CPB Netherlands Bureau for Economic Policy Analysis

CPB Discussion Paper | 316

The effects of higher teacher pay on teacher retention

Evidence from regional variation in teacher salaries

Marc van der Steeg
Sander Gerritsen
Sonny Kuijpers

The effects of higher teacher pay on teacher retention
Evidence from regional variation in teacher salaries

Marc van der Steeg*, Sander Gerritsen**, and Sonny Kuijpers***

December 2015

Abstract

This paper investigates the effects of higher teacher pay for secondary school teachers on their teacher retention decision and enrollment in additional schooling. We exploit variation in teacher pay induced by the introduction of a new remuneration policy. This policy provided schools in an urbanized region with extra funds to place a larger share of teachers in a higher salary scale. We exploit this policy in an IV-setup to estimate the effects of higher teacher pay on our outcomes. The main finding is that we find no effects of higher teacher pay on the probability to stay in the teaching profession. The policy however succeeded in keeping a slightly larger share of teachers in the targeted region. In addition, our findings suggest that the policy increased teachers' enrollment in bachelor or master degree programs from 2.3% to 3.2%. This finding is consistent with the setup of the policy in which one of the criteria for placement in a higher salary scale is that teachers would obtain extra qualifications or gain extra expertise.

JEL Codes: I21, I22, I28

Keywords: differences-in-differences, instrumental variables, impact evaluation, teacher pay, teacher retention, teacher schooling.

We would like to thank DUO and Centerdata for kindly providing the data to us. During the project we received valuable comments from Dinand Webbink (Erasmus School of Economics), Nienke Ruijs (University of Amsterdam), seminar participants at Maastricht University and seminar participants at CPB Netherlands Bureau for Economic Policy Analysis.

* CPB Netherlands Bureau for Economic Policy Analysis, m.w.van.der.steeg@cpb.nl

** CPB Netherlands Bureau for Economic Policy Analysis, s.b.gerritsen@cpb.nl

*** CPB Netherlands Bureau for Economic Policy Analysis, s.kuijpers@cpb.nl

1. Introduction

Many countries face teacher shortages, especially in regions where there are better outside options for teachers, higher costs of living and higher shares of low-SES pupils. (e.g. Clotfelter et al., 2008; Greaves & Sibieta, 2014). Policymakers respond to these shortages with various policies to attract more people into the teaching profession or to retain a higher share of teachers. Among these policies, a higher teacher pay is one of the most widely used.[1] It is not immediately clear whether higher teacher pay increases teacher retention rates. The choice to stay in the teaching profession is often motivated by factors other than salary. Studies by Hanushek et al. (2004) and Clotfelter et al. (2011) suggest that the effects of teacher pay on teacher retention are very modest compared to the effect of pupil characteristics. Teachers prefer not to work in schools with high shares of disadvantaged children. Moreover, while there is a large literature that suggests that higher teacher pay increases teacher retention (e.g. Murnane et al., 1989; Dolton & Van der Klaauw, 1995; Hanushek et al., 1999; Imazeki, 2005; Reed et al., 2006; Gilpin, 2011), very few of these studies exploit (plausibly) exogenous variation in teacher pay. Most existing studies of effects of teacher pay on teacher retention exploit across-district variation in teacher salaries.[2] Estimates derived from these studies will be biased for instance if districts with unobserved positive (negative) attributes of teachers offer higher (lower) salaries.

The studies of Hendricks (2014) and Clotfelter et al (2008) are notable exceptions. Hendricks (2014) uses detailed panel data from the state of Texas to estimate the effects of higher teacher pay on teacher retention in a differences-in-differences type of setting. Controlling for changes in district and local labour market characteristics, he finds that a 1% increase in teacher pay reduces the turnover rate by 1.4% and that this effect is largest for inexperienced teachers. Clotfelter et al (2008) exploit both within- and between-school variation in teacher pay caused by the introduction of subject-specific retention bonuses for teachers in public secondary schools with either high-poverty rates or low test scores. Controlling for time-varying school, district and labour market characteristics, they find that 1800 USD retention bonuses led to a relative reduction of turnover rates at targeted schools by 17%.

[1] Other policies are writing-off student loans in exchange for a commitment to teach, subsidies for housing and the expansion of alternative certifications (Hanushek et al., 2004).

[2] This also holds for studies looking at effects on entry decisions into teaching (e.g. Manski, 1987; Dolton, 1990; Wolter & Denzler, 2003; Chevalier et al., 2007) or on pupil test scores (e.g. Dolton & Marcenaro-Gutierrez, 2011).

We contribute to this small literature by exploiting regional variation in teacher pay induced by the introduction of a Dutch teacher pay policy in 2009. The policy provided schools in an urbanized region with relatively large shares of disadvantaged pupils with additional funds to place a larger share of their teachers in a higher salary scale. Nearly 20% of all teachers in the targeted region were given the perspective of a 17% salary increase through placement in a higher salary scale. We use this policy as an instrument for higher teacher pay in an IV-strategy to estimate its effects on retention as a teacher.

Our paper differs from that of Hendricks (2014) and Clotfelter et al. (2008) in an important way. Whereas these studies look at the effects of salary increases on teacher turnover at targeted schools (Clotfelter et al.) or regions (Hendricks), we focus on effects on retention in the teaching profession. That is, we investigate whether salary increases affects teachers' decisions to stay in the teaching profession.[3] Our data allows us to track all Dutch teachers in the Netherlands from 1995-2014, such that we can exactly determine when a teacher leaves the educational labor market. In addition, we look at the effects of higher teacher pay on teachers' decisions to complete more schooling as this was one of the criteria for being placed in a higher salary scale. This outcome has not been studied by Clotfelter et al. (2008) and Hendricks (2014).

Our main findings are as follows. First, we find no effects of being placed in a higher salary scale on teacher retention. That is, we do not find that a higher salary leads to a higher probability to stay in the teaching profession. Second, we find that teachers switch somewhat less from treatment to control regions because of the new remuneration policy, but that this does not affect our results found for retention in the teaching profession. Hence, the policy succeeded in keeping a slightly larger share of teachers in the targeted region. These positive effects are however small relative to the costs of the policy.[4] Third, we find that the policy has a positive impact on teachers' enrollment in degree programs. Our estimates suggest an increase in this probability from 2.3% to 3.2%. This finding is consistent with the setup of the policy in which one of the criteria for placement in a higher salary scale was that teachers would obtain extra qualifications or expertise.

Taken our findings and those of Clotfelter et al. and Hendricks together, we conclude that a higher teacher pay might not be effective in increasing retention rates in the teaching

[3] We use a dummy that equals 1 if the teacher is in the teaching profession and 0 if she is out. Clotfelter et al. and Hendricks use a dummy that equals 1 if the teacher works at a school or region and 0 if she exits this school or region. In that case the teacher can still teach but at another school or region not covered in the dataset.
[4] The policy cost on average 50 million euro per year and around 0.4% of teachers decided not to switch from treatment region to control region because of the policy, see also section 7.

profession, but might be an appropriate policy tool to decrease turnover rates in specific schools or regions, especially in schools or regions with high shares of disadvantaged pupils. In addition, it can be used as an financial incentive to increase participation in follow-up teacher training.

This paper proceeds as follows. Section 2 gives details of the regional teacher pay policy and context. Section 3 discusses our data. Section 4 describes the empirical strategy. Main results are presented in section 5. Section 6 deals with heterogeneous treatment effects. Section 7 concludes and provides a discussion of the results.

2. Institutional background and the regional teacher pay policy

The regional teacher pay policy was introduced in 2009. Secondary schools in the Dutch Randstad region received additional funds from the government to place a substantially larger share of their teachers in a higher salary scale. Figure 1 shows the Randstad region in the Netherlands in dark grey. The Randstad region covers around 40% of all schools and pupils in the Netherlands, indicating a relatively densely populated area. It is a relatively urbanized area that contains the four biggest cities of the Netherlands.

Figure 1: Randstad region within the Netherlands

The motivation for the Randstad policy was to reduce the relatively large wage differential between teaching and other jobs in the Randstad region in order to mitigate (future) teacher shortages. In addition, the policy had to compensate for more difficult working conditions in Randstad schools due to more disadvantaged pupil populations. Heyma et al. (2006) show that the regional wage differential to jobs outside teaching for female teachers is -30% in Amsterdam (situated in the Randstad), whereas in the rural province of Friesland (not situated

in the Randstad) this is -5%. In addition, the share of lessons not given by a certified teacher for the subject is about a quarter higher in the Randstad region, i.e. 28% versus 22%.[5] The share of pupils that lives in so-called 'poverty problem accumulation areas'[6] is more than twice as large in the Randstad schools (i.e. 24% versus 10%, see Table 1 in the next section).

There are three salary scales for secondary school teachers in the Netherlands: LB (low), LC (middle), and LD (high) with maximum (i.e. end-of-scale) gross monthly salaries of 3784, 4413 and 5022 euro, respectively. Starting salaries are roughly equal at around 2500 euro. Most of the teachers are in the low salary scale. Being placed in the mid salary scale instead of the low salary scale gives the perspective of a 17% higher salary end-of-scale, which is equal to 7200 euro in gross terms per year.

The goal of the policy was to place 39% of the 'LB'- teachers in the Randstad in the LC scale by 2014. Schools outside the Randstad also received some additional funding, but much less than the schools inside the Randstad. They could only place 10% of the 'LB'-teachers in the mid salary scale LC. This implies a difference of 29 percentage points in the growth of the share to be placed in the mid salary scale at the cost of the low salary scale between the Randstad and the non-Randstad region.

In total, 290 million euro was made available to the Randstad schools over the period 2009-2014 to achieve this goal, that is, a little less than 50 million euro per year on average.[7] By 2014, however, the realized difference between Randstad and non-Randstad schools was 18 percentage points and not the originally targeted 29 percentage points (see also figure 2 in section 4). The Ministry of Education stated in a letter to parliament that the budget turned out to be insufficient to cover the structural extra wage costs of the Randstad policy (Ministry of Education, 2015). In addition, due to concerns about how the additional funding was spent, it was confirmed that the extra budget was used for placing more teachers in higher salary scales and not for other purposes. We have investigated this in Section 5.3, and have found

[5] Based on year pre-treatment year 2008. Being certified for a certain lesson depends on two things. First, the teacher should have a teacher education degree in the subject of that lesson. Second, the teacher should have a master degree in the subject of that lesson when the lesson is taught in the upper years of secondary education.
[6] This is a zip code area that meets the following three criteria: 1) the share of people with low incomes exceeds 15%, 2) the share of people being welfare recipient exceeds 13% and 3) the share of non-western immigrants exceeds 7%.
[7] The goal was publicly monitored by a website. National goals have been translated into goals per school. As an additional incentive for schools to spend the additional funding on placement of teachers in higher salary scales the extra funding for the second half of the policy period only became available upon reaching intermediate targets for 2011.

no evidence that the funding has been spent in ways other than in placing teachers in higher salary scales.

