Essays On The Effectiveness And Production Of Teacher Inputs

Christopher Hein

Department of Quantitative Social Science

UCL Institute of Education

Thesis submitted for the degree of Ph.D.

June 2015

Declaration

I hereby declare that, except where explicit attribution is made, the work presented in this thesis is entirely my own.

Word count (excluding appendices and references): 48,197

Signed: Date:

Acknowledgements

This thesis would not have been if not for a number of people. Most importantly I would like to thank my parents, Prim and Klaus, and all my family for instilling in me the never-ending interest in wanting to know "why" the world is the way it is.

A PhD thesis is a lot of work. At times motivation will drop and it is important not to lose focus completely. I therefore would like to thank again my family and my many friends here in London, back home in Germany and everywhere else around the globe for letting me bore them with my PhD woes and joys. I would also like to thank my research room colleagues for helping me come to grips with my research, be it in our student seminars, in the research room or somewhere in a pub or restaurant in the Bloomsbury area. A special "thank you" goes to my wonderful girlfriend Lu for supporting me with lots of love and good food!

A PhD thesis also requires a lot of thought. Most of this took place while walking around Regent's Park and I am very thankful for its existence and long opening hours, especially during the summer. During, what is nearly a four-year thought process I have been guided by three different supervisors. My thanks therefore go to Dr Rebecca Allen, Professor John Micklewright and Dr Andrew Jenkins.

A PhD also costs money. I would like to thank my parents again for the funding, as well as my late grandfather, Albert Hein, for paying the cost of my tuition fees. I am also very thankful to have been given the opportunity to work part-time as a Study Service Assistant in the UCL Institute of Education Library. Not only did the salary cover my living expenses, but I also learnt a lot of useful things, for example how to maximise the efficiency (yes, this thesis is in Economics of Education after all) of searching literature and doing so in various different ways.

Finally, although this thesis has my name on it, it is yours, too. I hope it makes you proud.

Abstract

This thesis conducts cross-country analyses using data from all inhabited continents to examine the support of common expectations based on either Neo-classical Economics or popular beliefs. The first two chapters use SACMEQ data from sub-Saharan Africa.

The first chapter argues that changes in class size trigger a number of mechanisms affecting how the pupils' household, school leaders, teachers and peers behave. These behaviours are highly context-specific and may counterbalance or exacerbate one another. It finds that the main threat to a pupil's achievement is sharing the teacher with more peers, but that household behaviours can mitigate or even outweigh this threat.

The second chapter examines the conditional correlation of observable teacher characteristics and pupil achievement. It argues and demonstrates that previous research using the same data does not sufficiently address the teacher-pupil matching problem and that lacking to do so leads to very different conclusions. The chapter categorises the available observable teacher characteristics as proxies for either subject-matter or pedagogic competency and examines their complementarity by adding interactions between the individual proxies of these two competencies. The evidence suggests these two competencies are substitutes in six of ten countries.

The third chapter uses OECD TALIS 2013 data to explore the connection between teachers' workload and their job satisfaction. It applies a production function approach that combines both Top-down and Bottom-up approaches. It finds that the effect of teachers' workload measured in hours is negligible. But evidence of the effect of teachers' perceptions of their workplace from the English sub-sample provides clear evidence that the workplace matters.

Table of Contents

List of Figures

List of Tables

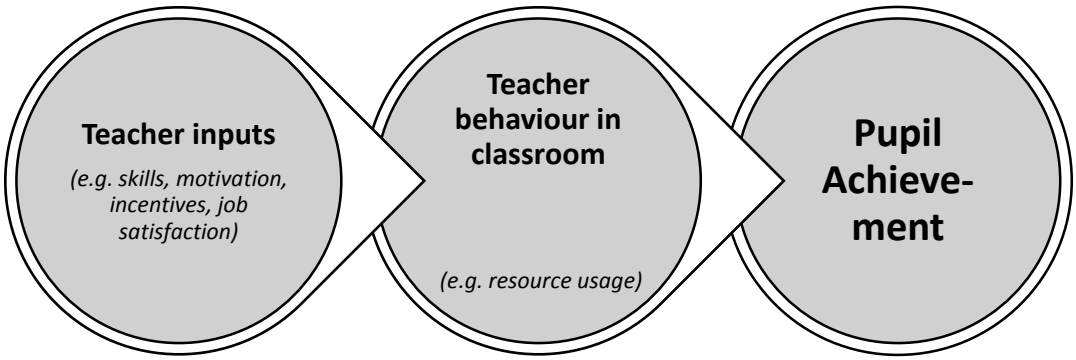
Introduction

As the title of this thesis suggests, this thesis focuses on teachers. Millions of teachers around the world are preparing their country's next generations for adult life and to compete in an increasingly global labour market. At least since Hanushek's seminal "Failure of input-based schooling policies" (2003) in which he argues that school policies should favour investment in teachers, it is generally accepted among education economists and policy makers that teachers are the most important among the school resources. More recent research from the US and the UK further substantiates his claim and shows that the total effect of a teacher on a pupil's achievement test score, i.e. a teacher's total *effectiveness*, ranges between 8 and approximately 20 percent of a standard deviation (cf. Nye et al, 2004; Rockoff, 2004; Hanushek et al, 2005; Rivkin et al, 2005; Kane et al, 2006; Aaronson et al, 2007; Clotfelter et al, 2007; Slater et al, 2011).

In this thesis I apply a theoretical framework in which a teacher's effect on a pupil's achievement depends on the teacher's *inputs*, for example the teacher's skills, job satisfaction, intrinsic motivation, as well as incentives to extrinsically motivate the teacher. Thus, this definition extends typical definitions of teacher inputs that tend to be restricted to teachers' observable characteristics. I assume that these inputs cause the heterogeneity of behaviour patterns that pupils face, which ultimately affects their achievement (see Figure 1).

This theoretical framework is inspired by previous research into teacher effectiveness by economists as well as psychologists. Economists tend to explore teacher effectiveness focusing on teachers' observable characteristics, such as their teaching experience, kind and duration of teacher training, academic qualifications or gender and their conditional correlation with pupil achievement test scores. The field of educational psychology has a longstanding tradition in researching individuals' motivation. Various definitions of motivation exist, but one of the most well-known is by Deci and Ryan (1985), according to which motivation can be intrinsic or extrinsic. Applied to teaching, teachers may be intrinsically motivated if they teach simply for the joy of teaching. In contrast, examples for extrinsic motivation to teach are a teacher's salary and public recognition. This definition of extrinsic motivation is very closely related to the economic notion of an incentive, which entails extrinsic motivators that can be manipulated by policy makers.

Figure I-1: Schema of theoretical framework



From the field of occupational psychology, this framework lends the notion of job satisfaction as an important teacher input. According to Kahneman and Krueger (2006) job satisfaction measures the quality of an individual's perceived experience at work. Evidence from across the social sciences demonstrates that job satisfaction is important for the efficiency of organisations. For example, job satisfaction is negatively associated with burnout (cf. Prosser et al, 1997; Kalliath and Morris, 2002; Piko, 2006). To psychologists burnout is a symptom characterised by individuals feeling overwhelmingly exhausted, detached from their job, cynical and ineffective (Maslach et al, 2001). Burnt out individuals are therefore likely to be less productive in their jobs, for example due to extended periods of absence, and run the risk of incurring social costs in the form of health care expenses (Faragher et al, 2005). Fischer and Sousa-Poza (2009) find that job satisfaction predicts both subjective and objective measures of health and others find that job satisfaction predicts both workers' intentions to leave their employer and actual turnover rates (for example see Hellman, 1997; Lambert et al, 2001; Kalliath and Morris, 2002; Sousa-Poza and Sousa-Poza, 2007). Thus it is important to examine how changes in a teacher's working environment affect variation in his or her job satisfaction in order to maximise teacher's productivity and thereby the efficiency of the education system.

The third chapter of this thesis therefore investigates the production of teachers' job satisfaction and focuses on the conditional correlation of teachers' weekly workload and their job satisfaction. The first two chapters use this theoretical framework to examine (i) the effect of teacher behaviour in the classroom on pupil achievement, specifically teachers' usage of classroom resources as well as of themselves as a teaching resource (ii) the effect of teachers' subject-matter and pedagogic skills on pupil achievement.

In this thesis I use data from a total of 46 education systems covering all inhabited continents. The first two chapters use data from the Southern and Eastern African Consortium for the Monitoring of Educational Quality (SACMEQ) collected in the year 2000 (SACMEQ II) and 2007 (SACMEQ III). As the name suggest, SACMEQ data focuses on a large region within Africa and contains more countries from the African continent than other, more prominent cross-country datasets such as the PISA and TIMSS datasets. SACMEQ data is also much less frequently used than the latter. In total, SACMEQ II and III consists of 15 education systems, Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania, Uganda, Zambia, Zanzibar and Zimbabwe[1].

In between these two time periods the education systems of Kenya, Tanzania, Zambia and Zanzibar especially underwent large expansions, as substantial proportions of these countries' school-aged population were not attending school. In other words, these countries lagged far behind reaching their Millennium Development Goal of universal primary education in the year 2000 (cf. UNESCO Global Monitoring Report, 2011). Other countries, such as Botswana, Mozambique and Swaziland, also needed to expand their education systems, but rather than suddenly abolishing all direct school fees and thereby encouraging a demand-shock of pupils wanting to attend schools, these latter countries gradually expanded access to their education systems. Some SACMEQ countries, such as Namibia, the Seychelles and South Africa had already reached, or were very close to universal primary education by 2000.

In each wave SACMEQ samples multiple grade 6 pupils per school, whereby the schools are registered government or non-government schools. SACMEQ surveys consist of different parts: Head teachers are surveyed on their demographics and

[1]In this thesis I use data from all except Zimbabwe as this data has the reputation of being unreliable due to the political system.

those of the school, as well as their own teaching. Similarly teachers are surveyed about their demographics and teaching. Finally the pupils are asked to fill in a questionnaire about themselves and their family. Pupils are tested in maths and reading in wave two, and in a third subject 'health/science' in wave three. Teachers are also tested in the subject they teach. For example if a teacher teaches maths, reading and 'health/science' in wave three, he or she will be tested in these three subjects.

The third chapter uses data from the OECD Teaching and Learning International Survey (TALIS) collected in 2013, which is the second round after the first in 2008. In contrast to SACMEQ, TALIS data samples multiple teachers per school in 33[2] high and middle-income countries. TALIS 2013 surveys both head teachers and teachers about the teaching and learning process. In doing so it collects data on characteristics of the school, as well as demographics, attitudes, beliefs and behaviours of the teachers and the heads. I use TALIS data as it is the first cross-country dataset of teachers to collect data on teachers' job satisfaction and it is therefore ideal for my research interest.

The three chapters also share the same methodological approach. As both SACMEQ and TALIS data are observational, and although each data has a nested structure of either pupils or teachers in schools, I cannot obtain causal estimates to answer my research questions. I therefore employ a cross-country comparative approach. This should not be mistaken with a multiple case study approach. According to Yin (2009) a case study is "[a]n empirical inquiry about a contemporary phenomenon (e.g. a "case"), set within its real-world context […]" (Yin, 2009, p.18), which focus on the understanding of the cases by examining the context and other complex conditions related to them (Yin, 2009). Thus a multiple case study approach here would require an in-depth understanding of each of the educations systems sampled in this thesis, which would go not only beyond the scope of this thesis, but also beyond the scope of most case study research.

As the data used here is correlational the main advantage of cross-country comparisons is to improve validity of the study. Instead of applying a model to a sample from one country only, the focus here lies on the cross-country patterns. For example, a consistent correlation of teachers' weekly workload and their job satisfaction across all TALIS countries if observed would add more weight to the

[2] I do not use data from the US, as this sample does not meet the OECD threshold response rate for schools and teachers.

claim that teachers' job satisfaction is in fact affected by their weekly workload. This study does not focus on pooled models, but does include some to contrast findings, because such models average over the entire sample of countries and could mask heterogeneity across countries. In addition, although pooled models often account for unobserved differences between levels, which in this thesis is the country level, this thesis acknowledges the fact that each country has its own social, cultural and economic context that will affect country-level estimates. And reporting these individual estimates for each country is more likely to be of use to policy makers than pooled estimates averaging across a range of countries.

Another common feature of the three chapters is their *challenging common expectations*. In each paper I set out to find empirical support for a traditional economic theory or a common belief. The first chapter for example addresses the debate on class size. In a nutshell, reducing class size implies high costs, as increasing numbers of teachers are needed. On the other hand, smaller class sizes are supposed to allow pupils to learn better. The existing evidence though does not fully support this notion, which suggests that the mechanisms of class size are not fully understood. The first chapter argues that a change in class size will have a compositional effect, changing classroom and school composition as well as peer dynamics, and an effect on how school resources and teachers are used. Moreover these underlying mechanisms relate to human behaviour. Thus the effect these mechanisms have on pupil achievement will depend on the sample's specific (country) context i.e. the preferences of households, teachers and school leaders as well as the incentives they are faced with.

I explore this using data from nine countries from the sub-Saharan Eastern and Southern Africa region that are part of the SACMEQ datasets sampled in the year 2000 and 2007. I find that the effect of an increased amount of pupils sharing school, classroom resources or teachers varies in magnitude and direction both within and across the nine countries. This underlines that the aggregate effect of a change in class size depends heavily on the individual country context. Evidence obtained from pooled estimates shows that when these country contexts are accounted for, sharing the teacher with an increased number of pupils is the main threat to a pupil's achievement. Yet household support behaviours such as providing the child with extra tuition in particular can counterbalance or even outweigh having to share the teacher with more peers.

The second chapter explores why teachers differ in their quality and challenges the assumption that the laws of human interaction are universal, as Neo-Classical Economics might lead one to believe. If this were the case then teachers in developing countries, for example in sub-Saharan Africa should differ in their quality due to the same variables, so that evidence from the US should be applicable for policy guidance in these countries. But teacher labour markets in sub-Saharan southern and eastern Africa are likely to operate differently, due to the various financial and human capital restraints these countries face, leading to a very heterogeneously skilled teacher labour force. Similarly, because of the different incentive structure in these countries, a different kind of teacher may be attracted into the profession, who may behave differently. For example, better trained teachers may be worse teachers in the classroom, because they are trying to progress up the career ladder or aspire to a career in the government, and therefore do not use their skills to the pupils' advantage.

At the same time I argue that previous research in sub-Saharan Africa exploring this issue do not address adequately the non-random matching of pupils to teachers. I therefore follow Clotfelter et al (2006), who argue that allowing for pupil-fixed effects addresses this issue sufficiently. I demonstrate that previous research using the same data does not sufficiently address the teacher-pupil matching problem and that failing to do so leads to very different conclusions. The evidence also shows that even within schools, pupils and teachers are not matched randomly. Finally, the pattern of the findings of my pupil-fixed effects models resembles the pattern emerging from the US, in that there are no consistent predictors across countries. Yet each country has its own combination of statistically significant predictors, so that US evidence is not a suitable guide for policy in these countries.

Especially in England it is commonly perceived, as a quick Google search of the keywords "teacher workload" will show, that teachers work too many hours[3], that most of the administrative tasks are unnecessary[4], and that the high workloads are

[3] http://www.telegraph.co.uk/education/educationopinion/11243368/Teacher-workload-at-unacceptable-levels.html ,[Last accessed: 24/05/15]

[4] https://www.gov.uk/government/news/government-pledges-to-reduce-teacher-workload ,[Last accessed: 24/05/15]

a key factor for newly qualified teachers to quit the teaching profession[5,6]. One would therefore expect that a teacher's workload would be strongly correlated with his or her job satisfaction. In the third chapter I explore this specific aspect using the OECD TALIS data from 32 countries. I embed this analysis in a production function that unifies the two prevailing streams of literature, namely Top-Down and Bottom-Up approaches (cf. Diener, 1984). Thus it combines the notion that a teacher's job satisfaction might be determined by the environment he or she is in (Bottom-Up) with the notion that individuals process their environment differently (Top-Down). Based on this production function I identify potential groups of stakeholders that could both directly affect teachers' job satisfaction as well as their workload and thereby bias endogenously bivariate OLS estimates of the workload – job satisfaction relationship. These groups of stakeholders are the teachers themselves, the pupils, the parents, the other teachers in the school, the head teacher and the physical resources available in the school. In addition, I consider the association of not only total weekly workload, but also explore heterogeneity of effects for five different activities teachers do while working; face-to-face teaching hours, time spent planning, marking, doing administrative tasks as well as the time they are engaged in school leadership activities.

I find that for the majority of the 32 education systems estimates are not statistically significant and those that are, are of negligible magnitude. In other words there is no convincing evidence that teachers' job satisfaction varies as a result of the amount of hours they work. Yet, I do find that certain aspects of the teachers' working environment explain substantial variation in their job satisfaction. Using data on teachers' perceptions of their workplace available in the English sub-sample, I find that teachers value well behaved pupils, sufficient autonomy to do their job and scope to progress as a teacher. An effective school leadership team is similarly important to them, but supportive parents, school leaders providing clear vision and direction, scope to progress to higher pay and the perception of receiving fair pay given their performance are also statistically significant, but increasingly less important. The obtained estimates range from 0.09 to 0.39 SD.

[5] https://www.the-newshub.com/uk-politics/i-quit-the-ever-increasing-workload-of-the-british-teacher ,[Last accessed: 24/05/15]

[6] http://www.dailymail.co.uk/news/article-3020580/More-40-new-teachers-leave-profession-12-months-Excessive-workload-blamed-number-quit-triples-six-years.html ,[Last accessed: 24/05/15]

Before moving on to the first chapter, I would like to make a remark regarding the structure of this thesis. Each of the three chapters themselves was conducted as a small research project with the intention to be published. Thus, in the chapters I refer to them as "papers". The second chapter of this thesis has already been released as a working paper in the IoE Department of Quantitative Social Science working paper series in a previous form, but has since undergone substantial changes before entering this thesis and does not include the contribution of my previous supervisor Rebecca Allen and is solely my own work. Also, as the chapters are to be published, and as they do not all share the same underlying dataset, I do not introduce the reader to the data in a separate chapter. Instead I introduce the reader to the relevant aspects of the data within each chapter.

Chapter 1

Exploring mechanisms of class size effects on achievement in Sub-Saharan Africa: the role of teachers, physical resources and households

1.1. Introduction

Reducing class size is a common item on the agenda for policy makers to debate in developed countries. In a nutshell, reducing class size implies high costs, as increasing numbers of teachers are needed. On the other hand, smaller class sizes are supposed to allow pupils to learn better. Evidence on the effect of class size on achievement is predominantly from developed economies. The most robust findings are either from the Tennessee STAR randomised experiment in the 1980s in the US or from adaptations of Angrist and Lavy's (1999) paper applying Maimonides' rule in Israel. Angrist and Lavy (1999) find significant effects in favour of smaller classes, but not for every examined cohort. Adaptations of this approach in Poland and France also find such effects, but of much smaller, nearly negligible magnitude (cf. Jakubowski and Sakowski, 2006; Gary-Bobo and Mahjoub, 2006). In contrast Asadullah (2005) finds a positive effect in favour of larger class size in Bangladesh. Evidence from Tennessee STAR again suggests results in favour of smaller class sizes and long-term effects seem to exist until entry into higher education (eg. Nye et al., 1999, 2000 and 2001). Hanushek (2003) though notes that the advantage of continuously being in a smaller class does not increase over time. Thus, even in the light of this evidence, the underlying mechanisms of the effects of class size are not fully understood.

This paper argues that a change in class size will have a compositional effect and an effect on how school resources and teachers are used. The compositional effect can be divided into two parts: First it is likely to change existing peer dynamics as literally individuals will either be removed or added to a class depending on the direction of change in class size. Second, if households care about the quality of their children's education, they will react to changes in class size and reconsider the kind of school they send their children to, how much academic support such as extra tuition and help with homework to provide their children with, and the amount of education to invest in their children.

A change in class size also implies that a teacher or any other resource in a school needs to be shared by a different number of pupils. How these resources are shared depends on head teacher and teacher behaviour. For example in the classroom, the teaching methods a teacher applies decide how the available resources in the classroom are shared among the pupils.

Thus a change in class size triggers various mechanisms, which on aggregate may exacerbate or counteract each other. Heterogeneity of findings across countries is to be expected here as the individual country contexts will affect household preferences, i.e. the extent to which households support their children, as well as the kind of teachers and school leaders who select into the teaching profession and the kind and quality of resources and teachers available in individual schools. This paper therefore reports findings both for individual countries and contrasts a pooled model that accounts for unobserved heterogeneity across countries.

This paper uses a rich secondary dataset of pupils matched to teachers and schools, in nine of the fifteen member countries of the Southern and Eastern African Consortium for the Monitoring of Educational Quality (SACMEQ), which was collected in the years 2000 and 2007. The paper estimates the association of usage of school, classroom resources and teachers as well as of receiving extra tuition and help with homework, using multivariate linear regressions. Estimates are reported separately for each country and are contrasted with a pooled model that accounts for unobserved between-country differences. The findings indicate that the effect of an increased amount of pupils sharing school, classroom resources or teachers varies in magnitude and direction both within and across the nine countries and underlines that a class size aggregate effect of a change in class size depends heavily on the individual country context, i.e. the preferences of households, teachers and school leaders as well as the incentives they are faced with. The pooled estimates show that when these country contexts are accounted for, sharing the teacher with an increased number of pupils is the main threat to a pupil's achievement. Yet household behaviours such as providing the child with extra tuition in particular can counterbalance or even outweigh having to share the teacher with more peers.

The paper is structured as follows: The next section provides the reader with the relevant theory and literature before Section 1.3 introduces the reader to the SACMEQ data and the context of the sampled countries. Section 1.4 discusses the multivariate regressions applied to the data, before reporting the obtained estimates in section 1.5, Section 1.6 concludes.

1.2. Theory and literature

Traditionally human capital accumulation is formalised in an education production function (cf. Hanushek, 1979), whereby educational achievement A of individual i in class j and school k is a function of the individual P, often referred to as innate

ability, his or her family background H, schooling S, which in turn is a function of physical school resources R, teachers T, school leadership L and the respective peer group G.

$$A_{ijk} = f(P_i, H_i, S_k(R_{jk}, T_{jk}, L_k, G_{jk})) \quad (1)$$

The following will show how a demand shock for education can affect each of the constituent inputs of the education production function, beginning with the household level. The discussion will include a number of factors that may affect the class-size mechanisms as well as pupil achievement, i.e. potential sources of endogeneity, which I will return to when discussing the modelling strategy in section 1.4. As this research uses data from sub-Saharan Africa, the following discussion will use examples from the African context for illustration.

1.2.1 The household

An influx of pupils into the school a pupil is attending will present the child's household with a number of decisions:

First, households may be prompted to differentiate schooling investments for their children, either through amount of schooling or through attendance at schools of differing quality. As Gandhi Kingdon (2002) notes, parents may have differing preferences for sons compared to daughters. In the Indian context for example it is usual that a daughter's parents-in-law reap the majority of the profit of her productivity, thereby providing a disincentive for parents to invest in daughters' education. Dickerson et al (2013) find that in the African countries focused on in this paper, girls tend to be enrolled in schools where pupils perform better in maths. Further they find that the gap in achievement between girls and boys is much larger in regions "where the role of the women is confined to the home", i.e. in areas with high fertility, low levels of education among women and Islam is more prevalent (cf. Davis-Kean, 2005 for socio-economic status and race in the US).

Also a perceived deterioration of educational quality will increase opportunity costs of education and encourage pupil absenteeism and drop out in order to participate either in the labour market or in home production. Especially as a child's age increases, the more suitable it becomes to work in the labour market, which will increase opportunity costs further and provide a disincentive to participate in education. These opportunity costs may also be exacerbated if the labour market depends on school-aged children prior to the shock. In this case the abolition of school fees may function as a negative supply shock to the labour market

increasing wages. Governments can, of course, make this choice obsolete by enforcing anti child-labour legislation.

Parents may also decide to move their children into private school. Watkins (2004) reports that private schools founded without government intention are known to cater for the poor who are willing to pay, because these schools signal better transparency and accountability to these groups. Oketch et al (2010b) explore the popularity of such schools in Nairobi, Kenya and conclude that contrary to Watkins' argument of household preferences, the driving factor is a lack of supply of regular public schools. In general school choice will also affect the peer dynamics in schools, as instead of the number of pupils, this behaviour changes the socio-cultural composition of a school.

Households may also be prompted to adapt the academic support they provide to their children, such as helping their children with their homework or paying for extra tuition. Evidence from Norway and the US suggests a connection between class size and household support. Bonesronning (2004) finds in Norway that households tend to decrease their effort when class size increases from low levels. Datar and Mason (2008) find in a sample of US kindergarten pupils that an increase in class size is associated with an increase in parent-financed activities and a reduction in parent-child interaction, but there is no significant association with parent-school interaction. Again, there is no corresponding evidence from sub-Saharan Africa.

1.2.2 The role of the school leader and teacher

Within schools, technical effectiveness of available physical resources (such as textbooks, chairs, tables, blackboards, class libraries, and so on) and teachers arises from behaviour patterns of school leaders and teachers. School leaders make strategic decisions how to equip classrooms with resources and allocate teachers and classes to them. In other words, the effect of a resource on a pupil's achievement is a function of how the resource is allocated to the pupil.

In the sub-Saharan context multi-grade teaching and teaching in shifts are common practice in some countries (cf. Mulkeen, 2009). Multi-grade classrooms contain pupils of more than one school grade and thus increase class size compared to single grade classes. Little (2004) notes that multi-grade teaching often occurs out of necessity rather than as a pedagogic choice, for example when demand for education outstrips supply in order to grant pupils access to education.

Governments rarely acknowledge its existence, so that curricula are not adapted accordingly and teachers are expected to cover these and meet exam expectations as if the classroom only consisted of the one grade (cf. Little, 2004). Evidence is mostly from developed countries such as the US and suggests no cognitive disadvantages for pupils educated in multi-grade classrooms. Evidence from developing country contexts is sparse. Jarousse and Mingat (1991) find that multi-grade classes outperform single grade classes in both Burkina-Faso and Togo, but the indicative quality for the southern and eastern African context more than a decade later must be questioned.

In contrast, teaching in shifts provides a means to reduce class size. Here schools allocate specific time slots to proportions of the school's total enrolment. Although this approach is popular among policy makers for reducing unit costs of resources, it is criticised for implying a reduction in the amount of time a pupil is taught. Mulkeen (2010) notes that in Zambia teaching in shifts means a school day may only be 2.5 hours long. Also pupils and teachers allocated to afternoon shifts may be less productive because they are tired and are attending school in the heat. Corresponding evidence though does not show a clear negative association of teaching in shifts and achievement. Again, robust evidence is sparse, but while Lee and Zuze (2011) find a negative association in sub-Saharan southern and eastern African countries and Michaelowa (2001) in francophone Africa, Bray (2008) reports no significant differences. Bray suggests that although pupil-teacher contact time may be reduced per day, it is possible to impose more days of school per week, such as school on Saturdays. Also, international comparisons show that regular contact time varies substantially across countries, suggesting that the crucial variable is how a given amount of time is used rather than the total amount of time per se (cf. Bray, 2008).

In the classroom a teacher's strategic use of resources arises from the teaching methods he or she applies while creating learning environments. For example, if there are not enough textbooks for all pupils due to an increase in class size, teachers may opt to 'preach to the pupils from the front of the class', also known as ex-cathedra teaching, or use group work methods (cf. Cuseo, 2007 for the case of higher education in the US). Similarly, the effect of a blackboard in the classroom on the pupils' achievement depends on how the teacher uses the blackboard, i.e. what he or she writes on the board or whether he or she uses it at all.

A further key skill teachers require to design effective learning environments is being able to identify the learning needs of each individual pupil. Yet addressing these fully may require a variety of teaching methods, which may not be possible given school and class size or available resources.

These two key skills, knowledge of and being able to implement teaching methods and identifying pupils' learning needs can arguably be acquired through training. Sub-Saharan African countries have a history of an undersupply of trained teachers even before signing the MDGs, although there was stark variation between countries. For example only 29 percent of all teachers were trained in Namibia in 1999, whereas 98 of all teachers were trained in Kenya (UN Institute for Statistics, 2010 cited in UNESCO EFA Global Monitoring Report 2011 Statistical Annex[7]). In the light of increased demand, countries may hire less well trained or untrained teachers on temporary contracts whose demography can vary from parents, over village elders to trainee teachers, whose pedagogic skills, particularly those concerning effective teaching methods, will vary (cf. Fyfe, 2007). Vegas and de Laat (2003) find that in Togo pupils of regular teachers outperform those of contractual teachers, despite contractual teachers having the same amount of years of education. In contrast in Niger, Bourdon et al (2005) find no significant difference of teachers' contractual status on pupil achievement. Hein and Allen (2013) also find no clear link between teachers' observable characteristics and teacher quality in twelve SACMEQ countries using data from the year 2000.

In addition, teachers' motivation and attitudes towards their pupils are important (cf. Michaelowa, 2002; Bennell, 2004). If, for example, there is a societal expectation that women perform particular social roles such as housekeeping, childbearing and child rearing, this may influence the teacher's behaviours towards female pupils (cf. Kazeem et al, 2010). Evidence by Dickerson et al (2013) based on data from 19 countries in Western, Eastern and Southern Africa in cross-sectional data from 2000 and 2002 is consistent with no discrimination of female pupils. But with a sudden increase in school-goers, teachers' motivation and attitudes towards their pupils may change.

7 Link has expired: file name "Statistical tables-2011-Longer version-Final-Website.xls", downloaded on the 15th of November 2012, is available on request from the author.

1.2.3 The peer group

Changes in class size are also likely to change the composition of a pupil's respective peer group and thus affect the existing peer dynamics as literally pupils are either added or removed from the class depending on the direction of the change in class size. Evidence on peer effects is again mainly from developed countries and some evidence suggests that strategically grouping pupils according to gender (e.g. Hanushek et al., 2003; Lavy and Schlosser, 2011) or ability (Lavy et al., 2011; Duflo et al., 2011) may be beneficial for pupil achievement. There is no evidence on the size and direction of corresponding peer effects from the countries chosen here, nor on the prevalence of schools strategically assigning pupils to classes.

Also, from a teacher's perspective a change in peer dynamics may complicate being able to identify pupils' needs. Also teachers need to successfully channel a change in peer dynamics through their classroom management skills to maximise the pupils' learning.

1.2.4 Geographical variation of household, school leader and teacher behaviour

In sub-Saharan Africa research has shown that the more remotely a school is located, the more likely it will struggle to attract well trained staff and will be less well equipped in respect to physical school resources such as schoolbooks, classrooms, etc (cf. Mulkeen, 2010). Teachers in sub-Saharan Africa are also known to lack adequate incentives to teach in increasingly remote areas (cf. Mulkeen, 2010). Apart from lower motivation, teacher absenteeism rates, especially in rural areas, are also attributed to teacher's health conditions for example due to HIV, local weather conditions such as floods, and administrative factors such as needing to travel some distance to collect their salary (cf. Das et al, 2005; Bennell, 2005; Chaudhury, 2006). Standards of living also vary substantially by geographical location (Sahn and Stifel, 2003). In rural areas, farming may require as many hands as possible to manage the workload, thus functioning as a disincentive to attend or do well at school. Also, indirect school costs may be higher in rural areas in respect to commuting expenses, either monetary or non-monetary in the form of time spent.

Another dimension of unequal distribution of resources is between public and private schools. Especially high fee-paying private schools may have access to more and better quality resources and teachers than public schools. Further there is cooperation of schools and aid organisations that may reduce transaction costs for

access to number and quality of resources, teachers and leadership (cf. Rose, 2009; Ulleberg, 2009).

1.3. The SACMEQ data

1.3.1 Context

The data used are from the second and third wave of SACMEQ, the Southern and Eastern Africa Consortium for the Monitoring of Educational Quality. The second wave was collected in the year 2000, the third in 2007. There are fifteen member countries taking part in total, of which nine are used here. Among these countries, some, such as the Seychelles, Mauritius, Malawi, South Africa and Namibia had already reached, or were close to, enrolling all primary school aged children. In other words, these countries were close to meeting the second of the UN Millennium Development Goals. Others were far off meeting this goal as they had school fee policies in place preventing large proportions of the school-aged population from attending school. But whereas Botswana and Mozambique decided to gradually abolish these fees to enable increasing amounts of pupils to attend school, other SACMEQ countries such as Kenya in 2003, mainland Tanzania and Tanzania-Zanzibar (from now on Zanzibar) in 2001 and Zambia in 2002 opted to abruptly abolish all direct school costs. This sudden abolition of fees led to a substantial increase in primary school enrolment, threatening to increase school and class size against a backdrop in supply of schools, physical resources and teachers, with potentially disastrous effects on pupil achievement. For the remainder of this paper, Kenya, Tanzania, Tanzania-Zanzibar and Zambia will be referred to as group 1.

The rise in enrolment is shown in Table 1.3.1, where between 1999 and 2008, adjusted net enrolment for primary school[8] increases by 19 percentage points in Kenya. In Zambia the corresponding increase is 26 percentage points and the overall average increase for mainland Tanzania and Zanzibar is reported to be 39 percentage points (UNESCO Institute for Statistics, 2014). It must be added at this point that in contrast to Kenya and Zambia, the Tanzanian government decided to stagger the increased demand between 2002 and 2006 so that 100 percent primary school net enrolment was to occur in 2006. The Tanzanian government decided to do so by granting access to primary schools to certain additionally eligible birth cohorts (Government of the United Republic of Tanzania, 2001).

8 The proportion of children of the official primary school age enrolled in primary or secondary education.

Table 1.3.1: Adjusted Net Enrolment Rate for Primary School

		1999	2008	Difference
Policy change	**Kenya**	63	82	19
	Tanzania	49	98	39
	Zambia	71	97	26
	Unweighted Average	**61**	**92**	**28**
No policy change	**Mauritius**	93	97	4
	Malawi	99	97	-2
	Namibia	88	87	-1
	Seychelles	93[1]	94	1
	South Africa	97	96	-1
	Unweighted Average	**94**	**94**	**0**

Source: UNESCO Institute for Statistics http://data.uis.unesco.org/ last accessed on the 10th of April 2014; 1 data from year 2000

The second half of Table 1.1.1 contains five countries that did not receive such a policy change and are Mauritius, Malawi, Namibia, the Seychelles and South Africa, which from now on will be referred to as group 2. These countries already had achieved very high levels of adjusted net enrolment for primary school, which changed only slightly during the same period (Table 1.3.1). Of these countries South Africa and the Seychelles have a long history of compulsory, free primary education (cf. Government of Seychelles Ministry of Education, 2001; Government of the Republic of South Africa, 2005). Namibia and Mauritius implemented free primary education policies in the early 1990s and Malawi did so in 1994 (cf. Government of the Republic of Mauritius, 2001; Government of the Republic of Namibia, 2002).

Apart from differing by experiencing the demand shock to their primary school systems, World Bank data of these nine countries indicates (see appendix for details) that the group 2 countries are typically much wealthier, but in general grow slower economically and in terms of population, and have higher HIV prevalence rates than the group 1 countries. Group 2 is also more heterogeneous, and includes small holiday island nations as well as large territorial states with both small and large populations. Malawi appears to be an outlier. Although classified here as a group 2 country, its growth in population is similar to the group 1 countries. Yet it has the smallest growth in GDP per capita and by far the lowest rate of persistence to grade 6 (see appendix).

1.3.2 The data

SACMEQ data is heavily used by member governments as an additional data source to complement their administrative data and they welcome the data collection process as a means of local capacity building.

In each wave SACMEQ samples grade 6 pupils attending *registered* government or non-government schools. Thus the sampled data may not be representative for all schools if, as is the case in Kenya, private schools exist that are not registered with the government. Oketch et al (2010a) highlight that of their sample of low cost private schools in Nairobi, one quarter are not registered. But it is unknown how large this potential bias is for the entire country. The participating governments also have some leeway in the definition of the target population. For example, the Tanzanian samples only contain individuals in government schools. In short, it is hard to ascertain the actual corresponding populations and the reader should keep this in mind.

Pupils are sampled in two stages. In the first stage schools are stratified by region and number of grade 6 pupils, whereby the latter is truncated into two categories, small and large. Schools are then selected with probability proportional to the number of their respective grade 6 pupils. In the second wave, a maximum of 20 pupils are sampled at random in each school in all countries except Namibia and Mauritius where a maximum of 40 pupils are sampled and the Seychelles, where all grade 6 pupils in the selected schools are sampled. In the third wave the maximum sample size for the second stage is increased in all countries apart from the Seychelles where again all pupils in grade 6 are sampled in the chosen schools. This time a maximum of 50 pupils are sampled in Namibia, Mauritius, Kenya and Zanzibar and a maximum of 25 pupils in all other countries. According to SACMEQ's technical documents and country reports the response rates of pupils in the second wave are 89 percent in Kenya, 83 percent in Malawi, 93 percent in Mauritius, 92 percent in Namibia, 96 percent in the Seychelles, 85 percent in South Africa, 77 percent in Tanzania, 75 percent in Zambia and 83 percent in Zanzibar (Onsomu et al, 2005). In the third wave response rates are 91 percent in Kenya, 79 percent in Malawi, 89 percent in Mauritius, 91 percent in South Africa (Moloy and Chetty, 2010; Milner et al, 2011; Sauba and Lutchmiah, 2011; Wasanga et al, 2012). Response rates are not available for the other countries. Table 1.4.1 reports the achieved sample size in each country. In order to ease comparability, individuals and schools in educational districts that are not sampled in both waves

are omitted here. This leads to a total sample size of 55,328 pupils nested in 2,630 schools.

Table 1.3.2: Composition of sample used

	Schools			Pupils		
	SACMEQ II	**SACMEQ III**	**Total**	**SACMEQ II**	**SACMEQ III**	**Total**
Kenya	184	193	377	3,299	4,436	7,735
Tanzania	127	140	267	1,937	2,996	4,933
Zambia	155	84	239	2,410	2,614	5,024
Zanzibar	145	143	288	2,514	2,791	5,305
Malawi	140	139	279	2,333	2,781	5,114
Mauritius	89	72	161	1,685	1,622	3,307
Namibia	270	267	537	5,048	6,398	11,446
Seychelles	24	24	48	1,484	1,480	2,924
South Africa	125	309	434	2,366	7,134	9,500
Total	**1259**	**1371**	**2630**	**23,076**	**32,252**	**55,328**

SACMEQ surveys consist of different parts: both pupils and teachers are tested on similar, but not identical tests in all tested subjects. These are maths and reading in wave two and an additional 'health/science' test in wave three. Head teachers are surveyed on their demographics and those of the school, as well as their own teaching. Similarly teachers are surveyed about their demographics and teaching. Finally the pupils are asked to fill in a questionnaire about themselves and their family. In their technical documents SACMEQ acknowledge that item non-response exists 'occasionally'. In these cases, if less than 15 percent of the respective item is missing, SACMEQ impute the missing values by replacing continuous variables with the mean and categorical variables with the mode of the respective lowest level of aggregation in the data. As will be seen in the following multivariate regression analysis, the loss of observations due to missing data is negligible.

SACMEQ pupil and teacher tests scores are based on multiple-choice items and are estimated using Rasch models. To make SACMEQ test score measures comparable over time, the tests in each subject in each wave contain common items. Rasch models (cf. Rasch, 1960), also known as one-parameter item response models, are commonly used to estimate competency in a specific subject and are also applied in

the OECD PISA studies. Compared to the two and three parameter item response models, Rasch models neither account for guessing nor that items may differ in their quality to discriminate between higher and lower abilities (cf. van der Linden and Hambleton, 1997). Baird et al (2011) criticise the validity of PISA studies because of the way pilot studies are conducted. Pilot studies are intended to identify items in which certain samples differ substantially from expected levels, so that such items can be improved or removed for the main study. As PISA tests are not piloted in every participating country, it cannot be established whether all items used are valid. Kreiner and Christensen (2013) reanalyse PISA reading skills data from 2006 and provide evidence indicating that the Rasch models used by PISA are neither valid within nor across countries and that the produced rankings of countries are not robust. This may also be an issue in SACMEQ, as the extent of piloting is unknown to the author. Further the language of testing may be a problem. SACMEQ tests are conducted in English, but English typically will not be the mother tongue of every sampled pupil within countries. Similarly children from wealthier households or from urban centres may come into contact with English sooner than poorer children. Also the age at which English is introduced as medium of instruction can differ across countries. In the light of these potential flaws in the provided Rasch scores, this research therefore assumes that their validity holds.

For the purpose of the analysis here I generate several variables. The first variable captures the number of physical school resources. This summative measure is based on whether or not the head teacher reports the presence of a school library, a hall, a staff room, a head teacher's office, a store room, a first aid kit, a sports ground, water supply, electricity, a telephone, a fax machine, a school garden, a typewriter, a duplicator, radio and tape recorder.

The second variable captures the *number of classroom resources*. This summative measure is based on 6 items that measure whether or not the classroom has a wall chart, cupboard, bookshelves, classroom library, a desk and chair for the teacher as well as a blackboard and chalk. Although this variable is a teacher-level variable, analysis of variance shows that in all countries except the Seychelles (35 percent), Mauritius (63 percent) and Zanzibar (77 percent), at least 90 percent of the variation lies between schools.

The third variable captures the *number of pupil materials*. This summative measure is obtained from 8 items in the pupil questionnaire that indicate whether

the pupil has an exercise book, a notebook, a pencil, a sharpener, an eraser, a ruler, a ballpoint pen and a file.

The fourth generated variable represents *socioeconomic status* and captures the sum of three aspects: First, the level of parental education. On a scale from 1, representing no education to 6, representing post-secondary education a pupil can achieve a maximum of 12 for both parents. The second aspect covers the quality of walls, floor and roof of the house the pupil lives in. The corresponding items are whether or not the pupil lives in a house with stone or cement walls, carpeted or tiled floors and a tiled roof. The third aspect covers a list of ten household items, which are a newspaper, magazine, cassette player, radio, car, motorcycle, bicycle, water, electricity and a table. Thus, the final SES variable is the sum of these three aspects and can range from 0 to 25. For example, a pupil whose parents have both attended post-secondary education, lives in a house with concrete walls, carpeted floor and tiled roof and owning all ten household items will score at the top of the SES measure. This measure implicitly assumes that owning a radio is equal to one additional level of parental education, but any other reweighting of these aspects is arbitrary in the absence of suitable theory. Further, this measure assumes that prices for the ten household items remain constant. Also cassette players may become out-dated and disposed of in the 2007 wave due to widespread availability and affordability of CD players.

SACMEQ provides its own measure of socioeconomic status, which is estimated using Rasch models. Dolata (2005) describes the complex procedure. The main difference between the latter and the generated measure here are that SACMEQ's measure is estimated pooling all member countries together across all three waves and standardised to the SACMEQ mean in the year 2000. Analogously to the test score measure, this latent SES measure assumes that each item has the same quality to discriminate between higher and lower socioeconomic status, and households with the same level of SES have the same probability of agreement with each item. Thus it ignores that prices for items in the house or building materials may differ across countries and over time, or that the severity of barriers for parents to reach higher levels of education may differ across countries. My manually generated socioeconomic status measure is therefore preferred for both its simplicity and similarity with SACMEQ's analogue: The Pearson correlation coefficient is 0.95 when both waves and all countries are pooled together; the correlation coefficient is ≥ 0.93 within each sampled country.

As discussed in the theory section, school leaders may decide to apply multigrade teaching, teaching in shifts or both. Using official school enrolment as a proxy for school size will be misleading if the school in fact teaches in shifts, because it is of interest here how physical school resources and teachers are allocated to pupils when the sampled grade 6 pupils attend school. I therefore use the maximum number of pupils attending school during the same shifts as the sampled grade 6 pupils as a proxy for school size.

In SACMEQ both teachers and head teachers are asked corresponding questions on grade 6 class size, yet they refer to the official enrolment figures instead of a headcount of pupils in the actual rooms the grade 6 pupils are taught. Thus, SACMEQ data on class size represents grade 6 enrolment divided by the number of grade 6 classes, which does not differ by subject, thus suggesting that class composition remains unchanged for both subjects. Dividing school size per shift by the number of classrooms the schools claim to have, which may be located in temporary, permanent or open-air spaces, gives an idea of the actual class size grade 6 pupils may be in. This measure is preferred to grade 6 enrolment as a proxy of actual class size the pupils are likely to experience.

As a change in class size implies that a teacher or any other resource in a school needs to be shared by a different number of pupils, the variables of interest are ratios of the corresponding number of pupils over the number of teachers or school and classroom resources. Similarly, pupil-level resources may also be shared among pupils in a classroom. Although having generated a variable capturing the number of pupil resources available to a sampled pupil, it is not possible to compute an analogue ratio of pupils per class over pupil-level resources in the classroom. In order to do so, data on all pupils in a class would be necessary. It is possible to compute a ratio of the number of sampled grade 6 pupils per school by the number of sampled pupils' resources for each school. Micklewright et al (2012) though show that measures based on a sample rather than on the population will be subject to sampling bias in a regression attenuating the coefficient. Furthermore, Micklewright et al (2012) show that the bias increases as sample size per school decreases. If multigrade teaching is common, this variable is also subject to an additional bias, as it then must be assumed that the sampled grade 6 pupils are representative of all pupils in their classroom. I therefore do not compute a ratio corresponding to pupil-level resources.

All in all, SACMEQ data is quite suitable to investigate class size mechanisms and their effect on pupil achievement. SACMEQ provides repeated cross sectional data from before and after the abolition of direct school fees in Kenya, Tanzania, Zanzibar and Zambia as well as from five 'untreated countries' in the same geographical region. Many questions in both waves of data used are identical or contain the same items so that many variables can be studied over time. Recall the education production function in equation (1), where pupil achievement is a function of the pupil, the household he or she lives in and the school, which in return is a function of available resources, teachers, leadership and a peer effect. The available SACMEQ data covers household behaviour, pupil characteristics and school resources. Corresponding variables are the number of physical school resources, the number of physical classroom resources, the number of pupil materials, the number of teachers per school, household SES, pupils' age, gender, days absent, grade repetition, whether or not they receive extra tuition or help with their homework. Based on these data it is possible to compute variables indicating resource usage in the form of the number of pupils per resource, such as pupils per teacher, pupils per school resource and pupils per classroom resource.

The data also has a few shortcomings. First, the data does not provide headcounts of pupils per school per shift, or class size per shift and instead relies on school enrolment numbers. Thus proxy variables for school and class size must be computed based on school enrolment figures per shift and therefore cannot account for pupil absenteeism. Neither does the data cover all individuals in a school, but only grade 6. Thus, if multigrade classrooms are common it will not be possible to capture peer effects sufficiently. Following the literature on peer effects outlined in section 1.3, data on the average SES of the respective classrooms, or the proportion of female pupils in them would be useful, but is not available. Although it is possible to generate corresponding proxies for these variables from the sampled pupils, this would make the assumption that the composition of the grade 6 pupils in a school is also representative for other grades in the school. This though may not be the case especially in the second wave in the treated countries where lower grades may contain poorer pupils who were unable to attend school prior to the policy change. Also the argument of classical measurement error on such variables as described earlier applies too (cf. Micklewright et al, 2012). Thirdly, the data does not provide information on whether the parents moved the sampled pupils from one school to another or on the type of school the pupils attended previously.

1.3.3 Descriptive statistics

This section documents changes over the 2000-2007 period in the key explanatory variables (see Table 1.3.3). The narrative here will focus on the group 1 countries and discuss the extent to which the variables of interest change in the expected directions. The reader is referred to the appendix for more detailed information.

Physical school and classroom resources

Given the theory outlined above one would expect to observe a relatively fixed supply of physical school resources and teachers to prevail among all group 1 countries, and thus only small statistically significant or no significant changes. Indeed the data indicates that this expectation generally holds for both groups of countries. Table 1.3.3 also shows that the average number of teachers per school remains unchanged in all countries except Zanzibar. Contrary to expectation this education system appears to have invested in hiring teachers so that there are approximately 8.9 teachers more per school on average than in 2007. Furthermore the group 1 countries appear to experience a slight decrease in number of school resources, which are significant in Kenya (-0.6), Tanzania (-0.8) and Zambia (-1.8). These reductions may be due to wear and tear of items that due to prevailing policy prioritisation of ensuring access, have not been replaced. Similarly the average number of classroom resources remains unchanged in Kenya and Zambia, but increase slightly in Tanzania (by 1.1 items) and Zanzibar (by 0.3 items). In respect to the average number of pupil resources (see appendix) there are small significant increases in Tanzania by 0.8 materials, Zambia by 0.4 materials and Zanzibar by 1.1 materials.

In contrast to the group 1 countries, there is no theory to predict the direction of corresponding changes among the group 2 countries, thus heterogeneity of magnitude and direction of changes is likely. The data presented here supports this expectation. For example while there is an increase in the average number of school resources per school in Namibia by 1.1 items, there is a decrease by 0.9 items in Malawi, but no significant change among the other untreated countries. Similarly there is a decrease in pupil resources in Mauritius by 0.4 items, whereas there are significant increases in all other untreated countries. Overall though the supply of school resources and teachers in this group of countries also remains fairly stable.

Table 1.3.3: Change over time (2000 to 2007) in key variables

		Constituents of key explanatory variables						Key explanatory variables				Achievement	
		School size	**Class size /shift**	**School resources / school**	**Class resources [1]**	**Teachers /school**	**Help with home-work**	**Extra tuition**	**Pupils/ school resource**	**Pupils/ class resource[1]**	**Pupils/ teacher**	**Maths**	**Reading**
Group 1	**KEN**	156***	9.6***	-0.6*	0.2	0.5	-1.3	-17.5***	39.1***	2.5	9.6***	-6	-3
		(7)	(1.4)	(0.3)	(0.1)	(0.9)	(1.9)	(3.8)	*(7.2)*	(1.3)	*(1.3)*	(6)	(7)
	TAN	168***	13.0**	-0.8**	1.3***	-1.2	-27.9***	-47.3***	83.8***	-10.4***	16.0***	24***	33***
		(20)	(4.7)	(0.3)	(0.2)	(2.1)	(3.2)	(3.9)	*(16.4)*	(3.2)	*(4.1)*	(6)	(7)
	ZAM	-75***	-31.1	-1.8***	-0.1	-4.0	6.6**	-41.5***	1.7	-2.5	2.3	0	-6
		(12)	(32.0)	(0.5)	(0.2)	(2.6)	(2.2)	(4.1)	*(7.2)*	(7.3)	*(4.8)*	(4)	(6)
	ZAN	127***	0.9	-0.1	0.3*	8.9*	-0.5	-39.5***	31.7	-4.7	-0.8	8	56***
		(21)	(4.6)	(0.3)	(0.1)	(3.5)	(2.2)	(3.8)	*(18.8)*	(4.5)	*(0.9)*	(4)	(5)
	Raw average	**94**	**-1.9**	**-0.8**	**0.4**	**1.0**	**-5.8**	**-36.4**	**39.1**	**-3.8**	**6.8**	**6**	**20**
Group 2	**MAL**	-33	5.0	-0.9***	-0.0	-4.2	12.3***	-65.4***	145.6**	3.8	16.0***	14***	5
		(27)	(5.0)	(0.3)	(0.2)	(2.6)	(4.2)	(3.9)	*(50.2)*	(4.7)	*(4.8)*	(4)	(4)
	MAU	-117***	-15.6	-0.6	0.7***	0.1	-3.5	-6.1*	-10.3	-4.6	-6.9**	23*	27**
		(11)	(11.6)	(0.3)	(0.2)	(1.8)	(2.8)	(2.6)	*(5.5)*	(3.3)	*(2.8)*	(11)	(10)
	NAM	15**	1.2	1.1**	-0.2	0.7	8.2***	-23.8***	-17.6***	0.9	-1.5*	40***	48***
		(5)	(2.2)	(0.4)	(0.2)	(0.9)	(1.2)	(3.3)	*(4.5)*	(2.5)	*(0.7)*	(5)	(6)
	SEY	-182***	-2.9	0.2	0.7***	-4.7	6.1***	-18.6**	-15.6	-0.7	-2.4*	-4	-7
		(10)	(2.1)	(0.4)	(0.2)	(5.8)	(1.2)	(5.5)	*(8.3)*	(0.4)	*(1.0)*	(8)	(10)
	SOU	-12	-7.5	0.3	-0.2	0.7	5.1*	-44.3***	-4.0	2.6*	-1.3	11	3
		(8)	(8.6)	(0.5)	(0.2)	(1.0)	(2.2)	(3.8)	*(7.5)*	(1.1	*(0.8)*	(9)	(11)
	Raw average	**-66**	**-4.0**	**0.0**	**0.2**	**-1.5**	**5.6**	**-31.6**	**19.6**	**0.4**	**0.8**	**17**	**15**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001; 1 estimates for maths classrooms reported, as these are practically identical for reading classrooms. Raw average = unweighted average of estimates of each group*

School and class size

Having established that the supply of physical resources in schools has remained fairly constant among both groups of countries and that the number of teachers per school has only changed significantly in Zanzibar, the Table shows that, as expected, the group 1 countries here experience a strong increase in primary school enrolment. Estimates reveal an increase in average enrolment per shift by 156 pupils in Kenya, a corresponding increase of 168 pupils on average in Tanzania and an increase by 127 pupils in Zanzibar. In contrast average enrolment per shift decreases in Zambia in the same time period by 75 pupils. One of the reasons for the decrease in school enrolment per shift in Zambia may be that schools have changed the number of shifts taught. Estimates for the number of shifts are reported in the appendix and appear to be very stable only decrease slightly in Tanzania and Zanzibar. These suggest that the observed changes in school enrolment per shift in Zambia might be driven by either an increase in number of schools, or by a change in the number of pupils allocated to shifts.

