How does Public Service Motivation affect performance in schools?

Lotte Bøgh Andersen (lotte@ps.au.dk),

Eskil Heinesen (esh@rff.dk) and

Lene Holm Pedersen (LeHo@kora.dk)

Abstract

The literature expects public service motivation (PSM) to affect performance; but most studies of this relationship use subjective performance data. This article therefore investigates the association between PSM and the performance of Danish teachers using an objective performance measure (the students' final examination marks). Combining survey data and administrative register data in a multilevel dataset, we are able to control very robustly for the specific characteristics of the students (n=5,631), the schools (n=85), and other characteristics of teachers (n=694) than PSM. We find that PSM is positively associated with examination marks, but that the relationship is more complex than expected when we look at the PSM dimensions. The results imply that PSM can be relevant for performance improvements although attention should be given to the employees' specific understanding of the desirable for society and others.

Key words: Public service motivation, performance, examination marks, teachers

Paper presented at the 2012 APPAM conference

Introduction

As the provision of public service is community-oriented in nature, 'an individual's orientation to delivering services to people with the purpose of doing good for others and society' – also known as public service motivation (PSM) – has attracted considerable interest among public management scholars (Hondeghem and Perry, 2009:6; Perry, Hondeghem & Wise, 2010). Indeed, it has been a central driving force behind PSM research that PSM is expected to result in better performance in public organizations (Perry and Wise, 1990). The reasoning behind this is that public service motivated individuals are expected to invest more effort in providing public service, because they want to do good for others and society. However, this outcome has proved difficult to verify empirically. Existing studies of the relationship between PSM and performance indicate that there is a positive association (Petrovsky & Ritz, 2010; Vandenabeele 2009; Leisink and Steijn 2009; Naff and Crum 1999), but the heavy reliance on self-reports and cross-sectional data in these studies make causal inference difficult (Brewer, 2008; Perry, Hondeghem and Wise, 2010: 685; Petrovsky and Ritz 2010). Recent studies of social desirability bias in PSM research further highlight that it can be problematic to use the same survey to measure PSM and performance, because social desirability may lead the same individuals to have a positive bias both when they report their performance and when they answer questions on PSM (Kim and Kim 2012). This indicates that we should try to obtain objective performance data when we analyse the association between PSM and performance. To be objective, a measure of performance must involve a precise assessment of a given dimension of performance and an external process to verify its accuracy (Andrews, Boyne and Walker, 2006: 16). It is, however, very challenging to find such a measure which can be linked to public employees' motivations and effort, because public service delivery is often the results of an aggregate effort where several

employees work together, and because public organizations often have multiple and unclear goals (Dixit 2002). It is like finding a needle in a hay-stack. But due to the extremely high quality of the Danish administrative registers, it is possible to do this for Danish school teachers by combining the data from the administrative registers on the students' socio-demographic background and final examination marks with survey data on the teachers' PSM. The present study therefore provides a unique analysis of the relationship between PSM and very objectively measured performance. The fact that children are taught in the same classes by different teachers in different subjects (Danish, math etc.) means that we can control for student specific characteristics and teacher selection effects by including student fixed effects in the analyses. Controlling for student and class specific variation ensures that the estimates of the associations are not biased by school, student or class level confounding. We take account of the multilevel structure of our data by clustering standard errors on classes. We present estimates both of the effects of the aggregate PSM construct and of the PSM dimensions and thus test whether these dimensions are differently associated with performance. Our research design enables us to contribute to answer one of the most central questions in PSM research: Is PSM and individual performance positively related, also when performance is objectively measured?

Our key argument is that PSM 'fuels' the types of behavior which the individual sees as "doing good for others and society". If what the organization rates as high performance is also seen by the individuals as the desirable to society and others, we expect PSM to be positively associated with performance. We also argue that if we disaggregate PSM into its dimensions, some of the differences in what "doing good" means can be captured, and different patterns of association between the PSM dimensions and performance can be expected.

Our main contribution is that we test the PSM-performance association very robustly, using objective performance data. To do this, the article proceeds as follows. We first discuss the literature on PSM and individual performance and formulate specific hypotheses followed by a presentation of the data and methods. We then present and discuss the results, and the article concludes with a discussion of its contributions, limitations and implications.

Theory: Public service motivation and individual performance

Already in 1990, Perry and Wise (1990) hypothesized that PSM is positively related to individual performance. They argued that individuals will be motivated to perform well when finding their work meaningful. Accordingly, they expected that task identity and perceived task significance contribute to performance, and that individuals with high PSM derive these job characteristics from producing public service (ibid.: 371). It can, however, be discussed whether this is the case for all jobs. Delivery of public services is not a natural part of all jobs, and PSM can only be expected to increase performance for jobs which allow the individuals to do "good for others and society" (Hondeghem & Perry, 2009: 6). There should, in other words, be a PSM fit which can be defined as "comparability between the needs of individuals to serve the public interest and the environmental conditions in their organisation which affect the fulfilment of these altruistic motives" (Taylor 2008, 71-2). When there is PSM fit, existing studies find a positive association between public service motivation and job satisfaction (Steijn 2008; Andersen & Kjeldsen, in press; Kjeldsen & Andersen, in press), and the interpretation in the person-job fit literature is that the employees' needs, desires and preferences are then met by the jobs that they perform (Kristof-Brown et al. 2005: 306).

The PSM construct consists of four separate dimensions - 'attraction to public policy making', 'commitment to the public interest', 'compassion' and 'self-sacrifice' - which add up in an aggregate construct called PSM. When the individual dimensions can be expected to have different consequences, there is reason to pay attention to the causes and consequences of the dimensions as well as of the overall PSM construct. Perry and Wise (1990) implicitly assume that performance for public services is always what is good for others and society and that a consensual understanding exists of what one should do to obtain this. We agree that individuals with high PSM are likely to make great efforts to deliver service to people with the purpose of doing good for others and society, but different individuals do not necessarily see the same actions as beneficial to others and society. PSM might also affect behavior in a way that is beneficial to some interests and not to others (Andersen & Serritzlew, 2012). In line with this, Brewer (2008: 146) has pointed out that some PSM dimensions may be more strongly linked to individual and organizational performance, while others are weakly related, unrelated or even inversely related. If PSM only fuels behavior which the *individual employee* sees as desirable (Andersen et al., 2012a), there will only be a positive association between PSM and performance if the investigated performance measure corresponds with the individual's understanding of the desirable for society and others. Investigating the PSM-performance relationship is therefore difficult, because different individuals may have different conceptions of what the public interest is. This makes it very important to discuss how performance is measured.

Performance is a contested concept. According to Boyne (2003), it is difficult to define what high performance is for public services, because different stakeholders seldom agree on what the goal is. Even when a consensual performance measure exists, testing the effect of PSM on performance is difficult because employee performance is affected by an

array of factors other than PSM, e.g. pay structure, education level, professionalism, private or public ownership, other motivational variables, and institutional context (Rainey & Steinbauer, 1999; Perry, Mesch & Paarlberg, 2006; Flynn, 2007). Additionally, because of low efficacy, it is a very conservative test to use outcome measures - such as student examination marks - in studies of individual employee performance (Miller & Whitford, 2007). These difficulties are perhaps an important explanation of why individual performance has typically been measured as self-reported performance. Self-reported performance measures have several merits, and the most important advantage is that it is possible to obtain it from virtually any group of employees, even when tasks are hard or even impossible to measure (Kim 2005). However, such measures also have limitations. Some employees may thus have biased perceptions of their own performance, and the tendency to inflate one's own performance may be correlated with the tendency to inflate PSM (Kim & Kim, 2012). Self-reported performance also often leaves the definition of performance to the individual employees. This implies that a self-reported performance measure may rest on very different conceptions of what high performance actually is. One solution is to measure performance as the supervisors' performance appraisals and promotions of the employees (Alonso and Lewis 2001), but this introduces a potential supervisor bias, because supervisors may favor employees with high PSM and promote them more easily or award them in performance appraisals by giving them more credit for their contributions (Wright and Grant (2010: 695). It is therefore highly relevant to supplement the existing PSM-performance literature with a study using an objective performance measure.

The performance measures used in this paper are based on student examination marks in Danish public schools. Although examination marks are a standardized and widely accepted measure of outcome (Chubb and Moe 1990; Andersen & Mortensen, 2010; Meier

and O'Toole 2011), it is important to stress that Danish schools also have other important objectives than the academic qualifications measured in the final exams (such as promoting a "well-rounded development of the individual student" (Ministry of Children and Education, 2012)). Almost all public organizations do, however, have multiple objectives, and academic achievement is among the most important goals in Danish schools (Law no. 998, issued August 16 2010).

The discussion of PSM fit implies that the nuances of the PSM-performance relationship may depend on the context of the organization, and a description of Danish schools is therefore necessary to form specific expectations concerning PSM, its dimensions and students' examination marks. In Denmark, the 98 municipalities are responsible for providing primary and lower secondary education (from preschool class to 9th grade, age 6-15). About 85 percent of Danish children attend public schools which are free, and this study focuses on public schools alone. Expenditure on public schools is financed by municipal taxes, primarily income tax, but extensive grants and equalization schemes eliminate the greater part of financial inequalities between municipalities. At the end of compulsory schooling (9th grade) there is an examination in 11 different subjects (see the data section below for details). The examination is closely related to what is taught in the different subjects in grades 7-9, especially in grade 9, and students typically take the examination seriously. Some examinations are written, and some are oral. Oral examinations are conducted by the teacher and an external examiner. Marks for most written examinations are also given by the teacher and an external examiner, but in some subjects they are given by two external examiners (and not by the teacher). External examiners are appointed by the Ministry of Children and Education. Students also receive marks for the year's work, which are given by

the teacher (analyses of these marks are not shown in this paper, because they do not qualify as an objective performance measure due to the lacking external control).

Given that teaching concerns doing good for students and society, we argue that for Danish teachers PSM has a positive effect on the person-job fit. This means that we expect the aggregate measure of PSM to be positively associated with student examination marks. However, this does not necessarily apply for all the PSM dimensions.

