Bureaucratic representation in Dutch secondary schools: what is the role of gender in student performance?

Pieterse, Sanne

Universiteit Leiden

Bureaucratic representation in Dutch secondary schools: what is the role of gender in student performance?

Sanne Pieterse
S1130307
Master's Thesis for Public Administration
International and European Governance

Faculty of Governance and Global Affairs
Leiden University
Supervised by dr. P. van den Bekerom

9-6-2023, Leiden

Abstract

The effect of bureaucratic representation on organisational performance has for decades been a subject of interest for scholars. Many scholars have debated the merits and effects of passive and active representation. Drawing on the literature, this thesis aims to add to existing research by focusing on the influence of bureaucratic representation on student performance in Dutch secondary schools. It will focus specifically on the dimension of gender and aims to research whether the gender of a teacher has an effect on the performance of students of the same gender. The study includes two moderator variables. Both the stratification of management and the gender division of the entire educational staff could have an effect on the relationship between teachers and student performance.

This study was conducted using a Large-N design that includes all Dutch secondary schools. Data on Dutch secondary schools, collected by the Dienst Uitvoering Onderwijs, was used to perform a multilevel regression analysis to test the hypotheses. This study reveals that only the moderator variable 'number of educational staff' has an impact on the relationship between teachers and student performance. There was no support for the other hypotheses.

Keywords: bureaucratic representation, discretion, secondary education, gender, organisational performance

Foreword

This thesis is written to complete the Master's program Public Administration: International and European Governance at Leiden University. It describes the role that bureaucratic representation in secondary education plays in student performance.

Writing this thesis gave me the chance to delve deeply into a specific subject. By reading relevant literature I had the opportunity to see the perspective on the subject of bureaucratic representation develop. Scholars build onto each other's works, establish their own theories, formulate new hypotheses and discard the old ones, and incorporate current findings into the subject. It was interesting to see that even though bureaucratic representation has been researched for decades, scholars still manage to gain new insights and discover possible benefits for society.

Though I have done quantitative research before, it was never at the level that was required for this thesis. Though challenging, it also gave me new insights into the difficulties that come with quantitative research. I had to figure out the best way to measure variables with only the limited data that was available and then try and present it in a way that makes sense to the reader.

In this foreword I would like to express my gratitude to my thesis supervisor dr. Petra van den Bekerom for all her advice and for answering my endless questions. I would also like to thank my friends Nina, Bjorn and Moniek for their encouragement and Heske and Gina for their help and support.

Leiden, June 2023

Sanne Pieterse

Table of contents

List of figures

List of tables

List of tables in the appendix

1. Introduction

The importance of a good education to fight inequality has been widely acknowledged. It is one of the key elements for mobility in a society and could help bridge the gender pay gap between men and women (Walker et al, 2019: 6; WEF, 2020). Despite a gradual decline of wage inequality in the Netherlands (CBS, 2022), at 13.5% it remains higher than the EU average (Eurostat, 2023).

On the other hand, there also exists a significant gender disparity in academic performance, as for more than two decades a higher proportion of women than men have pursued tertiary education (CBS, 2023a). This disparity in performance begins at an early age and starts to increase around the age of twelve. Female students consistently outperform male students in all areas except calculus (Zumbuehl and Dillingh, 2020: 6-7). One of the causes of this disparity could be the dynamics between teachers and students: the idea that the presence of a teacher of the same gender can have a positive effect on the performance of a student. Explanations for this link differ, but can roughly be divided into two categories. One group of explanations concerns passive effects: the mere presence of a teacher of the same gender or race as a role model that a student can identify with, can lead to more confidence and motivation. A second group of explanations concerns active effects: teacher biases can lead to a different treatment of students based on their demographic background (Dee, 2004).

The notion that the background of teachers impacts student performance is part of the concept of bureaucratic representation. Bureaucratic representation refers to the idea that the composition of the bureaucracy of public organisations should mirror that of its clients (Grissom et al, 2015). The bureaucratic workforce can hold significant influence on the implementation of public policies; bureaucrats often exercise discretion in decision-making as organisations cannot make rules for every eventuality. Research has suggested that a representative bureaucracy can greatly benefit groups that are traditionally underrepresented or disadvantaged (Riccucci and Van Ryzin, 2017), focusing on race, ethnicity and gender.

The idea of representative bureaucracy can also be applied to schools. Multiple studies show that bureaucratic representation in schools has a positive effect on student performance (Grissom et al, 2015). A study by Meier (1993) suggests that Latino students benefit from being taught by a Latino teacher. Likewise, a study by Dee (2004) showed improvements in student performance when students were taught by a teacher of their own race. Keiser et al (2002) demonstrated an increase in maths scores for female students when being taught by a female teacher.

Examining this mechanism could benefit minority students and improve social mobility.

1.1 Research question

This thesis aims to discover the impact of bureaucratic representation on student performance, specifically focusing on the dimension of gender. The study will target secondary schools in particular, as they publish the required data. The research question is formulated as follows:

To what extent has bureaucratic representation of gender an effect on student performance in secondary schools?

1.2 Justification

Though the impact of bureaucratic representation has been studied widely, many of these studies have been conducted in the United States. The question remains whether the environment and conditions that link the relationship between bureaucratic representation and organisational performance holds true in other countries (Song, 2018). Differences in governmental structures between countries might limit the impact of bureaucratic representation. It is not always possible to transfer the same ideas to another country, as cultural and institutional differences might hinder the relationship or even be counter-productive (Dauda, 1990). The results of previous studies have been mixed: a study comparing the impact of management on organisational output in the United States and Denmark showed that institutional context does matter (Meier et al, 2015), while a study conducted in Korea shows a positive link between bureaucratic representation and educational performance (Song, 2018). Studying the effects of bureaucratic representation in Dutch secondary schools can shed light on the effectiveness of trying to strive for representation in bureaucratic organisations in the Netherlands in general, and secondary schools in particular.

Furthermore, the subject of this thesis is of societal relevance. Researching the impact of bureaucratic representation on student performance can help schools develop ways to address the gender disparity in academic outcomes between male and female students. In addition, answering this question could potentially benefit minority students and improve social mobility in the Netherlands.

1.3 Thesis outline

This thesis is structured into multiple chapters. The following chapter provides an overview of the existing research that has already been conducted on the topic of bureaucratic representation. These insights will serve to develop several hypotheses. Chapter three will detail the methodology employed to address the research question. It includes the study design, data collection methods and operationalisation. Chapter four presents the initial results obtained from the data analysis and examines the findings. In chapter five, the impact and implications of the findings will be discussed.

2. Theoretical framework

In order to examine the role of bureaucratic representation in education, it is necessary to consider the valuable insights collected from decades of prior research on this subject. This study will rely on the knowledge and findings that have been previously established. This chapter will provide a comprehensive overview of earlier research which will be used to develop the hypotheses required to answer the research question.

2.1. Introducing representative bureaucracy

The concept of representative bureaucracy was first introduced by Kingsley (1944) in his book titled: 'Representative Bureaucracy: An Interpretation of the British Civil Service'. It is founded on the idea that organisational performance will improve with a bureaucracy that is representative of a society's population. Since then, this idea has become an important topic of discourse on good governance and public administration. These discussions have intensified in recent decades, with increased attention for diversity and equality (Rainey, 2009: 108). Bureaucrats are no longer seen as neutral implementers of policy, but as agents filling a gap between their organisation and its clients (Groeneveld and Van de Walle, 2010). A representative bureaucracy can improve the effectiveness of a public organisation and potentially advance greater social equality in the long term (Peters, 2010: 85-86).

2.1.1. Defining bureaucratic representation

One of the major problems with the theory of representative bureaucracy is its lack of consistency in defining what exactly representative bureaucracy means, making analysis of the subject more difficult. Therefore, it is essential to define clear characteristics of bureaucratic representation. Contemporary research on the subject tends to focus mostly on race and gender, while other characteristics like class/income, religion, region, and education are given less attention (Kennedy, 2014: 396; Ding et al, 2021: 1005).

Many scholars have suggested their own definitions of the concept of representative bureaucracy (e.g. Kingsley, 1944; Grissom et al, 2015; Riccucci and Van Ryzin, 2017). This study will adopt the definition put forward by Bradbury and Kellough (2008: 697-698):

"The theory of representative bureaucracy suggests that diversity within the public workforce, especially in terms of characteristics such as race and ethnicity will help to ensure that the interests of diverse groups are represented in policy formulation and implementation processes."

However, this definition does not specifically address the dimension of gender that this study intends to explore. Therefore, Bradbury and Kellough's definition will be slightly modified:

"The theory of representative bureaucracy suggests that diversity within the public workforce, especially in terms of characteristics such as race, ethnicity and gender will help to ensure that the interests of diverse groups are represented in policy formulation and implementation processes."

Another issue encountered when discussing bureaucratic representation is the distinction between passive and active representation. While Frederick Mosher was not the first to broach the subject of bureaucratic representation, his book 'Democracy and the Public Service' (1968) has been very influential in subsequent discourse in public administration (e.g. Lim, 2006; Grissom et al, 2015). Mosher was the first to differentiate between passive and active representation, which has become an important distinction in later discussions on bureaucratic representation. Passive representation is characterised by Mosher as symbolically significant for the public, as it demonstrates that the public sector is open to anyone. He defines passive representation as:

"[...] the source of origin of individuals and the degree to which, collectively, they mirror the total society" (Mosher, 1968: 12).

In contrast, Mosher regards active representation less favourably. Minority bureaucrats may prioritise their minority group over other groups; special interests of minority groups would dominate the interests of the general public. Advantages for one group may mean disadvantaging another group (Kennedy, 2014: 397). However, Mosher acknowledges that minority bureaucrats will not necessarily represent the interests of their group, and measures can be implemented to prevent partiality. He provides the definition of active representation as:

"[...] an individual is expected to press for the interests and desires of those whom he is presumed to represent, whether they be the whole people or some segment of the people" (Mosher, 1968: 12).

2.1.2 Linking passive and active representation

An important factor in the relationship between passive and active representation is the assumption that minority bureaucrats will hold the same beliefs as the minority group they are representing (Bradbury and Kellough, 2008). Bradbury and Kellough (2008) attempt to find evidence for this link by measuring the similarities between values and attitudes held by minority bureaucrats and minority citizens. Their findings suggest that there is a lot more similarity between the attitudes of minority citizens and minority bureaucrats, compared to non-minority bureaucrats. This supports one of the assumptions of active representation: demographic characteristics like race and ethnicity influence the attitudes of both citizens and bureaucrats.

In a subsequent literature review, Bradbury and Kellough (2011) found clear evidence of an association between passive and active representation: the interests of minorities are more likely to be represented in organisations with a larger share of minorities. However, it is difficult to determine whether this is the result of active representation, as other sources could also lead to increased benefits for the minority group.

A significant area of research in the field of bureaucratic representation concerns the circumstances under which passive representation can lead to active representation. Though there has been extensive research conducted discussing the link between representative bureaucracy and organisational performance, much is unclear as to under which conditions organisational output is actually impacted by bureaucratic representation (Keiser et al, 2002; Ding et al, 2021). Thompson (1976: 213-217) describes the difficulties associated with linking the two concepts: institutional factors like work socialisation, uncertainty and peer pressure, and a limited presence of minority bureaucrats in senior management roles can be an obstacle to active representation. Meier (1975: 541) likewise observes that the more significant decisions are made primarily at the higher management levels of an organisation, making bureaucratic representation at a lower level inconsequential. Hierarchical organisations leave less room for bureaucrats to influence organisational output.

An important condition for linkage between passive and active representation to occur is discretion (Meier and Stewart, 1992; Meier, 1993; Keiser et al, 2002): bureaucrats cannot

influence organisational output without having the ability to make independent decisions. A study by Meier and Bohte (2001) shows an increase in organisational performance when employees have greater discretion, as they can use their expertise without being limited by organisational structures. Bureaucrats wield considerable power by exercising discretion in decision-making regarding public policies. Higher level managers in organisations make the major decisions, while street-level bureaucrats can influence the implementation of public programs (Selden, 1997a: 13-15).

Henderson (1978) notes the importance of political support from other minorities, either inside the organisation or through advocacy groups. Without this support, minority bureaucrats are unlikely to become active representatives.