The observed changes in school size lead to the question how the average class size that grade 6 pupils experience changes. According to the data the average number of grade 6 classes per school remains unchanged in all countries except Tanzania where the number increased by 0.38 to 2.16 classes, and in Zambia where the number decreased by 0.67 to 2.2 classes (tables not shown). Assuming multigrade classrooms, this research uses the number of pupils per class per shift to approximate the class size the sampled grade 6 pupils experience. Interestingly, estimates change significantly only for Kenya (9.6 pupils) and Tanzania (13 pupils) and these are the only two countries where the average number of pupils per classroom and shift changes over time. Zambia is also interesting, because despite the decrease in the average number of classrooms per school noted above, the average number of pupils per classroom per shift has not changed significantly. Summarising, SACMEQ data shows that school size does not increase in each group 1 country and that schools compensate for these changes by varying the number of pupils per shift.

Resource usage

The resource usage variables combine the respective number of pupils and divide these by the respective number of school or classroom resources or the number of teachers in the school. One would expect to observe significant changes in the corresponding pupil-resource allocation ratios (see Table 1.3.3). Interestingly, the

changes in numerator and denominator for school resources appear to have counterbalanced each other in Zambia, Zanzibar, Mauritius and the Seychelles, as there are only significant increases in Kenya (+39.1 pupils per school resource), Tanzania (+83.8 pupils per school resource) and Malawi (+145.6 pupils per school resource), and a significant decrease in Namibia by 17.6 pupils per school resource. The table further indicates that pupil teacher ratios increase in Kenya by 9.6 pupils per teacher and in Tanzania and Malawi by 16 pupils per teacher. In contrast, this ratio decreases significantly in Namibia by 17.6 pupils per teacher.

Analogously, in respect to allocation of pupils to classroom resources one would expect to observe changes in Kenya and Tanzania, which are the only countries where on average the expected number of pupils per classroom increases significantly, and Zanzibar, Mauritius and South Africa where there are significant changes in the average number of resources in reading classrooms. Again there appears to be some counterbalancing effects of changes in the numerators and denominators so that the average number of pupils per classroom resource only increases in South Africa by 2.6 and decreases in Tanzania by 10.4.

Household support

In respect to receiving extra tuition the data shows a consistent negative pattern across all sampled countries ranging from -6.1 percent to -65.4 percent in Malawi. Yet there is no such consistent pattern regarding receiving help with homework. This household behaviour increases significantly in all group two countries except Mauritius as well as Zambia in group 1. In Tanzania this household behaviour reduces, but there is no significant change in any of the other countries.

Pupil achievement

The final piece in the puzzle now is to examine the change in pupil achievement. Following the common belief that net enrolment shocks, go hand-in-hand with a decrease in achievement one would expect to observe a reduction in achievement among the group 1 countries. The data though does not support this expectation, as in none of the sampled nine countries does average pupil achievement reduce significantly. Instead Tanzania experiences an increase in average achievement in reading by 33 points and in maths by 24 points and pupils in Zanzibar outperform their peers in the year 2000 by 56 points in reading. Among the group 2 countries, average achievement in maths increases by 27 points in reading and 23 points in maths in Mauritius, by 48 points in reading and 40 points in maths in Namibia,

and pupils appear to outperform their peers in the year 2000 in maths by 14 points in Malawi.

The reader should be reminded at this stage that the bivariate analyses presented above so far are not linked to pupil achievement, so that it is unknown at this point whether and to which extent the documented changes translate into changes in pupil achievement. Multivariate analyses allow resource usage and school composition to covary and estimate their correlation with pupil achievement. The next section discusses the multivariate regression models used.

1.4. Method

Ideally this section should explore all possible mechanisms of class size effects on achievement described in Section 1.3. Yet, due to data limitations it is not possible to explore peer effects or a household's school choice (see section 1.3). This section therefore focuses on two specific mechanisms, usage of resources and teachers in the school and household support.

I pool both waves so that the obtained estimates exploit variation both within and across points in time for each country. I report standardized coefficients below in Tables 1.5.1a to 1.5.3, also known as beta coefficients to ease cross-country comparison, as the dependent variable is likely to have a different variance in each country. Standardized coefficients also allow direct comparison of each predictor as standardization removes the scale of each variable. I obtain the standardized coefficients by standardizing the dependent variable and each of the continuous variables in the model to the corresponding national mean and national standard deviation. I do not standardise the dummy variables on the right hand side, Thus the coefficients on the continuous variables indicate the change in pupil achievement in national standard deviations when the corresponding predictor changes by a national standard deviation; the coefficients on the dummy variables indicate the difference in achievement when the variable is "1" compared to when the variable is zero.

I model the effect of these two mechanisms on achievement by estimating achievement A of individual i in class j, school k as a linear function of a vector of household support H, resource usage R and a vector of control variables C

$$A_{ijkt} = \alpha + \beta_1 H_i + \beta_2 R_{ijk} + \beta_3 C_{ijk} + u_{ijk} \qquad (2).$$

The variables in vector H are whether the pupil receives help with his or her homework and whether the pupil receives extra tuition. Vector R contains access ratios presented in section 1.5.3 as well as the number of pupil-level resources. As described earlier, SACMEQ does not provide adequate information to compute a variable corresponding to the allocation of pupils to pupil-level resources in a given classroom. Nevertheless, I include the number of pupil-level resources available *to a sampled individual*, as especially the effectiveness of how teachers use classroom resources is likely to depend on the extent to which a pupil has all necessary books and writing materials. Also, if pupil-level resources were shared among the pupils in a class, ceteris paribus an increase in class size would reduce the number of these resources available to each pupil.

The control variables included in vector C are closely linked to the discussion in section 1.2 and play an important role in order to interpret both resource and teacher usage and household support in the intended way as proxies for school leader, teacher and household behaviour. First, I control for the school's geographical location. Holding this dimension constant removes differences in teacher and head teacher behaviour due to geographical location, such as teachers being absent so they can travel to collect their monthly pay. Furthermore, holding constant geographical location implies that the supply of schools is distributed evenly across the country and counterbalances geographical asymmetries in the distribution of resources and teachers within them. I then control for school type as public and private schools differ in the contracts they can design. I also hold constant the number of years of pre-service and the number of in-service training days of teachers to ensure that changes in classroom resource usage are not driven by duration of teacher training. Analogously, I control for whether the head teacher has received leadership training.

The support parents can offer their child is conditional on the household's budget. SACMEQ does not have information on household income, so that this is proxied by socioeconomic status. Households also need to decide whether to trade-off education for child labour or the child contributing to home production. The first proxy included in the model for these decisions are the number of days the sampled child was absent in the month prior to data collection. This measure also incorporates days absent due to illness or other reasons, i.e. is measured with error and therefore is not included as a main variable of interest for household behaviour. A similar variable capturing opportunity costs in the long-term is the

pupil's age, which is also included. As households may also differentiate their support for sons and daughters, I include a variable indicating gender. Finally, the vector of control variables also contains a time-dummy in order to remove any unobserved change over time.

Although SACMEQ offers a large number of variables there are a few that are intentionally not included. For example teacher subject competency scores are not included in the model in order to assure comparability of findings across countries, as these data are not available for every country for every point in time. I also do not control for teacher and head teacher gender and academic education, as it is unlikely that these variables equip individuals with the necessary skills to teach or manage and lead a school. Another observable characteristic often noted in the literature is work experience. Literature from the US (cf. Rockoff, 2004; Rivkin et al, 2005; Hanushek, 2005) suggests that teachers improve only in the initial years of their career. Hein and Allen (2013) though do not find a clear improvement of teachers in SACMEQ countries even in the initial years. I therefore do not control for the number of years of teacher or head teacher experience.

The estimates obtained from equation *(2)* are first reported separately for each country, which allows exploring how the specific country context frames school leader, teacher and household behaviour. Thus the findings from these models are likely to be very heterogeneous, as households, school leaders and teachers may differ in their preferences or will be exposed to different incentives. In some countries for example, pupils may receive extra tuition not because they are struggling at school, but rather to give them an additional advantage. Thus the correlation of this household behaviour may positive or negative depending on the existing household preferences in the respective country. Teachers' behaviour may also vary not only due to the duration of their teacher training, both pre and in-service, but also the quality of the training they received, which the equation *(2)* when applied to each country separately cannot account for. I therefore pool the individual country samples together and include country dummies, which accounts for all unobserved country differences.

Nevertheless all findings presented here are correlational despite the large number of control variables, as confounders such as or other household support behaviours are unobserved. It is hard to eyeball the aggregate effect of the omitted variable bias here; teachers might interact differently with pupils of differing ability: pupils of differing ability might be grouped non-randomly into classes, so that teachers

and pupils are not matched randomly. Similarly households may invest less support in more able children. If this is the case, estimates of household support should be biased down. It is also possible that the variables of interest are measured with error, which will attenuate the estimates. The reader should also be aware that in the absence of experimental data, the direction of causality cannot be guaranteed to run from the right to the left in equation *(2)*.

1.5. Findings

1.5.1 Estimates for each country

Resource usage

As expected, the findings in Tables 1.5.1a to 1.5.2b are indeed very heterogeneous and thus underline the importance of the individual country context. Among the group 1 countries, most estimates corresponding to the number of pupils sharing a teacher, school or classroom resources are not associated with achievement. For reading achievement the only significant correlations are for the allocation of pupils to teachers in Kenya and Zambia, where a one standard deviation increase in this predictor is associated with a 0.20 and 0.05 SD decrease in achievement respectively. For maths achievement, the estimate for the sharing of teachers in Kenya is very similar to that for reading. But the estimates show that for both Kenya and Zambia a standard deviation increase in pupils per classroom resources is associated with an increase in pupil achievement by 0.06 and 0.05 SD.

Table 1.5.1a: Final model for reading achievement (group 1 countries)

		Kenya	Tanzania	Zambia	Zanzibar
Control variables	**Wave 2**	0.03 *(0.50)*	0.40*** *(5.99)*	-0.26*** *(-3.43)*	0.44*** *(8.88)*
	School location	-0.02 *(-0.74)*	-0.09** *(-3.18)*	-0.08* *(-2.45)*	-0.10*** *(-4.25)*
	Pre-service teacher training	0.08** *(2.74)*	-0.04 *(-1.49)*	0.04 *(1.23)*	0.03 *(1.07)*
	In-service teacher training	-0.00 *(-1.25)*	-0.00 *(-0.58)*	-0.00 *(-0.23)*	-0.00*** *(-3.58)*
	No leadership training (head)	0.09 *(0.80)*	-0.07 *(-0.11)*	0.06 *(0.91)*	0.03 *(0.57)*
	Government school	-0.39** *(-3.06)*	N/A	-0.47 *(-1.79)*	-0.62*** *(-6.06)*
	Pupil SES	0.20*** *(9.50)*	0.22*** *(10.71)*	0.23*** *(8.13)*	0.16*** *(7.21)*
	Female pupil	-0.11*** *(-4.42)*	-0.22*** *(-6.31)*	-0.13*** *(-3.99)*	-0.01 *(-0.17)*
	Pupil age	-0.19*** *(-10.88)*	0.01 *(0.26)*	-0.09*** *(-4.02)*	-0.03* *(-2.13)*
	Has repeated at least one grade	-0.17*** *(-5.33)*	-0.19*** *(-3.97)*	-0.36*** *(-6.59)*	-0.38*** *(-9.69)*
	Days absent last month	-0.09*** *(-5.79)*	-0.15*** *(-7.34)*	-0.09*** *(-4.31)*	-0.08*** *(-5.22)*
Use of	**School resources**[1]	0.01 *(0.29)*	0.06 *(1.93)*	0.05 *(1.21)*	-0.04 *(-1.65)*
	Teachers[1]	-0.20*** *(-6.07)*	-0.07 *(-1.52)*	-0.05* *(-2.21)*	0.00 *(0.07)*
	Classroom resources[1]	0.01 *(0.29)*	0.03 *(1.10)*	0.04 *(1.69)*	0.01 *(0.14)*
	Number of pupil-level resources	0.01 *(0.23)*	0.03 *(1.26)*	0.07** *(2.96)*	0.10*** *(4.82)*
Household support	**Receives extra tuition**	0.03 *(0.41)*	0.11 *(1.92)*	-0.00 *(-0.07)*	0.12** *(2.65)*
	Receives help with homework	0.15** *(3.29)*	0.03 *(0.67)*	0.12* *(2.15)*	0.16*** *(3.82)*
	N	7260	4701	3198	4886

*Notes: Estimates in national standard deviations; T-statistic in parentheses; 1 ratio of pupils per resource or teacher; * p<.05, ** p<.01, *** p<.001*

Table 1.5.1b: Final model for reading achievement (group 2 countries)

		Malawi	Mauritius	Namibia	Seychelles	South Africa
Control variables	**Wave 2**	-0.02 *(-0.22)*	0.38*** *(5.63)*	0.26*** *(5.06)*	-0.13 *(-1.89)*	-0.17* *(-2.38)*
	School location	0.03 *(0.40)*	0.00 *(-0.01)*	-0.08*** *(-3.54)*	0.01 *(0.33)*	-0.12** *(-2.85)*
	Pre-service teacher training	0.06 *(1.48)*	0.02 *(0.69)*	0.10*** *(4.22)*	0.07 *(1.08)*	0.07** *(2.65)*
	In-service teacher training	0.00 *(0.29)*	0.00 *(0.58)*	-0.00*** *(-4.49)*	-0.00 *(-1.38)*	-0.01 *(-1.61)*
	No leadership training (head)	-0.21* *(-2.37)*	0.10 *(1.46)*	-0.01 *(-0.21)*	0.03 *(0.46)*	0.01 *(0.22)*
	Government school	-0.11 *(-1.15)*	-0.14 *(-1.77)*	-0.52*** *(-3.86)*	-1.03*** *(-9.38)*	0.05 *(0.43)*
	Pupil SES	0.11*** *(4.22)*	0.25*** *(12.00)*	0.30*** *(13.99)*	0.19*** *(7.88)*	0.40*** *(19.22)*
	Female pupil	-0.22*** *(-6.41)*	0.14*** *(4.41)*	0.03*** *(3.93)*	0.22*** *(14.18)*	0.12*** *(6.03)*
	Pupil age	-0.10*** *(-4.99)*	0.04* *(2.00)*	-0.09*** *(-7.86)*	0.04* *(2.06)*	-0.10*** *(-7.25)*
	Has repeated at least one grade	-0.21*** *(-5.44)*	-0.72*** *(-12.22)*	-0.31*** *(-15.36)*	-0.24* *(-2.39)*	-0.28*** *(-10.86)*
	Days absent last month	-0.06** *(-3.04)*	-0.12*** *(-5.46)*	-0.04*** *(-4.20)*	-0.10*** *(-5.22)*	0.05* *(2.05)*
Use of	**School resources**[1]	-0.04 *(-1.33)*	0.09** *(2.71)*	-0.05 *(-1.77)*	0.01 *(0.18)*	-0.08*** *(-3.47)*
	Teachers[1]	-0.14*** *(-4.06)*	-0.11 *(-1.29)*	-0.03 *(-1.21)*	0.07** *(3.03)*	-0.11*** *(-3.59)*
	Classroom resources[1]	0.02 *(0.84)*	0.10 *(1.17)*	-0.07** *(-2.77)*	-0.02 *(-0.82)*	-0.00 *(-0.07)*
	Number of pupil-level resources	0.03 *(1.45)*	0.04 *(1.16)*	0.09*** *(5.30)*	0.12*** *(4.42)*	0.04* *(2.42)*
Household support	**Receives extra tuition**	-0.02 *(-0.31)*	0.47*** *(7.77)*	-0.09** *(-2.71)*	0.22*** *(3.72)*	-0.10 *(-1.84)*
	Receives help with homework	-0.01 *(-0.29)*	0.00 *(0.10)*	-0.03 *(-1.00)*	0.22* *(2.25)*	0.07* *(2.02)*
	N	4822	3037	10813	2921	8683

*Notes: Estimates in national standard deviations; T-statistic in parentheses; 1 ratio of pupils per resource or teacher; * p<.05, ** p<.01, *** p<.001*

Table 1.5.2a: Final model for maths achievement (group 1 countries)

		Kenya	Tanzania	Zambia	Zanzibar
Control variables	**Wave 2**	0.00 *(0.04)*	0.14*** *(5.05)*	-0.11** *(-2.90)*	0.16*** *(5.00)*
	School location	-0.02 *(-0.58)*	-0.09* *(-2.47)*	-0.07* *(-2.20)*	-0.13*** *(-4.97)*
	Pre-service teacher training	0.01 *(0.37)*	-0.09** *(-2.77)*	0.03 *(1.21)*	0.03 *(1.17)*
	In-service teacher training	-0.03 *(-1.39)*	-0.01 *(-0.70)*	-0.01 *(-0.35)*	-0.00 *(-0.09)*
	No leadership training (head)	0.00 *(0.10)*	-0.01 *(-0.28)*	0.00 *(0.14)*	0.00 *(0.07)*
	Government school	-0.10** *(-2.94)*	N/A	-0.10 *(-1.76)*	-0.13*** *(-6.45)*
	Pupil SES	0.15*** *(6.77)*	0.20*** *(8.23)*	0.20*** *(6.22)*	0.13*** *(4.97)*
	Female pupil	-0.16*** *(-10.99)*	-0.16*** *(-8.33)*	-0.09*** *(-4.51)*	-0.04* *(-2.13)*
	Pupil age	-0.20*** *(-11.12)*	0.00 *(0.08)*	-0.09*** *(-4.19)*	-0.03 *(-1.50*
	Has repeated at least one grade	-0.07*** *(-3.92)*	-0.09*** *(-4.53)*	-0.16*** *(-6.97)*	-0.12*** *(-7.28)*
	Days absent last month	-0.07*** *(-4.44)*	-0.11*** *(-3.32)*	-0.07*** *(-3.82)*	-0.06*** *(-3.83)*
Use of	**School resources**[1]	-0.02 *(-0.89)*	0.05 *(1.44)*	0.05 *(1.32)*	-0.05 *(-1.63)*
	Teachers[1]	-0.19*** *(-4.91)*	-0.03 *(-0.57)*	-0.05 *(-1.80)*	0.04 *(1.04)*
	Classroom resources[1]	0.06* *(2.30)*	0.01 *(0.44)*	0.05* *(2.26)*	-0.01 *(-0.70)*
	Number of pupil-level resources	0.04* *(1.98)*	0.03 *(1.33)*	0.07** *(2.87)*	0.06*** *(4.26)*
Household support	**Receives extra tuition**	0.04 *(1.44)*	0.06* *(2.22)*	0.00 *(0.08)*	0.07* *(-2.44)*
	Receives help with homework	0.03 *(1.57)*	0.01 *(0.65)*	0.03 *(1.74)*	0.04* *(2.34)*
	N	7272	4609	3214	4786

*Notes: Estimates in national standard deviations; T-statistic in parentheses; 1 ratio of pupils per resource or teacher; * p<.05, ** p<.01, *** p<.001*

Table 1.5.2b: Final model for maths achievement (group 2 countries)

		Malawi	Mauritius	Namibia	Seychelles	South Africa
Control variables	**Wave 2**	0.07 *(1.44)*	0.18*** *(5.05)*	0.15*** *(5.99)*	-0.06 *(-2.00)*	-0.05 *(-1.31)*
	School location	0.03 *(0.46)*	0.01 *(0.16)*	-0.06** *(-2.60)*	0.01 *(0.66)*	-0.11*** *(-3.53)*
	Pre-service teacher training	0.02 *(0.60)*	0.02 *(0.79)*	0.08*** *(3.56)*	0.05 *(0.78)*	0.08** *(2.76)*
	In-service teacher training	0.03 *(1.10)*	0.00 *(0.08)*	-0.05*** *(-3.32)*	-0.04 *(-1.14)*	-0.05* *(-2.42)*
	No leadership training (head)	-0.07 *(-1.71)*	0.06 *(1.65)*	0.01 *(0.44)*	0.02 *(0.89)*	-0.01 *(-0.37)*
	Government school	-0.05 *(-1.27)*	-0.04 *(-1.49)*	-0.12*** *(-3.77)*	-0.17*** *(-11.83)*	0.01 *(0.55)*
	Pupil SES	0.09*** *(3.43)*	0.25*** *(12.02)*	0.30*** *(13.68)*	0.20*** *(7.99)*	0.39*** *(16.31)*
	Female pupil	-0.14*** *(-7.69)*	0.07*** *(4.42)*	-0.00 *(-0.57)*	0.19*** *(12.55)*	0.03** *(3.01)*
	Pupil age	-0.09*** *(-4.40)*	0.04* *(1.98)*	-0.09*** *(-8.49)*	0.04 *(2.00)*	-0.07*** *(-5.40)*
	Has repeated at least one grade	-0.09*** *(-4.68)*	-0.29*** *(-12.21)*	-0.15*** *(-13.58)*	-0.04 *(-1.69)*	-0.13*** *(-10.69)*
	Days absent last month	-0.04* *(-2.19)*	-0.12*** *(-5.48)*	-0.04*** *(-4.01)*	-0.10*** *(-5.31)*	0.05 *(0.75)*
Use of	**School resources**[1]	-0.01 *(-0.26)*	0.12*** *(4.06)*	-0.04 *(-1.75)*	-0.05* *(-2.05)*	-0.08** *(-3.05)*
	Teachers[1]	-0.15*** *(-4.25)*	-0.24** *(-2.68)*	-0.05 *(-1.78)*	0.08** *(3.29)*	-0.12** *(-2.96)*
	Classroom resources[1]	0.04 *(1.20)*	0.22* *(2.53)*	-0.03 *(-1.13)*	0.04 *(1.53)*	-0.01 *(-0.27)*
	Number of pupil-level resources	0.03 *(1.25)*	0.05 *(1.22)*	0.08*** *(5.72)*	0.14*** *(4.06)*	0.04* *(2.03)*
Household support	**Receives extra tuition**	0.03 *(0.96)*	0.18*** *(7.92)*	-0.04* *(-2.63)*	0.10** *(3.38)*	-0.03 *(-1.11)*
	Receives help with homework	-0.01 *(-0.30)*	-0.00 *(-0.13)*	-0.00 *(-0.11)*	0.03 *(1.12)*	0.01 *(0.75)*
	N	4799	3042	10877	2865	8605

*Notes: Estimates in national standard deviations; T-statistic in parentheses; 1 ratio of pupils per resource or teacher; * p<.05, ** p<.01, *** p<.001*

Among the group 2 countries, estimates are more heterogenous. For example, while a one standard deviation increase in pupils sharing a teacher is associated with a

0.14 and 0.11 SD decrease in reading achievement in Malawi and South Africa, it is not associated with a change in achievement in Mauritius and Namibia, but is associated with a 0.07 SD increase in the Seychelles. Corresponding estimates for maths achievement are very similar. In addition, increasing the number of pupils sharing a maths teacher reduces pupil achievement (-0.24 SD). Analogously, while a standard deviation increase of the number of pupils sharing school resources is not associated with a change in pupil reading achievement in Malawi, Namibia and the Seychelles, it is associated with a 0.09 SD increase in Mauritius and an 0.08 SD decrease in South Africa. For maths achievement the estimates again are similar for Mauritius and South Africa, but in addition, the estimated effect for the Seychelles is a statistically significant 0.05 SD decrease.

As described above, equation *(2)* also includes the number of pupil-level resources. In Section 1.3 I describe that due to data limitations this variable is only available for the sampled pupils and it is not possible to obtain a variable of the number of pupils sharing a pupil-level resource in a given class that is not subject to classical measurement error. Nevertheless this variable is still interesting, for if these resources were to be shared among pupils in a class, an increase in class size ceteris paribus would lead to a decrease in this variable. The results indeed suggest that a reduction in the number of pupil-level resources can have an impact on achievement: For example a one standard deviation decrease is associated with a reduction in achievement in the Seychelles by 0.14 and 0.12 SD. A similar but weaker pattern exists also in Zambia, Zanzibar, Namibia and South Africa, but pupil level resources are not associated with pupil achievement in any subject in Tanzania, Malawi and Mauritius.

Household support

In respect to the correlation of household support behaviours with pupil achievement the data suggest different patterns for group 1 and 2 countries. Among group 1, receiving help with his or her homework is significantly associated with an increase in the pupil's reading achievement by 0.15 SD in Kenya, 0.16 SD in Zanzibar and 0.12 SD in Zambia. In contrast receiving help with homework is only associated with maths achievement in Zanzibar (0.04 SD increase). Receiving extra tuition on the other hand is associated with an increase in achievement by 0.06 SD in Tanzania for maths only, and in Zanzibar by 0.12 and 0.07 SD in reading and maths. Among group 2 receiving help with his or her homework is only associated with an increase in reading achievement in the Seychelles (0.22 SD) and

South Africa (0.07 SD), but it is not associated with maths achievement in any of these five countries. Receiving extra tuition on the other hand appears to be more consistent and is associated with increases in achievement of up to 0.47 SD in Mauritius. Namibia stands out as with estimates of a reduction in maths (0.09 SD) and reading (0.04 SD) achievement in respect to receiving extra tuition. The estimates may be due to the correlational nature of the multivariate analysis applied, so that the direction of causality may also be from the left hand side of the equation to the right. Thus, these estimates might indicate that in Namibia pupils, whose achievement is lower on average, receive extra tuition.

1.5.2 Pooled model

Although the reported estimates above demonstrate the important role the country context plays in framing human behaviour, the main aim of this paper is to gain a deeper understanding of class size mechanisms in general. As described in section 1.4 I therefore complement the analysis above with a pooled model. The estimates reported in Table 1.5.3 show that across the sampled countries the usage of classroom resources is not associated with pupil achievement. The usage of school resources is only associated with a reduction in reading achievement (0.03 SD), but not in maths. Yet as above the usage of teachers, i.e. pupil-teacher-ratios are associated with a reduction in achievement in both subjects (0.03 SD in maths and 0.04 SD in reading).

Table 1.5.3: Pooled model

		Maths	Reading
Control variables	**MAL[2]**	-0.97***	-0.75***
		(-23.37)	*(-17.47)*
	MAU[2]	-0.19**	-0.62***
		(-3.17)	*(-11.57)*
	NAM[2]	-1.09***	-0.74***
		(-29.83)	*(-18.92)*
	SEY[2]	-0.64***	-0.42***
		(-11.36)	*(-6.73)*
	SOU[2]	-0.96***	-0.87***
		(-20.05)	*(-18.51)*
	TAN[2]	-0.12**	0.31***
		(-2.99)	*(7.51)*
	ZAM[2]	-1.14***	-0.86***
		(-29.15)	*(-19.59)*
	ZAN[2]	-0.75***	-0.32***
		(-19.6)	*(-8.34)*
	Wave 2	0.08***	0.10***
		(3.29)	*(4.25)*
	School location	-0.04***	-0.07***
		(-3.93)	*(-5.48)*
	Pre-service training	0.04***	0.07***
		(3.24)	*(5.27)*
	In-service training	-0.02***	-0.02***
		(-3.42)	*(-3.74)*
	No leadership training (head)	-0.01	0.01
		(-0.69)	*(0.24)*
	Government school	-0.26***	-0.25***
		(-5.17)	*(-5.09)*
	Pupil SES	0.24***	0.29***
		(22.52)	*(28.78)*
	Female pupil	-0.10***	0.03**
		(-10.21)	*(2.87)*
	Days absent last month	-0.05***	-0.06***
		(-5.48)	*(-5.46)*
	Pupil has repeated at least one grade	-0.22***	-0.28***
		(-18.54)	*(-23.66)*
	Pupil age	-0.05***	-0.08***
		(-8.16)	*(-12.07)*
Use of	**School resources[1]**	-0.01	-0.03**
		(-1.45)	*(-2.83)*
	Teachers[1]	-0.03*	-0.04***
		(-2.5)	*(-3.21)*
	Classroom resources[1]	-0.01	-0.02
		(-0.9)	*(-1.24)*
	Number of pupil-level resources	0.06***	0.07***
		(8.39)	*(8.53)*
House-hold	**Receives extra tuition**	0.06***	0.04*
		(3.34)	*(2.16)*
	Receives help with homework	0.02	0.03*
		(1.26)	*(2.12)*
	N	49856	50321

*Notes: Estimates in standard deviations; T-statistic in parentheses; 1 ratio of pupils per resource or teacher; 2 estimates relative to Kenya; * p<.05, ** p<.01, *** p<.001*

Also the number of pupil-level resources is associated with achievement in both subjects and implies that any sharing of these resources with a pupil's peers would harm the achievement of pupil sharing his or her pens, books and rulers, while benefiting the pupil receiving these resources.

In the previous section the results showed that household behaviour could be associated with achievement both positively and negatively, which could be related to differing household preferences across the sampled countries. For example pupils may receive extra tuition not because they are struggling academically, but because their parents want to give them an additional advantage at school. The pooled model removes these differences. The findings show that in this sample receiving extra tuition is associated with an increase in achievement in maths by 0.06 SD and by 0.04 SD in reading. On the other hand pupils receiving homework only seem to profit in respect to their reading achievement.

The second question this sub-section wants to shed light on is the extent to which households can exacerbate or mitigate the effect of class size mechanisms located in the school. The findings of the pooled model show that households can indeed play an important role. Ceteris paribus, the findings suggest that the negative effect a standard deviation increase in class size can have on a pupil's reading achievement through the usage of school resources and the teacher can be mitigated by supporting the pupil through extra tuition and helping them with their homework. For maths achievement, supporting the learner with extra tuition could even outweigh the negative effect of having to share the teacher with one standard deviation more peers.

1.5.3 Summary

The findings in sections 1.5.1 and 1.5.2 show that the effect of an increased amount of pupils sharing school, classroom resources or teachers varies in magnitude and direction both within and across the nine countries. This underlines that the aggregate effect of a change in class size depends heavily on the individual country context, i.e. the preferences of households, teachers and school leaders as well as the incentives they are faced with. The pooled estimates show that when these country contexts are accounted for, sharing the teacher with an increased number of pupils is the main threat to a pupil's achievement. Yet household behaviours such as providing the child with extra tuition in particular can counterbalance or even outweigh having to share the teacher with more peers.

1.6. Conclusion

This paper explores the underlying mechanisms of class size effects and argues that a change in class size has a compositional effect and an effect on how school resources and teachers are used. The compositional effect combines two parts: First it is likely to change existing peer dynamics as literally individuals will either be removed or added to a class depending on the direction of change in class size. Second, if households care about the quality of their children's education, they will need to reconsider the kind of school they send their children to, how much academic support such as extra tuition and help with homework to provide their children with, and the amount of education to invest in their children.

A change in class size is also likely to require teachers to use the given physical school resources differently, i.e. apply different teaching methods when designing learning environments. Similarly school leaders will need to reconsider how they allocate resources to classrooms, and pupils and teachers to them. In this framework individuals do not need to behave optimally, so the directions of these mechanisms can be positive, negative, or zero and therefore require empirical estimation.

It uses sub-Saharan African data of pupils in grade 6 from nine SACMEQ countries collected in the year 2000 and the year 2007. The paper employs multivariate linear regression to estimate the association of the school resource, classroom resource and teacher usage as well as households' support towards their children with pupil achievement in maths and reading. The findings underline that the aggregate effect of changes in class size depend heavily on the specific country context as well as the directions and magnitudes of the individual underlying mechanisms. support the adopted framework, which expects estimates for the variables of interest to vary in magnitude and direction both within each country and between them. For example, estimates for the number of pupils sharing a teacher stand out especially, as these are associated with substantial reductions in Kenya at around 20 percent of a standard deviation, Malawi 14 and 16 percent of a standard deviation and South Africa at 10 and 12 percent of a standard deviation, but an increase of approximately 7 percent of a standard deviation in the Seychelles. These estimates indicate that teachers themselves are valuable resources that in Kenya, Malawi and South Africa, are unable to use themselves effectively when class size increases. At the same time the number of pupils

sharing a teacher is not associated with a change in learning in Tanzania, Zanzibar, Mauritius and Namibia.

Estimates from a pooled model show that when these country contexts are accounted for, sharing the teacher with an increased number of pupils is the main threat to a pupil's achievement. Yet household behaviours such as providing the child with extra tuition in particular can counterbalance or even outweigh having to share the teacher with more peers.

Although this paper focuses on the sub-Saharan African context, it also has implications for class size studies in general. The findings here indicate that household support has the potential to counterbalance the effect of class size mechanisms located in schools or may be the main driving force. Furthermore the magnitude of household support effects is highly context specific and should be addressed accordingly. Thus Hanushek's (2003) observation that the gap between pupils in smaller and larger classes in the Tennessee STAR randomised control trial does not increase over time could be explained by households adapting their support over time. Similarly studies based on Maimonides' Rule (Angrist and Lavy, 1999) should consider household support. Even without such information, providing longitudinal data is available, value-added approaches using lagged achievement may circumvent this issue.

Chapter 2

Why do teachers differ in their quality? Pupil-fixed effects estimates from twelve sub-Saharan countries

2.1. Introduction

There is common agreement that teachers matter in respect to pupil learning outcomes, and that there is substantial variation in estimated teacher quality in many countries. Empirically, teacher quality is usually explored in two ways. The first estimates the overall effect a teacher has on his or her pupils' achievement in a subject. Such evidence is predominantly from the US and where researchers have access to large longitudinal datasets of pupils matched to teachers and schools. This data has enabled researchers to estimate the 'total' effect of a teacher on his or her pupil's achievement and to isolate this estimate from school and pupil-level variables. The evidence shows between 0.06 and 0.15 SD of a pupil's achievement in a subject is associated with their teacher and teachers appear to matter more for maths achievement than for reading achievement (cf. Nye et al, 2004; Rockoff, 2004; Hanushek et al, 2005; Rivkin et al, 2005; Kane et al, 2006; Aaronson et al, 2007; Clotfelter et al, 2007). Estimates from the UK are slightly higher suggesting that approximately 0.17 to 0.19 SD of pupils' GCSE achievement can be associated with the teacher (Slater et al, 2011).

Another line of research focuses on explaining why teachers differ in their effectiveness by estimating the association of teacher observable characteristics with pupil achievement (e.g. Vegas and De Laat, 2003; Bourdon et al, 2005; Aaronson et al, 2007; Metzeler and Woessmann, 2012). This allows exploring the skills teachers use when teaching and gauging their importance.

This paper aims to explore this particular aspect of teacher quality in sub-Saharan Africa by examining the association of observable teacher characteristics and pupil achievement. Specifically, this paper focuses on characteristics that relate to teachers' *subject-matter competency*, such as teachers' academic education and their achievement in a subject-matter test in the subjects they teach, and their *pedagogic competency*, such as their pre- and in-service teacher training as well as their teaching experience.

To date, developing countries lack the rich, high-quality datasets containing multiple observations of teachers matched to pupils that are available in the US. Thus the existing estimates for the African countries in particular are obtained from simple statistical methods and are therefore subject to more sources of bias than the US estimates.

In addition, although the US estimates are more robust, teacher labour markets in sub-Saharan southern and eastern Africa are likely to operate differently, leading to a very heterogeneously skilled teacher labour force, so that US estimates will not be suitable to inform policy in developing country contexts. Most strikingly, in developing countries teacher training and continual professional development may not be available to all teachers due to a lack of adequate financial and human resources. Especially in countries that are expanding their education systems to meet their Millennium Development Goals, many are forced to hire large numbers of new, potentially less well trained and qualified teachers due to a lack of choice. The financial constraints may also affect the quality of the skills teachers acquired when they were attending school, their teacher training and the professional development they have received. Mulkeen (2010) reports that teaching in sub-Saharan countries tends to offer small salary increases that are determined by years of experience until a plateau is reached. Mulkeen reports that in Zambia and the Gambia a teacher's salary is also partly determined by the qualifications he or she holds. Thus a different kind of people might be attracted into teaching in these countries compared to those in developed economies. In short, teachers in these countries may vary substantially in their quality and it is important to know which kind teachers' skills explain teacher quality, i.e. a change in pupil achievement.

In this paper I use data from the Southern African Consortium for the Monitoring of Educational Quality (SACMEQ) collected at two points in time either in the year 2000/2002 or 2007 from twelve countries. The data are cross-sections of primary school pupils in grade 6 matched to schools and their respective teachers. In order to explore the association of teacher observable characteristics with pupil achievement I exploit the fact that pupils are tested in reading and maths at the same point in time, the scores on each of these tests are on the same scale, and that the respective teachers in each subject are also observed. I follow Clotfelter et al (2006) and allow for pupil-fixed effects, so that I associate the difference in the observable characteristics of the teacher who teaches a pupil maths and the teacher who teaches a pupil reading with the difference in a pupil's maths and reading achievement score. In the SACMEQ countries used here pupils either have the same or two different teachers in maths and reading. I refer to estimating *teacher quality* when pupils are taught by different teachers. In the case when pupils are taught by the same teacher in both subjects, most teacher variables are the same for both subjects; hence the difference between them is zero. But in SACMEQ, as in other developing countries, teachers also sit an achievement test

in the subject they teach. If a teacher teaches both subjects, he or she is tested in both subjects. Thus I can estimate teachers' *differential teaching ability* as a function of the difference in their subject-matter competency. As the pupils are tested at the same point in time, all variables that do not differ between these two subjects for a given pupil are removed in the differencing. The method thereby accounts for all school-level variables, the non-random matching of pupils to teachers, and all subject-invariant characteristics at the pupil level known or unknown, observed or unobserved.

Although the focus of this study lies on cross-country patterns to gauge whether these patterns differ from those obtained from developed economy contexts such as the US, there is a second motivation for reporting estimates separately for each country. This reason is methodological as previous research by Fehrler et al (2009) and Zuze (2010) use the same data. I argue that these studies do not sufficiently address potential biases.

My findings support the pattern emerging from the US, that there is no consistent pattern of teacher characteristics associated with pupil achievement. Furthermore, my findings suggest that the studies by Fehrler et al (2009) and Zuze (2010) who use the same SACMEQ data indeed suffer from endogeneity, as my findings do not reproduce their results. Moreover a number of findings are counterintuitive. As it is reasonable to assume that teachers need to combine both their subject-matter and pedagogic competency I then move on to examine the interaction of the two kinds of competencies. I do so by adding interactions between proxy variables of each kind of competency and find again that in most cases the interactions are not statistically significant. Nevertheless the evidence suggests that these two competencies are substitutes rather than compliments in six of the ten examined countries, i.e. pupil achievement suffers when these competencies are combined.

The remainder of this paper is structured as follows: Section 2.2 introduces the reader to the existing literature and contrasts the substantive findings from developed country contexts with the few estimates from African contexts. It follows with an overview of the basic statistical methods employed so far. Section 2.3 then introduces the reader to the SACMEQ data and the implications arising for estimating the effect of observable teacher characteristics on pupil achievement. In Section 2.4 I present the method applied and discuss the associated strengths and weaknesses and argue that the pupil-fixed effects models applied here are the best

way, given the data, to eliminate as many potential sources of bias as possible. Section 2.5 presents my findings and the last section concludes.

2.2. Literature

To the best of my knowledge, the existing literature so far tends to estimate the association of teacher characteristics with pupil achievement in an ad hoc manner without specifying why certain characteristics should affect pupil achievement. For example, one might assume that the teacher effect T could be a function of their subject-matter competency W and their pedagogic competency D. I conceive teachers' subject-matter competency as a function of their academic qualifications Q, and their pedagogic competency a function of their teaching experience, X, their pre- and in-service teacher training S, incentives M that motivate teachers to differ in their behaviour as well as their intrinsic motivation O.

$$T_j = g(W_j(Q_j), D_j(X_j; S_j), M_j, O_j) \qquad (1)$$

Apart from the observable characteristics mentioned in the previous paragraph, teacher achievement tests in the subjects is a proxy of W; teachers' contractual status may incentivise them to differing behaviour. For example, teachers employed on temporary contracts may be more effective as they are trying to secure a permanent position.

So far previous research approximates equation *(1)* assuming additive linearity, i.e. T_j is the sum of teacher j's academic qualifications, teacher training, teaching experience, etc. Following equation *(1)* it is then reasonable to explore the compatibility of W_j and D_j. For example teachers with higher subject-matter competency may have a stronger impact on pupil achievement the better their pedagogic skills. I will return to this issue in section 3.

2.2.1 What does previous research find?

To aid the reader with this sub-section I will follow equation *(1)* and contrast the evidence from the developed countries with findings from the African context.

In respect to subject-matter competency, teachers in developed countries do not tend to be tested in the subject-matter of the subjects they teach, so that researchers are limited to estimating the association of teachers' academic qualifications and their pupils' achievement. Aaronson et al (2007) find no significant impact of academic qualifications. By contrast, Clotfelter et al (2006) find a significant impact for teachers holding graduate degrees compared to

bachelor's only and for teachers holding state licensure. In respect to pedagogic competency, Hanushek et al (2005) find that different kinds of teaching qualifications or passing scores on state licensure exams do not significantly predict pupil achievement.

In respect to teachers' pedagogical competency, Aaronson et al (2007) find that teachers holding a certificate in bilingual education are associated with a reduction in pupil achievement by 0.08[9] SD. There is conflicting evidence on how teachers' work experience affects pupil achievement. Rockoff (2004) finds an effect of 0.15 to 0.18 SD for a ten-year increase in work experience for pupils' reading achievement. With pupils' maths computation, he finds no experience effect after 8 years, and for pupils' vocabulary and maths concepts test scores teachers do not seem to improve after being on the job for six years. In contrast Hanushek et al and Rivkin et al (both 2005) find that teachers do not improve after being on the job for five years. On the other hand Clotfelter et al (2007) find that a standard deviation increase in teachers' work experience is associated with an increase in pupil achievement ranging from 0.06 to 0.12 SD in maths and 0.04 to 0.09 SD in reading.

Corresponding estimates from sub-Saharan Africa are sparse as only few African datasets of teachers matched to pupils are available. Most notable are the Southern African Consortium for the Monitoring of Educational Quality (SACMEQ) datasets, from mainly English speaking sub-Saharan East African countries, and Program for the analysis of education systems of states and member-governments of the Conference of Ministers of Education of Countries sharing the French language **(**PASEC) datasets from French speaking West African countries. Both are repeated cross-sectional datasets and the findings vary according to the countries sampled.

The literature from Africa suggests that incentives such as differing contractual status may play a role. Bourdon et al (2005) find that pupils of contractual teachers do not differ significantly in their achievement from pupils taught by regular civil servant teachers in Niger. In contrast, Vegas and De Laat (2003) find that pupils of regular civil servant teachers outperform pupils of contractual teachers in Togo. In respect to proxies of pedagogic skills, Vegas and De Laat find a negative impact of teacher experience. Fehrler et al (2009) conduct a cross-country analysis of 21 countries form French (PASEC I) and English (SACMEQ II) speaking African

[9] Computed from data provided in Aaronson et al (2007).

countries collected between 1996 and 2002. They find that academic qualifications are not significant in the French-speaking countries, but have an effect on pupil test scores of between 0.02 and 0.04 standard deviations in the English speaking ones. Further they find that teacher tests have an impact of between 0.21 and 0.32 standard deviations in SACMEQ countries, but are not significant in PASEC countries[10]. Zuze (2010) analyses SACMEQ II and TIMSS data for Botswana and finds that both teacher academic education and teacher qualification are not significant in SACMEQ and none of the teacher characteristics are significant in the TIMSS data. The only highly significant predictor of pupil achievement in the SACMEQ data is the teacher test. Thus according to equation (1) the literature suggests that teachers' subject-matter competency is more important than teachers' pedagogic competency in the SACMEQ data.

As will be shown below, these findings must be considered of inferior quality compared to those from the US, as the methods these authors use cannot convincingly remove many sources of bias.

2.2.2 Which methods are used?

The typical economic approach to estimating the effect of teacher observable characteristics on pupil achievement is to apply an education production function, according to which the achievement A of pupil i with teacher j in school k, is a function of the pupil P, the pupil's household background H, the teacher T, the pupil's peers C, school and classroom resources R, and school leadership L (cf. Hanushek, 1979).

$$A_{ijk} = f(P_i, H_i, T_j, C_k, R_{jk}, L_k) \qquad (2).$$

This model is often empirically approximated using OLS regression, which assumes additive linearity of the input vectors in equation *(2)*.

$$A_{ijk} = \alpha + \beta_1 P_i + \beta_2 H_i + \beta_3 T_j + \beta_4 C_k + \beta_5 R_{jk} + \beta_6 L_k + \varepsilon_{ijk} \qquad (3).$$

When using observational data, such as SACMEQ, the T are likely to be correlated with the other right hand side variables. This is likely to be the case due to selection processes. For example teachers may prefer to apply to work in certain schools and schools may have preferences to hire certain teachers. Similarly governments might allocate teachers to schools where there is the greatest need for them. Within schools school leaders may decide to allocate their best teachers to

[10] The reason for this could be due to the nature of the tests used in the different data sets. Whereas SACMEQ test the teachers on similar tests as the pupils are presented with, PASEC test teachers with a fictitious pupil's French dictation and the teachers are required to identify the mistakes.

certain classes. These two selection effects may run in opposing directions. Thus, it is important that the other inputs in (3) are sufficiently controlled for. But even when equation (3) is estimated directly using all observable data available for the right-hand-side input vectors, as Fehrler et al (2009) and Zuze (2010) do, the β_3 are likely to be subject to endogeneity as one does not tend to observe all elements of these input vectors, such as teachers' intrinsic motivation. Econometrically, the error term in equation (3), ε, then contains unobserved variables, U, at all three levels

$$\varepsilon_{ijk} = U_i + U_j + U_k \qquad (4).$$

Alternatively one can extend this simple approach by simultaneously modelling the non-random matching of pupils and teachers as a function of parental school choice (cf. Vegas and De Laat, 2003; Bourdon, 2005). Again these matching approaches can only model matching on observable characteristics thereby ignoring any possible matching on unobservable characteristics, which may result in misleading findings.

The key to addressing these endogenous biases in estimates of teacher characteristics is to have multiple observations of teachers and / or pupils. With longitudinal data on pupil achievement it is possible to include the pupil's prior achievement in equation *(3)* as an element of P_i. Such 'value-added' approaches usually claim to simultaneously control for pupil ability, as well as for prior motivation of the pupil, family involvement and prior knowledge (cf. Todd and Wolpin, 2003). Following this line of argumentation, including prior achievement also controls for the effect of past teachers on a pupil's achievement. In other words estimates of the impact of observable teacher characteristics T observed at a given point in time will contain the effect of other teachers if prior achievement is not controlled for and will therefore be biased. With longitudinal data only variables that change between these two points in time are potential sources for bias.

Alternatively one might exploit the hierarchical structure of the data. One could for example exploit the fact that teachers teach more than one pupil. By assuming a teacher-fixed effect, i.e. that the teacher's effect on achievement is the same for all of his or her pupils, one can estimate the average pupil achievement associated with the sampled teacher while allowing for pupil-level characteristics to influence pupil achievement. But as the teacher-fixed effects approach cannot estimate the association of variables at the teacher level, a second simultaneous regression is needed to estimate the association of teacher observable characteristics and the

teacher fixed-effects estimates (Aaronson et al, 2007). The disadvantage of this approach is that the teacher-fixed effects may also contain the effect of the school level. Thus, the estimates on T might be subject to endogeneity if one cannot sufficiently control for school-level factors. Aaronson et al (2007) also demonstrate that the number of observations for each teacher is crucial for the precision of the teacher-fixed effects estimates, and to minimise measurement error.

Metzeler and Woessmann (2012) propose an alternative method to estimating the effect of teacher characteristics on pupil achievement. They use cross-sectional data on pupil-teacher matches in Peru, where the pupils are observed in reading and maths. They estimate equation *(3)* separately for each subject simultaneously as seemingly unrelated regressions allowing these two equations to be correlated for each pupil. The advantage of this method is that it accounts for each input in equation *(2)* to have a differential effect on pupil achievement in the two subjects. The disadvantage of this approach is that in order to interpret teacher estimates as causal, one needs to assume that all input vectors in equation *(2)* are sufficiently controlled for.

In contrast to Metzeler and Woessmann (2012) Clotfelter et al (2006) argue that allowing for pupil-fixed effects addresses non-random matching of pupils to teachers. Algebraically, allowing for pupil-fixed effects, when pupils are observed in two subjects at the same point in time,

$$A_{i1k} = \alpha_1 + \beta_1 P_i + \beta_2 H_i + \beta_3 T_1 + \beta_4 C_k + \beta_5 R_{jk} + \beta_6 L_k + \varepsilon_{i1k} \qquad (5a)$$

$$A_{i2k} = \alpha_2 + \beta_1 P_i + \beta_2 H_i + \beta_3 T_2 + \beta_4 C_k + \beta_5 R_{jk} + \beta_6 L_k + \varepsilon_{2k} \qquad (5b),$$

can be achieved for example by solving equation *(5b)* for $\beta_1 P_i + \beta_2 H_i$ and replacing this in equation *(5a)*. This removes all variables that do not differ within pupils. Then the difference in pupil achievement in the two subjects is explained by the difference in intercepts and the difference in subject-level variables T and R

$$A_{i1} - A_{i2} = (\alpha_1 - \alpha_2) + \beta_3(T_1 - T_2) + \beta_5(R_1 - R_2) + (\varepsilon_{i1} - \varepsilon_{i2}) \qquad (6).$$

In summary, US researchers have access to rich longitudinal datasets that link teachers to pupils and annual measures of pupil scores, which allow them to address most sources of bias. The estimates produced in these papers can be considered the 'gold-standard' when it comes to estimating the effect of teachers on pupil achievement in the absence of data on random allocation of pupils to teachers, and have made a substantial contribution to our understanding of teacher labour markets. In contrast, developing countries do not have access to this

kind of data, so that simpler methods of data analysis prevail that cannot address the various sources of bias sufficiently.

2.3. Method

In this research I exploit the fact that in SACMEQ data, pupils and their respective teachers are observed at a given point in time in two subjects, which has not been done so far with this data to my knowledge. I apply a variation of equation *(6)* that can be summarised as

$$A_{i1} - A_{i2} = (\alpha_1 - \alpha_2) + \gamma(T_1 - T_2) + \mu(Z_1 - Z_2) + (\varepsilon_{i1} - \varepsilon_{i2}) \qquad (7)^{11},$$

where the difference in maths ($j=1$) and reading ($j=2$) achievement A of pupil i is a linear function of the difference in intercepts α, the difference in teacher characteristics, as well as the difference in a vector of control variables Z, and the difference in errors ε.

As I am examining differences within pupils, the fact that pupils are sampled at two points in time for each country does not play a role, as the sampled pupils are different individuals. Algebraically, the factor time is subsumed within the pupil-fixed effects. Thus, I can pool pupils sampled at both points in time for each country. Especially as some of the examined countries' education systems undergo substantial expansions, a rehiring of new teachers as well as a redistribution of teachers may accompany this. In other words the non-random distribution of teachers to pupils is likely to change. This might result in pupils being exposed to more different teachers and will thus improve precision of the estimated coefficients as well as of the corresponding standard errors, whereby the latter should reduce the chance of making a type-II error.

In the SACMEQ countries used here pupils either have the same teacher or two different teachers for maths and reading. I therefore estimate equation *(7)* twice for each country; once for the sample of pupils taught by the same teacher and then for those taught by different teachers, providing the sample sizes of each pupil type is sufficient. I refer the reader to section 2.4 where I describe the sample used. I refer to estimating *teacher quality* when pupils are taught by different teachers. In the case when pupils are taught by the same teacher in both subjects, most variables are the same for both subjects, hence the difference between them is zero and these variables are dropped. In SACMEQ, as in other developing countries, teachers also sit an achievement test in the subject they teach. If a teacher teaches both subjects,

[11] For simplicity I have changed the naming of the coefficients in equation (6).

he or she is tested in both subjects. Thus I can estimate teachers' *differential teaching ability* as a function of the difference in their observed subject-matter competency.

