ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Master Thesis Policy Economics

School accountability and teacher absence

Abstract

This study investigates the relationship between the qualification of a primary school given by the Inspectorate of Education and sickness absence of teachers in the Netherlands in the period 2016-2019. Inspection is a tool for school accountability. The expectation is that school accountability causes more distrust, which causes more job stress, which will lead to more teacher burnout and more sickness absence of teachers. This research will conduct three methods, namely Ordinary Least Squares, Fixed Effects and a Regression Discontinuity Design using data from the Inspectorate of Education and the Education Executive Agency. The estimates show that a deterioration in the qualification has a statistically significant association with sickness absence of teachers in 2016-2019. It will result in an increase in sickness absence of teachers of approximately 19.3%. However, the sensitivity analyses suggest reverse causality. Therefore, it is most plausible that the direction of the relationship is from sickness absence of teachers to the qualification of a school given by the Inspectorate of Education.

Author: Lieke Spekschoor
Student ID number: 537514
Supervisor: prof. dr. H.D. Webbink
Second assessor: prof. dr. A.C. Gielen
Date final version: November 30, 2020

Table of Contents

1. Introduction

School accountability is increasingly used in the world, whereby inspection is a tool for accountability (Figlio & Loeb, 2011; Learmonth, 2000; Bowers & McIver, 2000). School accountability is the process of evaluating school performance based on pupil performance measures. The aim of school accountability is to increase pupil achievement. However, there are also unintended consequences of school accountability. There is further research needed about these unintended consequences. For example, the influence of school accountability on health outcomes like the sickness absence of teachers (Figlio & Loeb, 2011)

The theory predicts a positive relationship between school accountability and sickness absence of teachers. School accountability might cause more distrust (e.g. Dworkin & Tobe, 2014). Distrust might cause more job stress (e.g. Dworkin & Tobe, 2014). More job stress implies more teacher burnout and more sickness absence of teachers (e.g. Balyer et al, 2017).
However, there is limited empirical evidence. For example, previous studies about the relationship between school accountability/inspection and sickness absence of teachers have almost exclusively focused on the United States and the United Kingdom (e.g. Dworkin et al, 2003, Troman & Woods, 2001).
Next to the limited empirical evidence, there is also mixed empirical evidence. For example, the expectation is that if a school accountability system leads to more stress, more teachers will leave the teaching profession as a consequence (Perryman & Calvert, 2020). However, Sun et al (2017) investigated that the school accountability system No Child Left Behind in the United States does not change the average national rate of teachers that leave the teaching profession.

Due to the limited and mixed empirical evidence, this research will look at the unintended consequences of school accountability on sickness absence of teachers in the Netherlands, whereby inspection is used as a tool for accountability. The focus is on primary schools in the period 2016-2019. Consequently, the research question is: *What is the relationship between the qualification of a primary school given by the Inspectorate of Education and sickness absence of teachers in the Netherlands in the period 2016-2019?*

To answer the research question, this research will conduct three methods. These are Ordinary Least Squares, Fixed Effects and a Regression Discontinuity Design. Furthermore, to examine heterogeneity, the sickness absence of teachers in large schools will be compared with sickness absence of teachers in small schools. Additionally, this research will also investigate the

sickness absence of teachers in schools that have a deterioration in the qualification from ‘basis’/’sufficient’ to ‘weak’/’insufficient’.

To briefly summarize the results, a deterioration in the qualification from ‘good’ to ‘basis’/’sufficient’ or from ‘basis’/’sufficient’ to ‘weak’/’insufficient’ or from ‘weak’/’insufficient’ to ‘very weak’ has a statistically significant association with sickness absence of teachers in 2016-2019. It will result in an increase in sickness absence of teachers of approximately 19.3%. However, there is a concern about reverse causality, which implies that a higher sickness absence of teachers will lead to a deterioration in the qualification. Therefore, some robustness checks will be presented.

The remainder of the paper is organized as follows. Section 2 provides the context regarding the inspection of schools in the Netherlands. Section 3 provides the theory and empirical evidence, especially about the relationship between school accountability/inspection and sickness absence of teachers. Section 4 explains the data that is used in this research. Section 5 explains the empirical methodology. Section 6 presents the results. Section 7 presents the results of the sensitivity analyses. Section 8 concludes and discusses the findings.

2. Context

This section contains an overview of the school inspection system in the Netherlands. The supervision framework and inspection framework are discussed.

2.1. Dutch Inspectorate of Education

The Dutch Inspectorate of Education is part of the Ministry of Education, Culture and Science and is responsible for the inspection and qualifications of schools (Inspectorate of Education, 2020a). The Dutch Inspectorate of Education supervises for example the financial policy of schools. The way the money was spent has to be illustrated in the annual report of the schools (Government of the Netherlands, 2020a). The Inspectorate of Education controls whether the school spend the government money properly (Government of the Netherlands, 2020b).

Since the 1st of August in 2017, supervision by the Inspectorate of Education has been changed. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b).

2.1.1 Supervision framework

The supervision framework has three different phases in primary and secondary education. The first phase is the risk analysis. Every year, the risk of all schools will be analyzed. The Inspectorate of Education will look at three different things. These are the outcomes of a school, potential problems and information from annual reports. The purpose of the risk analysis is to determine which school will receive the qualification 'basis' and for which school there is further research needed before determining the qualification. When a school with qualification 'basis' scores below the norm on learning outcomes for one year according to the risk analysis, there will be preventive supervision. The school will be informed about their outcomes, where the purpose is to make the governing body of the school aware of the need to take preventive measures. When a school with qualification 'basis' scores below the norm on learning outcomes for two years according to the risk analysis, there will be more preventive supervision. The governing body of the school will receive a warning that the school will receive the qualification 'weak' or 'very weak' if the learning outcomes are also below the norm next year. If it is needed, the Inspectorate of Education will do a quality inspection. When it turns out that there are shortcomings, measures have to be taken to prevent that the school receives the qualification 'weak' or 'very weak'.

Schools that have not received the qualification ‘basis’ will go to the second phase. This is the quality inspection. There will be an investigation regarding possible shortcomings. The purpose of this phase is to determine the qualification. The investigation can be a visit to the school or verifying the analysis of the governing body of the school. If it turns out that there are no shortcomings, the qualification is ‘basis’. If it turns out that there are shortcomings, the qualification is ‘weak’ or ‘very weak’.
Schools that have received the qualification ‘weak’ or ‘very weak’ will go to the third phase. In this phase there are interventions by the supervisor or the minister. The seriousness of the shortcomings will determine which interventions have to be taken. The purpose of every intervention is to remove the shortcomings, so that the school receives the qualification ‘basis’ (Inspectorate of Education, 2012).

There is a framework for primary education that determines the qualifications. There are 45 indicators. These indicators are classified in five domains: ‘outcomes’, ‘teaching-learning process’, ‘special needs provision and guidance’, ‘quality assurance’ and ‘statutory regulations’. There are ten indicators called norm indicators. In general, a primary school will receive the qualification ‘very weak’ when the school scores insufficient on the domain ‘outcomes’ and two or more norm indicators of the ‘teaching-learning process’ and ‘special needs provision and guidance’ are insufficient (Inspectorate of Education, 2012).

Since October 2007, the Inspectorate of Education is obliged to visit all the schools one time every four years, even if there are no risks at those schools. If the Inspectorate of Education has to visit a school with the qualification ‘basis’, the Inspectorate of Education will only look at the domains ‘outcomes’, ‘quality assurance’ and ‘special needs provision and guidance’ (Inspectorate of Education, 2012).

2.1.2 Inspection framework

The inspection framework came into effect from the 1st of August in 2017. Every year the framework will be updated, so that the framework corresponds to the current law and regulations. The inspection framework consists of an annual performance analysis and a four-yearly school inspection (Inspectorate of Education, 2020c).

The purpose of the annual performance analysis is to detect possible risks in school quality and financial management of governing bodies and schools. When there is no presumption of risk,

the analysis will be closed. When there is a presumption of risk, the Inspectorate of Education will do an expert analysis. An expert will give a judgement about the seriousness and size of the risks. If the risks are huge and have to be taken seriously, there will be a meeting with the governing body of the school. The purpose of the meeting is to make a judgement about the extent to which the governing body of the school effectively combats the risks. After the meeting, the Inspectorate of Education will decide whether there is a reason to do a quality inspection (Inspectorate of Education, 2020c). The Inspectorate of Education will investigate schools where there is the presumption that they offer insufficient quality or show risks regarding financial management (article 11 Wet op het onderwijstoezicht; Inspectorate of Education, 2020c).

The Inspectorate of Education is obliged to visit all the schools every four years, this is the four-yearly school inspection. This investigation starts with an expert analysis, meeting with the governing body of the school, plan of investigation and presentation of the school. After that, there will be an investigation at the level of the school and at the level of the governing body of the school. This will result in a concept report. There will be a feedback meeting and a final meeting with the school, which will result in a final report (Inspectorate of Education, 2020c).

The inspection framework consists of five quality areas: ‘educational process’, ‘school climate’, ‘learning outcomes’, ‘quality assurance and ambition’ and ‘financial management’. Every quality area consists of standards, for example ‘didactics’ in the quality area ‘educational process’. For every standard, the basis quality is determined by the Inspectorate of Education. Whether a school scores ‘sufficient’ or ‘insufficient’ on a particular standard is determined by whether the school meets the basis quality. It is also possible that schools have their own ambitions and purposes in addition to the basis quality. When these additional ambitions and purposes are convincing and the school meets the basis quality, then the score on this particular standard might be ‘good’ (Inspectorate of Education, 2020c).

The final qualification is ‘good’, if the school scores on all standards at least ‘sufficient’, the standard ‘quality culture’ is ‘good’, at least two standards of ‘educational process’ and/or ‘school climate’ are ‘good’ and all the legal requirements are met.
The final qualification is ‘sufficient’, if the standards ‘development perspectives’, ‘didactics’, ‘safety’ and ‘results’ are ‘sufficient’ and no more than one standard in the area of ‘educational process’ is ‘insufficient’.

The final qualification is ‘insufficient’, if the standard ‘development perspectives’ or ‘didactics’ or ‘safety’ or ‘results’ is ‘insufficient’ and at least two other standards in the areas ‘learning outcomes’, ‘educational process’ or ‘school climate’ are ‘insufficient’.

The final qualification is ‘very weak’, if the standard ‘results’ is ‘insufficient’ and at least one of the standards of ‘development perspectives’ or ‘didactics’ or ‘safety’ is ‘insufficient’ (Inspectorate of Education, 2020c).

3. Literature Review

This section provides the theory and empirical evidence, especially about the relationship between school accountability/inspection and sickness absence of teachers.

3.1. School accountability

The government provides a lot of funding for schools. Therefore, the government needs to know how well schools are doing. However, there is a principal-agent problem in schooling. It is difficult for stakeholders, for example policy makers, to monitor the activities of schools. The stakeholders are the principals and the teachers are the agents. The teachers might behave in a way that is contrary to the interests of the stakeholders. A school accountability system can overcome the principal-agent problem in schooling. School accountability is the process of evaluating school performance based on pupil performance measures. This gives incentives to teachers to concentrate on the subjects and materials that are being measured. This provides the stakeholders with independent information how well schools are doing (Figlio & Loeb, 2011).

A school accountability system can generate incentives through direct government intervention (explicit) or incentives through the provision of information (implicit). In the case of direct government intervention, a school accountability system generates rewards and/or sanctions to schools on the basis of pupil performance on standardized tests. There is for example the possibility of a bonus to high-performing schools or threats of closing low-performing schools. In the case of the provision of information, accountability ratings help the community to observe the school performance. This provides community pressure (Figlio & Loeb, 2011).

The intended consequence of school accountability is to increase pupil achievement. However, there are also unintended consequences of school accountability. For example, teaching to the test, selective assignment of students to special education programs or eating to the test (Figlio & Loeb, 2011). The question remains whether there are unintended consequences of a school accountability system on teacher absence and student absence. This subject will be discussed in the next paragraphs.

3.2. School accountability and sickness absence of teachers

3.2.1 Distrust

According to Bryk and Schneider (2002), there are two forms of relational trust. These are organic trust and contractual trust. In a social system with organic trust, individuals give their

trust unconditionally, because they believe in the rightness of the system and all the other individuals. In a social system with contractual trust, there is a contract which will define the basic actions that have to be taken by the parties of the contract (Bryk & Schneider, 2002).

When a school accountability system comes into place, there is a shift from organic trust to contractual trust. Prior to the school accountability system, there was the belief that teachers would teach in the best interests of their pupils. This is organic trust. However, in the case of a school accountability system, the belief is that teachers would act in their own best interests and not in the best interests of their pupils. This is contractual trust, which can also be called as distrust (Dworkin & Tobe, 2014).

3.2.2 Job stress

With organic trust, there is belief that teachers would teach in the best interests of their pupils. Consequently, the teachers receive job autonomy in the classroom. Only if there is clear evidence that a teacher has failed to teach in the best interests of the pupils, there will be an intervention and the teacher might lose their job. But if this is not the case, teachers could assume that there is job security. With a school accountability system, there is the belief that teachers would act in their own best interests and not in the best interests of their pupils. Consequently, teachers cannot longer expect to have classroom autonomy and job security. This causes more stress (Dworkin & Tobe, 2014). Furthermore, the school accountability system provides more pressure on the test scores of the pupils, this causes more stress by teachers and principals (Dworkin et al, 2003). Teachers have to follow a strict curriculum to increase the test scores of the pupils. This constant and increasing pressure on teachers causes teacher stress, whereby the school accountability is a major cause of the teacher stress (Haberman, 2004).

3.2.3 Sickness absence of teachers

Job stress will lead to more teacher burnout and changes trust between teachers and other stakeholders in schools (Dworkin & Tobe, 2014). Accountability systems are based on externally imposed criteria. This is in contrast with teacher empowerment, where the teachers have job autonomy. It is believed that teacher empowerment lowers teacher burnout, which implies that less teacher empowerment leads to more teacher burnout (Balyer et al, 2017). Less teacher empowerment implies less job autonomy, which causes more stress (Dworkin & Tobe, 2014). Thus, less job autonomy can lead to more stress and more teacher burnout (Balyer et al, 2017). More teacher burnout implies higher sickness absence of teachers.

3.2.4 Conceptual model

Based on the previous sections, Figure 1 presents a conceptual model which shows the relationship between school accountability and sickness absence of teachers.

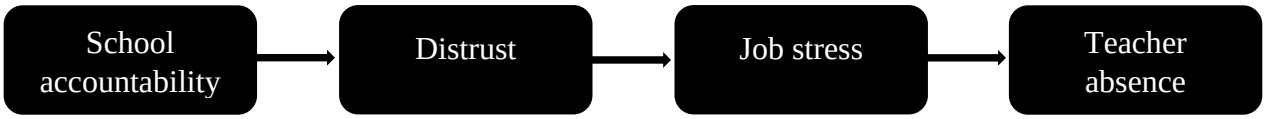

Figure 1. Relationship between school accountability and sickness absence of teachers

This hierarchy of distrust in a school accountability system can generate more job stress and therefore higher burnout rates, which implies more sickness absence of teachers. Higher burnout rates can generate again less trust, which heightens job stress and turns into burnout and more sickness absence of teachers. Thus, this can become a vicious cycle (Dworkin & Tobe, 2014).

