

Verkenning van de mogelijkheden van het gebruik van administratieve data voor het project onderwijs-arbeidsmarkt

Citation for published version (APA):

Borghans, L., & Kriechel, B. (2008). Verkenning van de mogelijkheden van het gebruik van administratieve data voor het project onderwijs-arbeidsmarkt. (ROA Technical Reports; No. 1). Maastricht: Researchcentrum voor Onderwijs en Arbeidsmarkt, Faculteit der Economische Wetenschappen.

Document status and date:

Published: 01/01/2008

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Verkenning van de mogelijkheden van het gebruik van administratieve data voor het Project Onderwijs-Arbeidsmarkt

Lex Borghans
Ben Kriechel

ROA-TR-2008/1
April 2008

**Researchcentrum voor
Onderwijs en Arbeidsmarkt**

P.O. Box 616

6200 MD Maastricht

The Netherlands

E-mail: <mailto:secretary@roa.unimaas.nl>

Internet: <http://www.roa.unimaas.nl>

Maastricht University

Faculty of Economics and Business Administration

ISBN-978-90-5321-477-0
Sec08.058.pdf

Inhoud

	Bladzijde
Voorwoord	i
1 Introductie	1
2 Beschrijving van de data	2
3 Lonen	7
4 Arbeidsmobiliteit	13
5 Vooruitzichten en conclusies	15

Voorwoord

In dit werkdocument wordt verslag gedaan van een verkennend onderzoek naar de mogelijkheden van het gebruik van administratieve data voor het Project Onderwijs-Arbeidsmarkt. Een aantal onderdelen van het onderzoek zijn gepubliceerd in het rapport *De arbeidsmarkt naar opleiding en beroep tot 2012* (ROA-R-2007/4) dat het ROA in november 2007 heeft uitgebracht, of opgenomen in het bijbehorende Arbeidsmarktinformatiesysteem (AIS).

Het Project Onderwijs-Arbeidsmarkt wordt gefinancierd door het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW), het Ministerie van Landbouw, Natuur en Voedselkwaliteit (LNV), het Centrum voor Werk en Inkomen (CWI), de vereniging kenniscentra beroepsonderwijs bedrijfsleven COLO, Randstad Nederland, de Raad voor Werk en Inkomen (RWI) en het Uitvoeringsinstituut Werknemersverzekeringen (UWV).

1 Introductie

Al sinds lange tijd zijn de Enquête Beroepsbevolking (EBB) van het CBS en de schoolverlatersenquête (SIS) van het ROA de belangrijkste databronnen voor zowel de beschrijving van de actuele arbeidsmarktsituatie als de prognoses, zoals iedere twee jaar naar buiten worden gebracht in de Arbeidsmarkt naar Opleiding en Beroep. Hoewel deze bronnen voorzien in een grote hoeveelheid cruciale gegevens over de arbeidsmarkt naar opleiding en beroep en met name de positie van schoolverlaters hierbij, zijn er ook duidelijke lacunes in de beschikbare data. De meest in het oog springende witte vlekken in de beschikbare data zijn stroomgegevens over de arbeidsmarkt en informatie over lonen voor personen die al langere tijd het onderwijs hebben verlaten.

Stromen op de arbeidsmarkt zijn vooral van belang voor prognoses van de vervangingsvraag. Aan de basis van voorspellingen hoeveel vraag naar arbeid er zal ontstaan vanwege het vertrek van werkenden, liggen de actuele ontwikkelingen in deze stromen op de arbeidsmarkt. De EBB is in de eerste plaats een cross-sectioneel onderzoek, wat impliceert dat alleen het huidige werk van werkenden wordt beschreven, maar dat er geen informatie beschikbaar is over de vraag of een respondent het afgelopen jaar van baan is veranderd en om welke baan het daarbij ging. Voor een deel heeft de EBB sinds enige tijd overigens wel een panelstructuur gekregen. Dit wil zeggen dat een deel van de respondenten meerdere keren wordt geïnterviewd zodat veranderingen in beeld kunnen worden gebracht. Deze panelgegevens worden door het CBS echter nog niet benut.

Daarnaast bevat de EBB geen vragen over het inkomen van de respondenten. In de schoolverlaterenquêtes is wel een inkomensvraag opgenomen, maar er is dus geen informatie beschikbaar over het inkomen van mensen die al langere tijd werkzaam zijn. Hierdoor ontbreken ook gegevens waarmee de ontwikkeling van het salaris gedurende de loopbaan in beeld kunnen worden gebracht. Dergelijke gegevens zijn van belang voor modellen over de ontwikkelingen op de arbeidsmarkt, maar ook voor de beschrijving van de actuele situatie op de arbeidsmarkt zou het nuttig zijn

als ook meer informatie over de perspectieven op langere termijn in een beroep geschetst zouden kunnen worden.

