd  | 5,6  | 3,5   | 6,2  | 2,7   | 7,0  | 3,8   |
| Gekwalificeerd       | 84,4 | 89,6  | 83,7 | 91,4  | 82,0 | 90,0  |
| <b>Correctie 1</b>   |      |       |      |       |      |       |
| Ongekwalificeerd     | 10,3 | 7,0   | 10,5 | 6,4   | 10,9 | 6,2   |
| Laag gekwalificeerd  | 5,7  | 3,5   | 6,4  | 2,8   | 7,0  | 3,7   |
| Gekwalificeerd       | 84,0 | 89,5  | 83,1 | 90,8  | 82,1 | 90,1  |
| <b>Finale weging</b> |      |       |      |       |      |       |
| Ongekwalificeerd     | 16,3 | 10,5  | 15,3 | 10,2  | 15,2 | 9,7   |
| Laag gekwalificeerd  | 5,0  | 3,1   | 6,2  | 2,8   | 6,5  | 3,9   |
| Gekwalificeerd       | 78,7 | 86,4  | 78,5 | 87,1  | 78,3 | 86,4  |
| <b>Aantallen</b>     |      |       |      |       |      |       |
| Ongewogen            | 1544 | 1471  | 1509 | 1493  | 1528 | 1465  |
| Gewogen              | 1525 | 1490  | 1518 | 1484  | 1514 | 1480  |

Gegeven de “universele” weging op het hoogst behaalde onderwijsniveau op 23-jarige leeftijd, wordt geraakt aan een beleidsmatig zeer relevant kenmerk dat rechtstreeks samenhangt met het bereikte onderwijsniveau: de mate waarin jongeren het onderwijs wel of niet gekwalificeerd verlaten. In deze paragraaf wordt daarom gekeken naar de impact van de wegingen op het aantal ongekwalificeerde schoolverlaters.

<sup>14</sup> Tot de gekwalificeerden hoort iedereen die een diploma of getuigschrift haalde van minstens het zesde jaar in het secundair onderwijs. Heeft men dit niet, maar heeft men wel een getuigschrift van de leerovereenkomst of een studiekwalificatiegetuigschrift van het deeltijds beroepssecundair onderwijs, of rondde men het vijfde (laatste) leerjaar van het buso af, dan hoort men bij de laaggekwalificeerden. Al de anderen beschouwen we hier als ongekwalificeerd.

## Wegingscoëfficiënten SONAR cohortes

In bovenstaande tabel wordt voor elk van de cohorten, apart voor mannen en vrouwen, de verdeling naar basiskwalificaties weergegeven. Zowel de ongewogen als gewogen cijfers staan vermeld, alsook de cijfers na correctie van de eerste stap in de weging (op de verdeling regio en geslacht).

Wat opvalt is de grote stabiliteit in de cijfers over de drie cohorten. En niet louter in de gewogen cijfers: noch bij mannen, noch bij vrouwen blijken er in de ongewogen of de half gecorrigeerde data significante verschillen tussen de cohortes te bestaan in wel of niet gekwalificeerd zijn, en indien niet-gekwalificeerd, of men minstens een lage kwalificatie behaalde. De weinige verschillen na weging verdoezelen dus geen reële verschillen die bestonden voor de weging.

De veralgemening van de cijfers van mensen geboren in 1978 naar de geboortecohorten twee jaar vroeger en later, lijkt op basis van beide bovenstaande toetsen dus geen vertekening te introduceren noch een verschuiving in de kenmerken te verbergen.

### **4. Conclusie**

Op basis van de verdelingen op de drie socio-demografische indicatoren zijn in dit verslag wegingsfactoren afgeleid. Hiermee wordt (hoofdzakelijk) enerzijds een correctie op de regionale verdeling gemaakt, wat wellicht vooral het gevolg is van de samenstelling van de steekproef, en anderzijds een correctie op de selectieve uitval zoals blijkt uit de ondervertegenwoordiging van lager geschoolden. De afwijkingen van de SONAR-steekproeven ten opzichte van de bevolking op die vlakken worden door de toepassing van de wegingscoëfficiënten kleiner. Door deze wegingscoëfficiënten te hanteren bij de analyses kunnen de belangrijkste vertekeningen door onevenwichten in de steekproef opgevangen worden, en sluiten uitspraken over “23-jarigen” of “26-jarigen” beter aan bij de realiteit.



