**Nederlandse samenvatting PhD thesis Essays on Teacher Quality and Coaching, Marc van der Steeg, Augustus 2016**

**Motivatie en onderzoeksvragen**

Een groeiende literatuur laat consistent zien dat de kwaliteit van leraren een bepalende factor is voor de prestaties van leerlingen. Leerlingen die worden toegewezen aan een leraar met een standaarddeviatie hogere kwaliteit winnen 0.1 tot 0.3 standaarddeviatie meer in termen van prestaties op cognitieve toetsen (zie bv. Rockoff, 2004; Rivkin et al., 2005; Aaronson et al., 2007; Kane & Staiger, 2008; Hanushek & Rivkin, 2010). In een recent paper van Chetty et al. (2014) wordt bovendien gevonden dat de invloed van leraren zich ook uitstrekt tot uitkomsten later in het leven. Een jaar in een klas met een leraar met een 1 standaarddeviatie hoger dan gemiddelde kwaliteit houdt verband met een toename in de kans op deelname aan college van 0.8 procent, met 0.4 procent in de kans om werk te hebben en met 350 dollar per jaar hoger inkomen, beiden op 28-jarige leeftijd.

Het is alleen minder duidelijk wat nu precies de kwaliteit van leraren bepaalt en hoe de kwaliteit kan worden bevorderd. Observeerbare kenmerken als opleidingsniveau en ervaring – opvallend genoeg vaak de belangrijkste determinanten van de beloning van leraren - blijken slechts een beperkt gedeelte van de variatie in lerarenkwaliteit te verklaren, zie overzichtsstudies van Hanushek en Rivkin (2006) en Harris en Sass (2011). Anders gezegd, het grootste gedeelte van de verschillen in de kwaliteit van leraren bevindt zich binnen leraren met hetzelfde opleidings- en ervaringsniveau. Deze lage verklarende kracht van observeerbare factoren heeft er in de afgelopen jaren toe geleid dat onderwijsonderzoekers hun focus zijn begonnen te verleggen naar de vraag welke docentpraktijken en –vaardigheden ertoe doen en in hoeverre verschillende soorten meetinstrumenten van docentkwaliteit voorspellend zijn voor leerwinst van leerlingen. Een recent groot onderzoeksproject dat zich op deze onderzoeksvragen werpt is het Measures of Effective Teaching project in de VS. Hoofdstuk 2 van dit proefschrift sluit aan op deze smalle maar groeiende literatuur en onderzoekt in hoeverre een gedetailleerd observatie-instrument gericht op het meten van pedagogische, didactische en organisatorische vaardigheden van leraren voorspellend is voor leerwinst van leerlingen.

Gegeven het belang van leraren voor onderwijskwaliteit is het niet vreemd dat beleidsmakers in de hele wereld allerhande beleidsmaatregelen inzetten om de kwaliteit en kwantiteit van leraren te verhogen. Investeringen in in scholing van leraren en in een hogere beloning van leraren behoren tot de beleidsmaatregelen die frequent worden ingezet. De evidentie onder deze beleidsmaatregelen levert een gemengd beeld op en overtuigende effectstudies zijn over het algemeen nog vrij schaars. Dit komt vooral door een gebrek aan experimentele of quasi-experimentele variatie in deze beleidsinterventies die benut kan worden om effecten betrouwbaar te schatten. Hoofdstuk 3 en 4 van dit proefschrift dragen bij aan het vergaren van meer kennis over de effectiviteit van beleidsmatige investeringen in het stimuleren van scholingsdeelname onder leraren respectievelijk een hogere beloning van leraren.

Naast de stevige evidentie dat leraren ertoe doen voor het succes van leerlingen is er een beperktere maar groeiende literatuur die erop wijst dat coaching of mentoring van jongeren gedurende het onderwijs positief kan uitpakken voor de schoolloopbanen van jongeren. Zie bv. Lavecchia (2014) voor een overzicht van enkele van deze studies. Hoofdstuk 5 van dit proefschrift voegt toe aan deze literatuur door een experimentele evaluatie van een intensieve coaching aanpak voor studenten aan de onderkant van het middelbaar beroepsonderwijs.

De vier hoofdonderzoeksvragen die in vier achtereenvolgende hoofdstukken van dit proefschrift aan de orde komen in dit proefschrift zijn de volgende:

1. In welke mate is de evaluatiescore op een gedetailleerd observatie-instrument van leraren voorspellend voor de leerwinst die hun leerlingen boeken? (hoofdstuk 2)
2. Wat is het effect van scholingsvouchers voor leraren op hun deelname aan en afronding van hoger onderwijsopleidingen? (hoofdstuk 3)
3. Wat is het effect van een hogere beloning van leraren op hun kans op behoud voor het lerarenberoep? (hoofdstuk 4)
4. Wat is het effect van een intensieve coaching aanpak voor studenten aan de onderkant van het middelbaar beroepsonderwijs op de kans op voortijdig schoolverlaten? (hoofdstuk 5)