As described above, equation *(6)* nets out all confounders that are constant for each pupil at the given point in time with only differences at the classroom/subject level remaining. In other words, the estimates produced should contain less sources of bias than any previous estimates for these countries. Yet applying equation *(6)* to the SACMEQ data will not necessarily produce causal estimates of T for a number of reasons. In the following I will describe which reasons these are and how, if possible, these issues can be addressed with the data used. In other words, I will explain which variables are included in Z_j and which should be included, but cannot.

Endogeneity

Evidence suggest that teachers may not be able to maximise the use of their pedagogic and subject-matter competency. This may be due to a number of variables, which, if these also directly affect pupil achievement, will lead to an endogenous bias in the coefficients of the teacher variables (cf. Wooldridge, 2002).

First of all, while teaching, teachers will need to use classroom resources, such as school books. I therefore include the number of classroom resources as an element of Z_j. Second, teachers might use their subject-matter and pedagogic skills more effectively with certain pupils (cf. Dee, 2005; Rawal and Kingdon, 2010). In SACMEQ it is possible to generate variables indicating whether the teacher is the same gender as the pupil and I include this variable in Z_j. In the African context, the age of pupils differs substantially, as pupils do not all start school at the expected age and neither do all progress through the education system at the expected speed. I therefore also include the age difference between the pupil and the teacher in Z_j. Teachers may also be able to use their competencies more effectively with pupils of differing ability or prior knowledge. This could be addressed if pupils' prior achievement in both subjects were available. But as SACMEQ only observes pupils at one point in time, and no administrative data with a pupil's performance on another test in the same subject is available, this potential source of bias cannot be addressed. Fourth, larger class sizes might be harder to manage, due to more sources of noise and commotion. This potential bias could be addressed by including class size in Z_j, but as class size does not vary within pupil in any country in the SACMEQ data, i.e. the maths and reading classes have the same number of pupils, it is not possible to do so.

Apart from potential sources of endogeneity at the pupil and classroom level, the effectiveness of teachers' competencies may also differ with teachers' motivation. For example, if more highly motivated teachers receive better training and more motivated teachers teach better, there is reason to believe that omitting the difference in teachers' motivation could be a source of endogeneity, if this has a direct effect on pupil achievement. Aslam and Kingdon (2011) explore the correlation of teacher classroom practices with pupil achievement using data from Pakistan and point out that apart from indicating how teachers teach, these variables may also proxy teacher motivation. SACMEQ contains questions on teachers' teaching values, goals and practices, but the questions are designed in a way that does not generate sufficient variation in the data for two reasons: First, the respective variables are based on three-level Likert items corresponding to 1 "Not very important", 2 "Of some importance" and 3 "Very important". Second, the questions themselves, to the eye of an educator, do not provoke variation in the data. For example, in respect to the goals of reading, the following have to be scored according to their importance:

- Making reading enjoyable
- Extending pupils' vocabulary
- Improving word attack skills
- Improving pupils' reading comprehension
- Developing a lasting interest in reading

It does not surprise that these items create very little variation, as these are all core goals of learning to read.

Another source of endogeneity lies in different contractual status of teachers. As noted in equation *(1)*, teachers' effectiveness T may vary as a result of economic incentives. For example, teachers employed on temporary contracts might be more effective as they are trying to secure a permanent position. Similarly teachers on temporary contracts might also differ systematically in their observable characteristics, by for example being less well trained as their peers on permanent contracts. In SACMEQ information on teachers' contractual status is not available at the teacher level, but only at the school level. Here the head teacher is asked to report the number of teachers either on permanent or temporary contracts. Due to the pupil-fixed effects strategy in the cross-sectional data, the school level is differenced out, but it is likely that an endogenous bias remains in the teacher estimates.

Apart from endogeneity, which will bias the magnitude and direction of the coefficients, the applied method is likely to face additional challenges due to measurement error on the continuous variables, the teacher test scores and their years of experience.

Attenuation

Observable teacher characteristics are potentially subject to random measurement error. If this is the case, observed teacher variables such as their years of experience will contain not only the true value but also an orthogonal disturbance. This disturbance will reduce the strength of the association between the difference of years of experience and the difference in pupil achievement. In other words the estimated mean association will be biased towards zero. The disturbance will also inflate the standard error associated with this estimate and could therefore encourage type-II errors. Common measurement error methods such Structural Equation Models (cf. for example Bentler and Chou, 1987) depend on multiple covariates of the "true" variable, for example repeated measurements of years of experience. These models estimate a latent, error free variable based on these multiple observations that is included in the estimation. In SACMEQ no such multiple observations of right-hand-side variables are available.

To summarise, employing equation *(5)* with the SACMEQ data would purge the teacher estimates from more potential sources of endogeneity than any previous research using SACMEQ data, but a number of sources remain. Some of these potential biases can be addressed, so that I include a vector of control variables Z_j for each subject that contains the teacher's gender, the difference in the teacher's and pupil's age, whether the teacher is of the same gender as the respective pupil, whether the pupil receives extra tuition in the subject and the number of classroom resources.

The main variables of interest here, T, are teachers' academic education and their achievement in a subject-matter test in the subjects they teach, as proxies of teachers' *subject-matter competency*, and their pre- and in-service teacher training as well as their teaching experience, as proxies for their *pedagogic competency*. As mentioned in section 2.2, it is also interesting to examine the interaction of these two kinds of competencies. In order to do so I run a separate model whereby I add interactions of the test score, as the proxy of subject-matter competency, with each of the three variables representing pedagogic competency discussed above (one year or more pre-service training, no in-service training, years of experience) to the vector X.

As the method described by equation *(7)* holds constant everything that does not vary within pupil, it is also possible to pool all countries together. I do not do so, as the obtained estimates from such a model will average across the entire sample, which may mask the cross-country heterogeneity or be dominated by large sample sizes. As the main foci of this study is to explore how useful evidence from developed countries is for individual countries, as well as to demonstrate the bias in previous research that provided estimates for separate countries, reporting pooled models does not add any value here and may even be misleading.

2.4. Data

2.4.1 Context and quality of the data

The data used are from the second and third wave of SACMEQ, the Southern and Eastern Africa Consortium for the Monitoring of Educational Quality. The second wave was collected in the year 2000, the third in 2007. There are fifteen member countries taking part in total, of which twelve are used here. These countries are Botswana, Kenya, Lesotho, Malawi, Mozambique, Namibia, the Seychelles, Swaziland, Tanzania, Uganda, Zambia and Tanzania-Zanzibar (from now on referred to as Zanzibar). In between these two time periods the education systems of Kenya, Tanzania, Zambia and Zanzibar especially underwent a large expansion, as substantial proportions of these countries' school-aged population were not attending school. In other words, these countries lagged far behind reaching their Millennium Development Goal of universal primary education in the year 2000 (cf. UNESCO Global Monitoring Report, 2011). Other countries, such as Botswana, Mozambique and Swaziland, also needed to expand their education systems, but rather than suddenly abolishing all direct school fees and thereby encouraging a demand-shock of pupils wanting to attend schools, these latter countries gradually expanded access to their education systems.

In each wave SACMEQ samples grade 6 pupils attending *registered* government or non-government schools. Thus the sampled data may not be representative for all schools if, as is the case in Kenya, private schools exist that are not registered with the government. Oketch et al (2010a) highlight that of their sample of low cost private schools in Nairobi, one quarter are not registered. But it is unknown how large this potential bias is for the entire country. The participating governments also have some leeway in the definition of the target population. For example, the

Tanzanian samples only contain individuals in government schools. In short, it is hard to ascertain the actual corresponding populations and the reader should keep this in mind.

Pupils are sampled in two stages. In the first stage schools are stratified by region and number of grade 6 pupils, whereby the latter is truncated into two categories, small and large. Schools are then selected with probability proportional to the number of their respective grade 6 pupils. In the second wave, a maximum of 20 pupils are sampled at random in each school in all countries except Namibia and Mauritius where a maximum of 40 pupils are sampled and the Seychelles, where all grade 6 pupils in the selected schools are sampled. In the third wave the maximum sample size for the second stage is increased in all countries apart from the Seychelles where again all pupils in grade 6 are sampled in the chosen schools. This time a maximum of 50 pupils are sampled in Namibia, Mauritius, Kenya and Zanzibar and a maximum of 25 pupils in all other countries. According to SACMEQ's technical documents and country reports the response rates of pupils in the second wave are 89 percent in Kenya, 83 percent in Malawi, 93 percent in Mauritius, 92 percent in Namibia, 96 percent in the Seychelles, 85 percent in South Africa, 77 percent in Tanzania, 75 percent in Zambia and 83 percent in Zanzibar (Onsomu et al, 2005). In the third wave response rates are 91 percent in Kenya, 79 percent in Malawi, 89 percent in Mauritius, 91 percent in South Africa (Moloy and Chetty, 2010; Milner et al, 2011; Sauba and Lutchmiah, 2011; Wasanga et al, 2012). Response rates are not available for the other countries. Table 1 reports the achieved sample size in each country. In total 81,185 pupils and 12,489 teachers are sampled in these twelve countries. The sample sizes range from 2,964 pupils and 217 teachers in the Seychelles to 11,446 pupils and 1,784 teachers in Namibia.

Table 2.1: Sample size of teachers and pupils

Country	Abbreviation	Pupils	Teachers	Sample used
Botswana	BOT	7,190	877	7,186
Kenya	KEN	7,704	836	7,698
Lesotho	LES	7,395	546	7,383
Malawi	MAL	5,114	626	5,103
Mozambique	MOZ	6,480	2,845	6,422
Namibia	NAM	11,446	1,784	11,387
Seychelles	SEY	2,964	217	2,961
Swaziland	SWA	7,169	869	7,166
Tanzania	TAN	7,035	1,065	7,029
Uganda	UGA	7,920	891	7,891
Zambia	ZAM	5,463	724	5,429
Zanzibar	ZAN	5,305	1,209	5,247
	Total	**81,185**	**12,489**	**80,902**

SACMEQ surveys consist of different parts: Head teachers are surveyed on their demographics and those of the school, as well as their own teaching. Similarly teachers are surveyed about their demographics and teaching. Finally the pupils are asked to fill in a questionnaire about themselves and their family.

Pupils are tested in maths and reading in wave two and in a third subject 'health/science' in wave three. Teachers are also tested in the subject they teach. For example if a teacher teaches maths, reading and 'health/science' in wave three, he or she will be tested in these three subjects. The pupil and teacher tests are similar, but not identical. SACMEQ pupil and teacher tests scores are based on multiple-choice items and are estimated using Rasch models (cf. Rasch, 1960). To make SACMEQ test score measures comparable over time, the tests in each subject in each wave contain common items. The test scores in both subjects are standardised to a SACMEQ-wide mean of 500 and a standard deviation of 100 in wave 2. Thus the scores in both subjects are on the same scale, so that taking the difference between these two, as implied by equation (6) is valid.

Rasch models (ibid), also known as one-parameter item response models, are commonly used to estimate competency in a specific subject and are also applied in the OECD PISA studies. Compared to the two and three parameter item response models, Rasch models neither account for guessing nor that items may differ in their quality to discriminate between higher and lower abilities (cf. van der Linden and Hambleton, 1997). Baird et al (2011) criticise the validity of PISA test scores because of the way pilot studies are conducted. Pilot studies are intended to identify items in which certain samples differ substantially from expected levels, so that such items can be improved or removed for the main study. As PISA tests are

not piloted in every participating country, it cannot be established whether all items used are valid. Kreiner and Christensen (2013) reanalyse PISA reading skills data from 2006 and provide evidence indicating that the Rasch models used by PISA are neither valid within nor across countries and that the produced rankings of countries are not robust.

These weaknesses may also exist in SACMEQ, as the extent of piloting is unknown to the author. Further the language of testing may be a problem. SACMEQ tests are conducted in English, but English typically will not be the mother tongue of every sampled pupil within countries. Similarly children from wealthier households or from urban centres may come into contact with English sooner than poorer children. Also the age at which English is introduced as medium of instruction can differ across countries. In the light of these potential flaws in the provided Rasch scores, this research therefore assumes that their validity holds.

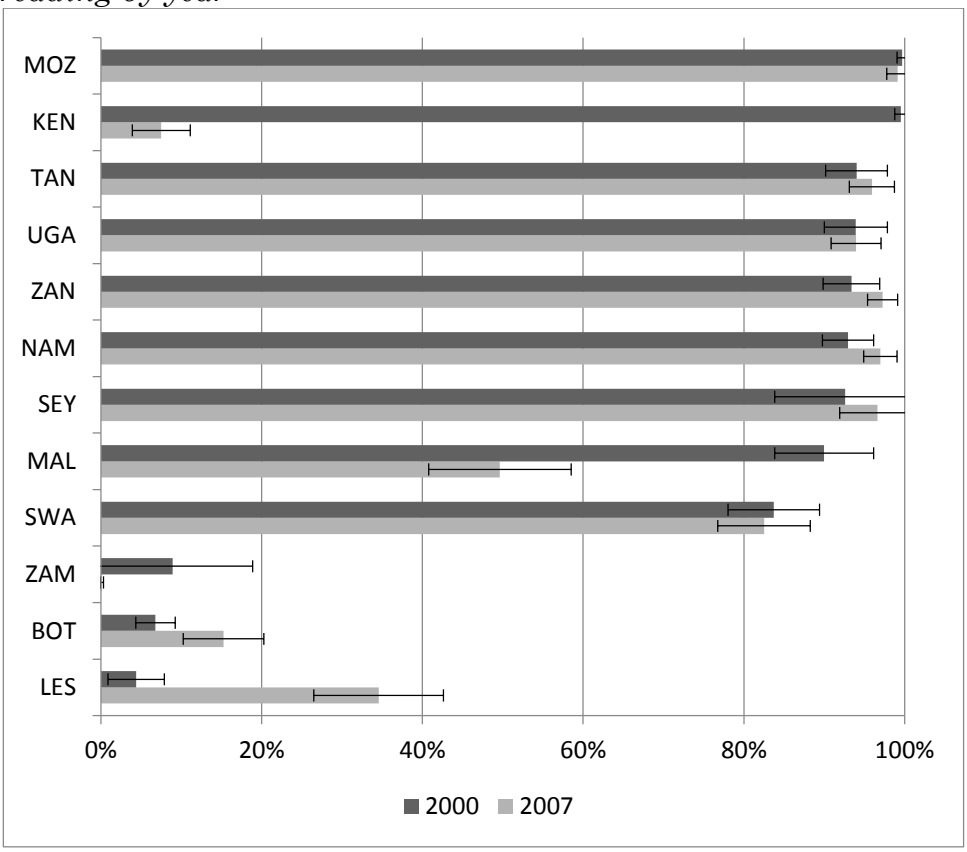
In their technical documents SACMEQ acknowledge that item non-response exists 'occasionally'. In these cases, if less than 15 percent of the respective item is missing, SACMEQ impute the missing values by replacing continuous variables with the mean and categorical variables with the mode of the respective lowest level of aggregation in the data. In the case of this research, for most of the data used here the proportion of missing data is negligible or at maximum 5 percent in a given country. Only in Uganda there appears to have occurred a problem when surveying maths teachers in 2007. Here approximately 15 percent of observations are missing information on these teachers' gender and subject-matter test, and approximately 9 percent of observations in Uganda lack information on the number of classroom resources and the number of years of teacher experience. For all cases with missing data I include missing dummies. For continuous variables, such as teacher test scores, I replace the missing information with the country mean for the respective year; for binary variables I categorise missing observations as "0" and include corresponding "missing data" dummies to correct for this arbitrary classification of the missing information. The only variables that are not imputed are the dependent variables, the pupils' achievement in maths and reading. Thus the descriptive statistics and the findings reported below are based on the number of observations in the column titled "sample used" in Table 1. Comparing this column to the "total sample size" shows that the former, which indicates the number of observations without missing observations on either of the pupil scores, that missing observations on the pupil scores are negligible.

2.4.2 Which teachers are the pupils facing?

As the title suggests, this sub-section describes the observable characteristics of the teachers the sampled pupils are exposed to. Thus the unit of analysis throughout this sub-section is the pupil.

Figure 2.1 shows the percentage of pupils for each country that are taught by different teachers in the two sampled subjects. It shows two patterns. First, over 90 percent of pupils are taught maths and reading by different teachers in Mozambique, Tanzania, Uganda, Zanzibar, Namibia, and the Seychelles in both SACMEQ II and III. In contrast a vast majority of pupils is taught by the same teacher in both subjects in Zambia, Botswana and Lesotho. Second, this graph shows that especially in Kenya, Malawi and Lesotho, there are substantial changes in how teachers are allocated to pupils. In Lesotho 4 percent of the sampled pupils are taught by the different teachers in the year 2000, while 35 percent of the pupils sampled in 2007 have different teachers in the two subjects. The data shows a similar change in Malawi, but in the opposite direction. Here 90 percent of the pupils sampled in the year 2000 have different teachers in maths and reading. Yet in 2007 this number has dropped to 50 percent. On the other hand, while practically all pupils in the year 2000 have different teachers in Kenya, only 7 percent of pupils sampled in 2007 appear to have different teachers. This latter change might be due to the extreme demand-shock experienced by the Kenyan education system. Thus, in order to absorb the increased amount of pupils' teachers who used to teach one subject might have been reassigned to teach more subjects.

Figure 2.1: Percentage of pupils taught by different teachers in maths and reading by year



Note: Error bars indicate 95 percent confidence interval. Underlying standard errors account for clustering by school. Countries are sorted by values in the year 2000.

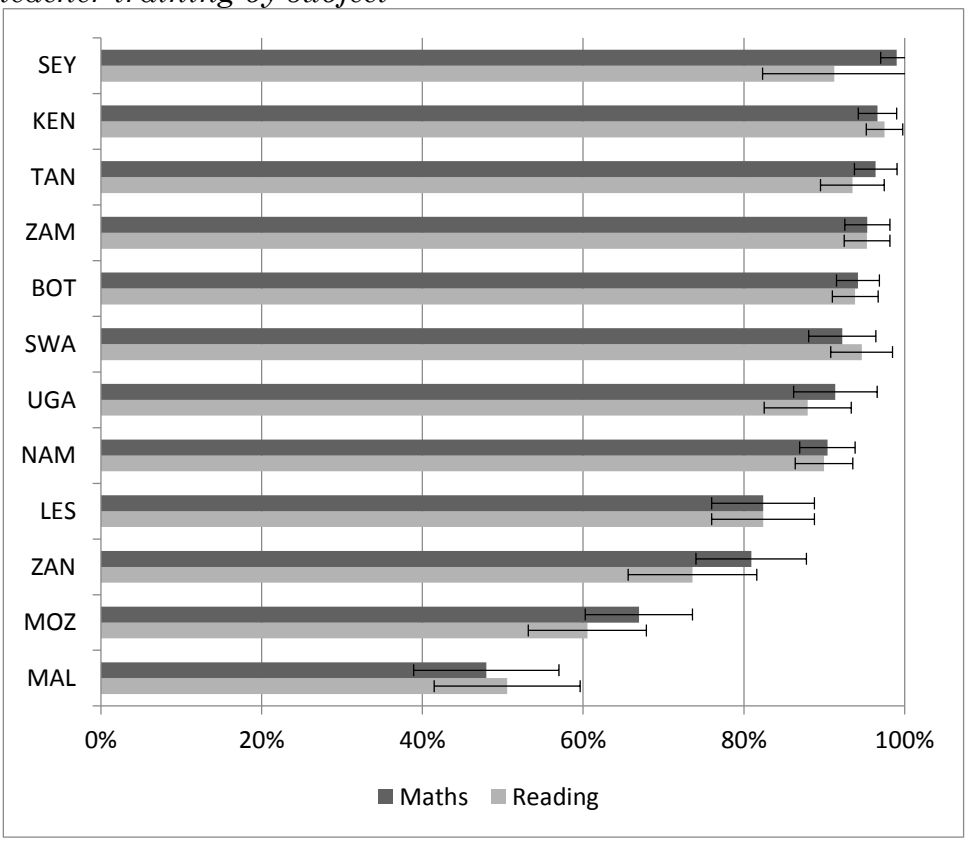
As pointed out above, it is possible to run two models for each country, one for pupils taught by the same teacher and the other for pupils taught by different teachers, if the sample size for any of these two groups of pupils is large enough. I will therefore only report estimates for any of the two models if the sample size of pupils is at least 1,000[12] to ensure I have sufficient statistical power. Thus I will report estimates for *teachers' differential teaching ability* as a function of the difference in their subject-matter competency, for Botswana, Kenya, Lesotho, Malawi, Swaziland and Zambia. Analogously, I will report estimates for *teacher quality* for all countries except Botswana and Zambia.

The following paragraphs move the focus to describing the observable characteristics of the teachers that the sampled pupils face. For brevity I follow equation *(1)* and limit the focus to variables that relate to teachers' subject-matter competency and pedagogic competency.

[12] Although this threshold may appear arbitrary, sample sizes of approximately 1,000 individuals are typical in political science, for example when predicting elections and the popularity of political parties.

Due to the numerous constraints developing countries face especially in African contexts, not all teachers receive the full pre-service training that their country offers. Figure 2.2 therefore shows the percentage of pupils whose maths or reading teachers have experienced at least one year of teacher training. The graph shows that in eight countries 90 percent or more pupils have a teacher who has at least been trained for a year, this is not the case in Lesotho, Zanzibar, Mozambique and Malawi. In the latter three countries it appears that the amount of pre-service training a teacher receives correlates with the subject they teach as in Zanzibar and Mozambique the percentage of pupils whose teacher has received at least on year of pre-service training is higher in maths than for reading.

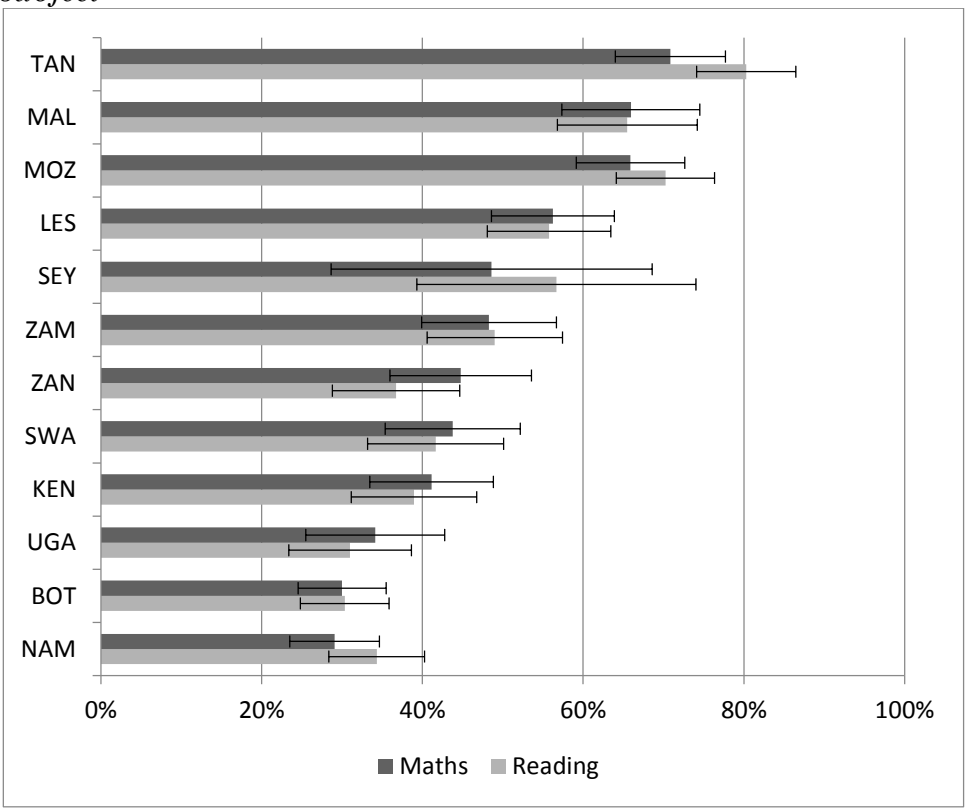
Figure 2.2: Percentage of pupils whose teacher has more than 1 year of teacher training by subject



Note: Error bars indicate 95 percent confidence interval. Underlying standard errors account for clustering by school. Countries are sorted by values of Maths teachers. The reported estimates are the intercepts obtained from linear regressions for each subject in each country that account for changes over time and missing data dummies.

Analogously, not all teachers receive in-service training. Figure 2.3 illustrates the proportion of pupils whose teachers have experienced at least one day of in-service training. The graph shows that in none of the sampled countries are the figures above 80 percent. Most strikingly less than 40 percent of pupils are taught by teachers who have received any in-service training in Uganda, Botswana, and Namibia in both subjects, as well as Zanzibar and Kenya for reading.

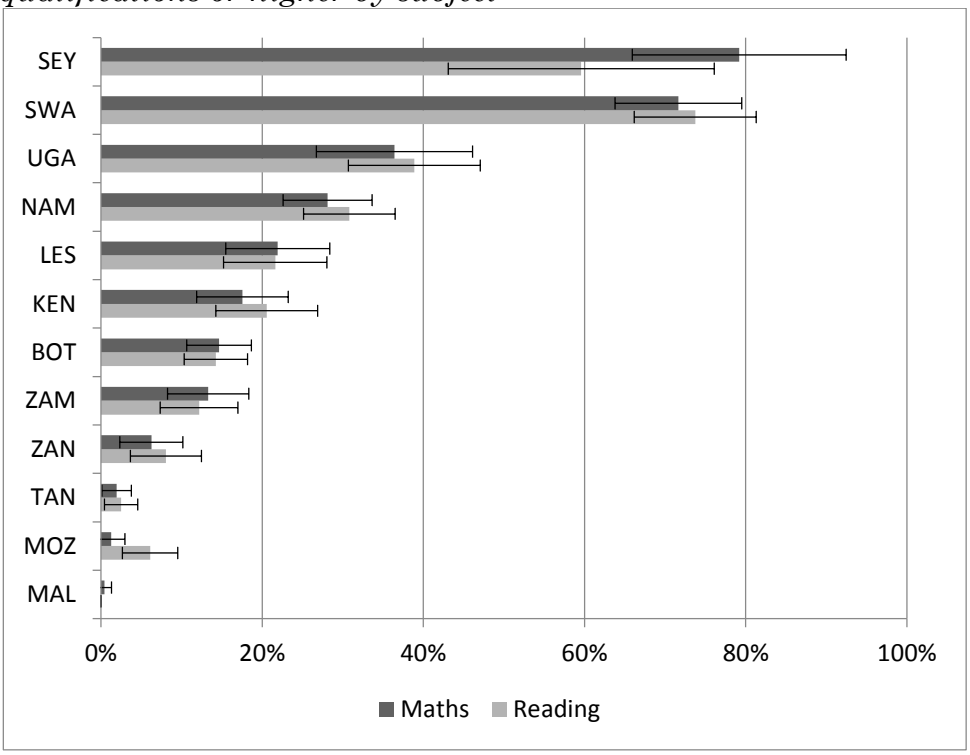
Figure 2.3: Percentage of pupils whose teacher has no in-service training by subject



Note: Error bars indicate 95 percent confidence interval. Underlying standard errors account for clustering by school. Countries are sorted by values of Maths teachers. The reported estimates are the intercepts obtained from linear regressions for each subject in each country that account for changes over time and missing data dummies.

This pattern continues in respect to teachers' academic qualifications. Whereas in developed countries teachers usually hold undergraduate or even postgraduate qualifications, developed countries tend to face a lack of adequately qualified teachers and therefore need to recruit individuals willing to teach, but who do not hold qualifications from higher education. Figure 2.4 shows the percentage of pupils whose teacher holds university entry qualifications or above. The graph shows that while in the Seychelles and Swaziland a majority of pupils are taught by teachers with such qualifications, in all other sampled countries this is not the case, especially in Malawi, Mozambique and Tanzania.

Figure 2.4: Percentage of pupils whose teacher holds university entry qualifications or higher by subject



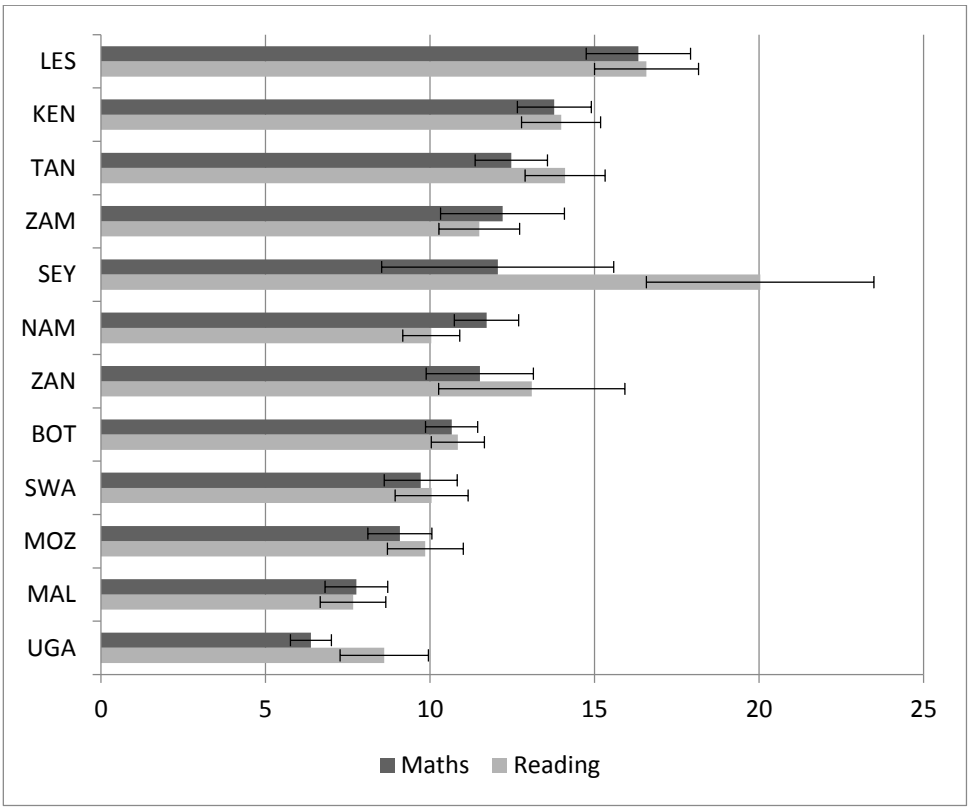
Note: Error bars indicate 95 percent confidence interval. Underlying standard errors account for clustering by school. Countries are sorted by values of Maths teachers. The reported estimates are the intercepts obtained from linear regressions for each subject in each country that account for changes over time and missing data dummies.

According to the World Bank data from the year 2007, the three countries among these twelve with the lowest GDP per capita are Malawi, Mozambique and Uganda. Thus, one would expect that these countries lack sufficient resources to train their teachers as well as the other countries. Comparing Figures 2.2 to 2.4, Malawi and Mozambique follow this expected pattern and rank last in Figures 2.2 and 2.4, and respectively at the top in Figure 2.3. Uganda does not follow this expected pattern and ranks 7th in Figure 2.2, 3rd from the bottom in Figure 2.3 and 3rd in Figure 2.4, which suggests that teachers in this country are better trained and qualified as one may expect.

I now move the focus to the left hand side of equation (6). Figure 2.7 documents mean pupil achievement scores for both subjects by country. The SACMEQ test scores are scaled to a SACMEQ-wide mean of 500 and a standard deviation of 100 points. The figure shows, as one would expect, that of these twelve countries six countries better than the mean in both subjects. In Uganda, pupils on average appear to achieve slightly above the mean in maths, but not in reading. Figure 2.7 also shows substantial differences between countries. Pupils in Malawi, for

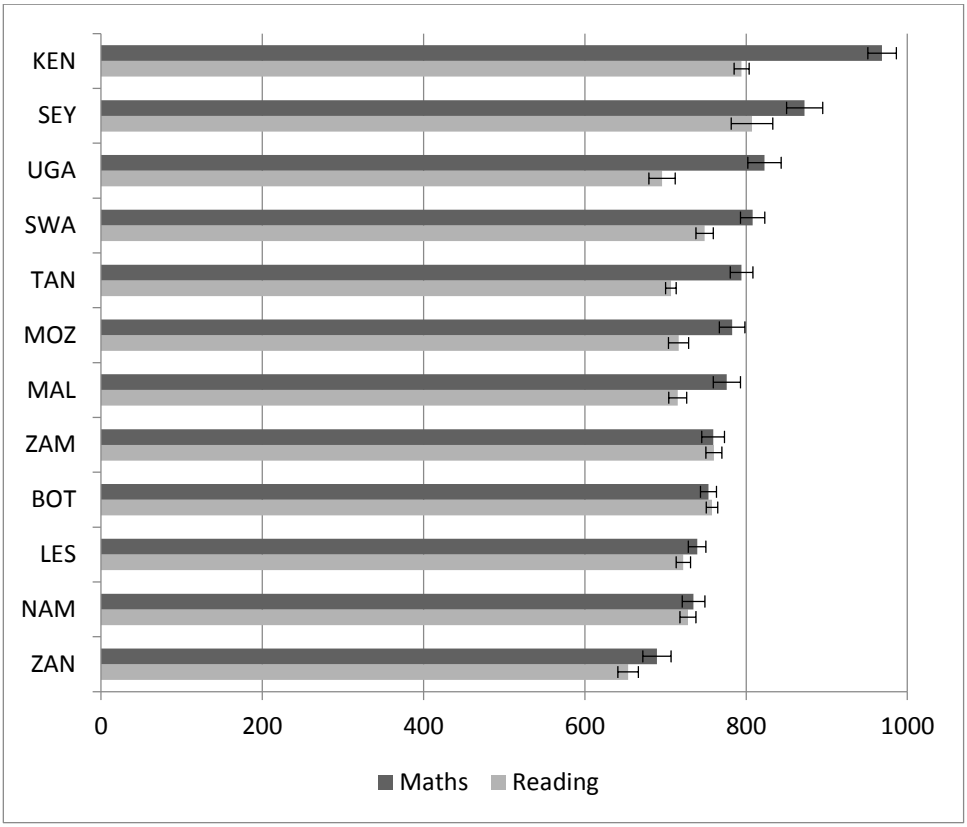
example, are approximately 1.5 SD (153 test points) behind their peers in the Seychelles in reading.

Figure 2.5: Mean years of teaching experience of pupils' maths and reading teachers



Note: Error bars indicate 95 percent confidence interval. Underlying standard errors account for clustering by school. Countries are sorted by values of Maths teachers. The reported estimates are the intercepts obtained from linear regressions for each subject in each country that account for changes over time and missing data dummies.

Figure 2.6: Mean subject-matter test score of pupils' maths and reading teachers

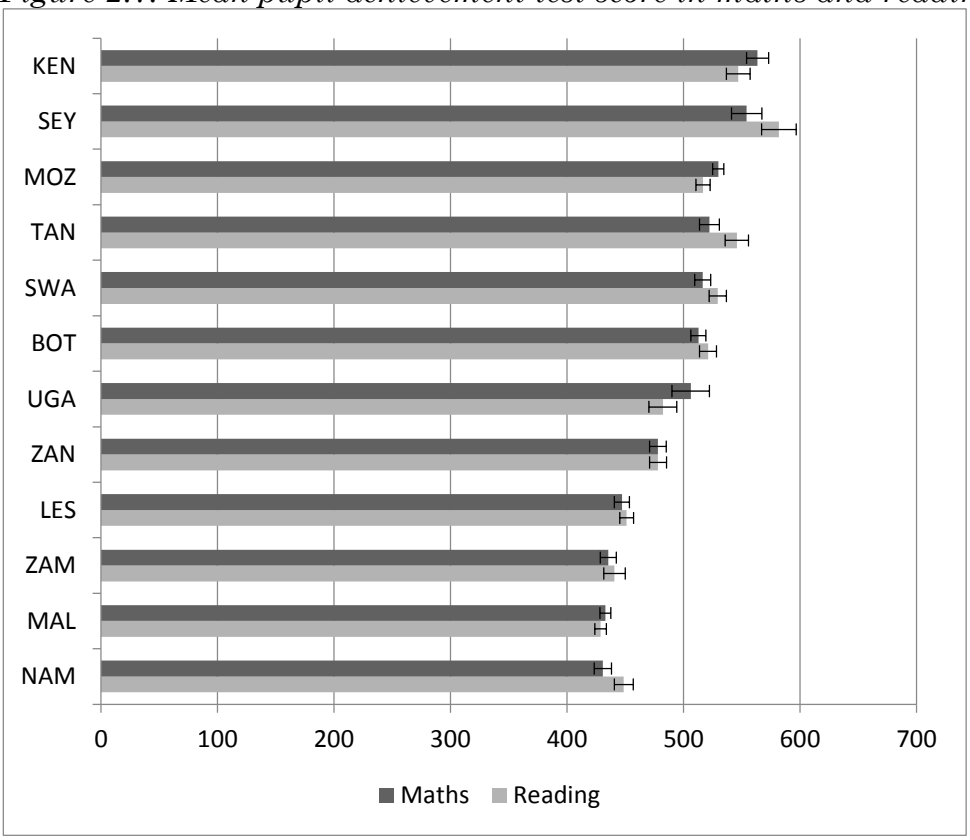


Note: Error bars indicate 95 percent confidence interval. Underlying standard errors account for clustering by school. Countries are sorted by values of Maths teachers. The reported estimates are the intercepts obtained from linear regressions for each subject in each country that account for changes over time and missing data dummies.

As already mentioned, pupils in each country are either taught by two different teachers or by the same teacher in both subjects. Table 2 therefore reports the average difference in pupil maths and reading achievement separately, for these two groups of pupils. As before the estimates are scaled in SACMEQ test points. Negative estimates indicate that pupils tend to do better in reading than in maths, positive estimates indicate that pupils tend to do better in maths. The data shows that, apart from Zambia and Zanzibar, the magnitude of this difference in achievement is closer to zero when comparing pupils with different teachers to pupils with the same teacher in both subjects. For example in Kenya, pupils with different teachers in maths and reading tend to do 17 points better in their reading than in maths, compared to an 11 point gap for pupils with the same teacher. In Lesotho and Malawi pupils with two different teachers tend to do better in one subject, but this is not the case for those pupils with the same teacher in both subjects. This suggests that the variation between teachers, i.e. teacher quality may be larger than the variation within a teacher, i.e. differential teaching ability.

In Zanzibar on the other hand, pupils with two different teachers in maths and reading tend to do equally well on both tests.

Figure 2.7: Mean pupil achievement test score in maths and reading



Note: Error bars indicate 95 percent confidence interval. Underlying standard errors account for clustering by school. Countries are sorted by pupils' performance in maths. The reported estimates are the intercepts obtained from linear regressions for each subject in each country that account for changes over time. As these variables are the dependent variables of my models, I do not mean – impute these variables.

Table 2.2: Mean difference in maths and reading achievement for pupils with the same or different teachers

Country	Different Teacher	N	Same Teacher	N
BOT	-20*** *(-4.14)*	857	-7*** *(-4.23)*	6329
KEN	17*** *(7.69)*	3749	11*** *(3.35)*	3949
LES	-22** *(-2.80)*	1507	-3 *(-1.34)*	5876
MAL	4* *(2.16)*	3424	-2 *(-0.22)*	1679
MOZ	13*** *(5.99)*	6400	11*** *(4.26e[15])*	22
NAM	-18*** *(-10.97)*	10853	-15*** *(-4.09)*	534
SEY	-28*** *(-8.72)*	2803	-19 *(-1.75)*	158
SWA	-14*** *(-5.95)*	5848	-7 *(-1.26)*	1318
TAN	-25*** *(-10.13)*	6748	-3 *(-0.26)*	281
UGA	25*** *(4.37)*	7457	8 *(0.52)*	434
ZAM	-4 *(-0.47)*	122	-6** *(-2.61)*	5307
ZAN	0 *(0.07)*	4967	-13* *(-2.29)*	280

*Note: Reported estimates are scaled in SACMEQ test score points, with a SACMEQ-wide mean of 500 and a corresponding SD of 100. T-statistic in parentheses. * = sig. at 95% C.I., ** = sig. at 99% C.I., *** = sig. at 99.9% C.I. Negative estimates imply that pupils tend to do better in reading than in maths, positive coefficients imply that pupils tend to do better at maths than reading.*

2.5. Findings

Having introduced the reader to the SACMEQ data and the kind of teachers the pupils in the sampled countries face, this section now analyses how these observable teacher characteristics correlate with the within-pupil difference in maths and reading. The main focus of this paper is to examine whether there are cross-country patterns and whether these patterns differ from US studies. In order to compare estimates across countries it is crucial that all estimates are on the same scale. Although the SACMEQ pupil and teacher test scores are on the same SACMEQ-wide scale, the variances are likely to differ across individual countries so that a one unit increase will be of different relative importance in each country. It is important to note at this point that I estimate the models reported in this section by computing the difference in pupil achievement and the individual subject-level variables manually before running equation *(6)* as a multiple linear regression. I therefore re-standardise all continuous variables in equation *(6)* to their national mean, coded as 0, and their respective national standard deviations as 1. Thus, the obtained estimates for these variables indicate the change in the dependent variable in national standard deviations, when the corresponding independent variable changes by one standard deviation. Another advantage of standardising is that it allows comparing the magnitude of the estimates obtained here with those of previous research[13]. One of these continuous variables that I standardise is the number of years of teachers' teaching experience. Although a "year" is a scale of measurement common to all education systems, it is convenient to be able to compare the relative importance of estimates within each country. I therefore also standardise the number of years of teachers' teaching experience.

At the same time, the models estimated here include dummy variables. These are not standardised. Because the dependent variable is standardised, corresponding estimates represent the change in national standard deviations of the dependent variable when the respective dummy variable is "on".

Although the model applied here examines differences within pupil, the coefficients can be interpreted in the same way as if the model would regress, for example, pupil achievement in maths to the maths teacher's test score. In order to aid readability of my findings I will report my findings in the following example for the association of teacher test scores: "teacher test scores are associated with an

[13] Alternatively, I could standardise to the entire SACMEQ sample used. I prefer to treat each country as a different sample rather than as a sub-sample of the SACMEQ region. I also find it more intuitive to think of the effect of national SD changes in X on nationally standardised Y, rather than the effect of a SACMEQ-wide SD change in X on SACMEQ-standardised Y in a given sub-sample.

increase in *'d'* standard deviations". This statement though implies that a standard deviation change in the difference between a pupil's maths and reading teacher's subject-matter test score is associated with a change in the difference between the pupil's maths and reading achievement by *'d'* standard deviations.

As described previously in section 2.2, the existing literature to date includes teacher observable characteristics in an ad hoc manner without explicitly specifying why these variables should be included and what they should proxy. In this paper I have therefore hypothesised how and why teacher observable characteristics may be related with pupil achievement in equation *(1)*. Following this equation these variables can relate to teachers' subject-matter, pedagogic competency, to incentives extrinsically motivating teachers to behave in certain ways as well as teachers' intrinsic motivation (see for example Deci and Ryan, 1985 for extrinsic and intrinsic motivation). In the absence of suitable data for incentives and teacher motivation, the following sub-sections will now concentrate on the former two kinds of inputs, teachers' competencies.

2.5.1 The association of teacher observable characteristics: Subject-matter competency

As stated in the introduction, the main contribution of this paper is applying a pupil-fixed effects approach to estimating teacher quality, which reduces the number of sources of endogenous bias compared to previous research conducted in the African context. Thus, as the estimates are purged of more confounding factors more consistent trends may appear across countries. Additionally, because not all teachers in these countries have enjoyed a duration of teacher training or have achieved a level of academic education similar to developed countries such as the US or the UK, one might expect to also observe strong effect sizes. Yet the data does not support this expectation and some findings are counterintuitive as I will discuss below.

For example, in respect to teachers' subject-matter test scores one would expect a positive association, i.e. one would assume that the better a teacher has mastery of the content he or she teaches, the better they should be able to teach. Only the findings for the Seychelles (0.05 SD) and Zanzibar (0.08 SD) support this assumption (see Table 2.3). Yet in Uganda a one standard deviation increase in maths and reading teachers' test scores is associated with a 0.06 SD decrease in the gap in pupils' maths and reading achievement, thus suggesting that teachers with better mastery of their subject appear to be worse teachers. Where pupils have the same teacher in both subjects, I estimate whether the difference in a

teacher's subject-matter competency explains variation in their differential teaching ability (Table 2.4). The data suggests that only in Lesotho is there a significant relationship; but again the coefficient is negative (-0.03 SD).

Table 2.3: OLS estimates of teacher observable characteristics and teacher quality (equation (6) for pupils with different teachers)

	Predictors (ΔX_j; ΔZ_j)	KEN	LES	MAL	MOZ	NAM
Subject-matter competency	**Teacher Test score**	0.03 *(1.52)*	-0.02 *(-0.82)*	0.01 *(0.67)*	0.01 *(1.07)*	0.01 *(1.01)*
	Graduated Primary school[1]	-0.35* *(-2.00)*	-0.02 *(-0.31)*	0.04 *(0.20)*	0.14* *(2.14)*	0.04 *(1.50)*
	Attended secondary school[1]	0.03 *(0.83)*	0.21*** *(3.21)*	0.25 *(1.44)*	-0.01 *(-0.32)*	0.01 *(0.62)*
Pedagogic competency	**1yr or more teacher training**	0.03 *(0.52)*	0.24** *(3.14)*	0.06 *(1.76)*	-0.00 *(-0.06)*	-0.12*** *(-3.84)*
	No in-service training	0.03 *(0.73)*	-0.06 *(-1.26)*	-0.08* *(-2.26)*	0.01 *(0.36)*	-0.03 *(-1.69)*
	Years of experience	-0.01 *(-0.36)*	-0.01 *(-0.59)*	0.02 *(1.31)*	-0.01 *(-0.94)*	-0.01 *(-1.24)*
Control variables	**Same gender as pupil**	0.13*** *(4.52)*	0.05 *(1.17)*	0.02 *(0.44)*	0.04 *(1.61)*	0.03* *(2.10)*
	Teacher Gender	-0.01 *(-0.46)*	-0.09 *(-1.81)*	0.16*** *(4.05)*	0.03 *(1.04)*	0.03 *(1.83)*
	# Classroom resources	0.01 *(0.24)*	-0.01 *(-0.44)*	0.02 *(1.41)*	0.00 *(0.11)*	-0.01 *(-1.05)*
	Constant	0.00 *(0.50)*	0.00 *(0.04)*	-0.06*** *(-3.09)*	-0.01 *(-0.97)*	0.00 *(0.03)*
	N_i	**3749**	**1507**	**3424**	**6400**	**10853**

Table 2.3: continued

	Predictors (ΔX_j; ΔZ_j)	SEY	SWA	TAN	UGA	ZAN
Subject-matter competency	**Teacher Test score**	0.05* *(2.50)*	0.01 *(1.03)*	0.01 *(0.62)*	-0.06*** *(-4.29)*	0.08*** *(4.70)*
	Graduated Primary school[1]	0.05 *(0.53)*	-0.08 *(-1.61)*	-0.17* *(-2.44)*	-0.11** *(-3.15)*	0.55** *(5.37)*
	Attended Secondary school[1]	-0.08 *(-1.80)*	-0.07 *(-1.47)*	-0.11 *(-1.75)*	-0.04 *(-1.73)*	0.17** *(4.28)*
Pedagogic competency	**1yr or more teacher training**	0.24*** *(3.43)*	-0.08 *(-1.64)*	-0.09** *(-3.16)*	0.13** *(2.60)*	0.03 *(0.68)*
	No in-service training	-0.09** *(-2.74)*	-0.04 *(-1.65)*	0.02 *(0.78)*	0.03 *(0.93)*	0.02 *(0.78)*
	Years of experience	0.05* *(2.31)*	0.00 *(0.08)*	0.01 *(0.52)*	0.01 *(0.63)*	-0.09*** *(-5.91)*
Control variables	**Same gender as pupil**	0.20*** *(5.32)*	0.07** *(3.08)*	0.11*** *(5.65)*	-0.04 *(-1.40)*	0.02 *(0.94)*
	Teacher Gender	-0.10* *(-2.21)*	0.06* *(2.55)*	0.05* *(2.05)*	-0.01 *(-0.18)*	-0.01 *(-0.39)*
	# Classroom resources	0.03 *(1.29)*	-0.01 *(-0.67)*	-0.01 *(-0.69)*	0.02 *(1.74)*	-0.00 *(-0.07)*
	Constant	-0.06** *(-2.92)*	-0.01 *(-0.53)*	0.02 *(0.28)*	0.03* *(2.11)*	-0.03 *(-1.85)*
	N_i	**2803**	**5848**	**6748**	**7457**	**4967**

*Note: Estimates reported indicate change in national standard deviations. Continuous variables are nationally standardised to mean zero and SD=1, thus effect sizes are comparable across variables within countries and across countries. T-statistic in parentheses. [1] compared to "Graduated Secondary school and above". * = sig. at 95% C.I., ** = sig. at 99% C.I., *** = sig. at 99.9% C.I.*

Table 2.4: OLS estimates of teacher test scores and their differential teaching ability (equation (6) for pupils with the same teacher)

Predictors (ΔX_j; ΔZ_j)	**BOT**	**KEN**	**LES**	**MAL**	**SWA**	**ZAM**
Teacher Test score	-0.02 *(-1.68)*	-0.04 *(-1.94)*	-0.03* *(-2.18)*	0.02 *(0.73)*	-0.03 *(-0.81)*	-0.01 *(-0.81)*
# Classroom resources	0.00[a] *(0.0)*	-0.04 *(-1.91)*	0.00[a] *(0.0)*	0.00[a] *(0.0)*	0.00[a] *(0.0)*	1.27** *(2.72)*
Constant	0.04*** *(3.34)*	0.01 *(0.30)*	0.02 *(1.45)*	0.07** *(2.58)*	0.06* *(2.15)*	0.02 *(1.04)*
N_i	**6329**	**3949**	**5876**	**1679**	**1318**	**5307**

*Note: Estimates reported indicate change in national standard deviations. Continuous variables are nationally standardised to mean zero and SD=1, thus effect sizes are comparable across variables within countries and across countries. T-statistic in parentheses. [a] variable is constant for all observations and is therefore automatically dropped. * = sig. at 95% C.I., ** = sig. at 99% C.I., *** = sig. at 99.9% C.I.*

The second proxy for teachers' subject-specific competency is the level of their academic education. As this variable, and all the other variables discussed below only vary between teachers, the reported findings refer to samples where pupils have two different teachers for each subject. I report estimates for pupils whose teachers graduated from primary school or whose teachers attended secondary school compared to pupils whose teachers have university-entry qualifications or above. As teachers in developed countries usually hold an undergraduate or even a postgraduate qualification, one is tempted to expect the teachers holding university entry qualifications or above should outperform their peers who do not by a large margin. Yet the data shows that the teachers holding university entry qualifications and above do not appear to outperform their peers who have attended secondary school. Also they only appear to outperform their peers who have graduated from primary school by 0.35 SD in Kenya, 0.17 SD in Tanzania and 0.11 SD in Uganda. Instead the estimates suggest that teachers with university entry qualifications might be outperformed by their academically less well-educated peers by 0.14 SD in Mozambique and by more than half a standard deviation (0.55 SD) in Zanzibar, for primary school graduates, and by 0.21 SD in Lesotho and 0.17 SD in Zanzibar for secondary school attendees. SACMEQ data does not allow delving deeper into the causes for these findings. Thus it is impossible to explore whether have certain skills, such as using a more simpler vocabulary while teaching which may aid pupils' comprehension of the content, or whether the reference category teachers differ in their motivation, for example because they are pursuing a career in the ministry and are using teaching as an entry path.

2.5.2 The association of teacher observable characteristics: Pedagogic competency

In this sub-section I shift the focus to proxies of teachers' pedagogic competency. In this study I consider teachers' pre-service and in-service training, as well as their years of teaching experience as sources of pedagogic competency. Again one could expect that teachers with at least one year of pre-service training and those with in-service training should outperform their peers that do not meet these criteria. As above there are no consistent patterns across countries and some counterintuitive findings. In respect to teachers' with at least one year of pre-service training, although findings for Lesotho and the Seychelles (0.24 SD) and Uganda (0.13 SD) support the initial expectation, estimates for Namibia (-0.12 SD) and Tanzania (-0.09 SD) indicate that the teachers who do not meet this threshold outperform their peers who do. In respect to in-service training, teachers who do not have any in-service training appear to be 0.08 and 0.09 SD less effective in Malawi and the Seychelles, which supports the initial expectation, there is no such support in the other ten countries.