## **5. Addendum : wegingscoëfficiënten op 26 jaar**

### 5.1. Toevoeging: wegingscoëfficiënten cohorte 76(26)

Tot slot wordt hier nog een gelijkaardige weging uitgewerkt voor wie op 26-jarige leeftijd bereikt werd van de cohorte jongeren geboren in 1976<sup>15</sup>. Omdat het hier een tweede golf, een tweede enquête bij wie reeds drie jaar voordien bevraagd werd, betreft, kan verwacht worden dat deze groep respondenten selectiever is dan de groep die op 23-jarige leeftijd nog deelnam aan het onderzoek.

In een eerste tabel wordt een overzicht gegeven van de mate waarin de respondenten op 26 jaar op dezelfde manier verdeeld zijn over de basiskenmerken (gemeten op 23 jaar; ongewogen gegevens), en de mate waarin de kans op deelname op 26 jaar (aantal bevraagden op 26 op de totale groep die op 23 jaar bevraagd werd) al dan niet samenhangt met die kenmerken.

**Globale verdeling naar geslacht, regio en onderwijsniveau op 23 jaar, voor c76(26)**

|                                  | Volledige groep op 23 |       | Bereikte groep op 26 |       | Deelname<br>26 / 23 |
|----------------------------------|-----------------------|-------|----------------------|-------|---------------------|
| <b>Geslacht</b>                  |                       |       |                      |       |                     |
| Mannen                           | 1544                  | 51,2% | 1053                 | 51,1% | 68,2%               |
| Vrouwen                          | 1471                  | 48,8% | 1007                 | 48,9% | 68,5%               |
| <b>Regio</b>                     |                       |       |                      |       |                     |
| Grote stad                       | 332                   | 11,0% | 234                  | 11,4% | 70,5%               |
| Regionale stad                   | 414                   | 13,7% | 275                  | 13,3% | 66,4%               |
| Kleine stad                      | 575                   | 19,1% | 380                  | 18,4% | 66,1%               |
| Niet-stedelijke gemeente         | 1695                  | 56,2% | 1171                 | 56,8% | 69,1%               |
| <b>Onderwijsniveau (op 23)</b>   |                       |       |                      |       |                     |
| Geen diploma secundair onderwijs | 252                   | 8,4%  | 137                  | 6,7%  | 54,4%               |
| Beroepssecundair onderwijs       | 677                   | 22,5% | 422                  | 20,6% | 62,3%               |
| Technisch secundair onderwijs    | 517                   | 17,2% | 357                  | 17,4% | 69,1%               |
| Algemeen secundair onderwijs     | 474                   | 15,8% | 314                  | 15,3% | 66,2%               |
| Hoger onderwijs één cyclus       | 723                   | 24,0% | 543                  | 26,4% | 75,1%               |
| Universitair onderwijs           | 365                   | 12,1% | 280                  | 13,6% | 76,7%               |

Noot: de gemiddelde deelnamekans 26/23 (= bruto respons-rate) bedraagt 68,3%

De kansen op deelname zijn quasi dezelfde voor mannen en vrouwen, lijken enigszins te variëren met de regio (hoewel globaal niet statistisch significant), en nemen duidelijk toe met het onderwijsniveau dat men op 23 jaar bereikt had (sterk statistisch significant). De oorspronkelijke oververtegenwoordiging van de niet-stedelijke gemeenten bij de enquête op 23-jarige leeftijd, blijft aanwezig op 26 jaar door de relatief hoge respons bij diezelfde respondenten. De ondervertegenwoordiging van de lagere onderwijscategorieën wordt met de enquête op 26 nog verder versterkt door de hogere respons bij de hoogst geschoolden.

In een tweede tabel blijkt dat het toepassen op de 26-jarigen van de wegingscoëfficiënten die uitgewerkt werden voor de groep van 23-jarigen niet helemaal doeltreffend is (zie tweede kolom in

15 Voor het opzet en het verloop van de enquêtering van c76(26) kan het technisch gedeelte geconsulteerd worden van Belet et al. (2004a).

## Wegingscoëfficiënten SONAR cohortes

onderstaande tabel). Vooral op het vlak van het onderwijsniveau bekomen we voor de 2060 26-jarigen te kleine groepen met lagere onderwijsniveaus (in vergelijking met de bevolkingsgegevens). Daarom wordt hier bij de groep 26-jarigen de procedure van schatten van wegingscoëfficiënten overgedaan. Daarbij worden dezelfde kenmerken gehanteerd: geslacht, en regio en onderwijsniveau op 23 jaar. Met het longitudinale karakter van de enquête zijn er uiteraard nog heel wat meer mogelijkheden om de selectiviteit van de uitval te analyseren en te corrigeren, maar dat valt buiten de doelstelling hier, waar we ons beperken tot een eenvoudige weging van de groep op 26 jaar. De afleiding van de wegingscoëfficiënten voor de bereikte groep van 26-jarigen hier volgt de gebruikte methode op 23 jaar.