2.1.3 Effect of bureaucratic representation on organisational performance

An important aspect of the concept of bureaucratic representation is the idea that increased representativeness can have beneficial results for organisational performance. Numerous scholars have tried to identify potential factors contributing to this relationship.

Selden (1997a: 6-7) outlines five benefits of a representative bureaucracy. First of all, it demonstrates the values of the government by promoting equal access to influential positions. Secondly, minority bureaucrats have a greater understanding of the needs of minority groups and might be more responsive to those needs. Thirdly, an increased representation of minorities in higher management positions can influence agenda-setting and prioritise issues that are of specific importance to them. Fourthly, underrepresented minority groups may be more cooperative with public organisations as the number of minorities increases. And lastly, the recruitment of minorities expands the pool of potential candidates.

In a study on the Farmers Home Administration, Selden (1997b) also finds a link between the representativeness of the bureaucracy and the output of the policy process. Active representation leads to increased responsiveness to minority groups as. She concludes that:

"A bureaucracy that employs a cross-section of society is likely to ensure that the interests of all groups are considered in the policy process." (Selden, 1997b: 37)

Meier (2019: 45) suggests that organisations can benefit from a representative bureaucracy if they provide extrinsic benefits. Representative organisation will be more effective in engaging with minority groups.

As Riccucci and Van Ryzin (2017) show, even without active representation, bureaucratic representation can be useful for promoting equality. The presence of minority bureaucrats alone can have a positive effect on society by giving the government a greater legitimacy in the eyes of its citizens.

Lim (2006: 195-196) describes several mechanisms where passive representation can benefit minority groups. The positive impact of minority bureaucrats does not need to result from partiality, but from both direct and indirect sources. These direct sources are effects stemming from their own behaviour, as a shared background also means shared values. Indirect effects can come from influencing the behaviour of their fellow bureaucrats by disapproving or restraining bias from non-minority bureaucrats. Choi and Rainey (2010) suggest that bureaucrats perceive that racial diversity improves organisational performance, but only when it involves effective management practices. A study by Hong (2016) on the effects of increased diversity on the organisational performance of the U.K. police force shows that it improved both the integrity of the organisation and affected the perception of minority citizens.

Though previous studies (e.g. Meier, 1993; Keiser et al, 2002; Bradbury and Kellough, 2011) have demonstrated that minority representation in bureaucracies can benefit minority groups, this does not necessarily confirm that this is the result of active representation. Even without taking action, minority bureaucrats can have a significant influence on the behaviour of their colleagues and clients (Lim, 2006: 197).

2.1.4 The role of symbolic representation

A recent strand of research focuses on the idea of symbolic representation: minority groups can benefit from passive representation even without minority bureaucrats taking purposeful actions. Citizens are more likely to trust the organisation and perceive the organisation as legitimate, which leads to a higher willingness to cooperate (Theobald and Haider-Markel, 2009). This effect is visible among both racial and ethnic minorities (Theobald and Haider-Markel, 2009) and women (Riccucci et al, 2014; Riccucci et al, 2016). The increased perception of legitimacy can also lead to higher satisfaction with the services of public organisations (Riccucci and Van Ryzin, 2017).

A recent meta-analysis by Ding et al (2021: 1013) even suggests there is little difference in the effects of active and symbolic representation on improving organisational performance.

2.2 Bureaucratic representation in education

Studies on bureaucratic representation have shown the effects of representation on the performance of public organisations. This mechanism can also be applied to schools, as both teachers and administrators are considered bureaucrats. Meier and Bohte (2001: 459) present several ways minority teachers can influence the performance of minority students: teachers are the ones to decide which students gain from extra educational benefits; minority teachers can serve are role models to minority students; and minority teachers might be better able to instruct minority students.

Several studies (Riccucci et al, 2014; Riccucci et al, 2016) show that women are more likely to be responsive to symbolic representation; this study will therefore mainly focus on the effect of representation on female students.

2.2.1 Effect of teachers on student performance

In a study of 67 Florida school districts, Meier and Stewart (1992) examine bureaucratic representation and its impact on the educational performance of minority students. They conclude that minority bureaucrats can impact both public policy outputs, and outcomes. Another notable finding was the greater impact of street-level bureaucrats compared to management-level bureaucrats. Teachers with discretionary decision-making have more influence on the public policy outcomes.

This conclusion is supported by a study by Thomas S. Dee (2004): both black and white students benefit from a teacher of their own race. A possible explanation for this effect is passive representation: simply the presence of a minority teacher as a role model can improve student performance. This means that there are advantages to the recruitment of minority teachers, though with the drawback that this could negatively impact the educational results of non-minority students (Dee, 2004; Dee, 2005). Whether this is the case, is still under contention. A study by Meier et al (1999) shows that non-minority students are not negatively affected by the presence of minority teachers.

Several studies also find that minorities do better at exams when the number of minority teachers increases (Meier, 1993; Pitts, 2005). This leads to the following hypothesis:

Hypothesis 1: *a higher number of female teachers will have a positive effect on the performance of female students*

2.2.2 Moderating effect of representation in management

Keiser, Wilkins, Meier and Holland (2002: 557) fill in a gap in research on the role of gender in bureaucratic representation, by trying to identify under which specific conditions passive representation of gender will lead to active representation. They identify seven conditions that affect the likelihood of passive representation turning into active representation. The first two conditions are always required for linkage to occur. First, there must be room for discretion in how bureaucrats exercise their role, so they have the opportunity to influence organisational outcomes. Second, the policy issue that bureaucrats are concerned with must be a gendered issue; it must be of specific importance to women.

Keiser et al (2002: 557) also identify five other environmental factors that play a role in linking passive representation to active representation: some bureaucratic organisations have a specific focus on advocacy, which makes it easier for bureaucrats to champion issues important for women, making active representation more acceptable. Organisations with a strong hierarchical structure make linking passive and active representation less likely. However, this can be somewhat mitigated by stratification: a greater number of women in higher management positions might lead to a more favourable environment for linkage. Bureaucrats might also be more likely to become active representatives when there is a critical mass of women in an organisation. Finally, the presence of professionals in a bureaucratic organisation can lead to linkage, as professionals integrate their own values into the organisation.

This study will include the factor 'stratification' to measure whether an environmental factor can affect the relationship between bureaucratic representation and student performance. Data on secondary school personnel collected for this study shows that female administrators only make up about a third of total management personnel. As there is clearly a gender discrepancy, measuring the impact of an increase in the number of female administrators can give insights in the importance of women in high management positions. When women gain a position in higher management, they are able to support their own gender by adopting policies that allow female teachers to advocate for female students. This study will test for this moderating effect with the following hypothesis:

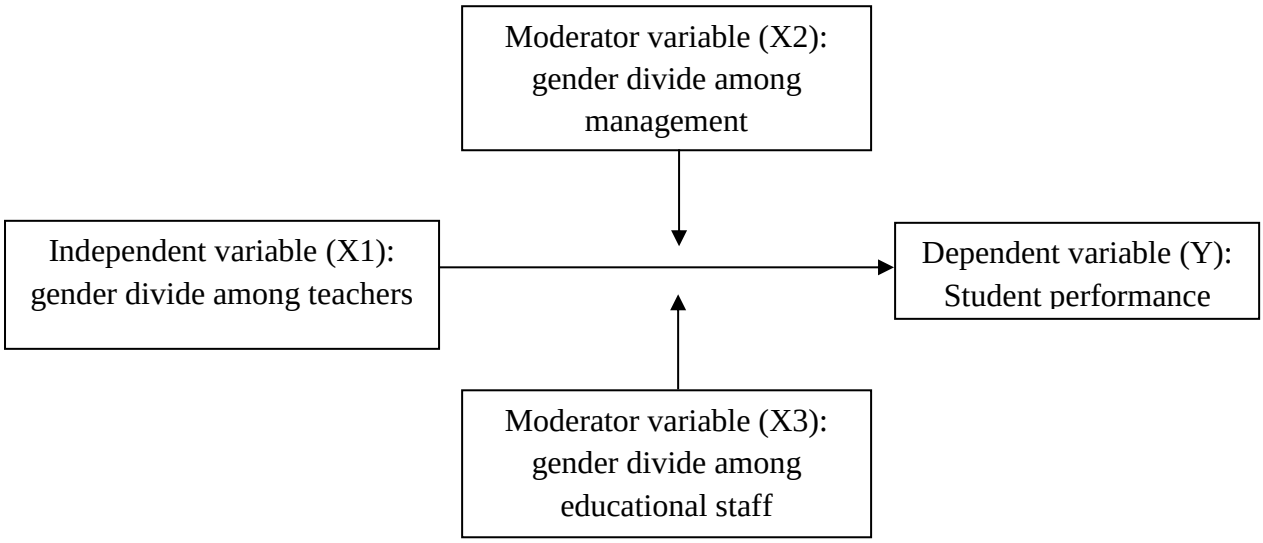
Hypothesis 2: *a higher number of female managers will increase the positive effect of female teachers on the performance of female students*

2.2.3 Moderating effect of educational staff

Not only the gender distribution among teachers and managers, but also the distribution among all educational staff can influence organisational performance. Non-teaching staff can still serve as role models for female students (Keiser et al, 2002: 558). A higher number of women working in secondary schools make it more likely that decisions that will benefit women are positively received and more effectively implemented (O'Toole and Meier, 2015: 248). Female teachers might have more opportunities to support female students, which could benefit the relationship between female teachers and student performance:

Hypothesis 3: *a higher number of educational staff in general will increase the positive effect of female teachers on the performance of female students*

FIGURE 1: conceptual model of the relationship between the dependent, independent and moderator variables



3. Research design

This study aims to examine the relationship between bureaucratic representation and student performance in secondary education. The following chapter will describe the steps that will be taken to answer the research question. It includes the study design and data collection methods and detail the operationalisation of the different variables. To test the hypotheses and answer the research question, this study will include four variables.

The dependent variable is student performance, which will be measured by examining the number of students who have passed their final exams in secondary education. The independent variable is measured by the gender division of the teachers. Two additional moderator variables will be included: the number of women in management positions and the gender distribution of the entire educational staff. The complete codebook can be found in Appendix A.

3.1 Study design

To answer the research question, the appropriate study design is a quantitative study. The benefit of a quantitative study is that it allows for the inclusion of data from all secondary schools, making it more convenient to compare schools. It increases the reliability of the study and makes statistical analysis feasible (Babbie, 2010: 24). It is non-experimental, as the variables will not be manipulated. This study only aims to examine the relationship between the variables.

This study will use a longitudinal research design that measures the same variable at different points in time. It takes a multilevel approach, as it includes nested data of observations per year within the same schools.

One of the advantages of a longitudinal study over a cross-sectional study is its ability to show changes over a period of time. Though it is not well-suited to observe changes in behaviour at the individual level, it can detect changes at a collective level (Taris, 2000: 6). This is appropriate, as this study uses aggregated data instead of individual-level data. This specific longitudinal study is a trend study; it measures the same research unit at multiple points in time (Babbie, 2010: 107).

This study will focus on all secondary schools in the Netherlands, using data from the past five school years: from 2017-2018 until 2021-2022. It is a large-N design, as it includes a large number of observations. The advantage of a large-N design is that even a relatively weak relationship between variables can be detected (Toshkov, 2016: 200).

TABLE 1: number of secondary schools and students in their exam year, per year

	2017-2018	**2018-2019**	**2019-2020**	**2020-2021**	**2021-2022**
Total schools	538	535	534	536	533
Total students	199.721	196.721	193.502	187.245	177.425

3.2 Data collection

This study will rely on secondary quantitative data that has been collected by the Dienst Uitvoering Onderwijs (Duo). The research units for this study are the secondary schools.

The data on student performance was reported to Duo by the secondary schools (Duo, 2022a). Duo has data available from school years 2015-2016 to 2021-2022, which includes the different types of education, the number of exam candidates and the percentage that passed their exams, and the gender of the students (Duo, 2022b).

The data (Duo, 2022c; Duo, 2022d) on the teaching staff - including teachers, supporting personnel and management - has been reported to Duo by the secondary schools, as is legally required. It encompasses personnel information ranging from 2011 to 2021 and is updated annually. The reference date is 1 October. The database includes a large amount of data on the staff at secondary schools, such as age, gender and types of education. It also distinguishes between the number of full-time equivalent (fte) employees and the number of people (Duo, 2022e; Duo, 2022f). For full details on the specific content of the database, refer to Appendix B.