3.2.5 Influence of school accountability on sickness absence of teachers

A school accountability system destroys trust and can generate more job stress. Job stress will lead to more teacher burnout (e.g. Dworkin & Tobe, 2014). More teacher burnout implies higher sickness absence of teachers. This is an unintended consequence of a school accountability system, where a school accountability system might cause more sickness absence of teachers. In this section, empirical evidence of the relationship between school accountability and sickness absence of teachers will be discussed.

A large number of existing studies in the broader literature have examined the relationship between a school accountability system and sickness absence of teachers in the United States. For example, Dworkin et al (2003) administered a survey questionnaire to all the teachers in a school district in Houston. This district has received awards from the Texas Education Agency for the performance of their pupils. The researchers concluded that a democratic personnel policy diminishes the teacher burnout. A school accountability system can make democratic schooling more problematic because there is control over the curricular decision-making in a school accountability system. This implies that a school accountability system increases teacher burnout (Dworkin et al, 2003). Furthermore, using 30 years of survey data in Houston, Dworkin (2009) concluded that changes in a school reform by for example the school accountability system No Child Left Behind exacerbated job stress and teacher burnout. Because teachers are often blamed for not working hard enough to raise pupils standardized achievement test scores.

Thus, job stress is a main determinant of a burnout and school accountability exacerbated job stress (Dworkin, 2009). Additionally, looking at three school reforms in the United States since 1983, each wave exacerbated burnout of teachers in a Texas school district, which employs 12000 teachers (Dworkin, 2001).

Next to research in Houston (Texas), Ryan et al (2017) did a survey among 1866 public school teachers in Pennsylvania, Maryland and Connecticut. This study identifies that an accountability system results in more teacher stress and more teacher burnout (Ryan et al, 2017). Furthermore, Von der Embse et al (2016b) investigated the relationship of accountability on teacher stress and school climate in four states in the United States. Using data from 6428 teachers, the researchers concluded that accountability generates more pressure and more teacher stress (Von der Embse et al, 2016b). Additionally, Berryhill et al (2009) investigated whether primary school teachers in South Carolina perceived accountability policies as having unintended consequences for their well-being. Traditionally, the standardized test scores of South Carolina have been the lowest of the United States. The researchers concluded that primary school teachers perceived accountability policies as having unintended consequences for their well-being (Berryhill et al, 2009). Moreover, Richards et al (2018) uses data from interviews with 28 teachers from the Midwest of the United States. Several teachers documented that the accountability system took control out of the hands of the teachers, which led them feel that there is no trust (Richards et al, 2018). Furthermore, Saeki et al (2018) uses data from 541 kindergarten (primary schools) across three states. The researchers concluded that accountability policies may have a negative influence on school environment and teacher stress (Saeki et al, 2018). Moreover, Mathison & Freeman (2006) did naturalistic research among teachers from three different school districts in the Upstate New York. An accountability system results in changes in the nature of a teacher's work and the nature of their workplace. It might also have unhealthy effects, for example stressed teachers (Mathison & Freeman, 2006). Furthermore, Shernoff et al (2011) did a qualitative study of the sources and impact of stress among urban teachers in three low-performing primary schools in high poverty urban neighborhoods in a large Midwestern city. Teachers described pressure and stress around accountability policies with three subdomains. First of all, teachers experienced stress, pressure and personal responsibility due to the fact that they have to raise pupils' test scores. Second, teachers in an underperforming school experienced intense supervision by the principal and local superintendent as a source of stress. Third, teachers experienced stress due to the fact that they are forced to teach to the test. Teachers feel limited control. For example, there is more a

focus to pupils that are most likely to score high on standardized tests (Shernoff et al, 2011). Additionally, Webb (2006) did a yearlong case study in a Cypress Elementary School. With a school accountability system, teachers are responsible for the pupils' outcomes. There was more teacher stress when risk factors in pupils' lives generates worse pupils' outcomes, for example emotional and physical abuse, alcoholism, drugs and divorce (Webb, 2006).

Next to research in the United States, Van Droogenbroeck et al (2014) uses survey data of 1878 Flemisch teachers who were 45 years of age or older. They investigated that paperwork and accountability are the most important stressors and causes of teacher job dissatisfaction (Van Droogenbroeck et al, 2014). Furthermore, Perryman & Calvert (2020) uses survey data of teacher education graduates of the Institute of Education in London. Since the Education Reform Act in 1988, the work of teachers in England has been increasingly regulated with more accountability in education. There is a problem in the United Kingdom that teachers leave the profession within five years, whereby accountability plays an important role (Perryman & Calvert, 2020).

Regarding the relationship between job stress and sickness absence of teachers, Smulders (2004) uses data from the Netherlands Working Conditions Survey (NEA) 2003. There are more than 10,000 employees in the Netherlands surveyed about questions regarding sickness absence. Workload and job stress play a big role in the cause of sickness absence of teachers in primary education, this is in 46,2% of the cases. In secondary education, workload and job stress play also a big role in the cause of sickness absence of teachers, this is in 48,1% of the cases. In general, workload and job stress play a big role in the cause of sickness absence in a lot of jobs, this is in 27,9% of the cases. As just mentioned, this is even higher for teachers. Job stress and workload can generate different complaints. The most important complaints are cardiovascular disease, mental health problems (for example burnout), tiredness, stomach problems, intestinal problems and skin problems (Smulders, 2004).

As previous research stated, a school accountability system can cause more teacher stress and more sickness absence of teachers, for example teacher burnout.

According to Von der Embse et al (2016a), strengthening teachers' efficacy may be an important factor to help teachers deal with the teachers' stress that is coming from school

accountability. Von der Embse et al (2016a) came to this conclusion when he did research among a sample of 1242 teachers in one state in Southeastern United States.

It is important to take into consideration that the level of teacher burnout is dependent on the extent to which the school accountability threatens job security, according to Dworkin & Tobe (2014). Using a sample of teachers collected between 2002 and 2012, a survey was administered in Houston. An accountability system such as No Child Left Behind increases stress levels for teachers and the burnout rates. However, if the threats of an accountability system such as losing jobs were not realized, stress and burnout rates diminished. But, if there are budgetary constraints in a school district resulting in closing programs and losing jobs, the accountability system will affect job security, which will generate more stress and more burnout (Dworkin & Tobe, 2014).

If a school accountability system can cause more teacher stress, teachers will leave the teaching profession as a consequence (Perryman & Calvert, 2020). The expectation is that teachers will especially drive away from teaching in disadvantaged schools. In general, pupils from disadvantaged schools have a low test score, and teachers are blamed for not working hard enough to raise pupils standardized achievement test scores. However, Sun et al (2017) investigated this expectation by using a difference-in-difference approach. While there is a lot of research that concludes that accountability leads to more teacher stress, Sun et al (2017) investigated that the school accountability system No Child Left Behind in the United States does not change the average national rate of teachers that transfer between schools (for example drive away from disadvantaged schools) or leave the teaching profession.

Overall, a school accountability system can cause more teacher stress and sickness absence of teachers, for example teacher burnout. Whereby strengthening teachers' efficacy may be important in dealing with the teacher stress and if the threats of an accountability system such as losing jobs were not realized, stress and burnout rates diminished. The question remains whether a school accountability system can cause more teachers to transfer between schools or leave the teaching profession, because of conflicting research.

3.2.6 Inspection

Inspection is a tool for accountability (Learmonth, 2000; Bowers & McIver, 2000). In the Netherlands, there is accountability by the Inspectorate of Education (Inspectorate of

Education, 2020a). Most inspectorates have the ambition to realize 'school improvement through inspection'. An example of an inspectorate is the Office for Standards in Education, Children's Service and Skills (Ofsted) in the United Kingdom (Ehren & Visscher, 2006). The Ofsted inspection has to contribute to the policy aim in the United Kingdom, which indicates an improvement in the quality of education (Rosenthal, 2004). However, inspection seems to generate unintended consequences (Ehren & Visscher, 2006). An example of an unintended consequence is teacher stress. Teacher stress is a problem in the United Kingdom. The cause of teacher stress lies in the increase in government intervention in education, for example school inspection by Ofsted. Stress is a term for teachers to describe the pressures, anxiety and dilemmas in their work (Hepburn & Brown, 2001). The cost of teacher stress is not quantifiable (PA Consulting 2000; Alexander, 1999).

A large number of existing studies in the broader literature have examined the unintended consequences of an Ofsted inspection in the United Kingdom. For example, Troman & Woods (2001) investigated primary teachers' stress. Under an Ofsted inspection, there is distrust of teachers. This is a source of stress. Furthermore, Case et al (2000) did a qualitative study in the field about the impact of Ofsted inspection on primary teachers. Teachers feel intimidated and stressed by the inspection process (Case et al, 2000). Additionally, Phillips et al (2007) uses a questionnaire in a population of 290 head teachers and principals. An example of a case that generates more stress is an Ofsted inspection (Phillips et al, 2007). Moreover, according to Bubb & Earley (2004) teachers perceive additional pressure. For example, due to increasingly difficult parents and students, poor financial rewards, low public esteem, lack of control over their job and Ofsted inspections (Bubb & Earley, 2004). Furthermore, Learmonth (2000) concluded that an inspection is seen as punitive, and this generates a climate of fear leading to stress and anxiety among teachers (Learmonth, 2000).

Next to research in the United Kingdom, Alkutich & Abukari (2018) did research among four private schools in Dubai. They came to the conclusion that a school inspection is effective and benefiting teaching and learning. However, there are concerns about some parts of the inspection which has a negative effect on the benefits. These are issues like fear and anxiety. This can lead to stress which can affect teacher effectiveness (Alkutich & Abukari, 2018).

Next to stress, Perryman (2007) finds a negative emotional impact of an Ofsted inspection of teachers that is not only stress and workload. There is a loss of power and control, fear, anger

and disaffection (Perryman, 2007). The only positive emotion to an inspection is 'relaxation' which is experienced sometimes after the inspection, according to Tunç et al (2015). This contrasts with Quintelier et al (2019), who examined the teachers' emotions with regard to different moments during the school inspection visit in Belgium (Flanders). Emotions of joy were mostly reported. While emotions of anger and sadness were reported less frequently (Quintelier et al, 2019).

Chapman (2002) uses a case study approach in ten recently inspected secondary schools. Some teachers held the Ofsted inspection responsible for stress related illness (Chapman, 2002). This corresponds to Smulders (2004), which came to the conclusion that job stress plays a big role in the cause of sickness absence of teachers, for example teacher burnout. However, Bowers & McIver (2000) finds that an Ofsted inspection is not the main factor that causes sickness absence of teachers in the United Kingdom. Workload and paperwork were mostly associated with teachers' illness. Around 10% of the respondents associated bullying by schools' managers with teachers' illness and 13,8% mentioned Ofsted inspections that contributes to teachers' illness (Bowers & McIver, 2000).

According to Fitz-Gibbon & Stephensen-Forster (1999), the most stressed teachers are the most excellent teachers. If sickness absence of teachers is a sign of nervousness because of incompetence, the expectation is that the sickness absence of teachers is related to the rating of the school given by the inspector. Because the most stressed teachers are the most excellent teachers, there is no relationship between the amount of sickness absence of teachers and the rating of the school given by the inspector (Fitz-Gibbon & Stephensen-Forster, 1999).

3.2.6.1 Before the inspection

Prior research suggests that a lot of stress exists before an inspection. First of all, Chapman (2001) investigated the impact of an Ofsted inspection on classroom changes. The anxiety and stress of a teacher is the worst during the period before an inspection (Chapman, 2001). Second, Rosenthal (2004) did empirical research to the direct effect of the Ofsted system on the observed exam performance of secondary schools. There is a small but well-determined adverse, negative effect associated with the Ofsted inspection for the year of the inspection. Stress and preparation for the inspection might be causes for this negative effect (Rosenthal, 2004). Thirdly, Ferguson et al (2000) concluded that inspection creates high levels of stress. Primary school teachers

were interviewed. Stress of the preparation before an inspection is being worse than the inspection itself, according to the teachers (Ferguson et al, 2000).

Fitz-Gibbon & Stephenson-Forster (1999) asked to headteachers if there were any stress-related absences before the Ofsted inspection. The average of absences before inspection was 15.1 staff days (Fitz-Gibbon & Stephenson-Forster, 1999). When interviewing teachers, a number of teachers had become ill prior to the inspection (Troman & Woods, 2001).

Tunç et al (2015) use a qualitative research model to understand teacher emotions. 38 primary school teachers were interviewed in Turkey. The perception of teachers about inspections are negative. Emotions experienced before inspection are pressure, hurry, stress, concern, tension, anxiety, worry and uncertainty (Tunç et al, 2015). A notice of inspection generates anxiety and stress to those inspected, according to teachers. The anxiety is higher for schools with a long history of low attainment or other problems. A short notice or unannounced inspection should reduce the stress and workload. As a consequence, Ofsted has decreased the notice of an inspection in 2000 (Matthews & Sammons, 2004). MacBeath (2006) noted that a shorter notice of an Ofsted inspection avoid schools for unnecessary inspection preparation and lower levels of stress. Shorter notice should help inspectors to see schools as they really are (MacBeath, 2006).

However, after decreasing the notice of inspection by Ofsted, anxiety before the inspection remains a source of concern, but the benefits of an inspection outweigh the negative effects (Matthews & Sammons, 2004).

3.2.6.2 During the inspection

In general, when looking at empirical studies which investigates the effect of inspections, an inspection visit has a positive effect on the policy of the school and the behavior, but there are some situations where the effect on learning outcomes is negative. An example is stress during the inspection. Thus, an inspection visit seems to have a positive effect on the quality improvement of schools, even though it does not lead to higher grades and there is stress (Wolf & Janssens, 2005). The stress and energy for only a few hours of an Ofsted inspection seemed to be detrimental to the teachers and the pupils, according to Brown et al (2000). A more interactive inspection process would make teachers happier. The revised framework for Ofsted already has a more interactive inspection process (Brown, et al, 2002).

Fitz-Gibbon & Stephenson-Forster (1999) asked headteachers if there were any stress-related absences during the Ofsted inspection. The average of absences during inspection was 2.3 staff days (Fitz-Gibbon & Stephenson-Forster, 1999). When interviewing teachers, a number of teachers had become ill and were absent when the inspection took place (Troman & Woods, 2001).

Tunç et al (2015) use a qualitative research model to understand teacher emotions. The perception of teachers about an inspection is negative. Emotions experienced during inspection are stress, tension, anger, loss of strength, humiliation and lack of self-efficacy (Tunç et al, 2015). Quintelier et al (2018) did some qualitative research with interviews. When inspectors are inadequately informed, arrogant or disrespectful, respondents tend to reject feedback. When negative feedback is seen as unfair, negative emotions get in the way with feedback acceptance. A negative emotion might be anger or sadness. Teachers might be angry, because of an increase in stress and pressure (Quintelier et al, 2018).