### **Globale verdeling naar geslacht, regio en onderwijsniveau op 23 jaar, voor c76(26)**

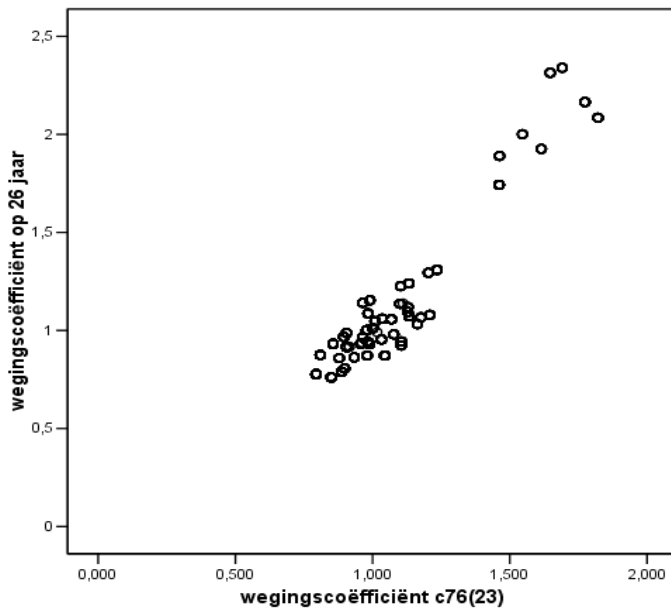
|                                  | <b>23-jarigen<br/>(gewogen)</b> |       | <b>26-jarigen<br/>Na weging 23</b> |       | <b>26-jarigen<br/>Na weging 26</b> |       |
|----------------------------------|---------------------------------|-------|------------------------------------|-------|------------------------------------|-------|
| <b>Geslacht</b>                  |                                 |       |                                    |       |                                    |       |
| Mannen                           | 1525                            | 50,6% | 1031                               | 50,5% | 1042                               | 50,6% |
| Vrouwen                          | 1490                            | 49,4% | 1010                               | 49,5% | 1018                               | 49,4% |
| <b>Regio</b>                     |                                 |       |                                    |       |                                    |       |
| Grote stad                       | 371                             | 12,3% | 258                                | 12,6% | 261                                | 12,7% |
| Regionale stad                   | 436                             | 14,5% | 287                                | 14,1% | 301                                | 14,6% |
| Kleine stad                      | 638                             | 21,2% | 421                                | 20,6% | 440                                | 21,4% |
| Niet-stedelijke gemeente         | 1571                            | 52,1% | 1075                               | 52,7% | 1058                               | 51,4% |
| <b>Onderwijsniveau (op 23)</b>   |                                 |       |                                    |       |                                    |       |
| Geen diploma secundair onderwijs | 400                             | 13,3% | 219                                | 10,8% | 273                                | 13,3% |
| Beroepssecundair onderwijs       | 605                             | 20,1% | 375                                | 18,4% | 413                                | 20,1% |
| Technisch secundair onderwijs    | 489                             | 16,3% | 336                                | 16,5% | 334                                | 16,3% |
| Algemeen secundair onderwijs     | 473                             | 15,7% | 314                                | 15,4% | 323                                | 15,7% |
| Hoger onderwijs één cyclus       | 659                             | 21,9% | 495                                | 24,3% | 450                                | 21,9% |
| Universitair onderwijs           | 382                             | 12,7% | 294                                | 14,5% | 261                                | 12,7% |

Noot:  $N = 2060$  in c76(26); bij weging 23: nog 2041

Die aangepaste wegingscoëfficiënten corrigeren de afwijkingen bij de groep van 26-jarigen zodanig dat ze opnieuw beter aansluiten bij de groep van alle Vlaamse jongeren van 26, toch op het vlak van de gebruikte socio-demografische kenmerken (zie derde kolom in bovenstaande tabel). Voor de exacte wegingscoëfficiënten voor c76(26) verwijzen we naar de bijlagen.