De afgelopen jaren heeft het CBS gewerkt aan het beschikbaar en bruikbaar maken van een groot aantal administratieve bronnen die voor een deel in deze lacunes kunnen voorzien. Deze gegevensbronnen zijn samengebracht in het banenbestand van het Sociaal Statistisch Bestand (SSB-Banen). In deze notitie wordt verslag gedaan van de verkenningen die hebben plaatsgevonden om te onderzoeken op welke wijze de gegevens in dit nieuwe administratieve bestand gebruikt kunnen worden voor het Project Onderwijs-Arbeidsmarkt (POA).

In paragraaf 2 geven we een beschrijving van het SSB-Banen en de wijze waarop we de gegevens in dit bestand hebben gebruikt. In paragraaf 3 gaan we in op loonsgegevens die met deze bron beschikbaar zijn gekomen. In paragraaf 4 komt de arbeidsmobiliteit aan de orde. We sluiten – paragraaf 5 – af met de belangrijkste conclusies over de mogelijkheden die deze nieuwe gegevensbron biedt.

2 Beschrijving van de data

Het CBS heeft in de afgelopen jaren een aantal administratieve bestanden gekoppeld om zo te komen tot een nieuw onderzoeksbestand waarin uiteenlopende sociaal-economische gegevens worden samengebracht. Dit Sociaal Statistisch Bestand biedt in principe informatie over de gehele Nederlandse bevolking en wordt daarom ook wel de virtuele volkstelling genoemd.

De basis van het SSB wordt gevormd door de gemeentelijke basisadministratie (GBA). De GBA bevat demografische informatie van alle inwoners van Nederland. Op basis van of het SOFI-nummer of informatie over leeftijd, geslacht en huisadres kunnen andere gegevensbronnen aan dit basisbestand worden gekoppeld. Zo zijn in het banenbestand van het SSB gegevens beschikbaar met informatie over de loonbelasting afkomstig van de belastingdienst (Fibase), de “Verzekeringsadministratie werknemers” (VZA), en de Enquête Werkgelegenheid en Lonen

(EWL) van het CBS. In principe bevat dit bestand daarmee voor alle Nederlanders die hebben gewerkt gegevens over het bedrijf waar men werkt en het loon dat men heeft verdiend.

De data bevatten gegevens op het niveau van banen, dat wil zeggen dat per baan een observatie met de begin- en einddatum bestaat. Per kalenderjaar wordt voor elke baan het fiscaal loon weergegeven. Het is dus een *event-history* opbouw die in aparte jaren onderverdeeld wordt. Wij hebben voor de verkenning gebruikt gemaakt van de bestanden van de jaren 1999-2003.

Grote voordelen van deze bestanden zijn dat ze vrijwel de gehele Nederlandse beroepsbevolking bevatten en dat de gegevens koppelbaar zijn over de tijd. Hierdoor heeft het bestand dus een panel-karakter en kan de ontwikkeling van mensen over de tijd worden gevolgd.

Op dit moment zijn er nog geen gegevens beschikbaar over zelfstandigen en freelancers, maar het CBS werkt aan aanvullende bestanden over deze groepen, zodat over enige tijd ook deze informatie kan worden toegevoegd.

Voor de onderzoeksmogelijkheden van het SSB is het ook van belang dat een koppeling gemaakt kan worden met enquêtes van het CBS. In de kern van het SSB zit relatief weinig informatie over zeer veel mensen. Door het SSB te koppelen aan enquêtebestanden kan voor een deelpopulatie extra informatie worden verkregen. Doordat het SSB vrijwel de gehele bevolking omvat is het verlies bij een dergelijke koppeling gering, terwijl normaal gesproken de overlap tussen twee enquêtes vrijwel leeg zal zijn.

Praktische Problemen

De gegevens in het SSB zijn in eerst instantie verzameld vanwege bepaalde administratieve doelstellingen en niet vanuit het perspectief van een bepaalde wetenschappelijke probleemstelling. Het werken met administratieve gegevens

brengt daardoor een aantal praktische problemen met zich mee. De inhoud van de beschikbare gegevens wijkt vaak af van wat vanuit onderzoeksperspectief gewenst zou zijn; afhankelijk van administratieve procedures kunnen gegevens van bepaalde groepen systematisch ontbreken; terwijl de kwaliteit van de data afhangt van de functie die de betreffende gegevens hadden in het administratieve proces. De door het CBS voorbereide administratieve gegevens zijn hier geen uitzondering, maar hebben wel het voordeel dat zij al enkele bewerkingsstappen achter de rug hebben. Zo heeft het CBS door de combinatie van gemeentelijke basisadministratie en de Fibase of VZA ervoor gezorgd dat slechts consistente en relevante informatie vanuit deze administratieve bestanden ter beschikking gesteld worden. In vergelijking met administratieve bestanden die in andere landen beschikbaar zijn, lijken de Nederlandse gegevens van goede kwaliteit te zijn.

Ook is door de combinatie met de grootschalige Enquête Werkgelegenheid en Lonen (EWL) voor ongeveer de helft van de cases belangrijke informatie over de omvang van een aanstelling, maar ook informatie over het bedrijf of de organisatie beschikbaar. Omdat in de EWL grote bedrijven sterk oververtegenwoordigd zijn, is deze informatie zonder verdere weging niet representatief. Om die reden hebben we bij de analyses waar we hier verslag van doen vooralsnog geen gebruik gemaakt van deze informatie.