The additional funds were given to the schools in addition to the regular lump-sum funding that schools receive from the government. Teacher salaries are paid out of this lump-sum funding and take up the largest share of expenditures. The lump-sum also covers salaries of non-teaching personnel such as management and supporting staff, material costs and maintenance costs of school buildings. School leaders decide in which salary scale teachers are placed. This is often based on teacher qualifications (i.e. whether the teacher has a master degree or not) and specific expertise. In the Randstad policy, one of the additional criteria for being placed in a higher teacher salary scale was that teachers would complete extra schooling. This could be i) extra training or expertise in a pedagogical-didactical area, ii) an additional qualification that allows a teacher to teach in two subjects or more or iii) a master degree in the particular subject being taught. Since 2008, teachers are stimulated to obtain this additional schooling by applying for publicly financed schooling vouchers that allow them to follow additional education. These vouchers consist of a financial contribution to teachers to cover tuition costs of a bachelor or master degree program and a contribution to their schools to finance a substitute teacher while the teacher is on study leave.[8]

3. Data

3.1 Data sources and variables

The data for our analysis come from various sources. Information on teacher retention and teacher salary come from two files: the *Mirror* and *Functiemix* datafiles. *Mirror* has been produced and provided to us by an executive agency of the Dutch Ministry of Education (i.e. *DUO, Dienst Uitvoering Onderwijs*). This dataset contains information on the working status of all Dutch teachers in the period 1995-2013. It indicates for instance whether a teacher works in a particular year and, if so, at which school she works. *Functiemix* has been provided by *CenterData* and is similar to *Mirror*, except that it contains more information on teachers and spans a shorter time period, 2006-2014. For our analysis we use both datasets. For checks on our identification strategy we use *Mirror* as this covers a longer time period. For estimation of the effects we use *Functiemix* as it contains more detailed information and hence more covariates. Throughout the paper we refer to *Mirror* as our long sample and

[8] We refer to Van der Steeg and Van Elk (2015) for more details on this teacher schooling voucher scheme.

Functiemix as our estimation sample. All information in these files is measured in October of a particular year.

Information on teacher's schooling decisions comes from another data source of DUO, the *Teacher Schooling Voucher* file. This file gives information at the school level on the share of teachers that applied for a teacher schooling voucher in a particular year. Data are available for the years 2008-2013. This means we have no data on the pre-treatment period. This is because the teacher schooling voucher was not introduced until 2008. From DUO we also obtained information on additional school characteristics such as the share of pupils with a disadvantaged background and school size. This file is called the *school-pupil characteristics file*. From Statistics Netherlands we obtained some additional information on local labor market conditions, that is, unemployment rates in Randstad and non-Randstad regions.

To obtain one main estimation sample, we merged the latter two files with the *Functiemix* file at the school level. We also carried out a few steps to get rid of noise in our data. We refer to Appendix A.1 for a detailed explanation of this procedure. Our estimation sample finally consists of 480,600 observations for which we have full information on teacher retention. The observational unit is a teacher by year.

Main outcomes

In our main analysis we use two outcomes: a dummy for teacher retention and the share of teachers that applied for a schooling voucher. The first is given at the individual teacher level, the latter at the school level (see above).

Teacher retention equals 1 at time *t* if a teacher is observed working at *t* and *t-1*. It equals 0 at time *t* if a teacher is observed working at *t-1* but not at *t*. Teacher retention is reported missing at time *t* if the teacher is not observed working at *t-1*. In that case we cannot identify the retention status of the teacher. Note that by defining this variable as such, we lose the first year of our data.

The share of teachers that applied for a schooling voucher is given for every school *s* and year *t*.

Main independent variable

The main independent variable is a dummy that equals 1 if the teacher is observed in the mid or high salary scales LC or LD and 0 if the teacher is observed in the low salary scale LB. In

the next section we describe our empirical instrumental variable strategy to identify the effect of being placed in a higher salary scale on teacher retention.

It is important to note that we will look at the effects of being placed in a higher salary scale rather than at the effect of having a higher salary since the latter would not capture the full treatment given to teachers. Teachers are promoted by the Randstad policy rather than being given a higher salary in itself. Hence, a higher salary alone does not capture the future career prospects of being promoted to a higher salary scale. Taking salary as independent variable would invalidate the exclusion restriction as we cannot distinguish salary effects from the effects of possible future career prospects of being promoted.

Covariates

As covariates we use a number of school and teachers characteristics. Teacher covariates include age, teaching load, and gender. School covariates include school size, pupil population growth, and the share of pupils from disadvantaged neighborhoods.[9]

Other outcomes

In our analysis we also use other outcomes necessary to support the assumptions underlying our identification strategy. They include the unemployment rate, the pupil-teacher ratio, school board finances, the number of new teachers, the share of lessons given by a certified teacher and a dummy that indicates whether a teacher switches between Randstad and non-Randstad schools. These outcomes will be discussed when they are exploited.

3.2 Construction of main estimation sample

We have a full sample of 480,600 observations for which we have full information on teacher retention. If we use this sample in our analysis, then we compare the whole Randstad region with the rest of the Netherlands. These two regions might not be very similar. The setup of the Randstad policy allows us to create more similar regions. This can be done by selecting schools around the geographical cutoff that separates the two regions. The idea is that (teachers in) schools will become more similar if we select schools closer to the border of the Randstad. To create such a sample, we have selected 53 municipalities. They comprise the first two rings of municipalities around the geographical cutoff. Taking these two rings is based on the consideration that i) the municipalities are close to the Randstad border and ii)

[9] A pupil from a disadvantaged neighborhood is defined here as a pupil living in a so-called poverty problem accumulation area. See footnote 4 for a description of the criteria used for identifying these areas.

the treatment and control group would comprise a large enough sample size to estimate effects.[10] Note that the biggest four cities in the Randstad (Amsterdam, Rotterdam, Den Haag and Utrecht) are not included in this sample as they do not lie at the border. This local sample will be our main estimation sample. Table A.1 in the Appendix gives the list of the selected municipalities and figure A.1 provides a map.

3.3 Descriptive statistics

Table 1 shows descriptive statistics for our local and full sample. Panel A shows statistics for all years pooled together, panel B shows statistics for the pre-treatment year 2008. For each sample, statistics are given for the treatment group (Randstad) and control group (Outside Randstad). We observe similar patterns in panels A and B, except that in 2008 there are no significant salary differences between our control and treatment group. In the full sample there are statistically significant differences between the groups in terms of teacher and school characteristics. Randstad teachers are a bit younger, are more likely to be female, and have a somewhat smaller assignment size compared to non-Randstad teachers. In addition, Randstad schools are smaller in size and have more disadvantaged children than non-Randstad schools. In our local sample these differences disappear. This is what we would expect if we select schools in regions closer to the border of the Randstad; they should become more similar.

[10] Taking three rings would mean including big cities like Amsterdam and Utrecht in the treatment group but not in the control group. This would create less similar comparison groups. Taking only one ring would decrease the sample size substantially.

Table 1 Descriptive statistics for local and full sample, all years pooled (panel A) and pre-treatment year (panel B)

	Local sample			Full sample		
Variable	Outside Randstad (Control)	Randstad (Treatment)	P-value	Outside Randstad (Control)	Randstad (Treatment)	P-value
A: all years pooled						
Main outcome variables						
Retention as a teacher (a)	0.934	0.930	0.176	0.936	0.925	0.000
Teacher applies for schooling voucher (b)	0.030	0.029	0.483	0.027	0.025	0.094
Main independent variable						
Teacher in mid or high salary scale (a)	0.393	0.494	0.000	0.404	0.516	0.000
Covariates						
Teacher's age in years	45.29	45.09	0.609	45.85	45.06	0.000
Teacher's assignment size in FTE	0.822	0.823	0.912	0.826	0.809	0.000
Female teacher	0.484	0.479	0.686	0.479	0.501	0.000
School size (b)	1,524	1,512	0.950	1,54	1,399	0.105
Yearly school population growth (b)	0.016	0.011	0.507	0.007	0.011	0.351
Pupils from disadvantaged neigborhood (b)	0.083	0.091	0.679	0.098	0.238	0.000
Number of observations	61,611	58,882	120,493	279,149	201,451	480,600
Number of schools	78	74	152	350	287	637
B: pre-treatment year (2008)						
Main outcome variables						
Retention as a teacher (a)	0.916	0.915	0.895	0.926	0.910	0.000
Teacher applies for schooling voucher (b)	0.032	0.023	0.077	0.027	0.022	0.010
Main independent variable						
Teacher in mid or high salary scale (a)	0.357	0.340	0.344	0.357	0.361	0.686
Covariates						
Teacher's age in years	44.24	44.11	0.770	44.70	44.06	0.002
Teacher's assignment size in FTE	0.828	0.830	0.785	0.833	0.813	0.000
Female teacher	0.469	0.466	0.802	0.463	0.489	0.000
School size (b)	1,476	1,506	0.879	1,530	1,340	0.032
Yearly school population growth (b)	-.005	.003	0.315	-.005	-.004	0.867
Pupils from disadvantaged neighborhood (b)	0.094	0.097	0.906	0.111	0.245	0.000
Number of observations	7,736	7,645	15,381	33,721	25,002	58,723
Number of schools	76	72	148	346	270	616

Notes: a) weighted by the assignment size of teachers in FTE's, b) school averages.

4. Empirical Strategy

4.1 Instrumental Variables Framework

We are interested in the effect of the treatment, i.e. being placed in a higher teacher salary scale, on our outcomes. To estimate the treatment effect one could use the following specification:

$$(1)\ Y_{ist} = \alpha_0 + \alpha_1 HS_{ist} + \alpha_2 \boldsymbol{X_{ist}} + \theta_{ist}$$

in which Y_{ist} represents the outcome of teacher *i* in school *s* in year *t*, HS_{ist} represents a dummy that equals 1 if teacher *i* in school *s* at time *t* is in scale LC or higher (and 0 if in LB), $\boldsymbol{X_{ist}}$ is a set of controls, e.g. school fixed effects and teacher characteristics, and θ_{ist} is the error term which captures unobservable determinants of the outcome. For our outcome teacher retention we relate the retention decision at time t, which is measured with respect to t-1, to the conditions they were exposed to at t-1.[11] The parameter of interest is α_1, which represents the effect of being placed in a higher teacher salary scale on the outcome.