Data about secondary schools is obtained from the database ‘Hoofdvestigingen vo’ (Duo, 2023), which is updated monthly. It contains information on the location, denomination and the types of education offered.

3.3 Operationalisation

Data on educational staff is publicised as both fte and number of people. Theories on bureaucratic representation give no clear lead on which one is preferable. A higher amount of fte leads to more contact hours between students and teachers, so representation might have more impact on student performance. On the other hand, just the passive presence of female staff might already have an effect. As a robustness check, both the amount of fte and the number of female staff will be included.

3.3.1 Dependent variable: student performance

The variable student performance will be measured by looking at the number of students that have passed their final exams. This data is collected from the database 'Examenkandidaten 2017-2022' (Duo, 2022a). Standardised testing is a common way to measure student performance, as it allows for easy comparison between schools (Song, 2018: 351). All schools are required to adhere to the same standards.

The requirements for passing the exams vary by type of education, but the final exam grade is the average of the school exam and central exam grades. The school exam grade is determined by a combination of the exam results from the upper years of secondary school. The school decides the contents and when they will be taken (Duo, n.d.). Central exams are taken at the end of secondary education - usually in May - and are meant to assess the knowledge a student has acquired and retained. Both grades contribute equally to the final grade. Some subjects do not have a central exam, in which case the school exam grade is the final exam grade (DUO, n.d.).

Due to Covid-19 measures, the central exams could not take place in 2020 and the final exam result is based solely on the school exam grade (Rijksoverheid, 2020). In both 2021 and 2022 the central exams did take place but the requirements for passing were more flexible, leading to a higher success rate (Rijksoverheid, 2022; RTLNieuws, 2022).

This study will specifically focus on data from the five school years 2017-2018 to 2021-2022; the school years will be coded from '0' to '4'.

TABLE 2: total number of male and female candidates, per year

	2017-2018	**2018-2019**	**2019-2020**	**2020-2021**	**2021-2022**
Candidates:					
male	99588	97733	96623	93813	99554
female	100133	98988	96879	93432	88870
total	199721	196721	193502	187245	177425
Graduates:					
male	91836	90217	95088	89147	83858
female	91634	90832	96000	88525	83546
total	183470	181049	191088	177672	167405

3.3.2 Independent variable: gender divide among teachers

The impact of the number of female teachers on student performance will be measured by assessing the proportion of women; specifically, the ratio of female teachers to female students in every school. The information on teaching staff in secondary schools is collected from two databases: 'Onderwijspersoneel vo in aantal personen' (Duo, 2022c) and 'Onderwijspersoneel vo in aantal fte' (Duo, 2022d).

The database shows the information on teaching staff based on the astronomical year instead of the school year. The reference date is 1 October, half a year before the final test and therefore not aligned with the data on student performance, which is collected per school year. The data from the first half of the school year will be used to measure the variable of the teaching staff. These are the five years from 2017 to 2021. This study will focus on the influence of gender on student performance. The variable 'gender' is divided into either 'male' or 'female'. Teaching staff can be divided in both the amount of fte and the number of people that are employed.

3.3.3 Moderator variable: representation in management

Based on the assumption of Keiser et al. (2002: 557) that active representation will increase if there are more women in higher positions, this study will include this moderating effect by analysing the distribution of male and female personnel in management positions, both in fte and number of people. The data on management personnel in secondary schools is obtained from two databases: 'Onderwijspersoneel vo in aantal personen' (Duo, 2022c) and 'Onderwijspersoneel vo in aantal fte' (Duo, 2022d).

Due to privacy regulations, the database does not disclose the number of management personnel if this is less than five. As a consequence, not all data from secondary schools can be included in the analysis. This means that the data included is more likely to be comprised of schools with larger management bodies. It is also biased towards extreme cases, as schools with an unequal gender distribution in the management body are included, while schools with an equal distribution are not. To address this bias, the variable will be measured in two ways:

- a *limited* variant will solely include schools with a sufficiently large management body to report both male and female managers. The number of cases will be smaller, but it will eliminate the bias towards extremes.
- a *comprehensive* variant will include all available data. When only the total number of managers is provided, it will be assumed that the gender distribution is equal. This will lead to a larger number of cases, but also greater uncertainty.

TABLE 3: number of schools included in the limited and comprehensive variant, per year

	2017-2018	2018-2019	2019-2020	2020-2021	2021-2022
Schools included (limited)	39	41	47	43	46
Schools included (comprehensive)	210	205	208	191	193

3.3.4 Moderator variable: gender divide among educational staff

The final variable to be measured is the gender distribution among the entire educational staff. This does not only include teachers and managers, but also supporting staff and trainee teachers, both in number of fte and number of people. The data is obtained from two databases: 'Onderwijspersoneel vo in aantal personen' (Duo, 2022c) and 'Onderwijspersoneel vo in aantal fte' (Duo, 2022d).

3.3.5 Control variables

A large number of other factors can influence student performance. To establish nonspuriousness, this study will include six control variables:

Class size: some research suggests that class size can influence student performance (Arias and Walker, 2004). This variable will be measured by calculating the ratio of teachers (in fte) to students in their exam year.

School denomination: the Dutch school system includes both government-funded public schools, and special schools which offer education based on a specific religion, ideology or pedagogic concept (CBS, n.d.). The different ways education is organised in public and special secondary schools could affect student performance. Secondary schools will be coded as either 'public' or 'special'.

Income: some studies suggest that a student's performance could be affected by their socio-economic background (CPB, 2020). To measure its possible impact, this study uses the standardised disposable income per municipality divided into five categories (CBS, 2020): less than 27 000, 27 000 - 31 000, 31 000 - 35 000, 35.000 - 39 000, and more than 39 000. These are coded 0 to 4, with the first category as the reference category.

Region: the Netherlands has both urban and rural areas, which may impact student performance. Urban and rural schools can differ in staff size, resources and student background (McCracken and Barcinas, 1991). Secondary schools will be categorised in one of five categories - coded 0 to 4 - based on the population density of the municipality in which they

are located (CBS, 2023b; CBS, 2023c): not urban, little urban, moderately urban, highly urban and very highly urban, with the first category as the reference category.
Teaching experience: several studies show that a teacher's experience can have a positive effect on student performance (Coenen et al, 2018). Older teachers are likely to have more experience, so the average age of teachers will be included as a control variable.
Type of education: secondary schools in the Netherlands can choose to offer either a single type or multiple types of education, and this distinction might impact student performance. To categorise schools, each school is designated as either 'general' or 'specialised'.

The data required to measure these control variables is collected from the following databases: 'Examenkandidaten 2017-2022' (Duo, 2022a), 'Onderwijspersoneel vo in aantal personen' (Duo, 2022c), 'Onderwijspersoneel vo in aantal fte' (Duo, 2022d) and 'Hoofdvestigingen vo' (Duo, 2023).

TABLE 4: operationalisation of dependent and independent variables

Dependent variable	*Graduation percentage of female students*
Independent variable	*Number of female students per female teacher*, both in fte and number of people
Moderator variables	1. *Number of female managers*, both in fte and number of people
	limited: solely includes schools with a sufficiently large management body to report both male and female managers
	comprehensive: includes all available data
	2. N*umber of female educational staff*, both in fte and number of people
Control variables	*Class size*: the ration of teachers (in fte) to students in their exam year
	School denomination: schools are designated as either 'public' or 'special'
	Income: standardised disposable income per municipality, divided into five categories: less than 27 000, 27 000 - 31 000, 31 000 - 35 000, 35.000 - 39 000, and more than 39 000
	Region: population density of the municipality in which schools are located, divided into five categories: not urban, little urban, moderately urban, highly urban and very highly urban
	Teaching experience: average age of teachers, both female and male
	Type of education: schools are designated as either 'general or 'specialised'

4. Results

The previous chapter detailed the steps that need to be taken to test the hypotheses. This chapter will show the results of following these steps by using the data science software program Stata. The first section of this chapter shows the descriptive statistics of the variables included in this study. The second section will focus on the correlations between the variables. The third section will feature a multilevel regression analysis by using a mixed-effects linear regression model, which will test the impact of the ratio of female teachers on the performance of female students. It will also show the impact of the moderator variables on the relationship between the dependent and independent variable. The syntax for this analysis is shown in Appendix C. For an overview of a similar analysis measuring the effect of bureaucratic representation on the performance of male students, see Appendix D.

4.1 Descriptive statistics

Descriptive statistics provide some insights into the basic characteristics and the distribution of the data. This can be useful for summarising and exploring the individual variables (Babbie, 2010: 426; Healey, 2021: 7). All independent variables were measured in both fte and number of people.

TABLE 5: descriptive statistics for all variables included in the analysis

	N	Mean	St. Dev.	Var.	Min	Max
Dependent variable						
1. Graduation percentage of female students (%)	2 671	94.15	5.46	29.83	0	100
Independent/moderator variables in fte						
2. Number of female students per female teacher	2 489	3.14	.93	.86	0	8.70
3a. Number of female manager (limited)	217	6.99	2.15	4.63	3.04	13.61
3b. Number of female managers (comprehensive)	1 006	3.47	2.35	5.51	0	13.61
4. Number of female educational staff	2 502	79.56	50.51	2551.47	2.88	323.78
Independent/moderator variables in people						
5. Number of female students per female teachers	2 489	2.27	.64	.42	0	5
6a. Number of female managers (limited)	215	7.62	2.36	5.55	5	15
6b. Number of female managers (comprehensive)	1 007	3.74	2.56	6.57	0	15
7. Number of female educational staff	2 502	111.43	69.61	4845.97	5	429
Control variables						
8. Class size	2 486	2.46	.57	.33	.17	6.18
9. School denomination	2 550	.70	.46	.21	0	1
10. Income	2 570	1.79	.81	.65	0	4
11. Region	2 625	2.77	1.14	1.29	0	4
12. Average age of female teachers	2 550	41.82	6.38	40.71	0	55.78
13. Type of education	2 750	.23	.42	.18	0	1

Note: rounded to the second decimal

This table shows the number of observations per variable, the mean, the standard deviation and the range of the data. The dependent variable (Y) shows the mean at 94.15% with a standard deviation of 5.46%. The values fall close to the mean, within a relatively small range of outcomes. The independent variable (X) has a higher standard deviation, which means that the data is more spread out.

The moderator variable 'number of female managers' is measured with both a limited and comprehensive variant. The limited variant involves a smaller sample size, but eliminates the bias towards extreme cases. The comprehensive variant has a larger number of cases, but relies on more assumptions. The table shows a high similarity between the standard deviations of the two different versions, suggesting that the measurements of both variants are reliable.

The moderator variable 'number of female educational staff' has a very high standard deviation and variance, suggesting that the data is widely spread out. The number of educational staff differs greatly between schools.

4.2 Correlation analysis

A correlation analysis is used to determine the relationship between two variables (Babbie, 2010: 436). The measure of association used in this analysis is the Pearson's correlation coefficient. The results are shown in table 6 on the next page.

The table does not show a significant correlation between the number of female teachers and the performance of female students. Interestingly, it does show a moderate positive correlation between the number of female managers and the number of female educational staff, suggesting that an increase in one variable will lead to an increase in the other variable. A similar analysis measuring the relationship between the number of male managers and male educational staff, gives a comparable result (see table A2 in Appendix D).