In respect to teachers' teaching experience one might expect that, because not every teacher in each of these countries has been as highly educated academically, nor do all teachers enjoy the same amount of in-service and pre-service teacher training, there may be returns to teaching experience throughout the teaching career and not only for the initial years as the literature from the US suggests. Interestingly the findings here follow the pattern of the US. In ten of the twelve countries is there no statistically significant association. Only in the Seychelles do teachers appear to improve throughout their career so that on average a standard deviation increase in maths and reading teachers' experience is associated with a 0.05 SD increase in the difference in a pupil's maths and reading achievement. In contrast this association is negative (-0.09 SD) in Zanzibar, which may suggest for example that with increasing experience teachers lose their motivation or "burn out" (cf. Maslach et al, 2001). In order to test whether the statistically insignificant estimates for the ten countries may be due to the fact that, as in the US, teachers in their initial years are less effective than their more experienced peers, I replace the continuous measure of teaching experience with a dummy indicating teacher in their first 5 years of their career (table not reported). The findings suggest that these teachers are less effective by 0.1 SD in Uganda. In contrast this group of teachers outperforms their peers by 0.14 SD in Zanzibar. This counterintuitive finding may represent a cohort effect, such that these relatively inexperienced teachers have, for example, enjoyed better quality teacher training.

2.5.3 Illustrating the selection bias

The main motivation driving the choice of method is to reduce selection bias in the estimators of interest. In this section I will therefore compare my findings reported above with two models, one for each subject, that include school-fixed effects which control for all school-level variables, observable and unobservable. In other words, these models control for the non-random matching of pupils and teachers to school, but do not control for the non-random matching within schools. As each of these models refers to one subject only, each contains all pupils regardless whether they are taught by the same or by different teachers in both subjects. The results of these models are reported in Table 2.5a and 2.5b.

Table 2.5a: Within-school estimates of teacher quality for maths

Predictors (X_j)	**BOT**	**KEN**	**LES**	**MAL**	**MOZ**	**NAM**
Teacher Test score	0.01 *(0.35)*	0.04 *(1.28)*	-0.03 *(-0.65)*	-0.05 *(-0.25)*	0.02 *(0.69)*	-0.04 *(-1.03)*
Graduated Primary school[1]	-0.16* *(-1.99)*	-0.79*** *(-6.21)*	0.17 *(1.35)*	[a]	-0.26 *(-1.58)*	0.09 *(1.02)*
Attended Secondary school[1]	-0.03 *(-0.57)*	-0.03 *(-0.45)*	0.19 *(1.9)*	0.45 *(1.5)*	0.03 *(0.31)*	0.01 *(0.11)*
1yr or more teacher training	0.14 *(1.77)*	-0.10 *(-0.96)*	0.33*** *(4.0)*	-0.11 *(-0.98)*	0.09 *(1.52)*	-0.21 *(-1.84)*
No in-service training	0.06 *(1.14)*	-0.08 *(-0.89)*	-0.23*** *(-3.26)*	-0.39 *(-1.36)*	-0.01 *(-0.27)*	0.11 *(1.82)*
Years of experience	-0.08* *(-2.44)*	-0.39*** *(-6.98)*	-0.25*** *(-4.36)*	-0.24*** *(-3.04)*	-0.03 *(-0.89)*	-0.22*** *(-5.96)*
Constant	-1.02*** *(-6.82)*	-1.88** *(-2.92)*	-1.23*** *(-3.62)*	-0.43 *(-0.86)*	-0.36* *(-2.48)*	-0.89*** *(-5.11)*
N	**7186**	**7698**	**7383**	**5103**	**6422**	**11387**

Table 2.5a continued

Predictors (X_j)	SEY	SWA	TAN	UGA	ZAM	ZAN
Teacher Test score	-0.02 *(-0.25)*	0.19 *(1.12)*	0.04 *(0.79)*	-0.07 *(-0.87)*	-0.02 *(-0.71)*	0.11* *(2.23)*
Graduated Primary school[1]	-1.13*** *(-4.49)*	[a]	0.73* *(2.13)*	0.52 *(1.77)*	-0.04 *(-0.44)*	0.00 *(0.01)*
Attended Secondary school[1]	-0.55* *(-2.42)*	0.89*** *(8.10)*	0.31 *(1.47)*	0.38 *(1.61)*	-0.02 *(-0.33)*	-0.13 *(-1.21)*
1yr or more teacher training	0.39 *(1.72)*	[a]	-0.02 *(-0.20)*	-0.39 *(-0.58)*	-0.05 *(-0.62)*	-0.13 *(-1.04)*
No in-service training	0.18 *(0.87)*	0.03 *(0.18)*	0.23* *(2.02)*	-0.29 *(-0.87)*	0.06 *(0.99)*	-0.08 *(-0.96)*
Years of experience	0.15 *(0.22)*	-0.22 *(-1.28)*	0.18* *(2.32)*	-0.36*** *(-3.73)*	-0.09 *(-1.71)*	0.00 *(-0.01)*
Constant	-3.75* *(-2.11)*	-1.43* *(-2.23)*	-0.82 *(-1.68)*	0.65 *(0.94)*	0.04 *(0.22)*	0.11 *(0.40)*
N	**2961**	**7166**	**7029**	**7891**	**5429**	**5247**

*Note: Estimates reported indicate change in national standard deviations. Continuous variables are nationally standardised to mean zero and SD=1, thus effect sizes are comparable across variables within countries and across countries. T-statistic in parentheses. [1] compared to "Graduated Secondary school and above". * = sig. at 95% C.I., ** = sig. at 99% C.I., *** = sig. at 99.9% C.I. Estimates conditional on the number of days the pupil was absent in the month prior to testing, whether the pupil has repeated a grade, pupil SES, Teacher gender, whether the teacher is the same gender as the pupil, the number of classroom resources and the survey wave in which the pupil was observed. [a] variable is constant for all observations and is therefore automatically dropped.*

Table 2.5b: Within-school estimates of teacher quality for reading

Predictors (X_j)	BOT	KEN	LES	MAL	MOZ	NAM
Teacher Test score	0.02 *(1.03)*	0.08* *(2.00)*	0.04 *(0.78)*	0.09 *(0.97)*	-0.02 *(-0.65)*	0.04 *(1.20)*
Graduated Primary school[1]	-0.09 *(-1.31)*	-0.05 *(-0.28)*	-0.08 *(-0.71)*	-0.58 *(-1.41)*	0.02 *(0.14)*	-0.08 *(-0.49)*
Attended Secondary school[1]	-0.04 *(-0.94)*	0.04 *(0.58)*	-0.24* *(-2.11)*	-0.52* *(-2.16)*	0.10 *(1.36)*	-0.08 *(-1.62)*
1yr or more teacher training	0.09 *(1.06)*	-0.62*** *(-3.86)*	-0.11 *(-0.6)*	-0.01 *(-0.05)*	-0.14** *(-2.9)*	-0.09 *(-0.74)*
No in-service training	0.03 *(0.77)*	-0.06 *(-0.83)*	-0.15 *(-1.48)*	-0.29 *(-1.24)*	0.03 *(0.63)*	-0.07 *(-1.05)*
Years of experience	-0.09** *(-3.05)*	-0.40*** *(-4.91)*	-0.30*** *(-4.18)*	-0.21 *(-1.38)*	-0.17*** *(-4.39)*	-0.19*** *(-4.35)*
Constant	-1.08*** *(-6.02)*	-2.03* *(-2.62)*	-0.75 *(-1.68)*	-0.13 *(-0.19)*	-0.43*** *(-3.28)*	-0.69** *(-3.08)*
N	**7190**	**7704**	**7395**	**5114**	**6480**	**11446**

Table 2.5b continued

Predictors (X_j)	SEY	SWA	TAN	UGA	ZAM	ZAN
Teacher Test score	0.02 *(0.25)*	-0.46*** *(-3.56)*	0.04 *(0.70)*	0.02 *(0.95)*	-0.01 *(-0.43)*	-0.01 *(-0.27)*
Graduated Primary school[1]	-0.52 *(-1.45)*	[a]	0.61 *(1.43)*	-0.27* *(-2.42)*	-0.13 *(-1.50)*	0.43* *(2.16)*
Attended Secondary school[1]	-0.05 *(-0.34)*	0.90*** *(3.58)*	0.47 *(1.26)*	-0.04 *(-0.47)*	-0.10 *(-1.26)*	0.23* *(2.29)*
1yr or more teacher training	0.18 *(0.48)*	[a]	0.19 *(1.43)*	-0.04 *(-0.27)*	0.03 *(0.37)*	0.14 *(1.25)*
No in-service training	0.24 *(1.61)*	0.91** *(2.69)*	0.04 *(0.42)*	0.07 *(0.70)*	0.17* *(2.25)*	-0.01 *(-0.17)*
Years of experience	-0.10 *(-0.37)*	-1.04*** *(-6.67)*	0.01 *(0.10)*	-0.20*** *(-4.16)*	-0.11 *(-1.86)*	-0.13 *(-1.77)*
Constant	-2.90*** *(-3.63)*	-2.96*** *(-5.48)*	-1.18** *(-2.88)*	0.03 *(0.04)*	-0.59** *(-2.92)*	-0.61** *(-2.60)*
N	**2964**	**7169**	**7035**	**7920**	**5463**	**5305**

*Note: Estimates reported indicate change in national standard deviations. Continuous variables are nationally standardised to mean zero and SD=1, thus effect sizes are comparable across variables within countries and across countries. T-statistic in parentheses. [1] compared to "Graduated Secondary school and above". * = sig. at 95% C.I., ** = sig. at 99% C.I., *** = sig. at 99.9% C.I. Estimates conditional on the number of days the pupil was absent in the month prior to testing, whether the pupil has repeated a grade, pupil SES, Teacher gender, whether the teacher is the same gender as the pupil, the number of classroom resources and the survey wave in which the pupil was observed. [a] variable is constant for all observations and is therefore automatically dropped.*

If the pupil-teacher allocation were random within school, the estimates from both the school-fixed and the pupil-fixed effects models should be the same. Yet at first glance it becomes clear that this is not the case, as a completely different picture emerges. This finding strongly supports the hypothesis that previous research such as by Fehrler et al (2008) or Zuze (2010) does not sufficiently address the pupil-teacher matching problem. Zuze (2010) applies a multilevel approach, as do Fehrler et al (2008), who also apply a country-fixed effects specification. Although the latter include many covariates for the pupil, teacher and school level, the results shown in Tables 5a and 5b suggest that the allocation of pupils to teachers contains a substantial unobservable part. Furthermore, comparing the school-fixed effects models to the pupil-fixed effects models reported above strongly supports the assumption that even within schools, the pupil-teacher matching process is not random and ignoring this will lead to completely different conclusions if such findings were to be used for the basis of education policies.

Apart from differing from the pupil-fixed effects models, the estimates of the school-fixed effects models differ by subject. This pattern supports previous findings from the US (cf. Nye et al, 2004; Rockoff, 2004; Hanushek et al, 2005; Rivkin et al, 2005; Kane et al, 2006; Aaronson et al, 2007; Clotfelter et al, 2007;

Slater et al, 2011). The estimates here though suggest that years of teaching experience are the most consistent predictor for teacher quality; the coefficients are statistically significant at or below the five percent level in 7 countries for both subjects – but these are not necessarily the same countries in both models. The effect sizes in the school-fixed effect models are also surprisingly large. In Swaziland for example, a standard deviation increase in teaching experience is associated with 1.04 SD decline in pupil reading achievement.

2.5.4 Interactions of teachers' subject-matter and pedagogic competencies

I now return to the pupil-fixed effects models. Having grouped the available teacher characteristics into proxies of either teachers' subject-matter or pedagogic competency (see equation *(1)*), I now explore the interaction of these two kinds of competencies. I add a vector of interactions. Specifically these are interactions of the test score, as the proxy of subject-matter competency, with each of the three variables representing pedagogic competency discussed above (one year or more pre-service training, no in-service training, years of experience). A statistically significant and positive finding indicates complementarity, as combining these two competencies maximises pupil achievement. In contrast, a statistically significant negative estimates indicates the two competencies are substitutes, as trading off one competency in favour of the other maximises pupil achievement. Arguably, teaching involves combining both competencies, so that policy should strive to avoid these from being substitutes.

The findings of these interactions are reported in the lower half of Table 2.6, the corresponding direct effects are reported in the top half of said table. Table 2.6 reports estimates for the ten countries that have sufficient data of pupils taught by different teachers for maths and reading. Only these samples can be used, because when pupils have the same teacher for both subjects, the only variable that differs between subjects is the teacher's respective subject-matter test score. Comparing the estimates for the interaction of teacher test scores and their years of experience, in 4 of these ten countries the combination of these two competencies does not appear to affect pupil achievement. In contrast in Mozambique and Swaziland (both 0.03 SD) these two competencies appear to fit together well, whereas in Malawi, the Seychelles and Uganda (each -0.04 SD) and Namibia (-0.02 SD) there appears to be a bad fit. In respect to the interaction of teacher test scores and having at least one year of pre-service training, findings suggest that these two competencies fit well in Kenya (0.16 SD), but in Mozambique and Tanzania this

interaction is associated with a reduction in pupil achievement by 0.08 and 0.11 SD respectively.

Table 2.6: Interaction effects of Teacher Test score with proxies of teachers' pedagogic competency (for pupils with different teachers)

		KEN	LES	MAL	MOZ	NAM
Direct effect	**Teacher Test Score**	0.04*	-0.01	0.02	0.00	0.01
		(1.97)	*(-0.48)*	*(0.99)*	*(0.32)*	*(0.83)*
	1yr or more Teacher Training	-0.06	0.23**	0.06	-0.02	-0.12***
		(-0.70)	*(3.07)*	*(1.68)*	*(-0.76)*	*(-3.83)*
	No In-service training	0.03	-0.05	-0.08*	0.01	-0.03
		(0.79)	*(-1.04)*	*(-2.48)*	*(0.42)*	*(-1.61)*
	Years of teaching experience	0.00	-0.01	0.00	-0.01	-0.01
		(0.22)	*(-0.70)*	*(0.23)*	*(-0.69)*	*(-1.24)*
Interaction effect	**Test Score*Ex-perience**	-0.03	-0.01	-0.04***	0.03*	-0.02*
		(-1.67)	*(-0.34)*	*(-3.50)*	*(2.46)*	*(-2.37)*
	Test Score*1yr or more training	0.16*	0.10	0.06	-0.08***	-0.02
		(2.36)	*(1.41)*	*(1.73)*	*(-3.26)*	*(-0.68)*
	Test Score*No In-service training	0.01	-0.04	0.12***	-0.06**	0.03
		(0.21)	*(-0.98)*	*(3.99)*	*(-2.58)*	*(1.88)*
	N_i	**3749**	**1507**	**3424**	**6400**	**10853**

*Note: Estimates reported indicate change in national standard deviations. Continuous variables are nationally standardised to mean zero and SD=1, thus effect sizes are comparable across variables within countries and across countries. T-statistic in parentheses; * = sig. at 95% C.I., ** = sig. at 99% C.I., *** = sig. at 99.9% C.I.*

Table 2.6 continued

		SEY	SWA	TAN	UGA	ZAN
Direct effect	**Teacher Test Score**	0.05*	0.02	0.01	-0.05***	0.08***
		(2.50)	*(1.23)*	*(0.40)*	*(-4.24)*	*(4.62)*
	1yr or more Teacher Training	0.21**	-0.08	-0.10**	0.14**	0.03
		(2.91)	*(-1.78)*	*(-3.26)*	*(2.73)*	*(0.76)*
	No In-service training	-0.09**	-0.04	0.02	0.03	0.02
		(-2.93)	*(-1.88)*	*(0.67)*	*(1.18)*	*(0.67)*
	Years of teaching experience	0.06*	0.00	0.01	0.02	-0.09***
		(2.38)	*(0.03)*	*(0.63)*	*(1.38)*	*(-5.63)*
Interaction effect	**Test Score*Experience**	-0.04*	0.03*	0.01	-0.04***	0.01
		(-2.28)	*(2.29)*	*(0.76)*	*(-3.22)*	*(0.54)*
	Test Score*1yr or more training	0.02	-0.00	-0.11***	-0.04	0.03
		(0.21)	*(-0.14)*	*(-3.63)*	*(-1.43)*	*(0.74)*
	Test Score*No In-service training	0.00	0.02	-0.03	-0.04	-0.05*
		(0.02)	*(0.73)*	*(-1.13)*	*(-1.60)*	*(-1.98)*
	N_i	**2803**	**5848**	**6748**	**7457**	**4967**

*Note: Estimates reported indicate change in national standard deviations. Continuous variables are nationally standardised to mean zero and SD=1, thus effect sizes are comparable across variables within countries and across countries. T-statistic in parentheses. * = sig. at 95% C.I., ** = sig. at 99% C.I., *** = sig. at 99.9% C.I. Apart from the interactions included in this model, the same variables are included as reported in table 4, but these are omitted here to aid readability.*

In respect to the interaction of teacher test scores and not having any in-service training, one would expect that if these two competencies were complimentary to one another the obtained coefficient should be negative. This is the case in 5 countries, but the effects are only statistically significant at $p \leq 0.05$ in Mozambique and Zanzibar (-0.06 SD and -0.05 SD). In contrast there appears to be a bad fit of these two competencies in Malawi (0.12 SD).

2.5.5 Summary

In summary, the findings challenge the expectation that teacher effectiveness differs due to the same variables consistently across these sampled countries. This pattern mirrors the US findings. On the other hand, each country appears to have its own combination of crucial variables, which underlines the importance of each country's respective context. This supports the notion the US findings are not suitable policy guidance. A number of individual estimates themselves are counterintuitive. For example returns to teachers' subject-matter competency being negative on average rather than positive or not statistically different from zero, or that teachers with primary school education appear to outperform their peers with university entry qualifications or above by more than half a standard deviation, as in Zanzibar, are surprising. These findings warrant further investigation, which the scope of this paper as well as the SACMEQ data itself do not allow. Finally, the findings suggest that in six countries teachers' subject-matter and pedagogic competency lacks compatibility.

2.6. Conclusion

This paper examines teacher quality for twelve sub-Saharan Southern and Eastern African countries. In contrast to previous studies around the globe that tend to examine the conditional correlation of observable teacher characteristics and pupil achievement in an ad hoc manner, i.e. they do not specify how and why certain variables should affect pupil achievement, I classify teacher observable characteristics as proxies of either subject-matter competency, pedagogic competency or incentives. In this study I focus on the conditional correlation of teacher tests scores and teachers' academic education as proxies for subject-matter competency, and teachers' pre and in-service training as well as years of experience as proxies of their pedagogic competency.

The second and main contribution of this paper is the estimation method applied. Previous studies exploring teacher quality in the African context apply standard multivariate OLS or multilevel models with observational data, some of which use

the same data as used in this study. These models cannot adequately address the non-random matching of pupils to teachers. Neither can they observe all potential confounders, so that their estimates are subject to endogeneity. In this paper I exploit the fact that pupils are tested in reading and maths at the same point in time, the scores on each of these tests are on the same scale, and that the respective teachers in each subject are also observed. I follow Clotfelter et al (2007) who argue that pupil-fixed effects sufficiently address the non-random matching problem, thus in the cross-sectional SACMEQ data I associate the difference in a pupil's maths and reading teacher's observable characteristics with the difference in a pupil's maths and reading achievement score. In essence, this method does not explore the absolute importance of teachers for pupil achievement. It rather explores whether observable characteristics, or as conceived here, different competencies explain the variation in teacher quality. The advantage of this method is that all potential confounders that do not vary within pupil are eliminated, these being the school level and many pupil and household variables. As time does not play a role in this approach, I combine two time-points of SACMEQ data, the first collected in the year 2000/2, the second in 2007, which effectively doubles the available sample size thereby increasing statistical precision.

In the SACMEQ countries used here pupils in each country are either predominantly taught by the same or by different teachers in both subjects. I refer to estimating *teacher quality* when pupils are taught by different teachers. In the case when pupils are taught by the same teacher in both subjects, most variables are the same for both subjects, hence the difference between them is zero. In SACMEQ, as in other developing countries, teachers also sit an achievement test in the subject they teach. If a teacher teaches both subjects, he or she is tested in both subjects. Thus I can estimate teachers' *differential teaching ability* as a function of the difference in their subject-matter competency.

As my estimates in this study are purged of more confounding factors than any previous research using African data, more consistent trends may appear across countries. Additionally, because not all teachers in these countries have enjoyed teacher training or have achieved a level of academic education similar to developed countries, one might expect to also observe strong effect sizes. Interestingly the findings do not support this expectation. Instead the findings support the pattern emerging from the US, that there is no consistent pattern of teacher characteristics associated with pupil achievement.

I compare the findings from my pupil-fixed effects models with two school-fixed effects models, one of reach subject, to demonstrate the effect of not sufficiently addressing the non-random allocation of pupils to teachers. Comparing these two types of models shows that even within schools, pupil-teacher matching is not random. Lacking to address this sufficiently will lead to completely different conclusions, especially should these estimates be used as a basis for policy making.

Moreover a number of findings are counterintuitive. For example, in Zanzibar teachers who have only graduated from primary school outperform teachers holding 'university entry qualification or above' by more than half a standard deviation. Also, in the same country, there appears to be a negative association of years of teaching experience and pupil achievement (-0.09 SD), which may suggest that more experienced teachers become less motivated or might be 'burning out'. In respect to teachers' differential teaching ability, the estimates here suggest that only in one of the six countries, where a sufficient number of pupils has the same teacher for both subjects, does the difference in the teachers' subject-matter test score explain the difference in pupil achievement in maths and reading. This country is Lesotho, where again there is a negative (-0.03 SD) rather than a positive association, which one might expect.

As it is reasonable to assume that teachers need to combine both their subject-matter and pedagogic competency I then move on to examine the compatibility of the two kinds of competencies. I do so by adding interactions of the test score, as the proxy of subject-matter competency, with each of the following proxies of pedagogic competency:

1. one year or more pre-service training,
2. no in-service training,
3. Years of experience.

From the perspective of complementarity, positive estimates indicate the competencies are complimentary, whereas negative estimates indicate that the two competencies are substitutes, i.e. that maximising pupil achievement requires a trade-off of one competency for the other. Again in most cases the estimates suggest a combination of these two competencies has a no statistically significant effect on pupil achievement. But in some cases combining these two competencies appears to have negative effects on pupil achievement. The interaction of teacher test scores and years of experience is associated with a decrease in pupil

achievement by 0.04 SD in Malawi, the Seychelles and Uganda, and by 0.02 SD in Namibia. The interaction of teacher test scores and pre-service training is associated with a decrease in pupil achievement by 0.08 SD in Malawi and 0.11 SD in Tanzania. Again in Malawi the interaction of teacher test scores and teachers who do not have received any in-service training is associated with an increase in pupil achievement by 0.12 SD and thus suggesting that the skills teachers acquire in in-service training hinder their effectiveness the better they master the subject-matter of the subject they teach.

Although most of the findings in this paper are not statistically significant, one should not conclude that teachers do not matter. Instead this paper adds additional weight to the evidence from the developed countries such as the US that teachers vary in their quality primarily due to their unobservable characteristics, and this appears to also be the case in sub-Saharan Africa.

Chapter 3

What is the connection between teachers' workload and their job satisfaction? Evidence from 32 secondary education systems

3.1. Introduction

Growing evidence from around the world shows us that teachers play a key role in explaining pupils' achievement (cf. Hein and Allen, 2013; Slater et al, 2011; Clotfelter et al, 2007; Aaronson et al, 2007; Kane et al, 2006; Rivkin et al, 2005; Nye et al, 2004). This paper takes one step back from explaining teacher effectiveness and investigates the production of teachers' wellbeing at work.

According to Kahneman and Krueger (2006) job satisfaction measures the quality of an individual's perceived experience at work and is therefore a suitable proxy for wellbeing at work. Evidence from across the social sciences demonstrates that job satisfaction is important for the efficiency of organisations. For example, job satisfaction is associated with burnout (cf. Prosser et al, 1997; Kalliath and Morris, 2002; Piko, 2006). To psychologists burnout is a symptom characterised by individuals feeling overwhelmingly exhausted, detached from their job, cynical and ineffective (Maslach and Schaufeli, 2001). Burnt out individuals are therefore likely to be less productive in their jobs, for example due to extended periods of absence, and run the risk of incurring social costs in the form of health care expenses (Faragher et al, 2005). Fischer and Sousa-Poza (2009) find that job satisfaction predicts both subjective and objective measures of health and others find that job satisfaction predicts both workers' intentions to leave their employer and actual turnover rates (for example see Hellman, 1997; Lambert et al, 2001; Sousa-Poza and Sousa-Poza, 2007).

This paper examines the association of teachers' workload, proxied by the amount of hours teachers work per week, and their job satisfaction. But why should teachers' workloads affect their job satisfaction, when classical microeconomic theory leads one to believe that individuals are maximising their utility by perfectly choosing the amount of time they work? First, it is likely that individuals in real life face contractual constraints that require them to work a certain amount of hours or only financially reward a certain amount of hours worked. Contracts may also specify how much time individuals need to spend on specific tasks while working, which may differ from the individual's choice should these constraints not exist. Secondly Kahneman and Krueger (2006) emphasize that job satisfaction should not be mistaken for a measure of utility gained from work, but rather as a measure of the quality of the experience an individual has when working in their job.

Surprisingly, there is no evidence on the nature of the workload-job satisfaction relationship for teachers. Thus not only is the strength of the association unknown, but also whether or not the association is linear or even monotonic. The only known relevant research consists of Judge and Watanabe (1993) and Sousa-Poza and Sousa-Poza (2000). Both studies use labour force data and therefore may contain teachers, but cannot focus exclusively on teachers. Both studies test a linear relationship and find no statistically significant relationship.

Yet their indicative quality for the teaching professions across the globe must be questioned for a number of reasons. First of all both studies use samples of the workforce containing a wide range of professions, so that their averaged estimates may be very different from that of the teaching profession at that time. The correlations may also change over time and individuals may gain satisfaction from certain tasks they perform while working, but experience a reduction from performing others. For example, teachers' job satisfaction may rise or fall differently by the amount of time he or she spends planning lessons, actually teaching, marking, being involved with school leadership, doing administrative chores, counselling pupils or organising and running extracurricular activities for them. In short, very little can be said about how workload affects job satisfaction of contemporary teachers, and this paper fills this gap and reports findings for 32 high and middle-income countries.

From the perspective of statistically establishing causal relationships an ideal scenario would randomise the amount of hours teachers work, either 'naturally' due to a policy, or due to an intervention. Such data, especially internationally comparable data, is not available, so that in the existing observational data, such as the data used here, the amount of hours teachers work will be the result of an unknown process. Statistically, this means that bivariate OLS estimates of the workload-job satisfaction relationship are likely to be subject to *endogeneity*, as unobserved variables may not only be correlated with teachers' workload, but also with their job satisfaction. The magnitude and direction of this potential bias is unknown and requires empirical estimation (cf. Wooldridge, 2002).

This paper addresses this potential endogeneity problem theoretically and empirically. It first develops a production function for teachers' job satisfaction based on the existing literature. This production function helps to identify which the unobserved confounders in the bivariate OLS regressions may be. Then the

paper uses data collected in 2013 for the OECD's Teaching and Learning International Survey (TALIS), a cross-sectional survey of samples of colleagues in schools, collected in 32 national and sub-national education systems. TALIS 2013 also surveys the head teachers and therefore comprises data on characteristics of the school, as well as demographics, attitudes, beliefs and behaviours of both teachers and their head teachers. The paper exploits the richness of the TALIS 2013 data, which allows including a range of potential confounding factors at both the school and individual level. The preferred model exploits the nested structure of the data with teachers nested in schools and removes all potential school-level confounders, whether observed or unobserved, by introducing school-fixed effects while controlling for potential observed individual-level confounders.

The main aim of this paper is to uncover cross-country patterns among estimates of the workload-job satisfaction relationship. The findings show that across the different kinds of workload examined here, the majority of the 32 individual education systems do not have statistically significant estimates. When the estimates are significant, at or below the 5 percent level, they tend to be of negligible magnitude. Although a one-hour change is a small change in teachers' workload in most cases, even a one standard deviation change in total weekly workload for example is associated at maximum with an 8 percent SD decrease in job satisfaction. The findings also support the expectation that teachers may gain job satisfaction from some activities and experience a reduction from others.

In the light of all the heterogeneity and statistically insignificant estimates, I run pooled models in order to obtain estimates averaged across the 32 education systems whilst accounting for between-country differences. Evidence from these models supports the assumption that bivariate estimates of the workload-job satisfaction relationships proxy school and teacher-level variables. Here the observable school-level variables appear to be associated positively with job satisfaction; so are also the observable teacher-level variables, but the unobservable school-level variables appear to be associated negatively with job satisfaction. The estimates from the pooled models also indicate that a one hour increase in school leadership activities increases job satisfaction, could be approximately counterbalanced by a one hour increase in marking (-0.5 percent SD) and administrative workload (-0.6 percent SD) and could outweigh a two-hour increase in face-to-face teaching workload, which also reduces job satisfaction.

The paper then exploits additional data collected in England regarding the teachers' perceptions of different stakeholders of their school, as well as their perception of their education systems' national context, such as social prestige, accountability system, and fairness of salary and career opportunities. As argued above, bivariate estimates of the workload-job satisfaction relationship may also proxy these perceptions. The findings support this assumption. At the same time, there is no evidence for the teaching profession regarding the extent to which these perceptions are associated with teachers' job satisfaction. This paper fills this gap by presenting these findings.

The remainder of the paper is structured as follows: The next section outlines the production function framework applied and which factors may be potential sources of endogeneity for estimates of the workload-job satisfaction relationship. The third section introduces the reader to the TALIS 2013 data. The fourth section describes the econometric method applied to deal with the potential sources of endogeneity, before the fifth section explores the association of hours worked both in general as well as on different tasks and teachers' job satisfaction among the 32 sampled education systems. The sixth section then focuses on England and the additional variables available, explores how the estimates of the workload-job satisfaction relationship change when controlling for teachers' perceptions of the stakeholders in their school and of features of the education system, before presenting and discussing the estimated associations for these perception variables with job satisfaction. The seventh section concludes.

3.2. Identifying sources of endogeneity

3.2.1. A production function for teachers' job satisfaction

In order to analyse the workload-job satisfaction relationship, this paper applies a production function approach to teachers' job satisfaction, which can be summarised as

$$S_{ij} = f(X_i, P_j, H_j, T_j, C_j, R_j) \qquad (1),$$

where the job satisfaction S of an individual i in school j is a function of the individual teacher X, the pupils P, their parents H, the head teacher T, the colleagues C and the equipment available in the building R.

The first advantage of equation *(1)* is that it combines the two predominant kinds of approaches to analysing subjective wellbeing, top-down and bottom-up (cf. Diener, 1984). Top-down approaches assume that differences in subjective wellbeing are due to individuals perceiving their environment differently, because of differences in their personality, goals and aspirations. In equation *(1)* these are elements of X_i.

Bottom-up approaches on the other hand focus on the environment an individual is in, assuming that "[…] there are basic and universal human needs, and that if one's circumstances allow a person to fulfil these needs, he or she will be happy" (Diener et al, 1999, p.278). In other words, bottom-up approaches model variation in satisfaction as a function of the conditions an individual is living in. Early bottom-up approaches in the US in the 1970s and 1980s use data on the general public to explore the association of life events, such as marriage or education, and demographic characteristics such as age, gender, ethnicity, with overall life satisfaction and find that at most 20 percent of the variation can be explained by these variables (Diener et al, 1999). Other research specifically examines individuals' job satisfaction and therefore explains job satisfaction as a function of characteristics of the workplace an individual is in. Hulin (1985) assumes that job satisfaction results in the balance between an individual's inputs into the job, i.e "pains", and the pay-offs, i.e. "pleasures", he or she receives. In other words an increase in a "pleasure" should, ceteris paribus, lead to an increase in job satisfaction, and vice-versa for the "pains". Evidence suggests that "pleasures" go beyond pecuniary rewards and also are emotional such as self-efficacy[14], or the feeling to do good to society and being in a job that has social prestige and career prospects (cf. Sousa-Poza and Sousa-Poza, 2000; Caprara et al, 2006; Skaalvik and Skaalvik, 2010). Other bottom-up approaches use workforce samples to examine the correlation of an individual's perceived working conditions. Notably autonomy, leadership, cooperation among co-workers appear to explain an individual's job satisfaction (cf. Judge and Watanabe, 1993; Perie and Baker, 1997; Ma and MacMillan, 1999; Sousa-Poza and Sousa-Poza, 2000; Bogler, 2001; Kim, 2002; Roelen et al, 2008; Kalisch et al, 2011; Collie et al, 2012 for pupil behaviour on teachers' job satisfaction). Again in equation *(1)* these are elements of X_i.

[14] Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments (Bandura, 1997, p.3).

Equation *(1)* also accounts for school-level factors indicated by subscript *j*. Michaelowa (2002) explores the extent to which class size and resources such as schoolbooks and teacher training influence teacher job satisfaction in French speaking Saharan and Sub-Saharan African countries. She argues in favour of increasing the availability of teaching equipment such as schoolbooks as these are associated with the strongest positive effects and are much cheaper than increasing teacher salaries or reducing class size. In contrast, the OECD TALIS 2013 official report (OECD, 2014) explores job satisfaction of school leaders. The authors find that the only two variables that predict head teachers' job satisfaction in nearly all 32 countries are mutual respect as a dimension of school climate, and a high workload and responsibility[15].

Yet although bottom-up approaches lead one to believe that an individual's workplace substantially affects his or her job satisfaction, the evidence regarding the proportion of an individual's job satisfaction associated with his or her workplace is limited to the teaching profession based on data from the 2008 sweep of TALIS (OECD, 2009). In their report, the OECD conducts an analysis of variance of a single-item indicator of job satisfaction on a four-point Likert scale ranging from "strongly agree" to "strongly disagree" pooling all available countries. They find that 4 percent of the variation lies between countries and 6 percent of the variation lies between schools, leaving 90 percent of the variation within schools.

3.2.2. Potential sources of endogeneity

The underlying principle of microeconomic theory is that individuals need to make choices. In contrast to factory-floor workers, teachers, like many other white-collar professionals nowadays, are not bound by strict working-time agreements with their employers. Thus the amount of time the teachers work will contain a substantial degree of choice; teachers will of course be bound to teach a specific amount of lessons due to contractual obligations, but teachers may choose how much time they invest in planning and marking lessons for example. Teachers may also vary the amount of time they work if they are incentivised[16] to do so. Equation (1) suggests that teachers may face such incentives from the other stakeholders in their school or from factors within the individual him or herself. Statistically,

[15] The authors separately test head teacher demographics, school characteristics, dimensions of school climate and perceived barriers to head teachers' effectiveness. Thus the authors assume that these groups of variables are not correlated.

[16] Offered a pecuniary or non-pecuniary reward to motivate a change in an individual's behaviour.

bivariate estimates of the workload-job satisfaction relationship will be subject to endogeneity if these incentives affect teachers' job satisfaction directly in addition to their workload. In the following paragraphs I outline which factors may be both such incentives and sources for endogeneity:

At the school level, for example, poor pupil behaviour may reduce teachers' job satisfaction and also affect their workload as they may need to spend more time carefully planning lessons for classes prone to disruption. Also, "good" head teachers may motivate their staff to put in more hours, whereas "demanding" head teachers may oblige teachers to work longer. Simultaneously, teachers may experience an increase in their job satisfaction in the former case, but experience reduced job satisfaction from being pressured in the latter case.

Similarly to the example of the head teacher above, there are a number of other factors that may affect teachers' workloads, but where it is unclear at the outset in which direction such factors directly affect teachers' job satisfaction. 'Demanding' parents may also require teachers to work longer in order to accommodate the time for parent-teacher meetings, which may directly affect teachers' job satisfaction negatively. On the other hand, strong parental involvement in school decision-making may relieve teachers from additional work. Teachers may directly gain job satisfaction from this if they appreciate the parents' involvement, but the opposite may also be the case.

Teachers may take longer planning lessons with alternative resources if the preferred resources are not available. Not having their preferred resources at their disposal may frustrate and directly reduce their job satisfaction. Also teachers lacking the skills they need to perform their job might take longer to prepare lessons or mark pupils' work. This could make teachers feel overwhelmed and reduce their job satisfaction directly. Further, the extent to which teachers cooperate with their colleagues may play a role: for example teachers may need to spend less time preparing lessons if they share lesson plans among each other, but more staff or team meetings may imply longer working hours. If teachers feel teamwork is a burden, they may experience a direct reduction in their job satisfaction.

In contrast, more senior teachers may experience more respect from their colleagues and part-timers may experience less stress, both of which may directly affect their job satisfaction positively in addition to affecting their workload.

Similarly teachers with higher levels of intrinsic motivation, i.e. they teach for enjoyment rather than for a separable reward are likely to trade-off more leisure time in favour of working and may also have higher levels of job satisfaction (cf. Deci and Ryan, 2000).

As these potential sources of endogeneity may bias the bivariate estimates in different directions, empirical estimation is required to establish its magnitude and direction. In order to reduce this bias it is important to control for as many of these potential confounders, both at the teacher and school level, as possible. But before explaining the estimation strategy, I will introduce the reader to the data used in this research.

3.3. The TALIS 2013 data

3.3.1. Introduction

The OECD TALIS datasets are repeated cross-sections of national and sub-national education systems. The 2013 round is the second after the first in 2008, and contains samples of 33 education systems in total. The OECD requires response rates of teachers and schools both to exceed 75 percent per country in order to be included in their official report (OECD, 2014); of the 33 sampled education systems in 2013, only the US sample does not meet this requirement. This paper follows the OECD's standard and therefore also does not use the US data. The response rates are high in TALIS compared to surveys such as the Teachers' Workload Diary Survey in England that has response rates around 15 percent (cf. DfE UK, 2014). The average response rate across all 32 sampled education systems is 92 percent for schools and 90 percent for teachers.

As hinted above, TALIS contains representative samples of schools and multiple teachers within them. The participating countries were instructed by the OECD to sample 200 schools and a minimum of 20 teachers per school. Some participating governments oversampled the number of teachers; for some education systems the sampling strategy was adapted to account for high numbers of schools with less than 20 teachers. A school was included in the data if at least half of the sampled teachers responded. On average there are 16 responding teachers per school. Alberta, Chile, Denmark and Iceland form outliers where on average approximately 9 to 11 teachers per school are sampled. Table 3.1 also shows that the achieved sample sizes of teachers across all 32 education systems used here is 103,862 which is equivalent to an average sample size of 3425 teachers in each

education system. Analogously, 6462 schools are sampled in total, which equals an average of 202 schools in each education system.

Table 3.1: Sample size of teachers, schools and corresponding response rates

	Number of sampled schools	Number of sampled teachers	Effective sample size (non-missing)	School response rate (%)	Average response rate of teachers in schools (%)	Overall participation (%)
Australia	123	2,059	1,892	81	87	70
Brazil	1,070	14,291	13,258	97	94	91
Bulgaria	197	2,975	2,948	99	97	96
Chile	178	1,676	1,541	91	93	85
Croatia	199	3,675	3,628	99	96	95
Czech Republic	220	3,219	3,199	100	98	98
Denmark	148	1,649	1,582	81	77	62
Estonia	197	3,129	3,056	100	99	99
Finland	146	2,739	2,684	99	91	90
France	204	3,002	2,815	82	75	61
Iceland	129	1,430	1,217	95	80	76
Israel	195	3,403	3,225	98	86	85
Italy	194	3,337	3,271	98	90	88
Japan	192	3,484	3,470	96	99	95
Korea	177	2,933	2,819	89	88	78
Latvia	116	2,126	2,088	80	96	77
Malaysia	150	2,984	2,957	75	97	73
Mexico	187	3,138	3,098	96	91	87
Netherlands	127	1,912	1,778	81	75	61
Norway	145	2,981	2,774	73	80	58
Poland	195	3,858	3,819	100	97	97
Portugal	185	3,628	3,583	93	92	86
Romania	197	3,286	3,247	100	98	98
Serbia	191	3,857	3,818	96	97	92
Singapore	159	3,109	3,095	100	99	99
Slovak Republic	193	3,493	3,454	99	96	95
Spain	192	3,339	3,249	97	91	88
Sweden	186	3,319	3,155	96	87	84
Sub-national education systems						
Abu Dhabi (UAE)	166	2, 433	2,220	89	83	74
Alberta (Canada)	182	1,773	1,719	94	93	87
England (UK)	154	2,496	2,341	75	83	63
Flanders (Belgium)	168	3,129	3,043	84	89	75
All education systems	**6,462**	**103,862**	**100,043**	**92**	**90**	**83**

Source: OECD TALIS 2013 report, 2014, Table A.2 and TALIS microdata

TALIS 2013 surveys both head teachers and teachers about the teaching and learning process. In doing so it collects data on characteristics of the school, as well as demographics, attitudes, beliefs and behaviours of the teachers and the heads. As equation (1) suggests that the relationship of hours worked and teachers job satisfaction may be confounded also by pupils', parents' and co-workers' characteristics and behaviours, the fact that TALIS only surveys teachers and their heads may appear insufficient. Yet the hierarchical data structure of multiple teachers observed in the same school allows one to address this apparent shortcoming, as will be seen below.

3.3.2. Measurement of hours worked and job satisfaction

TALIS 2013 collects data on the amount of hours teachers work in two different ways. The first asks teachers to indicate the total hours they worked on all tasks both in and out of school including evenings and at the weekend in the last full working week prior to data collection *in one single question*. The second method asks teachers to indicate *separately* how many hours he or she spent working on a range of eleven tasks, starting with face-to-face teaching and including planning, marking, etc., during the last full working week prior to data collection.

Yet the average time teachers work in total using the sum over the eleven tasks differs significantly from the *single question* for each of the sampled education systems. In each case the sum over the individual tasks indicates a larger teachers' workload in each education system: on average the sum over the individual tasks is 8.2 hours higher (cf. Micklewright et al, 2014). It is hard to ascertain why these numbers differ and whether this is a problem. In their TALIS 2013 report for the Department for Education in England, Micklewright et al (2014) conclude that average teachers' workload reported in TALIS for England is lower, but similar to that reported in the Teachers' Workload Diary Survey conducted in England in 2013. They also note that the data collected in the workload diaries should be more precise and that needing to recall information in TALIS may bias estimates downwards. For the case of this paper, I conclude that both ways TALIS employs to measure the number of hours teachers work are likely to be biased.

The individual activity variables also contain many more missing observations (for example school leadership 14 percent; administrative tasks 9.7 percent) than the single question (2.8 percent). Although teachers were instructed to write "0" if they did not do a specific activity, the pattern of missingness in each country suggests that these respondents indeed did not engage in this activity, as the number of

respondents varies from item to item. Yet if I were to simply replace missing observations with the number zero that would imply making an assumption. Instead I mean-impute separately for each country and for each activity and thereby replace the missing values with the expected value. Thus, the variable indicating the sum of the different activities combines both raw and imputed data.

In respect to teachers' job satisfaction, TALIS 2013 contains eight items that are asked in each country. Each of these items is on a four-point Likert scale ranging from 1 representing "strongly disagree" to 4 representing "strongly agree". The items are worded as follows:

1. The advantages of being a teacher clearly outweigh the disadvantages.
2. If I could decide again, I would still choose to work as a teacher.
3. I would like to change to another school if that were possible.
4. I regret that I decided to become a teacher.
5. I enjoy working at this school.
6. I wonder whether it would have been better to choose another profession.
7. I would recommend my school as a good place to work.
8. All in all, I am satisfied with my job.

Based on these 8 items the OECD has created a continuous summary measure of teachers' job satisfaction. Following Oshagbemi (1999) multi-item indicators of job satisfaction are preferable to single item equivalents (see for example Sousa-Poza and Sousa-Poza, 2000), as the latter must be assumed to have lower levels of reliability.

Combining these items into a proxy measure for an individual's overall job satisfaction faces two challenges. First it is unknown whether two individuals indicating the same response to the same item truly feel the same. Kahneman and Krueger (2006) note that this is unlikely as the two individuals will have different experiences and may also be comparing themselves to different groups of people. One method to deal with this problem is to use vignettes, by which one can benchmark an individual's responses to a common reference point (cf. King, 2005). Such vignettes are not available in the TALIS data, so that it must be assumed that the same response to an item by different individuals is comparable.

The second challenge to creating an overall measure of job satisfaction with these items is cultural variation. Asian countries such as Japan for example are known

for their modesty. Thus individuals may underreport the degree to what they can accomplish a task when being interviewed about their self-competence. Similarly there may be cultural factors influencing teachers' responses to the job-satisfaction items.

The OECD's summary measure is the average of teachers' satisfaction with *the teaching profession* (items a, b, d, f, h) and their satisfaction with their *current school* (items c, e, g, i, j). In order to explain the OECD's procedure to obtain a variable comparable across countries, I will use teachers' satisfaction with the teaching profession as an example:

First the OECD randomly selects an equal number of observations from each education system and weights them equally, before performing a confirmatory factor analysis (CFA) using items a, b, d, f, h on this sample. Based on the estimates of this CFA, they then predict the factor scores for all observations in the entire dataset. The obtained variable is then rescaled so that the value '10' represents the mid-point of the underlying Likert-scale. Thus, any value above 10 implies average agreement with all underlying items.

In their TALIS 2013 Technical Report, the OECD (2014) shows that the underlying items a, b, d, f, h meet common standards of reliability, i.e. the items relate to an underlying scale, as measured by Cronbach's alpha, as the estimates for each education system exceed 0.80. The OECD (2014) also shows that a CFA of these five items run separately in each country all meet common standards of validity as all 32 models achieve satisfactory goodness of fit.

This whole procedure implies that there is no systematic variation in teachers' responses to the items across education systems, which may exist for example due to translation of the items, or due to cultural norms that affect how individuals in a given education systems respond to the items. The OECD (2014) tests whether the slopes and intercepts of the underlying items in the CFA are the same across countries and reports that the data rejects this hypothesis. Thus one must be aware of these differences when comparing levels of job satisfaction across countries. Interestingly, the OECD (2014) does not consider a multilevel CFA approach as an alternative. Multilevel CFAs yield 'within-group' estimators, thus it could be possible to purge the individual item loadings from systematic variation by country. As not to exceed the scope of this paper, I do not conduct multilevel

CFAs and assume that the variables provided by the OECD are satisfactory proxies.

3.3.3. Descriptive statistics

As already described above, the OECD's measure of teachers' job satisfaction is constructed to range from 5.88 to 14.05 where the value '10' is equal to the midpoint of the Likert scale used for the underlying items. Estimates in Table 3.2 are above the value of '10' for all the 32 education systems. Teachers in Mexico appear to be the most satisfied with their job (mean = 13.3) and teachers in the Slovak Republic appear to be the least satisfied with their job (mean = 11.2). Nevertheless, the estimates for the Slovak Republic indicate substantial job satisfaction. To the lay reader these estimates might appear surprising as the job of a teacher will be subject to different legal contexts; the English education system is characterised by its decentralised quasi-market, which stands in stark contrast to highly centralised education systems such as Japan. Similarly social expectations will differ across countries creating differing degrees of social pressure on teachers. All in all, the high levels of teachers' job satisfaction here resonate with the findings of Sousa-Poza and Sousa-Poza (2000) and are consistent with the literature exploring the connection of wealth and happiness. The latter suggests that individuals compare their own situation to others in their country or to their own past experiences (cf. Clark et al 2008). On the other hand general job satisfaction across countries for this one particular profession is to be expected, as individuals who are very dissatisfied with their job are likely to quit the profession (cf. Hellman, 1997; Lambert et al, 2001; Sousa-Poza and Sousa-Poza, 2007) and thus would not be sampled.

A more detailed analysis exploring the association of the school with teachers' job satisfaction and weekly workload can be found in the appendix. For brevity this section limits the description of the data to the absolute minimum, i.e. the dependent as well as the most relevant independent variables that are at the core of this paper.

Table 3.2: Job satisfaction of teachers in TALIS countries

	Mean
Australia	12.3
Brazil	11.9
Bulgaria	11.7
Chile	12.3
Croatia	11.9
Czech Republic	11.5
Denmark	12.6
Estonia	11.3
Finland	12.4
France	12.0
Iceland	12.1
Israel	12.5
Italy	12.2
Japan	11.3
Korea	10.9
Latvia	11.5
Malaysia	12.8
Mexico	13.3
Netherlands	12.2
Norway	12.2
Poland	11.8
Portugal	11.9
Romania	12.0
Serbia	12.1
Singapore	11.3
Slovak Republic	11.2
Spain	12.4
Sweden	11.4
Sub-national education systems	
Abu Dhabi (UAE)	11.8
Alberta (Canada)	12.5
England (UK)	11.9
Flanders (Belgium)	12.7
Average (all education systems)	**12.0**

Table 3.3 (below) shows the distribution of teachers' weekly working time across the five different tasks that are of interest here, which are face-to-face teaching, lesson planning, marking, school leadership and administrative tasks. To aid comparability across countries, the table reports the percentage of teachers' total weekly workload. The final eighth column reports the average total weekly workload in hours. The table shows that on average teachers in Finland work the least number of hours (approximately 36 hours) and their peers in Malaysia work approximately 60 hours. The unweighted average shows that across all these 32 education systems the average weekly workload is approximately 46 hours. The table also shows that across these education systems teachers spend approximately 64 percent of their time working on the six tasks. Again there is substantial variation, as teachers in Italy, the Netherlands and Finland only spend 50 percent of their weekly workload on these tasks. In contrast teachers in Alberta, Canada, spend around 78 percent of their weekly workload engaged in these six tasks. Across all education systems, teachers spend on average around 34 percent of their weekly workload teaching in the classroom. The second most time-intensive task is lesson planning, followed by marking.

Table 3.3: Percentage of total hours worked on specific tasks

	Teaching	Lesson planning	Marking	School manage-ment	Admin work	Sum of tasks	Average weekly workload (hours)
Abu Dhabi (UAE)	35	13	9	5	4	72	54.5
Alberta (CAN)	44	12	9	4	4	78	57.7
Australia	31	12	8	4	5	66	49.8
Brazil	42	12	9	4	3	75	54
Bulgaria	31	13	7	3	2	60	44.4
Chile	44	10	7	4	4	74	53.2
Croatia	33	16	6	3	1	62	45.4
Czech Republic	30	14	7	4	2	61	42.4
Denmark	31	13	6	2	1	58	43
England (UK)	33	13	10	3	4	68	50.8
Estonia	35	11	7	3	1	60	43.9
Finland	34	8	5	2	1	53	35.9
Flanders (BEL)	32	10	7	2	1	55	40
France	31	12	9	2	1	58	39.9
Iceland	32	12	5	2	2	58	42.2
Israel	30	9	7	3	3	56	43.9
Italy	29	8	7	2	2	53	36.3
Japan	29	14	8	4	5	66	58
Korea	31	13	6	7	4	66	53.3
Latvia	32	11	8	5	2	62	44.1
Malaysia	28	11	12	5	8	71	60.2
Mexico	38	10	7	5	3	67	49
Netherlands	28	8	7	3	2	53	40
Norway	25	11	9	3	2	55	39.6
Poland	31	9	8	3	1	56	42
Portugal	35	14	16	4	3	78	57.2
Romania	27	13	7	4	1	56	41.8
Serbia	31	13	6	4	1	59	43.4
Singapore	28	14	14	4	3	69	55.3
Slovak Republic	33	12	6	3	2	60	43.8
Spain	31	11	10	2	3	61	42.9
Sweden	29	11	8	4	1	59	44.4
Unweighted average	**32**	**12**	**8**	**4**	**3**	**64**	**46.4**

3.3.4. Summary

The purpose of this section was to introduce the reader to the data and to see whether the nature of the data has any methodological implications. In respect to the former, section 3.3.3 begins to explore how characteristics of education systems are associated with teachers' job satisfaction and workload. The following analysis will now explore how teachers' job satisfaction is produced using OLS regression to approximate equation *(1)*.

3.4. Method

For causal estimation, in an ideal world, the amount of hours teachers work would be subject to randomisation, either 'naturally' due to a change in policy, or due to an intervention. Such data though is not available. Thus in the available TALIS 2013 data described in the previous section, the amount of hours teachers work will be the result of an unknown, non-random process. Statistically, this means that bivariate OLS estimates of the workload-job satisfaction relationship may proxy the effect of other unobserved variables, if these are associated directly with job satisfaction as well as their workload. The magnitude and direction of this possible bias, often referred to as *endogeneity,* is unknown and requires empirical estimation (cf. Wooldridge, 2002).

In section 3.2, I conceive teachers' job satisfaction as a good that is 'produced' as a function of the teacher him or herself, the physical environment of the school and the other stakeholders in the particular school (see equation *(1)*). Following this production function, I then hypothesise which factors may both affect teachers' workloads and their job satisfaction, and thus would lead to endogeneity in bivariate estimates of the workload-job satisfaction relationship.