Hopkins et al (2016) investigated the teachers' view on school evaluation and inspection using interviews. Techers are stressed when the inspector sits in the classroom and evaluates them. Thus, there are negative experiences of inspection. However, the teachers have found an approach that will reduce the stress and help them perform optimally. Not only the inspection should be taken into account, but also the school's own self-evaluation. Because, self-evaluation is conducted in a supportive way, while the impact of an inspection might be damaging (Hopkins et al, 2016).

3.2.6.3 After the inspection

Fitz-Gibbon & Stephenson-Forster (1999) asked to headteachers if there were any stress-related absences after the Ofsted inspection. The average of absences after inspection were 28 staff days (Fitz-Gibbon & Stephenson-Forster, 1999).

Tunç et al (2015) use a qualitative research model to understand teacher emotions. The perception of teachers about inspection are mostly negative. Emotions experienced after inspection are relaxation, stress, sorrow, anger, fury, meaninglessness and disappointment. The only positive emotion is 'relaxation' which is experienced sometimes after inspection (Tunç et al, 2015). Scanlon (1999) uses survey and interview data to investigate the impact of an Ofsted inspection. Headteachers are more positive about the inspection than the teaching staff. There

might be a post-Ofsted blues. This is for example the belief that the inspection was unfair and there are feelings of depression. Schools that are falling behind might get special measures. The period after an inspection at a special measure school might be traumatic because teachers feel shock, disappointment and disillusionment. Special measures might increase workload, which lead to stress. This indirectly implies that the school with the worst teachers will receive special measures and have the most stress (Scanlon, 1999). This is in contrast with Fitz-Gibbon & Stephensen-Forster (1999), which concluded that there is no relationship between the amount of sickness absence of teachers and the rating of a school given by the inspector. Because the most stressed teachers are the most excellent teachers, not the worst teachers (Fitz-Gibbon & Stephensen-Forster, 1999).

Overall, an inspection can cause more teacher stress and sickness absence of teachers, for example teacher burnout. This is the same relationship as the relationship between school accountability and sickness absence of teachers. This is expected because an inspection is a tool for accountability.
In the case of an inspection, research is inconclusive about the relationship between the amount of sickness absence of teachers and the rating of a school given by the inspector. Furthermore, there is conflicting research about the emotions regarding an inspection, for example stress. Moreover, research is inconclusive about whether an inspection generates more stress related illness of teachers.

3.2.7 School characteristics

School characteristics can have an influence on sickness absence of teachers.

First of all, connections between the staff are lacking in a large school. People do not know each other. This can lead to distrust and burnout (Dworkin & Tobe, 2014). In general, employees in larger schools have a higher sickness absence than employees in smaller schools (Smulders, 2004).

Second, a school that has a lot of pupils that speak a foreign language has a higher sickness absence of teachers. Because in general, people that speak a foreign language have a disadvantage (Roelandt et al, 1990). The more disadvantaged pupils, the higher the sickness absence of teachers (Dupont, 2007).

Thirdly, a school that is located in an impulse area has a higher sickness absence of teachers (Dupont, 2007). An impulse area is an area where there is high unemployment and there are low incomes. These factors play a role in the emergence of disadvantages in education (Education Executive Agency, 2019a). The more disadvantaged pupils, the higher the sickness absence of teachers (Dupont, 2007).

Fourth, a school with a lot of teachers that are older has a higher sickness absence of teachers. Because in general, older employees have a higher sickness absence than younger employees (Smulders, 2004; Dupont, 2007).

Fifth, a school with a lot of female teachers has a higher sickness absence of teachers. Because in general, females have a higher sickness absence than males (Smulders, 2004; Dupont, 2007). A reason for this is that female teachers experience more stress (Phillips et al, 2007).

Sixth, a school with a lot of teachers that work on a permanent basis has a higher sickness absence of teachers. Because in general, employees that work on a permanent basis have a higher sickness absence than employees that work on a temporary basis (Smulders, 2004).

3.3. School accountability and student absence

An unintended consequence of school accountability is selective assignment of students to special education programs. For example, classify pupils with low achievement as learning disabled. A reason for this selective assignment is that students in special education programs may be excluded from the accountability system. This gives the incentive to move low-performing students to the special education programs in order to look better at the performance measures. This gives less incentives to support the students in the special education programs (Figlio & Loeb, 2011).

Next to the reclassification of pupils, Figlio (2006) came to the conclusion that schools tend to assign more severe punishments to low-performing students than to high-performing students, for example suspension of students. This gap is substantially higher during a standardized test, which is used to evaluate schools in Florida. This implies that schools have the incentive to suspend the low-performing students during the test in order to look better at the performance measures (Figlio, 2006).

Thus, a school accountability system can cause more suspensions of low-performing students in order to look better at the performance measures.

3.4. Contribution to existing knowledge

School accountability is increasingly used in the world, and there is further research needed about the unintended consequences of a school accountability system. For example, the influence of school accountability on health outcomes (Figlio & Loeb, 2011). Inspection is a tool for accountability (Learmonth, 2000; Bowers & McIver, 2000).

Previous studies about the relationship between a school accountability system and sickness absence of teachers have almost exclusively focused on the United States, particularly in Houston (Texas) (e.g. Dworkin et al, 2003). However, there is also research conducted about this topic in Belgium and the United Kingdom (Van Droogenbroeck et al, 2014; Perryman & Calvert, 2020). Regarding inspections, previous studies about the relationship between inspections and sickness absence of teachers have almost exclusively focused on the United Kingdom (e.g. Troman & Woods, 2001). However, there is also research conducted about this topic in Dubai, Turkey and Belgium (Alkutich & Abukari, 2018; Tunç et al, 2015; Quintelier et al, 2019). Therefore, this research will be focused on a country outside the United States, Belgium, United Kingdom, Dubai and Turkey, namely in the Netherlands.

Most of the previous research used a survey as their research method, where teachers were asked how they perceive the school accountability system/inspection. This is focused on opinions and feelings. In this research, the sickness absence percentage will be used as health outcome. The Education Executive Agency gathers this percentage from the payroll administration and the governing body of the school. This is not just an opinion or a feeling (Education Executive Agency, 2020a).

A lot of previous research concluded that accountability leads to more teacher stress (e.g. Ryan et al, 2017). This corresponds to the problem in the United Kingdom, where teachers leave the profession within five years, whereby accountability is a crucial factor (Perryman & Calvert, 2020). However, this is in contrast with Sun et al (2017). Because Sun et al (2017) investigated that the school accountability system No Child Left Behind in the United States does not change the average national rate of teachers that transfer between schools or leave the teaching profession. While the expectation is that a school accountability system leads to more stress,

and more teachers leaving the teaching profession. Due to this inconsistency, this research will look at the unintended consequences of school accountability on the sickness absence of teachers, whereby inspection is used as a tool for accountability.

The only positive emotion to an inspection is 'relaxation' which is experienced sometimes after inspection, according to Tunç et al (2015). This contrasts with Quintelier et al (2019), who concluded that emotions of joy were mostly reported. While emotions of anger and sadness were reported less frequently (Quintelier et al, 2019). To clarify this inconsistency, this research will focus on the actual unintended consequences of an inspection.

Chapman (2002) concluded that some teachers held the Ofsted inspection responsible for stress related illness. However, Bowers & McIver (2000) investigated that only 13.8 % of the teachers mentioned Ofsted inspection as a contribution to their illness. This means that an inspection is not always associated with stress related illness. This raises questions about the actual stress related illness of teachers due to an inspection. Therefore, this research will focus on the unintended consequence of an inspection on sickness absence of teachers.

In general, an inspection visit seems to have a positive effect on the quality improvement of schools, even though it does not lead to higher grades and there is stress (Wolf & Janssens, 2005). Anxiety before the inspection remains a source of concern, but the benefits of an inspection outweigh the negative effects (Matthews & Sammons, 2004). However, Alkutich & Abukari (2018) concluded that a school inspection is effective and benefiting teaching and learning. However, there are concerns about some parts of the inspection which has a negative effect on the benefits. This raises questions about the actual benefits and negative effects of an inspection. Therefore, this research will focus on the negative effect of an inspection on the health of teachers.

A notice of inspection leads to stress to those inspected. The anxiety is higher in schools with a long history of low attainment or other problems (Matthews & Sammons, 2004). Special measures might increase workload, which generates stress. This indirectly implies that schools with the worst teachers will receive special measures and have the most stress (Scanlon, 1999). This contrasts with Fitz-Gibbon & Stephensen-Forster (1999), who concluded that there is no relationship between the amount of sickness absence of teachers and the rating of a school given by the inspector. Because the most stressed teachers are the most excellent teachers, not the

worst teachers (Fitz-Gibbon & Stephensen-Forster, 1999). Therefore, this research will focus on the relationship between a qualification of a primary school given by the Inspectorate of Education and sickness absence of teachers.

4. Data

This section explains the data that is used to perform the analysis, whereby the analysis focuses on the relationship between a qualification of a primary school given by the Inspectorate of Education and sickness absence of teachers.

In order to observe the sickness absence of teachers, this research uses data from the Education Executive Agency. In this research, the sickness absence percentage will be used in order to measure the sickness absence. The Education Executive Agency gathers this percentage from the payroll administration and the governing body of the school (Education Executive Agency, 2020a). There is data available about the sickness absence of the teachers and management per BRIN-number in the period 2016-2019, whereby every school has a particular BRIN-number (Education Executive Agency, 2019c; Education Executive Agency, 2020d; Education Executive Agency, 2020a). For simplicity, I will call this sickness absence of teachers.

In order to observe the qualifications of primary schools, this research uses data from the Inspectorate of Education. The qualifications given by the Inspectorate of Education per BRIN-number are available for the period 2012-2021. In most years, there is an update of the list of qualifications twice a year, whereby the 1st of September and the 1st of April are the reference dates each year (Inspectorate of Education, 2012-2021). It is important to keep in mind that supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b).

As just explained in paragraph 3.2.7, school characteristics can have an influence on sickness absence of teachers. To get rid of this selection bias, it is needed to control for these characteristics when investigating the relationship between the qualification of a primary school given by the Inspectorate of Education and sickness absence of teachers (Angrist & Pischke, 2015). The data source of all the control variables is the Education Executive Agency. The first variable is the number of pupils per school. There is data available per BRIN-number for the period 2010-2020 (Education Executive Agency, 2019d). The second variable is the percentage of pupils per school that speak a foreign language. There is data available per BRIN-number for the period 2010-2019 (Education Executive Agency, 2019e). There are 2 codes. Code 1 is used for the pupils that have a Dutch cultural background or the pupils that have parents from Suriname, Aruba or the Dutch Antilles. Code 2 is used for pupils that do not have a Dutch

cultural background (Education Executive Agency, 2020b). For every BRIN-number, there is data available about the number of pupils that have code 1 or code 2 (Education Executive Agency, 2019e). The third variable is a dummy variable which indicates whether the school is located in an impulse area. If the school is located in an impulse area, the dummy gets value 1. If the school is not located in an impulse area, the dummy gets value 0. There is data available per BRIN-number for the period 2014-2020 (Education Executive Agency, 2019f). The fourth variable is the average age of teachers per school. There is data available per BRIN-number for the period 2011-2019 (Education Executive Agency, 2020c). The fifth variable is the percentage of female teachers per school. There is data available per BRIN-number for the period 2011-2019 (Education Executive Agency, 2020c). The sixth variable is the average full-time equivalent (fte) per school. There is data available per BRIN-number for the period 2011-2019 (Education Executive Agency, 2020c).

Furthermore, for the years 2015-2019, there is data available about how a school scores on particular indicators/standards. These scores on indicators/standards are used by the Inspectorate of Education to determine the qualification of a school (Inspectorate of Education, 2015-2019). It is important to keep in mind that supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b).

5. Methodology

This section explains the empirical methodology. Due to the fact that there is only data available about the sickness absence of the teachers per BRIN-number in the period 2016-2019, this research will focus on the relationship between the qualification of a primary school given by the Inspectorate of Education and the sickness absence of teachers in the period 2016-2019.

5.1. Variables Ordinary Least Squares and Fixed Effects

The dependent variable is sickness absence of teachers of a particular school i in year t (Y_{it}). In this research, the sickness absence percentage of teachers per BRIN-number will be used in order to measure the sickness absence (Education Executive Agency, 2019c; Education Executive Agency, 2020d; Education Executive Agency, 2020a). The sickness absence percentage indicates which part of the working time is lost in a particular year because of absence of the teachers. This percentage shows the relationship between the amount of absence and the working capacity (Education Executive Agency, 2020e). In some cases, there are more schools that use the same BRIN-number. For example, when two schools belong to the same association, but have different locations. Due to the fact that there is only data available about the sickness absence percentage per BRIN-number, the schools that use the same BRIN-number are excluded from the analysis.

The independent variable is the qualification of a school i given by the Inspectorate of Education in year t (Q_{it}). In most years, there is an update of the list of qualifications twice a year, whereby the 1st of September and the 1st of April are the reference dates each year (Inspectorate of Education, 2012-2021). A reference date is a date at which the qualifications are retrieved from the database (Inspectorate of Education, 2020e). When using the 1st of April in a particular year as a reference date, the qualifications are given by the Inspectorate of Education between the 1st of September the year before and the 1st of April in that particular year. To investigate the relationship between a qualification of a school given by the Inspectorate of Education and the sickness absence of teachers in a particular year, the focus is on the school qualification given by the Inspectorate of Education on the 1st of April in that particular year (Inspectorate of Education, 2016; Inspectorate of Education, 2017a; Inspectorate of Education, 2018a; Inspectorate of Education, 2019b). Because stress is experienced before, during and after the inspection, which might cause more sickness absence of teachers (e.g. Tunç et al, 2015). Another important aspect is that parents who have to make a school choice for their

children will look at the list of qualifications on the 1st of April to decide the school for their children in September. Therefore, it might be the case that the list of qualifications published on the 1st of April is more important than the list of qualifications published on the 1st of September. This might create more stress and will lead to more sickness absence of teachers. Table 1 clarifies the dates chosen in this research, in particular the timing of the qualification, whereby a particular year is called year X and the year before is called year X-1.

Table 1. Timing of the qualification (1st of April, year X)

Variables	**Period**
Qualification with reference date 1st of April in year X	From the 1st of September in year X-1 until the 1st of April in year X
Sickness absence of teachers in year X	From the 1st of January in year X until the 31st of December in year X

Figure 2 clarifies the relationship between the qualification given by the Inspectorate of Education and sickness absence of teachers.