TABLE 6: correlations for all variables in the analysis

	1	2	3a	3b	4	5	6a	6b	7	8	9	10a	10b	10c	10d	10e	11a	11b	11c	11d	11e	12	13
1	1.00																						
2	-.01	1.00																					
3a	-.03	-.04	1.00																				
3b	-.02	-.18*	1.00*	1.00																			
4	-.01	-.10*	**.49***	**.45***	1.00																		
5	-.00	.96*	.00	-.13*	-.07*	1.00																	
6a	-.08	-.01	.91*	.61*	**.17***	.02	1.00																
6b	-.07*	-.14*	.65*	.84*	**.35***	-.11*	1.00*	1.00															
7	-.00	-.06*	**.46***	**.40***	.99*	-.05*	**.14***	**.31***	1.00														
8	-.00	.78*	.00	-.04	-.04	.85*	-.01	-.04	-.04	1.00													
9	.06*	.10*	.09	-.01	.06*	.07*	.10	-.04	.08*	.03	1.00												
10a	-.03	.02	.18*	.06	-.00	.04*	.13	.04	-.01	-.01	.01	1.00											
10b	-.05*	-.09*	.02	-.05	.04*	-.09*	-.05	.01	.04	-.10*	-.08*	-.10*	1.00										
10c	.03	.07*	-.05	.04	.04	.07*	.08	.04	.05*	.04*	-.01	-.13*	-.63*	1.00									
10d	.03	.05*	-.08	.01	-.09*	.05*	-.06	-.03	-.09*	.10*	.09*	-.05*	-.23*	-.30*	1.00								
10e	.03	-.00	.05	.03	-.01	-.01	-.14*	-.06	-.02	.00	.09*	-.03	-.14*	-.18*	-.06*	1.00							
11a	.03	-.03	-.13	-.05	-.13*	-.01	-.05	-.02	-.14*	-.01	-.06*	-.03	.05*	-.00	-.03	-.04	1.00						
11b	.07*	.12*	-.21*	-.05	-.04	.13*	-.23*	-.09*	-.03	.14*	.07*	-.06*	-.09*	.07*	.03	.07*	-.07*	1.00					
11c	.04*	.12*	.10	.01	.08*	.13*	.19*	-.00	.09*	.16*	.06*	-.06*	-.11*	-.05*	.18*	.16*	-.08*	-.17*	1.00				
11d	.02	.06*	.02	.06	.13*	.06*	-.14*	.03	.14*	.06*	-.02	.20*	-.06*	.11*	-.02	-.02	-.13*	-.28*	-.30*	1.00			
11e	-.10*	-.21*	.04	-.01	-.09*	-.24*	.14*	.05	-.11*	-.27*	-.05*	-.09*	.28*	-.04*	-.11*	-.13*	-.12*	-.26*	-.28*	-.41*	1.00		
12	-.03	-.07*	-.08	.04	-.10*	-.06*	-.07	-.04	-.11*	-.01	-.07*	.04*	.04	-.01	-.01	.01	.02	.05*	.02	-.05*	.03	1.00	
13	.15*	-.19*	-.18*	-.09	-.45*	-.19*	.11	-.06*	-.45*	-.12*	-.06*	.08*	-.07*	-.03	-.02	.03	-.00	-.09*	-.14*	-.00	.07*	-.02	1.00

Note: significance level*=p<0.05; rounded to the second decimal

TABLE 6: continued

	Dependent variable
1	Graduation percentage of female students
	Independent variables in fte
2	Number of female students per female teacher
3a	Number of female managers (limited)
3b	Number of female managers (comprehensive)
4	Number of female educational staff
	Independent variables in number of people
5	Number of female students per female teacher
6a	Number of female managers (limited)
6b	Number of female managers (comprehensive)
7	Number of female educational staff
	Control variables
8	Class size
9	School denomination
10a	Income: less than 27 000
10b	Income: 27 000 - 31 000
10c	Income: 31 000 - 35 000
10d	Income: 35 000 – 39 000
10[e]	Income: more than 39 000
11a	Region: not urban
11b	Region: little urban
11c	Region: moderately urban
11d	Region: highly urban
11e	Region: very highly urban
12	Teaching experience
13	Type of education

4.3 Multilevel regression analysis

A multivariate analysis allows for measuring the relationship between more than two variables. The variables are analysed with the use of data science program Stata.

Model 0 (on the next page) is the base model and includes only the dependent variable and the time variable. The intraclass correlation coefficient is 0.1708. This indicates that 17.08% of the variation in the dependent variable ‘graduation percentage of female students’ can be explained by the factor time. The other 82.92% can be explained by differences between the schools. A low intraclass correlation coefficient means there is little variance between the groups: different school years do not have much of an impact on student performance. A multilevel analysis could be gratuitous.

Model 1 adds the independent, moderator and control variables to the model. For the variable ‘number of female managers’ only the comprehensive variant is included, as the

analysis showed that including both variants would be redundant. The model has been split into an a and b version, to account for the measurements of staff in both fte and number of people.

TABLE 7: multilevel regression analysis of the graduation percentage of female students

	Model 0	**Model 1a**	**Model 1b**
Year	.878*	.847*	.859*
Number of female students per female teacher		.379	.434
Number of female managers (comprehensive)		-.079	-.091
Number of female educational staff		.005	.002
Class size		.209	.226
School denomination		.632	1.053*
Income: 27 000 - 31 000		-.082	-.053
Income: 31 000 - 35 000		-.250	-.161
Income: 35 000 - 39 000		.213	.475
Income: more than 39 000		.058	-.055
Region: little urban		.529	.492
Region: moderately urban		.581	.329
Region: highly urban		-.264	-.077
Region: very highly urban		-1.597	-1.806
Teaching experience		-.208*	-.213*
Type of education		4.160*	3.825*
Constant	91.513*	98.049*	98.402*
N	2 671	1 000	997

Note: significance level*=p<0.05; rounded to the third decimal
Model 1a shows personnel in fte, Model 1b shows number of people
Reference categories are ‘Income: less than 27 000’ and ‘Region: not urban’

Models 1a and 1b show a coefficient of 0.379 and 0.434 respectively, a moderate positive effect: an increase of one in the variable ‘number of female students per female teacher’ will increase the graduation percentage by ≈0.4%. The number of female teachers has a moderately positive influence on the student performance of female students.

The effects of both moderator variables are less than 0.10 and considered very weak. The moderator variable ‘number of female managers’ has a weak negative effect on student performance: an increase of one female manager in either fte or number of people leads to a decrease of the graduation percentage of 0.08% or 0.09% respectively. The moderator variable ‘number of female educational staff’ has a coefficient of 0.005 or 0.002; an increase of one has a negligible amount of effect on the graduation percentage. The results are also not statistically

significant, and could be the result of chance. As the model does not show a statistically significant effect of the number of female teachers on the performance of female students, the first hypothesis can be discarded.

Model 2 and 3 include the moderator variables to measure the interaction effects on the relationship between the dependent and independent variable. The model is again split into an a and b version, to account for the measurements of staff in both fte and number of people.

TABLE 8: multilevel regression analysis of the graduation percentage of female students including moderator variables

	Model 2a	**Model 2b**	**Model 3a**	**Model 3b**
Year	.849*	.852*	.850*	.850*
Number of female students per female teacher	.292	.334	-.317	-.212
Number of female managers (comprehensive)	-.076	-.132	.	.
Number of female managers (dummy)	-.256	-.208	.	.
Number of female educational staff	.	.	-.025*	-.016*
Number of female students per female teacher*Number of female managers	.010	.022	.	.
Number of female students per female teacher*Number of female educational staff	.	.	**.011***	**.010***
Class size	.150	.165	-.617	-.729
School denomination	.673	1.105*	.562	.564
Income: 27 000 - 31 000	-.074	-.041	-.157	-.146
Income: 31 000 - 35 000	-.229	-.160	-.114	-.095
Income: 35 000 - 39 000	.136	.445	.327	.358
Income: more than 39 000	.068	-.037	.465	.436
Region: little urban	.602	.552	-.061	-.117
Region: moderately urban	.701	.418	-.505	-.547
Region: highly urban	-.119	.038	-1.210	-.1.235
Region: very highly urban	-1.502	-1.711	-2.117*	-.2.139*
Teaching experience	-.217*	-.222*	-.061	-.062*
Type of education	3.895*	3.627*	3.109*	3.150*
Constant	99.270*	99.497*	96.145*	95.918*
N	1 000	995	2 451	2 451

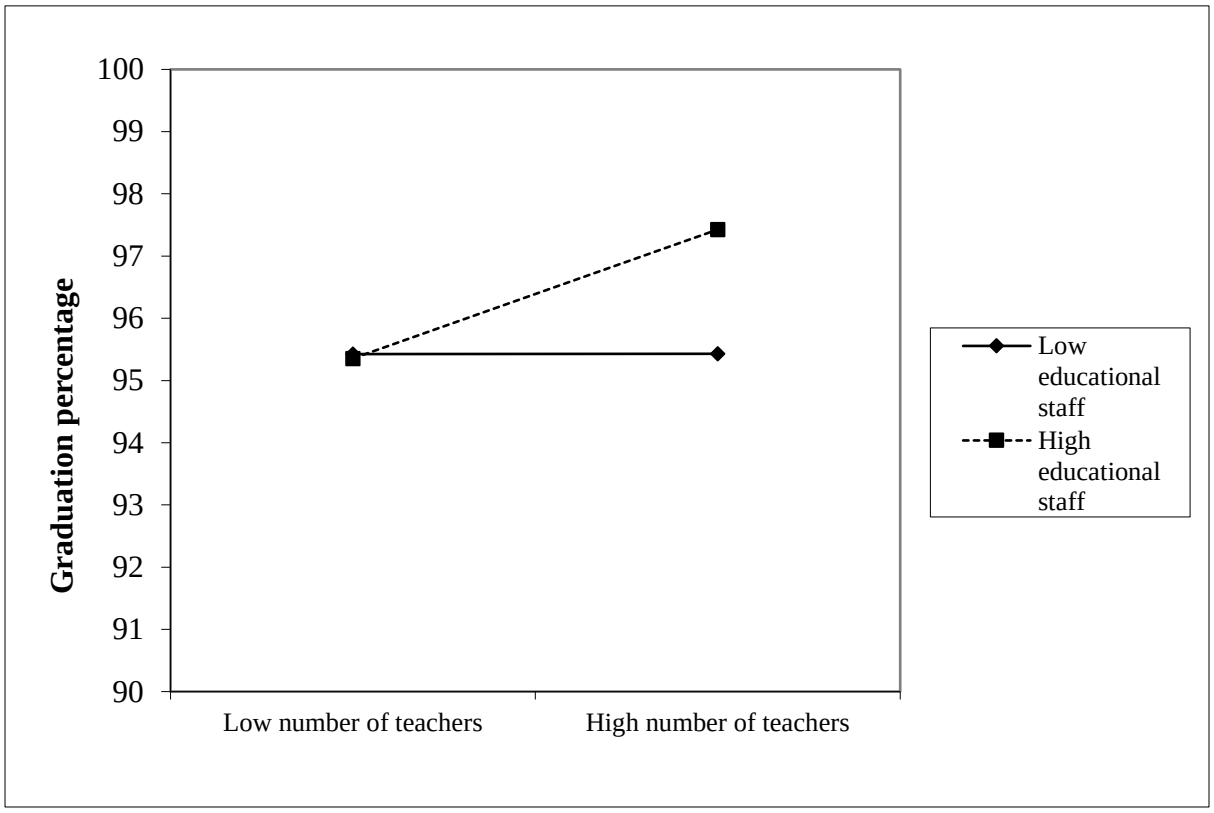
Note: significance level*=p<0.05; rounded to the third decimal
Models 2a and 3a show personnel in fte, Models 2b and 3b show number of people
Reference categories are ‘Income: less than 27 000’ and ‘Region: not urban’

The second model measure the effect of the interaction between the 'number of female students per female teacher' and 'number of female managers'. The coefficients of the variables show little differences, compared to model 1. The results show a weak interaction effect of the moderator variable. However, the result is not statistically significant, so the second hypothesis can be discarded.

Model 3 shows a weak but significant interaction effect of 0.011 and 0.010, supporting the third hypothesis. The moderator variable 'number of female educational staff' has a positive effect on the relationship between the independent and the dependent variable. A school with a higher number of female educational staff increases possible benefits of female teachers on female student performance.

This effect is depicted in figure 2, which shows the moderating effect of model 3a.

FIGURE 2: interaction plot for moderator variable 'number of educational staff'



Note: y-axis shows the graduation percentage between 90 and 100%; x-axis shows the number of female students per female teacher; graph line shows the number of female educational staff
Low is the mean minus standard deviation; high is the mean plus standard deviation

Interestingly, the independent variable 'number of female students per female teacher' on its own is not significant. The effect of the independent variable is contingent on the presence of the moderator.

A similar analysis focused on the effect of male teachers on male student performance shows the opposite effect (see table A4 in Appendix D). The moderator variable ‘number of male educational staff’ has a very weak but significant negative effect on the relationship between the dependent and independent variable. This means that a higher number of male educational staff lessens possible positive effects of male teachers on male student performance.

5. Discussion and conclusion

This first section of this chapter discusses the results of the analysis and its implication. The second section discusses the limitations of this study and the third section will summarise the main findings and answer the research question.