The TALIS 2013 data though does not contain all the variables I refer to in section 3.2.2. For example head teachers may, either through motivating ("good head teacher") or obliging ("demanding head teacher"), make their staff work longer hours. As such data for "good" or "demanding" head teachers is not available I am limited to using proxy variables. In this case TALIS 2013 contains data on head teachers' leadership style. Also as TALIS 2013 only surveys teachers and head teachers, there is no information obtained from the other stakeholders, such as the parents and pupils. There is also no information on the physical resources available in the school. Further there is no variable indicating a teacher's level of seniority. Instead I approximate this by including the teachers' age, time in the profession and time at the current school. TALIS 2013 does contain proxy

measures for the intensity of teacher collaboration with their peers, as well as teachers' self-efficacy. The latter proxies how competently teachers' believe they teach (cf. Bandura, 1997). I follow Albion (1999) who argues that self-efficacy can be conceived to measure how well prepared an individual is for a given task. I assume that teachers' self-efficacy in respect to teaching is not only a function of a teacher's formal education, teacher training, continuous professional development and work experience, but also a better indicator of the adequacy of a teacher's teaching skills. I therefore do not include a teacher's formal education, teacher training, and continuous professional development in any of the models specified below.

The model formulated in equation *(1)* does not make any assumptions regarding functional form. In the introduction I note that apart from not knowing the direction and magnitude of the workload-job satisfaction relationship for teachers, it is also unknown whether the relationship is linear or even monotonic. For example, teachers' job satisfaction may reach a maximum at a certain number of hours, or the relationship might be linear for only certain parts of the distribution of teachers' workload. Linear relationships have the advantage that they are easily interpretable and are therefore often preferred in the empirical literature. Yet modelling equation *(1)* using linear regressions implies making an assumption that a linear specification is a reasonable approximation. I test this assumption using three different methods, pooling all observations together. First, I test a linear specification, which yields a statistically significant estimate below the 5 percent level. I then include a quadratic term, to explore whether teachers' job satisfaction might be maximised at a certain number of hours worked, or if the relationship increases at a decreasing manner similarly to the relationship of work experience and wages. I find that a quadratic specification is not statistically significant at or below the 5 percent level. Secondly, I estimate the workload-job satisfaction relationship replacing the continuous total workload variable with fourteen dummy variables indicating 10-hour increments. I find no statistically significant differences across these dummy variables. Last, local polynomial regression strongly supports the linearity assumption.

I build the model in four steps addressing an increasing number of the potential sources of endogeneity by adding control variables in each step. As the main interest of this paper is to explore patterns across education systems, each of the following equations is applied separately to each of the 32 sampled education

systems. I consider two specifications of the following equations, the sum of the time spent on all activities[17] *(total workload)*, and the number of hours worked on five different activities: face-to-face teaching, lesson planning, marking, involvement in school leadership and administrative chores [18]. The latter specification relaxes the assumption that the kinds of activities teachers do while working do not matter. In other words this specification allows teachers to gain job satisfaction from one activity and experience a reduction in their job satisfaction from another. It is therefore preferred to a specification simply testing the total amount of hours worked per week. Nevertheless, both specifications will be reported in the following sections.

The first model represents the baseline and consists of the main variable(s) of interest M, which are either the total hours worked or the amount of hours spent on different activities during the teachers' working week. This baseline model is summarised in equation *(2):*

$$S_{ij} = \alpha + \beta_1 M_i + u_{ij} \qquad (2)$$

The second model additionally controls for part-time status, the teachers' involvement in mentoring, the degree of cooperation with their colleagues, whether or not they are on a permanent contract, their gender and their self-efficacy.

$$S_{ij} = \alpha + \beta_1 M_i + \beta_2 O_i + u_{ij} \qquad (3)$$

The third model additionally includes proxy measures of potential school level confounders. I control for the degree of distributional and instructional leadership, as well as the head teacher's gender, age as well as the ratio of teachers per administrative or pedagogic member of support staff. Thus M3 can be summarised as

$$S_{ij} = \alpha + \beta_1 M_i + \beta_2 O_i + \beta_3 \kappa_j + u_{ij} \qquad (4)$$

Although containing additional controls O and κ the estimates of the β_1 may still be confounded by unobserved variables at the school level (for example pupils, parents, physical resources) and individual level (for example teacher's

[17] As described in section 3.3, the values in this variable may be hybrids of raw reported data by the individual him or herself and imputed data. All in all, this variable is the sum of ten activities. I include a dummy for each of these activities, which is 1 if the value for this activity is imputed.

[18] I control for the amount of time spent on any other activities.

motivation). The final model therefore exploits the structure of multiple teachers observed in each school and introduces a dummy variable for each school, also known as school-fixed effects:

$$S_{ij} = \alpha + \beta_1 M_i + \beta_2 O_i + K_j + u_{ij} \qquad (5).$$

This comes with the drawback of not being able to estimate coefficients of school-level variables. On the other hand, school-fixed effects remove all between school-variation, i.e. the observable *as well as* known and unknown unobservable school-level confounders. Also, as already shown, despite removing all between-school variation in equation (4) the vast majority of the variation is still being exploited.

But although equation (4) is preferred, the obtained estimates cannot be considered causal, as potential unobserved confounders at the individual level such as teachers' motivation remain. More motivated teachers may have higher levels of job satisfaction and also work more hours, which will bias the obtained estimates upward. On the other hand the variables capturing the amount of time teachers work in total or spend on the different tasks may be measured with error. If this error is 'classical', the obtained estimates will be biased towards zero. Finally, as is always the case with observational data, the direction of effects must be assumed to go from the right to the left hand side of equations 2 to 5.

The main aim of this paper is to have comparable estimates across countries. Usually, if one has one single regression with independent variables measured on different scales, one can standardise these variables in order to put them on the same scale, which allows making reasonable comparisons of effect sizes. In the case of this research the focus lies on comparing estimates across education systems. Here the main variables of interest are all measured in hours, a metric common to all education systems. In section 3.3 though I note that regardless of the OECD's efforts, one still needs to make the assumption one can compare the level of teachers' job satisfaction across these education systems. By standardising teachers' job satisfaction separately for each education system, I relax this assumption as the dependent variable then is also on the same scale for all education systems. I therefore report regular regression coefficients that indicate the change of the dependent variable in standard deviations, when the workload changes by one hour.

3.5. Findings

Before reporting the findings I would like to start with a comment on their presentation. The findings are presented in Figures 3.7, and 3.10 to 3.14. Each of these illustrates the estimates for the effect of a change in teachers' total workload or specific kinds of workloads (also referred to as activities) as horizontal bars for models 1, 2 and 4 in descending order for each of the 32 sampled education systems. M1 refers to the baseline, bivariate OLS model, while M2 controls for teacher-level covariates, and M4 accounts for all possible school level variables by introducing school-fixed effects. The country estimates are ranked according to their estimate obtained from M4, which is the preferred model, as it controls for the largest number of potential confounders. But as the estimates in M4 will vary across the different kinds of workloads, the order the education systems are in will vary each time, and is simply intended to aid readability. The reported estimates indicate the marginal effect of an increase in teachers' total workload by one hour on their nationally standardised job satisfaction. This standardisation ensures that teachers' job satisfaction is on the same scale for each education system; at the same time teachers' total weekly workload is measured in hours, thus the obtained estimates are comparable in size across education systems.

The error bars indicate the 95 percent confidence interval. In their official TALIS report, the OECD (2014) apply a balanced repeated replication procedure (BRR) to estimate their standard errors (cf. Micklewright et al, 2014). I do not use this procedure, but instead use adaptations of the heteroskedasticity-robust estimators proposed by White (1980). These allow the errors to be correlated within school, but not between schools. The cluster-robust standard error procedure used in this research though yields very similar estimates of the standard errors to the BRR procedure and is less computationally intensive. In addition, I use post-stratification weights provided in TALIS to account for teachers differing in their probability of being sampled.

I consider parameter estimates to be statistically significant at the 5 percent level and below throughout the remainder of this paper. The sample sizes underlying each regression are the same as reported in the column headed "effective sample size" in Table 3.1. The "sample used" differs from the original sample size due to missing observations in the dependent variable and recoding teachers' workloads larger than the number of hours per week (168 hours) as missing. Missingness in

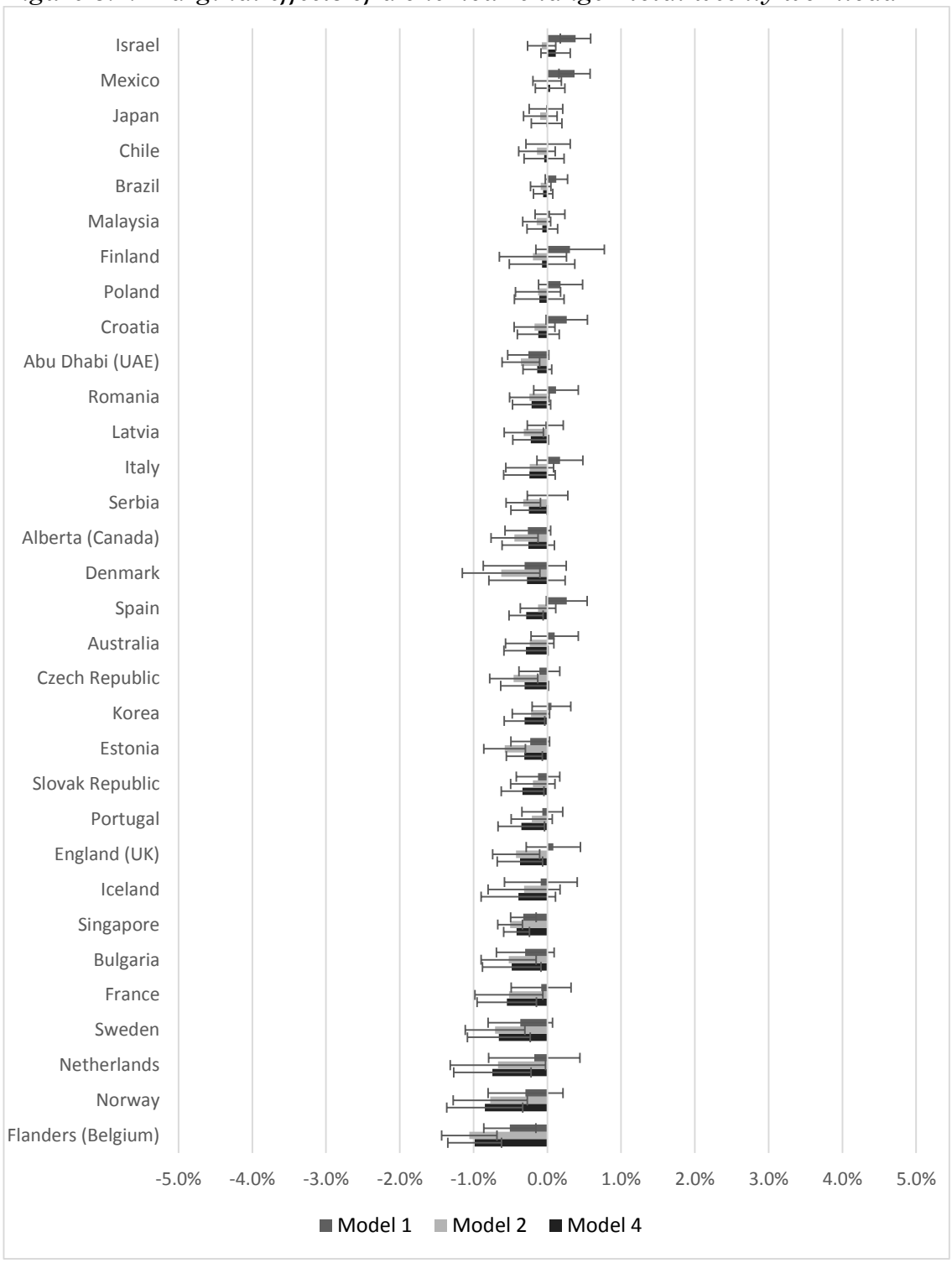
all other right hand side variables is addressed using dummy variables and mean imputation.

The reader should also be aware that the presented estimates may come across as small, as a one hour increase in face-to-face teaching for example may only be small compared to the standard deviation of hours in the respective education system. Table 3.3 shows the standard deviations associated with the different kinds of teachers' workloads. According to Table 3.3, the unweighted average standard deviation of face-to-face teaching is 12 hours. In contrast a standard deviation ranges between 8 and 14 hours for planning lessons, ranges between 4 and 14 hours for marking, ranges between 1 and 10 hours for leadership activities and ranges between 2 and 10 hours for administrative activities.

3.5.1. Does working longer reduce teachers' job satisfaction?

The findings in the baseline model (M1) shown in Figure 3.1 do not clearly support this expectation as only 17 of the 32 education systems show a negative association. This association though is only statistically significant at the 5 percent level in Singapore (-0.3 percent of a SD), Flanders and Denmark (each -0.5 percent of a SD). On the other hand, 15 education systems show a positive association, suggesting that on average teachers gain job satisfaction from working more hours. Among these 15 the estimates for Brazil (0.2 percent SD) Israel and Mexico (both 0.3 percent of a SD) are statistically significant.

Figure 3.1: Marginal effects of a one hour change - total weekly workload



Notes: 95 percent confidence intervals based on robust standard errors clustered by school. Countries ranked by M4 estimate. Reported coefficients measured in SD.

The adopted production function approach in this paper suggests that bivariate estimates of the workload-job satisfaction relationship might proxy teacher-level and school-level variables. When additionally controlling for the observable teacher-level variables (M2), there is a clear result: Figure 3.1 shows that the estimates for *all* education systems are reduced, i.e. become less positive or more

negative. I test the hypothesis that the joint effect of all the teacher-level covariates in the model, for example part-time status and self-efficacy, on job satisfaction is zero. The estimated coefficients from these tests are positive and statistically significant in each education system. I therefore conclude that jointly these variables are positively associated with teachers' workload (cf. Wooldridge, 2002). Furthermore, these findings suggest the workload job-satisfaction relationship may be negative as after controlling for the teacher-level covariates, estimates for 15 education systems are statistically significant; and for Israel and Mexico, the previously (M1) positive significant estimates are now not significantly different from zero.

Following equation *(1)* I move the focus to the school level in models 3 and 4. In addition to the variables in M2, M3 additionally controls for the degree of distributional and instructional leadership, the head teacher's gender and age, as well as the ratio of teachers per administrative or pedagogic member of support staff. Although F-tests[19] indicate that including these variables improves model fit significantly at the 5 percent level in 10 cases, interestingly the coefficients on workload remain virtually unchanged for each education system. For ease of readability these estimates are not shown in the graph. Nevertheless, these findings are surprising as they suggest that these particular variables are not correlated with teachers' workload.

In M4 I introduce school-fixed effects, which control for both the observed school-level variables in M3 and all unobservable school-level variables. Here the estimates increase in 20 cases, i.e. they become less negative or more positive. This suggests that the unobserved school-level variables appear to be negatively associated with teachers' job satisfaction. But in 12 cases the unobserved school-level variables appear to be positively associated with teachers' job satisfaction. Including school-fixed effects, the estimates for Abu Dhabi and Alberta that are statistically significant in M2 are not statistically significant any more. Yet the expected negative relationship of workload and teachers' job satisfaction remains statistically significant for 13 education systems and estimates range from -0.3 percent to -1 percent of a SD. These education systems are Spain, South Korea,

[19] STATA 13 cannot compute log-likelihood ratio tests when correcting for robust or clustered standard errors. I therefore use F-tests.

Estonia, the Czech Republic, Australia, Singapore, England, France, Denmark, Sweden, the Netherlands, Norway and Flanders[20].

The pattern in the adjusted[21] R^2 moving from the baseline to the preferred model shows that it is M2 and M4 that substantially improve model fit. On average across all 32 countries the adjusted R^2 is 0.4 percent in the baseline model (M1). With the introduction of the teacher-level covariates the adjusted R^2 increases on average to 13.6 percent. Including the observable school-level variables in M3 reduces the adjusted R^2 slightly in 11 countries and increases in the remaining countries, but these changes are only very small so that on average the adjusted R^2 is 13.8 percent and ranges between 8.5 percent in Latvia to 19.9 percent in Abu Dhabi. Including the school-fixed effects increases model fit in *every* country, the increase ranging from 2.5 percentage points in Poland to 10 percentage points in Brazil. On average the R^2 in M4 is 19.6 percent.

So far, I have described how the estimates change when moving from model to model. One of my research interests is to explore the extent to which the baseline estimates proxy unobserved school and teacher level variables. Figure 3.2 sheds light into this issue and shows a scattergram of the estimates obtained in the baseline (M1) on the X-axis and the preferred model (M4) on the Y-axis. It clearly shows that whereas 18 baseline estimates are positive the graph also shows that 29 M4 estimates are zero or negative. Also all but 3 data points are below the black 45 degree line, i.e. 29 M4 estimates are less positive or more negative than the baseline model. Thus the baseline estimates appear to be biased upward. The reader should though be reminded that the location of these 32 data points in the graph are subject to substantial uncertainty, as is indicated by the large confidence intervals in Figure 3.7. Because of the uncertainty attached to these data points one would expect them to be further spread out. Interestingly these data points are well approximated by a regression line, which achieves an R^2 of 0.6.

[20] In section 3.3, I discuss that TALIS asks two different questions from which one can generate an indicator for teachers' total weekly workload and that these variables are not the same. The variable used here is the sum over the individual activities teachers do during their working week. I run the preferred model using the single question asking teachers how many hours they worked in total in the last full week prior to data collection. Using the estimates from these two versions of M4 for all 32 education systems as data points, when summarising these data with a regression line I find that the estimates reported in Figure 3.7 are approximately on average 60 percent the size of the estimates on the single question. The R-squared of this regression line is 0.61, which indicates that a substantial proportion of variance is not accounted for. This is not particularly surprising as a large proportion of estimates in both M4 versions are not statistically significant. This large proportion of uncertainty around each data point is likely to cause the data points to be scattered rather than be allocated close to a regression line.

[21] The adjusted R2 accounts for the number of variables included in the model, i.e. the degrees of freedom. I estimate this statistic by including dummy variables for each school into M3.

One possible explanation for the findings shown in Figure 3.1 could be that there is a correlation between the amount of time teachers work on average in an education system and the estimated effect of an increase in workload on their job satisfaction. Thus, it could be that the significant estimates reported above occur in education systems in which teachers have a high average workload. I explore this aspect in Figure 3.3 where I plot average total teachers' workload (X-axis) by the estimated coefficients on total workload (Y-axis). Figure 3.3 does not suggest an association between these two variables. The 14 education systems with statistically significant estimates are also in the same range of average total teachers' workload as those that do not have significant estimates.

Figure 3.2: Correlation of Model 1 (baseline) and Model 4 (preferred) estimates – total weekly workload

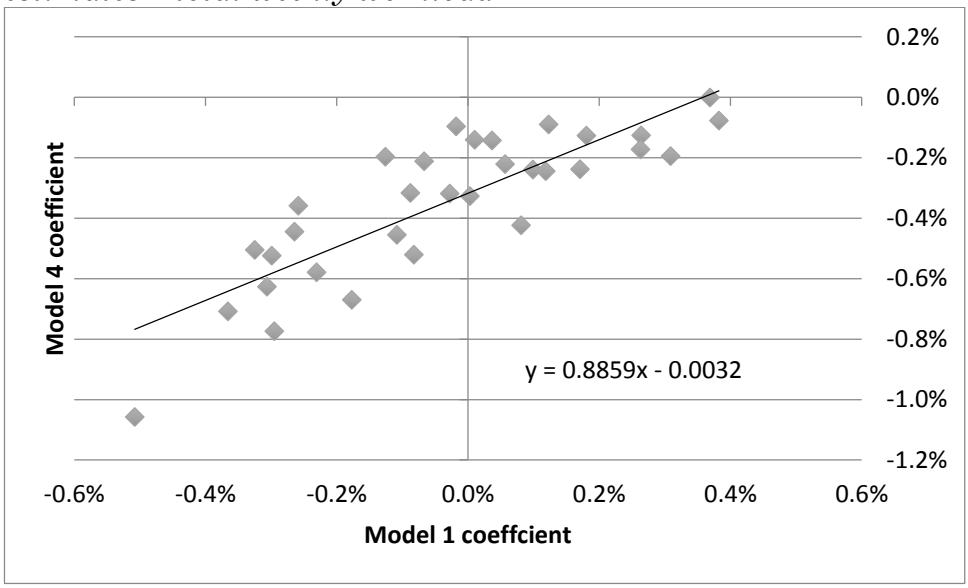
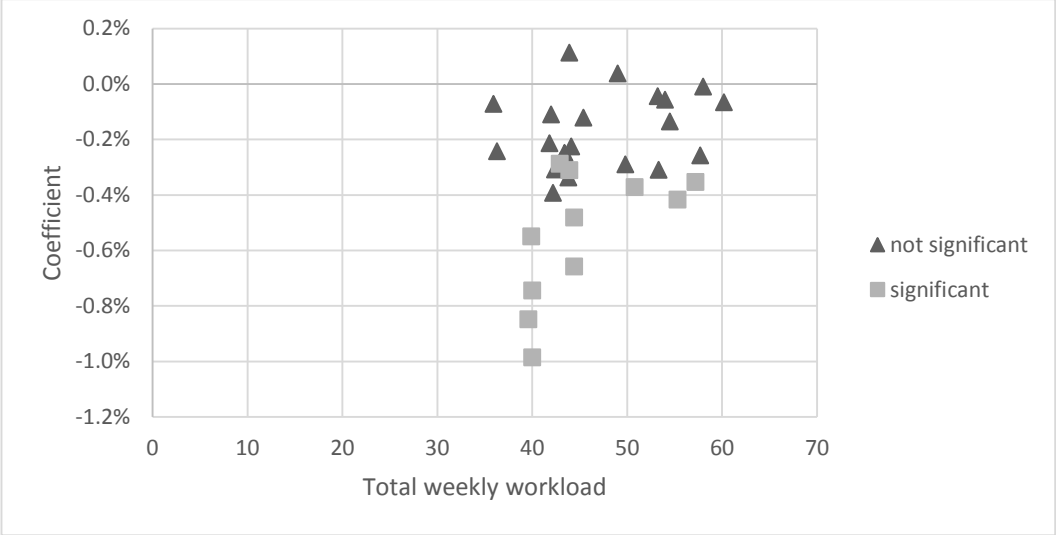


Figure 3.3: Correlation of average workload and estimates for the workload job-satisfaction relationship



Although all these estimates, even those that are statistically significant at conventional levels, are very small, the reader should be reminded that a change in total workload by one hour is a very small increase. The data shows, that the distributions of teachers' total weekly workload vary across education systems (Table 3.4). And the unweighted average across education systems is a standard deviation of 22 hours. Among those with a statistically significant estimate in M4, the standard deviation of teachers' workload is 16 hours in Latvia. Thus a teacher who works a standard deviation more than the average teacher will be expected to be 3.2 percent of a standard deviation less satisfied with their job. In Flanders a standard deviation is 8 hours, so that such an increase in total weekly workload is associated with an 8 percent decrease in job satisfaction. In short, even when comparing effect sizes for changes in workload by national standard deviations, the effect sizes are still very small.

In summary, this sub-section aimed to uncover cross-country patterns for the association of teachers' total weekly workload and their job satisfaction. It finds that support for equation (1), which suggests that bivariate estimates of this relationship are biased upward, because they proxy unobserved teacher and school level variables. Estimates from the preferred model (M4) also suggest that teachers suffer a reduction from an increased workload, but the effect sizes are very small. There is also substantial imprecision in most estimates, so that only 13 of the 32 cases are statistically significant.

Table 3.4: Standard deviation of teachers' workloads by country

	Face-to-face hours	**Planning**	**Marking**	**Leader-ship**	**Admin.**	**Total weekly hours**
Australia	12	7	7	10	5	20
Brazil	31	16	14	8	6	42
Bulgaria	11	8	5	4	4	17
Chile	17	10	6	4	5	26
Croatia	8	8	5	1	4	17
Czech Republic	8	7	4	3	4	18
Denmark	8	6	4	4	2	13
Estonia	13	7	6	4	4	24
Finland	11	5	4	2	4	13
France	4	7	5	1	2	12
Iceland	9	6	4	4	3	12
Israel	15	8	8	5	4	34
Italy	6	6	5	3	3	16
Japan	10	9	6	6	8	23
Korea	10	8	6	4	8	18
Latvia	16	9	8	4	5	27
Malaysia	14	10	12	9	10	46
Mexico	26	8	7	6	5	28
Netherlands	8	5	5	5	2	18
Norway	8	8	11	5	5	25
Poland	16	6	7	3	5	29
Portugal	8	13	13	7	10	17
Romania	12	10	6	3	4	31
Serbia	12	8	5	3	4	20
Singapore	10	7	9	4	6	23
Slovak Republic	10	9	5	5	5	22
Spain	11	7	10	4	3	18
Sweden	8	6	6	3	6	15
Abu Dhabi (UAE)	16	14	10	7	7	24
Alberta (Canada)	12	8	7	6	5	19
England (UK)	9	7	7	5	5	18
Flanders (Belgium)	8	7	5	2	3	16
Average	**12**	**8**	**7**	**4**	**5**	**22**

Note: Standard deviations computed from standard errors clustered by school.

3.5.2. Is there heterogeneity across different activities?

The findings for teachers' total workload presented in the previous sub-section give rise to the question whether there is heterogeneity in the impact of additional hours across different kinds of activities teachers do. This could explain the large number of non-significant estimates if the effects of different activities cancel out each other. For example teachers might gain job satisfaction from being involved in school leadership activities, but their job satisfaction may be reduced from an increase in time spent marking.

But before answering this question I will first present the estimates for teachers' face-to-face teaching, planning, marking, administrative and school leadership workload, will examine whether the estimates point in the expected directions and will explore how and to which extent teacher and school-level variables are proxied in the baseline estimates.

Do the estimates go in the expected directions?

Face-to-face teaching is the activity that is predominantly associated with schoolteachers, and it is also likely to be the activity that attracts people to the profession itself. In the previous sub-section I note that measures of workload may in fact proxy human emotions. In the case of face-to-face teaching teachers might feel pride and responsibility as they have a key role in ensuring their pupils have the skills they need to succeed in the future. One might therefore assume that teachers gain job satisfaction from this. As before with overall workload, the estimates for M1 (Figure 3.4) indicate mixed support for this expectation. Fourteen of the 32 cases show positive associations, of which only that of Finland (0.2 percent of a SD) is significant at conventional levels. Among the 18 remaining cases that show a negative association, the estimates for 6 education systems (Singapore, the Czech Republic, Norway and Denmark, Japan) are significant, ranging from -0.5 to -1.6 percent of a SD. This picture is similar to the baseline estimates for teachers' total workload where there are 19 cases with negative estimates and 13 with positive estimates.

When adding the available observable teacher-level variables the figure shows a shift in the estimates in both positive and negative directions. This suggests that these teacher-level covariates can be both positively and negatively associated with face-to-face teaching workloads. When controlling for these variables, Figure 3.4 shows that the estimates for Estonia (-1 percent of a SD) and Korea (-0.8 percent of

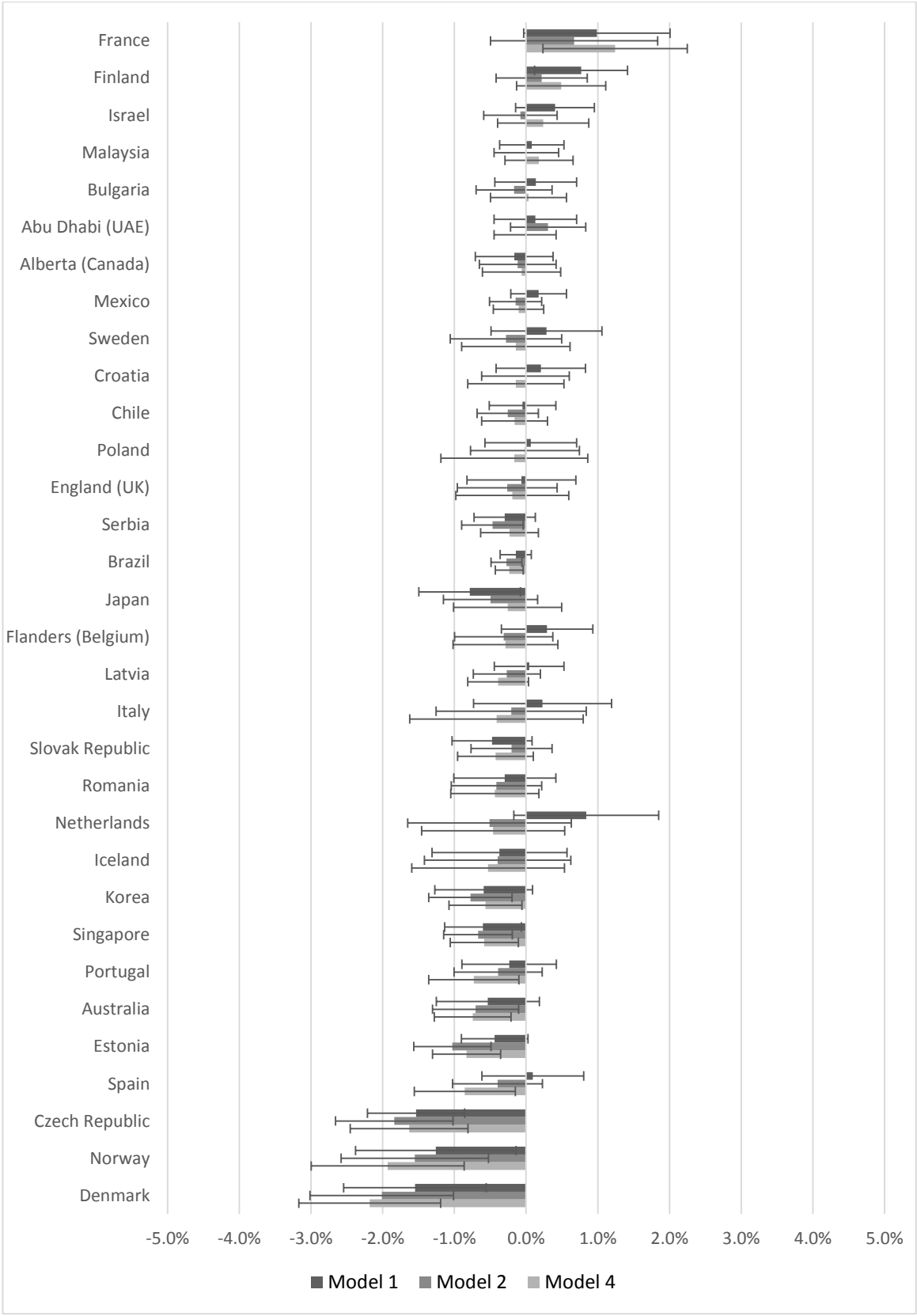
a SD) have become significant, but the estimates for Finland and Japan are no longer statistically different from zero.

As in the previous sub-section, when adding the available school-level variables, the estimates do not change. This is also the case for all other activities discussed here, which suggests that these particular variables are not associated with any of the activities examined. As before the estimates for M3 are not included in the graphs shown to aid readability.

In the final preferred model, in 19 cases the estimates increase, i.e. become more positive or less negative, and decrease, i.e. become more negative and less positive in the remaining 12. Analogously there is one country, France, with a statistically significant positive coefficient (1 percent of a SD) and Spain with a statistically negative coefficient (-0.9 percent of a SD). In contrast to the findings for overall workload, where the unobservable school level factors are associated negatively with teachers' workload, these estimates here suggest that these variables can be correlated positively and negatively with teachers' workload depending on the respective education system.

Planning lessons is another integral part of being a teacher. When planning lessons teachers need to develop a strategy and material with the aim to enable each of their pupils to achieve a specific learning outcome. As this process requires a certain degree of creativity from the teacher, one might expect that, similar to other creative activities such as art and music, teachers may gain job satisfaction from this activity. Estimates from M1 presented in Figure 3.5 though show that only 10 cases have positive estimates. Of these 10 education systems, only the estimate for Chile (0.8 percent SD) is statistically significant. When adding variables in models 2 to 4 the estimates change in both positive and negative directions, but of negligible magnitude. In none of the education systems do estimates become significant, and the estimate for Chile remains statistically significant in M4 with a mean of 0.9 percent SD. I conclude that there is little support that teachers' planning workload is associated with their job satisfaction, nor do the baseline estimates appear to proxy the school level and the available teacher-level data.

Figure 3.4: Marginal effects of a one hour change - face-to-face teaching workload



Notes: 95 percent confidence intervals based on robust standard errors clustered by school. Countries ranked by M4 estimate. Reported coefficients measured in SD.

As part of their job teachers may also be involved with *school leadership activities*. Following Harris and Muijs (2002) involving teachers in school leadership activities is commonly referred to as "Teacher Leadership", which apart from being given the opportunity also entails teachers being entrusted to take over certain responsibilities. One may therefore expect that being involved in school leadership activities will be associated positively with teachers' job satisfaction. Estimates in Figure 3.6 show strong support for this expectation, as that the baseline association is positive in 29 cases, of which 17 are statistically significant. When introducing the teacher-level variables the coefficients for all 32 education systems are reduced, thus indicating that the teacher-level variables are positively associated with teachers' leadership workload. In M2, coefficients that had been statistically significant in the baseline are now not statistically different from zero in Finland, Poland, Chile, Bulgaria, Australia, Singapore and Brazil. In the preferred model estimates are reduced further in 13 cases and 13 others increase suggesting that the unobservable school-level variables may either be associated positively or negatively with teachers' job satisfaction. In this model specification the estimates for Abu Dhabi (0.9 percent of a SD) and Poland (1.3 percent of a SD) become significant again compared to the corresponding estimates obtained in M2.

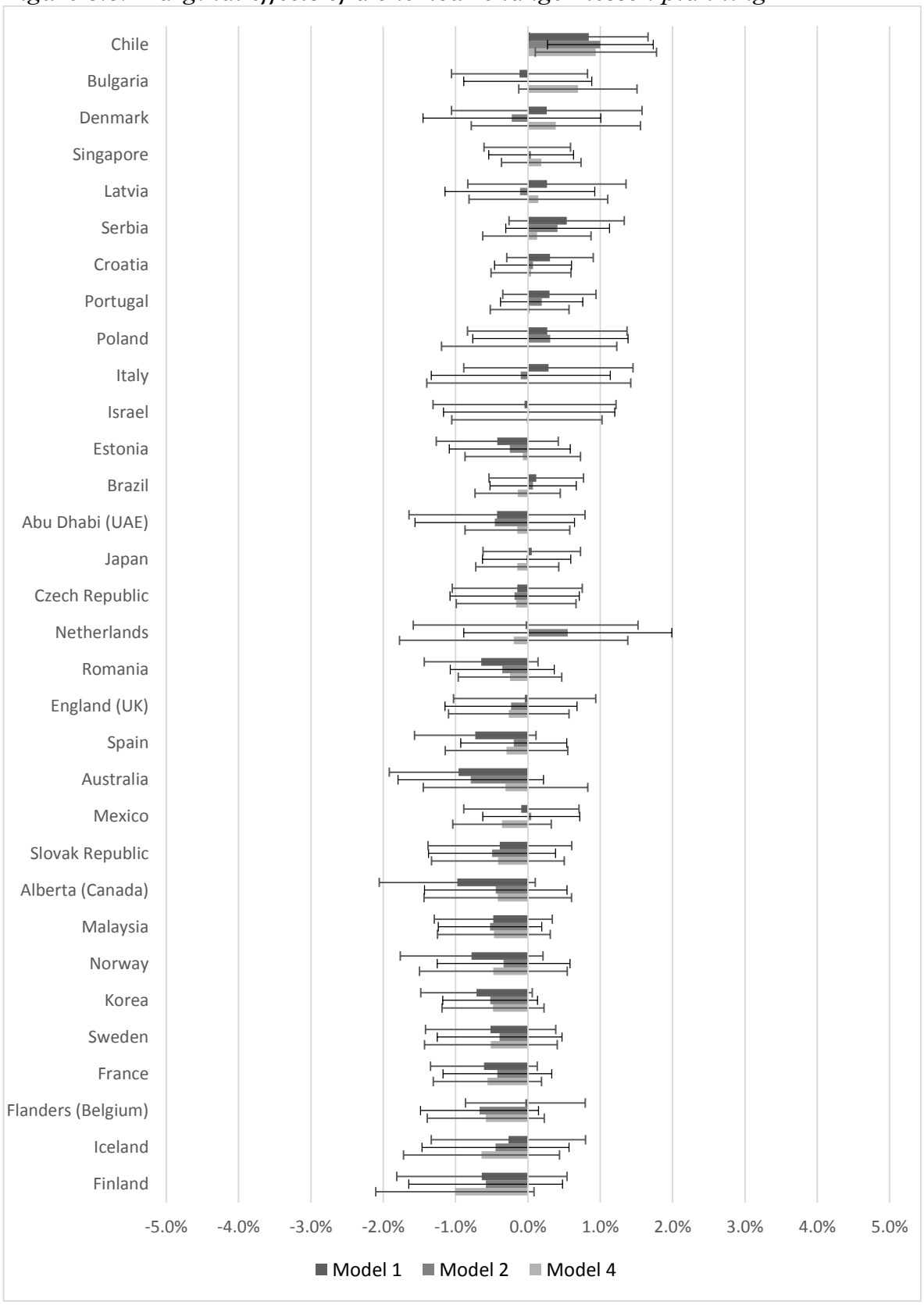
Another common activity teachers do during their job is to *mark students' work*. But in contrast to the previous activities one might expect that this repetitive work might be associated negatively with teachers' job satisfaction. The findings presented in Figure 3.7 appear to support this expectation, as in 25 of the 32 education systems the estimates are negative. In eight of these education systems the associations are statistically significant at the 5 percent level and range from -1 to -2.2 percent of a SD. When adding all available teacher-level variables in M2, which are jointly positively associated with job satisfaction, I conclude that they are positively correlated with marking workload, as the estimates are reduced in 18 cases (cf. Wooldridge, 2002). Here Norway (-0.9 percent of a SD) and Bulgaria (-1.5 percent of a SD), two education systems whose estimates in M1 are not significant become significant. In the final model estimates are also reduced in 17 cases, but in 15 cases the estimates increase, yet for none of the education systems do the estimates become statistically significant or cease being so. These changes nevertheless support equation 1 and the underlying assumption that the baseline estimates proxy, in this case, unobservable school-level characteristics.

The last activity examined here is teachers' *administrative workload*. According to the National Union of Teachers, a British teachers' union, administrative work is widely perceived to be "unnecessary"[22]. One might therefore expect a negative association with job satisfaction. Interestingly the estimates here presented in Figure 3.8 support this expectation in the baseline model. Here 23 education systems have a negative association, of which 9 are statistically significant and range from -1.2 to -3.2 percent of a SD. None of the positive associations are statistically significant. The available teacher-level variables also appear to be positively correlated with the administrative workload, as this variable is positively associated with job satisfaction and the coefficients on administrative workload is reduced for 22 cases, but the estimates increase in 10 cases. When controlling for these teacher-level variables the estimate for Poland becomes statistically significant and the estimate for Serbia is now not statistically different from zero. In the preferred model estimates increase for 15 cases and reduce for 17 cases, but in none of these cases do estimates become or cease to be statistically significant. Nevertheless, these changes support the assumption that the baseline estimates proxy, in this case, unobservable school-level characteristics.

In summary, the findings show mixed support for the expected directions of the respective estimates. The statistically significant estimates for marking, administrative and school leadership activities point in the expected direction for each activity. In contrast, I find the expected increase in teachers' job satisfaction when their face-to-face teaching workload increases in only one of the 8 significant cases, which is France. The fact that there are 7 countries with significant negative estimates suggests that although teachers may be drawn into the profession by the joy of teaching, they are more satisfied the less hours they do this activity. Yet, as for each kind of activity the majority of estimates are not statistically significant it is hard to make stronger claims.

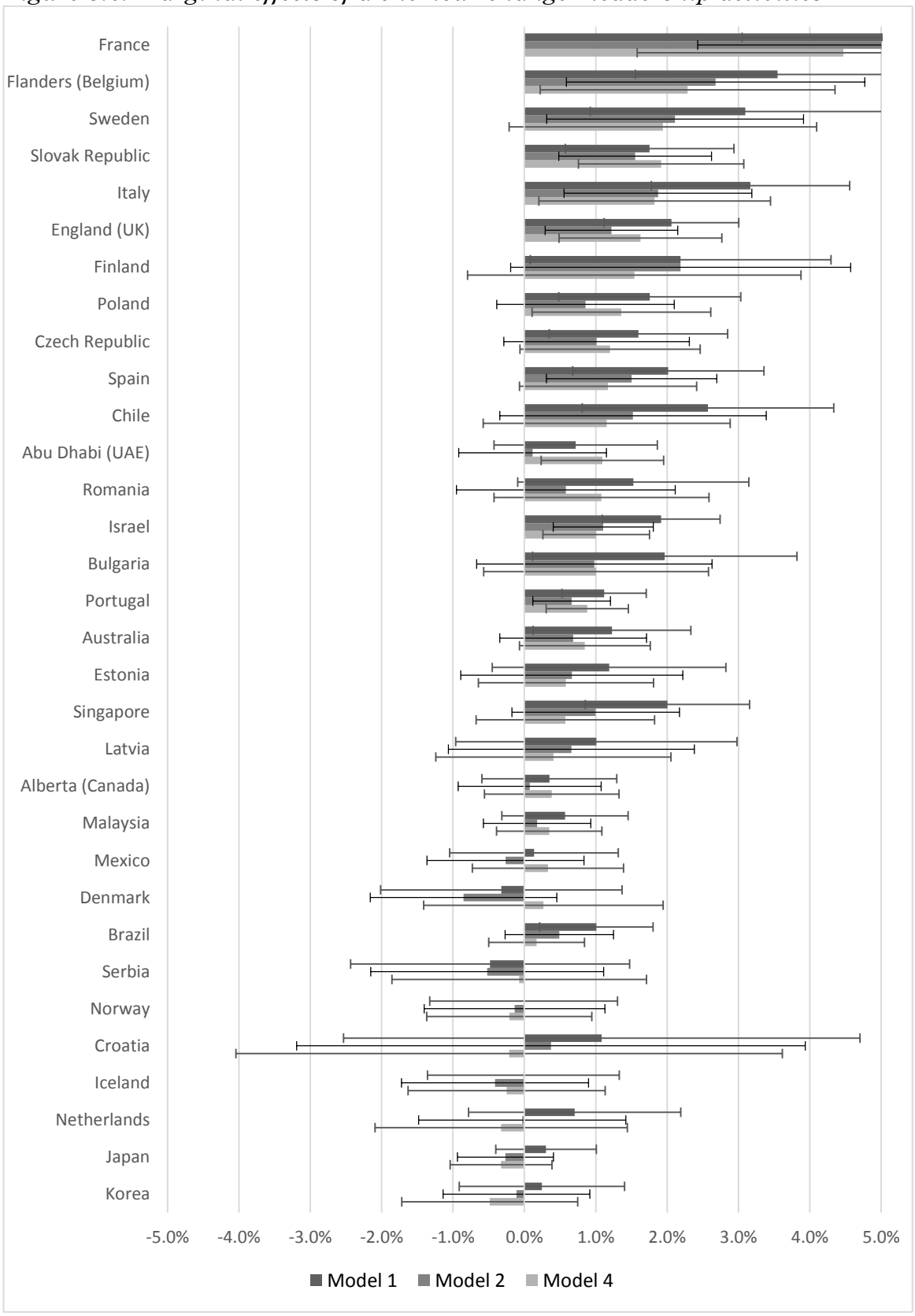
[22] http://www.teachers.org.uk/node/20645, last accessed on the 27th of January 2015

Figure 3.5: Marginal effects of a one hour change - lesson planning



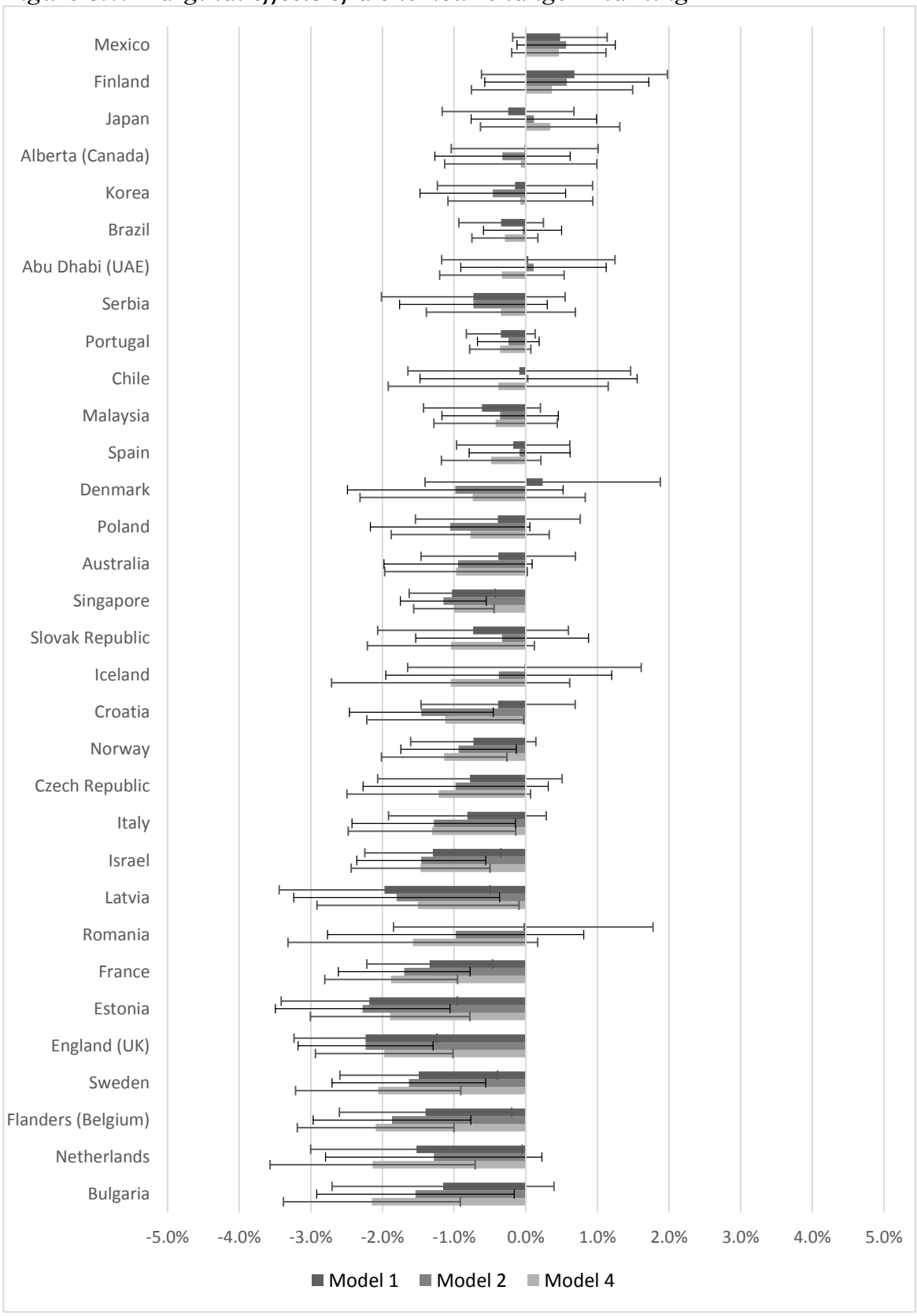
Notes: 95 percent confidence intervals based on robust standard errors clustered by school. Countries ranked by M4 estimate. Reported coefficients measured in SD.

Figure 3.6: Marginal effects of a one hour change - leadership activities



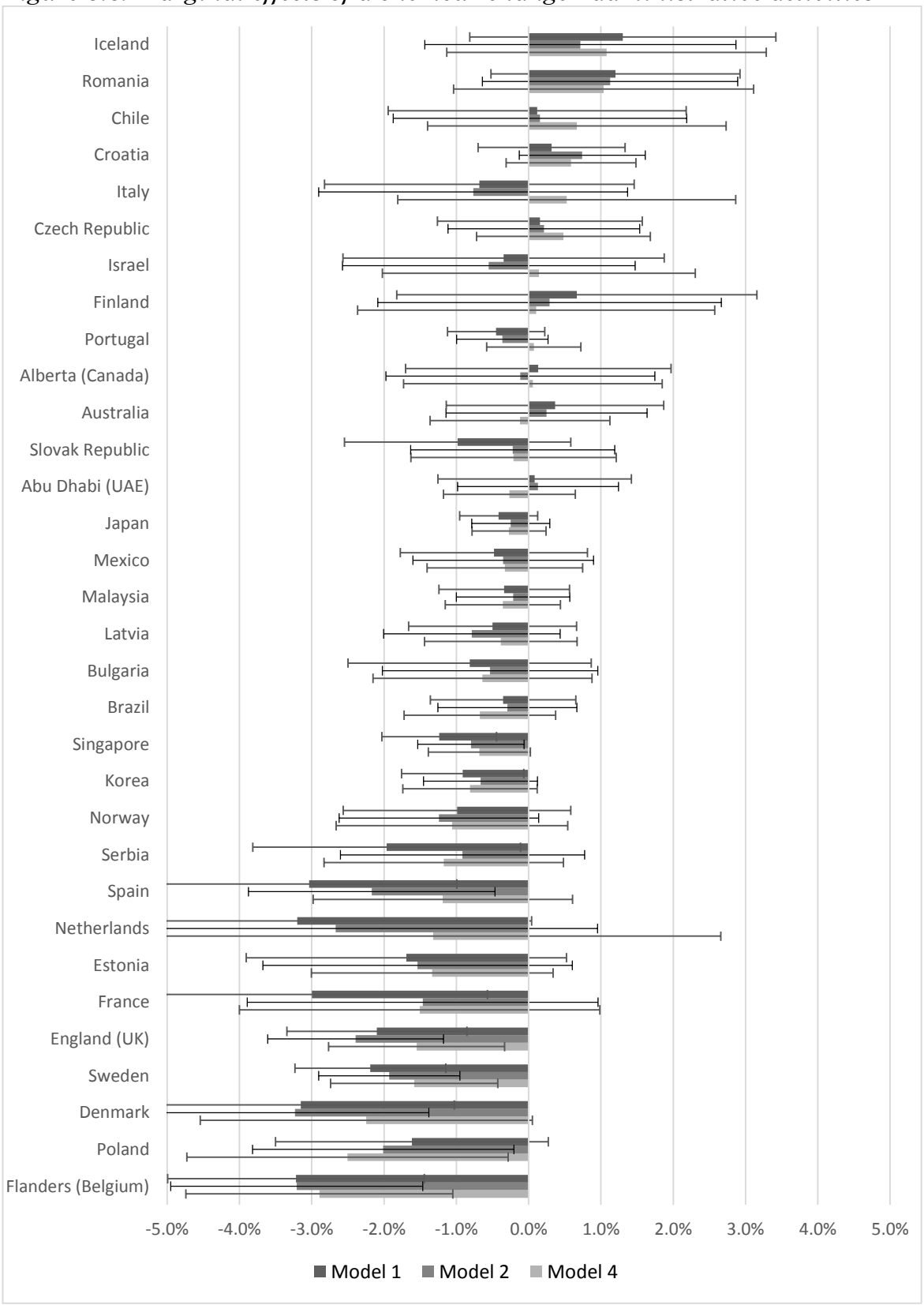
Notes: 95 percent confidence intervals based on robust standard errors clustered by school. Countries ranked by M4 estimate. Reported coefficients measured in SD. Estimates for France are 5.5% in M1, 5% in M2.

Figure 3.7: Marginal effects of a one hour change - marking



Notes: 95 percent confidence intervals based on robust standard errors clustered by school. Countries ranked by M4 estimate. Reported coefficients measured in SD.

Figure 3.8: Marginal effects of a one hour change - administrative activities



Notes: 95 percent confidence intervals based on robust standard errors clustered by school. Countries ranked by M4 estimate. Reported coefficients measured in SD.

How do baseline and preferred estimates correspond?

So far, the narrative in this sub-section has focused on whether the estimates point in the expected direction and how the estimates change moving from the baseline to the preferred model. In order to get a better understanding of the direction of the bias in the baseline estimates by not controlling for the available teacher-level and all school level covariates, I plot the baseline estimates for each of the different activities against the corresponding estimates obtained in the preferred model. As in Figure 3.2 there are 32 data points plotted in each graph (Figure 3.9-3.12) as well as a black 45 degree line. If data points are located on this line the estimates for both models 1 and 4 are the same. Accordingly if data points are above this line, the corresponding M4 estimate is larger than the baseline; if data points are below this line, the corresponding M4 estimates are lower than the baseline.