**Methoden**

Een belangrijke uitdaging in dit proefschrift is het identificeren van de oorzakelijke effecten van de (beleids-)interventies op de relevante uitkomstvariabelen. Dat is vaak lastig omdat er geregeld sprake is van selectie-effecten. Dat wil zeggen dat er vaak sprake is van niet-geobserveerde verschillen tussen degenen die wel en degenen die niet aan een bepaalde interventie deelnemen. Dat kunnen factoren zijn als motivatie en aanleg die tegelijkertijd van invloed zijn op de uitkomstvariabelen. Als er sprake is van dergelijke niet-geobserveerde verschillen zullen simpele vergelijkingen in uitkomsten tussen degenen die wel en niet de interventie hebben ondergaan een vertekend beeld geven van de daadwerkelijke effecten. De afgelopen decennia zijn in de empirische economische literatuur steeds vaker nieuwe empirische methoden toegepast om het endogeniteitsprobleem bij effectstudies te adresseren Angrist en Pischke (2010) spreken in dit verband van de ‘credibility revolution in empirical economics’. Het gaat hier om methoden die gebruik maken van experimentele of quasi-experimentele variatie in de toekenning van de interventies aan individuen. Door deze variatie ontstaan geloofwaardige controlegroepen van individuen die de interventie niet ondergaan hebben voor de individuen die de interventie wel ondergaan hebben en kunnen effecten betrouwbaarder worden geschat.

In dit proefschrift wordt een variëteit van deze empirische onderzoeksmethoden toegepast om de relevante oorzakelijke verbanden vast te stellen. Het betreft onder meer het benutten van een zogenoemde ‘fuzzy regressie-discontinuïteit’ (hoofdstuk 3), instrumentele variabele en verschil-in-verschillen schattingen (hoofdstuk 4), en benutting van experimentele variatie in een gerandomiseerd experiment (hoofdstuk 5). Ik verwijs graag naar het boek ‘Mostly harmless econometrics’ van Angrist en Pischke (2009) voor een uitgebreide beschrijving van deze methoden en de complicaties die zich hierbij kunnen voor doen en hoe daarmee om te gaan. In hoofdstuk 2 gaat het niet om een effectstudie maar om het bepalen van de voorspellende kracht van een meetinstrument voor prestaties van leerlingen op cognitieve toetsen. Om deze voorspellende kracht zo goed mogelijk vast te stellen wordt gebruik gemaakt van een regressie waarin het verband wordt geschat tussen de score op het meetinstrument en prestaties van leerlingen op taal- en rekentoetsen. Hierbij wordt gecontroleerd voor een uitgebreide set van relevante achtergrondkenmerken van leerlingen (waaronder de voorgaande scores van leerlingen en de sociaal-economische achtergrond) en klassen.

**Resultaten**

Hoofdstuk 2 onderzoekt de relatie tussen evaluaties van leraren, uitgevoerd door getrainde beoordelaars tijdens lesbezoeken, en leerlingprestaties in het basisonderwijs in een grote stad in Nederland. De evaluaties zijn gebaseerd op een gedetailleerd scoresysteem waarin leraren door getrainde observanten tijdens lessen gescoord worden op het wel of niet laten zien van 75 gedragsaspecten. Deze gedragsaspecten geven uiting aan organisatorische, didactische en pedagogische competenties die gerelateerd zijn aan opbrengstgericht lesgeven. De analyses laten zien dat de evaluatiescores van leraren significant voorspellend zijn voor de vooruitgang in leerlingprestaties van de leerlingen in hun klas. De gemiddelde voorspelde vooruitgang in toets-scores bij rekenen en spelling is 0.4 standaarddeviatie als een leerling is toegewezen aan een leraar uit het bovenste kwartiel van de gemeten vaardigheidsverdeling in plaats van een leraar uit het onderste kwartiel. Dit betekent dat een basisschoolleerling die twee jaar op rij een zwakke leraar krijgt toegewezen in plaats van een goede leraar, hierdoor een heel niveau lager terecht kan komen in het vervolgonderwijs, dus van bijvoorbeeld in potentie vwo-niveau naar havo-niveau. De resultaten suggereren dat met het evaluatie-instrument gedragsaspecten van leraren gemeten worden die ertoe doen voor leerlingprestaties. Vooral zwakkere leerkrachten kunnen met het observatie-instrument worden geïdentificeerd. De resultaten suggereren dat het evaluatie-instrument potentie heeft om gebruikt te worden voor ontwikkelings- en personeelsbeleid.