Self-sacrifice implies willingness to deliver services to others without tangible personal rewards or even with personal losses (Perry, 1996), and this PSM dimension represents the basic altruistic or pro-social origins of PSM (Kim and Vandenabeele, 2010). Self-sacrifice can be seen as the foundation on which the other three dimensions rest, and in contrast to the other three PSM dimensions the energy in the self-sacrifice dimension is not channeled in the direction of specific values. It is – in other words - the dimension which most purely can be seen as 'fuel' to obtain any kind of desirable end state for society and others (Andersen et al, 2012a). In its pure form, self-sacrifice is a matter of the degree to which personal rewards are put aside to do good for others. In the case of teaching, a teacher with high self-sacrifice may for example be willing to spend more evenings talking to parents, putting aside personal goods such as time with family and friends in order to do good for the students and their parents. Due to the extra effort given by the teacher, self-sacrifice is expected to be positively associated with student performance.

The three other PSM dimensions are compassion, 'attraction to public policy making', and 'commitment to the public interest', and these dimensions can be seen as based on affective, instrumental/rational and normative motives, respectively (Perry and Wise, 1990; Wise, 2000). The energy contained in these dimensions has more direction than self-

sacrifice in the sense that they are more closely associated with public values (Andersen et al., 2012a) as discussed for each of the dimensions below.

The affective motives behind the PSM dimension compassion are based on identification and emphasize an individual's commitment to or concern for the needs of specific individuals and groups. Given that affective bonding is the emotional basis of serving others, compassion can be seen as being based on the sense of oneness, or the feeling that one could be or could end up in the other person's situation, and it is this identification, which creates a willingness to do good for others (Kim and Vandenabeele 2010). If we apply this to motivation in the class-room, a teacher motivated by compassion identifies with a child or a group of children and is expected to channel this energy into improving both the everyday life of the children and their abilities. It could be argued that this would be more pronounced for children in a poor situation, but given that all children need help to learn, children in general are expected achieve better examination marks when their teacher has a higher level of compassion.

Instrumental/rational motives can also play a role for the delivery of public services. It is based on an understanding of how means and measures can be combined in order to contribute to the delivery of public services. In this case, exercising a particular behavior is an instrument. Participation in policy-making is one such behavior, which can be used to do good for others, as the laws and budgets which frame the delivery of public service can be influenced in order to contribute to the community or society, and it is this type of energy which is measured in the dimension 'attraction to public policy making'. The energy contained in this type of motivation is directed towards decision-making processes at an aggregate level. In the case of teaching, 'attraction to public policy making' is not expected to influence individual performance. 'Attraction to public policy making' may increase

participation in decision-making processes at the school, in the municipality or at the national level and thus increase organizational performance. However, the decisions made in these decision-making bodies will come back to the students in general, and hence it cannot be expected to cause variation in the students' performance in a given teacher's subject compared to the same students' performance in subjects taught by other teachers.

Norm-based motivation (such as the PSM dimension 'commitment to the public interest') is concerned with conforming to values and norms. Individuals are likely to internalize social norms and values regarding the appropriate behavior and societal contributions expected of them when participating in the delivery of public services. Therefore, they feel satisfied when they contribute to realizing these values (Andersen et al. 2012a). But what does "serving society" mean? We argue that this must be answered in a given context, because values can be conflicting (Andersen et al, 2012a & 2012b). It is both a weakness and a strength that the measurement of 'commitment to the public interest' (Perry 1996; Andersen & Pedersen 2012) does not make clear what the public interest is or who defines it. It is a strength, because the measure can be applied in many different contexts, but it is a weakness because we basically do not know the goal of the behavior linked to this type of motivation. Some may see equality as being in the public interest, whereas others argue that competitiveness and efficiency are more important for the public interest. This has important implications for how 'commitment to the public interest' can be expected to be associated to performance. In the case of teaching, the academic competence of the individual pupils is certainly valued by society. This is much evident, when the cross-country comparisons of student abilities are published, and performance in these comparisons is in many countries critiqued for being too low (Grek, 2009). Still, cross-curricular and social competences are also valued. There may be a trade-off. Some may argue that the students

cannot learn if the social climate in the class is not working, but time spent on improving social and inter-curriculum skills may not increase the examination marks as much as teaching aimed at improving the students' academic competences. If there is a trade-off, 'commitment to the public interest' may be negatively associated with performance. If there is not a trade-off, 'commitment to the public interest' is expected to be positively associated with student examination marks. The reason for the conflicting hypotheses in relation to the 'commitment to the public interest' dimension is thus that it is left to the individuals (here the teachers) to define what the public interest is, and we do not know whether there is a trade-off between academic skills and teachers' understanding of the public interest. In consequence, the association between 'commitment to the public interest' and examination marks may be positive or negative. The hypotheses are thus as follows:

Hypothesis 1: *Children taught by teachers with higher PSM get higher examination marks*

Hypothesis 2: *Children taught by teachers with higher self-sacrifice get higher examination marks*

Hypothesis 3: *Children taught by teachers with higher compassion get higher examination marks*

Hypothesis 4: *There is no association between the teachers' attraction to public policy making and the children's examination marks.*

Hypothesis 5a: *Children taught by teachers with higher commitment to the public interest get higher examination marks*

Hypothesis 5b: *Children taught by teachers with higher commitment to the public interest get lower examination marks*

Data

The investigation is based on three sources of data. First, a survey of all teachers at 85 schools conducted from December 2010 until June 2011. Second, to link each individual teacher with each individual student in each subject, we obtained information from these 85 schools on the distribution of students and teachers on classes in grades 7-9 for the three cohorts of students who completed 9th grade in 2009, 2010 or 2011. Different subjects are typically taught by different teachers, and we identified all teachers who taught the selected cohorts of students in nine different subjects for which there are exams: Danish, math, English, history, science, biology, geography, religion, and social studies. (There are also exams in German and French, but since each student chooses only one of these two subjects, they are taught in subject-specific classes which are different from the basic classes in which the other subjects are taught, and therefore, for practical reasons, we did not collect information about the link between students and teachers in these two subjects). In this dataset teachers are identified by names and initials providing a link to the teacher questionnaires which also contain this information, and each student is identified by a unique personal identification number which enables us to link to administrative register information on students, which is the third source of data. The register data contain each individual student's 9th grade examination marks (our performance measure), personal characteristics (such as gender, age, ethnic background), and socioeconomic variables for parents (including education, income, family structure, labour-market status, working experience, etc.).

The survey data from teachers was collected at school staff meetings, where the school principal agreed to let the teachers answer the questionnaire at the meeting. Teachers who were absent from the meetings received a questionnaire and a return envelope. The data from the questionnaires and from schools (about the links between students and teachers)

were collected by 10 research assistants at the Danish Institute for Governmental Research and Aarhus University. Student information with personal identification numbers enabled us to link to the administrative register data which was done with permission from the Danish Data Protection Agency.

The survey to teachers and data collection at schools were conducted at relatively large schools in order to maximize the number of teachers investigated for the given amount of resources available for the data collection. Thus, the 221 largest schools in Denmark were contacted and 38 per cent (85 schools from all parts of Denmark) participated in the study. Most of the remaining 62 per cent did not participate, because they did not have a staff meeting within the timeframe (December 2010 to June 2011) where there was available time for teachers to fill in the questionnaire. Using the administrative register data which are available for all schools, we have tested whether student characteristics and average exam marks of 9^{th} grade students are different for the 85 participating schools compared to the non-participating schools. Differences are small, but some are significant according to two-sample t tests, especially average exam marks are a little higher for participating schools (about 0.04 standard deviations in the distribution of individual marks) and the share of students who are children of immigrants is a little lower (4.7% compared to 6.0% for nonparticipants). The 85 schools are not representative for all Danish schools, because they are bigger. However, if effects of public service motivation are different at large schools compared to small schools, the fact that we have selected large schools might be an advantage in term of the generalizability to other countries, given that Danish schools are smaller than schools in most other countries (Little 2008).

The response rate among staff meeting participants at each school is very high, very close to 100 per cent (only a couple of teachers would not answer). After a review of the

data quality, where suspicious entries were deleted, 3,230 usable responses were retained. 1,383 of these teachers had students who took their final exams in one of the three investigated years, either in 7th, 8th or 9th grade, and taught them in one of the nine subjects which we focus on; 1,188 of these teachers had students from these cohorts in 9th grade. For the 2011 cohort analysed in this paper, the number of 9th grade teachers is 766. The Appendix contains a detailed description of the connection between the three sources of data and the number of observations. The survey questions regarding PSM were based on prior surveys (Perry 1996; Coursey and Pandey 2007; Andersen and Pedersen 2012), and the final questionnaire was adjusted after a pilot survey of 61 teachers in two schools. The scores on the PSM dimensions were calculated in a structural equation model (SEM), clustered on schools. The SRMR indicator is 0.05, indicating that the model has a good fit with data. PSM is an unweighted sum index of the four dimensions. The wording of the PSM items, the structural equation model and Cronbach's alphas for the PSM dimensions can be seen in the Appendix (Tables A.5 and A.6).

The most relevant personal characteristics of the teachers are gender, education, and age or years of teacher experience. Other studies find that females may have higher PSM (Perry, 1997; Pandey & Stazyk, 2008:102), and that age and PSM are correlated (ibid.). Education is relevant because the level of professionalism might affect both PSM and performance (Andersen & Pedersen 2012). Given that student fixed effects are included, we control for all characteristics of individual students, the classes, the schools and the municipalities (because this is constant in comparisons between the examination marks each student receives in different subjects).

The dependent variables in our analyses are examination marks in nine different subjects: Danish, math, English, history, science, biology, geography, religion and social

studies. The marks in Danish used in the analysis are an average of four individual marks (in reading, spelling, essay and oral) and math marks are an average of two individual written marks. The exams in Danish and math, and also the oral examinations in science and English, are mandatory. All students have to take two additional exams, a written exam in biology or geography, and a written exam in English or an oral exam in history, religion or social studies. The distribution of these exams on classes is decided by the Ministry of Children and Education. For students who take the written English exam, the English mark used in the analysis is an average of the marks for the (mandatory) oral exam and the written exam. The reason why we use average marks in case of several individual marks in the same subject (which is the case for Danish, math and, for some students, English) is that we focus on teacher effects, and teacher characteristics are of course constant within subjects for a given class. Marks are given according to a 7-point scale, but to make interpretation of results easier, we use standardized marks which have mean zero and standard deviation unity for each individual mark in a given subject in a given year (calculated for all students at the 85 schools, not just the estimation sample which is restricted to observations with teacher information).