Addendum : wegingscoëfficiënten op 26 jaar

**Samenhang wegingscoëfficiënten c76(26) met c76(23)**



De wegingscoëfficiënten op 23 jaar en 26 jaar hangen uiteraard sterk samen, maar blijken toch niet identiek (pearson correlatie van 0,93).

Het verband wordt hiernaast grafisch weergegeven.

De grootste afwijkingen doen zich juist voor bij de groepen van respondenten die de grootste positieve wegingen krijgen (reeds op 23 jaar, maar nog hogere weging op 26 jaar).

De groepen respondenten die op 23 jaar reeds een groter gewicht dienden te krijgen omdat ze ondervertegenwoordigd waren in de

gerealiseerde steekproef, worden in de beperktere groep van 26-jarigen nog meer opgetrokken opdat de eindverdelingen op de kenmerken beter zouden aansluiten bij de verdelingen in de volledige populatie.

Terwijl de wegingscoëfficiënten voor de hele groep 23-jarigen gemiddeld exact 1 bedraagt, ligt het gemiddelde bij de groep die ook op 26 aan het onderzoek deelnam, op 0,99, dus iets lager (het verschil met het iets hoger gemiddelde van 1,02 bij de groep die niet opnieuw onderzocht werd, is zelfs statistisch significant). Opvallend is eveneens de wat hogere spreiding bij de wegingscoëfficiënten die gebaseerd werden op de groep 26-jarigen, als gevolg van de zowel lagere als hogere scores.

**Wegingen toegepast op 26-jarigen van c76 (N=2060)**

|                  | <i>Minimum</i> | <i>Maximum</i> | <i>Gemiddelde</i> | <i>st.afw.</i> |
|------------------|----------------|----------------|-------------------|----------------|
| <b>Weging 23</b> | 0,79           | 1,82           | 0,99              | 0,20           |
| <b>Weging 26</b> | 0,76           | 2,34           | 1,00              | 0,30           |

## Bibliografie

Coppieters, P. et al. (2002) *Hoe maken Vlaamse jongeren de overgang van school naar werk? Basisrapportering cohorte 1976 (eerste golf)*. Eindrapport PBO97. Leuven: Sonar.

Belet, H. et al. (2003) *Hoe maken Vlaamse jongeren de overgang van school naar werk? Basisrapportering cohorte 1978 (eerste golf)*. Eindrapport PBO99: deel 1. Leuven: Sonar.

Belet, H. et al. (2004a) *Hoe maken Vlaamse jongeren de overgang van school naar werk? Basisrapportering cohorte 1976 (tweede golf)*. Eindrapport PBO99: deel 2. Leuven: Sonar.

Belet, H. et al. (2004b) *Hoe maken Vlaamse jongeren de overgang van school naar werk? Technisch rapport cohorte 1980 (eerste golf)*. Rapportering in het kader van het Steunpunt Loopbanen. Leuven: Sonar.

Billiet, J. (1996) *Methoden van sociaal-wetenschappelijk onderzoek: ontwerp en dataverzameling*. Leuven: Acco.

## Bijlagen

### Bijlagen

|  |    |
|--|----|
| Bijlagen.....  | 26 |
| Vergelijkbaarheid onderzoeksgegevens met bevolkingsgegevens..... | 27 |
| Wegingscoëfficiënten cohorte 76(23).....                         | 29 |
| Wegingscoëfficiënten cohorte 78(23).....                         | 30 |
| Wegingscoëfficiënten cohorte 80(23).....                         | 31 |
| Wegingscoëfficiënten cohorte 76(26).....                         | 32 |

### **Vergelijkbaarheid onderzoeksgegevens met bevolkingsgegevens**

Om tot de wegingscoëfficiënten te komen, worden de SONAR-gegevens vergeleken met bevolkingsgegevens, met name van de Volkstelling 2001. Hier komt de vergelijkbaarheid van beide databronnen aan bod, samen met de details over de indeling van de kenmerken.

De weging is gebaseerd op de vergelijking van drie socio-demografische kenmerken: geslacht, verstedelijkingsgraad en onderwijsniveau. De vergelijking beperkte zich tot deze kenmerken, aangezien deze het meest relevant en vergelijkbaar zijn. Andere kenmerken zoals bij voorbeeld de gezinssamenstelling werden niet gebruikt, omdat ze mogelijk grote vertekening zouden kunnen geven. Zo zal de gezinssituatie in de NIS-gegevens veel meer steunen op de officiële situatie, maar zullen de SONAR-enquêtegegevens vaker verwijzen naar de feitelijke situatie (cf. domicilie-adres dat niet overeenkomt met officiële woonsituatie).