In the previous section I find that for teachers' total weekly workload the baseline estimates appear to overstate the workload effect, i.e. all data points are below the 45 degree line. I find this general pattern for face-to-face teaching (17 data points below the line, see Figure 3.9), marking (21 data points below the line, Figure 3.10), leadership (27 data points below the line, Figure 3.11). Thus, the baseline estimates for these activities tend to be overstated (Figure 3.12), but this pattern is not as clear as for total weekly workload. In contrast, the estimates for administrative activities tend to be understated, as in 24 data points are above the 45 degree line meaning that the M4 estimate is more positive or less negative than the corresponding baseline estimate.

Figure 3.9: Correlation of baseline (M1) and preferred (M4) estimates – face-to-face teaching

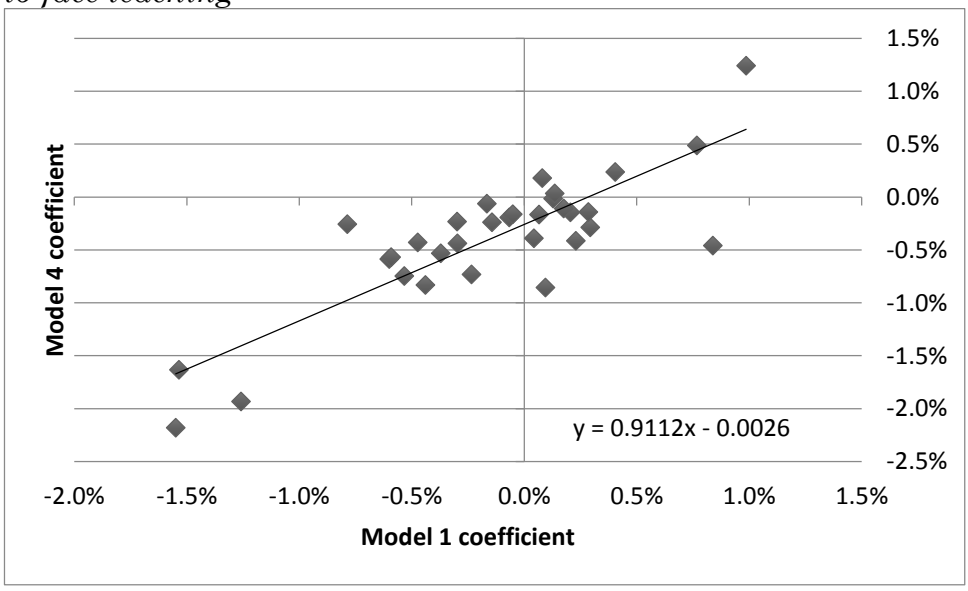


Figure 3.10: Correlation of baseline (M1) and preferred (M4) estimates - marking

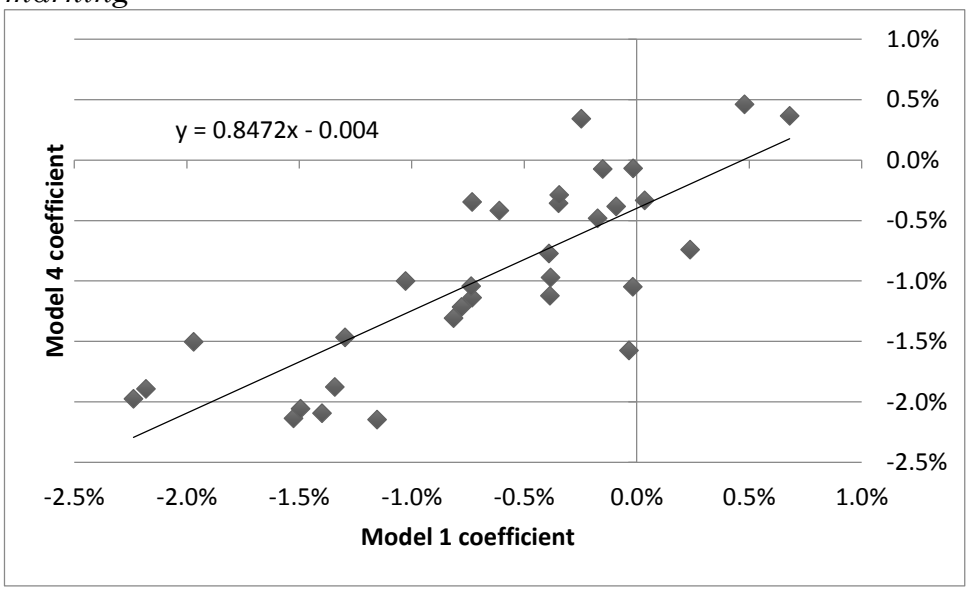


Figure 3.11: Correlation of baseline M1) and preferred (M4) estimates – school leadership activities

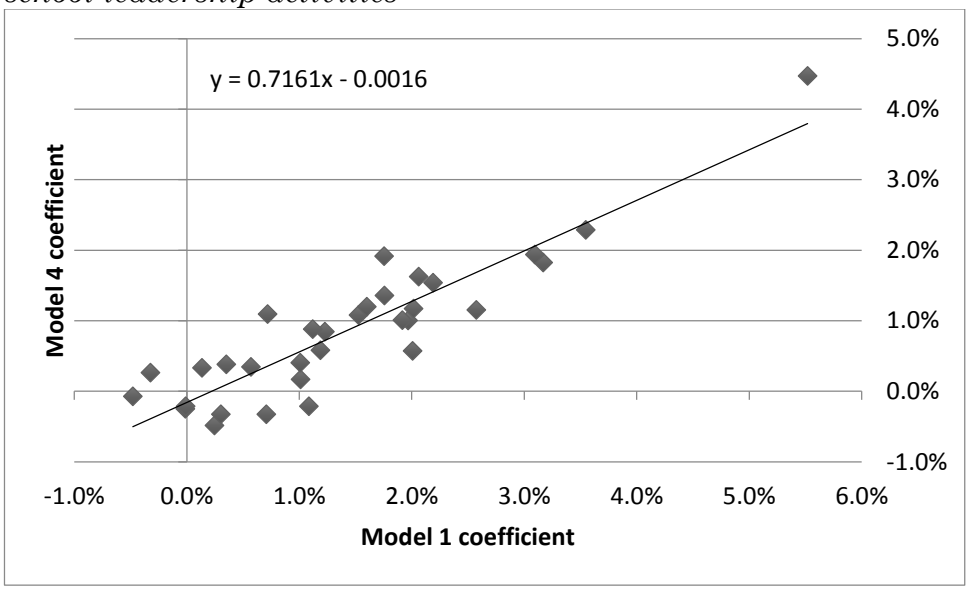
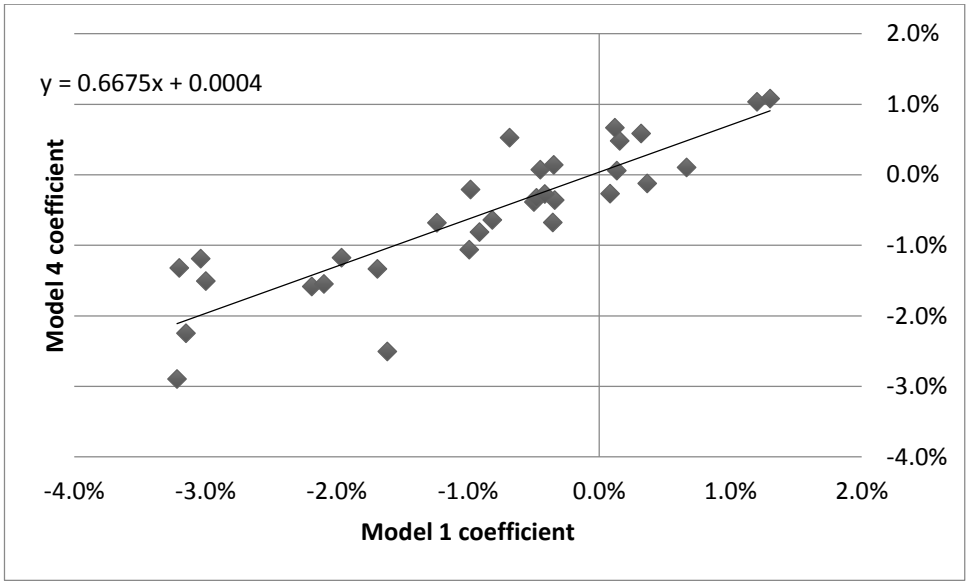


Figure 3.12: Correlation of baseline (M1) and preferred (M4) estimates – administrative activities



Do the effects of the different activities cancel out?

I now turn to the interplay of the individual activities. The estimates for teachers' total weekly workload encompass the effect of more activities than the just the five focused on in this research. Therefore I first need to estimate the joint effect of these five activities, i.e. their linear combination, on teachers' job satisfaction. I do so using the STATA 13 lincom command, which takes into account the variance and covariance of the parameter estimates of each of the five activities (face-to-face teaching, planning lessons, marking, leadership involvement and administrative tasks) to test the null that the effect of the sum of these five activities on teachers' job satisfaction is zero. The results can be found in the column titled "Sum of 5 activities" in Table 3.5, which summarises all the estimates reported in section 3.5 so far; the letter "p" indicates positive and an "n" indicates negative estimates, each statistically significant at the 5 percent level or below.

I first compare the estimates for teachers' total weekly workload with those of the linear combinations. The Table shows that there are 13 negative estimates for total weekly workload, but only 7 for the linear combinations. Yet, for 2 of these seven significant linear combination estimates, Japan and Mexico, the corresponding estimates for total weekly workload are not statistically significant. This implies that the extra activities included in total weekly workload add uncertainty and mask the joint effect of the 5 activities focused on here. In contrast the Table shows significant estimates for both total weekly workload and the linear combinations for the Czech Republic, Denmark, the Netherlands, England and Flanders.

Moreover, nine of the 13 cases with significant estimates for total weekly workload have insignificant estimates for the linear combinations. In six of these cases, Australia, Estonia, France, Latvia, Norway and Sweden, at least one of the individual activities is has a statistically significant estimate. In other words, the individual activities jointly cancel out.

Interestingly, I observe the opposite for Mexico and Japan. Here none of the individual activities has a statistically significant estimate, but jointly they are. England and Flanders are also interesting. The Table shows that teachers in these countries gain job satisfaction from school leadership activities, but their job satisfaction suffers a reduction from their marking and administrative workloads. The linear combinations of these estimates though are negative which suggests that a change in marking and administrative workloads is the driving force.

Table 3.5: Summary table of significant estimates at 95% confidence (M4)

Country	Total Work-load	Sum of activities	Face-to-face teaching	Planning	Marking	School Leader-ship	Admin.
Australia			n				
Brazil							
Bulgaria	n				n		
Chile				p			
Croatia							
Czech Republic		n	n				
Denmark		n	n				
Estonia	n		n		n		
Finland							
France	n		p		n	p	
Iceland							
Israel					n	p	
Italy					n	p	
Japan		n					
Korea	n						
Latvia	n				n		
Malaysia							
Mexico		n					
Netherlands	n	n					
Norway	n		n		n		
Poland							n
Portugal	n		n			p	
Romania							
Serbia							
Singapore	n		n		n		
Slovak Republic	n					p	
Spain	n						
Sweden					n		n
Abu Dhabi (UAE)						p	
Alberta (Canada)							
England (UK)	n	n			n	p	n
Flanders (Belgium)	n	n			n	p	n
Total # significant estimates	**13**	**7**	**8**	**1**	**11**	**8**	**4**

Note: "n" and "p" indicates a significant negative or positive estimate respectively.

Average estimates across education systems: a pooled model

With all this heterogeneity and imprecision among the findings reported above it is hard to make more general claims about the importance of teachers' workload for their job satisfaction. In order to obtain estimates averaged across the 32 education systems, I pool all 99,459 observations together and include country-dummies in

models 1 to 3; the school-fixed effects pick up the both the between-country and between-school variation in M4[23]. By doing so I obtain estimates that account for all the uncertainty in the data. I present the estimates in Table 3.5.

Table 3.6: The association of workloads and job satisfaction (pooled model)

Workload	Model 1 (baseline)	Model 2	Model 3	Model 4 (preferred)
Total weekly	0.1***	-0.2***	-0.2***	-0.1***
	(3.22)	*(-5.08)*	*(-5.07)*	*(-3.27)*
Face-to-face	-0.1	-0.3***	-0.2***	-0.2***
teaching	*(-1.27)*	*(-3.88)*	*(-3.82)*	*(-3.38)*
Planning	-0.1	-0.1	-0.1	-0.2
	(-1.14)	*(-0.66)*	*(-0.70)*	*(-1.33)*
Marking	-0.6***	-0.5***	-0.5***	-0.5***
	(-4.56)	*(-4.16)*	*(-4.18)*	*(-3.84)*
Leadership	1.2***	0.6***	0.6***	0.5**
	(8.32)	*(4.40)*	*(4.45)*	*(2.82)*
Administrative	-0.7***	-0.7***	-0.7***	-0.6***
	(-4.69)	*(-4.48)*	*(-4.48)*	*(-3.11)*
Sum of activities	-0.3	-0.8***	-0.9***	-0.1**
	(-1.22)	*(-3.57)*	*(-3.60)*	*(-3.07)*

Note: Estimates in percent of a standard deviation; T-statistic in parentheses

The findings in Table 3.6 generally point in the expected directions. The estimates for marking, administrative and leadership workloads are all statistically significant and point in the expected directions. Although the estimate for teachers' total weekly workload in M1 indicates a positive relationship (0.1 percent SD) from M2 onwards the estimates are negative (-0.1 percent SD), in other words in the expected direction. A similar pattern is observed for face-to-face teaching workload. Here the estimate in M1 is not statistically different from zero, but the estimate (-0.1 percent SD) is in the expected direction. From M2 onwards the coefficients are statistically significant. In contrast, although the estimates for planning are in the expected negative direction, they are not statistically significant for any of the four models.

Comparing the estimates from the baseline to the preferred model further supports the patterns observed earlier. The change between models 1 and 2 for total weekly workload, face-to-face teaching and leadership activities mirrors the cross-country pattern described above and thus adds further support to the assumption that the baseline estimates proxy teacher-level variables. But there is very little support

[23] The TALIS 2013 school identifier (IDSCHOOL) does not uniquely identify all schools, as some schools in different countries share the same number. I therefore manually compute a unique school identifier combining the country and school ID.

that school-level observed and unobservable variables accounted for in models 3 and 4 are a source of endogeneity in the respective baseline estimates.

But when comparing the change in the linear combination of the five activities (last row) across the four models a clearer pattern emerges. While in M1 the estimate is negative (-0.3 percent SD) but not statistically significant, when controlling for teacher-level covariates the estimate changes to -0.8 percent SD, which is statistically significant at the 0.1 percent level. Controlling for the available school-level covariates, the estimate remains statistically significant and decreases further to -0.9 percent SD. Another jump in coefficients occurs when adding school-fixed effects, which eliminates all observable and unobservable school-level confounders. Here the estimate is -0.1 percent SD, which implies two things. First, it suggests that the unobservable school-level variables appear to be negatively correlated with teachers' job satisfaction. Second, this estimate is of the same magnitude and direction as the coefficient on teachers' total workload, which suggests that these five activities drive the latter estimate.

I now return to the question whether the individual activities may cancel out each other. Comparing the estimates for the individual activities reported in the 6th column (M4) shows that the effect of a one hour increase in leadership workload is associated with a 0.5 percent SD increase in nationally standardised job satisfaction of teachers. The table shows that this change could be approximately counterbalanced by an increase in marking (-0.5 percent SD) and administrative workload (-0.6 percent SD). In contrast a one-hour increase in school leadership activities appears to be able to outweigh a two-hour increase in face-to-face teaching workload.

3.5.3. Summary

This section has focused on cross-country analyses of the workload-job satisfaction relationships and aimed to uncover the emerging cross-country patterns as well as exploring the extent to which corresponding bivariate estimates proxy teacher and school level variables as suggested by equation *(1)*.

The findings presented are marginal effects and represent the effect of a one-hour increase in workload on nationally standardised job satisfaction, so that the obtained estimates are comparable across countries. The findings show that for each of the different kinds of workload examined here, the majority of the 32 individual education systems do not have statistically significant estimates. When

the estimates are significant at or below the 5 percent significance level they tend to be of negligible magnitude. Although a one-hour change is a small change in teachers' workload in most cases, even a one standard deviation change in total weekly workload for example is associated *at the very maximum* with an 8 percent SD decrease in job satisfaction in the case of Flanders.

The findings also support the expectation that teachers may gain job satisfaction from some activities and experience a reduction from others. The statistically significant findings support the expectation that teachers experience a reduction in their job satisfaction from increases in their marking and administrative workloads. Yet although one might expect teachers to enjoy face-to-face teaching, the findings suggest that teachers are less satisfied with their jobs the more lessons they teach. The findings also show that teachers gain job satisfaction from being involved in school leadership. Lesson planning on the other hand appear to be a special case. Although one might expect teachers to enjoy this, as lesson planning involves being creative, this is only the case in Chile.

As the estimates from the cross-country models are heterogeneous and especially most estimates are subject to substantial imprecision, I pool all observed teachers together to obtain averaged estimates across all countries. I include country-dummies in models 1 to 3 to pick up the between-country variation; in M4 this between-country variation is included in the between-school variation that is accounted for by the school-fixed effects. Comparing the estimates for the individual activities obtained from the pooled version of M4 indicates that changes in the individual workloads could be compensated by changes in others. Particularly a one hour increase in school leadership activities could be approximately counterbalanced by a one hour increase in marking (-0.5 percent SD) and administrative workload (-0.6 percent SD). In contrast a one hour increase in school leadership activities appears to outweigh a two-hour increase in face-to-face teaching workload.

Interestingly the findings presented in section 3.5 show that the observable school-level variables are not correlated with teachers' workload. Instead the findings suggest that the unobservable variables are. In section 3.6 I will therefore move the focus to the English sample, which contains relevant additional data. First, the English sample is linked to the English National Pupil Database (NPD), so that extra administrative school-level information is available, which might be part of

the unobserved variables in the school-fixed effects in section 3.5. The English sample also contains information on teachers' perceptions of their workplace and the stakeholders in them as well as perceptions of their job in society. Although these variables represent the individual teachers' subjective perspective, they are nevertheless valuable as they relate to the school-level and are consistent with the adopted production function approach formulated in equation (1).

3.6. TALIS England 2013

3.6.1. The data

As already hinted above, TALIS England 2013 contains additional information that is of interest here. In the previous discussion of the cross-country data, the school-level variables included in M3 are head teachers' age, gender, self-reported leadership styles as well as teacher to administrative or pedagogic support staff ratios. Before my obtaining access to the English data, it was linked to the English National Pupil Database (NPD) so that it contains a few additional school-level variables:

- School type,
- Ofsted rating at the most recent inspection, as a measure of overall school quality as measured by the independent Office for Standards in Education (Ofsted). Ofsted classify schools into one of the following four categories: outstanding, good, satisfactory and inadequate.
- The percentage of pupils receiving free school meals (FSM) as an indicator of poverty,
- The percentage of pupils achieving 5 A* to C grades in their age 16 General Certificate of Secondary Education (GCSE),
- The average Key Stage 2 test score of pupils entering the school as an indicator of pupils' ability.

In short, these school level data cover a wider range of stakeholders in a school. The Ofsted rating in particular may be more valid than the head teachers' self-reported leadership styles, as this survey data may represent their aspirations rather than their true leadership styles. In spite of these advantages, the additional NPD variables may be weak proxies. For example, a poverty index such as the proportion of pupils receiving free school meals may be a poor indicator of the parental behaviour the teachers might be exposed to.

The English TALIS sample also contains information on teachers' perceptions of their workplace and the stakeholders in them as well as perceptions of their job in society. The data was collected as an additional question, teacher question 47, and is unique to England. Although these variables represent the individual teachers' subjective perspective, they complement the previous analysis reported in section 3.5 as they relate to the school-level. They are consistent with the adopted production function approach formulated in equation (1), because they combine both the bottom-up approach, that individuals' job satisfaction is the effect of external influences such as the working environment they are in, and the top-down approach, which assumes that individual's perceive the environment they are in differently due to different psychological factors. In equation (1) these variables are elements of X_{ij}.

For the selected items teachers are asked to rate their agreement with a statement on a 4-point Likert scale ranging from 1 representing "strongly disagree" to 4 representing "strongly agree". For the purpose of this research these items are collapsed to two categories representing agreement and disagreement. These statements are:

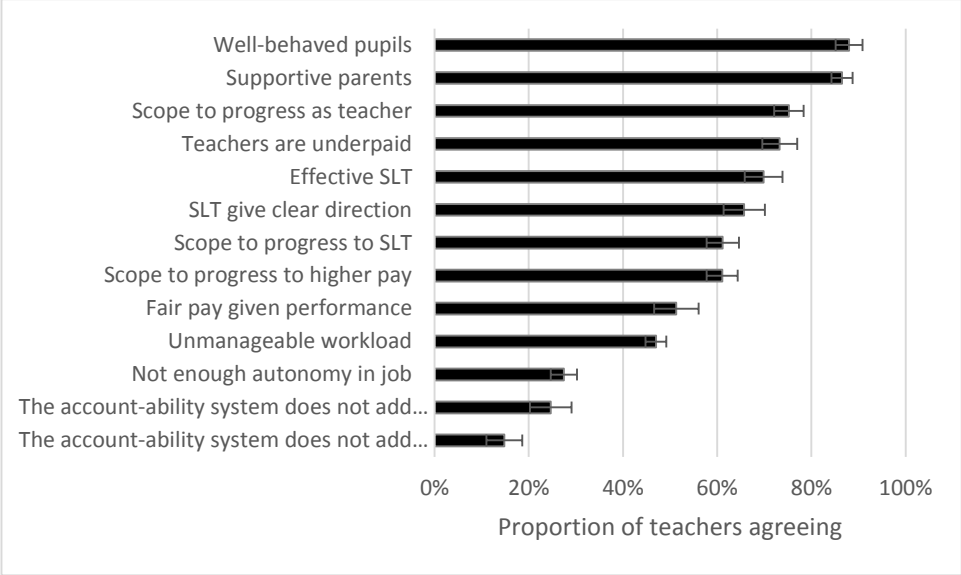
1. The students I teach are generally well behaved.
2. The school has an effective school management team.
3. The school management team gives clear vision and direction.
4. I do not have the autonomy I need to do a good job as a teacher.
5. Teachers are underpaid compared to other qualified professionals with similar levels of responsibility.
6. My own pay is fair given my performance.
7. The accountability system (Ofsted, league tables, etc.) does not add significantly to the pressure of my job.
8. Parents are supportive of my role as their children's teacher.
9. The accountability system does not add significantly to my workload.
10. I have scope to progress as a classroom teacher.
11. I have scope to progress into a leadership role.
12. I have scope to progress to a higher pay level.
13. My workload is unmanageable.

In this sample of teachers in England, 27 percent work in "outstanding", 44 percent in "good", 25 percent in "satisfactory" and approximately 4 percent work in

"inadequate" schools as deemed by Ofsted. In 2013 the English education system consists of a number of different kinds of schools, which can be summarised into three broader categories: Academies, Independent Schools and Maintained Schools. In this sample approximately 45 percent of teachers work in Academies, 6 percent in Independent (private) Schools and the remaining 49 percent in Maintained Schools. At the point of data collection, Maintained Schools could become Academies. Both school types are government-funded, but the latter are not administered by Local Authorities, enjoy more legal freedom, for example in respect to staffing and do not need to follow the National Curriculum. On average 18 (95 percent C.I. 16 to 21) percent of pupils in a school receive free school meals and can therefore be considered living in relative poverty. On average 63 (95 percent C.I. 59 to 67) percent of pupils achieve five or more A star to C grades in their GCSE exams at age 16.

Figure 3.13 shows the sample average proportions of teachers agreeing with the 13 statements listed above. The estimates are ranked according to their magnitude, and the error bars indicate the 95 percent confidence interval. I will integrate the results from further analyses into the discussion of this graph to show that teachers' perceptions correlate with characteristics of their workplace. In this sub-sample 47 percent of teachers agree that their workload is unmanageable. Further analysis shows that these teachers work approximately 3.7 hours more per week on average than their peers who disagree. This estimate appears to be robust to contractual status, as controlling for the latter reduces the estimate only slightly, so that teachers who believe their workload is unmanageable work 3.5 hours more per week than their peers who disagree.

Figure 3.13: Proportions of teachers in England agreeing with the following statements



Note: Standard errors adjusted for clustering by school.

Figure 3.13 also shows that 88 percent of teachers perceive their pupils to be well behaved. Additional analyses of the data indicate that this perception differs by the proportion of FSM pupils. On average teachers who agree that their pupils are well behaved tend to work in schools with 5 percentage points less FSM pupils. Similarly the proportion of teachers feeling supported by the pupils' parents also varies by the proportion of FSM pupils in the school. While 85 percent of teachers agree on average, agreement is highest in the second (90 percent) and first (87 percent) quartile of the distribution of FSM, i.e. in schools with lower proportions of poor pupils. Only 82 and 84 percent of teachers feel supported by their pupils' parents in the third and fourth FSM quartiles. These differences are significant at the 0.1 percent level.

Figure 3.13 also shows that 70 percent of teachers believe that their senior leadership team (SLT) is effective. I explore if this belief resonates with Ofsted ratings, as these are supposed to indicate the quality of a school. A chi-squared test suggests that while 95 percent of teachers in outstanding and good schools believe their SLT is effective, only 75 percent do so in satisfactory and inadequate schools. The reader should be reminded though that teachers may in part base their answers on their school's Ofsted rating.

The accountability system itself, of which Ofsted ratings are a central part, may also affect teachers' perceived workload and pressure. One would expect teachers in satisfactory and outstanding schools to be more affected than their peers in

better ranked schools, as satisfactory and inadequate schools will be desperate to achieve a better rating. The data shows that on average only a minority feel no additional pressure (12 percent) or workload (21 percent). As expected the percentage of teachers feeling this way is significantly lower in satisfactory and inadequate schools for both perceived workload (18 compared to 22 percent) and pressure (9 compared to 13 percent).

On average 73 percent of teachers believe they are underpaid compared to professions with similar levels of responsibility and 53 percent feel underpaid given their performance. I explore whether these beliefs differ across school types, as the pay teachers receive is likely to differ between public and private schools, and the recent Academy schools are more independent than traditional state schools. The data though suggests that these beliefs do not differ by school type.

3.6.2. Controlling for teachers' perceptions of school-level confounders

I now explore how the estimates for England reported in section 3.5 change, when adding the NPD variables and teachers' perceptions variables available in TALIS England 2013. I follow the stepwise approach outlined in section 3.4. I add a variation to M3, the model controlling for observed teacher and school characteristics, which I name M3b. Thus, this model controls for part-time status, the teachers' involvement in mentoring, the degree of cooperation with their colleagues, whether or not they are on a permanent contract, their gender and their self-efficacy at the teacher level, and the degree of distributional and instructional leadership, as well as the head teacher's gender, age as well as the ratio of teachers per administrative or pedagogic member of support staff *in addition to* the NPD variables described above. I also run M4b, which includes school-fixed effects, all the teacher-level variables included in M3b as well as the twelve dummy variables indicating teachers' perceptions of their job.

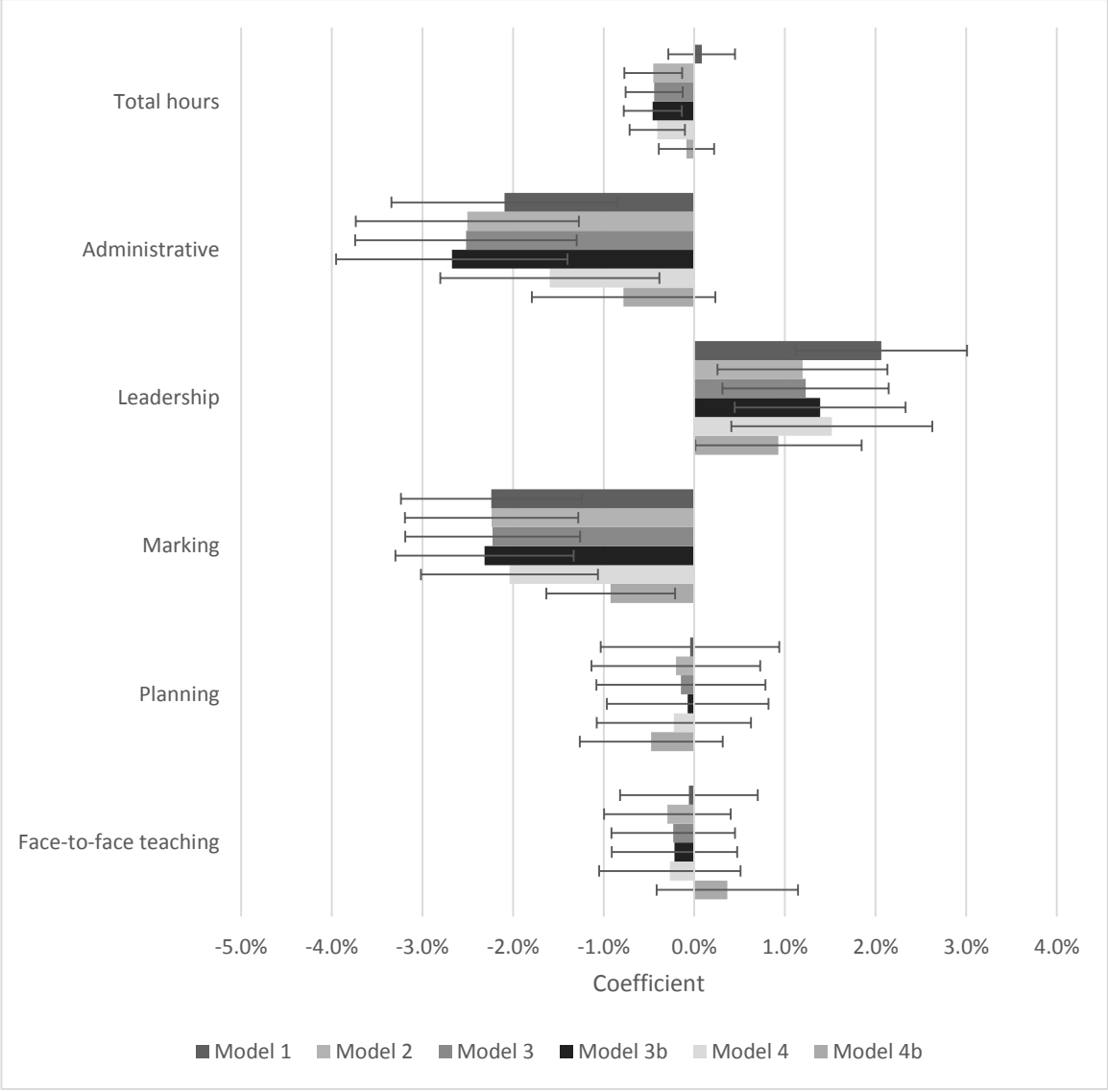
Figure 3.14 summarises the estimates from all six models for teachers' total workload and the individual activities. As before in section 3.5, the error bars indicate the 95 percent confidence interval. The fourth bar in descending order for each group of estimates, represents M3b; the last bar represents M4b. Comparing models 3 and 3b one can see that the estimates only vary slightly, as in model 3 the observable school-level variables are not correlated with teachers' workload. Moreover, none of the NPD variables are significantly correlated with job satisfaction. In model 3 only one school-level TALIS 2013 variable, the head

teachers' distributed leadership style, is significantly associated with teachers' job satisfaction. Here a one standard deviation increase in distributed leadership increases teachers' job satisfaction by 0.06 SD. This association though becomes statistically insignificant in model 3b. It may be the case because head teacher leadership styles proxy similar behaviours to the Ofsted rating variable. To test this, I remove the head teacher leadership style variables from the model. The estimates (not shown) remain identical, as still all NPD variables do not significantly predict teachers' job satisfaction[24].

Comparing models 4 and 4b model fit improves significantly at the 0.1 percent significance level. Jointly the teacher perception variables added in M4b are associated positively with teachers' job satisfaction. These perceptions appear to be negatively associated with teachers' total, administrative and face-to-face teaching workload, as their corresponding estimates are less negative or more positive in M4 and the estimates for total and administrative workload are not statistically significant anymore. In contrast these perceptions appear to be positively associated with leadership and planning workloads as the estimates in M4 are less positive or more negative. For example, the coefficient on teachers' marking workload falls in absolute size from -2 to -0.9 percent of a SD.

[24] Linear combination test yields a t-statistic of 1.41, and model fit does not improve significantly at the 5 percent level as the F-test yields a p-value of 0.06.

Figure 3.14: Comparing models for the workload-job satisfaction relationship in England



Note: 95 percent confidence interval based on standard errors that are adjusted for clustering by school.

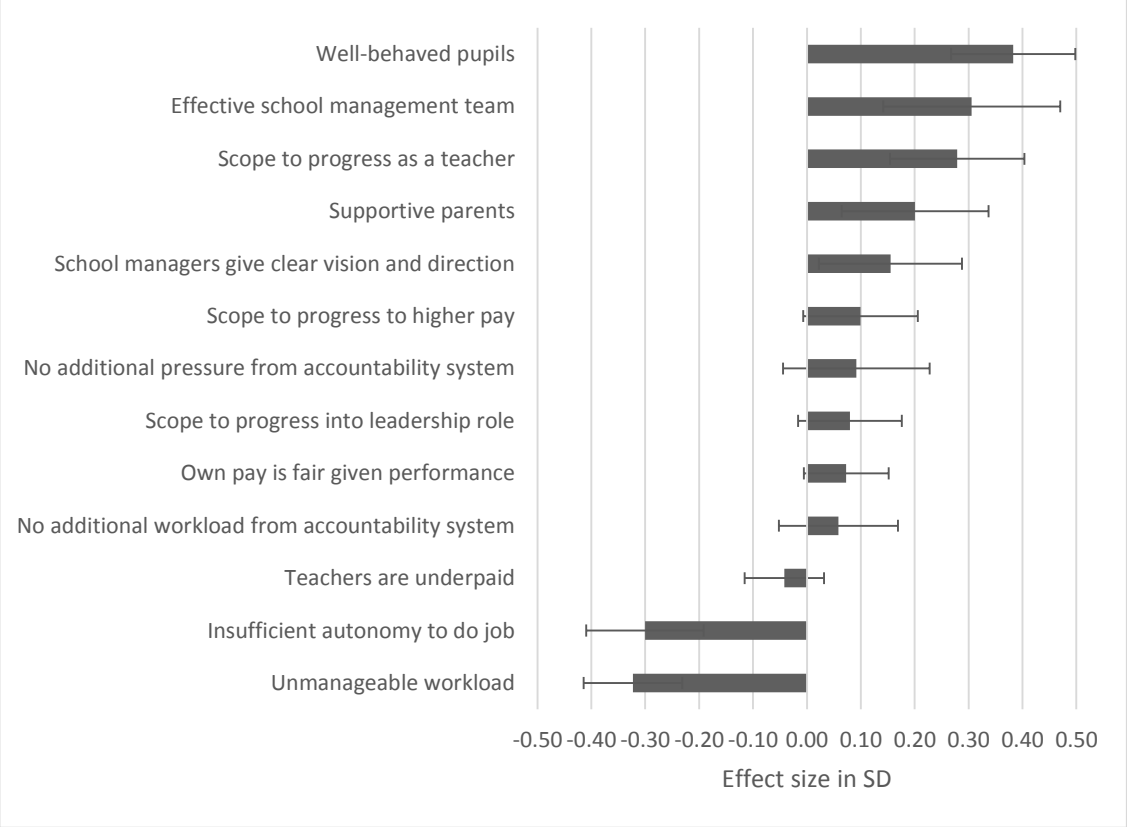
3.6.3. The association of teachers' perceptions of their job and their job satisfaction

Having established so far that the association of teachers' workload and their job satisfaction is weak at best, I now focus on teachers' perceptions of their job and how these are associated with their job satisfaction. The estimates reported in Figure 3.15 are obtained from M4b. As before, the error bars indicate the 95 percent confidence interval.

Model 4b explores the within-school variation, i.e. the variation between teachers in the same school. It controls for part-time status, the teachers' involvement in mentoring, the degree of cooperation with their colleagues, whether or not they are on a permanent contract, their gender and their self-efficacy at the teacher level.

As the variables in question are binary, the estimates indicate the average change in job satisfaction when agreeing with the respective statement, conditional on the other covariates. All the perception variables are added at once to form M4b. The reader should be reminded at this point that although the reported estimates are purged of any between-school variation, they still are only correlations and should not be considered causal. The estimates may for example be subject to endogeneity, as "grumpy" teachers may both indicate low levels of job satisfaction and tend to perceive their school and the other stakeholders within it more negatively.

Figure 3.15: The association of teacher perceptions and their job satisfaction in England



Note: 95 percent confidence interval based on standard errors that are adjusted for clustering by school. Estimates obtained from model 4b, which includes school-fixed effects as well as part-time status, the teachers' involvement in mentoring, the degree of cooperation with their colleagues, whether or not they are on a permanent contract, their gender and their self-efficacy.

As before in models M2 and M4, including teachers' perceptions again increases overall model fit in M4b. Compared to M4, the adjusted R^2 is 17 percentage points higher in M4b so that overall 37.6 percent of the variation in job satisfaction is explained in this sample.

I rank the estimates in ascending order according to their magnitude. This hierarchy suggests that how teachers perceive their job is valued in society, as indicated by the question "teachers are underpaid compared to other qualified

professionals with similar levels of responsibility", does not statistically significantly predict their job satisfaction. Instead the findings show that teachers appear to value well-behaved pupils (0.38 SD), having sufficient autonomy to do their job (0.30 SD), having the scope to progress as a teacher (0.28 SD) the most, as well as an effective school-leadership team (0.31 SD).

Further, teachers appreciate supportive parents (0.19 SD), a school management team that gives clear vision and direction (0.16 SD) and having the scope to progress to higher pay (0.13 SD). Interestingly, the accountability system appears to affect teachers' job satisfaction through causing a perceived greater workload (0.12 SD), but not through causing perceived increased pressure. Similarly perceived scope to progress into school leadership positions is not associated with a significant change in job satisfaction. The findings also show that although teachers who feel they are receiving fair pay given their performance are more satisfied with their job, it is the weakest of all the statistically significant predictors at an estimated 0.09 SD.

More importantly, the findings clearly show that teachers who believe their workload is unmanageable are 0.32 SD less satisfied with their job than their peers who disagree. As this estimate is obtained from model 4b, the estimate is not confounded by any school-level factors, as well as the actual amount of hours the teachers work, part-time status, the teachers' involvement in mentoring, the degree of cooperation with their colleagues, whether or not they are on a permanent contract, their gender and their self-efficacy.

3.6.4. Summary

Following the pattern of section 3.5 this section has focused on England and first exploits additional school level information obtained from a link with the National Pupil Database, which are school type, the proportion of pupils receiving free school meals, the school's Ofsted rating, the average pupil Key Stage 2 score and the proportion of pupils achieving 5 grade A* to C in their GCSE. Adding these NPD variables to those in M3, the former do not correlate significantly with job satisfaction and therefore do not change the estimates compared to the model omitting the NPD variables.

The England data also contains information on teachers' perceptions of their workplace and their role as a teacher in society. When adding these variables to the fourth model, which includes school-fixed effects and a number of teacher-level

characteristics and behaviours, significant estimates in the previous models disappear, which supports the assumption that teachers' perceptions of their workplace and people in them are sources of bias.

Moreover, the data shows that teachers' perceptions of their workplace, the people in it as well as of their status in society are strongly associated with their job satisfaction. The obtained estimates range from 9 to 39 percent of a standard deviation and are thus much stronger than the association of their workload. Most importantly teachers seem to value well behaved pupils, sufficient autonomy to do their job and scope to progress as a teacher. An effective school leadership team is similarly important to them, but supportive parents, school leaders providing clear vision and direction, scope to progress to higher pay and fair pay given a teacher's performance are also significant but increasingly less important. Last, the results clearly show that teachers who believe that their workload is unmanageable are 0.32 SD less satisfied than their peers and this estimate is robust to all potential school-level confounders as well as a number of teacher-level observable characteristics.

3.7. Conclusion

This paper is the first of its kind to explore the connection of teachers' workload and their job satisfaction. Before using cross-sectional, observational data from 32 education systems from the OECD's TALIS 2013 dataset, the paper applies a production function approach to job satisfaction. The main advantage of this approach is that it unifies the two prevailing streams of literature, namely Top-Down and Bottom-Up approaches (cf. Diener, 1984). Thus it combines the notion that a teacher's job satisfaction might be determined by the environment he or she is in (Bottom-Up) with the notion that individuals process their environment differently (Top-Down). Based on this production function I identify potential groups of stakeholders that could both directly affect teachers' job satisfaction as well as their workload and thereby bias endogenously bivariate OLS estimates of the workload – job satisfaction relationship. These groups of stakeholders are the teachers themselves, the pupils, the parents, the other teachers in the school, the head teacher and the physical resources available in the school.

The paper then applies this production function to 32 education systems sampled in the OECD's TALIS dataset collected in 2013. It applies a stepwise regression approach, increasing the number of variables of either the teacher or school-level in

4 steps. The first specification is a bivariate OLS model, the final preferred model exploits the fact that multiple colleagues are observed in the same school and contains school-fixed effects to capture observable and unobservable school-level variables as well as part-time status, the teachers' involvement in mentoring, the degree of cooperation with their colleagues, whether or not they are on a permanent contract, their gender and their self-efficacy.

This paper further contributes to the literature being the first to consider the association of not only total weekly workload, but also explores heterogeneity of effects for five different activities teachers do while working; face-to-face teaching hours, time spent planning, marking, doing administrative tasks as well as the time they are engaged in school leadership activities.

The findings presented are marginal effects and represent the effect of a one-hour increase in workload on nationally standardised job satisfaction, so that the obtained estimates are comparable across countries. The findings show that across the different kinds of workload examined here, the majority of the 32 individual education systems do not have statistically significant estimates. When the estimates are significant at or beyond the 5 percent significance level they tend to be of negligible magnitude. Although a one-hour change is a small change in teachers' workload in most cases, even a one standard deviation change in total weekly workload for example is associated at *maximum* with an 8 percent SD decrease in job satisfaction in the case of Flanders. The findings also support the expectation that teachers may gain job satisfaction from some activities (leadership) and experience a reduction from others (marking, administrative tasks).

As the estimates from the cross-country models are heterogeneous and especially most estimates are subject to substantial imprecision, this paper runs pooled models to obtain averaged estimates across all countries. The evidence indicates that a one hour increase in school leadership activities could be approximately counterbalanced by a one hour increase in marking (-0.5 percent SD change in job satisfaction) and administrative workload (-0.6 percent SD) and could outweigh a two-hour increase in face-to-face teaching workload.

The paper then moves its focus to England and uses additional administrative school-level information obtained from a link with the English National Pupil Database, which are school type, the proportion of pupils receiving free school

meals, the school's Ofsted rating, the average pupil Key Stage 2 score and the proportion of pupils achieving 5 grade A* to C in their GCSE. Adding these variables to part-time status, the teachers' involvement in mentoring, the degree of cooperation with their colleagues, whether or not they are on a permanent contract, their gender and their self-efficacy at the teacher level, and the degree of distributional and instructional leadership, as well as the head teacher's gender, age as well as the ratio of teachers per administrative or pedagogic member of support staff, the NPD variables do not correlate significantly with job satisfaction and therefore do not change the estimates compared to the model omitting the NPD variables.

In contrast to the negligible estimates for teachers' workloads and their job satisfaction, the English data suggests that teacher perceptions of their workplace and the people in them are strongly associated with their job satisfaction and estimates range from 9 to 39 percent of a standard deviation. Most importantly teachers value well behaved pupils, sufficient autonomy to do their job and scope to progress as a teacher. An effective school leadership team is similarly important to them, but supportive parents, school leaders providing clear vision and direction, scope to progress to higher pay and fair pay given a teacher's performance are also statistically significant, but increasingly less important. Moreover, the results clearly show that teachers who believe that their workload is unmanageable are 0.32 SD less satisfied than their peers and this estimate is robust to all potential school-level confounders as well as a number of teacher-level observable characteristics. Future research should focus on which unobserved teacher-level variables in the research presented here explain teachers perceiving their workload as unmanageable.

The policy implications of these findings are that changes in teachers' workloads only have a small impact on their job satisfaction. Yet the school level does appear to play an important role determining teachers' job satisfaction. The findings suggest that their school leaders play a central role, as they can help bridge the gap between parents and teachers so that the parents are supportive of their children's teachers. School leaders can also help incentivise the pupils to better behaviour, can help teachers develop professionally, and may be able to develop means to reduce teachers' perceived workload caused by accountability systems. The findings also show that school leaders and their leadership teams play an important role in determining their teachers' job satisfaction, and that teachers

particularly value effective school management teams that give a clear vision and direction.

Concluding remarks

In this final section of the thesis, I would like to take the opportunity to reflect on and highlight the contribution of the presented work as well as potential avenues for future research.

- The first chapter delves into explaining why even the most robust estimates for class size effects from around the world do not provide a conclusive picture. It is often argued that pupil achievement will decrease as class sizes increase. I challenge this widespread belief and argue that a change in class size will have both a compositional and a resource usage effect. The compositional effect consists of a change in peer dynamics, as pupils will, quite literally, be either added to or removed from the class, (depending on the direction of the change in class size), and a household effect, whereby parents may choose different schools or adapt their academic support for their children. Within schools physical resources such as schoolbooks, blackboards, desks and chairs and teachers need to be allocated to pupils, and this allocation process may change with the number of pupils. Drawing on data from a range of sub-Saharan African countries, I provide evidence that these different mechanisms can indeed have opposing directions and thus may cancel out each other. This can explain the observation that the gap between pupils in larger and smaller class sizes does not grow after the first year, as the households the children are in may adapt their behaviour (cf. Nye et al, 1999).

- The resource usage effect also has direct implications for the specification of education production functions (cf. Hanushek, 1979). Production functions have been used by economists for many years to relate inputs to outputs. Education production functions are adaptations of this approach to educational contexts. Thus it is understandable that physical resources such as schoolbooks are considered a separate input from all others. In contrast, I argue that physical classroom resources such as schoolbooks do not have an "innate" effect on pupil achievement of their own. Rather their effect depends on how the teacher uses them. Analogously the effect of school-level resources is the result of how school leaders use them. In short, education production functions should consider including physical school and

classroom resources as elements of the school leader or classroom teacher. Thus a change in how teachers allocate schoolbooks to pupils can explain why in Kenya over time schoolbooks do not appear to affect pupil achievement anymore (cf. Glewwe et al, 2009)

- The second chapter's contribution to the field is mainly methodological. In this chapter I argue that previous research analysing the association of teachers' observable characteristics and pupil achievement in sub-Saharan African contexts do not sufficiently address either the non-random matching of teachers to schools or the matching of pupils and teachers within them and argue that allowing pupil-fixed effects is the best way to address this issue. I contrast my estimates with school-fixed effects models for each subject that only account for the non-random matching of pupils and teachers to schools. The results from these different types of model contrast rather sharply. Two conclusions may be drawn from this: 1) pupil teacher-matching within schools is also a non-random process and 2) the findings from previous studies using the same data are not suitable bases for policy making. The emerging patterns of my pupil-fixed effects models also differ from the US patterns, which suggests that neither should US evidence automatically be used as a basis for policy in sub-Saharan African countries, because of their different contexts. Together with the US evidence, my findings contribute to an emerging picture, which suggests that although teachers matter, they differ in their quality predominantly due to their unobservable characteristics.

- The second chapter breaks new ground by categorising teachers' observable characteristics into proxies for two different kinds of teachers' skills and then explores whether they are complementary to one another or if they are substitutes. While the findings here focus on African countries, this kind of analysis could also be done in a developed country context, opening up a new avenue for research aimed at helping teachers maximise their effectiveness. Future research might extend to exploring interactions of these different teacher skills with different incentives as well as proxies for their motivation.

- The third chapter reports the first study to explicitly examine the association of teachers' workload and their job satisfaction and to do so

using cross-country data sampling 32 distinct education systems. In this chapter I formulate a production function that combines the two prevailing strands of research in occupational psychology. This approach combines the notion that a teacher's job satisfaction might be determined by the environment he or she is in (Bottom-Up) with the notion that individuals process their environment differently (Top-Down). Based on this production function I identify potential sources for endogeneity of bivariate OLS estimates of the workload – job satisfaction relationship when using observational data. These potential sources of bias are the teachers themselves, the pupils, the parents, the other teachers in the school, the head teacher as well as the physical resources available in the school. In addition, I consider not only the association of total weekly workload with job satisfaction, but also explore heterogeneity of effects for five different activities teachers do while working; face-to-face teaching hours, time spent planning, marking, doing administrative tasks as well as the time they are engaged in school leadership activities. Interestingly there is no consistent cross-country pattern supporting the expectation derived among others from the UK media that teachers' workload affects their job satisfaction to a worrying degree. In most countries the estimated effects are not significant, and those that are, appear to be of negligible magnitude.

- I then move on to using data of teachers' perceptions of their job and their workplace available only in the English sub-sample. I find while holding constant the school-level and a number of observed teacher characteristics that teachers who believe their workload is unmanageable are almost a third of a standard deviation less satisfied with their job than their peers who do not share this view. It is hoped that future research will explore further which teacher-level variables that could not be included in this research here affect teachers perceiving their workload as unmanageable in order to understand better how teachers' wellbeing at work can be improved.

- In addition, using the English data on teachers' perceptions of their workplace shows that the teacher's working environment, comprising their school and its staff, indeed affects their job satisfaction. I find that teachers value well-behaved pupils, sufficient autonomy to do their job and scope to progress as a teacher. An effective school leadership team is similarly

important to them, but supportive parents, school leaders providing clear vision and direction, scope to progress to higher pay and fair pay given a teacher's performance are also statistically significant, but of decreasing importance. It is to be hoped that future research will follow this avenue. The next round of OECD TALIS surveys could include similar questions in their main survey in order to compare these findings across countries. The UK School Workforce Census could also include such questions. This would enable monitoring of how policy changes affect teacher job satisfaction and provide an opportunity to explore causality. As will have been abundantly clear throughout this thesis, good data are essential if educational researchers are to robustly identify causal effects.

References

Aaronson, D., Barrow, L. and Sander, W. (2007). 'Teachers and student achievement in the Chicago public high schools'. *Journal of Labour Economics,* 25 (1), 95 - 135.

Albion, P. (1999). *Self-efficacy beliefs as an indicator of teachers' preparedness for teaching with technology*. Paper presented at the Proceedings of the 10th International Conference of the Society for Information Technology & Teacher Education (SITE 1999).

Angrist, J. D. and Lavy, V. (1999). 'Using Maimonides' rule to estimate the effect of class size on scholastic achievement'. *The Quarterly Journal of Economics,* 114 (2), 533-575.

Asadullah, M. N. (2005). 'The effect of class size on student achievement: evidence from Bangladesh'. *Applied Economics Letters,* 12 (4), 217-221.

Aslam, M. and Kingdon, G. (2011). 'What can teachers do to raise pupil achievement?'. *Economics of Education Review,* 30 (3), 559-574.

Baird, J., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T. and Daugherty, R. (2011). 'Policy effects of PISA'. *Report commissioned by Pearson UK*

Bandura, A. (1986). 'The Explanatory and Predictive Scope of Self-Efficacy Theory'. *Journal of Social and Clinical Psychology,* 4 (3), 359-373.

Bennell, P. (2004). 'Teacher motivation and incentives in Sub-Saharan Africa and Asia'.

Bennell, P. (2005). 'The impact of the HIV/AIDS epidemic on the education sector in Sub-Saharan Africa: a synthesis of the findings and recommendations of three country studies'. [Online],

Bentler, P. M. and Chou, C.-P. (1987). 'Practical issues in structural modeling'. *Sociological Methods & Research,* 16 (1), 78-117.

Bogler, R. (2001). 'The influence of leadership style on teacher job satisfaction'. *Educational Administration Quarterly,* 37 (5), 662-683.

Bonesrønning, H. (2004). 'The determinants of parental effort in education production: do parents respond to changes in class size?'. *Economics of Education Review,* 23 (1)*,* 1-9.

Bourdon, J., Frölich, M. and Michaelowa, K. (2005). *Broadening Access to Primary Education: Contract Teacher Programs and their Impact on Education Outcomes in Africa – An Econometric Evaluation for the Republic of Niger –*. Paper presented at the *Annual Conference of the Research Committee Development Economics of the German Economic Association on "Pro-Poor Growth"*. Kiel Institute of World Economics.