Figure 2. Relationship between the qualification given by the Inspectorate of Education with a reference date in year X and sickness absence of teachers in year X

It is important to keep in mind that supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b). There are three qualifications possible in the supervision framework, namely 'basis', 'weak' and 'very weak' (Inspectorate of Education, 2012). This framework is used for the years 2016 and 2017. In the dataset, 'basis' will receive the code 2, 'weak' will receive the code 3 and 'very weak' will receive the code 4. There are four qualifications possible in the inspection framework, namely 'good', 'sufficient', 'insufficient' and 'very weak' (Inspectorate of Education, 2020c). This framework is used for the years 2018 and 2019. In the dataset, 'good' will receive the code 1, 'sufficient' will receive the code 2, 'insufficient' will receive the code 3 and 'very weak' will receive the code 4.

Even though, 'insufficient' in the new inspection framework is not the same as 'weak' in the old supervision framework and 'sufficient' in the new inspection framework is not the same as

'basis' in the old supervision framework, the code is the same, because in my opinion a school that receives the qualification 'basis' in the old supervision framework and 'insufficient' in the new inspection framework, will feel this as a deterioration in the qualification (Inspectorate of Education, 2020d).

As just explained in paragraph 3.2.7, school characteristics can have an influence on sickness absence of teachers. To get rid of this selection bias, it is needed to control for these characteristics (Angrist & Pischke, 2015). Therefore, the control variables are the number of pupils of a school i in year t ($Pupils_{it}$), percentage of pupils of a school i in year t that speak a foreign language ($Foreign_{it}$), average age of teachers of a school i in year t (Age_{it}), percentage female teachers of a school i in year t ($Gender_{it}$), the average full-time equivalent (fte) of a school i in year t (FTE_{it}) and whether a school i is located in an impulse area in year t ($Area_{it}$) (Education Executive Agency, 2019g; Education Executive Agency, 2019h; Education Executive Agency, 2019i; Education Executive Agency, 2019m; Education Executive Agency, 2019e; Education Executive Agency, 2020c; Education Executive Agency, 2019j, Education Executive Agency, 2019k; Education Executive Agency, 2019l; Education Executive Agency, 2019b). If the school is located in an impulse area, the dummy gets value 1. If the school is not located in an impulse area, the dummy gets value 0 (Education Executive Agency, 2019j, Education Executive Agency, 2019k; Education Executive Agency, 2019l; Education Executive Agency, 2019b).

The reference date for the number of pupils is the 1st of October the year before. This implies that the 1st of October in 2015 is used for the number of pupils in 2016 (Education Executive Agency, 2019g). Consequently, the 1st of October in 2016 is used for the number of pupils in 2017 (Education Executive Agency, 2019h). Similarly, the 1st of October in 2017 is used for the number of pupils in 2018 (Education Executive Agency, 2019i). Also, the 1st of October of 2018 is used for the number of pupils in 2019 (Education Executive Agency, 2019m). This is also the case for whether the school is located in an impulse area (Education Executive Agency, 2019j, Education Executive Agency, 2019k; Education Executive Agency, 2019l; Education Executive Agency, 2019b).

5.2. Ordinary Least Squares (OLS)

$$Y_{it} = \beta_0 + \beta_1 Q_{it} + \beta_2 Pupils_{it} + \beta_3 Foreign_{it} + \beta_4 Age_{it} + \beta_5 Gender_{it} + \beta_6 FTE_{it} + \beta_7 Area_{it} + \varepsilon_{it} \quad (1)$$

Y_{it} = sickness absence of teachers of a school i in year t

β_0 = intercept

Q_{it} = qualification of a school i given by the Inspectorate of Education in year t

$Pupils_{it}$ = number of pupils of a school i in year t

$Foreign_{it}$ = percentage pupils that speak a foreign language of a school i in year t

Age_{it} = average age of teachers of a school i in year t

$Gender_{it}$ = percentage female teachers of a school i in year t

FTE_{it} = average full-time equivalent (fte) of a school i in year t

$Area_{it}$ = whether a school i is located in an impulse area in year t

ε_{it} = error term of a school i in year t

β_1 is the coefficient of interest, because the research is focused on the relationship between a qualification of a school given by the Inspectorate of Education and sickness absence of teachers.

Since there are observations of the same schools in different years, the standard errors will be clustered at school level.

The main assumption with OLS is ceteris paribus, meaning that schools that differ in the qualification given by the Inspectorate of Education (Q_{it}) are similar on unobserved factors (ε_{it}), because unobserved factors (ε_{it}) can influence the sickness absence (Y_{it}). When regressing equation 1, there might be a failure to include enough controls or the right controls. Because for example, some confounding factors are unobserved, and it is impossible to control for everything. This problem is called omitted variable bias. Therefore, there is no believe that the assumption will hold (Angrist & Pischke, 2015; Angrist & Pischke, 2009).

A solution is a fixed effects strategy, which is a strategy that controls for unobserved but fixed omitted variables. Thus, this strategy controls for time invariant variables (Angrist & Pischke, 2009). The fixed effects strategy will be explained in the next paragraph.

5.3. Fixed Effects

$$Y_{it} = \beta_0 + \beta_1 Q_{it} + \beta_2 Pupils_{it} + \beta_3 Foreign_{it} + \beta_4 Age_{it} + \beta_5 Gender_{it} + \beta_6 FTE_{it} + \beta_7 Area_{it} + f_i + f_t + \varepsilon_{it} \quad (2)$$

The variables used in equation 2 are the same as the variables used in equation 1. The difference between these two equations is that equation 2 includes a school i fixed effect (f_i) and a year t fixed effect (f_t). A school i fixed effect considers unobserved but fixed (time invariant) characteristics of a school. A year t fixed effect considers shocks that are common to various schools at the same point in time. An example of a year t fixed effect in this research is the fact that the supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b).

β_1 is the coefficient of interest, because the research is focused on the relationship between a qualification of a school given by the Inspectorate of Education and sickness absence of teachers.

Since there are observations of the same schools in different years, the standard errors will be clustered at school level. This equation will first be estimated with only the school i fixed effect (f_i). After that, the whole equation with school i fixed effect (f_i) and year t fixed effect (f_t) will be estimated.

With a fixed effects strategy, the fixed effects only consider the unobserved time invariant variables. Selection into the qualification of schools (Q_{it}) should be based on unobserved but fixed characteristics. This implies that the assumption of fixed effects is that the omitted variables are time invariant. Meaning that the unobserved factors (ε_{it}) that might affect sickness absence (Y_{it}) and the qualification of schools (Q_{it}) are time invariant. For many causal questions, the notion that the most important omitted variables are time invariant is not plausible. This means that controlling for time variant variables is important to prevent omitted variables bias. As all the control variables mentioned in paragraph 5.1 are time variant, all these variables have to be taken into account as control variables. When regressing equation 2, there might be a failure to include all the time variant variables. Because for example, some time variant variables are unobserved, and it is impossible to control for everything. This problem is called omitted variable bias. However, the omitted variable bias is smaller in the case of a fixed effects strategy compared with OLS, because there are no omitted time invariant variables with fixed effects. Next to the possible omitted variable bias, fixed effects estimates might be subject to

attenuation bias from measurement error. First of all, the qualification of a school might be persistent, meaning that a school who receives the qualification 'sufficient' this year is most likely to receive the qualification 'sufficient' next year. Second, measurement error often changes from year to year, meaning that the qualification of a school may be misreported this year but not next year. This implies that while the qualification of a school may be misreported for only a few schools in any single year, the observed changes in qualifications of schools from year to year may be mostly noise. Another problem with fixed effects is that controlling for fixed effects might remove good and bad variation, meaning that fixed effects remove not only the potential omitted variable bias, but also some useful information about the variable of interest (Angrist & Pischke, 2015; Angrist & Pischke, 2009).

5.4. Heterogeneity analyses

5.4.1 Small versus large schools

As discussed in paragraph 3.2.7, school characteristics can have an influence on sickness absence of teachers. For example, teachers in larger schools might have a higher sickness absence than teachers in smaller schools (Smulders, 2004). Because connections between the staff are lacking in a large school. This can lead to distrust and burnout (Dworkin & Tobe, 2014). Therefore, in this research the sickness absence of teachers in large schools will be compared with sickness absence of teachers in small schools. The dataset that is used for OLS and Fixed Effects will be sorted according to size of the schools (mean of the four years) and divided into two samples. The schools that have no value for the mean size of the school will be excluded from the analysis. The first sample contains of the large schools and the second sample contains of the small schools. For both samples, OLS and Fixed Effects will be conducted. Since there are observations of the same schools in different years, the standard errors will be clustered at school level.

5.4.2 Schools with qualification from 'basis'/'sufficient' to 'weak'/'insufficient'

A school that receives the qualification 'basis'/'sufficient' in the first year and 'weak'/'insufficient' in the second year, will have a deterioration in the qualification. The expectation is that this deterioration is much worse than a deterioration from 'good' to 'sufficient' for example. Therefore, in this research the sickness absence of teachers in schools that have a deterioration in the qualification from 'basis'/'sufficient' to 'weak'/'insufficient' will be investigated. The dataset that is used for OLS and Fixed Effects will be used for the schools that have a deterioration in the qualification from 'basis'/'sufficient' to

‘weak’/’insufficient’. For this sample, OLS and Fixed Effects will be conducted. Since there are observations of the same schools in different years, the standard errors will be clustered at school level.

When investigating the relationship between a qualification of a school given by the Inspectorate of Education and the sickness absence of teachers in a particular year for the whole sample with OLS and Fixed Effects, the focus is on the school qualification given by the Inspectorate of Education on the 1st of April in that particular year (Inspectorate of Education, 2016; Inspectorate of Education, 2017a; Inspectorate of Education, 2018a; Inspectorate of Education, 2019b). Therefore, this heterogeneity analysis focuses only on the difference in qualifications between the 1st of April each year. Even though, there is an update of the list of qualifications on the 1st of September (Inspectorate of Education, 2012-2021). Meaning that only the schools that have a deterioration from ‘basis’/’sufficient’ on the 1st of April in the first year to ‘weak’/’insufficient’ on the 1st of April in the second year are included in the analysis.

5.5. Regression Discontinuity Design

In this research, the sickness absence percentage of teachers per BRIN-number will be used in order to measure the sickness absence (Education Executive Agency, 2019c; Education Executive Agency, 2020d; Education Executive Agency, 2020a). The sickness absence percentage indicates which part of the working time is lost in a particular year because of absence of teachers. This percentage shows the relationship between the amount of absence and the working capacity (Education Executive Agency, 2020e). In some cases, there are more schools that use the same BRIN-number. For example, when two schools belong to the same association, but have different locations. Due to the fact that there is only data available about the sickness absence percentage per BRIN-number, the schools that use the same BRIN-number are excluded from the analysis.

Before the 1st of August in 2017, there was a supervision framework that determines the qualifications of the schools (Inspectorate of Education, 2020b). In general, a primary school will receive the qualification ‘very weak’ when the school scores insufficient on the domain ‘outcomes’ and two or more norm indicators of the ‘teaching-learning process’ and ‘special needs provision and guidance’ are insufficient (Inspectorate of Education, 2012). Due to data constraints, it is not possible to use a Regression Discontinuity Design to estimate the effect of the qualification ‘very weak’ before the 1st of August in 2017 on sickness absence of teachers.

Since the 1st of August in 2017, there is an inspection framework that determines the qualifications of the schools (Inspectorate of Education, 2020b). In general, a primary school will receive the qualification 'very weak', if the standard 'results' is 'insufficient' and at least one of the standards of 'development perspectives' or 'didactics' or 'safety' is 'insufficient' (Inspectorate of Education, 2020c). Therefore, only the schools that score 'insufficient' on 'results' will be included in this analysis. This new inspection framework came into effect in the schoolyear 2017-2018. This is the first period where there is data available about how a school scores on particular standards of the new inspection framework (Inspectorate of Education, 2019a). A Regression Discontinuity Design will be used to investigate the effect of the qualification 'very weak' of a school on the sickness absence of teachers. The running variable is the scores on standards, which determine the qualification of a school. The datafile about how a school scores on the standards in a particular schoolyear represents reports in the period from the 1st of September until the 31st of August (Inspectorate of Education, 2019e).

To investigate the relationship between the qualification 'very weak' given by the Inspectorate of Education and sickness absence of teachers in 2018, the focus is on how a particular school scores on the standards in the schoolyear 2017-2018 (Inspectorate of Education, 2019a; Education Executive Agency, 2020d). Similarly, to investigate the relationship between the qualification 'very weak' given by the Inspectorate of Education and sickness absence of teachers in 2019, the focus is on how a particular school scores on the standards in the schoolyear 2018-2019 (Inspectorate of Education 2019c; Education Executive Agency, 2020d).

$$Y_{it} = \beta_0 + \beta_1 D_{it} + \beta_2 C_{it} + \varepsilon_{it} \quad (3)$$

Y_{it} = sickness absence of teachers of a school i in year t

D_{it} = the qualification of a school i given by the Inspectorate of Education in year t (= 1 if the qualification is 'very weak', = 0 otherwise)

C_{it} = scores of a school i in year t on standards ('development perspectives', 'didactics', 'safety'). The school receives code 0 if the schools scores on none of the other standards ('development perspectives', 'didactics', 'safety') 'insufficient'. The school receives code 1 if the school scores on one of the other standards ('development perspectives', 'didactics', 'safety') 'insufficient'. The school receives code 2 if the schools scores on two of the other standards ('development perspectives', 'didactics', 'safety') 'insufficient'. The school receives

code 3 if the school scores on all the other standards ('development perspectives', 'didactics', 'safety') 'insufficient'.

ε_{it} = error term of a school i in year t

Cutoff = when the school receives the qualification 'very weak'. This is the case when the school scores 'insufficient' on 'results' and at least one of the standards of 'development perspectives' or 'didactics' or 'safety' is 'insufficient' (Inspectorate of Education, 2020c). This is the case when a school receives code 1, 2 or 3 on C_{it}.

β_1 is the coefficient of interest, because this Regression Discontinuity Design is focused on the relationship between the qualification 'very weak' of a school given by the Inspectorate of Education and sickness absence of teachers. Robust standard errors will be used.

The scores of a school on standards (C_{it}) determine whether the qualification of a school is 'very weak' (D_{it}). This implies that the scores of a school on standards (C_{it}) is the running variable. In this case, the running variable is discrete, meaning that the set of values that the running variable can take is countable. The discrete running variable can only take a finite number of values, namely 0, 1, 2 and 3. These points are called the mass points, which are values that are shared by many observations. The fact that the running variable is discrete instead of continuous has consequences for the assumptions of a Regression Discontinuity Design. With a continuous running variable, the continuity-based approach is used. While with a discrete running variable, the local randomization approach is used when there are only a few mass points. This is the case in this research because there are only four mass points. The local randomization approach is based upon two assumptions. The first assumption implies that treatment is assigned as good as randomly close to the cutoff, meaning that the Regression Discontinuity Design behaves like a randomized experiment close to the cutoff. In this research, the schools that scores 0 and 1 on the standards (C_{it}) are comparable to a randomized experiment, where the qualification 'very weak' of a school (D_{it}) is assigned as good as randomly. The second assumption implies that close to the cutoff the potential outcomes depend on the running variable solely through the treatment, but not directly. This means that the sickness absence of teachers (Y_{it}) of schools that score 0 and 1 on the standards depends on the scores of a school on standards (C_{it}) solely through the qualification 'very weak' of a school

(D_{it}). Thus, sickness absence of teachers (Y_{it}) is not directly related to the scores of a school on standards (C_{it}) for schools that score 0 and 1 on the standards. Far away from the cutoff, the sickness absence of teachers (Y_{it}) can be directly related to the scores of a school on the standards (C_{it}) (Cattaneo et al, 2017).