5.1 Discussion

To answer the research question of this thesis, three hypotheses were formulated and then tested using a multilevel regression analysis. Student performance was measured with the graduation rate per school, as standardised tests are often used in related studies (Song, 2018: 351). As a robustness check, the gender disparity of teaching staff was measured in both fte and number of people. The results showed little difference between the two measures, implying both measures are valid. Using models, it was possible to analyse data from all Dutch secondary schools and measure the impact of the gender of a teacher on student performance.

The results of the correlation analysis show a moderate positive correlation between the number of female managers and the number of female educational staff, suggesting that an increase in one of these variables is associated with an increase in the other variable. A similar relationship is present between the number of male managers and the number male educational staff (see table A2 in Appendix D). However, a correlation analysis does not indicate the direction of the relationship. A possible explanation for the positive correlation can be found in the literature: a higher number of female managers may contribute to an increase in the number of female educational staff. Managers often possess some discretionary powers (Selden, 1997a: 13-15), so it is possible these are utilised to implement policies benefitting the educational staff of their own gender.

The results of the first two models of the multilevel regression analysis were not statistically significant, thus the first and second hypothesis can be discarded. The third model did show significant results, providing support for the third hypothesis. It suggests that a higher number of female educational staff has a positive influence on the relationship between female teachers and student performance. A literature-based explanation for this result comes from O'Toole and Meier (2015: 248), who propose that a higher number of female staff can make it easier for managers to implement policies benefiting their group.

A similar analysis was conducted showing the effect of male teachers on the performance of male students and showed contrasting results. The model showed a very weak but statistically significant negative effect, meaning that a higher number of male educational staff negatively affected the relationship between the number of male teachers and male student performance. This is in line with conclusions from other studies (Riccucci et al, 2014; Riccucci et al, 2016), which demonstrate that women are more responsive to bureaucratic representation than men.

5.2 Limitations

Researchers in this subject commonly rely on aggregated data gathered at the organisational level, which poses challenges in measuring the attitude and behaviour of individuals (Bradbury and Kellough, 2011: 163). Aggregated data makes it difficult to differentiate between active and symbolic representation, as it is challenging to determine whether organisational outcomes are the result of actions taken by bureaucrats or by the reaction of citizens. Additionally, it is also difficult to ascertain that there are no other mechanisms at work (Theobald and Haider-Markel, 2009: 410). To control for spuriousness, this study included six control variables: class size, school denomination, income, region, teaching experience and type of education.

To increase reliability, this study includes all available data from Dutch secondary schools over a five-year period. Large sample sizes minimise the impact of outliers and decrease the margin of error. However, by necessity, these five years include several years of Covid-19 measures, which may have impacted teaching practices and student performance, potentially skewing the research results.

Several control variables have been included to increase internal validity. There may however be other possible confounding variables that fall beyond the scope of this study. For example, data on the educational level of a student's parents is not publicly available.

5.3 Conclusion

This thesis tried to determine the potential effect of bureaucratic representation on student performance by answering the following research question: *To what extent has bureaucratic representation of gender an effect on student performance in secondary schools?* Education is a key factor in social mobility and researching the role of bureaucratic representation in education could potentially benefit many students. Studies have shown the positive effects of the presence of minority teachers on the performance of minority students. However, the results

of the analysis for this study were not statistically significant. The first hypothesis had to be discarded.

Keiser et al (2002) identify stratification as an environmental factor linking passive and active representation, suggesting that a higher number of female managers could positively influence the relationship between teachers and student performance. Yet the results of this study did not show a statistically significant result and the second hypothesis had to be discarded.

The analysis did provide support for the third hypothesis, suggesting that a higher number of female educational staff has a positive influence on the relationship between the number of female teachers and student performance. Curiously, this effect is only present in the analysis of female student performance.

Overall, the results of this study showed only a very limited effect of bureaucratic representation of gender on student performance.

However, it does present several potential avenues of research. Results show a moderate correlation between the gender division of management personnel and educational staff for both men and women. A possible literature-based explanation is that managers may have a (subconscious) preference for staff of their own gender. Individual-level research is recommended to find a definitive answer.

Additionally, as the evidence for the effect of bureaucratic representation of gender on student performance overall is quite weak, an interesting avenue of research could be to examine a gendered educational area, like calculus. In primary school, boys consistently outperform girls in calculus and this trend continues in later years (Zumbuehl and Dillingh, 2020: 6-7). Keiser et al (2002: 556) argue that a gendered subject is a necessary condition for linkage between passive and active representation, so a study of a specifically gendered educational area could lead to different results.

Another possible avenue of future research is a focus on the effect of bureaucratic representation on the performance of students with a non-Western immigration background, which also shows a performance gap (Zumbuehl and Dillingh, 2020: 10-11). As there is no aggregated public data on the ethnic background of children and teachers, a study like this would need to be done on an individual level

Despite female students pursuing tertiary education in larger numbers than their male counterparts, the data in this study does not indicate a notable difference in graduation rates between genders. As this study only examines graduation percentages and does not consider

the level of education of a student, it is likely that female students participate in higher education more often. A possible future study could include the level of education as a variable, as this could play a role in how students are affected by representation. Many studies on bureaucratic representation are conducted in the United States where there are no educational levels in secondary schools, so research on this subject is quite limited.

Finally, this study only includes secondary schools. A similar study focused on kindergartens or primary schools might have different results, as younger children could be impacted differently by bureaucratic representation.

This study also has practical implications. The results show the benefits of a higher number of female educational staff on student performance, while the effect of a higher number of male educational staff is negligible. Though the relationship is quite weak, it should still be taking into consideration by school administrators.

Furthermore, the correlation between the gender division of management personnel and its effect on educational staff shows the importance of striving for equality in management positions. As the data shows, currently only about a third of management personnel is female. Increasing this number benefits female educational staff and makes the school administration a better reflection of the populace.

.

List of references

Articles and books

Arias, J.J. and Walker, D.M., 2004. Additional evidence on the relationship between class size and student performance. *The Journal of Economic Education*, 35(4), p311-329.

Babbie, E., 2010. *The practice of social research.* (12th ed.). Wadsworth: Cengage Learning.

Bradbury, M.D. and Kellough, J.E., 2008. Representative bureaucracy: exploring the potential for active representation in local government. *Journal of Public Administration Research and Theory*, 18(4), p697-714.

Bradbury, M.D. and Kellough, J.E., 2011. Representative bureaucracy: assessing the evidence on active representation. *The American Review of Public Administration*, 41(2), p157-167.

Choi, S. and Rainey, H.G., 2010. Managing diversity in U.S. federal agencies: effects of diversity and diversity management on employee perceptions of organizational performance. *Public Administration Review*, 70(1), p109-121.

Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H. and Van Klaveren, C., 2018. Teacher characteristics and their effects on student test scores: a systematic review. *Journal of Economic Surveys*, 32(3), p848-877.

Dauda, B., 1990. Fallacies and dilemmas: the theory of representative bureaucracy with a particular reference to the Nigerian public service 1950-1986. *International Review of Administrative Sciences*, 56(3), p467-495.

Dee, T.S., 2004. Teachers, race, and student achievement in a randomized experiment. *The Review of Economics and Statistics*, 86(1), p195-210.

Dee, T.S., 2005. A teacher like me: does race, ethnicity, or gender matter. *The American Economic Review*, 95(2), p158-165.

Ding, F., Lu, J. and Riccucci, N.M., 2021. How bureaucratic representation affects public organizational performance: a meta-analysis. *Public Administration Review*, 81(6), p1003-1018.

Grissom, J.A., Kern, E.C. and Rodriguez, L.A., 2015. The 'representative bureaucracy' in education: educator workforce diversity, policy outputs, and outcomes for disadvantaged students. *Educational Researcher*, 44(3), p185-192.

Groeneveld, S. and Van de Walle, S., 2010. A contingency approach to representative bureaucracy: power, equal opportunities and diversity. *International Review of Administrative Sciences*, 76(2), p239-258.

Healey, J.F., 2012. *Statistics: a tool for social research*. (9th ed.). Wadsworth: Cengage Learning.

Henderson, L.J., 1978. Administrative advocacy and black urban administrators. *The Annals of the American Academy of Political and Social Science*, 439, p68-79.

Hong, S., 2016. Representative bureaucracy, organizational integrity, and citizen coproduction: does an increase in police ethnic representativeness reduce crime? *Journal of Policy Analysis and Management*, 35(1), p11-33.

Keiser, L.R., Wilkins, V.M., Meier, K.J. and Holland, C.A., 2002. Lipstick and logarithms: gender, institutional context, and representative bureaucracy. *The American Political Science Review*, 96(3), p553-564.

Kennedy, B., 2014. Unraveling representative bureaucracy: a systematic analysis of the literature. *Administration & Society*, 46(4), p395-421.

Kingsley, J.D., 1944. *Representative bureaucracy: an interpretation of the British civil service*. Yellow Springs: Antioch Press.

Lim, H.-H., 2006. Representative bureaucracy: rethinking substantive effects and active representation. *Public Administration Review*, 66(2), p193-204.

McCracken, J.D. and Barcinas, J.D.T., 1991. Differences between rural and urban schools, student characteristics, and student aspirations in Ohio. *Journal of Research in Rural Education*, 7(2), p29-40.

Meier, K.J., 1975. Representative bureaucracy: an empirical analysis. *The American Political Science Review*, 69(2), p526-542.

Meier, K.J., 1993. Latinos and representative bureaucracy: testing the Thompson and Henderson hypotheses. *Journal of Public Administration Research and Theory*, 3(4), p393-414.

Meier, K.J., 2019. Theoretical frontiers in representative bureaucracy: new directions for research. *Perspectives on Public Management and Governance*, 2(1), p39-56.

Meier, K.J., Andersen, S.C., O'Toole, L.J., Favero, N. and Winter, S.C., 2015. Taking managerial context seriously: public management and performance in U.S. and Denmark schools. *International Public Management Journal*, 18(1), p130-150.

Meier, K.J. and Bohte, J., 2001. Structure and discretion: missing links in representative bureaucracy. *Journal of Public Administration Research and Theory*, 11(4), p455-470.

Meier, K.J. and Stewart Jr., J., 1992. The impact of representative bureaucracies: educational systems and public policies. *The American Review of Public Administration*, 22(3), p157-171.

Meier, K.J., Wrinkle, R.D. and Polinard, J.L., 1999. Representative bureaucracy and distributional equity: addressing the hard question. *The Journal of Politics*, 61(4), p1025-1039.

Mosher, F.C., 1968. *Democracy and the public service*. New York: Oxford University Press.

O'Toole, L.J. and Meier, K.J., 2015. Public management, context, and performance: in quest of a more general theory. *Journal of Public Administration Research and Theory*, 25(1), p237-256.

Peters, B.G., 2010. *The politics of bureaucracy*. (6th ed.). London: Routledge.

Pitts, D.W., 2005. Diversity, representation, and performance: evidence about race and ethnicity in public organizations. *Journal of Public Administration Research and Theory*, 15(4), p615-631.

Rainey, H.G., 2009. *Understanding and managing public organizations: essential texts for nonprofit and public leadership and management*. (4th ed.). New Jersey: A Wiley Imprint.

Riccucci, N.M., Van Ryzin, G.G. and Lavena, C.F., 2014. Representative bureaucracy in policing: does it increase perceived legitimacy? *Journal of Public Administration Research and Theory*, 24(3), p537-551.

Riccucci, N.M., Van Ryzin, G.G. and Li, H., 2016. Representative bureaucracy and the willingness to coproduce: an experimental study. *Public Administration Review*, 76(1), p121-130.

Riccucci, N.M. and Van Ryzin, G.G., 2017. Representative bureaucracy: a lever to enhance social equity, coproduction, and democracy. *Public Administration Review*, 77(1), p21-30.

Selden, S.C., 1997a. *The promise of representative bureaucracy: diversity and responsiveness in a government agency*. Armonk: M.E. Sharpe.

Selden, S.C., 1997b. Representative bureaucracy: examining the linkage between passive and active representation in the Farmers Home Administration. *The American Review of Public Administration*, 27(1), p22-42.

Song, M, 2018. Gender representation and student performance: representative bureaucracy goes to Korea. *The American Review of Public Administration*, 48(4), p346-358.

Taris, T.W., 2000. *A primer in longitudinal data analysis*. London: SAGE Publications.