1. Wat betreft geslacht nemen we aan dat zich geen probleem van vergelijkbaarheid stelt.
2. Voor de verstedelijkingsgraad speelt wel een verschil in de referentie-periode waarop de woonplaats vastgelegd werd. De regio van de SONAR-respondenten werd bepaald aan de hand van de NIS-code die overeenstemt met het steekproefpunt dat in de respondentnummers gecodeerd zit. Dit impliceert dat de veronderstelde woonplaats werd vastgelegd op basis van de steekproeftrekking, via het rijksregister, dus voorafgaand aan de enquête. De referentie-periode voor de regio ligt voor de respondenten dus ergens in de zomermaanden, terwijl die voor de Volkstelling ligt op 1 oktober. In de mate dat 23-jarigen veel verhuizen tussen die maanden (met verandering van domicilie) naar gemeenten met een andere verstedelijking, kan de vergelijking minder geldig worden. Beide databronnen zijn wel vergelijkbaar in het gebruik van het officiële adres, dat zou kunnen afwijken van het feitelijke adres.
3. Voor de verdeling van de bevolking over de onderwijscategorieën, werden bepaalde categorieën samengevoegd om een vergelijking met de SONAR-respondenten mogelijk te maken. Drie categorieën met een dubbelzinnige betekenis werden bovendien ook via een aantal schattingen verdeeld over andere categorieën: post-secundair hoger onderwijs (hoofdzakelijk naar bso, maar ook tso), hoger onderwijsdiploma maar geen informatie over welk soort hoger onderwijs, en één cyclus van hoger academisch onderwijs (beide verdeeld over diploma's secundair onderwijs).

| <b>Gebruikte categorie</b>       | <b>Samenstelling van de categorie</b>  |
|----------------------------------|--|
| Geen diploma secundair onderwijs | Samenvoeging van geen formeel onderwijs gevolgd, lager onderwijs en diploma's lager secundair beroeps, technisch en algemeen secundair onderwijs   |
| Beroepssecundair onderwijs       | Samenvoeging van beroepssecundair onderwijs met 80% van post-secundair hoger onderwijs, en 1/3 van onbepaald diploma hoger onderwijs.  |
| Technisch secundair onderwijs    | Samenvoeging van technisch secundair onderwijs met 20% van post-secundair hoger onderwijs, 1/3 van onbepaald diploma hoger onderwijs, en 10% van één cyclus van hoger academisch onderwijs |
| Algemeen secundair onderwijs     | Samenvoeging van algemeen secundair onderwijs met 1/3 van onbepaald diploma hoger onderwijs, en 90% van één cyclus van hoger academisch onderwijs  |
| Hoger onderwijs één cyclus       | Niet-academisch hoger onderwijs  |
| Universitair onderwijs           | Samenvoeging van cyclus twee van hoger academisch onderwijs met doctoraatsdiploma  |

4. Voor het onderwijsniveau speelt ook de referentie-periode een rol. Voor de SONAR-respondenten werd het onderwijsniveau berekend tot en met het laatste afgeronde schooljaar bij de enquête op 23, in de praktijk dus het schooljaar voorafgaand aan de enquête. Voor de vergelijking met de bevolkingsgegevens (situatie 1 oktober) geeft dit geen probleem, zolang geen volledig extra

## Bijlagen

schooljaar kan afgerond zijn. Voor c78(23) werden de observaties daarom expliciet beperkt tot het begin van het schooljaar (owv veldwerk tot in de zomer); voor de andere cohortes was dit niet nodig, aangezien het veldwerk niet zo lang liep.

5. De gegevens van de Volkstelling bevatten ook ontbrekende data. Voor alle jongeren geboren in 1978 zijn de verdelingen naar geslacht en regio volledig gekend. Maar voor het onderwijsniveau ontbreken de gegevens van een groep die met de enquête niet bereikt werd (die de vragenlijst niet terugzond) én een kleine groep die geen antwoord gaf op de vraag naar het diploma. Bij de afleiding van de wegingscoëfficiënten werd de vergelijking telkens volledig gemaakt naar geslacht en regio, en in de tweede stap beperkt tot de gekende verdeling van onderwijsniveau naar geslacht. De onbekende diploma's werden met andere woorden niet samengevoegd met een andere – lage – onderwijscategorie, maar volledig buiten beschouwing gelaten (om de schatting van het aantal laaggeschoolden niet de hoogte in te duwen).