Using cross sectional data and estimating this specification with OLS will probably yield a biased estimate of α_1 because of the endogeneity of HS_{ist}. Salaries are not randomly assigned to teachers. On the contrary, there are a lot of reasons why some teachers end up earning more than others. Teachers and their salaries often differ from each other in ways not observed by the researcher. For instance, better teachers with unobserved qualities could have been placed in higher salary scales by the school board in order to keep them in the teaching profession. In that case their unobserved qualities influence both their salary and their retention decision, causing any OLS-estimate to be biased.

We therefore use a two stage least squares (IV-)approach to address this endogeneity problem. We exploit the Randstad bonus as an instrument for being placed in a higher teacher salary scale. This bonus affected and benefited the teachers in the schools in the Randstad

[11] In that case the index of the right-hand side variables is t-1. For instance, the retention decision of a teacher in 2014 with respect to 2013 is related to the salary scale and school characteristics she is exposed to in 2013. Because for our main analysis we pool our data over 2007-2014, the choice of using yearly retention rates as our outcome measure may give rise to selection effects over time. We address these issues in sections 5.2 and 5.3.

after 2008, while the teachers in the schools in the other regions were unaffected by this bonus. The first stage in this framework is

$$(2)\ HS_{ist} = \beta_0 + \beta_1 RS_s * POST_t + \beta_2 RS_s + \beta_3 \boldsymbol{X_{ist}} + \tau_t + \varepsilon_{ist}$$

in which RS_s represents a dummy that equals one if school *s* resides in the Randstad (RS) region; $POST_t$ represents a dummy that equals 1 if the outcome is observed post treatment, i.e. in 2009 or thereafter (*t≥2009*) and 0 if the outcome is observed pre-treatment, i.e. in 2007 or 2008, and τ_t are year fixed effects. The parameter of interest is β_1, which represents the effect of the Randstad bonus on the probability of being placed in a higher teacher salary scale. Note that this first-stage equation is a basic differences-in-differences model in which HS_{ist} is the outcome. The second stage is

$$(3)\ Y_{ist} = \gamma_0 + \gamma_1 \widehat{HS_{ist}} + \gamma_2 RS_s + \gamma_3 \boldsymbol{X_{ist}} + \tau_t + \vartheta_{ist}$$

where $\widehat{HS_{ist}}$ is the predicted probability of equation (2). Estimates of parameter $\gamma_1$yield the causal effect of the treatment on the outcome if the regular IV-conditions apply (see below). The corresponding reduced form of equation (3) is

$$(4)\ Y_{ist} = \delta_0 + \delta_1 RS_s * POST_t + \delta_2 RS_s + \delta_3 \boldsymbol{X_{ist}} + \tau_t + \theta_{ist}$$

in which δ_1 represents the impact of the Randstad policy on the outcome. This can be considered as an intention-to-treat effect. We use this specification to look at the impact of the Randstad policy on the share of teachers that applied for a schooling voucher. We cannot use an IV-approach for this outcome variable because applying for a schooling voucher precedes being placed in a higher salary scale. As noted in section 2, one of the criteria for being placed in a higher salary scale is that teachers would obtain additional schooling.

4.2 Validity of the IV-approach

To apply this IV-setup three conditions should hold. First, the Randstad policy should have a significant impact on the probability of being placed in a higher salary scale, which means that the model should not suffer from a weak instrument problem. Second, receiving the Randstad policy treatment bonus should be independent of the error term (second stage

independence). Third, the Randstad policy should *only* have an effect on teacher retention via the increased probability of being placed in a higher salary scale (second stage exclusion restriction). We address the first two conditions below. After presenting our main results, we discuss possible selection effects like switching behavior of teachers that may bias our IV-estimates. Also, we discuss the validity of the exclusion restriction.

First stage relevance

The Randstad teacher pay policy, introduced in 2009, should have a significant impact on teacher salary in the Randstad region as compared to regions outside the Randstad. Figures 2a (full sample) and 2b (local sample) show that this is the case. The share of teachers in the mid or high salary scale (i.e. not in the low salary scale) increases substantially more in the treatment group than in the control group after the introduction of the Randstad policy. The difference increases to 18 percentage points in 2014.[12] The F-statistics of the first-stage regressions are above 200, largely exceeding the rule of thumb of 10 (Staiger & Stock, 1997). This shows that we do not suffer from a weak instrument problem; that is, the Randstad policy significantly increases the probability of being placed in a higher salary scale for teachers in the Randstad region. In addition, the figures show that the pre-treatment trends are rather similar across treatment and control regions. This means that salaries were not that different between the regions before introduction of the policy. From the figures it also becomes clear that the nearly 20 percentage points difference between the Randstad and non-Randstad does not coincide with the policy goal of a 29 percentage points difference. According to the Ministry of Education the 290 million euro turned out to be insufficient to achieve the policy goal, see section 2.

[12] The impact on average gross salary of all teachers is 2.2 percentage points by 2014, which is also highly significant. This implies that the 18% of teachers that were additionally placed in a higher salary scale due to the policy received approximately 13% more salary (=2.2/0.18). The perspective of being placed in a higher salary scale was a 17% higher salary.

Figure 2a: Development of share of teachers in mid or higher salary scales for control and treatment group, full sample

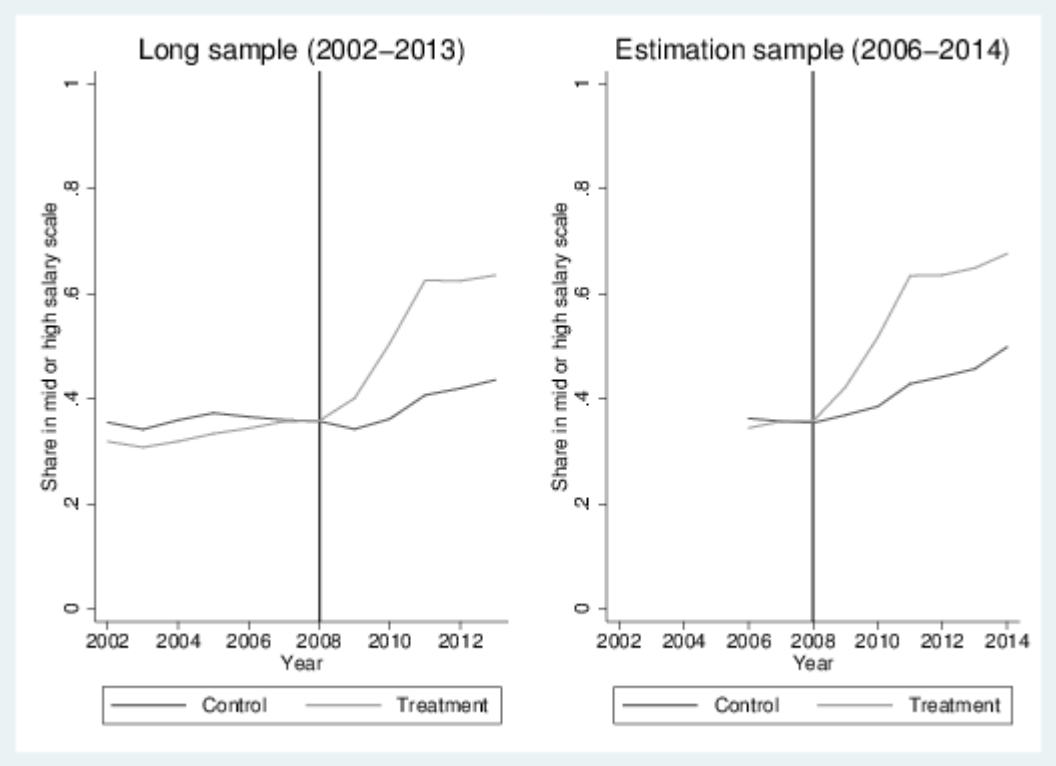
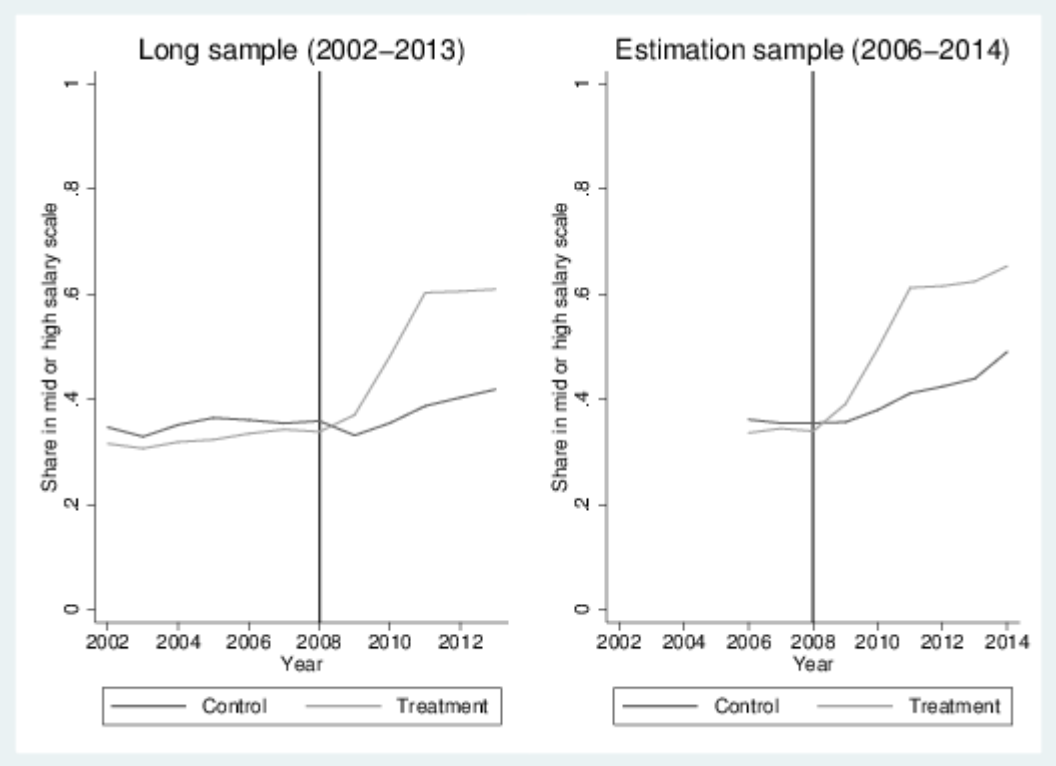


Figure 2b: Development of share of teachers in mid or higher salary scales for control and treatment group, local sample



Second stage independence

Second we address the second stage independence. This means that receiving the Randstad bonus should not be correlated with unobserved time-varying characteristics of teachers and their outcomes. As we compare Randstad teachers with teachers outside the Randstad over time, this comes down to a common trend assumption. That is, we assume that the trend in the outcome in the treatment group would have followed the same trend as that of the control group in absence of the treatment (i.e. receiving the Randstad policy). Although this assumption cannot be tested directly, its credibility can be strengthened by showing pre-treatment trends in the outcome variables for the treatment and control group. These trends should be similar and should not diverge until the introduction of the Randstad policy in 2009.[13] Our data allows us to investigate this for teacher retention but not for enrolment in schooling vouchers. We can test this with two datasets: a long dataset spanning the period 1995-2013 (*Mirror*) and a main estimation sample which spans the period 2007-2014 and has more covariates (*Functiemix)*. With the long sample we can test whether trends are similar before introduction of the policy as we have data that go back as far as 1995. The main estimation sample gives us only one year before the intervention as it only goes back to 2007. We will use both datasets to check whether trends are similar, but only use the main estimation sample to estimate the effects of the treatment in the next section.