It is unfeasible to estimate equation 3 when adding a school i fixed effect (f_i) to this equation, because there are too little observations.

6. Results

This section presents the results of the analyses that are described in chapter 5.

6.1. OLS and Fixed Effects

Table 2 shows the summary statistics per year of the variables used when conducting OLS and Fixed Effects. The mean of sickness absence of teachers is approximately 6% every year, namely 6.3% in 2016, 6.1% in 2017, 5.4% in 2018 and 5.6% in 2019. The mean of the qualification given by the Inspectorate of Education is 2.0 every year in the period 2016-2019. This indicates that the mean of the qualification is 'basis'/'sufficient'. When looking at the other variables, there are no big differences per year. For example, the age of teachers is approximately 42 year every year, namely 43.3 in 2016, 41.8 in 2017, 42.5 in 2018 and 41.3 in 2019. The number of schools per year are also approximately the same.

Table 2. Summary statistics OLS and Fixed Effects per year

	2016			2017			2018			2019		
Variables	(1) **Mean**	(2) **Standard Deviation**	(3) **Number of schools**	(4) **Mean**	(5) **Standard Deviation**	(6) **Number of schools**	(7) **Mean**	(8) **Standard Deviation**	(9) **Number of schools**	(10) **Mean**	(11) **Standard Deviation**	(12) **Number of schools**
Sickness absence of teachers (%)	6.3	4.8	5,893	6.1	4.6	5,986	5.4	35.2	5,926	5.6	59.8	5,885
Qualification (1-4)	2.0	0.1	5,878	2.0	0.2	5,800	2.0	0.2	5,767	2.0	0.3	4,664
Number of pupils	221.4	138.3	6,280	221.9	138.9	6,197	222.8	139.5	6,117	223.8	140.0	6,055
Pupils that speak a foreign language (%)	14.2	19.1	5,231	14.5	19.0	5,264	14.6	18.7	5,308	14.8	18.4	5,313
Age of teachers	43.3	7.5	5,315	41.8	10.3	5,257	42.5	7.7	5,280	41.3	10.0	5,297
Female teachers (%)	83.8	8.3	5,198	84.3	8.1	4,992	84.5	8.0	5,149	89.7	48.3	4,649
Full-time equivalent	0.7	0.1	6,275	0.6	0.2	6,275	0.6	0.1	6,275	0.6	0.16	6,275
Impulse area	0.2	0.4	6,280	0.2	0.4	6,166	0.2	0.4	6,117	0.2	0.4	6,055

Notes. This table presents summary statistics per year. The sickness absence percentage indicates which part of the working time is lost in a particular year because of absence of the teachers. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to one number after the decimal point.

Table 3 shows the summary statistics per qualification of the variables used when conducting OLS and Fixed Effects. The mean of sickness absence of teachers is 5.1% for the schools with qualification 1, 5.7% for schools with qualification 2, 8.3% for schools with qualification 3 and 9.3% for schools with qualification 4. When looking at the other variables, there are no big differences per qualification. For example, the full-time equivalent is around the 0.6-0.7 for every qualification, namely 0.7 with qualification 1, 0.6 with qualification 2, 0.6 with qualification 3 and 0.6 with qualification 4. However, the number of pupils differ somewhat per qualification and the number of observations per qualification differ. It is important to keep in mind that when conducting OLS and Fixed Effects the schools with different qualifications will be compared.

Table 3. Summary statistics OLS and Fixed Effects per qualification

	Qualification 1 ('good')			**Qualification 2 ('basis'/'sufficient)**			**Qualification 3 ('weak'/'insufficient')**			**Qualification 4 ('very weak')**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables	**Mean**	**Standard Deviation**	**Obser-vations**	**Mean**	**Standard Deviation**	**Obser-vations**	**Mean**	**Standard Deviation**	**Obser-vations**	**Mean**	**Standard Deviation**	**Obser-vations**
Sickness absence of teachers (%)	5.1	4.0	104	5.7	19.8	19,432	8.3	8.7	292	9.3	7.3	72
Number of pupils	281.9	180.0	108	214.3	130.9	20,649	170.7	102.1	310	139.4	71.2	76
Pupils that speak a foreign language (%)	15.4	18.3	103	13.2	17.7	17,416	14.1	15.1	238	17.7	21.8	59
Age of teachers	41.8	3.3	91	43.1	6.8	16,837	41.0	10.3	238	40.5	12.1	54
Female teachers (%)	94.1	22.6	77	84.8	17.6	16,268	87.9	30.0	218	81.1	19.4	46
Full-time equivalent	0.7	0.1	106	0.6	0.1	20,218	0.6	0.2	307	0.6	0.2	75
Impulse area	0.2	0.4	108	0.2	0.4	20,644	0.3	0.5	310	0.3	0.5	76

Notes. This table presents summary statistics per qualification for the period 2016-2019. The sickness absence percentage indicates which part of the working time is lost in a particular year because of absence of teachers. Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to one number after the decimal point.

Table 4 shows the relationship between the qualification of a school and sickness absence of teachers. When conducting OLS, the coefficient of interest is 2.056. This implies that when there is a deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak', the sickness absence of teachers will increase with 2.056 percentage point. This is statistically significant at an 1% level. When including school fixed effects, the coefficient of interest is 1.248. This implies that when there is a deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak', the sickness absence of teachers will increase with 1.248 percentage point. This is statistically significant at an 1% level. When including school fixed effects and year fixed effects, the coefficient of interest is 1.159. This implies that when there is a deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak', the sickness absence of teachers will increase with 1.159 percentage point. This is statistically significant at an 1% level. This estimate is most credible, because the omitted variable bias is smaller when including fixed effects compared with OLS.
Thus, when including fixed effects, the influence of the qualification of a school on sickness absence of teachers will decline. But it is all still statistically significant at an 1% level.

The mean of sickness absence of teachers is approximately 6%, according to Table 2. When the sickness absence of teachers increases with 1.159 percentage point (according to Table 4) when there is a deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak', the sickness absence of teachers will increase from for 6% to 7.159%. This is an increase of approximately 19.3%.

Table 4. Relationship between the qualification of a school and sickness absence of teachers

Variables	(1) OLS	(2) Fixed Effects	(3) Fixed Effects
Qualification (1-4)	2.056***	1.248***	1.159***
	(0.411)	(0.331)	(0.346)
Number of pupils	0.000	0.000	-0.002
	(0.001)	(0.008)	(0.009)
Pupils that speak a foreign language (%)	0.029***	0.089	0.115*
	(0.006)	(0.070)	(0.069)
Age of teachers	0.134***	0.071	0.033
	(0.032)	(0.075)	(0.079)
Female teachers (%)	-0.001	0.009**	0.015***
	(0.004)	(0.005)	(0.005)
Full-time equivalent	-5.178***	-6.235*	-5.876*
	(1.994)	(3.396)	(3.457)
Impulse area	0.884*	-1.376	-1.683
	(0.505)	(1.331)	(1.433)
Constant	-1.105	2.623	4.479
	(0.976)	(2.854)	(3.321)
Observations	13,836	13,836	13,836
R-squared	0.002	0.000	0.001
Number of schools	4,602	4,602	4,602
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

6.2. Heterogeneity analyses

6.2.1 Small versus large schools

Table A1 (appendix) shows the relationship between the qualification of a school and sickness absence of teachers for small schools. A deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak' has a statistically significant effect on sickness absence of teachers in small schools. The estimated effect is somewhat larger than the main estimate in Table 4, but the significance level declines when including fixed effects.

Table A2 (appendix) shows the relationship between the qualification of a school and sickness absence of teachers for large schools. A deterioration in the qualification from 'good' to

‘basis’/’sufficient’ or from ‘basis’/’sufficient’ to ‘weak’/’insufficient’ or from ‘weak’/’insufficient’ to ‘very weak’ has a statistically significant effect on sickness absence of teachers in large schools. The estimated effect is somewhat smaller than the main estimate in Table 4, but the significance level declines when including school fixed effects and year fixed effects.

6.2.2 Schools with qualification from ‘basis’/’sufficient’ to ‘weak’/’insufficient’

Table A3 (appendix) shows the relationship between the qualification of a school and sickness absence of teachers for schools who have a deterioration in the qualification from ‘basis’/’sufficient’ to ‘weak’/’insufficient’. A deterioration in the qualification from ‘basis’/’sufficient’ to ‘weak’/’insufficient’ has only a statistically significant effect on sickness absence of teachers in large schools when fixed effects are included. This estimate is most credible, because the omitted variable bias is smaller when including fixed effects compared with OLS. The estimated effect is quite similar to the main estimate in Table 4, but the significance level declines when including school fixed effects and year fixed effects.

6.3. Regression Discontinuity Design

Table 5 shows the summary statistics per score on the standards of the variables used when conducting a Regression Discontinuity Design. The mean of sickness absence of teachers is 7.1% for schools with score 0, 8.3% for schools with score 1 and 8.5% for schools with score 2. There are no schools that receive score 3. The qualification ‘very weak’ differs, but this makes sense, because the scores on standards determine the qualification. When looking at the other variables, there are no big differences per score on standards. For example, the number of pupils is approximately 177 for every score on standards, namely 171.5 for schools with score 0, 182.3 for schools with score 1 and 176.9 for schools with score 2. However, the number of observations per qualification differ. It is important to keep in mind that with a Regression Discontinuity Design left (score = 0) and right (score = 1) from the cutoff will be compared.

Table 5. Summary statistics Regression Discontinuity Design per score on standards

	Score 0			Score 1			Score 2		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables	**Mean**	**Standard Deviation**	**Obser-vations**	**Mean**	**Standard Deviation**	**Obser-vations**	**Mean**	**Standard deviation**	**Obser-vations**
Sickness absence of teachers (%)	7.1	5.2	94	8.3	6.5	12	8.5	5.9	48
Qualification 'very weak' (0-1)	0	0	96	1	0	12	1	0	52
Number of pupils	171.5	111.2	95	182.3	96.0	12	176.9	106.0	50
Pupils that speak a foreign language (%)	13.7	14.2	76	13.9	10.5	11	17.5	18.3	41
Age of teachers	41.3	8.1	69	38.9	13.9	11	42.6	7.7	40
Female teachers (%)	91.9	21.5	62	80.1	31.0	11	86.8	24.8	33
Full-time equivalent	0.6	0.1	94	0.6	0.2	12	0.6	0.1	51
Impulse area	0.3	0.5	95	0.4	0.5	12	0.3	0.5	50

Notes. This table presents summary statistics per score for the period 2018-2019. The scores are on three other standards: 'development perspectives', 'didactics' and 'safety'. The school receives code 0 if the school scores on none of the other standards 'insufficient'. The school receives code 1 if the school scores on one of the other standards 'insufficient'. The school receives code 2 if the schools scores on two of the other standards 'insufficient'. The school receives code 3 if the school scores on all the other standards 'insufficient'. There are no schools that receive code 3. The sickness absence percentage indicates which part of the working time is lost in a particular year because of absence of teachers. The qualification 'very weak' is a dummy variable, which gets the value 1 if the qualification is 'very weak', and 0 otherwise. Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to one number after the decimal point.

Table 6 shows the relationship between the qualification ‘very weak’ and the sickness absence of teachers in 2018 and 2019. When conducting a Regression Discontinuity Design, the coefficient of interest is 1.187. This implies that the sickness absence of teachers is 1.187 percentage point higher in the case a school scores ‘insufficient’ on ‘results’ and scores on one of the other standards (‘development perspectives’, ‘didactics’, ‘safety’) ‘insufficient’ in comparison to schools that score ‘insufficient’ on ‘results’ and ‘insufficient’ on none of other standards (‘development perspectives’, ‘didactics’ and ‘safety’). Meaning that schools that receive the qualification ‘very weak’ have a sickness absence of teachers that is 1.187 percentage point higher than schools that do not receive the qualification ‘very weak’. This is not statistically significant, which implies that the qualification ‘very weak’ has not a statistically significant effect on sickness absence of teachers.

Table 6. Relationship between the qualification ‘very weak’ and the sickness absence of teachers in 2018 and 2019

	(1)
Variables	**Regression Discontinuity Design**
Qualification (=1 if the qualification is ‘very weak’, =0 otherwise)	1.187
	(1.889)
Scores on other standards	-
Constant	7.120***
	(0.543)
Observations	106
R-squared	0.005
Number of schools	
School FE	

Notes. Robust standard errors in parentheses. The qualification ‘very weak’ is a dummy variable, which gets the value 1 if the qualification is ‘very weak’, and 0 otherwise. The scores are on three other standards: ‘development perspectives’, ‘didactics’ and ‘safety’. The school receives code 0 if the school scores on none of the other standards ‘insufficient’. The school receives code 1 if the school scores on one of the other standards ‘insufficient’. The school receives code 2 if the schools scores on two of the other standards ‘insufficient’. The school receives code 3 if the school scores on all the other standards ‘insufficient’. Scores on standards is omitted because of collinearity. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

7. Sensitivity analyses

This section presents the results of several robustness checks that are performed to assess the sensitivity of the results.

7.1. OLS and Fixed Effects: Supervision framework and inspection framework

Supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b). The supervision framework is used for the years 2016 and 2017 and the inspection framework is used for the years 2018 and 2019. As discussed in paragraph 5.1, 'insufficient' in the new inspection framework is not the same as 'weak' in the old supervision framework and 'sufficient' in the new inspection framework is not the same as 'basis' in the old supervision framework (Inspectorate of Education, 2020d). However, when investigating the relationship between a qualification of a school given by the Inspectorate of Education and the sickness absence of teachers in the period 2016-2019, 'insufficient' and 'weak' is considered to be the same and 'sufficient' and 'basis' is considered to be the same. As described in paragraph 5.3, the year fixed effect considers the fact that supervision by the Inspectorate of Education has been changed since the 1st of August in 2017.

To make sure that the change in supervision has no impact in this research, the relationship between a qualification of a school given by the Inspectorate of Education and the sickness absence of teachers will be investigated for two periods, namely 2016-2017 (supervision framework) and 2018-2019 (inspection framework). The dataset that is used for OLS and Fixed Effects will be used. This dataset will be split into two datasets. One dataset for the period 2016-2017 and one dataset for the period 2018-2019. For both datasets, OLS and Fixed Effects will be conducted. Since there are observations of the same schools in different years, the standard errors will be clustered at school level.