Een technische complicatie is de grote omvang van de data. Databronnen overschrijden, zelfs in de vorm van aparte jaarbestanden, met gemak het werkgeheugen van moderne computers, zodat een geheugenefficiënte selectie en voorbereiding van de data cruciaal is. De omvang van een jaarbestand moet teruggebracht worden op variabelen die strikt noodzakelijk zijn, en de combinatie van gegevens wordt slechts gedaan voor zover dit voor de analyse noodzakelijk is. Gezien de omvang van de bestanden kunnen bepaalde analyses ook met een steekproef uit het SSB worden uitgevoerd. Bij het maken van koppelingen tussen bestanden kan deze selectie echter niet vooraf gemaakt worden. Interessant is ook dat vrijwel iedere Nederlandse baan in de data gerepresenteerd is. Van alle

werknemers is daardoor ook informatie beschikbaar over al zijn collega's op het werk. Steekproeven zouden het benutten van deze mogelijkheden van de data ook bemoeilijken. Dit is echter een tijdelijke belemmering gezien de omvang van het werkgeheugen van computers gestaag toeneemt, maar ook de efficiënte voorbereiding van de data zeker een leerproces is.

Het CBS hanteert restricties voor het gebruik van administratieve data door onderzoekers ter bescherming van de privacy van de individuen waarover informatie in de bestanden is opgenomen. Dit is binnen het CBS geïstitutioniseerd in de vorm van gebruikerscontracten voor microdata. Analyse van de data vindt on-site of per remote-access plaats. Het eerste houdt in dat slechts in Heerlen of Voorburg toegang tot de data is, terwijl remote-access aan een apart ingerichte computer – zoals deze ook bij het ROA beschikbaar is – toegang tot de server van het CBS verkregen wordt. Dit heeft eerdere problemen met rekentijd en capaciteiten van de computer opgelost.

Selectie SSB

Bij het onderzoek naar de loonstructuur en mobiliteit hebben wij ons geconcentreerd op de belangrijkste baan per werknemer. Per werknemer en jaarbestand werd de combinatie van werknemer en werkgever gekozen met het hoogste jaarinkomen. Zo wordt voorkomen dat kleinere contracten voor dubbeltelling van werknemers en banen zorgen. Uitzondering op deze regel is – voor praktische redenen – de telling van de in- en uitstromen. Hier worden alle banen meegenomen. Binnen deze selectie worden slechts contracten meegenomen die een looptijd van minimaal 90 dagen in het jaar hebben en die minimaal een omvang van het minimumloon hadden. Hierdoor valt een groot aantal zeer kleine banen buiten de analyses.

Binnen de SSB zijn salaris, begin- en einddatum contract, gewerkte uren, leeftijd en geslacht, en bedrijf voor ons van belang. Het salaris in de data is, zoals eerder

gezegd, gebaseerd op het fiscale jaarinkomen. Met behulp van de duur van een contract en de gewerkte uren wordt dit omgerekend naar een maandloon.

Begin- en einddatum zijn van belang om zowel voor het actueel geldige contract de totale contractduur te bepalen, maar ook om het verloop van personeel te bepalen. Bij tellingen van personen (omvang van bedrijven) hebben we de donderdag in de derde week van september als ijkpunt genomen. Door een datum midden in het jaar te nemen worden administratieve problemen van contracten die op de eerste van de maand beginnen en vakantie-invloeden vermeden.

Gewerkte uren zijn slechts gedeeltelijk beschikbaar omdat deze vanuit de EWL toegevoegd worden. Terwijl de EWL een grote omvang heeft, beslaat zij slechts de helft van de beschikbare data. Gezien de EWL een groter gewicht op grotere bedrijven legt, zou het concentreren op de data die zowel omvang van het contract in werkuren als ook het salaris omvat de studie verstoren. Wij hebben voor deze ontwikkelstudie daarom ervoor gekozen dat de bovengenoemde selectie minimaal 90 werkdagen en een omvang van minimaal minimumloon voorlopig minder verstorend is, ook al worden hier deeltijd en voltijd contracten gezamenlijk bekeken, zonder onderscheid te kunnen maken.

Leeftijd en geslacht zijn voor de vervangingsvraag van groot belang. Doordat de uitstroom uit beroepen zowel verandert met leeftijd als ook verschillend voor mannen en vrouwen moet hiermee rekening gehouden worden. De informatie is afkomstig vanuit de gemeentelijke basisadministratie (GBA).

Het bedrijf of de organisatie waar een werknemer werkt wordt via een uniek identificatienummer weergegeven. Het is voor het onderzoek van belang dat gegevens naar het bedrijfsniveau geaggregeerd kunnen worden. Dit biedt bijvoorbeeld de mogelijkheid om de uitstroom per bedrijf te bepalen. Ook kan het gemiddelde lonen per bedrijf worden berekend, zodat bedrijven in percentilen van de loonverdeling kunnen worden ingedeeld.