We perform three analyses. First, we start with a graphical analysis. Figure 3a shows the trends in teacher retention rates for the control and treatment group for our full sample. The left (right) figure exploits the long (estimation) sample. A vertical is drawn at 2008, the last year before introduction of the policy. The figure shows that retention rates vary between 90 and 95%, and that the pre-trends between control and treatment group are rather similar. As the trends continue to be similar after introduction of the policy, they also suggest that there are no direct effects of the Randstad policy. Figure 3b does the same for our local sample. Hence this figure shows the trends when the dataset is limited to schools in the 53 border municipalities. The idea is that trends will become more similar if we select schools closer to the Randstad border. This seems to be confirmed as the lines lie closer to each other when compared to figure 3a.

[13] A reason why pre-trends could differ is when early announcement of the program would cause teachers to select themselves in Randstad schools before the start of the Randstad teacher pay policy. We think this is unlikely to be the case as information about the policy was not made public until spring of 2009.

Figure 3a: Development of retention rates for control and treatment group, full sample

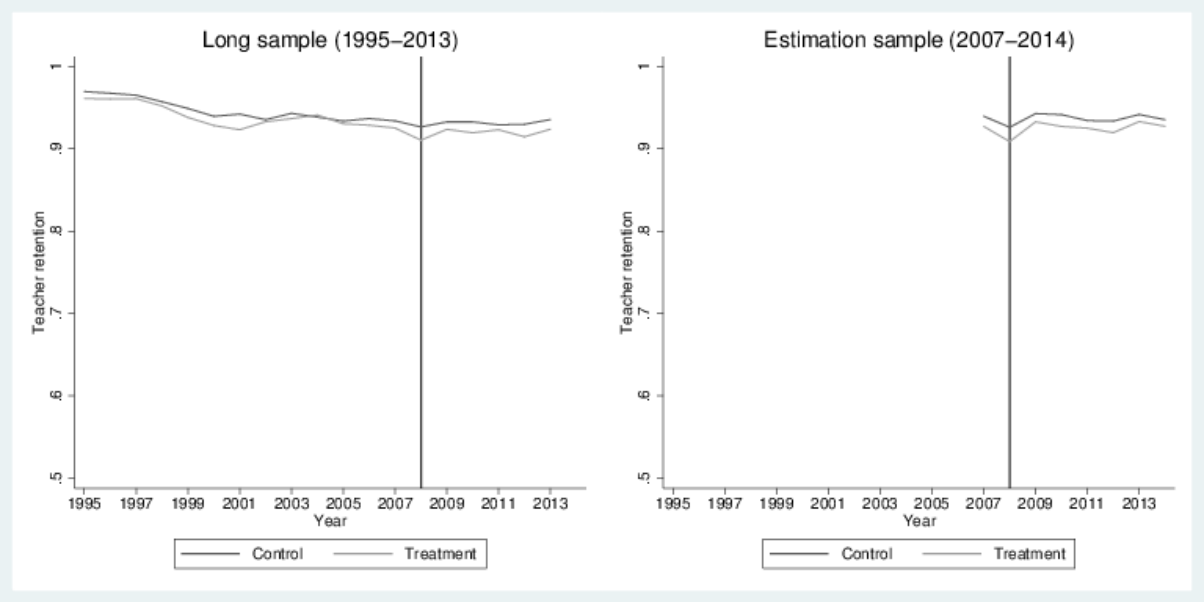
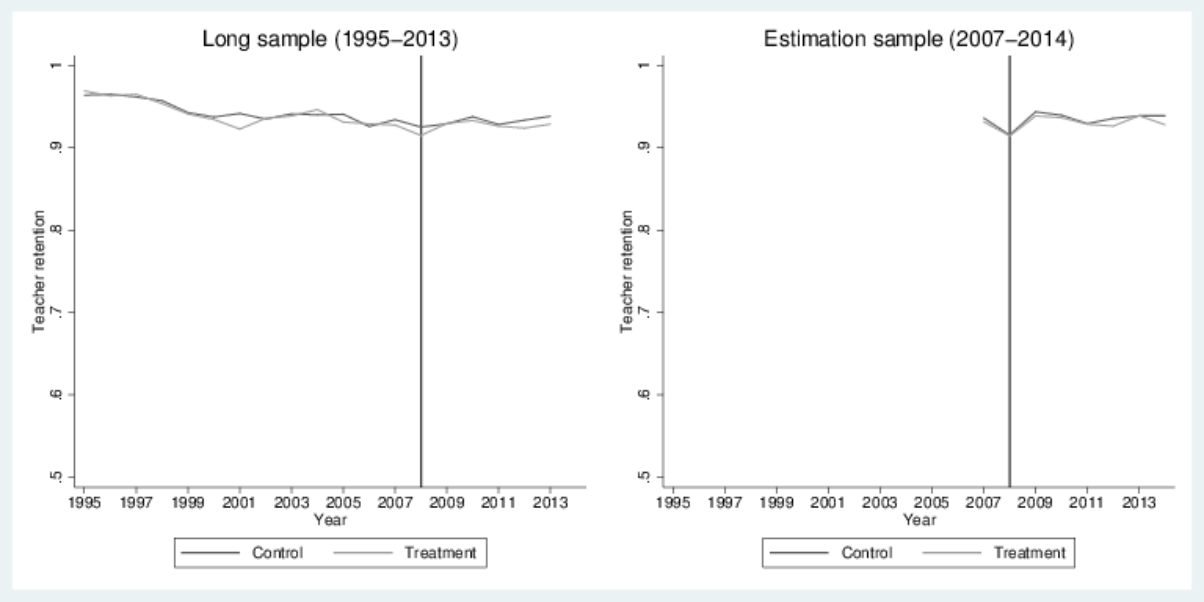


Figure 3b: Development of retention rates for control and treatment group, local sample



However, from this graphical analysis we cannot be certain yet that pre-trends are similar. To shed more light on the similarity of the pre-trends we perform a second analysis in which we statistically test whether trends are similar. We use the long sample and run two types of regressions.

First, we select the observations before 2008 and regress teacher retention on a constant, the set of available control variables, a linear time trend, a dummy for Randstad,

and the interaction of the time trend with the Randstad dummy. If pre-trends are similar, then the estimated coefficient for the interaction should be close to zero.

Second, we run a similar regression except that, instead of a linear time trend, we include dummies for the years and its interactions with the Randstad dummy. Hence, in this model we are more flexible and allow the time trends to deviate from each other in a non-linear way. If pre-trends are similar, then the estimated coefficients of the interactions should be close to zero.

Table 2 presents results of these regressions for our local and full sample. Panel A shows results of the model with a linear time trend, and panel B gives results of the model with a non-linear trend.

Table 2: Test on similarity of pre-treatment trends for teacher retention

	Local sample		Full sample	
	Estimate	Standard error	Estimate	Standard error
	(1)	(2)	(3)	(4)
Panel A: Linear trend				
year*treatment region	-0.000	0.000	0.000	0.000
Panel B: Non-linear trend				
1995*treatment region	0.016**	0.008	0.007*	0.004
1996*treatment region	0.005	0.009	0.006	0.004
1997*treatment region	0.011	0.008	0.010**	0.004
1998*treatment region	0.005	0.007	0.010**	0.004
1999*treatment region	0.006	0.008	0.004	0.004
2000*treatment region	0.009	0.008	0.004	0.005
2001*treatment region	-0.012	0.007	-0.005	0.005
2002*treatment region	0.012	0.009	0.012**	0.005
2003*treatment region	0.004	0.008	0.008*	0.004
2004*treatment region	0.016**	0.008	0.016***	0.004
2005*treatment region	-0.001	0.009	0.010**	0.005
2006*treatment region	0.014	0.012	0.007	0.005
2007*treatment region	0.004	0.007	0.006	0.004
Number of observations	206,654		907,950	

Notes: Every pair of columns in each panel represents the results of an OLS-regression. The odd columns give the estimates and the even columns give the standard errors. We control for the covariates as presented in table 1. The estimates in panel B represent deviations from the trend with respect to 2008 (2008=omitted category). Standard errors are adjusted for clustering at the school level. Significance levels: *** p<1%, ** p<5%, * p<10%.

Panel A shows that the pre-trends in the teacher retention rates between control and treatment group do not deviate from each other when using a linear time trend. The estimated coefficients for the interaction term are close to zero and insignificant in both columns.

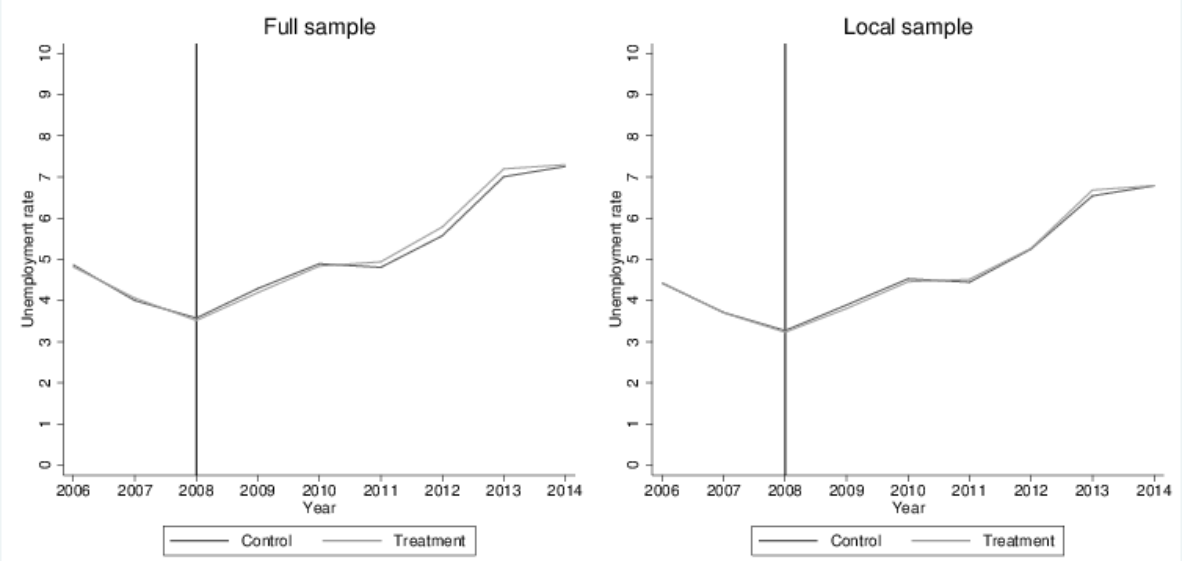
Panel B shows that there are small deviations in some years when we allow the trend to be non-linear. In the full sample we find that for some years the estimated coefficient of the

interaction is significant, suggesting that the pre-trend of the treatment group for these years deviate from that of the treatment group. However, when we limit the full sample to our border sample, the estimated coefficients are no longer significant anymore for most of the years.[14]