Bray, M. (2008). *Double-shift schooling: design and operation for cost-effectiveness*. (Vol. 90): Commonwealth Secretariat.

Caprara, G. V., Barbaranelli, C., Steca, P. and Malone, P. S. (2006). 'Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level'. *Journal of School Psychology,* 44 (6)*,* 473-490.

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K. and Rogers, F. H. (2006). 'Missing in Action: Teacher and Health Worker Absence in Developing Countries'. *Journal of Economic Perspectives*, 20 (1)*,* 91-116.

Clotfelter, C. T., Ladd, H. F. and Vigdor, J. L. (2006). *Teacher-student matching and the assessment of teacher effectiveness*. Cambridge, Mass.: National Bureau of Economic Research.

Clotfelter, C. T., Ladd, H. F. and Vigdor, J. L. (2007). 'Teacher credentials and student achievement: Longitudinal analysis with student fixed effects'. *Economics of Education Review,* 26*,* 673 - 682.

Collie, R. J., Shapka, J. D. and Perry, N. E. (2012). 'School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy'. *Journal of Educational Psychology,* 104 (4)*,* 1189.

Cuseo, J. (2007). 'The empirical case against large class size: adverse effects on the teaching, learning, and retention of first-year students'. *The Journal of Faculty Development,* 21 (1)*,* 5-21.

Das, J. and University of Cambridge. Faculty of Economics. (2005). *Teacher shocks and student learning: evidence from Zambia*. Cambridge: University of Cambridge, Faculty of Economics.

Datar, A. and Mason, B. (2008). 'Do reductions in class size "crowd out" parental investment in education?'. *Economics of Education Review,* 27 (6), 712-723.

Davis-Kean, P. E. (2005). 'The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment'. *Journal of Family Psychology,* 19 (2), 294.

Deci, E. L. and Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Dee, T. S. (2005). 'A teacher like me: Does race, ethnicity, or gender matter?'. *American Economic Review,* 158-165.

Department for Education. (2014). Teachers' workload diary survey 2013: Research report. London.

Dickerson, A., McIntosh, S. and Valente, C. (2013). 'Do the Maths: An Analysis of the Gender Gap in Mathematics in Africa'.

Diener, E. (1984). 'Subjective Well-being'. *Psychological Bulletin*, 95, 542-575

Diener, E. (2000). 'Subjective well-being: The science of happiness and a proposal for a national index'. *American psychologist,* 55 (1), 34.

Dolata, S. (2005). Construction and validation of the pupil socioeconomic status index for SACMEQ education systems: SACMEQ.

Duflo, E., Dupas, P. and Kremer, M. (2008). 'Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya'. *National Bureau of Economic Research Working Paper Series,* No. 14475.

Faragher, E. B., Cass, M. and Cooper, C. L. (2005). 'The relationship between job satisfaction and health: a meta-analysis'. *Occupational and environmental medicine,* 62 (2), 105-112.

Fehrler, S., Michaelowa, K. and Wechtler, A. (2009). 'The Effectiveness of Inputs in Primary Education: Insights from Recent Student Surveys for Sub-Saharan Africa'. *Journal of Development Studies,* 45 (9), 1545-1578.

Fischer, J. A. and Sousa-Poza, A. (2009). 'Does job satisfaction improve the health of workers? New evidence using panel data and objective measures of health'. *Health Economics,* 18 (1), 71-89.

Fyfe, A. (2007). *The use of contract teachers in developing countries: Trends and impact*: International Labour Organization.

Gandhi Kingdon, G. (2002). 'The Gender Gap in Educational Attainment in India: How Much Can Be Explained?'. *The Journal of Development Studies,* 39 (2), 25-53.

Gary-Bobo, R. J. M. and Mahjoub, M. B. (2006). *Estimation of Class-Size Effects, Using' Maimonides' Rule': The Case of French Junior High Schools*: CEPR Discussion Papers.

Glewwe, P., Kremer, M. and Moulin, S. (2009). 'Many children left behind? Textbooks and test scores in Kenya'. *American Economic Journal: Applied Economics,* 1 (1), 112-135.

Government of Seychelles Ministry of Education and Youth. (2002). Education for All National Action Plan 2001-2015: Meeting our collective commitments.

Government of the Republic of Kenya. (2005). Millennium Development Goals Status Report.

Government of the Republic of Kenya (2008). Education Statistical Booklet 2003-2007. In M. o. Education (ed). Nairobi.

Government of the Republic of Malawi. (2001). Malawi Education Sector: Policy and Investment Framework (PIF). In M. o. Education (ed).

Government of the Republic of Mauritius. (2001). Millennium Development Report. In M. o. Education (ed).

Government of the Republic of Namibia. (2002). Education For All (EFA) National Plan of Action 2002-2015. In S. a. C. Ministry of Basic Education (ed).

Government of the Republic of Namibia. (2004). Namibia 2004 Millennium Development Goals.

Government of the Republic of Namibia. (2008). 2nd Millennium Development Goals Report Namibia 2008.

Government of the Republic of South Africa. (2005). Millennium Development Goals Country Report.

Government of the Republic of Zambia. (2003). Strategic Plan 2003-2007. In M. o. Education (ed).

Government of the Republic of Zambia. (2005a). Educating the Nation: Strategic Framework for Implementation of Education For All. In E. S. Ministries (ed).

Government of the Republic of Zambia. (2005b). Millennium Development Goals Status Report.

Government of the United Republic of Tanzania. (2001). Education Sector Development Programme: Primary Education Development Plan (2002-2006). In B. E. D. Committee (ed).

Government of the United Republic of Tanzania. (2006). Education and Training Sector Development Programme: Primary Education Development Programme II (2007-2011). In B. E. D. Committee (ed).

Hanushek, E. A. (1979). 'Conceptual and Empirical Issues in the Estimation of Educational Production Functions'. *The Journal of Human Resources,* 14 (3), 351-388.

Hanushek, E. A. (2003). 'The failure of input-based schooling policies'. *The Economic Journal,* 113 (485), F64-F98.

Hanushek, E. A. (2005). *The market for teacher quality*. Cambridge, Mass.: National Bureau of Economic Research.

Hanushek, E. A., Kain, J. F., Markman, J. and Rivkin, S. (2003). 'Does peer ability affect student achievement?'. *Journal of Applied Econometrics,* 18 (5), 527-544.

Harris, A. and Muijs, D. (2004). *Improving schools through teacher leadership*: McGraw-Hill International.

Hein, C. F. and Allen, R. (2013). *Teacher Quality in Sub-Saharan Africa: Pupil-fixed effects estimates for twelve countries*: Department of Quantitative Social Science-Institute of Education, University of London.

Hellman, C. M. (1997). 'Job Satisfaction and Intent to Leave'. *The Journal of Social Psychology,* 137 (6), 677-689.

Hulin, C. L., Roznowski, M. and Hachiya, D. (1985). 'Alternative opportunities and withdrawal decisions: Empirical and theoretical discrepancies and an integration'. *Psychological Bulletin,* 97 (2), 233.

Jakubowski, M. and Sakowski, P. (2006). 'Quasi-experimental estimates of class size effect in primary schools in Poland'. *International Journal of Educational Research,* 45 (3), 202-215.

Jarousse, J. and Mingat, A. (1991). 'Les facteurs agissant sur les acquisitions des élèves à l'école primaire au Burkina Faso'. *IREDU-CNRS, Université de Bourgogne*.

Judge, T. A. and Watanabe, S. (1993). 'Another look at the job satisfaction-life satisfaction relationship'. *Journal of applied psychology,* 78 (6), 939.

Kahneman, D. and Krueger, A. B. (2006). 'Developments in the measurement of subjective well-being'. *The journal of economic perspectives,* 20 (1), 3-24.

Kalisch, B., Tschanen, D. and Lee, H. (2011). 'Does missed nursing care predict job satisfaction?'. *Journal of Healthcare Management,* 56 (2), 117.

Kalliath, T. and Morris, R. (2002). 'Job Satisfaction Among Nurses: A Predictor of Burnout Levels'. *Journal of Nursing Administration,* 32 (12), 648-654.

Kane, T. J., Rockoff, J. E. and Staiger, D. O. (2006). 'What does certification tell us about teacher effectiveness: evidence from New York City'. *NBER Working Paper Series,* (Number 12155).

Kazeem, A., Jensen, L. and Stokes, C. S. (2010). 'School Attendance in Nigeria: Understanding the Impact and Intersection of Gender, Urban-Rural Residence, and Socioeconomic Status'. *Comparative Education Review,* 54 (2), 295-319.

Kim, S. (2002). 'Participative management and job satisfaction: Lessons for management leadership'. *Public administration review,* 62 (2), 231-241.

King, G., Murray, C. J., Salomon, J. A. and Tandon, A. (2004). 'Enhancing the validity and cross-cultural comparability of measurement in survey research'. *American political science review,* 98 (01), 191-207.

Kreiner, S. and Christensen, K. B. (2013). 'Analyses of model fit and robustness. A new look at the PISA scaling model underlying ranking of countries according to reading literacy'. *Psychometrika,* 1-22.

Lambert, E. G., Lynne Hogan, N. and Barton, S. M. (2001). 'The impact of job satisfaction on turnover intent: a test of a structural measurement model using a national sample of workers'. *The Social Science Journal,* 38 (2), 233-250.

Lavy, V., Paserman, M. D. and Schlosser, A. (2012). 'Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom*'. *The Economic Journal,* 122 (559), 208-237.

Lavy, V. and Schlosser, A. (2011). 'Mechanisms and impacts of gender peer effects at school'. *American Economic Journal: Applied Economics,* 3 (2), 1-33.

Lee, V. E. and Zuze, T. L. (2011). 'School resources and academic performance in Sub-Saharan Africa'. *Comparative Education Review,* 55 (3), 369-397.

Little, A. W. (2004). 'Learning and teaching in multigrade settings'. *Background paper prepared for the Education for All Global Monitoring Report 2005, The Quality Imperative,*

Ma, X. and MacMillan, R. B. (1999). 'Influences of Workplace Conditions on Teachers' Job Satisfaction'. *The Journal of Educational Research,* 93 (1), 39-47.

Maslach, C., Schaufeli, W. B. and Leiter, M. P. (2001). 'Job Burnout'. *Annual Review of Psychology,* 52 (1), 397-422.

Metzler, J. and Woessmann, L. (2012). 'The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation'. *Journal of Development Economics,* 99 (2), 486-496.

Michaelowa, K. (2001). 'Primary Education Quality in Francophone Sub-Saharan Africa: Determinants of Learning Achievement and Efficiency Considerations'. *World Development,* 29 (10), 1699-1716.

Michaelowa, K. (2002). *Teacher job satisfaction, student achievement, and the cost of primary education in Francophone Sub-Saharan Africa*: HWWA Discussion Paper.

Micklewright, J., Allen, R., Jerrim, J., Vignoles, A., Jenkins, A., Illie, S., Bellarbre, E., Barrera, F. and Hein, C. (2014). Teachers in England's secondary schools: evidence from TALIS 2013. London: Department for Education.

Milner, G., Mulera, D. and Chimuzu, T. (2011). The SACMEQ III project in Malawi: A study of the conditions of schooling and the quality of education.

Moloi, M. Q. and Chetty, M. (2010). The SACMEQ III project in South Africa: A study of the conditions of schooling and the quality of education. Pretoria.

Mulkeen, A. (2010). 'Teachers in Anglophone Africa: Issues in Teacher Supply, Training, and Management'. *World Bank Publications*.

Nye, B., Hedges, L. V. and Konstantopoulos, S. (1999). 'The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment'. *Educational Evaluation and Policy Analysis,* 21 (2), 127-142.

Nye, B., Hedges, L. V. and Konstantopoulos, S. (2000). 'The effects of small classes on academic achievement: The results of the Tennessee class size experiment'. *American Educational Research Journal,* 37 (1), 123-151.

Nye, B., Hedges, L. V. and Konstantopoulos, S. (2001). 'Are Effects of Small Classes Cumulative? Evidence From a Tennessee Experiment'. *The Journal of Educational Research,* 94 (6), 336-345.

Nye, B., Konstantopoulos, S. and Hedges, L. V. (2004). 'How Large Are Teacher Effects?'. *Educational Evaluation and Policy Analysis,* 26 (3), 237-257.

Oketch, M., Mutisya, M., Ngware, M. and Ezeh, A. C. (2010b). 'Why are there proportionately more poor pupils enrolled in non-state schools in urban Kenya in spite of FPE policy?'. *International Journal of Educational Development,* 30 (1), 23-32.

Oketch, M., Mutisya, M., Ngware, M., Ezeh, A. C. and Epari, C. (2010a). 'Free primary education policy and pupil school mobility in urban Kenya'. *International Journal of Educational Research,* 49 (6), 173-183.

Onsomu, E., Nzomo, J. and Obiero, C. (2005). The SACMEQ II project in Kenya: A study of the conditions of schooling and the quality of education. Harare, Zimbabwe: SACMEQ.

Organisation for Economic Cooperation and Development. (2009). *Creating effective teaching and learning environments: First results from TALIS*. Paris: OECD.

Organisation for Economic Cooperation and Development. (2014). *TALIS 2013 Results: An International Perspective on Teaching and Learning*. Paris: OECD.

Organisation for Economic Cooperation and Development. (2014b). *TALIS 2013 Technical Report*. Hamburg: OECD.

Oshagbemi, T. (1999). 'Overall job satisfaction: how good are single versus multiple-item measures?'. *Journal of Managerial Psychology,* 14 (5), 388-403.

Perie, M. and Baker, D. P. (1997). 'Job Satisfaction among America's Teachers: Effects of Workplace Conditions, Background Characteristics, and Teacher Compensation. Statistical Analysis Report'.

Piko, B. F. (2006). 'Burnout, role conflict, job satisfaction and psychosocial health among Hungarian health care staff: A questionnaire survey'. *International Journal of Nursing Studies,* 43 (3), 311-318.

Prosser, D., Johnson, S., Kuipers, E., Szmukler, G., Bebbington, P. and Thornicroft, G. (1997). 'Perceived sources of work stress and satisfaction among hospital and community mental health staff, and their relation to mental health, burnout and job satisfaction'. *Journal of Psychosomatic Research,* 43 (1), 51-59.

Rasch, G. W. (1960). *Probabilistic Models for some Intelligence and Attainment Tests*: Copenhagen.

Rawal, S. and Kingdon, G. (2010). 'Akin to my teacher: does caste, religious or gender distance between student and teacher matter? Some evidence from India'. *IoE Working Papers, Department of Quantitative Social Science*.

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005). 'Teachers, Schools, and Academic Achievement'. *Econometrica,* 73 (2), 417-458.

Rockoff, J. E. (2004). 'The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data'. *The American Economic Review,* 94 (2), 247-252.

Roelen, C. A. M., Koopmans, P. C. and Groothoff, J. W. (2008). 'Which work factors determine job satisfaction?'. *Work: A Journal of Prevention, Assessment and Rehabilitation,* 30 (4), 433-439.

Rose, P. (2009). 'NGO provision of basic education: alternative or complementary service delivery to support access to the excluded?'. *Compare: A Journal of Comparative and International Education,* 39 (2), 219-233.

Sahn, D. E. and Stifel, D. C. (2003). 'Urban–rural inequality in living standards in Africa'. *Journal of African Economies,* 12 (4), 564-597.

Sauba, D. and Lutchmiah, B. (2011). The SACMEQ III project in Mauritius: A study of the conditions of schooling and the quality of education.

Skaalvik, E. M. and Skaalvik, S. (2011). 'Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion'. *Teaching and Teacher Education,* 27 (6), 1029-1038.

Slater, H., Davies, N. M. and Burgess, S. (2011). 'Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England*'. *Oxford Bulletin of Economics and Statistics,* no-no.

Sousa-Poza, A. and Sousa-Poza, A. A. (2000). 'Well-being at work: a cross-national analysis of the levels and determinants of job satisfaction'. *The Journal of Socio-Economics,* 29 (6), 517-538.

Sousa-Poza, A. and Sousa-Poza, A. A. (2007). 'The effect of job satisfaction on labor turnover by gender: An analysis for Switzerland'. *The Journal of Socio-Economics,* 36 (6), 895-913.

Taylor, S. and Spaull, N. (2013). 'The effects of rapidly expanding primary school access on effective learning: The case of Southern and Eastern Africa since 2000'. *Stellenbosch Economic Working Papers,*

Todd, P. E. a. W. K. I. (2003). 'On the Specification and Estimation of the Production Function for Cognitive Achievement'. *The Economic Journal,* 113 (485), F3--F33.

Ulleberg, I. (2009). The role and impact of NGOs in capacity development, *From replacing the state to reinvigorating education* (pp. 51). Paris: UNESCO International Institute for Educational Planning (IIEP).

United Nations (2011). Education For All Global Monitoring Report - Regional Overview Sub-Saharan Africa. In UNESCO (ed). Paris.

Van Der Linden, W. J. and Hambleton, R. K. (1997). 'Item response theory: Brief history, common models, and extensions', *Handbook of modern item response theory* (pp. 1-28): Springer.

Vegas, E. and De Laat, J. (2003). 'Do differences in teacher contracts affect student performance? Evidence from Togo'. *The World Bank,*

Wasanga, P. M., Ogle, M. A. and Wambua, R. M. (2012). The SACMEQ III project in Kenya: A study of the conditions of schooling and the quality of education. Nairobi, Kenya: Kenya National Examinations Council.

Watkins, K. (2004). 'Private Education And' Education for All'- or How Not to Construct an Evidence-Based Argument: a Reply to Tooley'. *Economic Affairs,* 24 (4)*,* 8--11.

White, H. (1980). 'Using least squares to approximate unknown regression functions'. *International Economic Review,* 149-170.

White, K. R. (1982). 'The relation between socioeconomic status and academic achievement'. *Psychological Bulletin,* 91 (3)*,* 461.

Wooldridge, J. (2002). *Introductory Econometrics*. (2 ed.): South Western.

Yu, G. and Thomas, S. M. (2008). 'Exploring school effects across southern and eastern African school systems and in Tanzania'. *Assessment in Education: Principles, Policy & Practice,* 15 (3)*,* 283-305.

Zhang, Y. (2006). 'Urban-Rural Literacy Gaps in Sub-Saharan Africa: The Roles of Socioeconomic Status and School Quality'. *Comparative Education Review,* 50 (4)*,* 581-602.

Zuze, T. L. (2010). 'Human Resource Inputs and Educational Outcomes in Botswana's Schools: Evidence from SACMEQ and TIMMS'. *Stellenbosch Economic Working Papers*, Number 16/10.

Appendix for chapter 1

Context

Table A1.1: GDP per capita in 2005 US Dollars

		2000	2007	Difference	% Difference
Group 1	**Kenya**	501	565	64	12.7
	Tanzania	304	405	101	33.1
	Zambia	562	669	107	19.0
	Unweighted Average	**456**	**546**	**90**	**21.6**
Group 2	**Malawi**	221	224	3	1.4
	Mauritius	4555	5488	933	20.5
	Namibia	3007	3937	930	30.9
	Seychelles	11492	13057	1566	13.6
	South Africa	4652	5706	1054	22.7
	Unweighted Average	**4786**	**5683**	**897**	**17.8**

Source: World Bank http://data.worldbank.org/indicator/all last accessed on the 27th of March 2014

Table A1.2: Population in millions

		2000	2007	% Difference
Group 1	**Kenya**	31.2	37.8	20.7
	Tanzania	34.0	41.1	20.9
	Zambia	10.1	12.1	19.9
	Unweighted Average	**25.1**	**30.3**	**20**
Group 2	**Mauritius**	1.2	1.3	6.2
	Malawi	11.3	13.7	21.1
	Namibia	1.9	2.1	9.6
	Seychelles	0.081	0.085	4.8
	South Africa	44.0	48.3	9.7
	Unweighted Average	**11.7**	**13.1**	**10**

Source: World Bank http://data.worldbank.org/indicator/all last accessed on the 27th of March 2014

Table A1.3: HIV prevalence

		2000	2007	Difference
Group 1	**Kenya**	8.9	6.4	-2.5
	Tanzania	7.8	5.8	-2
	Zambia	15.3	13.4	-1.9
	Unweighted Average	**10. 7**	**8.5**	**-2.1**
Group 2	**Mauritius**	0.9	1.2	0.3
	Malawi	15.8	12.4	-3.4
	Namibia	14.2	14.3	0.1
	South Africa	14.1	17.3	3.2
	Unweighted Average	**11.3**	**11.3**	**0.05**

Source: World Bank http://data.worldbank.org/indicator/all last accessed on the 27th of March 2014

Table A1.4: Proportion of age cohort reaching last grade of Primary School

		1999	2008	Difference
Group 1	Kenya	N/A	N/A	N/A
	Tanzania	69.7	73.9	4.2
	Zambia	66.3	53.1	-13.2
	Unweighted Average	**68.0**	**63.5**	**-4.5**
Group 2	Mauritius	98.5	96.2	-2.3
	Malawi	36.3	41.9	5.7
	Namibia	82.3	82.6	0.3
	Seychelles	96.1	92.0	-4.1
	South Africa	57.5	N/A	N/A
	Unweighted Average[1]	**78.3**	**78.2**	**-0.1**

Source: World Bank http://data.worldbank.org/indicator/all last accessed on the 27th of March 2014; 1 excludes South Africa

Detailed descriptive statistics of country contexts

Table A1.6: Average number of teachers per school

		2000	2007	Difference
Group 1	Kenya	15.6 (0.6)	16.0 (0.7)	0.5 (0.9)
	Tanzania	17.2 (1.9)	16.1 (1.0)	-1.2 (2.1)
	Zambia	23.5 (1.9)	19.5 (1.8)	-4.0 (2.6)
	Zanzibar	44.4 (2.4)	53.4 (2.5)	8.9* (3.5)
	Unweighted Average	**25.2**	**26.2**	**1.0**
Group 2	Malawi	21.9 (2.1)	17.6 (1.6)	-4.2 (2.6)
	Mauritius	24.3 (1.3)	24.4 (1.3)	0.1 (1.8)
	Namibia	18.8 (0.6)	19.5 (0.7)	0.7 (0.9)
	Seychelles	40.7 (3.8)	36.0 (4.3)	-4.7 (5.8)
	South Africa	19.2 (0.8)	19.9 (0.6)	0.7 (1.0)
	Unweighted Average	**25.0**	**23.5**	**-1.5**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Table A1.7: Average number of school resources per school

		Number of resources		
		2000	**2007**	**Difference**
Group 1	**Kenya**	7.2 (0.2)	6.6 (0.2)	-0.6* (0.3)
	Tanzania	5.3 (0.2)	4.5 (0.2)	-0.8** (0.3)
	Zambia	6.6 (0.4)	4.8 (0.3)	-1.8*** (0.5)
	Zanzibar	6.1 (0.2)	6.1 (0.2)	-0.1 (0.3)
	Unweighted Average	**6.3**	**5.5**	**-0.8**
Group 2	**Malawi**	4.3 (0.2)	3.4 (0.2)	-0.9*** (0.3)
	Mauritius	11.0 (0.2)	10.4 (0.3)	-0.6 (0.3)
	Namibia	8.5 (0.3)	9.6 (0.3)	1.1** (0.4)
	Seychelles	12.1 (0.2)	12.3 (0.3)	0.2 (0.4)
	South Africa	9.1 (0.4)	9.5 (0.3)	0.3 (0.5)
	Unweighted Average	**9.0**	**9.0**	**0.0**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001; The variable used is a summative measure based on whether or not the head teacher reports the presence of a school library, a hall, a staff room, a head teacher's office, a store room, a first aid kit, a sports ground, water supply, electricity, a telephone, a fax machine, a school garden, a typewriter, a duplicator, radio and tape recorder.*

Table A1.8: Average number of classroom resources

		Reading			Maths		
		2000	**2007**	**Difference**	**2000**	**2007**	**Difference**
Group 1	**Kenya**	2.7 (0.1)	2.8 (0.1)	0.1 (0.2)	2.7 (0.1)	2.8 (0.1)	0.2 (0.1)
	Tanzania	2.2 (0.1)	3.3 (0.1)	1.1*** (0.2)	2.0 (0.1)	3.2 (0.1)	1.3*** (0.2)
	Zambia	2.7 (0.1)	2.7 (0.1)	0.0 (0.2)	2.7 (0.2)	2.7 (0.1)	-0.1 (0.2)
	Zanzibar	2.2 (0.1)	2.5 (0.1)	0.3* (0.1)	2.1 (0.1)	2.4 (0.1)	0.3* (0.1)
	Unweighted Average	**2.2**	**2.5**	**0.3**	**2.4**	**2.8**	**0.4**
Group 2	**Malawi**	2.8 (0.1)	2.7 (0.1)	-0.1 (0.2)	2.7 (0.1)	2.7 (0.1)	-0.0 (0.2)
	Mauritius	4.2 (0.2)	4.9 (0.1)	0.7*** (0.2)	4.2 (0.2)	4.9 (0.1)	0.7*** (0.2)
	Namibia	3.5 (0.1)	3.3 (0.1)	-0.2 (0.1)	3.5 (0.1)	3.3 (0.1)	-0.2 (0.2)
	Seychelles	5.5 (0.1)	5.8 (0.1)	0.3 (0.2)	5.2 (0.2)	5.9 (0.1)	0.7*** (0.2)
	South Africa	4.2 (0.2)	3.8 (0.1)	-0.4* (0.2)	3.8 (0.2)	3.6 (0.1)	-0.2 (0.2)
	Unweighted Average	**4.0**	**4.1**	**0.1**	**3.9**	**4.1**	**0.2**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001; The variables used are summative measures based on whether or not the classroom has a wall chart, cupboard, bookshelves, classroom library, a desk and chair for the teacher as well as a blackboard and chalk.*

Table A1.9: Average number of pupil materials

		2000	2007	Difference
Group 1	Kenya	5.9 (0.1)	5.9 (0.1)	-0.0 (0.1)
	Tanzania	5.5 (0.1)	6.2 (0.1)	0.8*** (0.1)
	Zambia	4.2 (0.2)	4.6 (0.1)	0.4* (0.2)
	Zanzibar	5.5 (0.1)	6.5 (0.1)	1.1*** (0.2)
	Unweighted Average	**5.3**	**5.8**	**0.6**
Group 2	Malawi	4.6 (0.1)	5.5 (0.1)	1.0*** (0.1)
	Mauritius	6.7 (0.1)	6.3 (0.2)	-0.4* (0.2)
	Namibia	5.5 (0.1)	5.8 (0.1)	0.3*** (0.1)
	Seychelles	6.5 (0.1)	7.3 (0.1)	0.8*** (0.1)
	South Africa	5.3 (0.2)	6.5 (0.1)	1.3*** (0.2)
	Unweighted Average	**5.7**	**6.3**	**0.6**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001; The variable used is a summative measure based on whether the pupil has an exercise book, a notebook, a pencil, a sharpener, an eraser, a ruler, a ballpoint pen and a file.*

Table A1.10: Average school size per shift attended by grade 6 pupils and average number of shifts per school

		School size			Shifts	
		2000	2007	Difference	2000	2007
Group 1	Kenya	510 (5)	666 (6)	156*** (7)	1.0 (0.0)	1.0 (0.0)
	Tanzania	659 (18)	826 (9)	168*** (20)	1.2 (0.0)	1.1 (0.0)
	Zambia	498 (9)	424 (8)	-75*** (12)	2.8 (0.0)	2.8 (0.0)
	Zanzibar	896 (13)	1023 (16)	127*** (21)	2.0 (0.0)	1.7 (0.0)
	Unweighted Average	**641**	**735**	**94**	**1.75**	**1.65**
Group 2	Malawi	1198 (22)	1165 (16)	-33 (27)	1.1 (0.0)	1.1 (0.0)
	Mauritius	650 (8)	533 (7)	-117*** (11)	1.0 (0.0)	1.0 (0.0)
	Namibia	557 (4)	572 (4)	15** (5)	1.1 (0.0)	1.1 (0.0)
	Seychelles	691 (8)	509 (6)	-182*** (10)	1.0 (0.0)	1.0 (0.0)
	South Africa	692 (7)	680 (4)	-12 (8)	1.0 (0.0)	1.0 (0.0)
	Unweighted Average	**758**	**692**	**-66**	**1.04**	**1.04**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Table A1.11: Average grade 6 class size and average number of pupils per classroom per shift

		Grade 6 class size			Pupils per classroom per shift		
		2000	**2007**	**Difference**	**2000**	**2007**	**Difference**
Group 1	**Kenya**	37.0 (0.9)	44.5 (1.2)	7.5*** (1.5)	34.3 (0.8)	44.0 (1.2)	9.6*** (1.4)
	Tanzania	42.3 (1.7)	53.8 (2.6)	11.5*** (3.1)	67.5 (3.5)	80.5 (3.2)	13.0** (4.7)
	Zambia	36.9 (1.2)	46.6 (1.5)	9.7*** (1.9)	73.7 (31.9)	42.6 (2.0)	-31.1 (32.0)
	Zanzibar	49.8 (1.1)	55.4 (1.6)	5.6** (1.9)	58.8 (4.2)	59.7 (2.0)	0.9 (4.6)
	Unweighted Average	**41.5**	**50.1**	**8.6**	**58.6**	**56.7**	**-1.9**
Group 2	**Malawi**	56.6 (2.2)	66.3 (2.7)	9.7** (3.5)	84.0 (3.5)	89.0 (3.5)	5.0 (5.0)
	Mauritius	36.3 (0.8)	34.9 (0.6)	-1.4 (1.0)	42.0 (11.6)	26.3 (0.7)	-15.6 (11.6)
	Namibia	38.4 (0.8)	35.7 (0.5)	-2.7** (0.9)	33.9 (0.7)	35.1 (2.1)	1.2 (2.2)
	Seychelles	27.4 (0.9)	25.3 (0.8)	-2.1 (1.2)	24.3 (1.7)	21.4 (1.3)	-2.9 (2.1)
	South Africa	42.2 (1.4)	43.1 (0.9)	0.9 (1.6)	48.2 (8.5)	40.6 (1.4)	-7.5 (8.6)
	Unweighted Average	**40.2**	**41.1**	**0.9**	**46.5**	**42.5**	**-4.0**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Table A1.12: Average ratio of pupils to school resources and teachers

		Pupils per school resource			Pupils per teacher		
		2000	**2007**	**Difference**	**2000**	**2007**	**Difference**
Group 1	**KEN**	74.9 *(3.3)*	114.0 *(6.4)*	39.1*** *(7.2)*	33.3 *(0.8)*	42.9 *(1.0)*	9.6*** *(1.3)*
	TAN	125.3 *(8.6)*	209.2 *(13.9)*	83.8*** *(16.4)*	45.9 *(2.1)*	61.9 *(3.5)*	16.0*** *(4.1)*
	ZAM	78.1 *(4.5)*	79.8 *(5.7)*	1.7 *(7.2)*	32.0 *(2.7)*	34.4 *(4.0)*	2.3 *(4.8)*
	ZAN	168.0 *(10.1)*	199.7 *(15.9)*	31.7 *(18.8)*	20.8 *(0.6)*	20.0 *(0.6)*	-0.8 *(0.9)*
	Unweighted Average	**111.6**	**150.7**	**39.1**	**33.0**	**39.8**	**6.8**
Group 2	**MAL**	315.2 *(24.5)*	460.8 *(43.8)*	145.6** *(50.2)*	70.2 *(2.9)*	86.2 *(3.8)*	16.0*** *(4.8)*
	MAU	61.3 *(4.5)*	51.0 *(3.2)*	-10.3 *(5.5)*	28.4 *(2.8)*	21.6 *(0.7)*	-6.9** *(2.8)*
	NAM	81.8 *(3.9)*	64.2 *(2.2)*	-17.6*** *(4.5)*	31.5 *(0.6)*	30.0 *(0.4)*	-1.5* *(0.7)*
	SEY	57.5 *(6.9)*	41.9 *(4.6)*	-15.6 *(8.3)*	16.6 *(0.8)*	14.2 *(0.6)*	-2.4* *(1.0)*
	SOU	90.3 *(6.4)*	86.3 *(3.9)*	-4.0 *(7.5)*	36.1 *(0.7)*	34.8 *(0.4)*	-1.3 *(0.8)*
	Unweighted Average	**121.2**	**140.8**	**19.6**	**36.6**	**37.4**	**0.8**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Table A1.13: Average ratio of pupils per classroom resources

		Reading			Maths		
		2000	2007	Difference	2000	2007	Difference
Group 1	KEN	15.2 0.9	18.7 1.0	3.5** 1.3	15.8 0.8	18.3 1.0	2.5 1.3
	TAN	39.4 2.8	29.0 1.6	-10.4*** 3.2	38.9 2.8	28.4 1.5	-10.4*** 3.2
	ZAM	24.7 9.1	20.1 1.9	-4.6 9.3	22.6 7.0	20.1 1.9	-2.5 7.3
	ZAN	30.4 2.7	27.2 1.5	-3.2 3.0	33.1 4.2	28.4 1.6	-4.7 4.5
	Unweighted Average	**27.4**	**23.8**	**-3.7**	**27.6**	**23.8**	**-3.8**
Group 2	MAL	40.0 3.1	42.2 3.5	2.1 4.7	38.4 3.0	42.2 3.6	3.8 4.7
	MAU	10.3 3.2	5.7 0.3	-4.6 3.3	10.3 3.2	5.7 0.3	-4.6 3.3
	NAM	13.9 0.9	12.3 0.9	-1.6 1.3	13.2 0.7	14.1 2.4	0.9 2.5
	SEY	5.0 0.8	3.7 0.2	-1.3 0.8	4.4 0.3	3.7 0.2	-0.7 0.4
	SOU	12.2 1.6	12.5 0.7	0.3 1.7	11.3 0.7	13.8 0.8	2.6* 1.1
	Unweighted Average	**16.3**	**15.3**	**-1.0**	**15.5**	**15.9**	**0.4**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Table A1.14: Average socioeconomic status

		2000	2007	Raw Difference	Diff. in % S.D.
Group 1	Kenya	11.59 (0.23)	11.72 (0.21)	0.13 (0.30)	1.5
	Tanzania	9.70 (0.31)	9.73 (0.15)	0.03 (0.34)	0.4
	Zambia	11.61 (0.25)	11.24 (0.23)	-0.37 (0.34)	-4.3
	Zanzibar	10.21 (0.27)	10.95 (0.23)	0.74* (0.36)	8.6
	Unweighted Average	**10.78**	**10.91**	**0.13**	**1.55**
Group 2	Malawi	10.27 (0.26)	9.89 (0.17)	-0.37 (0.31)	-4.9
	Mauritius	16.59 (0.18)	16.46 (0.16)	-0.13 (0.24)	-2.0
	Namibia	11.82 (0.20)	12.59 (0.21)	0.78** (0.29)	8.2
	Seychelles	16.78 (0.20)	17.99 (0.15)	1.21*** (0.25)	19.3
	South Africa	13.69 (0.37)	14.50 (0.22)	0.81 (0.43)	7.2
	Unweighted Average	**13.83**	**14.29**	**0.46**	**5.56**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001; This variable is a summed score ranging from 0 to 25, combining parental education, quality of the house the pupil lives in and 10 household items. See section 1.4 for details.*

Table A1.15: Average proportion of pupils receiving extra tuition and help with their homework

		Extra tuition			Help with homework		
		2000	2007	Difference	2000	2007	Difference
Group 1	Kenya	87.7	70.2	-17.5***	85.5	84.2	-1.3
		(1.9)	(3.3)	(3.8)	(1.4)	(1.4)	(1.9)
	Tanzania	88.9	41.6	-47.3***	79.8	52.0	-27.9***
		(1.2)	(3.7)	(3.9)	(1.5)	(2.8)	(3.2)
	Zambia	55.0	13.5	-41.5***	82.1	88.7	6.6**
		(3.7)	(1.9)	(4.1)	(1.7)	(1.3)	(2.2)
	Zanzibar	55.9	16.4	-39.5***	82.0	81.5	-0.5
		(3.2)	(1.9)	(3.8)	(1.7)	(1.4)	(2.2)
	Unweighted Average	**71.9**	**35.4**	**-36.4**	**82.4**	**76.6**	**-5.8**
Group 2	Malawi	79.7	14.3	-65.4***	49.7	62.0	12.3***
		(3.4)	(1.9)	(3.9)	(3.1)	(2.8)	(4.2)
	Mauritius	86.9	80.8	-6.1*	87.4	83.8	-3.5
		(1.5)	(2.1)	(2.6)	(1.6)	(2.3)	(2.8)
	Namibia	44.7	20.9	-23.8***	84.3	92.5	8.2***
		(2.4)	(2.3)	(3.3)	(0.9)	(0.7)	(1.2)
	Seychelles	47.7	29.0	-18.6**	90.7	96.8	6.1***
		(3.3)	(4.4)	(5.5)	(1.1)	(0.6)	(1.2)
	South Africa	59.3	15.0	-44.3***	85.6	90.7	5.1*
		(3.3)	(1.9)	(3.8)	(2.0)	(0.8)	(2.2)
	Unweighted Average	**63.6**	**32.0**	**-31.6**	**79.5**	**85.2**	**5.6**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Table A1.16: Average number of days absent in month prior to survey and average pupil age in months

		Days absent			Age in months		
		2000	2007	Difference	2000	2007	Difference
Group 1	Kenya	1.96	1.29	-0.67***	168.4	165.1	-3.3**
		(0.10)	(0.11)	(0.15)	(0.8)	(0.8)	(1.1)
	Tanzania	2.36	2.06	-0.31	180.3	173.6	-6.7***
		(0.22)	(0.12)	(0.25)	(1.1)	(1.0)	(1.5)
	Zambia	2.50	2.54	0.04	166.2	168.6	2.4
		(0.12)	(0.12)	(0.17)	(1.4)	(1.0)	(1.7)
	Zanzibar	1.98	1.78	-0.20	179.1	169.7	-9.4***
		(0.11)	(0.12)	(0.15)	(0.8)	(0.7)	(1.1)
	Unweighted Average	**2.20**	**1.92**	**-0.29**	**173.5**	**169.3**	**-4.2**
Group 2	Malawi	2.00	1.66	-0.34	174.0	169.5	-4.5**
		(0.16)	(0.10)	(0.18)	(1.2)	(0.9)	(1.5)
	Mauritius	1.85	2.01	0.16*	135.9	136.7	0.8**
		(0.12)	(0.10)	(0.16)	(0.2)	(0.2)	(0.3)
	Namibia	1.49	0.97	-0.52***	166.4	163.2	-3.1***
		(0.09)	(0.06)	(0.11)	(0.7)	(0.6)	(0.9)
	Seychelles	0.86	1.70	0.85***	138.8	138.6	-0.2
		(0.07)	(0.08)	(0.11)	(0.2)	(0.1)	(0.2)
	South Africa	1.66	1.15	-0.51*	157.2	155.5	-1.7
		(0.15)	(0.15)	(0.21)	(0.9)	(0.5)	(1.0)
	Unweighted Average	**1.57**	**1.50**	**-0.07**	**154.4**	**152.7**	**-1.8**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Table A1.17: Average achievement by subject

		Reading			Maths		
		2000	2007	Difference	2000	2007	Difference
Group 1	Kenya	546 (5)	543 (5)	-3 (7)	563 (5)	557 (4)	-6 (6)
	Tanzania	542 (6)	575 (4)	33*** (7)	522 (5)	546 (4)	24*** (6)
	Zambia	442 (5)	436 (4)	-6 (6)	436 (4)	436 (3)	0 (4)
	Zanzibar	478 (4)	534 (4)	56*** (5)	478 (4)	486 (2)	8 (4)
	Unweighted Average	**502**	**522**	**20**	**500**	**506**	**6**
Group 2	Malawi	429 (3)	433 (3)	5 (4)	433 (2)	447 (3)	14*** (4)
	Mauritius	529 (7)	556 (7)	27** (10)	575 (8)	598 (8)	23* (11)
	Namibia	449 (4)	497 (4)	48*** (6)	431 (4)	471 (3)	40*** (5)
	Seychelles	582 (7)	575 (7)	-7 (10)	554 (6)	551 (5)	-4 (8)
	South Africa	486 (10)	489 (5)	3 (11)	480 (8)	492 (5)	11 (9)
	Unweighted Average	**495**	**510**	**15**	**495**	**512**	**17**

*Notes: Estimated means, Standard Errors clustered by school in parentheses; * p<.05, ** p<.01, *** p<.001*

Appendix for chapter 3

How do schools affect teachers' job satisfaction and workload? An additional exploration

As outlined in the introduction, bottom-up approaches lead one to believe that the school a teacher works in will affect his or her job satisfaction. The following paragraphs set out to explore the support for this assumption in the data.

As TALIS data contain multiple teachers observed in each school within a country, it is possible to use ANOVAs to explore the percentage of job satisfaction that lies between schools. Compared to the teachers' job satisfaction variable available in TALIS 2008, a single item on a 4-point Likert scale, the OECD's measure in TALIS 2013 is on a continuous scale and therefore more suitable for an analysis of variance. The OECD's measure is also a composite measure consisting of 9 indicators. According to Oshagbeni (1999) such multi-item measures have higher reliability than single-item measures in the same sample. Compared to the single-item measure in TALIS 2008, the 2013 measure might pick up differences in the variation between-schools more easily and yield higher results. The estimates in Table 3.2 show that only 9 to 22 percent of the variation in teachers' job satisfaction lies between schools. Compared to the estimated 6 percent between-school variation for the 28 participating education systems reported in the 2008 OECD TALIS report (OECD, 2009), the findings here are indeed higher. My findings suggest that although schools may be very heterogeneous in their composition of stakeholders, the vast majority of the variation in teachers' job satisfaction is associated with the teacher level. These findings are in stark contrast to wages, a traditional microeconomic variable, for which the variation is predominantly associated with time-invariant factors that vary across individuals rather than within individuals over time.

Table A3.1: Between-school variance in teacher's job satisfaction

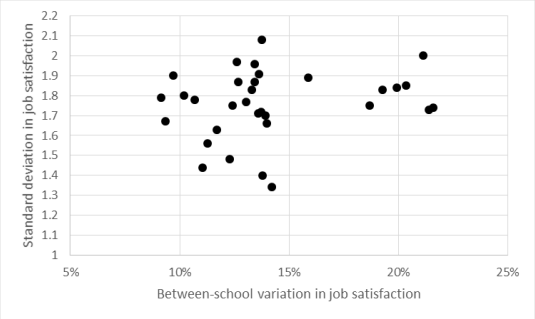
	Between-school variance (%)
Australia	13
Brazil	22
Bulgaria	19
Chile	21
Croatia	10
Czech Republic	14
Denmark	20
Estonia	11
Finland	9
France	14
Iceland	19
Israel	16
Italy	12
Japan	13
Korea	13
Latvia	11
Malaysia	14
Mexico	14
Netherlands	12
Norway	14
Poland	11
Portugal	13
Romania	14
Serbia	10
Singapore	9
Slovak Republic	12
Spain	14
Sweden	13
Sub-national education systems	
Abu Dhabi (UAE)	21
Alberta (Canada)	20
England (UK)	14
Flanders (Belgium)	13
Average (all education systems)	**–**

Estimates of the between-school variation are equivalent to the R-squared obtained from linear regression that includes a dummy variable for each school.

It may be that my estimates are biased upwards or even an artefact of selection processes. Non-random distribution of teachers to schools is likely to be the case here, as teachers will prefer to apply to different schools and schools will differ in which candidates they choose. Such processes may lead to teachers within schools being more similar than between schools, thus exacerbating the between-school variation.

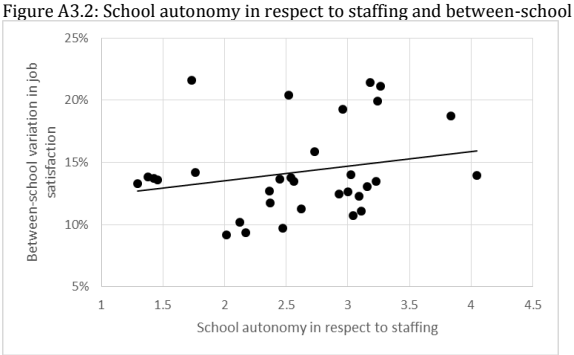
Schools may also affect the overall variance in teachers' job satisfaction in an education system. For example, the more schools are associated with teachers' job satisfaction, the larger the overall variance may be, as teachers are being subjected to an increased diversity of schools. Figure 3.1 shows a scatterplot with the percentage of variation explained by the school level on the x-axis and the standard deviation on the y-axis, with the 32 education systems as the observations. The graph does not support this expectation and instead suggests no association.

Figure A3.1: Does between-school variation in job satisfaction correlate with the variance in job satisfaction?



TALIS 2013 also contains information on the degree of school autonomy, from which the OECD create a summary measure for school autonomy in respect to staffing decisions. One might expect that the more schools and teachers can influence where they work, the more similar teachers will be in their characteristics. Top-down approaches would lead one to expect that teachers with similar characteristics should perceive their environment more similarly. Thus, the more autonomous schools are at hiring teachers, the larger the between school-variation should be in an education system. This correlation across education systems is shown in Figure A3.2 and tentatively supports this expectation, as the correlation is weakly positive.

Figure A3.2: School autonomy in respect to staffing and between-school variation in job satisfaction



In order to estimate the between-school variation in teachers' total weekly workload, I estimate ANOVAs for each country controlling for part-time status and gender, as females especially may have different preferences in respect to their working time than males, for example due to child-rearing commitments. Allowing for this, Figure A3.3 shows that, at the bottom end of the distribution, 6 percent of the variation in the total amount of hours teachers work on school matters in or outside the school building in Singapore and by 23 percent in Malaysia at the top end of the distribution. As with job satisfaction, these estimates suggest that only a minority of the variation in the time teachers work is associated with school-level factors, and that there may be many variables at the individual level driving differences in the their workload.

This leads to the question whether teachers work longer hours in education systems with higher-between school variation. Figure A3.4 though suggests a negative correlation, i.e. that teachers work less hours in education systems with higher between-school variation in teachers total weekly workload, but this correlation is negligible.

Figure A3.3: Percentage of between-school variation in teachers' total weekly workload

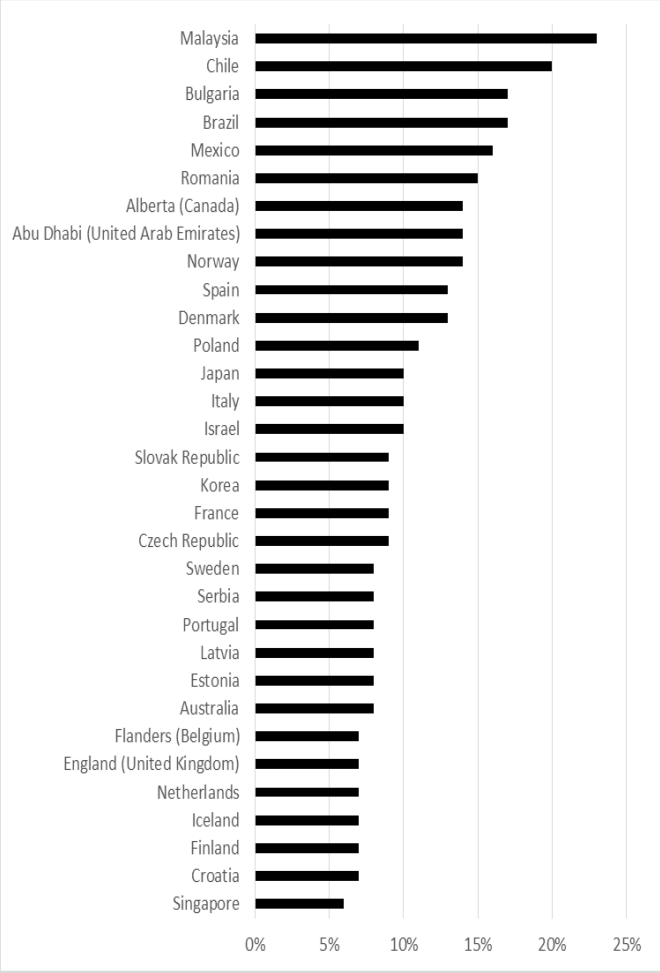
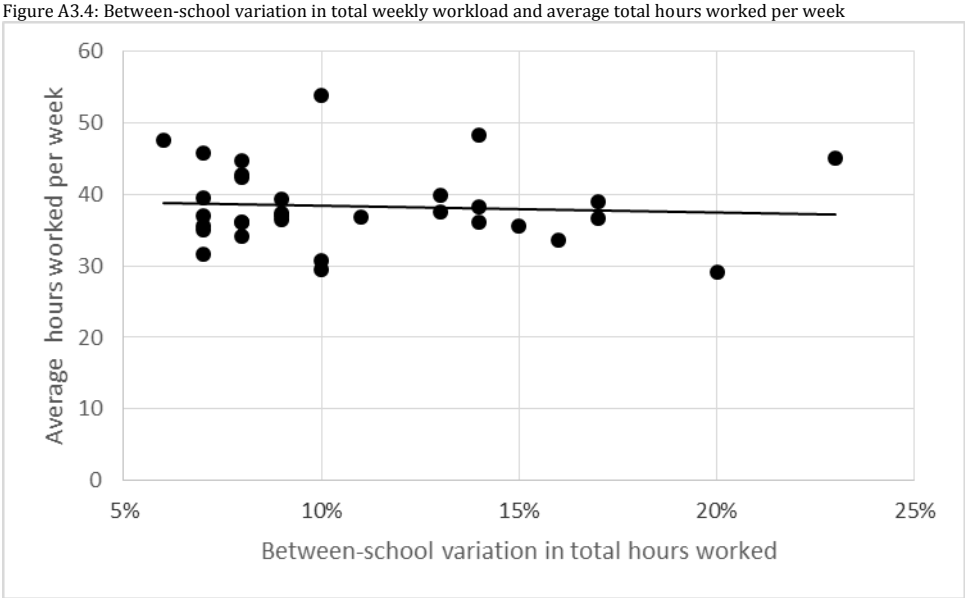


Figure A3.4: Between-school variation in total weekly workload and average total hours worked per week



The following tables report the estimates nationally standardised coefficients (national standard deviations) with T-statistics in parentheses. Estimates are statistically significant at the 5 percent level when T>1.96. The estimates are listed alphabetically by country in each table.