The descriptive statistics for the dependent variables (examination marks) are shown in the top rows of Table 1, which also lists the explanatory variables used in the analyses. The table is based on the student by subject observations used in the estimations. The corresponding number of observations for students, classes and teachers are detailed in Tables 2 and 3 with estimation results (and in the Appendix). The key explanatory variables are teachers' public service motivation (PSM) and its four dimensions. In all estimations we control for basic teacher characteristics: gender, qualifications and experience as teacher. For 51% of the observations the teacher is female, and for 25% the student is also female (we

include this cross term between teacher and student gender since effects of teacher gender may depend on student gender). For 69% of the observations the teacher has specific qualifications (from the teacher education) for teaching in the subject; 8% of teachers have a special teacher education (i.e. a shorter teacher education offered to persons who have relevant qualifications from other types of education), 2% do not have any teacher education, and the remaining 90% have a standard teacher education (the reference category). Teacher experience is 0-4 years for 18% of the observations and 5-9 years for 26% (and at least 10 years for the remaining 56%). For 22% of the student by subject observations the teacher in 9^{th} grade is different from the teacher in 8^{th} grade; for 39% there are at least two different teachers in grades 7-9; and for 6% there are three different teachers. Note that one cannot infer from these numbers that shift of teachers is less frequent from 7^{th} to 8^{th} grade than from 8^{th} to 9^{th} grade (in fact, it is not) since a few subjects are typically not taught in 7^{th} grade (see the Appendix) and in some cases the class has the same teacher in 7^{th} and 9^{th} grade, but another in 8^{th} grade.

The means of the nine subject dummy variables represent the observations' distribution on subjects. This distribution reflects that (almost) all students have exam marks in Danish, math, English and science, whereas each student only take the exam in two of the remaining subjects; see above. The remaining variables are characteristics of students and their parents from the administrative registers; these variables are used as controls in the OLS estimations.

[Table 1 here]

Methods

We did not obtain data for teacher characteristics before grade 7. This may in principle be a weakness since these earlier teacher characteristics may also affect skills at the end of 9th grade. However, in practice, this is not an important weakness because of our identification strategy (student fixed effects, see below), and also because the more basic skills taught in earlier grades are not very closely related to the examination at the end of 9th grade. A standard way to handle this kind of problem in the education production function literature is to estimate the effect of school inputs (in this case teacher characteristics) on achievement gains from an earlier to a later grade (the value added approach; see, e.g. Todd & Wolpin 2003). We are not able to use this strategy since we have no information on student academic achievement before the examination at the end of 9th grade. However, as we argue below, including student fixed effects is essential in order to take account of selection issues, and this will also take account of average (across subjects) lagged student achievement.

The main methodological challenge when investigating the causal effect of teacher characteristics, including credentials and motivation, on student outcomes is the potential bias that arises because the distribution of students and teachers on classes is not random. If, e.g., high-quality teachers sort into classes with more able students (in terms of unobservable characteristics), analyses that fail to address this sorting pattern would produce upward biased estimates of effects of teacher characteristics. Similarly, compensatory assignment (by policy-makers or school administrators) of high-quality teachers to classes with many disadvantaged students may result in downward biased estimates.

We address the problem of non-random sorting in terms of both observable and unobservable characteristics by including student fixed effects in the empirical models, which is possible because each student has examination marks in multiple subjects and different teachers in different subjects. This identification strategy is also applied in Clotfelter et al.

(2007c). More often student fixed effects are used in a different setting where longitudinal data on student test scores in a given subject are available for each student for multiple years (e.g. Rivkin et al. 2005; Clotfelter et al. 2007a, 2007b). In that case student fixed effects control for all time-invariant unobservable characteristics of students (such as ability or motivation) that could be correlated with teacher characteristics. In our application, the student fixed effects control for all unobservable student characteristics (such as overall ability and motivation) which are constant across subjects. To be specific, we estimate models of the form

$$y_{ijsk} = \alpha + T_{ijsk}\beta + \gamma_s + X_{ijsk}\delta + \mu_i + u_{ijsk} \,, \qquad (1)$$

where y_{ijsk} is the examination mark of student *i* in subject *s* taught by teacher *j* in school *k*, T_{ijsk} is a vector of teacher characteristics, γ_s are subject fixed effects, X_{ijsk} is a vector of interaction terms between student characteristics, teacher characteristics and subject dummy variables, μ_i are student fixed effects, u_{ijsk} is the error term, α is the constant term and β and δ are vectors of parameters. Note that even though it is not possible to include main effects of student characteristics (which do not vary by subject) in the model (because of the student fixed effects) it is possible to include interaction terms between student and teacher characteristics, and between student characteristics and subject dummy variables, since these will vary by subject for a given student. We include in all estimations interaction terms between dummy variables for subjects and student gender and immigrant status, since, e.g., boys typically have a comparative advantage in math. Interaction terms may also be included to investigate if some teacher characteristics have differential effects for different groups of students characterised by, e.g. parental background or gender.

We have data for teachers in three grades (7th, 8th and 9th). In principle we could include teacher characteristics for each grade as separate variables, but high correlation

in these variables may be present because many students have the same teacher in a given subject in all three years. In our main analysis we use characteristics of only 9th grade teachers, but include a dummy variable for shift of teachers between grades and cross terms between this variable and all teacher characteristics. In robustness checks we replace 9th grade teacher characteristics by average characteristics of teachers of grades 7-9.

If all subjects are taught in the same basic classes (which is the case in our data), student fixed effects will also take account of class-specific characteristics such as class size and peer-group characteristics. The institutional feature that different subjects are taught in the same basic classes allows us to get a more clear identification of teacher effects than in, e.g., Clotfelter et al. (2007c). Student fixed effects also control for variables at higher levels (e.g., schools and municipalities). Estimates of standard errors are robust taking account of clustering at classes.

In addition to the student fixed effects regression results, we also show, for comparison, results from OLS regressions where we include a wide range of controls (based on register data) for characteristics of students and their parents.

Results

Table 2 shows the main estimation results for the effects of PSM variables on examination marks using student fixed effects models. The first two estimations (columns) of Table 2 are for all exam marks, and the last two are for written exam marks only. Estimations (1) and (3) include the overall PSM index, whereas (2) and (4) include four separate variables for the dimensions of PSM. Table 2 shows parameter estimates for the key PSM variables for 9^{th} grade teachers and cross terms between these variables and a dummy variable which is equal to 1 if the teacher in grade 9 is not the same as in grade 8. We show in robustness checks

below that results are essentially unchanged if we replace this dummy with a dummy for whether there were at least two different teachers in grades 7-9 or if we replace characteristics of 9th grade teachers by average characteristics of 7th-9th grade teachers. All estimations in Table 2 control for teacher gender, education and experience, for cross terms between these teacher characteristics and the 'shift of teacher' dummy, and for subjects and cross terms between subjects and student gender and immigrant status.

For comparison, Table 3 shows the corresponding results using OLS (where a large set of controls for characteristics of students and their parents are included instead of the student fixed effects). Full estimation results of (1) and (3) in Tables 2 and 3, including parameter estimates of controls, are shown in Appendix Table A.4.

Estimation (1) in Table 2 shows a clearly significant and positive association between PSM of 9th grade teachers and their students' examination marks. The point estimate is 0.034. The standard deviation of the PSM index is 1.26 (see Table 1), so an increase in the PSM index of 1 standard deviation is associated with an increase of marks of about 0.04 standard deviations (since marks are standardized). The main effect of shift of teachers from 8th to 9th grade is not statistically significant (although the point estimate indicates a negative effect as one might expect). The coefficient of the cross term between PSM of the 9th grade teacher and the dummy for shift of teachers from 8th to 9th grade is not significantly different from zero. The point estimate is negative. In fact, it is not clear whether one should expect it to be negative or positive. A negative sign might be expected because a new 9th grade teacher can only affect students' learning in one year, which also applies for the potential effect of his or her PSM, and also because the 9th grade teacher's PSM may be a more noisy measure of a perhaps more relevant average PSM index of teachers at different grades. On the other hand,

one might expect a positive sign if PSM of the 9th grade teacher is especially important when teachers in a subject change between 8th and 9th grade.

Estimation (2) in Table 2 indicate very different effects of the four different PSM dimensions: The main effects of 'attraction to public policy making' and self-sacrifice are significant and positive, compassion has a marginally significant positive effect (it is significant at the 10% level, but not at 5%), whereas 'commitment to the public interest' is not significant (with a negative point estimate). The effects of the interaction terms with the teacher shift dummy are not statistically significant, but point estimates are rather large (of the same order of magnitude as the main effects) with the same signs as the main effects for 'commitment to the public interest' and self-sacrifice, but opposite signs for compassion and 'attraction to public policy making'. The main effects of 'attraction to public policy making' and self-sacrifice are about 0.06 and the standard deviation of these variables are about 0.5 (see Table 1), so the point estimates indicate that an increase of 1 standard deviation of these variables is associated with an increase of exam marks of about 0.03 standard deviations (if the class has the same teacher in grades 8 and 9). If teachers change between 8th and 9th grade, point estimates indicate that the effect of self-sacrifice is about twice as large, whereas the effect of 'attraction to public policy making' is about zero. Thus, the results indicate that the self-sacrifice dimension is especially important when teachers change.

Estimations (3) and (4) of Table 2 using only written exam marks produce the same overall pattern of results as estimations (1) and (2). In (3) the main effect of PSM is a little smaller than in (1), although not significantly so, and the cross term is larger numerically and marginally significant (at the 10% level). In (4) self-sacrifice is significant at the 5% level as in (2), but 'attraction to public policy making' is only marginally significant (at the 10% level) and its cross term is numerically large and significant.