For our third test we investigate whether the labor market conditions are different between Randstad teachers and non-Randstad teachers. Although we are quite confident that the previous tests support the common trend assumption in the local sample, which suggests that unobserved differences between teachers develop similarly between regions, we consider this third test as an extra check on the independence assumption. If labor market conditions develop more favorably for teachers in the Randstad than for those outside the Randstad, for instance if Randstad teachers have more outside options during recessions than non-Randstad teachers, then this might affect their retention decision differently. Figure 4 shows unemployment rates for Randstad and non-Randstad regions for our full and local sample. The figure suggests that there are no large differences between the regions with respect to employment: both the level and the development of unemployment look similar, both pre- and post-treatment.[15] When we perform statistical tests as in table 2, the null hypothesis of the similarity of the trends is not rejected. This suggests that labor market conditions do not develop differently, which may give extra support to the common trend assumption. In the next section we present our main results. Thereafter we continue with a discussion on possible selection effects and the exclusion restriction.

[14] We stress that with so many years and hence estimates, it is not unlikely that some estimated coefficients pop up significant at conventional significance levels when testing a true null hypothesis of no effect.
[15] The unemployment rates have been weighted by the size of the labor force by municipality.

Figure 4: Trends in unemployment rates for control and treatment group



5. Main results

5.1 Effects of placement in higher salary scale on teacher retention

Table 3 contains first stage, reduced form, OLS- and IV-estimates. The OLS and IV results show estimates of the effect of being placed in a higher (i.e. non-low) salary scale on teacher retention. The reduced-form (RF) estimates shows (intention-to-treat) estimates of the effect of the regional teacher policy on teacher retention. The first stage, also in the rows, represents the effect of the policy on the probability of being in a higher salary scale. The first (last) four columns exploit our local (full) sample. Odd columns include no controls except year-fixed effects. Even columns include school-fixed effects, teacher's age, gender, teaching load, school size, population growth, and the share of pupils from a disadvantaged neighborhood. Standard errors have been clustered at the school level.

The first stage estimates mirror the picture in figure 2. They are highly significant and around 0.16. They suggest that the Randstad policy led to a 15-17 percentage points increase in the probability of being in a higher salary scale.

The OLS-estimates in column (1) and (2) vary between 0.02 and 0.03 and are significant at the 1% level. They suggest that being placed in a higher salary scale leads to a 2-3 percentage points higher probability of being retained. These estimates cannot be interpreted causally, because of the endogeneity of being placed in a higher salary scale.

Teachers in higher pay scales differ from teachers in lower pay scales in ways not observed by the researcher.

The IV-results control for this endogeneity and show negative and insignificant estimates. The point estimate is -0.015 in column (3). Including controls in column (4) hardly changes the estimate. This is what one would expect when treatment and control groups are similar. Hence, based on these IV-results, we find no effect of being placed in a higher salary scale on teacher retention.

We continue by investigating whether the results found with our local sample can be replicated with the full sample in columns (5) to (8). In these regressions we are less confident about the validity of the second stage independence assumption (see previous chapter). The OLS-estimates are similar to those in columns (1) and (2). The IV-estimates are insignificant. The IV-estimate in column (7) is 0.022 but drops to 0.008 when including controls in column (8). This may reflect the fact that treatment and control regions are less similar in the full sample. The result in column (8) seems to replicate the result found with the local sample. We conclude that we find no effect of being placed in a higher salary scale on retention as a teacher.

Table 3: Estimates of the effect of a higher salary scale (OLS and IV) and of teacher pay policy (reduced form) on retention as a teacher

	Local sample				Full sample			
	OLS	OLS	IV	IV	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Effect on retention	0.028***	0.021***	-0.015	-0.018	0.026***	0.020***	0.022	0.008
	(0.003)	(0.003)	(0.031)	(0.029)	(0.002)	(0.002)	(0.017)	(0.017)
Reduced form			-0.002	-0.003			0.003	0.001
			(0.005)	(0.005)			(0.003)	(0.003)
First stage			0.154***	0.160***			0.155***	0.159***
			(0.013)	(0.011)			(0.008)	(0.007)
Number of observations	120,493	120,493	120,493	120,493	480,600	480,600	480,600	480,600
Controls								
teacher characteristics	no	yes	no	yes	no	yes	no	yes
school characteristics and school-fixed effects	no	yes	no	yes	no	yes	no	yes

Notes: In the even columns we include a set of controls. Included teacher covariates are: gender, age category and assignment size category in FTE's. School covariates include school size category, the share of disadvantaged pupils and school population growth. Standard errors in parentheses adjusted for clustering at the school level. All regressions include year-fixed effects. Significance levels: *** p<1%, ** p<5%, * p<10%.

5.2 Selection effects

In our identification strategy we compare teachers in Randstad regions with teachers in non-Randstad regions over time. As we pool our data over 2007-2014 and use yearly retention rates as our outcome measure, this may give rise to selection effects that might bias our estimates. Two things can happen, which both change the composition of the teacher population in the Randstad relatively to that of the non-Randstad regions.

First, the Randstad policy can lead to unobserved different inflow and outflow of teachers. After each year teachers enter and exit the teaching profession, causing the teacher population to change over time. If the Randstad policy causes other types of teachers to enter or exit the teaching profession in the Randstad regions than in non-Randstad regions, this may bias our estimates of the effect of the treatment on teacher retention. This can happen, for instance, if newly entering teachers who select themselves in schools in the Randstad because of the higher salary are more likely to leave the teaching profession.

Second, the Randstad policy might lead to switching behavior of existing teachers. Teachers outside the Randstad who favor higher salaries might leave their schools and move to schools inside the Randstad. In addition, Randstad teachers may become less willing to switch to schools outside the Randstad because of the higher salary. In the same way as above, this may bias our estimates if switchers differ in unobserved ways from non-switchers.

To address these two issues we perform four tests. The first three tests relate to the first issue, the fourth relates to the second.

First, we look at the effects of the Randstad policy on the number of new teachers per school. If the policy increases the attractiveness of the teaching job and hence the number of teachers, then this might be an indication of a changing teaching population. Panel A in table 3 shows reduced form estimates of the effect of the policy on the number of new teachers per school. We find no evidence in favor of an increased influx.[16]

Second, we investigate whether the policy changes the composition of the teacher population in the Randstad relative to the non-Randstad. Panel B investigates this issue by showing estimates of the effect of the policy on (observable) background characteristics of teachers. We do not find evidence in favor of a changing distribution in terms of age, gender

[16] This finding is supported by recent research among bachelor and master students in teacher training programs. It was found that these students seriously underestimate both starting and maximum wages for teachers, that is, by 15% and 40% respectively (Researchned, 2015). It thus seems that prospective teachers are unaware of improved career prospects for teachers.

or assignment size in the local sample; the estimated coefficients are insignificant. We also do not reject the null hypothesis of similarity of the pre-trends in these variables (not shown in table).

Third, we investigate whether the quality of teachers has changed because of the policy. Changes in the quality of teachers could hint at composition effects and may lead to biased estimates of effects on teacher retention if teacher quality (certification) is correlated with teacher retention. We use the share of lessons given by a certified teacher as a proxy for teacher quality. It has been found that being certified for the subject is positively correlated with pupil outcomes (Goldhaber&Brewer, 2000; Clotfelter et al., 2010). Panel C of table 3 shows estimates of the effect of the Randstad policy on the share of lessons given by a certified teacher. The statistically insignificant and close-to-zero estimates do not hint at composition effects in terms of teacher quality.

Fourth, we look at the effects of the Randstad policy on switching behavior of teachers. Although less than 1% of the teachers switches annually between control and treatment regions, we will investigate to what extent switching behavior has changed due to the Randstad policy. This analysis also sheds light on the question whether the Randstad bonus succeeds in keeping more teachers in the targeted region. One of the goals of the policy is to retain teachers in the Randstad region. If teachers switch less from treatment to control group because of the policy, this could be considered a success. In panel D of table 3 we show reduced form estimates of the effect of the policy on a dummy that equals 1 if a teacher switches from Randstad region to non-Randstad region or vice versa. We find a small albeit statistically significant effect of -0.4 percentage points ($p<0.05$). This shows that switching behavior of teachers decreased a bit due to the policy.[17] The estimate suggests that those who would have switched from Randstad regions to regions outside the Randstad in absence of the policy, now stick to the Randstad because of the higher salary. In the next section we show that switchers are more likely to exit the teaching profession. In our IV-setup we would then estimate a lower bound, because after introduction of the policy the treatment group will consist of a larger share of teachers with a higher probability of leaving the teaching profession. In the next section, we therefore investigate the sensitivity of our estimates with respect to switching behavior. In addition, we investigate the sensitivity of our estimates with respect to teacher's entries and exits, although the tests provided in this section do not hint at

[17] In this differences-in-differences setting the estimate would also have been negative if teachers in the control group would switch more often to the treatment group. Graphs of switching behavior, however, show that the effect is driven by the treatment group. It can be seen that switching behavior decreases in the treatment group relatively to that of the control group. Graphs are available upon request.

a changing distribution of teachers. The exclusion restriction of our IV-strategy will be discussed thereafter in Section 5.4.

Table 4: Reduced form estimates of impact of policy on various variables, check for composition effects

	Local sample		Total sample	
	(1)	(2)	(3)	(4)
A: Number of new teachers per school	-0.927	-0.517	-0.765	-0.702
	(1.521)	(1.089)	(0.722)	(0.537)
Number of observations	1,142	1,142	4,729	4,729
B: Background characteristics of teachers				
Age	0.167		-0.012	
	(0.242)		(0.126)	
Female	-0.009		-0.008***	
	(0.007)		(0.003)	
Assignment size	0.002		0.005**	
	(0.004)		(0.002)	
Number of observations	161,662		651,667	
C: Share of lessons given by a certified teacher	0.008	0.001	0.003	0.001
	(0.009)	(0.008)	(0.005)	(0.004)
Number of observations	62,154	62,154	247,368	247,368
D: Region switching	-0.004**	-0.004**	-0.004***	-0.004***
	(0.002)	(0.002)	(0.001)	(0.001)
Number of observations	111,082	111,082	442,893	442,893
Controls:				
teacher characteristics	No	Yes	No	Yes
school characteristics	No	Yes	No	Yes

Notes: Each cell is an OLS-regression. Standard errors are adjusted for clustering at the school level in panels B to D. In panel A robust standard errors are used. School-year average teacher and school covariates are used in panel A. The share of lessons given by a certified teacher is weighted by total number of lessons given. All regressions include year-fixed effects. Significance levels: *** p<1%, ** p<5%, * p<10%.