Theobald, N.A. and Haider-Markel, D.P., 2009. Race, bureaucracy, and symbolic representation: interactions between citizens and police. *Journal of Public Administration Research and Theory*, 19(2), p409-426.

Thompson, F.J., 1976. Minority groups in public bureaucracies: are passive and active representation linked? *Administration & Society*, 8(2), p201-226.

Toshkov, D., 2016. *Research design in political science.* New York: Palgrave.

Websites

CBS, n.d. *Denominatie (onderwijs)*. Centraal Bureau voor Statistieken. Retrieved from: https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen/denominatie--onderwijs-- (9-6-2023)

CBS, 2020. *Inkomen van huishoudens*. Centraal Bureau voor Statistieken. Retrieved from: https://longreads.cbs.nl/materiele-welvaart-in-nederland-2022/inkomen-van-huishoudens/ (9-6-2023)

CBS, 2022. *Loonverschil tussen mannen en vrouwen verder afgenomen*. Centraal Bureau voor Statistieken. Retrieved from: https://www.cbs.nl/nl-nl/nieuws/2022/17/loonverschil-tussen-mannen-en-vrouwen-verder-afgenomen (9-6-2023)

CBS, 2023a. *Al 23 jaar op rij meer vrouwen dan mannen in hoger onderwijs*. Centraal Bureau voor Statistieken. Retrieved from: https://www.cbs.nl/nl-nl/nieuws/2023/10/al-23-jaar-op-rij-meer-vrouwen-dan-mannen-in-hoger-onderwijs (9-6-2023)

CBS, 2023b. *Stedelijkheid*. Centraal Bureau voor Statistieken. Retrieved from: https://www.cbs.nl/nl-nl/nieuws/2019/44/meeste-afval-per-inwoner-in-minst-stedelijke-gemeenten/stedelijkheid (9-6-2023)

CBS, 2023c. *Gebieden in Nederland 2023*. Centraal Bureau voor Statistieken. Retrieved from: https://www.cbs.nl/nl-nl/cijfers/detail/85385NED (9-6-2023)

Duo, n.d. *Examenloket*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/zakelijk/voortgezet-onderwijs/examens-en-diplomas/examenloket/schoolexamen.jsp (9-6-2023)

Duo, 2022a. *Examens*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/open_onderwijsdata/voortgezet-onderwijs/aantal-leerlingen/examens.jsp [data set] (9-6-2023)

Duo, 2022b. *Aantal examenkandidaten en geslaagden in het voortgezet onderwijs*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/open_onderwijsdata/images/toelichting-06-examenkandidaten-en-geslaagden.pdf (9-6-2023)

Duo, 2022c. *Onderwijspersoneel vo in aantal personen*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/open_onderwijsdata/voortgezet-onderwijs/personeel/in-aantal-personen.jsp [data set] (9-6-2023)

Duo, 2022d. *Onderwijspersoneel vo in aantal fte*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/open_onderwijsdata/voortgezet-onderwijs/personeel/in-aantal-fte.jsp [data set] (9-6-2023)

Duo, 2022e. *Onderwijspersoneel in het vo in aantal fte*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/open_onderwijsdata/images/toelichting-02.-personeel-vo-in-fte.pdf [pdf] (9-6-2023)

Duo, 2022f. *Onderwijspersoneel in het vo in aantal personen*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/open_onderwijsdata/images/toelichting-01.-personeel-vo-in-personen.pdf [pdf] (9-6-2023)

Duo, 2023. *Adressen hoofdvestigingen*. Ministerie van Onderwijs, Cultuur en Wetenschap. Retrieved from: https://duo.nl/open_onderwijsdata/voortgezet-onderwijs/adressen/hoofdvestigingen.jsp [data set] (9-6-2023)

Data browser Eurostat, 2023. *Gender pay gap in unadjusted form*. Eurostat. Retrieved from: https://ec.europa.eu/eurostat/databrowser/view/sdg_05_20/default/table?lang=en [data set] (9-6-2023)

Rijksoverheid, 2020. *Nieuwsbericht: geen centrale examens dit jaar*. Rijksoverheid. Retrieved from: https://www.rijksoverheid.nl/actueel/nieuws/2020/03/24/besluit-centrale-eindexamens (9-6-2023)

Rijksoverheid, 2022. *Nieuwsbericht: extra aanpassing eindexamen voor middelbare scholieren*. Rijksoverheid. Retrieved from: https://www.rijksoverheid.nl/actueel/nieuws/2022/01/24/extra-aanpassing-eindexamen-voor-middelbare-scholieren (9-6-2023)

RTLNieuws, 2022. Schoolexamens komend jaar weer bijna als vanouds. *RTLNieuws*. Retrieved from: https://www.rtlnieuws.nl/nieuws/nederland/artikel/5348354/examens-middelbare-school-eisen (9-6-2023)

Walker, J., Pearce, C., Boe, K. and Lawson, M., 2019. *The power of education to fight inequality*. Oxfam. Retrieved from: https://www.oxfam.org/en/research/power-education-fight-inequality [pdf] (9-6-2023)

WEF, 2020. *Global Social Mobility Index 2020*. World Economic Forum. Retrieved from: https://www.weforum.org/reports/global-social-mobility-index-2020-why-economies-benefit-from-fixing-inequality (9-6-2023)

Zumbuehl, M. and Dillingh, R., 2020. *Ongelijkheid van het jonge kind*. Centraal Planbureau. Retrieved from: https://www.cpb.nl/sites/default/files/omnidownload/CPB-Notitie-Ongelijkheid-van-het-jonge-kind.pdf [pdf] (9-6-2023)

APPENDIX A: Codebook

Label	Code	Meaning	Values
BRIN Number	ID	Registration number of educational institution	
Year	YEAR	School year	0=2017-2018 1=2018-2019 2=2019-2020 3=2020-2021 4=2021-2022
Graduation percentage female	PERC_F	Graduation percentage of female students	
Graduation percentage male	PERC_M	Graduation percentage of male students	
Female students per female teacher (fte)	RATIO_FTE_F	Number of female students per female teacher in fte	
Female students per female teacher (pers)	RATIO_PERS_F	Number of female students per female teacher in number of people	
Male students per male teacher (fte)	RATIO_FTE_M	Number of male students per male teacher in fte	
Male students per male teacher (pers)	RATIO_PERS_M	Number of male students per male teacher in number of people	
Management in fte female lim	MAN_FTELIM_F	Female management personnel in fte (limited)	
Management in fte female comp	MAN_FTECOMP_F	Female management personnel in fte (comprehensive)	
Management in pers female lim	MAN_PERSLIM_F	Female management personnel in number of people (limited)	
Management in pers female comp	MAN_PERSCOMP_F	Female management personnel in number of people (comprehensive)	
Management in fte male lim	MAN_FTELIM_M	Male management personnel in fte (limited)	
Management in fte male comp	MAN_FTECOMP_M	Male management personnel in fte (comprehensive)	
Management in pers male lim	MAN_PERSLIM_M	Male management personnel in number of people (limited)	
Management in pers male comp	MAN_PERSCOMP_M	Male management personnel in number of people (comprehensive)	
Educational staff in fte female	TOT_FTE_F	Female educational staff in fte	
Educational staff in people female	TOT_PERS_F	Female educational staff in number of people	
Educational staff in fte male	TOT_FTE_M	Male educational staff in fte	
Educational staff in people male	TOT_PERS_M	Male educational staff in number of people	

Class size	CLASS_SIZE	Percentage of teachers (fte) per student	
School denomination	DENOMINATION	School denomination	0=Public 1=Private
Income	INCOME	Standardised disposable income per municipality	0=less than 27 000 1=27 000 - 31 000 2=31 000 – 35 000 3=35 000 – 39 000 4=more than 39 000
Region	REGION	Population density of the municipality in which the school is located	0=not urban 1=little urban 2=moderately urban 3=highly urban 4=very highly urban
Teaching experience female	AGE_F	Average age of female teachers	
Teaching experience male	AGE_M	Average age of male teachers	
Type of education	EDU_TYPE	Type of education	0=General 1=Specialised

APPENDIX B: Details of secondary school personnel data

Onderwijspersoneel in het vo in aantal fte

Inhoud

In dit bestand staat het onderwijspersoneel in het voortgezet onderwijs uitgedrukt in aantal fte. De aantallen worden weergegeven per onderwijstype, bestuur (bevoegd gezag), instelling (brin nummer), regio en functiegroep. Het bestand verschaft o.a. een overzicht van het aantal mannen en vrouwen, evenals de gemiddelde leeftijd en fte's van het personeel. De gegevens staan weergegeven over een oplopende tijdreeks vanaf 2011. Op de website zijn de definities te vinden van de kolomnamen en overige begrippen in het bestand.

In dit bestand staan 7 werkbladen:
Werkblad 1: fte naar onderwijstype
Werkblad 2: fte naar functiegroep
Werkblad 3: fte naar onderwijstype en functiegroep
Werkblad 4: fte naar regio
Werkblad 5: fte naar regio, onderwijstype en functiegroep
Werkblad 6: fte naar onderwijstype, bevoegd gezag en brin nummer
Werkblad 7: fte naar onderwijstype, bevoegd gezag, brin nummer en functiegroep

Bronnen

Schoolbesturen (bevoegde gezagen) zijn wettelijk verplicht de gegevens over onderwijspersoneel aan DUO te leveren. Veelal verzorgen salarisverwerkers deze leveringen namens de schoolbesturen.

Gegevensselectie

- Het bestand is een samenstelling van gegevens op 1 oktober over over een oplopende tijdreeks vanaf 2011.
- De gegevens betreffen personen met een betrekkingsomvang groter dan 0 fte op peildatum 1 oktober.
- De selectie is op basis van de reguliere formatie inclusief leraren in opleiding. Vervangers, gastdocenten, stagiairs en uitzendkrachten zijn niet meegeteld. Bovenschools personeel wordt wel meegeteld (zie brinnummer).
- De gegevens hebben betrekking op onderwijsinstellingen mét leerlingen en gegevens van instellingen zonder leerlingen die vallen onder een bevoegd gezag van minimaal 1 instelling met leerlingen.
- Bij meerdere dienstverbanden per persoon per brin nummer, wordt de functie met de grootste betrekkingsomvang als 'hoofdfunctie' gedefinieerd. Deze hoofdfunctie bepaalt de indeling van de persoon in een functiegroep.

Toelichting specifieke variabelen en waarden

- Vanwege de regelgeving m.b.t. de Algemene verordening gegevensbescherming (AVG) zijn aantallen < 5 vervangen door 1 sterretje (*). Als het aantal kleiner is dan zal de waarde van bijbehorende fte ook met een ster worden weergegeven.
- Niet alle instellingen hebben de nodige personeelsgegevens geleverd aan DUO. Dit is nadelig voor de vergelijkbaarheid van de cijfers over de verschillende jaren. Daarom is er besloten om landelijk en per regio een weegfactor te calculeren en hiermee bij te schatten voor de ontbrekende gegevens. De landelijke weegfactor is over de cijfers van de tabbladen 1 tot en met 3 toegepast. De regionale weegfactor is over de cijfers van de tabbladen 4 en 5 toegepast. Doordat er verschillende weegfactoren gebruikt zijn kunnen er kleine verschillen zijn tussen de totalen van tabbladen waar de landelijke weegfactor is gebruikt en de tabbladen waar de regionale weegfactoren zijn toegepast. Omdat de cijfers in de tabbladen 6 en 7 op instellingsniveau worden gepresenteerd, is er hier geen weegfactor toegepast.
- In gevallen waar het responspercentage binnen een regio onder de 50% ligt, wordt de weegfactor zodanig hoog dat de inhoud van gerelateerde cellen te onnauwkeurig wordt voor presentatie. Daarom is er besloten om de inhoud van deze cellen te vervangen door twee sterretjes (**). Zie Tabel 1 voor een overzicht van de totale dekkingsgraad per subsector.