The number of (student by subject) observations and the number of different students, classes and teachers in the estimation sample are shown at the bottom of Table 2, and these numbers are discussed in more detail in the Appendix. The number of student by subject observations is reduced considerably (by 44%) when the analysis is restricted to written marks since marks in many subjects are based on oral exams, but the number of different students and classes is reduced by only 4% and the number of different teachers by 27%. The reported R^2 of the regressions may seem small, but it is the R^2 of regressions using within student transformed variables, so it is a measure of the fraction of the variation in marks between subjects within individual students which the model can explain. Similar regressions using OLS and a full set of dummy variables for the individual students produce of course much higher values of R^2. For instance, the adjusted R^2 for such a model corresponding to (1) in Table 2 is 0.56. An F test strongly rejects the hypothesis that all 5630 student fixed effects are zero ($p<0.0001$).

For comparison, Table 3 shows results from OLS estimations corresponding to the student fixed effects estimations in Table 2. Here a large set of variables for characteristics of students and their parents are included; see Table 1 and Appendix Table A.4. The results for the coefficients of PSM and its dimensions are rather similar to the corresponding results of Table 2, especially for estimations (1) and (3). This is somewhat surprising since we would expect the OLS results to be much more inflicted by selection bias; see the discussion in the Methods section above. Comparing estimation (2) of Table 3 with (2) of Table 2, compassion is significant instead of self-sacrifice, but the interaction term between self-sacrifice and teacher shift has a very large and positive coefficient which is marginally significant (at the 10% level). The negative cross terms for ‘commitment to the

public interest' and 'attraction to public policy making' are larger than in Table 2, and the latter is statistically significant in both (2) and (4).

Table A.4 in the Appendix shows full estimation results for the regressions (1) and (3) of Tables 2 and 3. Not many variables for other 9th grade teacher characteristics than PSM are significant. In most cases the interaction effects between teacher characteristics and the variable for shift of teachers have the opposite sign of the main effects (as expected). The point estimates for the female teacher variable and its interaction with female student indicate that having a female teacher is an advantage for female students, but a disadvantage for males. However, none of the parameter estimates are statistically significant, except the interaction term in the FE model for written marks. Surprisingly, the estimates for subject-specific teacher qualifications and type of teacher education are not statistically significant. Less than five years of experience affects exam marks negatively (but the effect is only significant in the estimations for all marks, (1) and (3) in Table A.4), whereas there is no significant difference between having 5-9 years of experience and more than 9 years; these results are consistent with earlier findings, e.g. Clotfelter et al. (2007b). The coefficients of the subject-specific dummy variables are the effect for boys in the given subject compared to Danish (which is the reference category), and the sum of these coefficients and the corresponding cross terms with the female student indicator are the effects for girls. Most of these subject coefficients are clearly significant and reflect that boys obtain lower marks in Danish compared to other subjects, whereas girls obtain higher marks in Danish compared to other subjects. We also included interactions between subject and immigrant status (immigrant or child of immigrant), and many of the coefficients of these are also significant. In the OLS regressions (1) and (2) of Table A.4 (corresponding to (1) and (3) of Table 3) the coefficients for female and immigrant show that Danish marks for females are much higher

than for males, and that Danish marks for immigrants (and children of immigrants) are much lower than for ethnic Danes. Also, family background is very significant as expected. The following background variables are associated with higher exam marks: living in an intact family, high level of parental education, high income of father, parents having an occupation at the high or intermediate level.

Robustness checks

In the estimations in Tables 2 and 3 the 'shift of teachers' variable is a dummy for whether there was a change of teachers from 8^{th} to 9^{th} grade. We chose to use this variable in the main regressions due to an assumption that a change of teachers between 8^{th} and 9^{th} grade is far more important than a change from 7^{th} to 8^{th} grade when we consider effects on exam marks at the end of 9^{th} grade. However, as a robustness check we have replaced this variable with a dummy variable for whether the student had at least two teachers in the given subject during grades 7-9. As shown in Table 1 this is the case for 39% of the observations, whereas 'only' 22% are characterized by a shift from 8^{th} to 9^{th} grade, so it may be important to check if results are robust to the choice of shift variable. Results using the alternative 'shift of teachers' variable (both as a main effect and in the cross terms with PSM variables and other teachers characteristics) are shown in Table 4. Except for this change, the specification of the models estimated in Table 4 is the same as the student fixed effects models in Table 2. The estimates in (1)-(3) in Table 4 are very similar to the corresponding estimates in Table 2. None of the differences are statistically significant. The signs of the main effects of the four PSM dimensions are the same in column (4) of Tables 2 and 4, but the point estimates are numerically about twice as large in Table 4 for compassion, 'commitment to the public interest' and self-sacrifice. Similarly, the cross effects with the teacher shift variable are also

rather different for estimation (4) in the two tables: the signs change for the cross terms for 'commitment to the public interest' and self-sacrifice, and the compassion interaction effect becomes numerically larger and statistically significant.

Table 5 shows results of student fixed effects estimations identical to those in Table 2 except that the variables for PSM, its dimensions, and other teacher characteristics which are controlled for, are based on average values for teachers in grades 7-9 instead of only the 9th grade teachers. Averages are calculated as simple averages of (non-missing) teacher observations for grades 7-9. The estimates of Table 5 are very similar to the estimates of Table 2 for all four models. There are no significant differences. One may of course argue that it is better to use a weighted average of teacher characteristics so that observations for 9th grade teachers weigh higher than for 8th grade teachers, which in turn weigh higher than 7th grade teachers. This pattern would be expected because the rate of 'persistence of knowledge' from one grade to the next is supposed to be smaller than one, and also because the 9th grade teacher may be especially important in relation to preparation of students for the exam. We have tried to use different patterns of weights – e.g. 0.5, 0.3 and 0.2 for grades 9, 8 and 7, respectively – but, not surprisingly, the results are very similar to those in Tables 2 and 5.

Both main analyses and robustness tests clearly supported hypotheses 1 and 2. Children who were taught by teachers with higher PSM get higher examination marks, and when we look at the PSM dimensions, the same is the case for children taught by teachers with high self-sacrifice. Hypothesis 3, which expects children who were taught by teachers with higher compassion to get higher examination marks, is partially supported, given that the association is positive in all specifications and statistically significant in some of them. The same is the case for the association between 'attraction to public policy' and examination marks, and hypothesis 4 (which expects no association) is partially falsified. Finally,

hypothesis 5b (saying that children who were taught by teachers with higher 'commitment to the public interest' get lower examination marks) receives more support than 5a which expects a positive association between 'commitment to the public interest' and examination marks. Nothing final can, however, be concluded, because although the association is consistently negative, it is only statistically significant in one of the specifications.

Discussion of causality: Does PSM *affect* performance?

Our specific research question is whether PSM and individual performance is positively related, also when performance is objectively measured, and the result section above accordingly discusses associations, rather than effects. But the most interesting question, which also is in the title of the paper, is how Public Service Motivation affects performance in schools (and in other organizations). Given that we use a cross-sectional survey design, we cannot draw firm conclusions regarding causality due to potential endogeneity problems. We will, however, argue below that the two broad classes of rival explanations (reverse causality and omitted variables) are less problematic for this study than for many existing studies of PSM and performance.

The first rival explanation is that high performance may strengthen PSM, while low performance weakens PSM instead of the proposed effect of PSM on performance (Wright and Grant, 2010: 695). Although we do not have panel data with information about the same teachers' PSM over time, we still have panel data on their students' examination marks (for three years, 2009-2011). For the last year (2011), PSM is measured before the examination marks are given, and this is the analyses presented in this paper. If the correlation between PSM and examination marks is due to reverse causality, we would expect higher correlation between PSM and marks of earlier cohorts of students. However, when we

rerun the analyses using data for all three years, , the estimated PSM coefficients are smaller (approximately 30% smaller) compared to the estimations with only 2011 data (while the t-values are approximately the same, since the standard errors are smaller in the analyses with more observations). This indicates that reverse causality is not a problem. Note that we do find a (smaller) positive correlation between PSM measured in 2011 and performance (student marks) measured in earlier years, but this would be expected also without any reverse causality because PSM is probably rather persistent from year to year.

The second rival explanation is that there is a common cause of both PSM and performance. Potentially, this may be caused by non-random selection of students and teachers into classes, but as argued above our student fixed effects approach effectively controls for this. However, we cannot rule out that unobserved teacher characteristics might be correlated with both PSM and performance. Wright and Grant (2010: 695) thus argue that conscientiousness (together with supervisor biases which we will discuss below) is a very important omitted variable in many studies. The personality trait of conscientiousness refers to the degree to which individuals tend to be industrious, disciplined, goal oriented, and organized. Given that it is a robust predictor of job performance across a wide range of occupations, and given that there is reason to believe that conscientiousness will be positively associated with PSM, because a sense of duty and responsibility to others is one of the defining features of conscientiousness, Wright and Grant argue (2010: 695) that researchers should examine whether PSM predicts higher performance even after controlling for conscientiousness. A counter argument is, however, that this would be difficult and not necessarily fruitful, because of the conceptual overlap between the concepts, especially for the self-sacrifice dimension of PSM. Consequently, we have followed Brewer's (2008: 146) recommendation and have instead begun to unpack the PSM concept and strategically

explore sub-dimensional relationships, and this allows us to discuss more specifically what it is in PSM which is positively related to performance.

Although our results very robustly show a positive association between PSM and performance, the only dimension, which consistently has a positive, significant association with examination marks in all student fixed effects analyses, is self-sacrifice. The coefficients are consistently positive for 'attraction to public policy making' and compassion and consistently negative for 'commitment to the public interest', but the statistical significance of these associations cannot be robustly established. The findings concerning the PSM dimensions differ from existing studies (e.g. Leisink & Steijn 2009, Vandenabeele 2009, Andersen and Serritzlew 2012) which tend to find that 'commitment to public interest' is positively correlated with PSM. Our findings concerning the dimensions are not robust enough to make any final conclusions and (as argued above) the nuances of the PSM-performance relationship may depend on the context due to the importance of the PSM fit. Still, we argue that PSM research should take up this challenge and continue along this avenue of research and both look at overall PSM and separate PSM dimensions, respectively. The argument for maintaining the focus on PSM as the conceptual frame is that it has an overarching meaning, given that all the dimensions concern different types of orientations to delivering service to people with the purpose of doing good for others and society. This should not, however, blind us to the possibility that different PSM dimensions can affect different types of performance differently.