5.3 Sensitivity analysis

In this paragraph we test to what extent our estimates are sensitive to possible selection effects. First we address region switching, as this has been shown to be an issue. Thereafter we will address selection effects that might occur because of unobserved changes in the teacher population composition. Although the tests in the previous paragraph do not hint at a changing teacher population composition, we cannot be fully sure that unobserved characteristics of entering and exiting teachers develop differently over time in the Randstad than in the non-Randstad regions because of the policy. For all our sensitivity analyses we use our local sample and include the full set of controls.

First, we investigate the robustness of our results with respect to switching behavior of teachers between control and treatment regions. Columns (1) and (2) in table 5 show results of this sensitivity analysis. In column (1) we select the observations for which we have full information on teacher retention and switching, and run our IV-regression while controlling for (an indicator of) switching. By doing so we lose another year of our data as switching behavior is measured with respect to the previous period. For example, we investigate whether a teacher who switched in 2013 with respect to 2012 is retained in 2014 with respect to 2013. The estimate in column (1) is similar to that in column (4) of table 3. Including the switching variable hardly changes the IV-estimate. Switching in itself, however, seems not to be trivial. A teacher who switches between treatment and control region in a particular year has a 9 percentage points higher probability to drop out in the next year. We therefore also run an IV-regression in which we exclude all switchers from our estimation sample in column (2). This estimate is in the same order of magnitude as the previous IV-estimates. Hence our results seem to be robust to switching behavior.

We proceed by addressing possible unobserved changes in the teacher population composition. For our main estimation results we pool the data from 2007-2014, hence we do not distinguish between short and long term effects. However, short-run effects on teacher retention would hint at a changing distribution of teachers, such that estimates of medium-run effects could be biased. Pooling our data as we do in our main specification by including a post-treatment by Randstad interaction dummy would then render biased effect estimates. In columns (3) and (4) we therefore distinguish between short- and medium-run effects by running our IV-regression for years 2010-2011 (short run) and 2012-2014 (medium run) separately. Both short-run and medium-run effects are statistically insignificant and point estimates are (slightly) negative. The absence of short-run effects indicates that our main estimation results do not suffer from possible selection effects.

We continue with a final test on possible selection effects. We take the teacher population from the pre-treatment year 2008 for our local sample of border municipalities and follow this cohort over time. As such, we rule out the risk of selection effects due to a changing teacher population as we keep the estimation sample fixed. We investigate to what extent these 2008-teachers exit the teacher profession for post-treatment years 2009-2014. In columns (5)-(10) we show reduced form estimates of the effect of the Randstad policy on a

dummy that equals 1 if the teacher is observed working in the teaching profession in a particular year and 0 otherwise. That is, we look at their retention rate with respect to 2008.[18]

Again, we find no effects of the Randstad policy on this outcome variable. When we perform the same analysis for teacher cohorts from pre-treatment years 2003-2007, the results are similar and never significantly different from zero. Figure 5 shows these results graphically. In these graphs we show the development of retention rates for control and treatment group for these cohorts over time, i.e. for the 2003-cohort, 2004-cohort, etc. An advantage of taking a number of years before treatment is that we can investigate whether control and treatment group have the same pre-trend for this outcome variable. It can be observed that the retention rates of control and treatment group almost lie on top of each other and develop rather similarly over time before introduction of the policy. We have also empirically tested the similarity of the pretends and have found no evidence in favor of deviating trends.[19]

Table 5: Estimates of the effect of a higher salary scale (IV) and of teacher pay policy (RF) on retention as a teacher, local sample

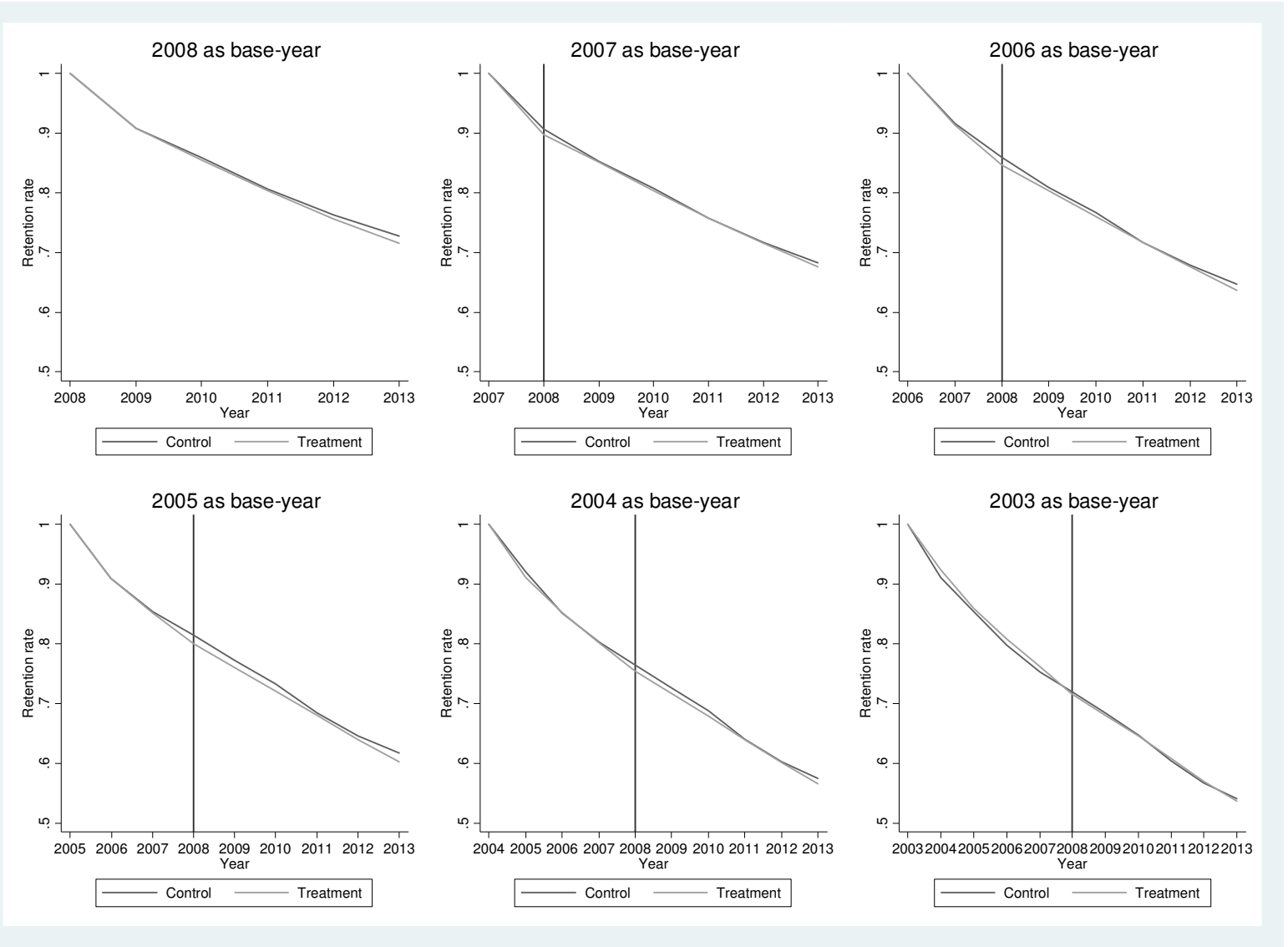
	Taking into account region switching		Short term (10-11)	Medium term (12-14)	1 - 6-year retention rate					
	IV	IV	IV	IV	RF	RF	RF	RF	RF	RF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					2009	2010	2011	2012	2013	2014
Effect on retention	-0.019	-0.014	-0.005	-0.021	-0.008	-0.013*	-0.012	-0.014	-0.011	-0.020
	(0.048)	(0.030)	(0.064)	(0.025)	(0.006)	(0.008)	(0.009)	(0.010)	(0.011)	(0.013)
Region switch	-.091***									
	(0.015)									
Reduced form	-0.003	-0.002	-0.000	-0.004						
	(0.008)	(0.005)	(0.006)	(0.005)						
First stage	0.165***	0.162***	0.090***	0.209***						
	(0.011)	(0.011)	(0.012)	(0.013)						
Number of observations	95,379	115,425	67,188	84,010	15,717	15,717	15,717	15,717	15,717	15,717
Controls:										
teacher characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
school characteristics and school-fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: Included teacher covariates are: gender, age category and assignment size category in FTE's. School covariates include school size category, the share of disadvantaged pupils and school population growth. Regressions in columns (1) to (4) include year-fixed effects. Standard errors in parentheses adjusted for clustering at the school level. Significance levels: *** p<1%, ** p<5%, * p<10%.

[18] Hence we look at the probability that a 2008-teacher is in the educational labor market after 1 year (in 2009) , 2 years (in 2010), 3 years (in 2011) etc. Note that this retention rate is different from a survival rate because we allow teachers to reenter the system after a drop out. Note also that we cannot use our IV-strategy when using this outcome variable, as we do not have information on teacher's salary after a teacher drops out.

[19] We performed tests like in table 2. Results available upon request.

Figure 5: Retention rates for different cohorts of teachers



5.4 Second stage exclusion restriction

In this section we address the second stage exclusion restriction of our IV-strategy. The increase in the probability of being placed in a higher salary scale should be the only channel through which the Randstad policy may have an impact on teacher retention. Is this the case? Although the Randstad policy is meant to increase the salaries of the teachers, it might be possible that the policy has not been fully used for this purpose. The extra funds have been given to autonomous schools that are in principle free to choose how to spend this additional money. Hence, instead of increasing teacher salaries, schools might spend it (partly) on other activities such as reducing the pupil-teacher ratio. This may be a threat to the exclusion restriction. For example, if the additional funds are used to hire new teachers to reduce the pupil-teacher ratio instead of promoting teachers, then this channel might affect the teacher retention decision. Reductions in class size may cause teachers to stay in the profession as these reductions might render the teaching profession more attractive. In that case the Randstad policy affects the retention decision via class size reductions. The second stage exclusion then fails. To address these types of issues we estimate the effect of the policy on a number of variables that can be considered channels through which the policy might affect the outcome. We estimate the effect of the policy on the pupil-teacher ratio (just discussed), the share of non-teaching personnel, the amount of money saved by the school board (i.e. yield of school board) and the share of school board expenditures not spent on personnel. The last three outcomes may be relevant if the extra funds are not given to teachers but saved or given to non-teaching personnel. Table 6 shows results of this analysis. All estimated reduced form effects are close to zero and insignificant. This gives support to the second stage exclusion restriction. Hence, we have no indications that the schools have spent the additional funds to destinations other than placement of teachers in higher salary scales.