Table A4 (appendix) shows the relationship between the qualification of a school and sickness absence of teachers in 2016-2017. A deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak' has a statistically significant effect on sickness absence of teachers in 2016-2017. The estimated effect is larger than the main estimate in Table 4.

Table A5 (appendix) shows the relationship between the qualification of a school and sickness absence of teachers in 2018-2019. A deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak' has not a statistically significant effect on sickness absence of teachers in 2018-2019 when including fixed effects. This estimate is most credible, because the omitted variable bias is smaller when including fixed effects compared with OLS.

7.2. OLS and Fixed Effects: Timing of the qualification

To investigate the relationship between a qualification of a school given by the Inspectorate of Education and sickness absence of teachers in a particular year, the focus is on the school qualification given by the Inspectorate of Education on the 1st of April in that particular year (Inspectorate of Education, 2016; Inspectorate of Education, 2017a; Inspectorate of Education, 2018a; Inspectorate of Education, 2019b). Because stress is experienced before, during and after the inspection, which might cause more sickness absence of teachers (e.g. Tunç et al, 2015). Another important aspect is that parents who have to make a school choice for their children will look at the list of qualifications on the 1st of April to decide the school for their children in September. Therefore, it might be the case that the list of qualifications published on the 1st of April is more important than the list of qualifications published on the 1st of September. This might create more stress and will lead to more sickness absence of teachers. In most years, there is an update of the list of qualifications twice a year, whereby the 1st of September and the 1st of April are the reference dates each year (Inspectorate of Education, 2012-2021). A reference date is a date at which the qualifications are retrieved from the database (Inspectorate of Education, 2020e). This implies for example that a school receives a qualification between the 1st of September and the 1st of April whereby the reference date is the 1st of April.

Due to the fact that stress is experienced before the inspection and that the 1st of April in a particular year is a reference date, it might be the case that the qualification with the 1st of April as a reference date in a particular year has an influence on the sickness absence of teachers the year before (e.g. Tunç et al, 2015; Inspectorate of Education, 2020e). Therefore, to investigate the relationship between a qualification of a school given by the Inspectorate of Education and the sickness absence of teachers, the focus is on sickness absence of teachers measured the year before the school qualification is given by the Inspectorate of Education on the 1st of April in a particular year (Inspectorate of Education, 2017a; Inspectorate of Education, 2018a;

Inspectorate of Education, 2019b, Inspectorate of Education, 2020f). Table 7 clarifies the dates chosen in this research, in particular the timing of the qualification, whereby a particular year is called year X and the year before is called year X-1.

Table 7. Timing of the qualification (1st of April, year X-1)

Variables	**Period**
Qualification with reference date 1st of April in year X	From the 1st of September in year X-1 until the 1st of April in year X
Sickness absence of teachers in year X-1	From the 1st of January in year X-1 until the 31st of December in year X-1

Figure 3 clarifies the relationship between the qualification given by the Inspectorate of Education and the sickness absence of teachers, whereby the timeline is taken into account.

Figure 3. Relationship between the qualification given by the Inspectorate of Education with a reference date in year X and sickness absence of teachers in year X-1

Because the data of sickness absence of teachers is only available for the period 2016-2019, the 1st of April 2017, the 1st of April 2018, the 1st of April 2019 and the 1st of April 2020 are used as reference dates (Education Executive Agency, 2019c; Education Executive Agency, 2020d; Inspectorate of Education, 2017a; Inspectorate of Education, 2018a; Inspectorate of Education, 2019b, Inspectorate of Education, 2020f).

It is important to keep in mind that supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b). The supervision framework is used for the year with the 1st of April 2017 as reference date. The inspection framework is used for the years with the 1st of April 2018, the 1st of April 2019 and the 1st of April 2020 as reference dates. The dataset that is used for OLS and Fixed Effects will be used, but with other reference dates for the qualification of schools. OLS and Fixed Effects will be conducted. Since there are observations of the same schools in different years, the standard errors will be clustered at school level.

Table A6 (appendix) shows the relationship between the qualification of a school and sickness absence of teachers using sickness absence of teachers measured the year before the school qualification is given by the Inspectorate of Education on the 1st of April in a particular year.
A deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak' has a statistically significant effect on sickness absence of teachers when using sickness absence of teachers measured the year before the school qualification is given by the Inspectorate of Education on the 1st of April in a particular year. The estimated effect when including school fixed effects and year fixed effects is quite similar to the main estimate in Table 4, but the significance level declines when including fixed effects. This estimate is most credible, because the omitted variable bias is smaller when including fixed effects compared with OLS.
This estimated effect suggests reverse causality, because the sickness absence of teachers is measured the year before the qualification is given by the Inspectorate of Education, which might imply that the direction of the relationship is from sickness absence of teachers to the qualification.

Because stress is experienced before, during and after the inspection, it might be the case that the qualification with the 1st of September in a particular year as a reference date has an influence on the sickness absence of teachers in that particular year (e.g. Tunç et al, 2015). Therefore, to investigate the relationship between a qualification of a school given by the Inspectorate of Education and the sickness absence of teachers in a particular year, the focus is on the school qualification given by the Inspectorate of Education on the 1st of September in that particular year (Inspectorate of Education, 2017a; Inspectorate of Education, 2017b; Inspectorate of Education, 2018b; Inspectorate of Education, 2019d). Table 8 clarifies the dates chosen in this research, in particular the timing of the qualification, whereby a particular year is called year X.

Table 8. Timing of the qualification (1st of September, year X)

Variables	**Period**
Qualification with reference date 1st of September in year X	From the 1st of April in year X until the 1st of September in year X
Sickness absence of teachers in year X	From the 1st of January in year X until the 31st of December in year X

Figure 4 clarifies the relationship between the qualification given by the Inspectorate of Education and the sickness absence of teachers.

Figure 4. Relationship between the qualification given by the Inspectorate of Education with a reference date in year X and sickness absence of teachers in year X

Because the data of sickness absence of teachers is only available for the period 2016-2019, the 1st of September 2016, the 1st of September 2017, the 1st of September 2018 and the 1st of September 2019 are used as reference dates (Education Executive Agency, 2019c; Education Executive Agency, 2020d; Inspectorate of Education, 2017a; Inspectorate of Education, 2017b; Inspectorate of Education, 2018b; Inspectorate of Education, 2019d).

It is important to keep in mind that supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education, 2020b). The supervision framework is used for the year with the 1st of September 2016 as reference date. The inspection framework is used for the years with the 1st of September 2017, the 1st of September 2018 and the 1st of September 2019 as reference dates. The dataset that is used for OLS and Fixed Effects will be used, but with other reference dates for the qualification of schools. OLS and Fixed Effects will be conducted. Since there are observations of the same schools in different years, the standard errors will be clustered at school level.

Table A7 (appendix) shows the relationship between a qualification of a school and sickness absence of teachers using sickness absence of teachers measured in the same year as the school qualification is given by the Inspectorate of Education on the 1st of September in that year.
A deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak' has a statistically significant effect on sickness absence of teachers when using sickness absence of teachers measured in the same year as the school qualification is given by the Inspectorate of Education on the 1st of September in that year. The estimated effect when including fixed effects is quite similar to the main estimate in Table 4. This estimate is most credible, because the omitted variable bias is smaller when including fixed effects compared with OLS.

Due to the fact that stress is experienced after the inspection, it might be the case that the qualification with 1st of September in a particular year as a reference date has an influence on the sickness absence of teachers the year after (e.g. Tunç et al, 2015; Inspectorate of Education, 2020e). Therefore, to investigate the relationship between a qualification of a school given by the Inspectorate of Education and the sickness absence of teachers in a particular year, the focus is on sickness absence of teachers measured in the year after the school qualification is given by the Inspectorate of Education on the 1st of September (Inspectorate of Education, 2016; Inspectorate of Education, 2017a; Inspectorate of Education, 2017b; Inspectorate of Education, 2018b). Table 9 clarifies the dates chosen in this research, in particular the timing of the qualification, whereby a particular year is called year X and the year after is called year X+1.

Table 9. Timing of the qualification (1st of September, year X+1)

Variables	**Period**
Qualification with reference date 1st of September in year X	From the 1st of April in year X until the 1st of September in year X
Sickness absence of teachers in year X+1	From the 1st of January in year X+1 until the 31st of December in year X+1

Figure 5 clarifies the relationship between the qualification given by the Inspectorate of Education and the sickness absence of teachers, whereby the timeline is taken into account.

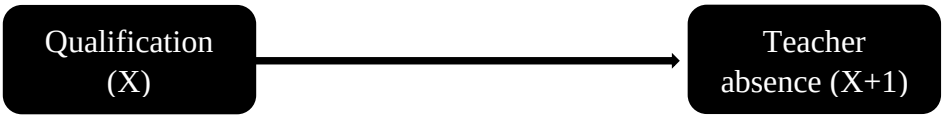

Figure 5. Relationship between the qualification given by the Inspectorate of Education with a reference date in year X and sickness absence of teachers in year X+1

Because the data of sickness absence of teachers is only available for the period 2016-2019, the 1st of September 2015, the 1st of September 2016, the 1st of September 2017 and the 1st of September 2018 are used as reference dates. (Education Executive Agency, 2019c; Education Executive Agency, 2020d; Inspectorate of Education, 2016; Inspectorate of Education, 2017a; Inspectorate of Education, 2017b; Inspectorate of Education, 2018b).

It is important to keep in mind that supervision by the Inspectorate of Education has been changed since the 1st of August in 2017. Before this date, there was a supervision framework. Since the 1st of August in 2017, there is an inspection framework (Inspectorate of Education,

2020b). The supervision framework is used for the years with the 1st of September 2015 and the 1st of September 2016 as reference dates. The inspection framework is used for the years with the 1st of September 2017 and the 1st of September 2018 as reference dates. The dataset that is used for OLS and Fixed Effects will be used, but with other reference dates for the qualification of schools. OLS and Fixed Effects will be conducted. Since there are observations of the same schools in different years, the standard errors will be clustered at school level.

Table A8 (appendix) shows the relationship between a qualification of a school and sickness absence of teachers using sickness absence of teachers measured in the year after the school qualification is given by the Inspectorate of Education on the 1st of September. A deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak' has not a statistically significant effect on sickness absence of teachers when including fixed effects. This estimate is most credible, because the omitted variable bias is smaller when including fixed effects compared with OLS.

This estimated effect suggests reverse causality, because the relationship is not significant anymore when sickness absence of teachers is measured the year after the qualification is given by the Inspectorate of Education, which might imply that the direction of the relationship is from sickness absence of teachers to the qualification. This is in contrast to the situation in which there is no reverse causality, because in that case the expectation is that there is still a significant relationship when the sickness absence of teachers is measured the year after the qualification is given.

7.3. Regression Discontinuity Design: Timing of the scores on standards

To investigate the relationship between the qualification 'very weak' given by the Inspectorate of Education and sickness absence of teachers in 2018, the focus is on how a particular school scores on the standards in the schoolyear 2017-2018 (Inspectorate of Education, 2019a; Education Executive Agency, 2020d). Similarly, to investigate the relationship between the qualification 'very weak' given by the Inspectorate of Education and sickness absence of teachers in 2019, the focus is on how a particular school scores on the standards in the schoolyear 2018-2019 (Inspectorate of Education 2019c; Education Executive Agency, 2020d).

The datafile about how a school scores on the standards in a particular schoolyear represents reports in the period from the 1st of September until the 31st of August (Inspectorate of Education, 2019e).

Therefore, it might be the case that the report about how a school scores on standards in the schoolyear 2017-2018 has an influence on the sickness absence of teachers in 2017. To investigate the relationship between the qualification 'very weak' given by the Inspectorate of Education and sickness absence in 2017, the focus is on how a particular school scores on the standards in the schoolyear 2017-2018 (Inspectorate of Education, 2019a; Education Executive Agency, 2020d). Similarly, it might be the case that the report about how a school scores on standards in the schoolyear 2018-2019 has an influence on sickness absence in 2018. To investigate the relationship between the qualification 'very weak' given by the Inspectorate of Education and sickness absence in 2018, the focus is on how a particular school scores on the standards in the schoolyear 2018-2019 (Inspectorate of Education, 2019c; Education Executive Agency, 2020d). The dataset that is used for the Regression Discontinuity Design will be used but linked to other years of sickness absence. A Regression Discontinuity Design will be conducted. Robust standard errors will be used.

Table A9 (appendix) shows the relationship between the qualification 'very weak' and the sickness absence of teachers in 2017 and 2018. The qualification 'very weak' has not a statistically significant effect on sickness absence of teachers.

8. Conclusion

In this research, the relationship between the qualification of a primary school given by the Inspectorate of Education and sickness absence of teachers in the Netherlands in the period 2016-2019 is investigated. The effect of the qualification given by the Inspectorate of Education on sickness absence of teachers is estimated by applying OLS, Fixed Effects and a Regression Discontinuity Design.

The results from OLS and Fixed Effects indicate that a deterioration in the qualification from 'good' to 'basis'/'sufficient' or from 'basis'/'sufficient' to 'weak'/'insufficient' or from 'weak'/'insufficient' to 'very weak' has a statistically significant effect on sickness absence of teachers in 2016-2019. It will result in an increase in sickness absence of teachers of approximately 19.3%. This estimated effect will be somewhat larger for small schools, larger in the period 2016-2017 (supervision framework) and somewhat smaller for large schools. This implies that there is an association between the qualification given by the Inspectorate of Education and sickness absence of teachers. There is no clear evidence that the direction of the relationship is from the qualification given by the Inspectorate of Education to sickness absence of teachers. The sensitivity analyses suggest reverse causality. Therefore, it is most plausible that the direction of the relationship is from sickness absence of teachers to the qualification given by the Inspectorate of Education.

The results from the Regression Discontinuity Design indicate that the qualification 'very weak' has no statistically significant effect on sickness absence of teachers in 2018 and 2019. The sensitivity analysis regarding the Regression Discontinuity Design did not raise any questions towards the credibility.

A number of limitations should be pointed out. First of all, there is data available about the sickness absence of teachers and management instead of data only about the sickness absence of teachers. Second, there is data available about the sickness absence of teachers per BRIN-number, not per school location. Therefore, the schools that use the same BRIN-number are excluded from the analyses. Third, as described in paragraph 5.3, the omitted variable bias is smaller in the case of a fixed effects strategy compared with OLS, because there are no omitted time invariant variables when including fixed effects. However, a fixed effects strategy still has shortcomings. For example, there might still be omitted variable bias due to omitted time variant

variables. Fourth, due to data constraints the Regression Discontinuity Design can only be conducted for two years.

A suggestion for further research is to investigate the direction of the relationship between the qualification given by the Inspectorate of Education and sickness absence of teachers more extensively, because the sensitivity analyses suggests reverse causality whereby the direction of the relationship is from sickness absence of teachers to the qualification given by the Inspectorate of Education. Furthermore, it is important to investigate why this relationship exists. A possible explanation is that a higher sickness absence of teachers will lead to lower quality of education and this might cause a deterioration in the qualification given by the Inspectorate of Education.