One of the key contributions in this study is that we measure performance objectively, using administrative register data. We therefore avoid social desirability bias which can be a problem in self-reported performance data (Kim and Kim, 2012), and we argue that it also prevents supervisor biases (and therefore another type of omitted variable

bias). Many of the studies linking PSM to performance have measured performance as the supervisors' performance appraisals and/or promotions of the employees (e.g. Alonso and Lewis 2001), and Wright and Grant (2010: 695) argue that another possible explanation for the identified positive association (except from confounding from conscientiousness) is that supervisors are biased in favor of employees with high PSM and award them in performance appraisals by giving them more credit for their contributions. As argued by Wright and Grant (ibid.), this may even skew objective performance measures if the supervisors offer employees with high PSM more resources and support than employees with low PSM. This is not, however, a serious problem in Danish schools, because rigid workload agreements between the teachers' union and the municipalities regulate how much time a given teacher get to his/her teaching, and there is very little other variation between teachers on the same school in the resources available for them for teaching. We agree with Wright and Grant (2010: 696) when they argue that randomized, controlled field experiments with interventions designed to increase PSM would be very desirable in further studies of the PSM-performance association, but we also argue that this study brings us a huge step forward by providing a test of the association between PSM (and its dimensions) and performance measured in a very objective way. Although we cannot say for sure whether PSM actually affects performance in schools, the research design with student fixed effects and objective performance data at least renders a positive effect more probable.

Conclusion

We set out to investigate whether students taught by teachers with higher PSM receive higher examination marks (i.e. perform better). Concerning the PSM dimensions, we hypothesized that compassion and self-sacrifice would be positively associated with performance, that

‘attraction to public policy making’ would not be associated with performance, and that ‘commitment to the public interest’ could both be positively and negatively associated with performance, depending on potential tradeoffs between teachers’ perceptions of the desirable for society and others. Our expectations were confirmed for overall PSM and for self-sacrifice, while the associations were positive (but not consistently significant) for compassion and ‘attraction to public policy making’ and negative for ‘commitment to the public interest’. While the results on the PSM dimensions are not conclusive, our finding concerning overall PSM is clear: Higher PSM is associated with higher performance.

Our test is strong in three ways. First, the method (student fixed effect regression) very robustly insures that student background and selection effects do not confound the results. Second, we analyze both overall PSM and the dimensions separately. Third and most importantly, we show that the positive association between PSM and performance identified in studies using self-reported performance measures can also be found in this study in which the performance measure is objective and based on register data. We thus contribute to establishing the link between PSM and performance much more robustly.

Obviously, our study also has limitations. While our approach takes account of important selection issues and our analysis does not indicate reverse causality, we cannot rule out that part of the estimated ‘effect’ of PSM on the objective performance measure may be due to unobserved teacher characteristics which are correlated with both PSM and performance. The study holds the institutional context constant (both at the school, municipal and country level), which strengthens the internal validity, but in future studies it would also be interesting to analyze the relationships between context, PSM and performance. In terms of external validity, the specific results do not necessarily apply to countries other than Denmark, but nothing indicates that the overall findings cannot be generalized analytically.

Still, it would be very interesting to apply the same approach for data in other countries. If possible, it would also increase generalizability to do a similar analysis for other occupations where the individual performance of different public employees can be compared for the same user.

Combined with other studies of the PSM-performance relationship, our findings strongly support that PSM is relevant for performance in organizations where the goal is linked to "the public good" as seen by the employees. We also think that the finding that the dimensions might have different associations with important performance indicators can be generalized, although we do not necessarily think that self-sacrifice is positively correlated with performance in all contexts. We know from other studies that 'commitment to the public interest' has often been shown to be positively correlated with performance, and that makes the negative coefficients in this paper the more interesting. Still, the key contribution is that we link PSM and performance more robustly, showing that PSM is also relevant for objectively measured performance.

Appendix.

Numbers of observations in data from schools, teacher survey and registers

Table A.1 shows the numbers of students, student-subject observations, teachers and classes for the three cohorts of students who completed 9th grade in 2009-2011. The first four columns show the numbers in the dataset collected from the 85 schools which links teachers to students. For 9th grade there are 16,858 students, 147,849 student-subject observations, 1,806 teachers and 833 classes. Since there are 9 subjects, the number of student-subject observations for 9th grade should be 9 times the number of student observations (i.e. 151,722), but the observed number (147,849) is about 2.5% smaller which is due to missing

or incomplete data on teachers in some subjects from a few schools. The problem with incomplete data from schools is larger for 7th and 8th grade, but the main reason for the rather low number of student-subject observations for especially 7th grade is the fact that not all subjects are taught in all three grades; thus religion and social studies are typically not taught in 7th grade (which is in accordance with recommendations from the Ministry of Children and Education). The last four columns of Table A.1 show the corresponding numbers when we only consider observations with matched teacher data from the survey. Thus, 1,188 (66%) of the 1,806 teachers who had 9th grade classes responded the questionnaire and could be linked to the students in the dataset. The 34% missing observations on 9th grade teachers are due to the fact that some teachers were not attending the staff meeting at which the questionnaire were handed out (and the response rate of these teachers is rather low), but another important reason is that some of the teachers who had 9th grade students in 2009 and 2010 were no longer at the school in 2011 (because of mobility to other jobs or retirement). Thus, it is not surprising that missing teacher survey observations is a more serious problem for 7th and 8th grade than for 9th grade, nor that the problem is less serious for 9th grade teachers in 2011 (the year of data collection), where we have survey data for 75% of the teachers (766 out of 1,017); see Table A.2.

[Tables A.1, A.2 and A.3 here]

Table A.3 shows the number of observations (with survey data on teachers) after merging with register data for students' marks and socioeconomic background. The register data contains all students who have marks for the year's work in one of the years 2009-2011. All students with exam marks also have marks for the year's work. Comparing the first column of Table A.3 with the last column of Table A.1 shows that about 4% of the student (and

student-subject) observations are lost due to missing register data (and as a consequence, about 1% of classes and matched teachers are lost). The missing register data may be due to students' school mobility in the months before the exam, dropouts, exemption from examination for special education students, or errors in registration of personal identification numbers at schools. Similar percentages of observations with missing register data apply if we consider the 2011 cohort separately; see the fourth column of Table A.3 and the last column of Table A.2. Comparing columns 1 and 2 (and similarly 4 and 5 for the 2011 cohort) in Table A.3 it will be seen that the number of student-subject observations is much smaller for examination marks than for marks for the year's work. This is because all students completing 9th grade will have marks for the year's work in each subject, whereas not all students' will take exams in all subjects, and especially because each student typically only has one exam mark in either biology or geography (but not in both of these subjects), and only one in either history, religion, social studies, or written English (see the data section). The number of student-subject observations is also reduced considerably when only subjects with written exams are considered; see the third and sixth columns of Table A.3.

Full estimation results for the main regressions

Table A.4 contains the full OLS and student fixed effects regression results corresponding to columns (1) and (3) of Tables 2 and 3.

[Table A.4 here]

PSM factor analysis

Table A.5 shows the structural equation model of the PSM dimensions, while Table A.6 shows Cronbach's Alphas together with the covariance between the PSM dimensions. The

only problematic dimension is (like in many other PSM studies, see Ritz 2011) 'attraction to public policy making' where the standardized factor loadings and squared multiple correlation coefficients are lower than for the other dimensions, and Cronbach's alpha is below the 0.7 threshold. RMSEA (Root Mean Square Error Of Approximation) is 0.027, CFI is 0.954, and TLI is 0.944, and this all indicates that the model has a good fit with the data.

[Tables A.5 and A.6 here]

References

Alonso, Pablo & Gregory B. Lewis (2001) Public service motivation and job performance. Evidence From the Federal Sector *American Review of Public Administration*, 31 (4): 363-380.

Andersen, L. B. & Pedersen, L. H. (2012) Public Service Motivation and Professionalism *International Journal of Public Administration* 35 (1): 46-57.

Andersen, L.B. & A.M. Kjeldsen (in press) Public Service Motivation, User Orientation and Job Satisfaction: A Question of Employment Sector? *International Public Management Journal.*

Andersen, L.B.; Beck Jørgensen, T; Kjeldsen, AM; Pedersen, LH; Vrangbæk, K. (2012a) Public Values and Public Service Motivation: Conceptual and Empirical Relationships. *American Review of Public Administration.* Early view. http://arp.sagepub.com/content/early/2012/03/26/0275074012440031.abstract

Andersen, Lotte Bøgh & Søren Serritzlew (2012): Does public service motivation affect the behavior of professionals? *International Journal of Public Administration* 35 (1): 19-29.

Andersen, Lotte Bøgh, Torben Beck Jørgensen, Anne-Mette Kjeldsen, Lene Holm Pedersen and Karsten Vrangbæk (2012b): Public Value Dimensions – Developing and Testing a Multidimensional Classification. *International Journal of Public Administration*, 35(11):715-728.

Andersen, Simon Calmar and Peter B. Mortensen (2010) Policy Stability and Organizational Performance: Is There a Relationship? *Journal of Public Administration Research & Theory* 20 (1): 1-22.

Andrews, Rhys, George A. Boyne and Richard M. Walker (2006) "Subjective and objective measures of organizational performance: An empirical exploration" pp. 14-334 in George A. Boyne, Kenneth J. Meier, Laurence J. O'Toole, Jr.,Richard M. Walker (eds.) *Public Service Performance: Perspectives on Measurement and Management* Cambridge: Cambridge University Press.

Boyne, G. A. (2003). What is public service improvement? *Public Administration* 81(2), 211-227.