Hoofdstuk 3 onderzoekt het effect van toewijzing van een lerarenbeurs op deelname en afronding van hoger-onderwijsopleidingen (bachelors of masters) door leraren. De effecten worden geschat door het benutten van een zogenoemde regressiediscontinuïteit in de kans om een beurs toegewezen te krijgen. De discontinuïteit is ontstaan doordat het aantal aanvragen in de eerste ronde veel hoger was dan het budget toeliet en de lerarenbeurzen vervolgens zijn toegewezen op volgorde van binnenkomst van de aanvragen. Het geschatte positieve effect van toewijzing van een lerarenbeurs op zowel de kans op deelname als op afronding ligt gemiddeld in de orde van 10 tot 20 procentpunt. Dit komt overeen met een relatief effect van 12-29 procent op deelname en 17-42 procent op afronding. De effecten op deelname en afronding zijn hoger dan gemiddeld voor aanvragers die nog niet gestart waren met de opleiding op het moment van aanvraag, maar lager dan gemiddeld voor aanvragers voor kortere opleidingen tot en met een jaar. Deze effecten zijn gemeten vijfeneenhalf jaar na aanvraag van de lerarenbeurs. De lerarenbeurs vormt grotendeels een substituut voor zowel financiering uit reguliere scholingsbudgetten van scholen als voor financiering uit eigen bijdragen van leraren. Er zijn aanwijzingen voor beperkte positieve effecten van de lerarenbeurs op de kans om in het onderwijs te blijven. Deze effecten lijken geconcentreerd bij leraren in het voortgezet onderwijs en bij leraren die nog niet gestart waren met de opleiding op het moment van de aanvraag.

Hoofdstuk 4 onderzoekt de vraag in hoeverre een hogere beloning voor leraren in het voortgezet onderwijs hun beslissing om leraar te blijven beïnvloedt. Om deze vraag te kunnen beantwoorden, wordt gebruik gemaakt van een beleidsmaatregel die heeft geleid tot regionale verschillen in de beloning van leraren in het voortgezet onderwijs in Nederland. Vanaf 2009 kregen scholen in de Randstad, in vergelijking met scholen buiten de Randstad, extra geld om meer leraren in een hogere salarisschaal te plaatsen. Het betrof een totale additionele bekostiging van 290 miljoen euro over de periode 2009-2014. Deze zogenoemde versterking van de Functiemix in de Randstad heeft ertoe geleid dat in 2014 bijna 20 procentpunt meer leraren in een hogere salarisschaal zijn geplaatst dan op scholen buiten de Randstad. Eenmaal in de hogere salarisschaal, kregen de leraren uitzicht op een 17 procent hogere beloning, ofwel 7200 euro bruto op jaarbasis. Doel van deze beleidsmaatregel was om de beloningsachterstand ten opzichte van banen buiten het onderwijs te verkleinen en (toekomstige) lerarentekorten in de Randstad te bestrijden. De belangrijkste bevinding van dit onderzoek is dat we geen effecten vinden van deze hogere beloning op de kans om leraar te blijven. We hebben de uittreedkans van leraren in de Randstad vergeleken met die van leraren buiten de Randstad. Elk jaar treedt ongeveer 7%  van de leraren uit het lerarenberoep, zowel binnen als buiten de Randstad. Dit percentage is na de invoering van de hogere beloning voor de leraren in de Randstad niet veranderd ten opzichte van leraren buiten de Randstad. De hogere beloning heeft er wel voor gezorgd dat een iets groter deel van de leraren in de Randstad blijft werken en niet kiest om elders een baan als leraar te aanvaarden. De jaarlijkse kans voor een leraar om te switchen van een school in de Randstad naar een school buiten de Randstad is gedaald met 0,4 procentpunt. Dit komt overeen met ongeveer 125 leraren per jaar die niet van regio veranderd zijn op een totaal van circa 30 duizend werkzame leraren in de Randstad. Onze bevindingen suggereren daarnaast dat het beleid heeft geleid tot meer deelname aan formele scholing. Het jaarlijkse aandeel leraren dat een aanvraag doet voor een lerarenbeurs, om zo deel te nemen aan een bachelor of masteropleiding, is door de regionale versterking van de Functiemix gestegen van 2.3 naar 3.2 procent. Deze bevinding is consistent met de opzet van het beleid, waarbij een van de overeengekomen criteria voor plaatsing in een hogere salarisschaal is dat leraren meer kwalificaties of extra expertise verwerven.

Hoofdstuk 5 onderzoekt het effect van intensieve coaching van studenten aan de onderkant van het middelbaar beroepsonderwijs op voortijdig schoolverlaten op basis gegevens van een gerandomiseerd experiment. Het coaching-programma bood studenten intensieve ondersteuning en advies bij hun studieactiviteiten, persoonlijke problemen en bij hun stages. Het coaching programma duurde 1 of 2 jaar. Studenten werden aselect toegewezen aan klassen en klassen werden aselect toegewezen aan de coachingaanpak. De belangrijkste bevinding is dat één jaar coaching voortijdige schoolverlaten met meer dan 40 procent reduceert: van 17 naar 10 procentpunt. Het tweede jaar coaching zorgde voor een verdere daling van het voortijdige schoolverlaten met 1 procentpunt. Het programma is het meest effectief voor studenten met een hoge ex-ante kans op uitval. Het betreft studenten die niet meer onder de kwalificatieplicht vallen, mannelijke studenten en studenten die niet bij beide ouders wonen. Een kosten-batenanalyse suggereert dat bij één jaar coaching de maatschappelijke baten groter zijn dan de kosten.