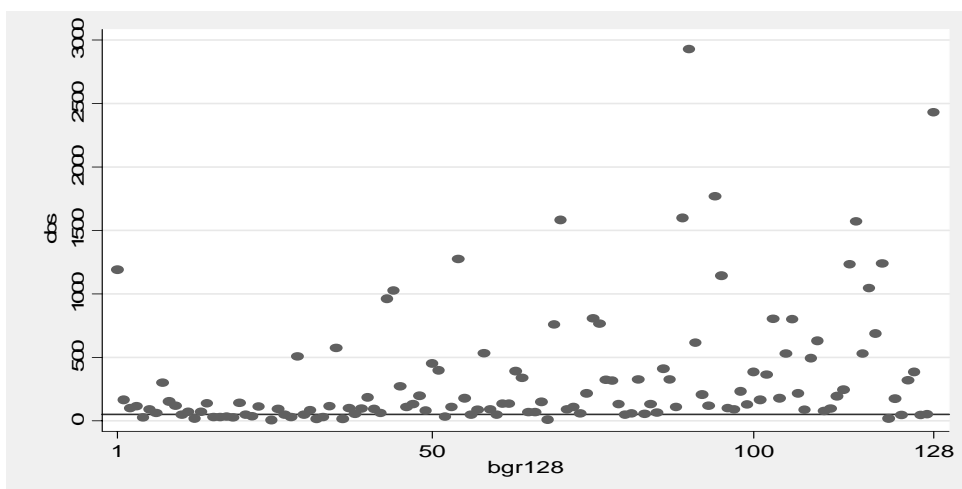
3 Lonen

Lonen vormen een belangrijke rol in het allocatieproces op de arbeidsmarkt. Tot nog toe waren wij door het ontbreken van loondata in de EBB niet in staat lonen in de prognoses mee te nemen. Door de beschikbaarheid van administratieve loondata en de koppelmogelijkheid met (enkele jaren van) de EBB is het echter mogelijk om ook naar de loon- en beloningsstructuur van de arbeidsmarkt te kijken.

Zo kan er een link gelegd worden tussen loonstructuur en de stromen van werkgevers. Maar ook kunnen loonniveaus en vooral verschillen tussen loonniveaus over beroepen of opleiding gebruikt worden om de methodiek van de substitutievraag te verfijnen.

Figuur 1

Aantal observaties in beroepsgroepen, 2002



Toelichting: Per beroepsgroep wordt het aantal observaties weergegeven. De lijn boven de y-as geeft de grens van 50 observaties weer.

Bron: Sociaal Statistisch Bestand (SSB) en EBB, 2002, eigen berekening

De beschikbaarheid van loongegevens betekent ook dat een beeld gegeven kan worden van de lonen per beroepsgroep en opleidingstype en de groei hiervan gedurende de loopbaan. Veranderingen in betaalde lonen zijn één van de indicaties van schaarste (of overschot) op de arbeidsmarkt. Door de verhouding tussen loonveranderingen over tijd en de schaarste op de arbeidsmarkt in het verleden te

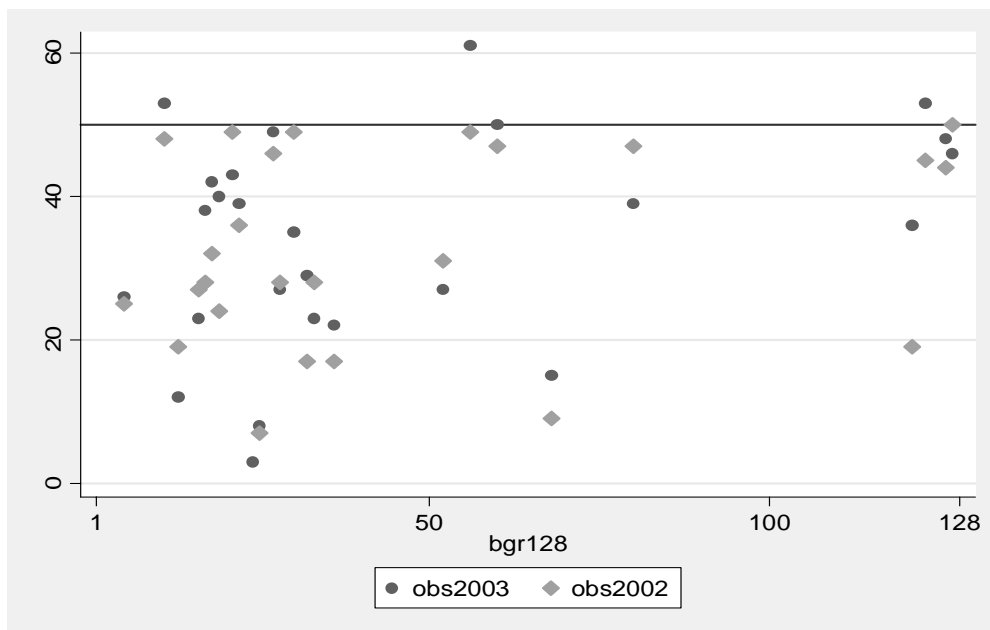
analyseren is het mogelijk om directe indicatoren van schaarste op basis van de loongegevens te ontwikkelen.