Depending on the direction of the relationship between sickness absence of teachers and the qualification of a school given by the Inspectorate of Education, a policy implication can be derived. It is most plausible that the direction of the relationship is from sickness absence of teachers to the qualification given by the Inspectorate of Education. As a consequence, it is important to combat sickness absence of teachers, because sickness absence of teachers might lead to lower quality of education. However, if the direction of the relationship is from the qualification given by the Inspectorate of Education to sickness absence of teachers, the finding that a deterioration in the qualification will result in an increase in sickness absence of teachers of approximately 19.3% raises questions about the effectiveness of the school inspection system in the Netherlands. A 19.3% increase in sickness absence of teachers will cost a lot of money. However, this has to weighted against the benefits. Therefore, a suggestion for further research is to investigate the benefits of the school inspection system in the Netherlands. After that, a cost-benefit analysis can be conducted to measure the effectiveness of the school inspection system in the Netherlands.

References

Alexander R. (1999). Inspection and education: the indivisibility of standards. In Cullingford, C. (1999). An inspector calls: Ofsted and its effect on school standards. Londen: Kogan Page.

Alkutich, M., & Abukari, A. (2018). Examining the benefit of school inspection on teaching and learning: a case study of Dubai private schools. *Journal of Education and Practice, 9*(5), 37- 48.

Angrist J.D. & Pischke J. (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Angrist, J.D. & Pischke, J. (2015). Mastering 'metrics: The path from cause to effect. Princeton University Press.

Balyer, A., Özcan, K., & Yildiz, A. (2017). Teacher Empowerment: School Administrators' Roles. *Eurasian Journal of Educational Research, 70*, 1-18.

Berryhill, J., Linney, J. A., & Fromewick, J. (2009). The Effects of Education Accountability on Teachers: Are Policies Too-Stress Provoking for Their Own Good? *International Journal of Education Policy and Leadership, 4*(5), 1-14.

Bowers, T., & McIver, M. (2000). Ill health retirement and absenteeism amongst teachers. Research Report No. 325.

Brown, M., Rutherford, D., & Boyle, B. (2000). Leadership for school improvement: the role of the head of department in UK secondary schools. *School Effectiveness and School Improvement, 11*(2), 237–258.

Bryk, A. S., & Schneider, B. (2002). Trust in schools: A core resource for improvement. Rose Series in Sociology. NY: Russell Sage Foundation.

Bubb, S., & Earley, P. (2004). Managing teacher workload: Work-life balance and wellbeing. Sage.

Case, P., Case, S., & Catling, S. (2000). Please show you're working: a critical assessment of the impact of OFSTED inspection on primary teachers. *British Journal of Sociology of Education*, *21*(4), 605-621.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2017). A Practical Introduction to Regression Discontinuity Designs. *Cambridge Elements: Quantitative and Computational Methods for Social Science - Cambridge University Press I.*

Chapman, C. (2001). Changing Classrooms Through Inspection. *School Leadership & Management*, *21*(1), 59-73.

Chapman, C. (2002). Ofsted and School Improvement: teachers' perceptions of the inspection process in schools facing challenging circumstances. *School Leadership & Management*, *22*(3), 257-272.

Dupont, E. (2007). Ziekteverzuim in het onderwijs: De balans tussen zakelijk en zorgzaam. Wolters Kluwer Nederland B.V.

Dworkin, A. G. (2001). Perspectives on teacher burnout and school reform. *International Education Journal*, *24*(2), 69–78.

Dworkin, A. G. (2009). Teacher burnout and teacher resilience: Assessing the impacts of the school accountability movement. In L.J. Saha & A. G. Dworkin (Eds.), New International Handbook of Teachers and Teaching. NY: Springer Publications, 491–509.

Dworkin, A. G., Saha, L. J., & Hill, A. N. (2003). Teacher burnout and perceptions of a democratic school environment. *International Education Journal*, *4*(2), 108–120.

Dworkin A. & Tobe P. (2014) The Effects of Standards Based School Accountability on Teacher Burnout and Trust Relationships: A Longitudinal Analysis. In: Van Maele D., Forsyth P. & Van Houtte M. Trust and School Life. Springer, Dordrecht.

Education Executive Agency. (2019a). Toelichting - Leerlingen bo – swv vestiging, impulsgebied, schoolgewicht. Retrieved September 10, 2020, from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/bo-sbo/bo-gewicht-impuls.jsp.

Education Executive Agency. (2019b, December 19). *Leerlingen bo – swv vestiging, impulsgebied, schoolgewicht – 2018-2019* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/bo-sbo/bo-gewicht-impuls.jsp.

Education Executive Agency. (2019c, September 6). *Verzuimkengetallen 2016-2018* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/publicaties/verzuim-personeel-po-vo/verzuim-personeel.jsp (as of June 8, 2020).

Education Executive Agency. (2019d). Leerlingen po – soort po, cluster, leeftijd – 2010-2020. Retrieved October 7, 2020, from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/po-totaal/bo-gewicht-leeftijd.jsp.

Education Executive Agency. (2019e, December 30). *Leerlingen bo per vestiging naar noat – 2010-2019* (Data file). Retrieved from https://onderwijsdata.duo.nl/dataset/poachtergrond/resource/5516f690-4aa9-43cf-90b3-14a652dd3710.

Education Executive Agency. (2019f). Leerlingen bo – swv vestiging, impulsgebied, schoolgewicht – 2014-2020. Retrieved October 7, 2020, from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/bo-sbo/bo-gewicht-impuls.jsp.

Education Executive Agency. (2019g, December 19). *Leerlingen po – soort po, cluster, leeftijd – 2015-2016* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/po-totaal/bo-gewicht-leeftijd.jsp.

Education Executive Agency. (2019h, December 19). *Leerlingen po – soort po, cluster, leeftijd – 2016-2017* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/po-totaal/bo-gewicht-leeftijd.jsp.

Education Executive Agency. (2019i, December 19). *Leerlingen po – soort po, cluster, leeftijd – 2017-2018* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/po-totaal/bo-gewicht-leeftijd.jsp.

Education Executive Agency. (2019j, December 19). *Leerlingen bo – swv vestiging, impulsgebied, schoolgewicht - 2015-2016* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/bo-sbo/bo-gewicht-impuls.jsp.

Education Executive Agency. (2019k, December 19). *Leerlingen bo – swv vestiging, impulsgebied, schoolgewicht - 2016-2017* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/bo-sbo/bo-gewicht-impuls.jsp.

Education Executive Agency. (2019l, December 19). *Leerlingen bo – swv vestiging, impulsgebied, schoolgewicht – 2017-2018* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/bo-sbo/bo-gewicht-impuls.jsp.

Education Executive Agency. (2019m, December 19). *Leerlingen po – soort po, cluster, leeftijd – 2018-2019* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/leerlingen-po/po-totaal/bo-gewicht-leeftijd.jsp.

Education Executive Agency. (2020a). Toelichting verzuimkengetallen personeel po en vo 2019. Retrieved October 6, 2020 from https://duo.nl/open_onderwijsdata/publicaties/verzuim-personeel-po-vo/verzuim-personeel.jsp.

Education Executive Agency. (2020b). NNCA uitwisselen. Retrieved September 21, 2020, from https://duo.nl/zakelijk/primair-onderwijs/leerlingenadministratie/nnca-uitwisselen-met-bron.jsp.

Education Executive Agency. (2020c, March 31). *Onderwijspersoneel po in personen 2011-2019* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/databestanden/po/onderwijspersoneel/po-personeel1.jsp.

Education Executive Agency. (2020d, September 25). *Verzuimkengetallen 2017-2019* (Data file). Retrieved from https://duo.nl/open_onderwijsdata/publicaties/verzuim-personeel-po-vo/verzuim-personeel.jsp.

Education Executive Agency. (2020e). Eindrapportage Verzuimonderzoek PO en VO 2019. Retrieved October 27, 2020 from https://duo.nl/open_onderwijsdata/images/eindrapportage-verzuimonderzoek-po-en-vo-2019.pdf.

Ehren M.C.M. & Visscher A.J. (2006). Towards a theory on the impact of school inspections. *British Journal of Educational Studies, 54*(1), 51-72.

Ferguson N., Earley P., Fidler B. & Ouston J. (2000). Improving Schools and Inspection: The Self-Inspecting School, London: Paul Chapman Publishing.

Figlio, D.N. (2006). Testing, crime and punishment. *Journal of Public Economics*, *90*(4-5), 837-851.

Figlio, D.N. & Loeb, S. (2011), School Accountability, Handbook of the Economics of Education, Volume 3, Chapter 8, Elsevier.

Fitz-Gibbon C.T. & Stephenson-Forster N.J. (1999). Is Ofsted helpful? In Cullingford, C. (1999). An inspector calls: Ofsted and its effect on school standards. Londen: Kogan Page.

Government of the Netherlands. (2020a). Overheidsfinanciering onderwijs. Retrieved September 8, 2020, from https://www.rijksoverheid.nl/onderwerpen/financiering-onderwijs/overheidsfinanciering-onderwijs.

Government of the Netherlands. (2020b). Toezicht en verantwoording in het basisonderwijs en voortgezet onderwijs. Retrieved September 8, 2020, from https://www.rijksoverheid.nl/onderwerpen/financiering-onderwijs/toezicht-en-verantwoording-in-het-onderwijs.

Haberman, M. (2004). Teacher burnout in black and white. *The new educator*, *1*(3), 153-175.

Hepburn, A., & Brown, S. D. (2001). Teacher stress and the management of accountability. *Human Relations*, *54*(6), 691-715.

Hopkins, E., Hendry, H., Garrod, F., McClare, S., Pettit, D., Smith, L., Burrel, H. & Temple, J. (2016). Teachers' views of the impact of school evaluation and external inspection processes. *Improving Schools*, *19*(1), 52-61.

Inspectorate of Education. (2012). Toezichtkader po/vo 2012. Retrieved September 8, 2020, from https://www.onderwijsinspectie.nl/onderwerpen/toezichtkaders/documenten/publicaties/2012/07/01/toezichtskader-po-vo-2012.

Inspectorate of Education. (2012-2021). Toezichtsarrangementen en oordelen 2012-2021. Retrieved November 3, 2020, from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/onderwijsdata/oordelen.

Inspectorate of Education. (2015-2019). Kwaliteitsindicatoren en standaarden 2015-2019. Retrieved October 8, 2020, from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/onderwijsdata/kwaliteitsindicatoren.

Inspectorate of Education. (2016, April 13). *Toezichtsarrangementen basisonderwijs sept. 2015 (met update apr. 2016)* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2016/4/13/toezichtarrangementen-basisonderwijs-2015-2016.

Inspectorate of Education. (2017a, April 12). *Toezichtsarrangementen basisonderwijs sept. 2016 (met update apr. 2017)* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2017/4/12/toezichtarrangementen-basisonderwijs-2016-2017.

Inspectorate of Education. (2017b, November 2). *Toezichtsarrangementen primair, special en voortgezet onderwijs sept. 2017 (versie 2)* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2017/11/02/arrangementen-2017-voor-primair-speciaal-en-voortgezet-onderwijs-v2.

Inspectorate of Education. (2018a, April 11). *Oordelen primair, speciaal en voortgezet onderwijs 1 april 2018* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2018/04/11/oordelen-po-so-vo-peildatum-1-april-2018.

Inspectorate of Education. (2018b, December 19). *Oordelen primair, special en voortgezet onderwijs 1 september 2018 (versie 2)* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2018/12/19/oordelen-primair-speciaal-en-voortgezet-onderwijs-1-september-2018-versie-2.

Inspectorate of Education. (2019a, March 15). *Standaarden PO, SO, VO en MBO 2017/2018* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2019/03/15/standaarden-po-so-vo-en-mbo-2017-2018.

Inspectorate of Education. (2019b, October 15). *Oordelen primair, special en voortgezet onderwijs 1 april 2019 (versie 2)* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2019/10/15/oordelen-primair-speciaal-en-voortgezet-onderwijs-1-april-2019-versie-2.

Inspectorate of Education. (2019c, December 20). *Standaarden PO, SO, VO en MBO 2018/2019* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2019/12/20/standaarden-po-so-vo-en-mbo-2018-2019.

Inspectorate of Education. (2019d, September 1). *Oordelen primair, special en voortgezet onderwijs 1 september 2019 (versie 2)* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2019/09/01/oordelen-po-s-vo-vo-1-september-2019-versie-2.

Inspectorate of Education. (2019e). Toelichting open data standaarden schooljaar 2018/2019 versie 1. Retrieved October 29, 2020 from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2019/12/20/toelichting-open-data-standaarden-2018-2019.

Inspectorate of Education. (2020a). The inspection of schools and governing bodies of schools. Retrieved September 8, 2020, from https://english.onderwijsinspectie.nl/inspection/inspection-of-schools-by-the-dutch-inspectorate-of-education/the-inspectorate%E2%80%99s-approach.

Inspectorate of Education. (2020b). Toezichtkaders. Retrieved September 8, 2020, from https://www.onderwijsinspectie.nl/onderwerpen/toezichtkaders.

Inspectorate of Education. (2020c). Onderzoekskader 2017 voor het toezicht op voorschoolse educatie en het primair onderwijs. Retrieved September 8, 2020, from https://www.onderwijsinspectie.nl/onderwerpen/onderzoekskaders/documenten/rapporten/2020/06/22/onderzoekskader-2017-po-en-vve.

Inspectorate of Education. (2020d). Oordelen door de inspectie. Retrieved October 7, 2020 from https://www.onderwijsinspectie.nl/onderwerpen/oordelen-van-de-inspectie-van-het-onderwijs.

Inspectorate of Education. (2020e). Toelichting open data oordelen op de peildatum 1 april 2020 versie 2. Retrieved October 29, 2020 from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/publicaties/2020/04/22/toelichting-open-data-oordelen-1-april-2020-versie-2.

Inspectorate of Education. (2020f, April 22). *Oordelen primair, special en voortgezet onderwijs 1 april 2020 (versie 2)* (Data file). Retrieved from https://www.onderwijsinspectie.nl/trends-en-ontwikkelingen/documenten/data-bestanden/2020/04/22/oordelen-po-s-vo-vo-1-april-2020-versie-2.

Learmonth, J. (2000). Inspection: What's in it for Schools? Psychology Press.

MacBeath, J. (2006). School inspection & self-evaluation: Working with the new relationship. Routledge.

Mathison S. & Freeman M. (2006). Teacher stress and high stakes testing: How Using One Measure of Academic Success Leads to Multiple Teacher Stressors. In: Lambert R. & McCarty C. Understanding Teacher Stress in an Age of Accountability. Information Age Publishing, Greenwich, Connecticut, 43-63.

Matthews, P., & Sammons, P. (2004). Improvement through Inspection: an evaluation of the impact of Ofsted‘s work. London: Ofsted/Institute of Education.

PA Consulting (2000). Better Accountability for Higher Education: Summary of a Review for the HEFCE. Report August 00/36.

Perryman, J. (2007). Inspection and emotion. *Cambridge journal of education*, *37*(2), 173-190.