Brewer, G. A. (2008). Employee and Organizational Performance. In J. L. Perry & A. Hondeghem (Eds.), *Motivation in Public Management. The Call of Public Service* (pp. 136-156). Oxford: Oxford University Press.

Clotfelter, C.T.; H.F. Ladd & J.L. Vigdor (2007a): How and why do teacher credentials *matter for student achievement?* NBER Working Paper 12828.

Clotfelter, C.T.; H.F. Ladd & J.L. Vigdor (2007b): Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26: 673-682.

Clotfelter, C.T.; H.F. Ladd & J.L. Vigdor (2007c): *Teacher credentials and student achievement in high school: a cross-subject analysis with student fixed effects*. NBER Working Paper 13617.

Dixit, Avinash (2002). "Incentives and Organizations in the Public Sector. An Interpretative Review", *Journal of Human Resources*, 37, pp. 696-727

DLF (2012) *Professionsetik og lærergerning - et baggrundspapir* [Professional ethics and the teacher job – a background paper, published by the Danish Union of Teacher] http://www.dlf.org/undervisning/profession+og+kompetence/professionsideal accessed 210912).

Flynn, N. (2007). *Public Sector Management* (5th ed.). London: Thousand Oaks, New Delhi: SAGE.

Freidson, E. (2001). *Professionalism. The third logic*. Cambridge: Polity Press.

Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe, *Journal of Education Policy*, 24 (1): 23-39.

Hondeghem, A. & Perry, J. L. (2009). EGPA symposium on public service motivation and performance: Introduction. *International Review of Administrative Sciences* 75, 5-9.

Kim, S. & Vandenabeele, W. 2010. 'A Strategy for Building Public Service Motivation Research Internationally', *Public Administration Review* 70, 701-9.

Kim, Seung Hyun & Sangmook Kim (2012) *Cultural Diversity and Social Desirability Bias in Measuring Public Service Motivation* Prepared for presentation at the 2012 EGPA Annual Conference, Bergen, Norway, September 5-8, 2012.

Kjeldsen, A.M. & Andersen, L.B. (in press) How Pro-social Motivation Affects Job Satisfaction. An international analysis of countries with different welfare-state regimes. *Scandinavian Political Studies.*

Krejsler, John (2005). Professions and their Identities: How to explore professional development among (semi-)professions, *Scandinavian Journal of Educational Research*, 49:4, 335-357

Kristof-Brown, A. L., Zimmerman, R. D. & Johnson, E. C. (2005). Consequences of individuals' fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Personnel Psychology* 58(2), 281–342.

Leisink, P. & Steijn, B. (2009). Public service motivation and job performance of public sector employees in the Netherlands. *International Review of Administrative Sciences* 75, 35-52.

Little, A.W. 2008. 'Size Matters for EFA Research Monograph No 26'. http://www.create-rpc.org/pdf_documents/PTA26.pdf (accessed May 23 2012).

Miller, G. J. & Whitford, A. B. (2007). The Principal's Moral Hazard: Constraints on the Use of Incentives in Hierarchy. *Journal of Public Administration Research and Theory* 17(2), 213-233.

Ministry of Children and Education (2012) *The Aims of the "Folkeskole"* http://eng.uvm.dk/Education/Primary-and-Lower-Secondary-Education/The-Folkeskole/The-Aims-of-the-Folkeskole (accessed October 9th 2012).

Naff, K. C. & J. Crum (1999), "Working for America: Does Public Service Motivation Make a Difference?" *Review of Public Personnel Administration*, 19 (4): 5-16.

Pandey, S. K. & Stazyk, E. C. (2008). Antecedents and Correlates of Public Service Motivation. In J. L. Perry & A. Hondeghem (Eds.), *Motivation in Public Management: The Call of Public Service* (pp.101-117). Oxford: Oxford University Press.

Perry, J. L. & Wise, L. R. (1990). The Motivational Bases of Public Service. *Public Administration Review*, 50, 367-373.

Perry, J. L. (1996). Measuring public service motivation: An assessment of construct reliability and validity. *Journal of Public Administration Research and Theory*, 6(1), 5-24.

Perry, J. L. (1997). Antecedents of Public Service Motivation. *Journal of Public Administration Research and Theory* 7(2), 181-197.

Perry, J. L., Hondeghem, A. and Wise, L. R. (2010), Revisiting the Motivational Bases of Public Service: Twenty Years of Research and an Agenda for the Future. Public Administration Review, 70: 681–690. doi: 10.1111/j.1540-6210.2010.02196.x

Perry, J. L., Mesch, D. & Paarlberg, L. (2006). Motivating Employees in a New Governance Era: The Performance Paradigm Revisited. *Public Administration Review* 66(4), 505-514.

Petrovsky, N. & Ritz, A. (2010). *Do Motivated Elephants Gallop Faster? An Assessment of the Effect of Public Service Motivation on Government Performance at the Individual and Organizational Levels.* Paper presented to the International Conference on Public Service Motivation (PSM) and Public Performance in a Globalized World, Huazhong University of Science & Technology, Wuhan China, July 15-17, 2010

Rainey, H. G. & Steinbauer, P. (1999). Galloping elephants: Developing elements of a theory of effective government organizations. *Journal of Public Administration Research*

Rasmussen, Jens (1993) *Lærerprofessionalisme*. Copenhagen: Unge pædagoger.

Ritz, A. (2011), Attraction to Public Policy-Making: A qualitative inquiry into improvements in PSM measurement. *Public Administration*, 89: 1128–1147.

Rivkin, S.; E.A. Hanushek & J.F. Kain (2005): Teachers, schools, and academic achievement. *Econometrica*, 73(2): 417-457.

Steijn, B. (2008). 'Person-environment Fit and Public Service Motivation', *International Public Management Journal* 11, 13-27.

Taylor, J. (2008). "Organizational Influences, Public Service Motivation and Work Outcomes: An Australian Study." *International Journal of Public Management* 11(1): 67-88.

Todd, P.E. & K.I. Wolpin (2003): On the specification and estimation of the production function for cognitive achievement. *Economic Journal*, 113: F3-F33.

Vandenabeele, W. (2009). The mediating effect of job satisfaction and organizational commitment on self-reported performance: more robust evidence of the PSM-performance relationship. *International Review of Administrative Sciences* 75, 11-34.

Wise, Lois Recascino (2000). The Public Service Culture. In *Public Administration Concepts and Cases,* 7th ed. Edited by Richard J. Stillman II, 342-353. Boston: Hougnton Mifflin.

Wright, B. E. and Grant, A. M. (2010), Unanswered Questions about Public Service Motivation: Designing Research to Address Key Issues of Emergence and Effects. *Public Administration Review*, 70 (5): 691–700.

Table 1. Summary statistics: student by subject observations for the 2011 cohort estimation sample

	count	mean	sd	min	max
All exam marks	24360	0.000	0.958	-3.149	1.955
Written exam marks	13660	0.004	0.942	-3.149	1.955
Teacher characteristics:					
PSM	24360	0.032	1.261	-4.650	3.089
Compassion	24360	-0.018	0.520	-2.683	0.649
Commitment to the public interest	24360	0.005	0.333	-1.435	0.594
Attraction to public policy making	24360	0.020	0.489	-1.007	1.111
Self-sacrifice	24360	0.025	0.504	-1.459	1.197
Female teacher	24360	0.511	0.500	0.000	1.000
Female teacher & student	24360	0.253	0.435	0.000	1.000
Qualifications in subject	24360	0.689	0.463	0.000	1.000
Special teacher education	24360	0.079	0.269	0.000	1.000
No teacher education	24360	0.020	0.140	0.000	1.000
Experience 0-4 years	24360	0.178	0.383	0.000	1.000
Experience 5-9 years	24360	0.260	0.439	0.000	1.000
Shift of teachers (from 8^{th} to 9^{th} grade)	24360	0.219	0.414	0.000	1.000
2+ teachers in grades 7-9 in the subject	24360	0.387	0.487	0.000	1.000
3 teachers in grades 7-9 in the subject	24360	0.059	0.235	0.000	1.000
Dummy variables for subjects:					
Danish	24360	0.184	0.388	0.000	1.000
Math	24360	0.182	0.386	0.000	1.000
English	24360	0.186	0.389	0.000	1.000
History	24360	0.040	0.196	0.000	1.000
Science	24360	0.166	0.372	0.000	1.000
Biology	24360	0.078	0.269	0.000	1.000
Geography	24360	0.085	0.278	0.000	1.000
Religion	24360	0.042	0.200	0.000	1.000
Social studies	24360	0.037	0.188	0.000	1.000
Variables for students and their parents:					
Female student	24360	0.495	0.500	0.000	1.000
Immigrant student	24360	0.029	0.167	0.000	1.000
Child of immigrant	24360	0.047	0.211	0.000	1.000
Mother not in register	24360	0.009	0.097	0.000	1.000
Father not in register	24360	0.039	0.193	0.000	1.000
Intact family	24360	0.647	0.478	0.000	1.000
One child in family	24360	0.396	0.489	0.000	1.000
Three children in family	24360	0.157	0.364	0.000	1.000
Four or more children	24360	0.042	0.202	0.000	1.000
Mother vocational education	24360	0.376	0.484	0.000	1.000
Mother further education	24360	0.355	0.478	0.000	1.000
Mother higher education	24360	0.088	0.284	0.000	1.000
Mother's education missing	24360	0.023	0.151	0.000	1.000
Father vocational education	24360	0.408	0.491	0.000	1.000
Father further education	24360	0.256	0.437	0.000	1.000
Father higher education	24360	0.120	0.326	0.000	1.000

	count	mean	sd	min	max
Father's education missing	24360	0.054	0.227	0.000	1.000
Log income mother	24360	5.387	0.979	0.000	8.101
Log income father	24360	5.505	1.421	0.000	8.777
Mother self employed	24360	0.043	0.202	0.000	1.000
Mother higher occupation	24360	0.167	0.373	0.000	1.000
Mother intermediate occupation	24360	0.249	0.432	0.000	1.000
Mother other occupation	24360	0.101	0.302	0.000	1.000
Mother unemployed/not in labour force	24360	0.129	0.336	0.000	1.000
Father self employed	24360	0.081	0.272	0.000	1.000
Father higher occupation	24360	0.211	0.408	0.000	1.000
Father intermediate occupation	24360	0.154	0.361	0.000	1.000
Father other occupation	24360	0.156	0.363	0.000	1.000
Father unemployed/not in labour force	24360	0.106	0.308	0.000	1.000

Reference categories are: Teacher experience at least 10 years; teacher education standard; mother (father) has no education beyond compulsory school (9^{th} grade); mother (father) has basic occupation.