Celgrootte

Om een betrouwbare schatting van het gemiddelde en de spreiding van de loonsverdeling in een beroep te kunnen maken is er een bepaalde minimale celvulling nodig. Het CBS hanteert hierbij als vuistregel doorgaans dat er minstens 50 observaties nodig zijn. In Figuur 1 en 2 bekijken we daarom in hoeverre de beroepsgroepen in de EBB aan deze ondergrenzen voldoen. Figuur 1 geeft het totaalbeeld, waarbij de streep de 50-observaties-grens weergeeft. Figuur 2 zoomt in op de kleine beroepsgroepen. In beide jaren hebben wij rond 20 procent van de beroepsgroepen die minder dan 50 observaties aan looninformatie heeft. In 2002 hebben 102 beroepsgroepen meer dan 50 observaties, ofwel 79.6 procent, terwijl in 2003 104 beroepsgroepen meer dan 50 observaties hebben, ofwel 81.25 procent.

Figuur 2

Aantal observaties in kleine beroepsgroepen, 2002 en 2003



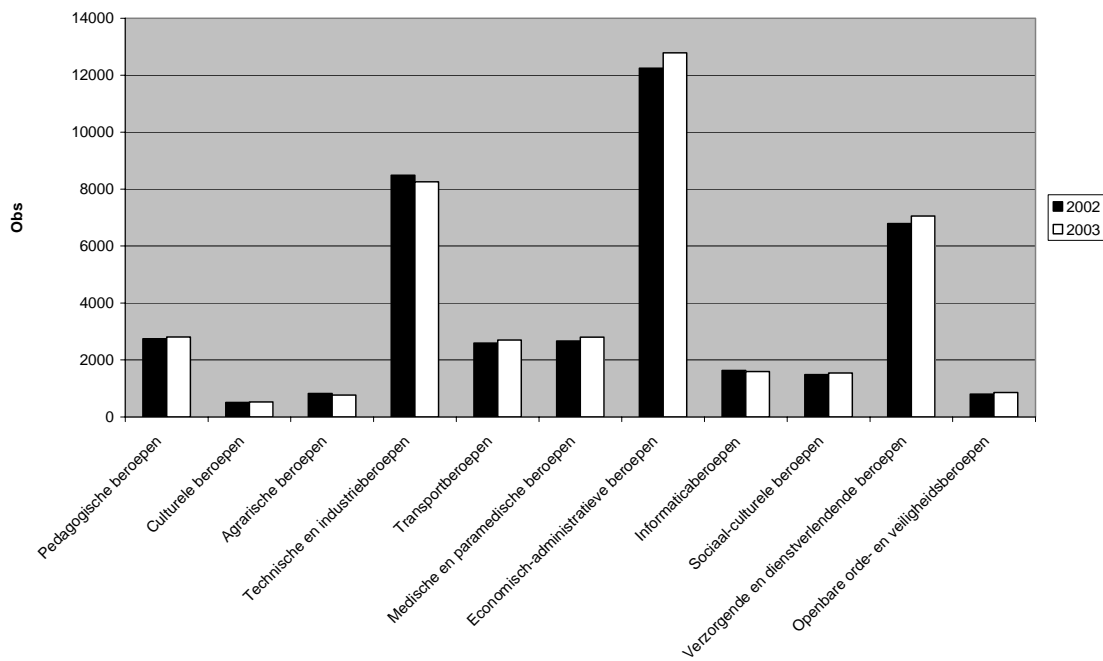
Toelichting: Per beroepsgroep wordt het aantal observaties weergegeven. Deze figuur geeft slechts de beroepsgroepen weer die of in 2002 of in 2003 minder dan 50 observaties hadden.

Bron: Sociaal Statistisch Bestand (SSB) en EBB, 2002 en 2003, eigen berekening

Op het niveau van beroepsklassen lijkt de celvulling geen probleem te vormen. In Figuur 3 wordt de omvang van de beroepsklassen weergegeven. De kleinste beroepsklasse in de BGR 12 indeling is die van Culturele beroepen. Deze omvatten 505 respectievelijk 529 observaties in de jaren 2002 en 2003.

Figuur 3

Aantal observaties per beroepsklasse (BGR12), 2002 en 2003



Toelichting: Per beroepsklasse (BGR 12) wordt het aantal observaties weergegeven.