Perryman, J., & Calvert, G. (2020). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *British Journal of Educational Studies*, *68*(1), 3-23.

Phillips, S., Sen, D., & McNamee, R. (2007). Prevalence and causes of self-reported work-related stress in head teachers. *Occupational medicine*, *57*(5), 367-376.

Quintelier, A., Vanhoof, J., & De Maeyer, S. (2018). Understanding the influence of teachers' cognitive and affective responses upon school inspection feedback acceptance. *Educational Assessment, Evaluation and Accountability*, *30*(4), 399-431.

Quintelier, A., Vanhoof, J., & De Maeyer, S. (2019). A full array of emotions: An exploratory mixed methods study of teachers' emotions during a school inspection visit. *Studies in Educational Evaluation*, *63*, 83-93.

Richards K. A. R., Hemphill, M. A., & Templin, T. J. (2018). Personal and contextual factors related to teachers' experience with stress and burnout. *Teachers and Teaching*, *24*(7), 768-787.

Roelandt, T., Martens, E. & Veenman, J. (1990). Achterstand van allochtonen in het onderwijs: sociaal milieu en migratie-achtergronden. *Mens en Maatschappij, 65*(2), 103-125.

Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of education review*, *23*(2), 143-151.

Ryan, S. V., Von der Embse, N. P., Pendergast, L. L., Saeki, E., Segool, N., & Schwing, S. (2017). Leaving the teaching profession: The role of teacher stress and educational accountability policies on turnover intent. *Teaching and Teacher Education*, *66*, 1-11.

Saeki, E., Segool, N., Pendergast, L., & von der Embse, N. (2018). The influence of test-based accountability policies on early elementary teachers: School climate, environmental stress, and teacher stress. *Psychology in the Schools*, *55*(4), 391-403.

Scanlon, M. (1999). The impact of OFSTED inspections. London: National Foundation for Educational Research for the National Union of Teachers.

Shernoff, E. S., Mehta, T. G., Atkins, M. S., Torf, R., & Spencer, J. (2011). A qualitative study of the sources and impact of stress among urban teachers. *School mental health*, *3*(2), 59-69.

Smulders, P. (2004). Ziekteverzuim: hoogte, oorzaken, aandoeningen, werkgebondenheid en maatregelen. Secundaire analyses Nationale Enquête Arbeidsomstandigheden 2003. TNO Arbeid.

Sun, M., Saultz, A., & Ye, Y. (2017). Federal policy and the teacher labor market: exploring the effects of NCLB school accountability on teacher turnover. *School Effectiveness and School Improvement*, *28*(1), 102-122.

Tunç, B., İnandı, Y., & Gündüz, B. (2015). Inspection and teachers' emotions: An emotional evaluation of inspection. *Journal of Human Sciences*, *12*(1), 554-568.

Troman, G., & Woods, P. (2001). Primary teachers' stress. Psychology Press.

Van Droogenbroeck, F., Spruyt, B., & Vanroelen, C. (2014). Burnout among senior teachers: Investigating the role of workload and interpersonal relationships at work. *Teaching and teacher education*, *43*, 99-109.

Von der Embse N. P., Sandilos, L. E., Pendergast, L., & Mankin, A. (2016a). Teacher stress, teaching-efficacy, and job satisfaction in response to test-based educational accountability policies. *Learning and Individual Differences*, *50*, 308-317.

Von der Embse, N. P., Pendergast, L. L., Segool, N., Saeki, E., & Ryan, S. (2016b). The influence of test-based accountability policies on school climate and teacher stress across four states. *Teaching and Teacher Education*, *59*, 492-502.

Webb. P. T. (2006) The Stress of Accountability: Teachers as Policy Brokers in a Poverty School. In: Lambert R. & McCarty C. Understanding Teacher Stress in an Age of Accountability. Information Age Publishing, Greenwich, Connecticut, 1-18.

Wet op het onderwijstoezicht. (2020, August, 1). Retrieved September 8, 2020, from https://wetten.overheid.nl/BWBR0013800/2020-08-01#Hoofdstuk3.

Wolf, de I., & Janssens, F. J. G. (2005). Effecten en neveneffecten van inspecties en prestatie-indicatoren in het onderwijs: een overzicht van empirische studies. *Pedagogische studiën*, *82*(6), 421-435.

Appendix

Table A1. Small schools: relationship between the qualification of a school and sickness absence of teachers

Variables	(1) **OLS**	(2) **Fixed Effects**	(3) **Fixed Effects**
Qualification (1-4)	2.266***	1.371**	1.320**
	(0.717)	(0.617)	(0.659)
Number of pupils	-0.002	0.002	-0.002
	(0.006)	(0.037)	(0.041)
Pupils that speak a foreign language (%)	0.038***	0.114	0.132
	(0.010)	(0.114)	(0.109)
Age of teachers	0.135***	-0.101	-0.149
	(0.046)	(0.123)	(0.146)
Female teachers (%)	0.004	0.020*	0.030**
	(0.007)	(0.011)	(0.013)
Full-time equivalent	-4.600	-0.693	-0.824
	(3.800)	(4.772)	(4.833)
Impulse area	1.484	1.235	1.111
	(1.037)	(2.376)	(2.743)
Constant	-2.553	4.289	6.688
	(1.828)	(6.878)	(8.712)
Observations	5,638	5,638	5,638
R-squared	0.002	0.001	0.001
Number of schools	2,088	2,088	2,088
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A2. Large schools: relationship between the qualification of a school and sickness absence of teachers

Variables	(1) OLS	(2) Fixed Effects	(3) Fixed Effects
Qualification (1-4)	1.820***	0.995***	0.884**
	(0.426)	(0.333)	(0.348)
Number of pupils	0.000	0.001	0.000
	(0.001)	(0.005)	(0.005)
Pupils that speak a foreign language (%)	0.023***	0.048	0.082**
	(0.008)	(0.034)	(0.038)
Age of teachers	0.144***	0.258**	0.219**
	(0.047)	(0.102)	(0.091)
Female teachers (%)	-0.014*	-0.006	-0.001
	(0.009)	(0.006)	(0.006)
Full-time equivalent	-5.636***	-14.23***	-13.19***
	(2.105)	(4.931)	(4.316)
Impulse area	0.472	-2.605***	-3.157***
	(0.416)	(0.891)	(0.885)
Constant	0.462	2.243	3.553**
	(1.091)	(1.684)	(1.693)
Observations	8,178	8,178	8,178
R-squared	0.004	0.002	0.004
Number of schools	2,499	2,499	2,499
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A3. Schools with qualification from 'basis'/'sufficient' to 'weak'/'insufficient': relationship between the qualification of a school and sickness absence of teachers

	(1)	(2)	(3)
Variables	**OLS**	**Fixed Effects**	**Fixed Effects**
Qualification (1-4)	-1.259	1.272**	1.125*
	(2.067)	(0.629)	(0.640)
Number of pupils	0.001	-0.017	-0.012
	(0.010)	(0.041)	(0.043)
Pupils that speak a foreign language (%)	0.226*	0.794***	0.771***
	(0.125)	(0.233)	(0.217)
Age of teachers	-0.726**	1.394***	1.374***
	(0.350)	(0.229)	(0.236)
Female teachers (%)	-0.043	0.032***	0.034***
	(0.037)	(0.008)	(0.009)
Full-time equivalent	-26.09	-0.451	-0.615
	(15.78)	(14.52)	(14.71)
Impulse area	-2.445	-	-
	(3.406)		
Constant	62.03*	-66.15***	-66.28***
	(31.95)	(12.31)	(12.27)
Observations	175	175	175
R-squared	0.140	0.721	0.727
Number of schools	103	103	103
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. When including fixed effects, impulse area is omitted because of collinearity. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Relationship between the qualification of a school and sickness absence of teachers in 2016-2017

Variables	(1) **OLS**	(2) **Fixed Effects**	(3) **Fixed Effects**
Qualification (1-4)	2.237***	1.643***	1.655***
	(0.495)	(0.621)	(0.624)
Number of pupils	-0.001**	0.007**	0.007**
	(0.000)	(0.004)	(0.003)
Pupils that speak a foreign language (%)	0.014***	-0.008	-0.005
	(0.004)	(0.025)	(0.026)
Age of teachers	0.169***	0.078	0.075
	(0.017)	(0.052)	(0.052)
Female teachers (%)	0.021**	-0.006	-0.005
	(0.008)	(0.019)	(0.019)
Full-time equivalent	0.724	-4.523**	-4.592**
	(0.836)	(1.817)	(1.820)
Impulse area	0.598***	-2.838***	-2.870***
	(0.174)	(0.971)	(0.971)
Constant	-8.151***	1.959	2.098
	(1.657)	(3.595)	(3.590)
Observations	7,525	7,525	7,525
R-squared	0.034	0.008	0.008
Number of schools	4,271	4,271	4,271
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A5. Relationship between the qualification of a school and sickness absence of teachers in 2018-2019

Variables	(1) **OLS**	(2) **Fixed Effects**	(3) **Fixed Effects**
Qualification (1-4)	1.899***	-0.089	-0.066
	(0.576)	(0.619)	(0.617)
Number of pupils	0.000	0.008**	0.007*
	(0.003)	(0.004)	(0.004)
Pupils that speak a foreign language (%)	0.047***	0.027	0.037
	(0.013)	(0.042)	(0.041)
Age of teachers	0.151**	0.080*	0.069
	(0.060)	(0.042)	(0.044)
Female teachers (%)	0.002	0.007***	0.010***
	(0.005)	(0.002)	(0.002)
Full-time equivalent	-9.788**	-4.187**	-4.393**
	(4.292)	(1.847)	(1.839)
Impulse area	1.175	-	-
	(1.083)		
Constant	0.493	2.036	2.603*
	(1.256)	(1.569)	(1.536)
Observations	6,311	6,311	6,311
R-squared	0.002	0.011	0.019
Number of schools	3,900	3,900	3,900
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. When including fixed effects, impulse area is omitted because of collinearity. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A6. Relationship between the qualification of a school and sickness absence of teachers using sickness absence of teachers measured the year before the school qualification is given by the Inspectorate of Education on the 1st of April in a particular year

	(1)	(2)	(3)
Variables	**OLS**	**Fixed Effects**	**Fixed Effects**
Qualification (1-4)	1.870***	1.083**	1.028**
	(0.379)	(0.430)	(0.430)
Number of pupils	0.000	0.009***	0.008***
	(0.000)	(0.002)	(0.002)
Pupils that speak a foreign language (%)	0.019***	0.011	0.019
	(0.004)	(0.018)	(0.018)
Age of teachers	0.093***	0.041	0.025
	(0.023)	(0.046)	(0.045)
Female teachers (%)	-0.003	0.003	0.007
	(0.005)	(0.004)	(0.005)
Full-time equivalent	-1.492	-5.508***	-5.567***
	(0.975)	(1.231)	(1.236)
Impulse area	0.520***	-0.770	-0.849
	(0.162)	(1.357)	(1.375)
Constant	-0.864	3.372**	4.082**
	(1.337)	(1.718)	(1.696)
Observations	12,048	12,048	12,048
R-squared	0.024	0.008	0.010
Number of schools	4,533	4,533	4,533
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A7. Relationship between the qualification of a school and sickness absence of teachers using sickness absence of teachers measured in the same year as the school qualification is given by the Inspectorate of Education on the 1st of September in that year

	(1)	(2)	(3)
Variables	**OLS**	**Fixed Effects**	**Fixed Effects**
Qualification (1-4)	1.650***	1.287***	1.229***
	(0.272)	(0.268)	(0.275)
Number of pupils	-0.001	0.001	-0.001
	(0.002)	(0.009)	(0.010)
Pupils that speak a foreign language (%)	0.030***	0.104	0.129*
	(0.006)	(0.076)	(0.075)
Age of teachers	0.139***	0.068	0.032
	(0.034)	(0.083)	(0.088)
Female teachers (%)	-0.001	0.010**	0.015***
	(0.005)	(0.005)	(0.006)
Full-time equivalent	-5.029**	-6.375*	-6.102*
	(2.093)	(3.639)	(3.682)
Impulse area	0.919*	-1.157	-1.370
	(0.529)	(1.484)	(1.598)
Constant	-0.590	2.294	4.083
	(1.370)	(3.237)	(3.675)
Observations	13,147	13,147	13,147
R-squared	0.002	0.001	0.001
Number of schools	4,578	4,578	4,578
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A8. Relationship between the qualification of a school and sickness absence of teachers using the sickness absence of teachers measured in the year after the school qualification is given by the Inspectorate of Education on the 1st of September in a particular year

	(1)	(2)	(3)
Variables	**OLS**	**Fixed Effects**	**Fixed Effects**
Qualification (1-4)	1.216***	0.278	0.183
	(0.406)	(0.352)	(0.379)
Number of pupils	-0.001	-0.001	-0.003
	(0.002)	(0.014)	(0.014)
Pupils that speak a foreign language (%)	0.031***	0.201	0.223
	(0.006)	(0.159)	(0.156)
Age of teachers	0.311	0.270	0.235
	(0.363)	(0.589)	(0.598)
Female teachers (%)	0.018	0.008	0.013
	(0.035)	(0.026)	(0.025)
Full-time equivalent	0.276	0.583	1.053
	(12.06)	(23.41)	(23.13)
Impulse area	0.945**	-0.506	-0.693
	(0.470)	(1.599)	(1.774)
Constant	-12.47	-10.17	-8.467
	(27.20)	(42.82)	(43.28)
Observations	14,486	14,486	14,486
R-squared	0.002	0.001	0.001
Number of schools	4,727	4,727	4,727
School FE		YES	YES
Year FE			YES

Notes. Clustered standard errors in parentheses. The qualification contains four values: 'good' (1), 'basis'/'sufficient' (2), 'weak'/'insufficient' (3) and 'very weak' (4). Impulse area is a dummy variable, which gets the value 1 if the school is located in an impulse area, and 0 otherwise. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.

Table A9. Relationship between the qualification 'very weak' and the sickness absence of teachers in 2017 and 2018

Variables	(1) **Regression Discontinuity Design**
Qualification (=1 if the qualification is 'very weak', =0 otherwise)	-0.495
	(1.546)
Scores on other standards	-
Constant	7.721***
	(0.604)
Observations	106
R-squared	0.001
Number of schools	
School FE	

Notes. Robust standard errors in parentheses. The qualification 'very weak' is a dummy variable, which gets the value 1 if the qualification is 'very weak', and 0 otherwise. The scores are on three other standards: 'development perspectives', 'didactics' and 'safety'. The school receives code 0 if the school scores on none of the other standards 'insufficient'. The school receives code 1 if the school scores on one of the other standards 'insufficient'. The school receives code 2 if the schools scores on two of the other standards 'insufficient'. The school receives code 3 if the school scores on all the other standards 'insufficient'. Scores on other standards is omitted because of collinearity. The result is at least rounded to three numbers after the decimal point. *** p<0.01, ** p<0.05, * p<0.1.