## Wegingscoëfficiënten SONAR cohortes

### **Wegingscoëfficiënten cohorte 76(23)**

#### **Overzicht van de wegingscoëfficiënten c76(23)**

|                                  | <b>Grote stad</b> | <b>Regionale stad</b> | <b>Kleine stad</b> | <b>Niet-stedelijke gemeente</b> |
|----------------------------------|-------------------|-----------------------|--------------------|---------------------------------|
| <b>Mannen</b>                    |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 1,820             | 1,615                 | 1,773              | 1,461                           |
| Beroepssecundair onderwijs       | 1,008             | 0,895                 | 0,982              | 0,810                           |
| Technisch secundair onderwijs    | 0,989             | 0,877                 | 0,964              | 0,794                           |
| Algemeen secundair onderwijs     | 1,128             | 1,000                 | 1,099              | 0,905                           |
| Hoger onderwijs één cyclus       | 1,105             | 0,980                 | 1,076              | 0,887                           |
| Universitair onderwijs           | 1,163             | 1,032                 | 1,133              | 0,934                           |
| <b>Vrouwen</b>                   |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 1,545             | 1,646                 | 1,690              | 1,463                           |
| Beroepssecundair onderwijs       | 0,904             | 0,963                 | 0,989              | 0,856                           |
| Technisch secundair onderwijs    | 1,129             | 1,202                 | 1,234              | 1,068                           |
| Algemeen secundair onderwijs     | 1,034             | 1,102                 | 1,131              | 0,979                           |
| Hoger onderwijs één cyclus       | 0,898             | 0,956                 | 0,982              | 0,849                           |
| Universitair onderwijs           | 1,104             | 1,176                 | 1,207              | 1,045                           |

Noot: coëfficiënten voor onderwijsniveau onbekend: mannen: kleine stad 1,111 ; niet-stedelijke gemeente 0,915 ; vrouwen: grote stad 1,016 ; niet-stedelijke gemeente 0,961

#### **Globale verdeling naar geslacht, regio en onderwijsniveau, voor c76(23)**

|                                  | <b>Bevolking</b> |       | <b>Origineel</b> |       | <b>Na weging</b> |       |
|----------------------------------|------------------|-------|------------------|-------|------------------|-------|
| <b>Geslacht</b>                  |                  |       |                  |       |                  |       |
| Mannen                           | 37325            | 50,6% | 1544             | 51,2% | 1525             | 50,6% |
| Vrouwen                          | 36482            | 49,4% | 1471             | 48,8% | 1490             | 49,4% |
| <b>Regio</b>                     |                  |       |                  |       |                  |       |
| Grote stad                       | 8717             | 11,8% | 332              | 11,0% | 371              | 12,3% |
| Regionale stad                   | 10595            | 14,4% | 414              | 13,7% | 435              | 14,4% |
| Kleine stad                      | 15635            | 21,2% | 575              | 19,1% | 638              | 21,2% |
| Niet-stedelijke gemeente         | 38860            | 52,7% | 1695             | 56,2% | 1571             | 52,1% |
| <b>Onderwijsniveau</b>           |                  |       |                  |       |                  |       |
| Geen diploma secundair onderwijs | 9348             | 13,3% | 251              | 8,4%  | 399              | 13,3% |
| Beroepssecundair onderwijs       | 14184            | 20,1% | 677              | 22,5% | 605              | 20,1% |
| Technisch secundair onderwijs    | 11465            | 16,3% | 517              | 17,2% | 489              | 16,3% |
| Algemeen secundair onderwijs     | 11097            | 15,7% | 474              | 15,8% | 473              | 15,7% |
| Hoger onderwijs één cyclus       | 15476            | 21,9% | 723              | 24,0% | 659              | 21,9% |
| Universitair onderwijs           | 8953             | 12,7% | 365              | 12,1% | 382              | 12,7% |