Table 2. Effects of 9th grade teacher PSM (and its four dimensions) on student examination marks: Student fixed effects regressions

	(1)	(2)	(3)	(4)
	All exam marks		Written exam marks	
PSM	0.0344***		0.0235*	
	(0.00945)		(0.00932)	
PSM * shift of teachers	-0.0131		-0.0338	
	(0.0142)		(0.0185)	
Compassion		0.0419		0.0243
		(0.0218)		(0.0273)
Commitment to the public interest		-0.0494		-0.0677
		(0.0429)		(0.0529)
Attraction to public policy making		0.0601**		0.0491
		(0.0208)		(0.0268)
Self-sacrifice		0.0629*		0.0639*
		(0.0270)		(0.0286)
Compassion * shift of teachers		-0.0321		-0.0581
		(0.0380)		(0.0413)
Public interest * shift of teachers		-0.0635		-0.0811
		(0.0823)		(0.0892)
Attr. to policy * shift of teachers		-0.0725		-0.121*
		(0.0369)		(0.0523)
Self-sacrifice * shift of teachers		0.0600		0.0457
		(0.0534)		(0.0593)
Shift of teachers (from 8^{th} to 9^{th} grade)	-0.0279	-0.0368	-0.0262	-0.0465
	(0.0395)	(0.0396)	(0.0526)	(0.0552)
Observations	24360	24360	13660	13660
Students	5631	5631	5422	5422
Classes	280	280	269	269
Teachers	694	694	509	509
R^2 (within student)	0.035	0.037	0.068	0.071

All estimations include controls for teacher characteristics and subjects, and interaction terms between dummy variables for subjects and students' gender and immigrant status; see Appendix Table A.4 for details.
Standard errors in parentheses – robust standard errors clustered on classes.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3. Effects of 9th grade teacher PSM (and its four dimensions) on student examination marks: OLS regressions

	(1)	(2)	(3)	(4)
	All exam marks		Written exam marks	
PSM	0.0330***		0.0249*	
	(0.00875)		(0.0110)	
PSM * shift of teachers	-0.0135		-0.0490*	
	(0.0147)		(0.0214)	
Compassion		0.0486*		0.0510
		(0.0232)		(0.0291)
Commitment to the public interest		-0.00781		-0.00584
		(0.0429)		(0.0572)
Attraction to public policy making		0.0526*		0.0535
		(0.0224)		(0.0321)
Self-sacrifice		0.0321		0.00222
		(0.0264)		(0.0330)
Compassion * shift of teachers		-0.0309		-0.0710
		(0.0416)		(0.0488)
Public interest * shift of teachers		-0.0841		-0.171
		(0.0865)		(0.107)
Attr. to policy * shift of teachers		-0.106*		-0.146*
		(0.0412)		(0.0656)
Self-sacrifice * shift of teachers		0.0893		0.103
		(0.0515)		(0.0640)
Shift of teachers (from 8th to 9th grade)	0.0120	-0.00446	0.0435	0.0210
	(0.0435)	(0.0440)	(0.0601)	(0.0637)
Observations	24360	24360	13660	13660
Students	5631	5631	5422	5422
Classes	280	280	269	269
Teachers	694	694	509	509
Adjusted R^2	0.164	0.164	0.182	0.183

All estimations include controls for student socioeconomic background, teacher characteristics and subjects, and interaction terms between dummy variables for subjects and students' gender and immigrant status; see Appendix Table A.4 for details.
Standard errors in parentheses – robust standard errors clustered on classes.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. Effects of 9th grade teacher PSM (and its four dimensions) on student examination marks: Student fixed effects regressions. Robustness check: Shift of teachers between 7th and 8th grade, and/or between 8th and 9th.

	(1)	(2)	(3)	(4)
	All exam marks		Written exam marks	
PSM	0.0311**		0.0247*	
	(0.0107)		(0.0105)	
PSM * shift of teachers	0.0000772		-0.0218	
	(0.0140)		(0.0175)	
Compassion		0.0431		0.0527
		(0.0242)		(0.0294)
Commitment to the public interest		-0.0595		-0.147*
		(0.0469)		(0.0574)
Attraction to public policy making		0.0570*		0.0315
		(0.0252)		(0.0292)
Self-sacrifice		0.0611*		0.105***
		(0.0301)		(0.0305)
Compassion * shift of teachers		-0.0198		-0.0945*
		(0.0366)		(0.0447)
Public interest * shift of teachers		-0.0187		0.113
		(0.0719)		(0.0862)
Attr. to policy * shift of teachers		-0.0312		-0.0170
		(0.0382)		(0.0442)
Self-sacrifice * shift of teachers		0.0424		-0.0516
		(0.0445)		(0.0515)
Shift of teachers (at least 2 teachers in the subject in $7^{th} – 9^{th}$ grade)	-0.00694	-0.00892	0.0286	0.0237
	(0.0339)	(0.0342)	(0.0494)	(0.0514)
Observations	24360	24360	13660	13660
R^2 (within student)	0.035	0.036	0.069	0.072

All estimations include controls for teacher characteristics and subjects, and interaction terms between dummy variables for subjects and students' gender and immigrant status.
Standard errors in parentheses – robust standard errors clustered on classes.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5. Effects of 9th grade teacher PSM (and its four dimensions) on student examination marks: Student fixed effects regressions. Robustness check: Teacher characteristics are defined as an (unweighted) average of characteristics of teachers in 7th, 8th and 9th grade

	(1)	(2)	(3)	(4)
	All exam marks		Written exam marks	
PSM	0.0362***		0.0263**	
	(0.00981)		(0.00964)	
PSM * shift of teachers	-0.0201		-0.0480*	
	(0.0172)		(0.0232)	
Compassion		0.0471*		0.0390
		(0.0225)		(0.0286)
Commitment to the public interest		-0.0525		-0.0935
		(0.0450)		(0.0561)
Attraction to public policy making		0.0582*		0.0462
		(0.0226)		(0.0280)
Self-sacrifice		0.0680*		0.0788**
		(0.0283)		(0.0283)
Compassion * shift of teachers		-0.0429		-0.0701
		(0.0460)		(0.0536)
Public interest * shift of teachers		-0.0551		-0.0956
		(0.0957)		(0.113)
Attr. to policy * shift of teachers		-0.0851		-0.145**
		(0.0435)		(0.0556)
Self-sacrifice * shift of teachers		0.0523		0.0405
		(0.0624)		(0.0689)
Shift of teachers (from 8th to 9th grade)	-0.0134	-0.0242	-0.0114	-0.0378
	(0.0408)	(0.0412)	(0.0551)	(0.0606)
Observations	24360	24360	13660	13660
R^2 (within student)	0.034	0.036	0.067	0.070

All estimations include controls for teacher characteristics and subjects, and interaction terms between dummy variables for subjects and students' gender and immigrant status.
Standard errors in parentheses – robust standard errors clustered on classes.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.1. Numbers of students, student-subject observations, teachers and classes for all three cohorts

	Dataset from schools				With matched teacher information			
Grade	7-9	7	8	9	7-9	7	8	9
Students	16858	12952	15415	16858	16797	12614	15212	16778
Student-subjects	364819	88041	128929	147849	246553	53843	86820	105890
Teachers	2454	1605	1842	1806	1383	930	1144	1188
Classes	2286	666	787	833	2254	648	777	829

Table A.2. Numbers of students, student-subject observations, teachers and classes for the 2011 cohort

	Dataset from schools				With matched teacher information			
Grade	7-9	7	8	9	7-9	7	8	9
Students	5895	5061	5539	5895	5895	4951	5450	5895
Student-subjects	133548	34318	46556	52674	95142	21759	32658	40725
Teachers	1442	836	984	1017	909	515	660	766
Classes	826	257	282	287	816	251	278	287

Table A.3. Numbers of 9th grade students, student-subject observations, teachers and classes after merging with register data for students' marks and socioeconomic background

	2009-2011 cohorts			2011 cohort		
	Marks for the year's work	Exam marks	Written exam marks	Marks for the year's work	Exam marks	Written exam marks
Students	16169	15986	15379	5679	5631	5422
Student-subjects	101598	63377	36088	39056	24360	13660
Teachers	1178	1129	972	747	694	509
Classes	820	813	784	281	280	269

Table A.4. Effects of 9th grade teacher PSM on student examination marks: Full OLS and Student FE regression results corresponding to columns (1) and (3) of Tables 2 and 3

	(1) Exam marks OLS	(2) Written marks OLS	(3) Exam marks FE	(4) Written marks FE
PSM	0.0330***	0.0249*	0.0344***	0.0235*
	(0.00875)	(0.0110)	(0.00945)	(0.00932)
PSM * shift of teachers	-0.0135	-0.0490*	-0.0131	-0.0338
	(0.0147)	(0.0214)	(0.0142)	(0.0185)
Female teacher	-0.0207	-0.0158	-0.0251	-0.00595
	(0.0284)	(0.0332)	(0.0256)	(0.0279)
Female teacher & student	0.0583	0.0522	0.0685**	0.0320
	(0.0313)	(0.0347)	(0.0234)	(0.0277)
Qualifications in subject	0.0180	0.0147	0.00375	0.00232
	(0.0214)	(0.0278)	(0.0216)	(0.0244)
Special teacher education	-0.0429	-0.0414	0.0253	0.0391
	(0.0466)	(0.0631)	(0.0447)	(0.0561)
No teacher education	-0.197	-0.156	-0.0885	-0.158
	(0.113)	(0.156)	(0.122)	(0.188)
Experience 0-4 years	-0.0604*	-0.0396	-0.0666*	-0.0577
	(0.0268)	(0.0354)	(0.0275)	(0.0350)
Experience 5-9 years	0.0171	-0.00887	0.0198	0.0305
	(0.0277)	(0.0364)	(0.0271)	(0.0294)
Shift of teachers (from 8^{th} to 9^{th} grade)	0.0120	0.0435	-0.0279	-0.0262
	(0.0435)	(0.0601)	(0.0395)	(0.0526)
Female teacher * shift of teachers	-0.0341	-0.0128	0.0468	0.0360
	(0.0459)	(0.0584)	(0.0441)	(0.0531)
Female teacher&student * shift	-0.0641	-0.0664	-0.0959*	-0.0897
	(0.0465)	(0.0507)	(0.0416)	(0.0498)
Qualifications in subject * shift	-0.0372	-0.0556	-0.0368	0.0404
	(0.0425)	(0.0588)	(0.0396)	(0.0488)
Experience 0-4 yrs. * shift	0.0545	-0.0145	0.0293	-0.00574
	(0.0500)	(0.0604)	(0.0516)	(0.0567)