Bron: Sociaal Statistisch Bestand (SSB) en EBB, 2002 en 2003, eigen berekening

Beroepen, lonen en loonontwikkeling

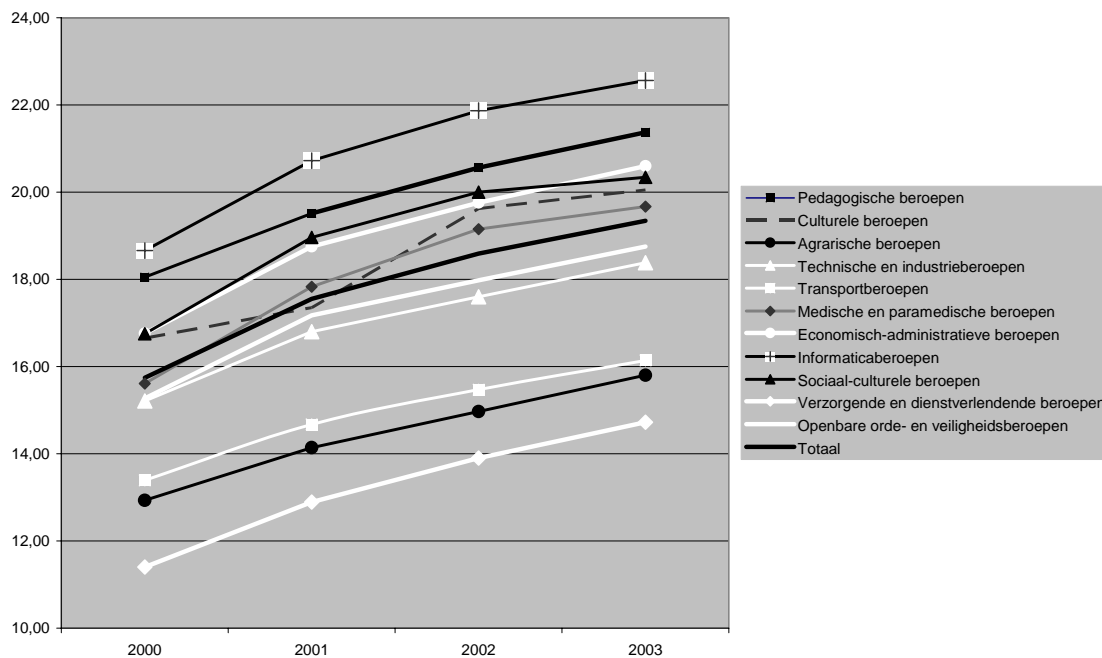
Lonen in beroepen zijn een weerspiegeling van vraag en aanbod, maar in belangrijke mate ook van de productiviteit van een werknemer in een beroep. Voor een werknemer is de absolute hoogte van het salaris van belang, maar ook het groeipotentieel vanuit een beroep.

Vanuit het Sociaal Statistisch Bestand samengesteld door het CBS op basis van sociale verzekering- en loonbelastinggegevens is een koppeling gemaakt met de informatie uit de Enquête Beroepsbevolking (EBB). Hierdoor kan per beroep een uurloon berekend worden. Ook is het mogelijk de salarisontwikkeling van een werknemer over tijd te volgen door meerdere achtereenvolgende jaren van de loongegevens aan de EBB te koppelen.

Voor de geaggregeerde beroepsgroepen geeft figuur 4 een overzicht over de gemiddelde uurlonen van werknemers in loondienst met een leeftijd tussen de 20 en 55 jaar. Duidelijk is de lagere loon te zien in de *verzorgende en dienstverlenende beroepen*, *agrarische beroepen*, en bij de *transportberoepen*. De hoogste uurlonen zijn te vinden bij de *informaticaberoepen*.

Figuur 4

Uurlonen per beroepsklasse, 2002-2003



Note: De figuur geeft de nominale uurlonen op basis van de jaarloon en de gewerkte jaaruren.

Bron: SSB en EBB (2000-2003). Eigen berekening.

Deze loonsverschillen over de beroepsgroepen kan verschillende oorzaken hebben. Ten dele ligt het aan de verschillen in onderliggende opleidingsniveaus. Terwijl bij

10

verzorgende en dienstverlenende beroepen weinig hoogopgeleiden werken, is dit bij de *informaticaberoepen* juist wel het geval. Ten tweede is de hoogte van het uurloon ook afhankelijk van de leeftijd van een werknemer, hoe ouder, hoe hoger de loon gemiddeld is.

Naast de absolute hoogte van het loon is ook de ontwikkeling van het salaris van belang. Voor de verschillende beroepen hebben wij de loonontwikkeling van een werknemer over een periode van vier jaren onderzocht. Hierbij is het niet van belang of de werknemer binnen de (oorspronkelijke) beroepsgroep blijft, of dat hij deze wisselt. Wij controleren hier enigszins voor doordat wij een onderscheid maken tussen werknemers die van werkgever zijn verandert.

Tabel 1 geeft het gemiddelde uurloon per beroepsgroep, voor baanwisselaars en werknemers die in de periode 2000-2003 niet van werkgever zijn verandert. Loongroei is opgesplitst naar deze twee categorieën. Het aandeel baanwisselaars is in de laatste kolom weergegeven.