## Bijlagen

### **Wegingscoëfficiënten cohorte 78(23)**

#### **Overzicht van de wegingscoëfficiënten c78(23)**

|                                  | <b>Grote stad</b> | <b>Regionale stad</b> | <b>Kleine stad</b> | <b>Niet-stedelijke gemeente</b> |
|----------------------------------|-------------------|-----------------------|--------------------|---------------------------------|
| <b>Mannen</b>                    |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 2,217             | 1,973                 | 1,130              | 1,451                           |
| Beroepssecundair onderwijs       | 1,359             | 1,209                 | 0,693              | 0,889                           |
| Technisch secundair onderwijs    | 1,450             | 1,290                 | 0,739              | 0,949                           |
| Algemeen secundair onderwijs     | 1,659             | 1,476                 | 0,846              | 1,086                           |
| Hoger onderwijs één cyclus       | 1,428             | 1,270                 | 0,728              | 0,934                           |
| Universitair onderwijs           | 1,241             | 1,104                 | 0,633              | 0,812                           |
| <b>Vrouwen</b>                   |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 2,346             | 2,054                 | 1,241              | 1,577                           |
| Beroepssecundair onderwijs       | 1,382             | 1,210                 | 0,731              | 0,929                           |
| Technisch secundair onderwijs    | 1,437             | 1,258                 | 0,760              | 0,966                           |
| Algemeen secundair onderwijs     | 1,659             | 1,453                 | 0,877              | 1,115                           |
| Hoger onderwijs één cyclus       | 1,309             | 1,146                 | 0,692              | 0,880                           |
| Universitair onderwijs           | 1,321             | 1,156                 | 0,698              | 0,888                           |

Noot: coëfficiënten voor regio onbekend: mannen: bso 0,904 ; aso 1,104 ; vrouwen: bso 0,945 ; ho1C 0,895

#### **Globale verdeling naar geslacht, regio en onderwijsniveau, voor c78(23)**

|                                  | <b>Bevolking</b> |       | <b>Origineel</b> |       | <b>Na weging</b> |       |
|----------------------------------|------------------|-------|------------------|-------|------------------|-------|
| <b>Geslacht</b>                  |                  |       |                  |       |                  |       |
| Mannen                           | 37325            | 50,6% | 1509             | 50,3% | 1518             | 50,6% |
| Vrouwen                          | 36482            | 49,4% | 1493             | 49,7% | 1484             | 49,4% |
| <b>Regio</b>                     |                  |       |                  |       |                  |       |
| Grote stad                       | 8717             | 11,8% | 239              | 8,0%  | 381              | 12,7% |
| Regionale stad                   | 10595            | 14,4% | 329              | 11,0% | 424              | 14,1% |
| Kleine stad                      | 15635            | 21,2% | 825              | 27,5% | 629              | 21,0% |
| Niet-stedelijke gemeente         | 38860            | 52,7% | 1605             | 53,5% | 1564             | 52,2% |
| <b>Onderwijsniveau</b>           |                  |       |                  |       |                  |       |
| Geen diploma secundair onderwijs | 9348             | 13,3% | 247              | 8,2%  | 398              | 13,3% |
| Beroepssecundair onderwijs       | 14184            | 20,1% | 651              | 21,7% | 604              | 20,1% |
| Technisch secundair onderwijs    | 11465            | 16,3% | 508              | 16,9% | 488              | 16,3% |
| Algemeen secundair onderwijs     | 11097            | 15,7% | 413              | 13,8% | 473              | 15,7% |
| Hoger onderwijs één cyclus       | 15476            | 21,9% | 732              | 24,4% | 659              | 21,9% |
| Universitair onderwijs           | 8953             | 12,7% | 451              | 15,0% | 381              | 12,7% |

## Wegingscoëfficiënten SONAR cohortes

### **Wegingscoëfficiënten cohorte 80(23)**

#### **Overzicht van de wegingscoëfficiënten c80(23)**

|                                  | <b>Grote stad</b> | <b>Regionale stad</b> | <b>Kleine stad</b> | <b>Niet-stedelijke gemeente</b> |
|----------------------------------|-------------------|-----------------------|--------------------|---------------------------------|
| <b>Mannen</b>                    |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 1,298             | 1,315                 | 1,357              | 1,433                           |
| Beroepssecundair onderwijs       | 0,820             | 0,831                 | 0,857              | 0,905                           |
| Technisch secundair onderwijs    | 0,861             | 0,872                 | 0,900              | 0,950                           |
| Algemeen secundair onderwijs     | 1,197             | 1,213                 | 1,252              | 1,322                           |
| Hoger onderwijs één cyclus       | 0,779             | 0,789                 | 0,814              | 0,860                           |
| Universitair onderwijs           | 0,852             | 0,863                 | 0,891              | 0,940                           |
| <b>Vrouwen</b>                   |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 1,655             | 1,350                 | 1,647              | 1,639                           |
| Beroepssecundair onderwijs       | 1,013             | 0,826                 | 1,008              | 1,003                           |
| Technisch secundair onderwijs    | 0,918             | 0,749                 | 0,913              | 0,909                           |
| Algemeen secundair onderwijs     | 1,309             | 1,068                 | 1,303              | 1,297                           |
| Hoger onderwijs één cyclus       | 0,949             | 0,774                 | 0,944              | 0,940                           |
| Universitair onderwijs           | 0,974             | 0,795                 | 0,970              | 0,965                           |