ONDERWIJSTYPE:

Het onderwijstype is onderverdeeld in de volgende categorieën:

- Het po (primair onderwijs): bao (basisonderwijs), sbao (speciaal basisonderwijs) en wec ((voortgezet) speciaal onderwijs).
- Het vo (voortgezet onderwijs).
- Het mbo (middelbaar beroepsonderwijs): bve (beroeps- en volwasseneducatie) en aoc (agrarisch opleidingscentrum). BRIN NUMMER:

Als het brin nummer van een instelling in de betreffende periode door een fusie is veranderd, dan wordt het personeel van die betreffende instelling voor alle jaren onder het brin nummer dat geldt na fusie vermeld. Als het brin nummer van een instelling in de periode door een splitsing is veranderd, dan wordt het personeel van die betreffende instelling voor alle jaren onder het brin nummer dat gold vóór de splitsing vermeld. Als in de kolom BRIN NUMMER de waarde 'bovenschools' staat vermeld, dan betreft het personeel in dienst van een bevoegd gezag dat door dit bevoegd gezag niet bij een specifieke instelling is ingedeeld. Doorgaans verricht dit personeel werkzaamheden voor meer dan één instelling of ter ondersteuning van het bevoegd gezag.

REGIO:

De Regionale Platforms Arbeidsmarkt heeft 34 RPA-gebieden afgeleid uit de 131 werkgebieden van de UWV WERKbedrijf (voorheen Centrums voor Werk en Inkomen). Deze gebieden zijn in dit bestand verder uitgesplitst naar 39 RPA-gebieden, waarbij de vijf grote steden (Amsterdam, Rotterdam, Den Haag, Utrecht en Almere) apart zijn benoemd.

FUNCTIEGROEP:
Van een aantal dienstverbanden is de functiegroep onbekend. Het is onmogelijk dat er geen onderwijzend personeel werkzaam is geweest bij een onderwijsinstelling (m.u.v. 'bovenschoolse brins'). Waar dit het geval bleek, achtte DUO de verdeling van de formatie voor die instelling onbetrouwbaar en is de formatie gehercodeerd naar de functiecategorie 'onbekend'.

VASTE EN TIJDELIJKE DIENST:
De aard van het dienstverband van personeel (in vaste of tijdelijke dienst) is opgenomen in dit bestand.

GESLACHT EN LEEFTIJD ONBEKEND:
Van een klein groepje personeel ontbreekt het geslacht en/of leeftijd in de aan DUO geleverde personeelsgegevens. Om toch een compleet overzicht te verschaffen worden deze cijfers apart getoond.

GEMIDDELDE LEEFTIJD:
De gemiddelde leeftijd is berekend op basis van de leeftijd op de peildatum (1 oktober).

LEEFTIJDSCATEGORIEËN:
Er zijn zeven leeftijdscategorieën gedefinieerd. De eerste categorie begint bij 'jonger dan 15 jaar' en loopt met stappen van 10 jaar uit tot de laatste leeftijdscategorie dat eindigt bij 'ouder 65 jaar'.

FTE CATEGORIEËN:
Er zijn drie fte categorieën gedefinieerd, namelijk: 1) personen met een betrekkingsomvang tussen de 0 en 0,5 fte's, 2) personen met een betrekkingsomvang tussen de 0,5 en 0,8 fte's, en 3) personen met een betrekkingsomvang die meer is dan
0,8 fte's.

Kwaliteit van de gegevens
Voor de volledigheid van de gegevens is DUO afhankelijk van leveringen door salarisverwerkers en bevoegde gezagen van scholen. Elk jaar ontbreken van één of meer instellingen gegevens, omdat deze niet zijn geleverd. Vanaf 2004 is het aantal scholen waarvan geen gegevens beschikbaar is, fors gestegen. Door nieuwe regelgeving is deze trend echter vanaf 2008 omgebogen. Tabel 1 verschaft een overzicht van de volledigheid van de gegevens vanaf 2011.

Tabel 1. Dekkingsgraad van de personeelsgegevens vanaf 2011 voor het vo

	2011	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
VO												
Aantal geleverde instellingen	590	588	602	599	599	607	610	611	616	605	615	583
Dekkingsgraad in % o.b.v. instellingen	97,2	96,7	98,5	98,5	94,3	99,5	99,4	99,4	99,4	97,7	99,2	94,2
Dekkingsgraad in % o.b.v. leerlingen	97,9	99,4	98,6	99,1	99,4	99,9	99,8	99,7	99,6	98,2	99,1	91,0

Onderwijspersoneel in het vo in aantal fte

Inhoud
In dit bestand staat het onderwijspersoneel in het voortgezet onderwijs uitgedrukt in aantal fte. De aantallen worden weergegeven per onderwijstype, bestuur (bevoegd gezag), instelling (brin nummer), regio en functiegroep. Het bestand verschaft o.a. een overzicht van het aantal mannen en vrouwen, evenals de gemiddelde leeftijd en fte's van het personeel. De gegevens staan weergegeven over een oplopende tijdreeks vanaf 2011. Op de website zijn de definities te vinden van de kolomnamen en overige begrippen in het bestand.

In dit bestand staan 7 werkbladen:
Werkblad 1: fte naar onderwijstype
Werkblad 2: fte naar functiegroep
Werkblad 3: fte naar onderwijstype en functiegroep
Werkblad 4: fte naar regio
Werkblad 5: fte naar regio, onderwijstype en functiegroep
Werkblad 6: fte naar onderwijstype, bevoegd gezag en brin nummer
Werkblad 7: fte naar onderwijstype, bevoegd gezag, brin nummer en functiegroep

Bronnen
Schoolbesturen (bevoegde gezagen) zijn wettelijk verplicht de gegevens over onderwijspersoneel aan DUO te leveren. Veelal verzorgen salarisverwerkers deze leveringen namens de schoolbesturen.

Gegevensselectie
Het bestand is een samenstelling van gegevens op 1 oktober over over een oplopende tijdreeks vanaf 2011.
De gegevens betreffen personen met een betrekkingsomvang groter dan 0 fte op peildatum 1 oktober.
De selectie is op basis van de reguliere formatie inclusief leraren in opleiding. Vervangers, gastdocenten, stagiairs en uitzendkrachten zijn niet meegeteld. Bovenschools personeel wordt wel meegeteld (zie brinnummer).
De gegevens hebben betrekking op onderwijsinstellingen mét leerlingen en gegevens van instellingen zonder leerlingen die vallen onder een bevoegd gezag van minimaal 1 instelling met leerlingen.
Bij meerdere dienstverbanden per persoon per brin nummer, wordt de functie met de grootste betrekkingsomvang als 'hoofdfunctie' gedefinieerd. Deze hoofdfunctie bepaalt de indeling van de persoon in een functiegroep.

Toelichting specifieke variabelen en waarden
Vanwege de regelgeving m.b.t. de Algemene verordening gegevensbescherming (AVG) zijn aantallen < 5 vervangen door 1 sterretje (*). Als het aantal kleiner is dan zal de waarde van bijbehorende fte ook met een ster worden weergegeven.

Niet alle instellingen hebben de nodige personeelsgegevens geleverd aan DUO. Dit is nadelig voor de vergelijkbaarheid van de cijfers over de verschillende jaren. Daarom is er besloten om landelijk en per regio een weegfactor te calculeren en hiermee bij te schatten voor de ontbrekende gegevens. De landelijke weegfactor is over de cijfers van de tabbladen 1 tot en met 3 toegepast. De regionale weegfactor is over de cijfers van de tabbladen 4 en 5 toegepast. Doordat er verschillende weegfactoren gebruikt zijn kunnen er kleine verschillen zijn tussen de totalen van tabbladen waar de landelijke weegfactor is gebruikt en de tabbladen waar de regionale weegfactoren zijn toegepast. Omdat de cijfers in de tabbladen 6 en 7 op instellingsniveau worden gepresenteerd, is er hier geen weegfactor toegepast.
In gevallen waar het responspercentage binnen een regio onder de 50% ligt, wordt de weegfactor zodanig hoog dat de inhoud van gerelateerde cellen te onnauwkeurig wordt voor presentatie. Daarom is er besloten om de inhoud van deze cellen te vervangen door twee sterretjes (**). Zie Tabel 1 voor een overzicht van de totale dekkingsgraad per subsector.

ONDERWIJSTYPE:
Het onderwijstype is onderverdeeld in de volgende categorieën:
Het po (primair onderwijs): bao (basisonderwijs), sbao (speciaal basisonderwijs) en wec ((voortgezet) speciaal onderwijs).
Het vo (voortgezet onderwijs).
Het mbo (middelbaar beroepsonderwijs): bve (beroeps- en volwasseneducatie) en aoc (agrarisch opleidingscentrum). BRIN NUMMER:
Als het brin nummer van een instelling in de betreffende periode door een fusie is veranderd, dan wordt het personeel van die betreffende instelling voor alle jaren onder het brin nummer dat geldt na fusie vermeld. Als het brin nummer van een instelling in de periode door een splitsing is veranderd, dan wordt het personeel van die betreffende instelling voor alle jaren onder het brin nummer dat gold vóór de splitsing vermeld. Als in de kolom BRIN NUMMER de waarde 'bovenschools' staat vermeld, dan betreft het personeel in dienst van een bevoegd gezag dat door dit bevoegd gezag niet bij een specifieke instelling is ingedeeld. Doorgaans verricht dit personeel werkzaamheden voor meer dan één instelling of ter ondersteuning van het bevoegd gezag.

REGIO:
De Regionale Platforms Arbeidsmarkt heeft 34 RPA-gebieden afgeleid uit de 131 werkgebieden van de UWV WERKbedrijf (voorheen Centrums voor Werk en Inkomen). Deze gebieden zijn in dit bestand verder uitgesplitst naar 39 RPA-gebieden, waarbij de vijf grote steden (Amsterdam, Rotterdam, Den Haag, Utrecht en Almere) apart zijn benoemd.

FUNCTIEGROEP:
Van een aantal dienstverbanden is de functiegroep onbekend. Het is onmogelijk dat er geen onderwijzend personeel werkzaam is geweest bij een onderwijsinstelling (m.u.v. ‘bovenschoolse brins’). Waar dit het geval bleek, achtte DUO de verdeling van de formatie voor die instelling onbetrouwbaar en is de formatie gehercodeerd naar de functiecategorie ‘onbekend’.

VASTE EN TIJDELIJKE DIENST:
De aard van het dienstverband van personeel (in vaste of tijdelijke dienst) is opgenomen in dit bestand.

GESLACHT EN LEEFTIJD ONBEKEND:
Van een klein groepje personeel ontbreekt het geslacht en/of leeftijd in de aan DUO geleverde personeelsgegevens. Om toch een compleet overzicht te verschaffen worden deze cijfers apart getoond.

GEMIDDELDE LEEFTIJD:
De gemiddelde leeftijd is berekend op basis van de leeftijd op de peildatum (1 oktober).

LEEFTIJDSCATEGORIEËN:
Er zijn zeven leeftijdscategorieën gedefinieerd. De eerste categorie begint bij 'jonger dan 15 jaar' en loopt met stappen van 10 jaar uit tot de laatste leeftijdscategorie dat eindigt bij 'ouder 65 jaar'.

FTE CATEGORIEËN:
Er zijn drie fte categorieën gedefinieerd, namelijk: 1) personen met een betrekkingsomvang tussen de 0 en 0,5 fte's, 2) personen met een betrekkingsomvang tussen de 0,5 en 0,8 fte's, en 3) personen met een betrekkingsomvang die meer is dan
0,8 fte's.

Kwaliteit van de gegevens
Voor de volledigheid van de gegevens is DUO afhankelijk van leveringen door salarisverwerkers en bevoegde gezagen van scholen. Elk jaar ontbreken van één of meer instellingen gegevens, omdat deze niet zijn geleverd. Vanaf 2004 is het aantal scholen waarvan geen gegevens beschikbaar is, fors gestegen. Door nieuwe regelgeving is deze trend echter vanaf 2008 omgebogen. Tabel 1 verschaft een overzicht van de volledigheid van de gegevens vanaf 2011.