Tabel 1

Baanwisselaars en loongroei, 2000-2003

Beroepsklasse	Uurloon	2000		% loongroei		% Verloop
		Baan-wisselaar	Blijver	Baan-wisselaar	Blijver	
Pedagogische beroepen	18.05	16.58	17.648	4.63	5.16	0.227
Culturele beroepen	16.65	15.84	16.85	2.73	5.08	0.270
Agrarische beroepen	12.93	12.66	13.19	3.62	4.59	0.219
Technische en industrieberoepen	15.21	14.57	15.6	4.23	4.07	0.211
Transportberoepen	13.39	12.59	13.98	4.87	4.08	0.222
Medische en paramedische beroepen	15.61	15.12	15.87	5.62	6.44	0.253
Economisch-administratieve beroepen	16.76	16.24	16.74	5.58	5.18	0.321
Informaticaberoepen	18.66	18.38	18.63	5.30	4.83	0.324
Sociaal-culturele beroepen	16.75	15.80	16.96	5.16	5.85	0.288
Verzorgende en dienstverlenende beroepen	11.40	11.33	11.77	5.44	6.43	0.303
Openbare orde- en veiligheidsberoepen	15.27	13.96	15.73	3.92	5.26	0.108
Overig	17.92	16.71	17.99	6.97	5.08	0.243

Note: De tabel geeft de gemiddelde loon, de jaarlijks loongroei tussen 2000 en 2003 en de proportie van baanwisselaars in een beroepsgroep voor werknemers. Werknemers zijn ingedeeld in beroepen op basis van de EBB 2000.

Bron: SSB (2000-2003) en EBB 2000. Eigen berekening.

Opvallend zijn de verschillen tussen de loongroei van baanwisselaars ten opzichte van blijvers. Het veranderen van werkgever valt in sommige beroepsgroepen positief uit voor de loonontwikkeling, terwijl het voor andere beroepen juist averechts werkt. De loongroei bij de *Pedagogische*, de *Culturele*, de *Agrarische*, de *Medische* en *Paramedische*, maar ook bij de *Verzorgende en dienstverlenende* en de *Openbare orde- en veiligheidsberoepen* is hoger voor werknemers die niet van bas veranderen. De interne carrière blijkt meer loongroei op te leveren dan een wisselen van werkgever over wellicht beroep. Deze beroepen vertonen ook een lager verloop van werknemers zoals in de laatste kolom van tabel 1 is weergegeven.

Tabel 2

Loongroei van 45 werknemers jonger 45 jaar (45-) en werknemers van 45 jaar en ouder (45+), 2000-2003

Beroepsklasse	Baanwisselaar	Blijver	Verloop
Pedagogische beroepen	2.06	3.23	9.51
Culturele beroepen	-0.04	0.24	7.14
Agrarische beroepen	-1.47	2.13	3.72
Technische en industrieberoepen	1.10	1.11	4.09
Transportberoepen	1.91	1.00	1.12
Medische en paramedische beroepen	3.21	1.49	9.81
Economisch-administratieve beroepen	1.63	1.36	12.12
Informatieberoepen	2.12	2.01	10.31
Sociaal-culturele beroepen	0.95	1.97	12.66
Verzorgende en dienstverlenende beroepen	-0.52	1.34	9.17
Openbare orde- en veiligheidsberoepen	4.21	2.28	-0.78
Overig	6.92	2.09	5.53

Note: De tabel geeft het verschil van het groei- en verlooppercentage van 45- t.o.v. 45+ weer.

Bron: SSB (2000-2003) en EBB 2000. Eigen berekening.

Beroepen waar een wisselen van werkgever wel hogere gemiddelde loonstijging met zich meebrengen, vertonen wel een hoger verloop. Duidelijk te zien bij bijvoorbeeld de *Economisch administratieve beroepen* en de *Informatieberoepen*.

Jongere werknemers hebben over het algemeen een duidelijk hogere verwachte loongroei dan oudere werknemers. Ook het verloop van werknemers is hoger onder jonger personeel dan onder oudere werknemers. Dit is duidelijk te zien in de laatste

twee kolomen van tabel 2. Werknemers die bij een bedrijf blijven in jonge jaren, de 45-ers, hebben een duidelijk hogere loongroei dan oudere werknemers. Toch is het niet bij alle beroepsgroepen zo dat de loongroei van jongere werknemers hoger is dan de loongroei van oudere werknemers; voor de groep van baanwisselaars is er bij de *Agrarische beroepen* en de *Verzorgende en dienstverlenende beroepen* een lagere loongroei voor jongere baanwisselaars dan voor oudere. Dit zou erop kunnen wijzen dat hier veel werknemers gedwongen wisselen wat meestal een lager salaris met zich meebrengt, terwijl oudere werknemers of beter beschermd zijn of geen nieuw werk meer zoeken.

4 Arbeidsmobiliteit

Doordat gegevens uit het SSB over de tijd gekoppeld kunnen worden komt er ook informatie beschikbaar met een panelkarakter. Op grond hiervan kan het personeelsverloop bij bedrijven worden geanalyseerd. In de huidige methodiek werd het verloop van personeel geschat op basis van een pseudo-panel waarin jaargangen van de EBB werden vergeleken. Deze cohort-componentenmethode werd ontwikkeld om vanuit de beschikbare jaarlijkse standcijfers in de beroepen en opleidingen een beeld te vormen over de netto stromen. De enquête beroepsbevolking is een steekproef van de beroepsbevolking waarvan in de huidige methodologie een pseudo-panel gecreëerd wordt. In principe zijn de feitelijke stroomgegevens voor een veel grotere populatie zoals die in het SSB zitten veel geschikter om de stromen op de arbeidsmarkt in beeld te brengen.