#### **Globale verdeling naar geslacht, regio en onderwijsniveau, voor c80(23)**

|                                  | <b>Bevolking</b> |       | <b>Origineel</b> |       | <b>Na weging</b> |       |
|----------------------------------|------------------|-------|------------------|-------|------------------|-------|
| <b>Geslacht</b>                  |                  |       |                  |       |                  |       |
| Mannen                           | 37325            | 50,6% | 1528             | 51,1% | 1513             | 50,6% |
| Vrouwen                          | 36482            | 49,4% | 1465             | 48,9% | 1479             | 49,4% |
| <b>Regio</b>                     |                  |       |                  |       |                  |       |
| Grote stad                       | 8717             | 11,8% | 358              | 12,0% | 376              | 12,6% |
| Regionale stad                   | 10595            | 14,4% | 479              | 16,0% | 441              | 14,7% |
| Kleine stad                      | 15635            | 21,2% | 630              | 21,0% | 626              | 20,9% |
| Niet-stedelijke gemeente         | 38860            | 52,7% | 1526             | 51,0% | 1549             | 51,8% |
| <b>Onderwijsniveau</b>           |                  |       |                  |       |                  |       |
| Geen diploma secundair onderwijs | 9348             | 13,3% | 275              | 9,2%  | 397              | 13,3% |
| Beroepssecundair onderwijs       | 14184            | 20,1% | 655              | 21,9% | 602              | 20,1% |
| Technisch secundair onderwijs    | 11465            | 16,3% | 538              | 18,0% | 487              | 16,3% |
| Algemeen secundair onderwijs     | 11097            | 15,7% | 372              | 12,4% | 471              | 15,7% |
| Hoger onderwijs één cyclus       | 15476            | 21,9% | 741              | 24,8% | 656              | 21,9% |
| Universitair onderwijs           | 8953             | 12,7% | 412              | 13,8% | 380              | 12,7% |

## Bijlagen

### **Wegingscoëfficiënten cohorte 76(26)**

#### **Overzicht van de wegingscoëfficiënten c76(26)**

|                                  | <i>Grote stad</i> | <i>Regionale stad</i> | <i>Kleine stad</i> | <i>Niet-stedelijke gemeente</i> |
|----------------------------------|-------------------|-----------------------|--------------------|---------------------------------|
| <b>Mannen</b>                    |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 2,085             | 1,926                 | 2,165              | 1,74                            |
| Beroepssecundair onderwijs       | 1,047             | 0,967                 | 1,087              | 0,875                           |
| Technisch secundair onderwijs    | 0,930             | 0,859                 | 0,965              | 0,778                           |
| Algemeen secundair onderwijs     | 1,094             | 1,011                 | 1,136              | 0,915                           |
| Hoger onderwijs één cyclus       | 0,944             | 0,872                 | 0,980              | 0,789                           |
| Universitair onderwijs           | 1,032             | 0,954                 | 1,072              | 0,863                           |
| <b>Vrouwen</b>                   |                   |                       |                    |                                 |
| Geen diploma secundair onderwijs | 2,001             | 2,314                 | 2,340              | 1,890                           |
| Beroepssecundair onderwijs       | 0,987             | 1,141                 | 1,154              | 0,932                           |
| Technisch secundair onderwijs    | 1,119             | 1,294                 | 1,309              | 1,057                           |
| Algemeen secundair onderwijs     | 1,060             | 1,226                 | 1,240              | 1,002                           |
| Hoger onderwijs één cyclus       | 0,806             | 0,932                 | 0,942              | 0,761                           |
| Universitair onderwijs           | 0,923             | 1,067                 | 1,079              | 0,872                           |

Noot: coëfficiënten voor onderwijsniveau onbekend: mannen: kleine stad 1,138 ; niet-stedelijke gemeente 0,917 ; vrouwen: grote stad 0,991 ; niet-stedelijke gemeente 0,936