Tabel 1. Dekkingsgraad van de personeelsgegevens vanaf 2011 voor het vo

	2011	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
VO												
Aantal geleverde instellingen	590	588	602	599	599	607	610	611	616	605	615	583
Dekkingsgraad in % o.b.v. instellingen	97,2	96,7	98,5	98,5	94,3	99,5	99,4	99,4	99,4	97,7	99,2	94,2
Dekkingsgraad in % o.b.v. leerlingen	97,9	99,4	98,6	99,1	99,4	99,9	99,8	99,7	99,6	98,2	99,1	91,0

APPENDIX C: Syntax

4.1 Descriptive statistics

Female

summarize PERC_F RATIO_FTE_F MAN_FTELIM_F MAN_FTECOMP_F TOT_FTE_F RATIO_PERS_F MAN_PERSLIM_F MAN_PERSCOMP_F TOT_PERS_F CLASS_SIZE DENOMINATION INCOME REGION AGE_F EDU_TYPE, detail

Male

summarize PERC_M RATIO_FTE_M MAN_FTELIM_M MAN_FTECOMP_M TOT_FTE_M RATIO_PERS_M MAN_PERSLIM_M MAN_PERSCOMP_M TOT_PERS_M CLASS_SIZE DENOMINATION INCOME REGION AGE_M EDU_TYPE, detail

4.2 Correlation analysis

Female

pwcorr PERC_F RATIO_FTE_F MAN_FTELIM_F MAN_FTECOMP_F TOT_FTE_F RATIO_PERS_F MAN_PERSLIM_F MAN_PERSCOMP_F TOT_PERS_F CLASS_SIZE DENOMINATION Income_0 Income_1 Income_2 Income_3 Income_4 Region_0 Region_1 Region_2 Region_3 Region_4 AGE_F EDU_TYPE, sig star(.05)obs

Male

pwcorr PERC_M RATIO_FTE_M MAN_FTELIM_M MAN_FTECOMP_M TOT_FTE_M RATIO_PERS_M MAN_PERSLIM_M MAN_PERSCOMP_M TOT_PERS_M CLASS_SIZE DENOMINATION Income_0 Income_1 Income_2 Income_3 Income_4 Region_0 Region_1 Region_2 Region_3 Region_4 AGE_M EDU_TYPE, sig star(.05)obs

4.3 Multilevel regression analysis

Female

Model 0

mixed PERC_F YEAR || ID:, var mle
estat icc

Model 1a

mixed PERC_F YEAR RATIO_FTE_F MAN_FTECOMP_F TOT_FTE_F CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_F EDU_TYPE || ID:, var mle

Model 1b

mixed PERC_F YEAR RATIO_PERS_F MAN_PERSCOMP_F TOT_PERS_F CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_F EDU_TYPE || ID:, var mle

Model 2a
mixed PERC_F YEAR c.RATIO_FTE_F##c.MAN_FTECOMP_F MAN_FTECOMP_DUMMY CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_F EDU_TYPE || ID:, var mle

Model 2b
mixed PERC_F YEAR c.RATIO_PERS_F##c.MAN_PERSCOMP_F MAN_PERSCOMP_DUMMY CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_F EDU_TYPE || ID:, var mle

Model 3a
mixed PERC_F YEAR c.RATIO_FTE_F##c.TOT_FTE_F CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_F EDU_TYPE || ID:, var mle

Model 3b
mixed PERC_F YEAR c.RATIO_PERS_F##c.TOT_PERS_F CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_F EDU_TYPE || ID:, var mle

Male
Model 0
mixed PERC_M YEAR || ID:, var mle
estat icc

Model 1a
mixed PERC_M YEAR RATIO_FTE_M MAN_FTECOMP_M TOT_FTE_M CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_M EDU_TYPE || ID:, var mle

Model 1b
mixed PERC_M YEAR RATIO_PERS_M MAN_PERSCOMP_M TOT_PERS_M CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_M EDU_TYPE || ID:, var mle

Model 2a
mixed PERC_M YEAR c.RATIO_FTE_M##c.MAN_FTECOMP_M MAN_FTECOMP_DUMMY CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_M EDU_TYPE || ID:, var mle

Model 2b
mixed PERC_M YEAR c.RATIO_PERS_M##c.MAN_PERSCOMP_M MAN_PERSCOMP_DUMMY CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_M EDU_TYPE || ID:, var mle

Model 3a
mixed PERC_M YEAR c.RATIO_FTE_M##c.TOT_FTE_M CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_M EDU_TYPE || ID:, var mle

Model 3b
mixed PERC_M YEAR c.RATIO_PERS_M##c.TOT_PERS_M CLASS_SIZE DENOMINATION Income_1 Income_2 Income_3 Income_4 Region_1 Region_2 Region_3 Region_4 AGE_M EDU_TYPE || ID:, var mle

APPENDIX D: List of tables showing results of analysis for male variables

TABLE A1: descriptive statistics for all variables included in the analysis

	N	Mean	St. Dev.	Var.	Min	Max
Dependent variable						
1. Graduation percentage of male students (%)	2 676	94.38	4.97	24.70	0	100
Independent/moderator variables in fte						
2. Number of male students per male teacher	2 481	3.21	.81	.66	.20	8.36
3a. Number of male manager (limited)	216	9.66	3.88	15.05	2.25	20.50
3b. Number of male managers (comprehensive)	1 006	5.82	3.79	14.37	0	24
4. Number of male educational staff	2 496	75.99	48.55	2357.53	2.3	282.48
Independent/moderator variables in people						
5. Number of male students per male teachers	2 481	2.73	.73	.53	.09	7.375
6a. Number of male managers (limited)	215	10.03	4.01	16.05	5	22
6b. Number of male managers (comprehensive)	1 007	6.08	3.93	15.43	0	24
7. Number of educational staff	2 496	88.04	54.99	3023.98	5	325
Control variables						
8. Class size	2 486	2.46	.57	.33	.17	6.18
9. School denomination	2 550	.70	.46	.21	0	1
10. Income	2 570	1.79	.81	.65	0	4
11. Region	2 625	2.77	1.14	1.29	0	4
12. Average age of male teachers	2 550	44.71	6.79	46.07	0	55.51
13. Type of education	2 750	.23	.42	.18	0	1

Note: rounded to the second decimal

TABLE A2: correlations for all variables in the analysis

	1	2	3a	3b	4	5	6a	6b	7	8	9	10a	10b	10c	10d	10e	11a	11b	11c	11d	11e	12	13
1	1.00																						
2	.00	1.00																					
3a	.08	.04	1.00																				
3b	.05	-.08*	1.00*	1.00																			
4	.04	-.12*	**.49***	**.62***	1.00																		
5	.03	.97*	.07	-.05	-.06*	1.00																	
6a	.04	.05	.94*	.87*	**.27***	.07	1.00																
6b	.07*	-.05	.83*	.89*	**.46***	-.01	1.00*	1.00															
7	.03	-.13*	**.47***	**.61***	1.00*	-.08*	**.26***	**.45***	1.00														
8	.01	.77*	.14*	.01	.02	.82*	.17*	.02	-.00	1.00													
9	.04*	-.01	.24*	.12*	.10*	-.01	.24*	.14*	.09*	.03	1.00												
10a	-.04*	-.07*	.10	.05	.02	-.06*	.08	.02	.02	-.01	.01	1.00											
10b	-.04*	-.09*	.19*	.03	.04*	-.09*	.17*	.04	.04*	-.10*	-.08*	-.10*	1.00										
10c	.02	.00	-.15*	.01	.06*	.02	-.12	.01	.06*	.04*	-.01	-.13*	-.63*	1.00									
10d	.04*	.15*	-.04	-.06	-.10*	.13*	-.03	-.03	-.10*	.10*	.09*	-.05*	-.23*	-.30*	1.00								
10e	.02	.04	-.07	-.03	-.06*	.01	-.16*	-.08*	-.06*	.00	.09*	-.03	-.14*	-.18*	-.06*	1.00							
11a	.04*	-.02	-.11	-.09*	-.12*	-.00	-.09	-.09*	-.12*	-.01	-.06*	-.03	.05*	-.00	-.03	-.04	1.00						
11b	.06*	.05*	-.16*	-.12*	-.02	.08*	-.17*	-.12*	-.04	.14*	.07*	-.06*	-.09*	.07*	.03	.07*	-.07*	1.00					
11c	.05*	.12*	.17*	.07*	.09*	.14*	.25*	.06	.08*	.16*	.06*	-.06*	-.11*	-.05*	.18*	.16*	-.08*	-.17*	1.00				
11d	.01	.03	-.06	.07*	.14*	.04*	-.13	.07*	.14*	.06*	-.02	.20*	-.06*	.11*	-.02	-.02	-.13*	-.28*	-.30*	1.00			
11e	-.10*	-.16*	.03	-.02	-.13*	-.21*	.04	-.01	-.11*	-.27*	-.05*	-.09*	.28*	-.04*	-.11*	-.13*	-.12*	-.26*	-.28*	-.46*	1.00		
12	-.04	.01	-.09	.01	.07*	.06*	-.11	-.04	.06*	.18*	-.06*	.06*	.02	.01	.00	-.04*	.02	.07*	.07*	-.03	-.05*	1.00	
13	.12*	.02	-.23*	-.13*	-.44*	-.01	-.04	-.05	-.44*	-.12*	-.06*	.08*	-.07*	-.03	-.02	.03	-.00	-.09*	-.14*	-.00	.07*	-.12*	1.00

Note: significance level*=p<0.05; rounded to the second decimal

TABLE A2: continued

	Dependent variable
1	Graduation percentage of male students
	Independent variables in fte
2	Number of male students per male teacher
3a	Number of male managers (limited)
3b	Number of male managers (comprehensive)
4	Number of male educational staff
	Independent variables in number of people
5	Number of male students per male teacher
6a	Number of male managers (limited)
6b	Number of male managers (comprehensive)
7	Number of male educational staff
	Control variables
8	Class size
9	School denomination
10a	Income: less than 27 000
10b	Income: 27 000 - 31 000
10c	Income: 31 000 - 35 000
10d	Income: 35 000 – 39 000
10e	Income: more than 39 000
11a	Region: not urban
11b	Region: little urban
11c	Region: moderately urban
11d	Region: highly urban
11e	Region: very highly urban
12	Teaching experience
13	Type of education

TABLE A3: multilevel regression analysis of the graduation percentage of male students

	Model 0	**Model 1a**	**Model 1b**
Year	.794*	.750*	.733*
Number of male students per male teacher		.436	.307
Number of male managers (comprehensive)		.016	.063
Number of male educational staff		.009	.005
Class size		-.169	-.109
School denomination		.628	.953*
Income: 27 000 - 31 000		.098	.208
Income: 31 000 - 35 000		-.211	-.001
Income: 35 000 - 39 000		-.024	.466
Income: more than 39 000		-.012	-.023
Region: little urban		.728	.754
Region: moderately urban		.986	.840
Region: highly urban		-.115	-.086
Region: very highly urban		-1.396	-1.609
Teaching experience		-.076	-.093
Type of education		2.954*	2.935*
Constant	91.997*	93.114*	94.130*
N	2 676	1 000	997

Note: significance level*=p<0.05; rounded to the third decimal
Model 1a shows personnel in fte, Model 1b shows personnel in number of people
Reference categories are ‘Income: less than 27 000’ and ‘Region: not urban’

TABLE A4: multilevel regression analysis of the graduation percentage of male students including moderator variables

	Model 2a	**Model 2b**	**Model 3a**	**Model 3b**
Year	.734*	.721*	.812*	.799*
Number of male students per male teacher	.726	.584	-.473	-.091
Number of male managers (comprehensive)	.318	.252	.	.
Number of male managers (dummy)	-.124	-.092	.	.
Number of male educational staff	.	.	**-.006***	**-.004***
Number of male students per male teacher*Number of male managers	-.080	-.060	.	.
Number of students per male teacher*Number of male educational staff	.	.	.005	.005
Class size	-.161	-.134	.193	-.262
School denomination	.698	1.014*	.403	.425
Income: 27 000 - 31 000	.059	.135	.016	-.046
Income: 31 000 - 35 000	-.212	-.037	.022	-.035
Income: 35 000 - 39 000	-.176	.372	.452	.325
Income: more than 39 000	-.193	-.143	.265	.144
Region: little urban	.873	.841	-.486	-.470
Region: moderately urban	1.244	1.020	-.571	-.582
Region: highly urban	.203	.125	-1.323	-.1.331
Region: very highly urban	-1.273	-1.494	-2.186*	-.2.219*
Teaching experience	-.090	-.090	.034	.044
Type of education	2.658*	2.694*	2.499*	2.405*
Constant	93.364*	93.635*	91.047*	90.566*
N	1 000	995	2 450	2 450

Note: significance level*=p<0.05; rounded to the third decimal

Models 2a and 3a show personnel in fte, Models 2b and 3b show personnel in number of people

Reference categories are ‘Income: less than 27 000’ and ‘Region: not urban’