Table 6: Test on the exclusion restriction: reduced form estimates of the effect of the Randstad policy on various outcomes

	Local sample		Full sample	
Effect Randstad policy on:	(1)	(2)	(3)	(4)
A: Pupil-teacher ratio	0.107	0.160	0.122	0.133
	(0.162)	(0.157)	(0.086)	(0.087)
Number of observations (school-year)	1,274	1,274	5,223	5,223
B: Share of non-teaching personnel	0.003	-0.002	-0.001	0.003
	(0.006)	(0.006)	(0.003)	(0.003)
Number of observations (employee-year)	224,745	224,745	904,902	904,902
Controls				
school-fixed effects	no	yes	no	yes
teacher and school covariates	no	yes	no	yes
C: Yield of school board (in %-points)	-0.750	0.338	-0.601	-0.225
	(-1.360)	(0.748)	(0.620)	(0.521)
Number of observations (board-year)	597	597	2,190	2,190
D: Share of expenses of the school board not spend on personnel	-0.001	-0.000	0.003	0.004
	(0.007)	(0.007)	(0.004)	(0.004)
Number of observations (board-year)	597	597	2,190	2,190
Controls				
School board-fixed effects	no	yes	no	yes

Notes: All regressions include year-fixed effects. Standard errors in parentheses are clustered at the school (panel A and B) or board (panel C and D) level. For pupil-teacher ratio (panel A) and share of non-teaching personnel (panel B) we include the same personnel and school covariates as in table 3.

5.5 Effects of the policy on teacher's schooling decisions

One of the criteria for being placed in a higher salary scale is that a teacher would obtain extra schooling. The policy should therefore lead to a higher share of teachers being enrolled in additional education. We do not have a direct measure for this outcome. We use the share of teachers that applied for a schooling voucher as a proxy for actual enrollment in degree programs among teachers. This seems to be a reliable proxy since Van der Steeg and Van Elk (2015) show that nine out of ten applicants actually start with the study they applied for. Table 7 contains reduced form estimates of the effects of the Randstad policy on this outcome. Columns (1) and (2) show the results for the full sample, columns (3) and (4) for the local sample. The results in the full sample suggest that the Randstad policy increased the probability of applying for a teacher schooling voucher by 0.5 percentage points. The estimates are statistically significant at the 5% level. In the local sample the point estimates are higher (0.9 percentage points), but marginally insignificant (p=0.11). With 2.3 % having applied for a schooling voucher before introduction of the policy, the estimated effect comes

down to a 39% (=0.009/0.023) increase in the probability of applying for a schooling voucher, and hence, in enrollment in a bachelor or master study of teacher education. This finding is consistent with the setup of the policy in which one of the criteria for placement in a higher salary scale was that teachers would gain extra qualifications or expertise.

It should be noted though that for this analysis we could not check the plausibility of the second stage independence, i.e. common trend, assumption as we have no data on enrollment in teacher schooling vouchers before 2008. This is because the teacher schooling voucher was not introduced until 2008.

Table 7: Reduced-form estimates of the effect of the Randstad policy on the share of teachers that applied for a schooling voucher

	Local sample		Full sample	
	(1)	(2)	(3)	(4)
Effect of Randstad policy	0.009*	0.009	0.005**	0.005**
	(0.006)	(0.006)	(0.002)	(0.002)
Number of observations (school-year combinations)	885	885	3649	3649
Controls				
school-fixed effects	no	yes	no	yes
teacher and school covariates	no	yes	no	yes

Notes: Each column is an OLS-regression. The even columns include the same set of control variables as in table 3, except for the fact that teacher covariates have been aggregated at the school level. Standard errors in parentheses adjusted for clustering at the school level. All regressions include year-fixed effects. Significance levels: *** p<1%, ** p<5%, * p<10%.

6. Heterogeneous treatment effects

In this section we investigate whether treatment effects for teacher retention differ by teacher's age and gender, and by school's population composition. We study age and gender effects because earlier literature suggests that young teachers (Gilpin, 2011; Hendricks, 2014; Hendricks, 2015) and male teachers (Dolton, 2006) are more sensitive to higher salary with respect to their retention decisions.[20] We study effects by school composition as it has been consistently shown that teachers in schools with a higher share of low-SES (or disadvantaged) pupils are less likely to be retained (e.g. Boyd et al., 2002; Hanushek et al, 2004; Bonhomme et al, 2015). A higher teacher pay might therefore affect these teachers differently than teachers in schools with lower shares of low-SES pupils.

[20] Hendricks (2014) however finds no differences in the sensitivity to higher wages across males and females.

Table 8 shows the results of this heterogeneous treatment effects analysis.[21] It shows estimates for three different age categories, for males and females, and for two groups of schools: one with more than 10% pupils from high-poverty areas and one with less than 10% pupils from high-poverty areas. All estimated effects are statistically insignificant and do not significantly differ from each other. Hence, we find no evidence for retention effects for these subgroups.

Table 8: Heterogeneous treatment effects of higher salary scale (IV) and Randstad policy (RF) on teacher retention

	Age			Sex		% of pupils from high-poverty areas	
	18-34	35-54	>55	Male	Female	<=10%	>10%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Retention as a teacher							
IV (a)	0.005	0.017	-0.061	0.001	0.017	0.013	0.023
	(0.039)	(0.017)	(0.039)	(0.023)	(0.020)	(0.021)	(0.030)
First stage	0.126***	0.182***	0.140***	0.139***	0.174***	0.162***	0.156***
	(0.011)	(0.009)	(0.009)	(0.007)	(0.009)	(0.010)	(0.010)
RF (b)	0.001	0.003	-0.008	0.000	0.003	0.002	0.004
	(0.005)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)
Number of observations	113,902	226,294	140,404	245,981	234,619	291,593	189,007

Notes: All models include the same set of controls as in the even columns in table 3. Standard errors in parentheses adjusted for clustering at the school level. Significance levels: *** p<1%, ** p<5%, * p<10%.
a) IV indicates effects of being placed in a non-low salary scale.
b) Reduced form indicates estimates of the effect of the teacher pay policy on the outcome of interest.

7. Conclusion and discussion

In this paper we have investigated the effects of higher teacher pay on teacher retention and teacher's schooling decisions in secondary education. We exploited variation in teacher pay induced by the introduction of a new remuneration policy. The policy provided schools in an urbanized region with extra funds to place a higher share of teachers in a higher salary scale. The salaries of teachers in the targeted regions were increased by approximately 13%. We used the regional variation in an IV-setup to estimate the effects of being placed in a higher on our outcomes. The setup of the new remuneration policy allowed us to create similar

[21] We use the total sample for this heterogeneity analysis to increase power. Results are similar when we use our local sample but estimates are much less precise due to the smaller sample size when splitting up the local sample by teachers' age sex and SES.

treatment and control groups by selecting (teachers in) schools around the geographical cutoff that separate treatment and control regions.

Our main findings are as follows. First, and most importantly, we find no effects of higher teacher pay on teacher retention. That is, we do not find that placement in a higher salary scale leads to a higher probability to stay in the teacher profession. Second, we find that the policy led to a small reduction in annual switching from treatment to control regions, but that this does not affect our results found for teacher retention. Hence, the policy succeeded in keeping a slightly larger share of teachers in the targeted region. However, these positive effects are small relative to the costs of the policy. The policy cost on average about 50 mln euro per year and around 0.4% of teachers per year decided not to switch from the treatment to the control region because of the policy. This would imply a cost of about 400k euro to prevent one teacher from switching from the treatment to the control region.[22] Third, we find that the policy has a positive impact on teachers' enrollment in additional schooling. Our estimates suggest that the policy increased teachers' enrollment in bachelor or master degree programs from 2.3% to 3.2%. This finding is consistent with the setup of the policy in which one of the criteria for placement in a higher salary scale is that teachers would complete extra schooling.

What do we learn from these results? First we discuss why we do not find effects for teacher retention. Although the realized salary increase of approximately 13% (and a prospect of a 17% increase end-of-scale) is by no means small, it may not have been large enough to increase teacher retention rates. This suggests that the retention decision is often motivated by factors other than salary. Studies by Hanushek et al. (2004) and Clotfelter et al. (2011) show that effects of teacher pay on retention are very modest compared to the effect of pupil characteristics. Teachers prefer not to work in schools with high shares of low-SES children (e.g. Boyd et al., 2002; Hanushek et al, 2004; Bonhomme et al., 2015). This suggests that salaries need to be increased substantially in order to increase retention rates of teachers in schools or regions with relatively high shares of low-SES children. It would be interesting to investigate what other policies could be more (cost-)effective. Policies one could think of are better guidance of starting teachers (i.e. induction programs) or opening up and investing in alternative routes to teaching to recruit highly talented people in hard-to-staff schools or regions, such as Teach for America.

Second, we discuss the effects of the new remuneration policy on teachers' enrollment in additional schooling programs and switching behavior. Our findings suggest that offering opportunities to be placed in a higher salary scale can induce the existing teacher

[22] Calculation is available upon request.

workforce to participate in additional schooling, and hence can be used as an incentive to get a better qualified teaching workforce. Furthermore, our results suggest that a higher teacher pay can be used to reduce switching out of a shortage region. This is consistent with the study of Hendrickx (2014) that also finds positive effects of higher teacher pay on retention rates at the regional level.

Taken these findings together, we conclude that a higher teacher pay may not be effective in increasing retention rates in the teaching profession, but might be effective in decreasing turnover rates in specific schools or regions, especially in schools or regions with relatively high shares of disadvantaged pupils.

Literature

Bonhomme, S., G. Jolivet, and E. Leuven, 2015, School characteristics and teacher turnover: assessing the role of preferences and opportunities, *Economic Journal* (forthcoming), retrieved from http://onlinelibrary.wiley.com/doi/10.1111/ecoj.12279/epdf

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff, 2002, Initial matches, transfers and quits: career decisions and the disparities in average teacher qualifications across schools, retrieved from http://www.teacherpolicyresearch.org/ResearchPapers/tabid/103/Default.aspx.

Chevalier, A., P. Dolton, and S. McIntosh, 2007, Recruiting and retaining teachers in the UK: An analysis of graduate occupation choice from the 1960s to the 1990s, *Economica*, vol. 74, pp. 69-96.