Tabel 3 geeft een beeld van de mediaan en het gemiddelde van de relatieve uitstroom per bedrijf per jaar, met de standaarddeviatie van dit uitstroomcijfer. De uitstroom schommelt tussen de jaren, waarbij de gemiddelde uitstroom in 2003 opmerkelijk lager is. Deze uitstroomcijfers zijn ook te vergelijken tussen grote en kleinere bedrijven. Grote bedrijven zijn in deze context bedrijven met meer dan 100 werknemers. De uitstroom bij grotere is lager dan bij kleine bedrijven.

Tabel 3

Uitstroom op verschillende delen van de gemiddelde loonverdeling per bedrijf, 1999-2003

	1999	2000	2001	2002	2003
Exit rate: Median	0.1905	0.2000	0.1961	0.1765	0.1343
Exit rate: Mean	0.2262	0.2366	0.2318	0.2160	0.1700
Exit rate: Standard Deviation	0.1675	0.1692	0.1744	0.1771	0.1539
Exit rate - groot: Median	0.1579	0.1697	0.1635	0.1455	0.1071
Exit rate - groot: Mean	0.1882	0.2008	0.1964	0.1817	0.1374
Exit rate - groot : Standard Deviation	0.1193	0.1225	0.1257	0.1328	0.1088
Exit big / Exit all firms	0.8320	0.8487	0.8473	0.8412	0.8082

Bron: CBS: SSB 1999-2003, eigen berekeningen

Toelichting: De uitstroom wordt berekend voor elk bedrijf. Alle verlopende contracten van een jaar worden gedeeld door het aantal werknemers op de derde donderdag in september van hetzelfde jaar.

Dergelijke uitsplitsingen kunnen worden gemaakt in allerlei dimensies met betrekking tot informatie die beschikbaar is in het SSB. Probleem is voornamelijk dat de hoeveelheid achtergrondkenmerken in het SSB beperkt is. Bovenstaande verschillen tussen bedrijven met hoge lonen en lage lonen zouden bijvoorbeeld het gevolg kunnen zijn van de loonsstructuur als zodanig, maar kunnen er ook op wijzen dat bedrijven met meer hoogopgeleide of oudere werknemers een lager verloop kennen. Op basis van het SSB zou een correctie voor leeftijd mogelijk zijn, maar voor achtergrondgegevens over opleiding en beroep is een koppeling met de EBB noodzakelijk. Een gevolg hiervan is dat slechts informatie beschikbaar is voor een deel van de populatie, waardoor een vergelijking met de collega's in het eigen bedrijf moeilijker wordt en ook vergelijkingen in de tijd moeilijker worden. In combinatie met gegevens uit de EBB weten we alleen het beroep van herkomst of het nieuwe beroep, maar nooit de combinatie van beide gegevens. Daardoor is niet vast te stellen of een persoon van beroep is veranderd. Het is alleen bekend dat de betreffende persoon van werkgever is veranderd. Dit kan betekenen dat hij hetzelfde beroep bij een nieuwe werkgever heeft, of dat hij bij een nieuwe werkgever daadwerkelijk ander werk heeft. Omgekeerd kunnen medewerkers binnen een bedrijf van beroep veranderen. Zonder panelgegevens van de EBB kunnen dergelijke veranderingen niet rechtstreeks worden waargenomen en moeten

dus modellen ontwikkeld en geschat worden om deze patronen uit de data af te leiden op basis van bepaalde veronderstellingen.

5 Vooruitzichten en conclusies

De analyses in deze notitie laten zien dat administratieve bronnen die bij het CBS beschikbaar zijn gekomen, zeer interessante aanvullingen kunnen bieden voor de gegevensbestanden die thans worden gebruikt voor het Project Onderwijs-Arbeidsmarkt. Vanwege het administratieve karakter van deze gegevens zijn er echter ook nog duidelijke beperkingen aan te wijzen.

In de eerste plaats bevatten de beschikbare administratieve bestanden minder en andere variabelen dan databestanden die gebaseerd zijn op enquêtes. Hierdoor zijn bijvoorbeeld alleen jaarlonen bekend en zijn er geen gegevens over het aantal gewerkte uren per week en de uurlonen. De EWL biedt op dit gebied een interessante aanvulling op het SSB. Om deze bron daadwerkelijk te kunnen benutten zal bij de analyses rekening gehouden moeten worden met de verschillen in dekkingsgraad van dit bestand naar bedrijfsomvang.

In het algemeen kan veel aanvullende achtergrondinformatie worden verkregen door koppeling van het SSB aan de EBB. Echter bij de analyses van stroomgegevens uit het SSB zijn deze mogelijkheden beperkt omdat iedere jaargang van het SSB aan een andere EBB-jaargang en dus andere steekproef wordt gekoppeld. Om zicht te krijgen op stromen tussen beroepen zal daarom een model moeten worden ontwikkeld dat de gegevens uit verschillende jaren op een consistente wijze bij elkaar brengt.