Experience 5-9 yrs. * shift	-0.00770	0.0177	-0.0463	-0.150*
	(0.0473)	(0.0764)	(0.0485)	(0.0626)
Special teacher education * shift	0.0574	0.0800	0.0584	0.0910
	(0.0712)	(0.103)	(0.0704)	(0.0894)
No teacher education * shift	0.106	-0.182	0.0318	-0.0557
	(0.153)	(0.182)	(0.137)	(0.226)
Math	0.209***	0.203***	0.226***	0.200***
	(0.0277)	(0.0285)	(0.0254)	(0.0264)
English	0.129***	0.262***	0.131***	0.266***
	(0.0251)	(0.0641)	(0.0246)	(0.0534)
History	0.187***		0.155***	
	(0.0514)		(0.0404)	
Science	0.118***		0.119***	
	(0.0335)		(0.0314)	
Biology	0.259***	0.248***	0.247***	0.251***
	(0.0403)	(0.0398)	(0.0405)	(0.0385)
Geography	0.288***	0.273***	0.295***	0.285***
	(0.0402)	(0.0409)	(0.0359)	(0.0348)
Religion	0.00763		-0.0231	
	(0.0511)		(0.0549)	
Social studies	0.162***		0.122*	
	(0.0467)		(0.0520)	
Math, female student	-0.422***	-0.421***	-0.447***	-0.439***
	(0.0271)	(0.0289)	(0.0244)	(0.0254)
English, female student	-0.297***	-0.508***	-0.306***	-0.455***
	(0.0261)	(0.0687)	(0.0251)	(0.0559)
History, female student	-0.373***		-0.310***	
	(0.0575)		(0.0475)	
Science, female student	-0.266***		-0.262***	
	(0.0347)		(0.0334)	
Biology, female student	-0.444***	-0.444***	-0.413***	-0.420***
	(0.0411)	(0.0429)	(0.0370)	(0.0397)

Geography, female student	-0.575***	-0.574***	-0.587***	-0.597***
	(0.0416)	(0.0419)	(0.0362)	(0.0355)
Religion, female student	-0.0473		-0.0689	
	(0.0714)		(0.0651)	
Social st., female student	-0.350***		-0.353***	
	(0.0596)		(0.0545)	
Math, immigrant	0.0305	0.0864	0.00699	0.0588
	(0.0484)	(0.0502)	(0.0464)	(0.0511)
English, immigrant	0.277***	0.0323	0.268***	0.174
	(0.0520)	(0.125)	(0.0467)	(0.129)
History, immigrant	0.286*		0.259*	
	(0.141)		(0.115)	
Science, immigrant	0.138*		0.108	
	(0.0571)		(0.0564)	
Biology, immigrant	-0.0536	0.00366	-0.120	-0.103
	(0.0714)	(0.0742)	(0.0688)	(0.0708)
Geography, immigrant	-0.142	-0.0822	-0.126	-0.0554
	(0.0866)	(0.0879)	(0.0771)	(0.0792)
Religion, immigrant	0.361***		0.276**	
	(0.0871)		(0.0883)	
Social studies, immigrant	0.0424		0.0348	
	(0.121)		(0.104)	
Female student	0.388***	0.390***		
	(0.0309)	(0.0342)		
Immigrant student	-0.355***	-0.404***		
	(0.0715)	(0.0760)		
Child of immigrant	-0.328***	-0.363***		
	(0.0491)	(0.0534)		
Mother not in register	0.380**	0.397**		
	(0.145)	(0.150)		
Father not in register	0.301*	0.248		
	(0.134)	(0.141)		

Intact family	0.0922***	0.0968***
	(0.0203)	(0.0226)
One child in family	-0.0809***	-0.102***
	(0.0211)	(0.0237)
Three children in family	-0.0825**	-0.0915**
	(0.0253)	(0.0291)
Four or more children	-0.0601	-0.0507
	(0.0541)	(0.0594)
Mother vocational education	0.125***	0.115**
	(0.0316)	(0.0369)
Mother further education	0.322***	0.302***
	(0.0337)	(0.0383)
Mother higher education	0.480***	0.493***
	(0.0444)	(0.0503)
Mother's education missing	-0.0388	-0.0958
	(0.0988)	(0.0998)
Father vocational education	0.139***	0.141***
	(0.0313)	(0.0332)
Father further education	0.272***	0.250***
	(0.0325)	(0.0362)
Father higher education	0.356***	0.335***
	(0.0425)	(0.0462)
Father's education missing	0.118	0.104
	(0.0849)	(0.0834)
Log income mother	0.0107	0.0139
	(0.0142)	(0.0159)
Log income father	0.0313*	0.0311*
	(0.0153)	(0.0158)
Mother self employed	0.0818	0.0553
	(0.0525)	(0.0566)
Mother higher occupation	0.191***	0.165***
	(0.0341)	(0.0365)

Mother intermediate occupation	0.0668*	0.0784*		
	(0.0274)	(0.0311)		
Mother other occupation	-0.0103	-0.0452		
	(0.0355)	(0.0390)		
Mother unemployed/not in labour force	-0.0878*	-0.0935*		
	(0.0396)	(0.0443)		
Father self employed	0.0538	0.0529		
	(0.0362)	(0.0418)		
Father higher occupation	0.230***	0.235***		
	(0.0345)	(0.0384)		
Father intermediate occupation	0.174***	0.164***		
	(0.0299)	(0.0343)		
Father other occupation	0.0333	0.0337		
	(0.0327)	(0.0383)		
Father unemployed/not in labour force	-0.00244	-0.0265		
	(0.0448)	(0.0471)		
Constant	-0.929***	-0.900***	0.00941	0.00502
	(0.134)	(0.144)	(0.0249)	(0.0292)
Observations	24360	13660	24360	13660
Adjusted R^2 (OLS) / R^2 (within student)	0.164	0.182	0.034	0.066

Reference categories are: Teacher experience at least 10 years; the subject is Danish; teacher education standard; mother (father) has no education beyond compulsory school (9th grade); mother (father) has basic occupation.

Standard errors in parentheses – robust standard errors clustered on classes.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Measurement statistics for the PSM dimensions. Structural Equation model (clustered by schools)

Dimensions and items (English translation in italics)	SFL	SMC (R^2)
Self-sacrifice		
Det er vigtigere for mig at gøre en forskel i forhold til samfundet end at opnå personlig vinding. *Making a difference in society means more to me than personal achievements*	0.584	0.341
Jeg mener, at man skal bidrage med mere til samfundet, end man modtager. *I feel people should give back to society more than they get from it*	0.546	0.298
Jeg er villig til at risikere at skulle tilsidesætte mine personlige behov for samfundets skyld. *I am willing to risk personal loss to help society*	0.858	0.736
Jeg er klar til at lide afsavn for samfundets skyld. *I am prepared to make sacrifices for the good of society*	0.834	0.695
Jeg sætter samfundsmæssige forpligtigelser over hensynet til mig selv. *I believe in putting duty before self*	0.785	0.616
Compassion		
Jeg bliver følelsesmæssigt berørt, når jeg ser mennesker i nød. *It is difficult for me to contain my feelings when I see people in distress.*	0.724	0.524
For mig er hensyntagen til andres velfærd meget vigtig. *For me, considering the welfare of others is one of the most important values*	0.709	0.502
Jeg bliver meget berørt, når jeg ser andre mennesker blive behandlet uretfærdigt. *I get very upset when I see other people being treated unfairly*	0.717	0.514
Jeg føler sympati overfor mindre privilegerede mennesker med problemer. *I feel sympathetic to the plight of the underprivileged*	0.674	0.454
Attraction to public policy making		
Jeg forbinder generelt politik med noget positivt. *I generally associate politics with something positive*	0.503	0.253
Jeg bryder mig ikke om politiske studehandler. *The give and take of public policy making doesn't appeal to me (reversed)*	0.381	0.146
Jeg har ikke særligt høje tanker om politikere. *I do not care much for politicians (Reversed)*	0.841	0.707
Commitment to the public interest		
Det er vigtigt for mig, at offentlige ydelser gavner samfundet som helhed. *It is important for me that public services contribute to the common good*	0.434	0.189
Jeg så helst, at offentligt ansatte gør det, der er bedst for hele samfundet, selvom det skulle gå ud over mine egne interesser. *I would prefer seeing public officials do what is best for the whole community even if it harmed my interests*	0.567	0.321
Det er vigtigt for mig at bidrage til det fælles bedste. *It is important for me to contribute to the common good*	0.734	0.539
Det er min borgerpligt at gøre noget, der tjener samfundets bedste. *I consider public service my civic duty.*	0.752	0.566

Note: SFL=standardized factor loading; SMC=squared multiple correlation coefficients. All standardized factor loadings and correlations are significant at p <.001 (adjusted for the 85 clusters (schools))

Table A.6: Cronbach's alphas and covariance between PSM dimensions

	Cronbach's alpha	2	3	4
1. Self-sacrifice	0.84	0.31	0.12	0.56
2. Compassion	0.80		0.01	0.38
3. Attraction to public policy making	0.57			0.10
4. Commitment to the public interest	0.70			