Clotfelter, C., H. Ladd, and J. Vigdor, 2008, Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina, *Journal of Public Economics*, vol. 92, no. 5-6, pp. 1352-1370.

Clotfelter, C., H. Ladd, and J. Vigdor, 2010, Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects, *Journal of Human Resources*, vol. 45, no. 3, pp. 655-681.

Clotfelter, C., H. Ladd, J. Vigdor, 2011, Teacher mobility school segregation and pay-based policies to level the playing field, *Education Finance and Policy*, vol. 6, no. 3, pp 399-438.

Dolton, P., 1990, The economics of UK teacher supply: the graduate's decision, *Economic Journal*, vol. 100, pp. 91-104.

Dolton, P., 2006, Teacher supply, in: Hanushek en Welch (eds), *Handbook of Economics of Education*, vol. 2, chapter 18.

Dolton, P., and W. van der Klaauw, 1995, Leaving teaching in the UK: a duration analysis, *Economic Journal*, vol. 105, pp. 431-444.

Dolton, P. and O. Marcenaro-Gutierrez, 2011, If You Pay Peanuts Do You Get Monkeys? A Cross-Country Analysis of Teacher Pay and Pupil Performance, *Economic Policy*, Vol. 26, No. 65, pp. 5-55.

Gilpin, G.A., 2011, Re-evaluating the effect of non-teaching wages on teacher attrition, *Economics of Education Review*, vol. 30, no. 4, pp. 598-616.

Goldhaber, D., and D. Brewer, 2000, Does teacher certification matter? High school teacher certification status and student achievement, *Educational Evaluation and Policy Analysis*, vol. 22, no. 2, pp. 129-145.

Greaves, E., and L. Sibieta, 2014, Estimating the effects of teacher pay on pupil attainment using boundary discontinuities, *IFS Working Paper*, W14/03.

Hanushek, E., J. Kain & S. Rivkin, 1999, Do Higher Salaries Buy Better Teachers?, *NBER Working Paper*, no. 7082.

Hanushek, E., J. Kain, and S. Rivkin, 2004, Why public schools lose teachers, *Journal of Human Resources*, vol. 39, no. 2, 326-354.

Hendricks, M., 2014, Does it pay to pay teachers more? Evidence from Texas, *Journal of Public Economics*, vol. 109, pp. 50‑63.

Hendricks, M., 2015, Towards an optimal teacher salary schedule: designing base salary to attract and retain effective teachers, *Economics of Education Review*, vol. 47, pp. 143-167.

Heyma, A., D. de Graaf, and C. van Klaveren, 2006, Exploratie van beloningsverschillen in het onderwijs 2001-2004, Stichting Economic Onderzoek (SEO).

Imazeki, J., 2005, Teacher salaries and teacher attrition, *Economics of Education Review*, vol. 24, no. 4, pp 431-449.

Manski, C., 1987, Academic ability, earnings, and the decision to become a teacher: Evidence from the National Longitudinal Study of the High School Class of 1972: University of Chicago Press.

Ministry of Education, 2015, Functiemix, Letter to parliament on May 29 2015.

Murnane, R., J. Singer, and J. Willett, J., 1989, The influences of salaries and "opportunity costs" on teachers' career choices: Evidence from North Carolina, *Harvard Educational Review*, vol. 59, no. 3, pp. 325-34.

Reed, D., K. Rueben, and E. Barbour, 2006, Retention of new teachers in California, retrieved from http://www.ppic.org/content/pubs/report/R_206DRR.pdf.

Researchned, 2015, Invulling en inrichting van een tegemoetkoming studiekosten lerarenopleidingen voortgezet onderwijs.

Rickman, D., H. Wang, and J. Winters, 2015, Adjusted state teacher salaries and the decision to teach, *IZA Discussion Paper*, no. 8984

Staiger, D., and J. Stock, 1997, Instrumental variables regression with weak instruments, *Econometrica*, vol. 65, no. 3, pp. 557-586.

Steeg, M. van der, and R. van Elk, 2015, The effect of schooling vouchers on higher education enrollment and completion of teachers: A regression discontinuity analysis, *CPB Discussion Paper*, no. 305.

Wolter, S., and S. Denzler, 2003, Wage elasticity of the teacher supply in Switzerland, *IZA Working Paper*, no. 733, Bonn.

Appendix

A.1 Data preparation

We took four steps to prepare the data for our analyses. First, we removed teachers that are employed in a Randstad school and in a non-Randstad school at the same time (499 observations, i.e. teacher-year combinations). For these teachers we cannot determine whether they belong to the control or treatment group. Second, we removed teachers for whom we have missing data on age and gender, or whose reported age is lower than 18 (1,882 observations). Imputing missing values on these covariates and including them in the estimation sample does not change results. Third, we removed data on (teachers in) schools in a particular year if more than 50% of the teachers drops out of school in that year (18,236 observations). In that case the school did not (correctly) provide the personnel data to the Ministry of Education (DUO). The Ministry of Education applies this criterion as well before using the data for calculating statistics. Fourth, we imputed missing values for school size and the share of disadvantaged pupils (300 observations).

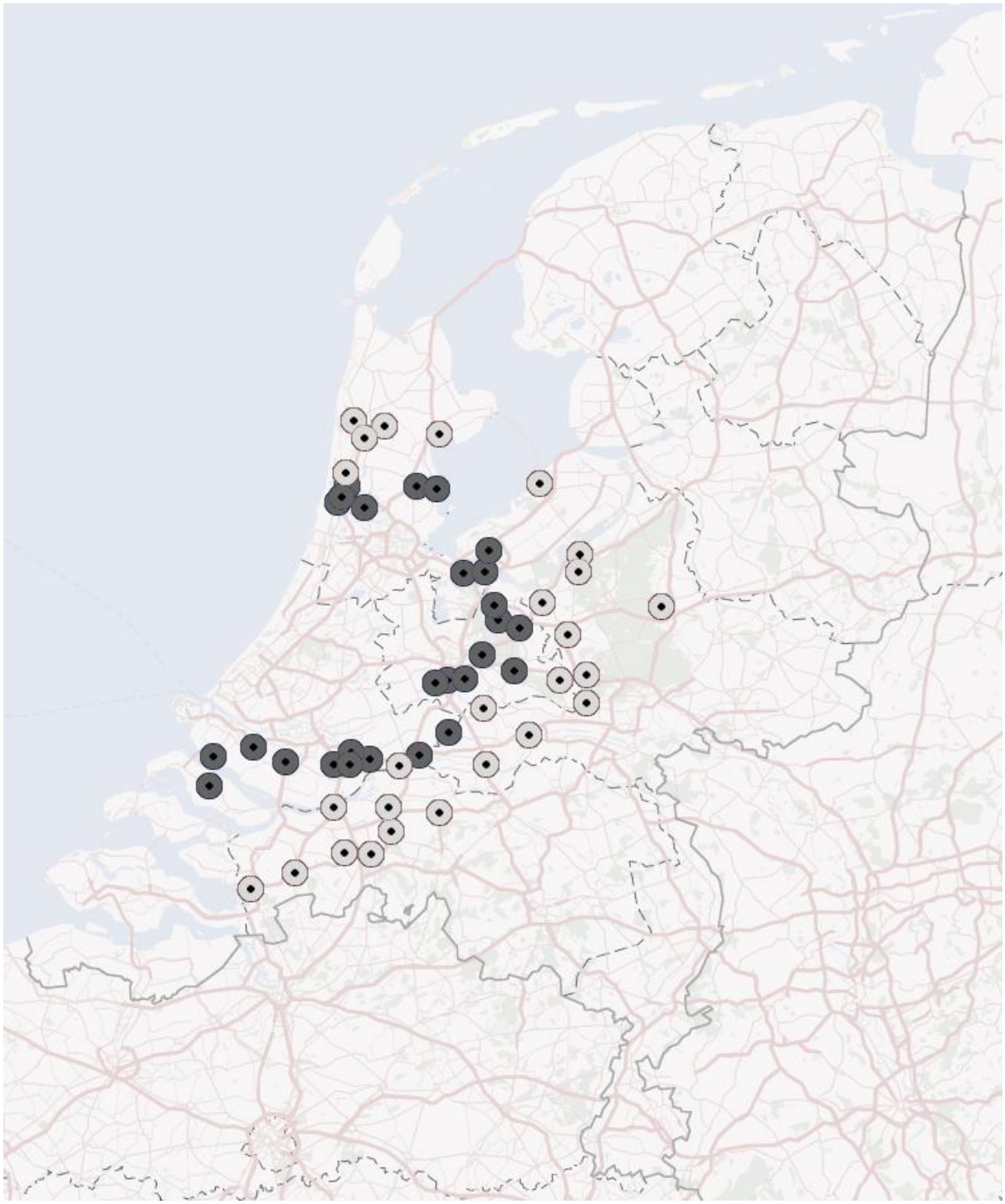
A.2 Construction of local sample

We selected 53 municipalities at the border of the Randstad for our local estimation sample. Table A.1 gives the list of the selected municipalities and figure A.1 provides a map. We selected the first two rings of municipalities around the Randstad border. Taking these two rings was based on the consideration that these municipalities are close to the border and that they would comprise a sample size large enough to estimate effects. Note that the biggest four cities in the Randstad (Amsterdam, Rotterdam, Den Haag and Utrecht) have not been included in the sample as they do not lie at the border.

Table A.1: 53 Municipalities in local sample

Control group	Treatment group
Alkmaar	Almere
Apeldoorn	Amersfoort
Barneveld	Baarn
Bergen (NH.)	Beverwijk
Bergen op Zoom	Dordrecht
Breda	Edam-Volendam
Castricum	Goeree-Overflakkee
Culemborg	Gorinchem
Ede	Heemskerk
Ermelo	Hellevoetsluis
Etten-Leur	Houten
Geertruidenberg	Huizen
Harderwijk	Ijsselstein
Heerhugowaard	Leerdam
Hoorn	Naarden
Lelystad	Nieuwegein
Moerdijk	Nissewaard
Nijkerk	Oud-Beijerland
Oosterhout	Papendrecht
Roosendaal	Purmerend
Tiel	Sliedrecht
Veenendaal	Soest
Waalwijk	Utrechtse Heuvelrug
Wageningen	Velsen
Werkendam	Zaanstad
Zaltbommel	Zeist
	Zwijndrecht

Figure A.1: Selection of 53 border municipalities for the local sample. Bullets in dark (light) grey are treated (control) municipalities.



Publisher:

CPB Netherlands Bureau for Economic Policy Analysis
P.O. Box 80510 | 2508 GM The Hague
T (070) 3383 380

December 2015 | ISBN 978-90-5833-708-5