Table A3.2: Chapter 3 – baseline Model (M1)

Country	Constant	Weekly workload
Australia	0.008	0.000
	(0.09)	*(-0.00)*
Brazil	0.003	0.002
	(0.08)	*(2.78)*
Bulgaria	-0.033	0.000
	(-0.43)	*(0.26)*
Chile	-0.048	0.002
	(-0.56)	*(1.31)*
Croatia	-0.066	0.002
	(-1.1)	*(1.53)*
Czech Republic	0.054	-0.001
	(0.87)	*(-1.08)*
Denmark	0.261	-0.005
	(2.37)	*(-2.05)*
Estonia	0.137	-0.003
	(2.4)	*(-1.95)*
Finland	-0.018	0.002
	(-0.24)	*(0.77)*
France	0.124	-0.002
	(1.65)	*(-1.3)*
Iceland	0.009	-0.001
	(0.1)	*(-0.45)*
Israel	-0.087	0.003
	(-1.56)	*(2.81)*
Italy	-0.044	0.001
	(-0.69)	*(0.86)*
Japan	0.010	0.000
	(0.15)	*(0.37)*
Korea	0.001	0.000
	(0.02)	*(0.37)*
Latvia	0.036	-0.001
	(0.67)	*(-0.45)*
Malaysia	-0.001	0.000
	(-0.02)	*(0.41)*
Mexico	-0.146	0.003
	(-2.99)	*(3.05)*
Netherlands	0.059	-0.002
	(0.57)	*(-0.65)*
Norway	0.158	-0.004
	(1.54)	*(-1.47)*
Poland	-0.043	0.002
	(-0.7)	*(1.1)*
Portugal	0.013	0.000
	(0.18)	*(-0.23)*
Romania	-0.044	0.001
	(-0.64)	*(0.84)*
Serbia	0.039	0.000
	(0.83)	*(-0.15)*
Singapore	0.207	-0.003
	(4.12)	*(-4.18)*
Slovak Republic	0.055	-0.001
	(0.91)	*(-0.63)*
Spain	-0.082	0.002
	(-1.36)	*(1.81)*
Sweden	0.193	-0.004
	(1.99)	*(-1.94)*
Abu Dhabi (UAE)	0.186	-0.002
	(2.25)	*(-1.78)*
Alberta (CAN)	0.129	-0.002
	(1.59)	*(-1.38)*
England (UK)	-0.005	0.000
	(-0.06)	*(0.28)*
Flanders (BEL)	0.232	-0.005
	(3.17)	*(-2.77)*

Table A3.3: Chapter 3 – Model 2

Variable	Australia	Brazil	Bulgaria	Chile	Croatia	Czech Republic	Denmark	Estonia
Constant	-3.281 *(-14.08)*	-3.127 *(-20.25)*	-3.316 *(-11.8)*	-3.237 *(-12.79)*	-4.007 *(-20.63)*	-2.673 *(-11.26)*	-3.721 *(-10.74)*	-2.956 *(-12.34)*
Cooperation with colleagues index	0.077 *(5.01)*	0.059 *(7.44)*	0.027 *(1.98)*	0.077 *(7.21)*	0.105 *(7.84)*	0.086 *(6.69)*	0.065 *(3.85)*	0.101 *(6.2)*
Teacher's self-efficacy	0.097 *(5.87)*	0.116 *(11.07)*	0.135 *(9.07)*	0.089 *(5.64)*	0.154 *(12.21)*	0.118 *(7.32)*	0.149 *(6.28)*	0.079 *(5.81)*
Teacher's age	0.002 *(0.66)*	0.002 *(1.25)*	-0.004 *(-1.06)*	0.012 *(3.16)*	0.006 *(2.31)*	0.000 *(-0.11)*	0.006 *(1.22)*	0.004 *(1.68)*
Experience at current school	0.009 *(1.51)*	0.000 *(0.15)*	-0.003 *(-0.85)*	-0.002 *(-0.45)*	0.000 *(-0.16)*	0.003 *(1.16)*	0.015 *(3.12)*	-0.001 *(-0.53)*
Total teaching experience	-0.007 *(-1.7)*	0.002 *(0.77)*	0.003 *(0.96)*	-0.004 *(-1.04)*	-0.002 *(-0.72)*	-0.007 *(-1.74)*	-0.017 *(-3.08)*	-0.001 *(-0.2)*
Female	0.028 *(0.46)*	0.000 *(-0.01)*	-0.060 *(-1.17)*	0.135 *(2.76)*	0.178 *(5.05)*	0.079 *(1.59)*	-0.005 *(-0.09)*	0.025 *(0.47)*
Full-time employment	0.002 *(0.02)*	0.022 *(0.71)*	-0.042 *(-0.47)*	0.049 *(0.75)*	-0.006 *(-0.1)*	-0.084 *(-1.46)*	0.086 *(0.91)*	0.074 *(1.7)*
Is mentee	-0.056 *(-0.87)*	0.006 *(0.09)*	0.027 *(0.39)*	0.056 *(0.5)*	0.021 *(0.43)*	0.148 *(2.04)*	0.015 *(0.2)*	0.033 *(0.51)*
Is mentor	0.352 *(4.51)*	0.081 *(2.6)*	0.277 *(4.47)*	-0.058 *(-0.48)*	0.246 *(3.16)*	0.179 *(2.1)*	0.077 *(0.47)*	-0.077 *(-0.71)*
Permanent contract	-0.059 *(-0.75)*	-0.050 *(-1.32)*	-0.106 *(-1.51)*	-0.020 *(-0.31)*	-0.062 *(-0.99)*	-0.055 *(-0.85)*	0.032 *(0.21)*	0.005 *(0.1)*
Total weekly workload	-0.003 *(-2.26)*	0.000 *(-0.34)*	-0.002 *(-1.37)*	0.000 *(0.06)*	-0.002 *(-1.93)*	-0.005 *(-3.05)*	-0.010 *(-4.05)*	-0.006 *(-4.42)*

Table A3.3 continued

Variable	Finland	France	Iceland	Israel	Italy	Japan	Korea	Latvia
Constant	-2.972 *(-13.15)*	-3.394 *(-11.56)*	-2.329 *(-6.78)*	-3.681 *(-17.72)*	-2.925 *(-11.36)*	-2.570 *(-10.82)*	-2.528 *(-8.52)*	-3.063 *(-9.71)*
Cooperation with colleagues index	0.061 *(4.29)*	0.112 *(6.4)*	0.100 *(5.25)*	0.082 *(5.27)*	0.083 *(6.94)*	0.069 *(5.09)*	0.023 *(1.88)*	0.083 *(4.58)*
Teacher's self-efficacy	0.134 *(11.32)*	0.128 *(8.33)*	0.047 *(2.34)*	0.116 *(9.61)*	0.127 *(9.12)*	0.122 *(9.03)*	0.091 *(9.47)*	0.106 *(4.76)*
Teacher's age	-0.001 *(-0.26)*	0.000 *(0.04)*	0.004 *(0.75)*	0.026 *(7.93)*	-0.009 *(-2.39)*	0.006 *(1.48)*	-0.004 *(-0.94)*	0.008 *(1.92)*
Experience at current school	0.003 *(0.72)*	0.012 *(3.42)*	-0.005 *(-0.91)*	-0.004 *(-1.21)*	0.011 *(3.55)*	0.004 *(0.96)*	-0.005 *(-1.37)*	-0.005 *(-1.8)*
Total teaching experience	-0.002 *(-0.39)*	-0.010 *(-1.87)*	-0.004 *(-0.66)*	-0.013 *(-3.46)*	0.000 *(-0.03)*	-0.011 *(-2.8)*	-0.001 *(-0.13)*	0.000 *(0.09)*
Female	0.050 *(1.07)*	0.086 *(2.13)*	0.268 *(4.12)*	-0.008 *(-0.15)*	-0.007 *(-0.16)*	-0.063 *(-1.84)*	-0.052 *(-1.22)*	0.115 *(1.55)*
Full-time employment	0.153 *(1.74)*	0.050 *(0.87)*	0.073 *(0.8)*	0.047 *(0.72)*	-0.088 *(-1.2)*	-0.466 *(-4.49)*	0.162 *(1.01)*	-0.035 *(-0.51)*
Is mentee	0.052 *(0.49)*	0.054 *(0.56)*	0.086 *(1.04)*	0.055 *(1.28)*	0.042 *(0.55)*	-0.017 *(-0.35)*	0.115 *(2.47)*	0.106 *(1.25)*
Is mentor	0.056 *(0.53)*	0.123 *(1.03)*	0.361 *(2.5)*	-0.067 *(-1.43)*	0.030 *(0.27)*	0.210 *(5.32)*	0.228 *(4.98)*	0.042 *(0.4)*
Permanent contract	-0.167 *(-3.01)*	-0.244 *(-2.57)*	-0.109 *(-1.24)*	-0.081 *(-1.47)*	0.016 *(0.27)*	-0.111 *(-2.17)*	-0.157 *(-2.49)*	0.028 *(0.34)*
Total weekly workload	-0.003 *(-1.31)*	-0.005 *(-2.61)*	-0.004 *(-1.68)*	-0.002 *(-1.57)*	-0.003 *(-1.84)*	0.000 *(-0.32)*	-0.002 *(-2.02)*	-0.003 *(-2.67)*

Table A3.3 continued

Variable	Malaysia	Mexico	Netherlands	Norway	Poland	Portugal	Romania	Serbia
Constant	-3.664 *(-6.82)*	-3.643 *(-16.39)*	-3.105 *(-7.23)*	-3.474 *(-12.76)*	-3.163 *(-13.7)*	-3.759 *(-17.13)*	-4.084 *(-13.57)*	-3.786 *(-20.58)*
Cooperation with colleagues index	0.057 *(5.31)*	0.070 *(7.49)*	0.095 *(4.1)*	0.113 *(7.06)*	0.047 *(2.85)*	0.065 *(5.00)*	0.062 *(4.73)*	0.093 *(9.78)*
Teacher's self-efficacy	0.168 *(12.96)*	0.134 *(9.7)*	0.180 *(9.76)*	0.126 *(5.99)*	0.134 *(8.91)*	0.132 *(9.34)*	0.138 *(7.00)*	0.150 *(12.45)*
Teacher's age	0.004 *(0.81)*	0.001 *(0.35)*	-0.004 *(-0.83)*	0.009 *(2.95)*	0.003 *(0.52)*	0.010 *(2.45)*	-0.007 *(-1.42)*	0.001 *(0.4)*
Experience at current school	0.006 *(1.77)*	0.008 *(2.69)*	0.007 *(1.73)*	0.006 *(1.59)*	0.004 *(1.05)*	0.018 *(4.23)*	0.006 *(1.97)*	0.003 *(1.15)*
Total teaching experience	-0.004 *(-0.85)*	0.001 *(0.24)*	-0.010 *(-1.76)*	-0.011 *(-2.25)*	-0.004 *(-0.73)*	-0.015 *(-3.04)*	0.009 *(1.67)*	0.004 *(1.3)*
Female	-0.058 *(-1.27)*	-0.015 *(-0.37)*	0.008 *(0.14)*	0.122 *(2.42)*	0.079 *(1.91)*	0.030 *(0.7)*	0.112 *(2.19)*	0.084 *(2.6)*
Full-time employment	0.151 *(1.02)*	-0.021 *(-0.41)*	0.164 *(2.6)*	0.043 *(0.57)*	-0.027 *(-0.44)*	-0.113 *(-1.24)*	-0.096 *(-1.34)*	0.069 *(1.46)*
Is mentee	0.069 *(1.57)*	0.025 *(0.4)*	0.079 *(1.05)*	0.063 *(0.98)*	0.160 *(3.15)*	0.053 *(0.72)*	-0.002 *(-0.03)*	-0.028 *(-0.61)*
Is mentor	0.103 *(2.87)*	0.017 *(0.33)*	-0.068 *(-1.08)*	0.277 *(1.89)*	0.126 *(2.12)*	0.139 *(1.69)*	0.117 *(1.69)*	0.135 *(1.86)*
Permanent contract	-0.439 *(-0.98)*	0.023 *(0.43)*	-0.222 *(-2.62)*	-0.239 *(-2.81)*	-0.242 *(-3.41)*	-0.035 *(-0.63)*	-0.081 *(-1.71)*	-0.362 *(-7.79)*
Total weekly workload	-0.001 *(-1.52)*	-0.001 *(-0.77)*	-0.007 *(-2.29)*	-0.008 *(-3.6)*	-0.001 *(-0.7)*	-0.002 *(-1.38)*	-0.002 *(-1.53)*	-0.003 *(-2.84)*

Table A3.3 continued

Variable	Singapore	Slovak Republic	Spain	Sweden	Abu Dhabi (UAE)	Alberta (Canada)	England (UK)	Flanders (BEL)
Constant	-2.965 *(-12.33)*	-3.086 *(-11.74)*	-3.325 *(-12.74)*	-3.373 *(-14.3)*	-4.350 *(-11.55)*	-2.616 *(-9.39)*	-3.295 *(-13.88)*	-2.508 *(-11.07)*
Cooperation with colleagues index	0.099 *(8.39)*	0.076 *(6.1)*	0.091 *(5.63)*	0.123 *(9.17)*	0.048 *(3.52)*	0.083 *(4.93)*	0.120 *(10.19)*	0.061 *(3.33)*
Teacher's self-efficacy	0.064 *(7.43)*	0.100 *(6.61)*	0.142 *(11.22)*	0.097 *(7.6)*	0.110 *(5.8)*	0.085 *(5.69)*	0.101 *(7.98)*	0.110 *(7.77)*
Teacher's age	0.014 *(3.65)*	0.011 *(2.78)*	0.001 *(0.18)*	0.009 *(3.08)*	0.017 *(3.15)*	0.004 *(1.03)*	-0.005 *(-1.33)*	0.005 *(1.44)*
Experience at current school	0.010 *(2.5)*	-0.002 *(-0.52)*	0.016 *(5.67)*	0.001 *(0.46)*	0.004 *(0.64)*	0.019 *(3.82)*	0.001 *(0.3)*	0.008 *(1.87)*
Total teaching experience	-0.001 *(-0.26)*	-0.009 *(-2.28)*	-0.014 *(-3.43)*	-0.009 *(-2.88)*	-0.006 *(-0.88)*	-0.009 *(-1.71)*	0.000 *(0.03)*	-0.012 *(-2.16)*
Female	-0.015 *(-0.38)*	0.060 *(1.17)*	0.113 *(3.34)*	0.131 *(3.42)*	-0.085 *(-1.51)*	-0.006 *(-0.11)*	-0.028 *(-0.66)*	0.042 *(0.99)*
Full-time employment	0.124 *(1.74)*	-0.164 *(-2.32)*	-0.045 *(-0.68)*	0.092 *(2.05)*	0.107 *(0.71)*	-0.115 *(-1.46)*	0.116 *(1.6)*	0.201 *(3.64)*
Is mentee	0.006 *(0.15)*	0.034 *(0.51)*	-0.033 *(-0.47)*	0.259 *(3.17)*	-0.039 *(-0.71)*	0.122 *(2.09)*	-0.015 *(-0.32)*	0.110 *(1.58)*
Is mentor	0.310 *(7.53)*	0.025 *(0.34)*	-0.118 *(-1.08)*	0.420 *(3.78)*	0.019 *(0.38)*	0.173 *(2.32)*	0.156 *(2.59)*	0.176 *(2.56)*
Permanent contract	-0.276 *(-5.24)*	-0.170 *(-3.12)*	-0.131 *(-2.46)*	-0.241 *(-3.86)*	-0.179 *(-2.87)*	-0.326 *(-4.08)*	-0.054 *(-0.66)*	-0.292 *(-4.8)*
Total weekly workload	-0.005 *(-6.47)*	-0.002 *(-1.04)*	-0.001 *(-1.15)*	-0.008 *(-4.21)*	-0.003 *(-2.65)*	-0.004 *(-2.57)*	-0.005 *(-3.05)*	-0.010 *(-5.6)*

Table A3.4: Chapter 3 – Model 3

Variable	Australia	Brazil	Bulgaria	Chile	Croatia	Czech Republic	Denmark	Estonia
Constant	-2.644 *(-5.79)*	-2.850 *(-12.75)*	-3.526 *(-7.7)*	-2.679 *(-7.7)*	-3.945 *(-13.00)*	-2.821 *(-9.05)*	-3.591 *(-7.21)*	-2.832 *(-8.27)*
Distributed Leadership Index	-0.051 *(-1.71)*	-0.038 *(-4.18)*	0.018 *(1.25)*	0.006 *(0.42)*	-0.005 *(-0.36)*	0.018 *(1.08)*	-0.031 *(-1.4)*	0.018 *(1.1)*
Instructional Leadership Index	-0.002 *(-0.13)*	0.028 *(2.79)*	0.017 *(0.64)*	-0.052 *(-2.4)*	-0.001 *(-0.09)*	0.000 *(-0.02)*	0.020 *(0.92)*	-0.030 *(-1.67)*
Teacher Administrative Support Personnel Ratio	0.011 *(0.98)*	0.005 *(0.88)*	-0.045 *(-0.7)*	-0.004 *(-0.33)*	-0.003 *(-0.65)*	0.004 *(0.32)*	0.006 *(0.38)*	0.002 *(0.31)*
Teacher cooperates with colleagues	0.082 *(5.31)*	0.063 *(8.31)*	0.030 *(2.11)*	0.079 *(7.7)*	0.107 *(7.91)*	0.083 *(6.35)*	0.066 *(3.83)*	0.100 *(6.25)*
Teacher Pedagogical Support Personnel Ratio	0.000 *(0.08)*	-0.004 *(-2.72)*	0.004 *(2.25)*	0.006 *(0.99)*	0.003 *(0.92)*	-0.005 *(-3.51)*	0.000 *(-0.02)*	-0.006 *(-1.27)*
Teacher's self-efficacy	0.095 *(5.93)*	0.113 *(10.72)*	0.133 *(8.79)*	0.092 *(5.79)*	0.155 *(12.29)*	0.120 *(7.4)*	0.150 *(6.27)*	0.080 *(5.87)*
Teacher's age	0.002 *(0.58)*	0.003 *(1.54)*	-0.004 *(-1.13)*	0.012 *(3.28)*	0.006 *(2.39)*	0.000 *(-0.09)*	0.007 *(1.3)*	0.004 *(1.57)*
Experience at current school	0.008 *(1.4)*	0.000 *(-0.01)*	-0.001 *(-0.46)*	-0.002 *(-0.35)*	-0.001 *(-0.33)*	0.003 *(1.14)*	0.015 *(3.33)*	-0.001 *(-0.58)*
Total teaching experience	-0.006 *(-1.53)*	0.002 *(0.84)*	0.004 *(1.05)*	-0.004 *(-0.98)*	-0.002 *(-0.69)*	-0.006 *(-1.66)*	-0.017 *(-3.07)*	0.000 *(-0.14)*
Female	0.043 *(0.7)*	0.003 *(0.1)*	-0.058 *(-1.13)*	0.131 *(2.75)*	0.179 *(5.09)*	0.076 *(1.55)*	-0.003 *(-0.04)*	0.028 *(0.51)*
Full-time employment	0.007 *(0.07)*	0.027 *(0.92)*	-0.068 *(-0.8)*	0.045 *(0.69)*	-0.012 *(-0.22)*	-0.087 *(-1.53)*	0.084 *(0.9)*	0.074 *(1.69)*
Head teacher is female	-0.136 *(-2.11)*	-0.035 *(-0.76)*	-0.101 *(-1.61)*	-0.026 *(-0.38)*	-0.065 *(-1.53)*	-0.083 *(-1.68)*	-0.052 *(-0.6)*	-0.013 *(-0.25)*
Is mentee	-0.052 *(-0.83)*	-0.005 *(-0.07)*	0.024 *(0.34)*	0.059 *(0.53)*	0.018 *(0.38)*	0.143 *(2.01)*	0.014 *(0.2)*	0.049 *(0.77)*
Is mentor	0.339 *(4.31)*	0.069 *(2.24)*	0.264 *(4.34)*	-0.037 *(-0.32)*	0.242 *(3.26)*	0.175 *(2.08)*	0.074 *(0.45)*	-0.085 *(-0.78)*
Permanent contract	-0.061 *(-0.79)*	-0.079 *(-2.1)*	-0.119 *(-1.67)*	-0.027 *(-0.43)*	-0.062 *(-0.97)*	-0.054 *(-0.83)*	0.030 *(0.19)*	0.005 *(0.09)*
Total weekly workload	-0.004 *(-2.48)*	0.000 *(-0.23)*	-0.002 *(-1.43)*	0.000 *(-0.05)*	-0.002 *(-1.79)*	-0.005 *(-2.94)*	-0.010 *(-4.06)*	-0.006 *(-4.29)*

Table A3.4 continued

Variable	Finland	France	Iceland	Israel	Italy	Japan	Korea	Latvia
Constant	-3.180 *(-8.17)*	-3.124 *(-8.64)*	-2.276 *(-4.69)*	-3.327 *(-10.33)*	-2.758 *(-7.04)*	-2.280 *(-8.43)*	-2.917 *(-7.46)*	-3.264 *(-6.71)*
Distributed Leadership Index	0.012 *(0.46)*	-0.030 *(-2.68)*	0.013 *(0.6)*	-0.015 *(-0.52)*	0.025 *(0.84)*	0.011 *(0.91)*	0.027 *(1.6)*	0.006 *(0.25)*
Instructional Leadership Index	-0.005 *(-0.29)*	0.018 *(1.21)*	-0.006 *(-0.35)*	-0.019 *(-1.02)*	-0.014 *(-0.91)*	-0.031 *(-1.88)*	0.003 *(0.21)*	-0.012 *(-0.57)*
Teacher Administrative Support Personnel Ratio	0.006 *(1.12)*	0.000 *(-0.05)*	0.010 *(0.79)*	0.016 *(1.06)*	-0.022 *(-3.1)*	-0.008 *(-0.96)*	-0.007 *(-0.45)*	0.018 *(1.41)*
Teacher cooperates with colleagues	0.061 *(4.23)*	0.109 *(6.29)*	0.100 *(5.42)*	0.081 *(5.18)*	0.085 *(7.15)*	0.070 *(5.12)*	0.024 *(1.92)*	0.080 *(4.44)*
Teacher Pedagogical Support Personnel Ratio	0.001 *(0.26)*	0.002 *(0.69)*	-0.027 *(-3.18)*	-0.004 *(-0.72)*	0.000 *(-0.61)*	-0.002 *(-1.1)*	0.000 *(0.13)*	0.016 *(2.11)*
Teacher's self-efficacy	0.134 *(11.35)*	0.129 *(8.46)*	0.050 *(2.54)*	0.116 *(9.65)*	0.124 *(8.73)*	0.123 *(9.19)*	0.092 *(9.69)*	0.111 *(4.96)*
Teacher's age	0.000 *(-0.11)*	0.000 *(-0.02)*	0.004 *(0.85)*	0.023 *(7.06)*	-0.008 *(-2.21)*	0.006 *(1.43)*	-0.004 *(-0.87)*	0.008 *(1.95)*
Experience at current school	0.003 *(0.73)*	0.010 *(2.88)*	-0.005 *(-0.89)*	-0.004 *(-1.08)*	0.011 *(3.42)*	0.005 *(1.06)*	-0.004 *(-1.14)*	-0.004 *(-1.6)*
Total teaching experience	-0.002 *(-0.45)*	-0.009 *(-1.6)*	-0.005 *(-0.82)*	-0.012 *(-3.06)*	0.000 *(-0.07)*	-0.011 *(-2.82)*	-0.001 *(-0.27)*	0.001 *(0.14)*
Female	0.049 *(1.06)*	0.082 *(2.05)*	0.266 *(4.08)*	-0.053 *(-0.98)*	-0.005 *(-0.11)*	-0.063 *(-1.84)*	-0.049 *(-1.16)*	0.117 *(1.62)*
Full-time employment	0.156 *(1.71)*	0.040 *(0.68)*	0.058 *(0.64)*	0.050 *(0.76)*	-0.067 *(-0.95)*	-0.478 *(-4.55)*	0.185 *(1.1)*	-0.031 *(-0.47)*
Head teacher is female	0.035 *(0.69)*	-0.106 *(-2.36)*	-0.143 *(-1.98)*	0.115 *(1.81)*	-0.040 *(-0.85)*	0.014 *(0.22)*	-0.065 *(-1.00)*	-0.035 *(-0.54)*
Is mentee	0.055 *(0.52)*	0.053 *(0.56)*	0.124 *(1.51)*	0.060 *(1.38)*	0.051 *(0.67)*	-0.014 *(-0.29)*	0.113 *(2.45)*	0.093 *(1.1)*
Is mentor	0.060 *(0.56)*	0.122 *(0.99)*	0.360 *(2.5)*	-0.052 *(-1.08)*	0.022 *(0.2)*	0.206 *(5.32)*	0.230 *(4.98)*	0.033 *(0.32)*
Permanent contract	-0.171 *(-3.09)*	-0.240 *(-2.61)*	-0.109 *(-1.23)*	-0.075 *(-1.39)*	0.000 *(0.01)*	-0.119 *(-2.31)*	-0.164 *(-2.55)*	0.031 *(0.38)*
Total weekly workload	-0.003 *(-1.46)*	-0.006 *(-2.8)*	-0.004 *(-1.84)*	-0.001 *(-1.33)*	-0.003 *(-2.00)*	0.000 *(-0.19)*	-0.002 *(-2.05)*	-0.003 *(-2.62)*

Table A3.4 continued

Variable	Malaysia	Mexico	Netherlands	Norway	Poland	Portugal	Romania	Serbia
Constant	-3.734 *(-6.56)*	-3.712 *(-13.25)*	-3.417 *(-5.75)*	-4.349 *(-9.61)*	-2.968 *(-9.3)*	-3.297 *(-8.97)*	-4.355 *(-12.83)*	-3.782 *(-12.9)*
Distributed Leadership Index	0.017 *(1.04)*	0.018 *(1.87)*	0.027 *(1.54)*	0.073 *(2.2)*	-0.003 *(-0.25)*	-0.005 *(-0.3)*	0.024 *(2.04)*	0.000 *(-0.05)*
Instructional Leadership Index	0.009 *(0.51)*	-0.015 *(-1.16)*	-0.009 *(-0.45)*	-0.012 *(-0.68)*	-0.019 *(-1.2)*	-0.025 *(-1.84)*	-0.008 *(-0.56)*	-0.004 *(-0.19)*
Teacher Administrative Support Personnel Ratio	-0.015 *(-1.56)*	-0.001 *(-0.29)*	0.015 *(1.1)*	0.003 *(0.25)*	0.004 *(0.57)*	-0.011 *(-1.13)*	0.000 *(-0.01)*	0.002 *(0.4)*
Teacher cooperates with colleagues	0.057 *(5.37)*	0.070 *(7.44)*	0.099 *(4.38)*	0.114 *(7.19)*	0.049 *(2.91)*	0.065 *(5.03)*	0.061 *(4.77)*	0.092 *(9.85)*
Teacher Pedagogical Support Personnel Ratio	0.000 *(-0.59)*	0.003 *(1.92)*	0.000 *(-0.04)*	0.003 *(0.67)*	0.007 *(3.09)*	-0.001 *(-0.64)*	0.001 *(0.9)*	0.001 *(1.08)*
Teacher's self-efficacy	0.167 *(13.17)*	0.135 *(9.74)*	0.180 *(10.01)*	0.125 *(5.98)*	0.134 *(9.13)*	0.132 *(9.4)*	0.138 *(7.22)*	0.150 *(12.52)*
Teacher's age	0.004 *(0.82)*	0.001 *(0.36)*	-0.004 *(-0.86)*	0.009 *(3.05)*	0.003 *(0.68)*	0.010 *(2.37)*	-0.007 *(-1.4)*	0.001 *(0.45)*
Experience at current school	0.005 *(1.59)*	0.008 *(2.56)*	0.007 *(1.54)*	0.006 *(1.54)*	0.003 *(0.84)*	0.017 *(4.16)*	0.006 *(2.15)*	0.003 *(1.19)*
Total teaching experience	-0.004 *(-0.8)*	0.001 *(0.19)*	-0.010 *(-1.64)*	-0.010 *(-2.21)*	-0.004 *(-0.75)*	-0.014 *(-2.95)*	0.009 *(1.55)*	0.004 *(1.24)*
Female	-0.030 *(-0.67)*	-0.014 *(-0.35)*	0.012 *(0.21)*	0.122 *(2.4)*	0.096 *(2.3)*	0.027 *(0.64)*	0.111 *(2.18)*	0.086 *(2.71)*
Full-time employment	0.130 *(0.88)*	-0.018 *(-0.36)*	0.162 *(2.61)*	0.036 *(0.46)*	-0.035 *(-0.59)*	-0.128 *(-1.43)*	-0.096 *(-1.36)*	0.067 *(1.41)*
Head teacher is female	-0.142 *(-2.68)*	0.019 *(0.37)*	-0.041 *(-0.49)*	-0.069 *(-0.87)*	-0.088 *(-1.94)*	-0.028 *(-0.54)*	0.086 *(1.59)*	-0.017 *(-0.39)*
Is mentee	0.072 *(1.61)*	0.028 *(0.46)*	0.080 *(1.05)*	0.065 *(1.02)*	0.154 *(3.02)*	0.061 *(0.83)*	-0.008 *(-0.11)*	-0.029 *(-0.64)*
Is mentor	0.106 *(2.95)*	0.025 *(0.47)*	-0.067 *(-1.02)*	0.322 *(1.98)*	0.127 *(2.14)*	0.146 *(1.79)*	0.119 *(1.71)*	0.139 *(1.88)*
Permanent contract	-0.484 *(-1.1)*	0.006 *(0.11)*	-0.232 *(-2.7)*	-0.242 *(-2.77)*	-0.233 *(-3.33)*	-0.031 *(-0.55)*	-0.085 *(-1.82)*	-0.362 *(-7.82)*
Total weekly workload	-0.001 *(-1.33)*	-0.001 *(-0.83)*	-0.007 *(-2.37)*	-0.008 *(-3.38)*	-0.001 *(-0.62)*	-0.002 *(-1.37)*	-0.002 *(-1.79)*	-0.003 *(-2.89)*

Table A3.4 continued

Variable	Singapore	Slovak Republic	Spain	Sweden	Abu Dhabi (UAE)	Alberta (Canada)	England (UK)	Flanders (BEL)
Constant	-3.451 *(-10.34)*	-3.028 *(-7.96)*	-3.266 *(-10.93)*	-3.260 *(-8.39)*	-4.734 *(-8.15)*	-2.554 *(-7.53)*	-3.570 *(-10.41)*	-2.592 *(-7.84)*
Distributed Leadership Index	0.034 *(2.41)*	0.014 *(0.83)*	-0.013 *(-1.41)*	-0.003 *(-0.17)*	0.011 *(0.57)*	-0.006 *(-0.48)*	0.033 *(2.88)*	0.007 *(0.51)*
Instructional Leadership Index	0.007 *(0.54)*	-0.023 *(-1.39)*	0.012 *(1.12)*	-0.009 *(-0.6)*	0.031 *(1.32)*	0.005 *(0.32)*	-0.013 *(-0.94)*	0.007 *(0.44)*
Teacher Administrative Support Personnel Ratio	-0.004 *(-0.29)*	0.019 *(1.62)*	0.006 *(0.59)*	0.006 *(1.03)*	-0.020 *(-1.43)*	-0.017 *(-1.2)*	0.017 *(0.85)*	-0.003 *(-0.54)*
Teacher cooperates with colleagues	0.100 *(8.58)*	0.078 *(6.2)*	0.089 *(5.45)*	0.123 *(9.1)*	0.051 *(3.9)*	0.090 *(5.27)*	0.119 *(10.13)*	0.062 *(3.38)*
Teacher Pedagogical Support Personnel Ratio	0.004 *(2.52)*	-0.001 *(-0.72)*	-0.001 *(-0.53)*	-0.003 *(-0.54)*	-0.003 *(-0.84)*	-0.004 *(-0.49)*	0.004 *(0.45)*	0.000 *(-0.31)*
Teacher's self-efficacy	0.064 *(7.55)*	0.101 *(6.71)*	0.143 *(11.35)*	0.098 *(7.73)*	0.113 *(6.26)*	0.086 *(5.8)*	0.102 *(7.9)*	0.110 *(7.69)*
Teacher's age	0.014 *(3.72)*	0.011 *(2.78)*	0.001 *(0.17)*	0.009 *(3.06)*	0.017 *(3.23)*	0.004 *(1.05)*	-0.005 *(-1.31)*	0.005 *(1.48)*
Experience at current school	0.011 *(2.69)*	-0.002 *(-0.74)*	0.016 *(5.49)*	0.001 *(0.43)*	0.001 *(0.24)*	0.017 *(3.44)*	0.002 *(0.54)*	0.008 *(1.75)*
Total teaching experience	-0.002 *(-0.37)*	-0.009 *(-2.24)*	-0.014 *(-3.21)*	-0.009 *(-2.9)*	-0.005 *(-0.85)*	-0.009 *(-1.65)*	0.000 *(-0.11)*	-0.011 *(-2.09)*
Female	-0.015 *(-0.39)*	0.065 *(1.27)*	0.111 *(3.34)*	0.128 *(3.37)*	-0.080 *(-1.25)*	-0.007 *(-0.14)*	-0.023 *(-0.55)*	0.048 *(1.12)*
Full-time employment	0.120 *(1.68)*	-0.165 *(-2.31)*	-0.035 *(-0.53)*	0.090 *(1.96)*	0.086 *(0.58)*	-0.104 *(-1.32)*	0.111 *(1.54)*	0.196 *(3.57)*
Head teacher is female	-0.077 *(-1.77)*	-0.057 *(-1.17)*	-0.037 *(-0.78)*	-0.025 *(-0.43)*	-0.068 *(-0.8)*	-0.081 *(-1.23)*	-0.082 *(-1.47)*	-0.111 *(-1.98)*
Is mentee	0.004 *(0.1)*	0.043 *(0.64)*	-0.036 *(-0.52)*	0.263 *(3.22)*	-0.033 *(-0.63)*	0.117 *(2.04)*	-0.016 *(-0.35)*	0.118 *(1.73)*
Is mentor	0.312 *(7.5)*	0.023 *(0.31)*	-0.121 *(-1.08)*	0.417 *(3.72)*	0.020 *(0.4)*	0.141 *(1.9)*	0.158 *(2.64)*	0.158 *(2.27)*
Permanent contract	-0.270 *(-5.02)*	-0.171 *(-3.16)*	-0.136 *(-2.53)*	-0.247 *(-3.92)*	-0.156 *(-2.68)*	-0.323 *(-4.16)*	-0.062 *(-0.74)*	-0.310 *(-4.94)*
Total weekly workload	-0.005 *(-6.37)*	-0.002 *(-1.05)*	-0.002 *(-1.33)*	-0.008 *(-4.15)*	-0.003 *(-2.57)*	-0.003 *(-2.33)*	-0.005 *(-2.98)*	-0.010 *(-5.41)*

Table A3.5: Chapter 3 – Model 4

Variable	Australia	Brazil	Bulgaria	Chile	Croatia	Czech Republic	Denmark	Estonia
Constant	-3.189	-2.882	-3.370	-2.967	-4.004	-2.725	-3.474	-2.922
	(-13.62)	*(-20.54)*	*(-12.62)*	*(-10.07)*	*(-21.91)*	*(-12.06)*	*(-9.73)*	*(-13.18)*
Cooperation with colleagues index	0.086	0.064	0.042	0.081	0.102	0.084	0.046	0.093
	(6.41)	*(8.27)*	*(3.16)*	*(7.15)*	*(7.64)*	*(6.08)*	*(2.55)*	*(7.15)*
Teacher's self-efficacy	0.084	0.099	0.130	0.078	0.162	0.122	0.136	0.078
	(5.69)	*(9.84)*	*(8.19)*	*(4.31)*	*(12.08)*	*(7.42)*	*(6.35)*	*(6.12)*
Teacher's age	0.006	0.005	0.000	0.010	0.006	0.002	0.007	0.004
	(1.47)	*(2.78)*	*(-0.14)*	*(2.09)*	*(2.36)*	*(0.49)*	*(1.66)*	*(1.88)*
Experience at current school	0.007	-0.001	-0.003	0.002	0.000	0.004	0.007	-0.002
	(1.15)	*(-0.26)*	*(-1.01)*	*(0.33)*	*(-0.04)*	*(1.29)*	*(1.69)*	*(-0.73)*
Total teaching experience	-0.009	0.001	0.002	-0.004	-0.001	-0.007	-0.012	-0.002
	(-2.44)	*(0.56)*	*(0.62)*	*(-0.89)*	*(-0.28)*	*(-2.1)*	*(-2.42)*	*(-0.58)*
Female	0.021	0.015	-0.052	0.103	0.164	0.108	0.050	0.043
	(0.29)	*(0.52)*	*(-1.03)*	*(1.94)*	*(4.61)*	*(2.22)*	*(1.02)*	*(0.85)*
Full-time employment	-0.075	0.021	-0.065	0.100	-0.055	-0.068	0.131	0.049
	(-0.62)	*(0.77)*	*(-0.71)*	*(1.41)*	*(-1.01)*	*(-1.27)*	*(1.53)*	*(1.16)*
Is mentee	-0.004	0.017	0.073	0.115	0.011	0.166	-0.041	0.087
	(-0.06)	*(0.27)*	*(0.98)*	*(1.01)*	*(0.22)*	*(2.44)*	*(-0.58)*	*(1.33)*
Is mentor	0.305	0.072	0.219	0.053	0.267	0.131	0.015	-0.022
	(4.49)	*(2.27)*	*(3.18)*	*(0.47)*	*(3.59)*	*(1.39)*	*(0.1)*	*(-0.2)*
Permanent contract	0.044	-0.119	-0.194	-0.136	-0.038	-0.116	0.100	-0.042
	(0.46)	*(-3.19)*	*(-2.68)*	*(-1.8)*	*(-0.6)*	*(-1.84)*	*(0.76)*	*(-0.75)*
Total weekly workload	-0.004	0.000	-0.001	0.001	-0.002	-0.004	-0.006	-0.004
	(-2.86)	*(-0.79)*	*(-0.62)*	*(0.7)*	*(-1.67)*	*(-2.3)*	*(-2.53)*	*(-3.01)*

Table A3.5 continued

Variable	Finland	France	Iceland	Israel	Italy	Japan	Korea	Latvia
Constant	-2.867 *(-13.46)*	-3.448 *(-10.57)*	-2.566 *(-7.73)*	-3.569 *(-15.81)*	-2.833 *(-9.01)*	-2.087 *(-7.68)*	-2.450 *(-6.09)*	-2.851 *(-9.24)*
Cooperation with colleagues index	0.056 *(4.15)*	0.127 *(7.01)*	0.094 *(4.73)*	0.070 *(3.46)*	0.081 *(5.52)*	0.074 *(5.35)*	0.034 *(2.65)*	0.082 *(4.55)*
Teacher's self-efficacy	0.134 *(12.96)*	0.133 *(8.6)*	0.061 *(3.22)*	0.124 *(9.88)*	0.130 *(8.48)*	0.111 *(8.13)*	0.095 *(9.32)*	0.092 *(4.27)*
Teacher's age	0.001 *(0.28)*	0.000 *(-0.05)*	0.008 *(1.5)*	0.013 *(3.54)*	-0.012 *(-2.97)*	0.005 *(1.34)*	-0.006 *(-0.97)*	0.005 *(1.28)*
Experience at current school	0.002 *(0.44)*	0.009 *(2.35)*	-0.004 *(-0.81)*	-0.003 *(-0.73)*	0.007 *(2.14)*	-0.002 *(-0.57)*	-0.013 *(-2.92)*	-0.001 *(-0.46)*
Total teaching experience	-0.004 *(-1.04)*	-0.009 *(-1.64)*	-0.007 *(-1.22)*	-0.006 *(-1.25)*	0.005 *(1.23)*	-0.010 *(-2.63)*	0.000 *(-0.01)*	0.001 *(0.14)*
Female	0.070 *(1.76)*	0.137 *(2.94)*	0.190 *(3.13)*	-0.060 *(-1.04)*	-0.026 *(-0.55)*	-0.060 *(-1.72)*	-0.004 *(-0.09)*	0.095 *(1.34)*
Full-time employment	0.058 *(0.66)*	0.086 *(1.58)*	0.048 *(0.51)*	0.032 *(0.46)*	-0.052 *(-0.7)*	-0.604 *(-5.97)*	0.230 *(0.83)*	-0.056 *(-0.88)*
Is mentee	0.135 *(1.49)*	0.086 *(0.95)*	0.045 *(0.5)*	0.005 *(0.12)*	-0.014 *(-0.15)*	0.041 *(0.86)*	0.115 *(2.23)*	0.043 *(0.53)*
Is mentor	0.079 *(0.77)*	0.127 *(1.12)*	0.385 *(2.73)*	-0.023 *(-0.45)*	0.133 *(1.03)*	0.151 *(3.63)*	0.196 *(3.74)*	0.174 *(1.76)*
Permanent contract	-0.148 *(-2.96)*	-0.193 *(-2.17)*	-0.055 *(-0.58)*	-0.095 *(-1.65)*	-0.031 *(-0.48)*	-0.127 *(-2.46)*	-0.192 *(-2.65)*	-0.002 *(-0.03)*
Total weekly workload	-0.002 *(-0.83)*	-0.005 *(-2.78)*	-0.005 *(-1.97)*	0.001 *(0.47)*	-0.003 *(-1.77)*	0.001 *(0.54)*	-0.003 *(-2.54)*	-0.002 *(-2.01)*

Table A3.5 continued

Variable	Malaysia	Mexico	Netherlands	Norway	Poland	Portugal	Romania	Serbia
Constant	-3.429 *(-6.35)*	-3.495 *(-14.83)*	-2.729 *(-5.65)*	-3.321 *(-13.33)*	-3.099 *(-13.51)*	-3.797 *(-15.04)*	-4.075 *(-13.42)*	-3.846 *(-21.71)*
Cooperation with colleagues index	0.057 *(4.95)*	0.058 *(6.25)*	0.075 *(2.74)*	0.091 *(5.69)*	0.053 *(3.15)*	0.065 *(5.68)*	0.048 *(3.8)*	0.096 *(9.35)*
Teacher's self-efficacy	0.153 *(10.88)*	0.131 *(9.33)*	0.196 *(6.97)*	0.118 *(5.23)*	0.123 *(7.87)*	0.131 *(8.43)*	0.145 *(7.84)*	0.149 *(11.98)*
Teacher's age	0.000 *(-0.01)*	0.000 *(-0.09)*	-0.003 *(-0.7)*	0.013 *(4.17)*	0.004 *(0.81)*	0.008 *(1.88)*	-0.006 *(-1.32)*	0.002 *(0.72)*
Experience at current school	0.007 *(1.72)*	0.006 *(1.62)*	0.006 *(1.37)*	0.004 *(1.02)*	0.003 *(0.8)*	0.013 *(3.16)*	0.006 *(1.92)*	0.000 *(0.1)*
Total teaching experience	-0.001 *(-0.25)*	-0.002 *(-0.55)*	-0.011 *(-2.42)*	-0.013 *(-2.67)*	-0.004 *(-0.78)*	-0.009 *(-2.01)*	0.008 *(1.49)*	0.005 *(1.71)*
Female	-0.002 *(-0.05)*	-0.005 *(-0.11)*	-0.029 *(-0.42)*	0.148 *(3.1)*	0.083 *(1.79)*	0.096 *(2.44)*	0.132 *(2.66)*	0.060 *(1.99)*
Full-time employment	0.091 *(0.64)*	0.003 *(0.07)*	0.190 *(2.52)*	0.041 *(0.6)*	0.023 *(0.39)*	-0.043 *(-0.53)*	-0.114 *(-1.53)*	0.019 *(0.43)*
Is mentee	0.077 *(1.59)*	0.010 *(0.17)*	0.011 *(0.15)*	0.080 *(1.37)*	0.137 *(2.52)*	0.045 *(0.61)*	-0.025 *(-0.38)*	-0.024 *(-0.47)*
Is mentor	0.094 *(2.38)*	0.040 *(0.7)*	-0.082 *(-1.06)*	0.295 *(1.71)*	0.133 *(2.18)*	0.175 *(2.24)*	0.100 *(1.59)*	0.145 *(2.07)*
Permanent contract	-0.324 *(-0.72)*	0.052 *(1.06)*	-0.326 *(-3.78)*	-0.199 *(-2.13)*	-0.261 *(-3.72)*	-0.033 *(-0.64)*	-0.137 *(-3.02)*	-0.381 *(-8.51)*
Total weekly workload	-0.001 *(-0.75)*	0.000 *(-0.1)*	-0.008 *(-3.27)*	-0.009 *(-3.81)*	-0.001 *(-0.56)*	-0.003 *(-1.78)*	-0.002 *(-1.4)*	-0.002 *(-1.92)*

Table A3.5 continued

Variable	Singapore	Slovak Republic	Spain	Sweden	Abu Dhabi (UAE)	Alberta (Canada)	England (UK)	Flanders (BEL)
Constant	-2.949 *(-12.59)*	-3.000 *(-11.51)*	-3.195 *(-10.98)*	-3.225 *(-11.7)*	-4.020 *(-13.82)*	-2.583 *(-8.84)*	-3.118 *(-12.00)*	-2.420 *(-8.92)*
Cooperation with colleagues index	0.096 *(7.79)*	0.065 *(5.3)*	0.104 *(5.78)*	0.133 *(9.29)*	0.067 *(7.29)*	0.088 *(4.94)*	0.120 *(10.23)*	0.072 *(3.7)*
Teacher's self-efficacy	0.069 *(8.34)*	0.106 *(6.93)*	0.129 *(9.7)*	0.094 *(6.85)*	0.096 *(6.36)*	0.071 *(4.63)*	0.107 *(7.34)*	0.102 *(7.13)*
Teacher's age	0.014 *(3.5)*	0.012 *(3.14)*	0.001 *(0.32)*	0.009 *(3.19)*	0.012 *(2.6)*	0.001 *(0.15)*	-0.005 *(-1.37)*	0.009 *(2.79)*
Experience at current school	0.011 *(2.8)*	-0.001 *(-0.35)*	0.007 *(2.1)*	0.002 *(0.73)*	0.002 *(0.41)*	0.012 *(2.72)*	-0.001 *(-0.31)*	0.003 *(0.81)*
Total teaching experience	-0.003 *(-0.62)*	-0.009 *(-2.26)*	-0.011 *(-2.16)*	-0.010 *(-3.49)*	0.001 *(0.21)*	-0.004 *(-0.64)*	0.003 *(0.77)*	-0.013 *(-2.86)*
Female	-0.022 *(-0.54)*	0.087 *(1.76)*	0.127 *(3.19)*	0.136 *(3.58)*	0.063 *(0.91)*	0.016 *(0.29)*	-0.027 *(-0.51)*	0.063 *(1.41)*
Full-time employment	0.144 *(1.96)*	-0.152 *(-2.1)*	0.116 *(1.67)*	0.094 *(1.99)*	-0.066 *(-0.49)*	-0.075 *(-1.00)*	0.110 *(1.24)*	0.198 *(3.46)*
Is mentee	0.039 *(0.97)*	0.083 *(1.22)*	-0.063 *(-0.94)*	0.336 *(4.16)*	0.003 *(0.06)*	0.102 *(1.71)*	-0.015 *(-0.29)*	0.094 *(1.55)*
Is mentor	0.309 *(7.7)*	0.122 *(1.63)*	-0.108 *(-1.16)*	0.344 *(3.1)*	0.074 *(1.93)*	0.099 *(1.26)*	0.179 *(2.94)*	0.150 *(1.94)*
Permanent contract	-0.256 *(-4.49)*	-0.168 *(-3.03)*	-0.130 *(-2.17)*	-0.243 *(-3.67)*	-0.099 *(-1.9)*	-0.375 *(-4.37)*	-0.100 *(-1.07)*	-0.279 *(-3.98)*
Total weekly workload	-0.004 *(-5.14)*	-0.003 *(-1.96)*	-0.003 *(-2.7)*	-0.008 *(-3.79)*	-0.001 *(-1.62)*	-0.002 *(-1.48)*	-0.005 *(-3.03)*	-0.010 *(-5.52)*

Table A3.6: Chapter 3 – Model 1 (different activities)

Variable	Australia	Brazil	Bulgaria	Chile	Croatia	Czech Republic	Denmark	Estonia
Constant	0.100 *(1.03)*	0.050 *(1.2)*	-0.028 *(-0.37)*	-0.040 *(-0.45)*	-0.069 *(-0.98)*	0.208 *(3.07)*	0.414 *(3.43)*	0.176 *(3.09)*
Face-to-face teaching (hours)	-0.005 *(-1.39)*	-0.002 *(-1.4)*	0.001 *(0.47)*	-0.001 *(-0.29)*	0.002 *(0.59)*	-0.014 *(-4.03)*	-0.016 *(-3.08)*	-0.005 *(-2.06)*
Planning (hours)	-0.010 *(-2.06)*	0.005 *(1.55)*	-0.001 *(-0.2)*	0.009 *(2.02)*	0.003 *(0.94)*	-0.002 *(-0.44)*	0.003 *(0.39)*	-0.004 *(-0.98)*
Team work (hours)	0.014 *(1.77)*	0.008 *(1.79)*	0.005 *(0.38)*	0.021 *(1.34)*	0.026 *(2.38)*	0.002 *(0.15)*	-0.017 *(-1.27)*	0.015 *(1.33)*
Marking (hours)	-0.004 *(-0.73)*	-0.003 *(-0.83)*	-0.011 *(-1.41)*	0.000 *(-0.02)*	-0.007 *(-1.07)*	-0.007 *(-1.05)*	0.002 *(0.29)*	-0.022 *(-3.29)*
Student counselling (hours)	-0.003 *(-0.35)*	-0.001 *(-0.39)*	-0.006 *(-0.42)*	-0.011 *(-0.76)*	-0.025 *(-2.27)*	0.012 *(1.11)*	-0.003 *(-0.25)*	0.005 *(0.49)*
School management (hours)	0.012 *(2.21)*	0.017 *(4.68)*	0.019 *(1.98)*	0.023 *(2.71)*	0.021 *(1.01)*	0.017 *(2.66)*	-0.003 *(-0.38)*	0.013 *(1.57)*
Admin duties (hours)	0.003 *(0.44)*	0.001 *(0.24)*	-0.010 *(-1.06)*	0.001 *(0.08)*	0.000 *(0.04)*	0.001 *(0.12)*	-0.032 *(-2.93)*	-0.016 *(-1.43)*
Interaction with parents (hours)	-0.006 *(-0.28)*	0.008 *(1.35)*	0.015 *(1.1)*	-0.004 *(-0.22)*	0.003 *(0.44)*	0.027 *(1.29)*	0.003 *(0.28)*	0.024 *(1.15)*
Extra-curricular activities (hours)	0.015 *(1.97)*	0.005 *(1.09)*	0.012 *(1.45)*	0.013 *(1.26)*	0.013 *(1.81)*	0.016 *(1.59)*	0.013 *(1.23)*	0.019 *(2.76)*
Other (hours)	-0.010 *(-1.19)*	-0.002 *(-0.52)*	0.004 *(0.37)*	-0.021 *(-2.16)*	0.002 *(0.25)*	-0.013 *(-1.9)*	-0.002 *(-0.31)*	0.003 *(0.32)*

Table A3.6 continued

Variable	Finland	France	Iceland	Israel	Italy	Japan	Korea	Latvia
Constant	-0.084 *(-1.1)*	-0.084 *(-0.82)*	0.032 *(0.29)*	-0.072 *(-1.12)*	-0.019 *(-0.21)*	0.117 *(1.56)*	0.074 *(1.07)*	0.028 *(0.48)*
Face-to-face teaching (hours)	0.008 *(2.41)*	0.010 *(1.98)*	-0.003 *(-0.58)*	0.003 *(1.01)*	0.002 *(0.46)*	-0.008 *(-2.3)*	-0.005 *(-1.56)*	0.001 *(0.2)*
Planning (hours)	-0.009 *(-1.37)*	-0.007 *(-1.76)*	-0.006 *(-1.18)*	0.004 *(0.63)*	0.003 *(0.48)*	0.000 *(0.14)*	-0.005 *(-1.25)*	0.002 *(0.3)*
Team work (hours)	0.021 *(1.2)*	0.023 *(1.91)*	-0.005 *(-0.3)*	0.017 *(1.74)*	-0.002 *(-0.33)*	0.012 *(2.18)*	0.032 *(2.94)*	0.023 *(1.46)*
Marking (hours)	0.006 *(0.83)*	-0.014 *(-3.18)*	0.002 *(0.18)*	-0.011 *(-2.21)*	-0.008 *(-1.4)*	-0.002 *(-0.48)*	-0.007 *(-1.03)*	-0.023 *(-2.77)*
Student counselling (hours)	-0.014 *(-0.9)*	0.018 *(1.14)*	0.019 *(1.09)*	0.024 *(2.64)*	0.011 *(1.44)*	-0.013 *(-2.05)*	0.006 *(1.16)*	0.010 *(1.09)*
School management (hours)	0.027 *(2.19)*	0.058 *(4.51)*	0.003 *(0.43)*	0.022 *(4.93)*	0.032 *(4.47)*	0.004 *(1.11)*	0.003 *(0.29)*	0.020 *(1.38)*
Admin duties (hours)	-0.010 *(-0.62)*	-0.040 *(-2.7)*	0.009 *(0.74)*	-0.010 *(-0.85)*	-0.007 *(-0.63)*	-0.004 *(-1.24)*	-0.011 *(-2.3)*	-0.006 *(-0.78)*
Interaction with parents (hours)	-0.042 *(-1.7)*	0.002 *(0.16)*	0.020 *(1.02)*	0.005 *(0.31)*	-0.026 *(-1.63)*	0.004 *(0.34)*	0.015 *(1.31)*	-0.011 *(-0.59)*
Extra-curricular activities (hours)	-0.001 *(-0.11)*	0.013 *(1.35)*	-0.001 *(-0.07)*	0.001 *(0.34)*	0.019 *(2.18)*	0.007 *(2.62)*	0.023 *(2.43)*	0.002 *(0.16)*
Other (hours)	0.010 *(0.92)*	-0.016 *(-1.85)*	-0.012 *(-1.38)*	-0.013 *(-1.99)*	0.003 *(0.6)*	0.002 *(0.52)*	-0.017 *(-2.44)*	-0.002 *(-0.22)*

Table A3.6 continued

Variable	Malaysia	Mexico	Netherlands	Norway	Poland	Portugal	Romania	Serbia
Constant	0.013 *(0.18)*	-0.137 *(-2.63)*	-0.039 *(-0.35)*	0.216 *(1.8)*	-0.033 *(-0.48)*	0.023 *(0.28)*	0.007 *(0.09)*	0.056 *(1.14)*
Face-to-face teaching (hours)	-0.001 *(-0.22)*	0.002 *(0.8)*	0.008 *(1.64)*	-0.012 *(-2.19)*	0.001 *(0.23)*	-0.001 *(-0.43)*	-0.003 *(-0.9)*	-0.003 *(-1.42)*
Planning (hours)	-0.005 *(-1.05)*	0.000 *(-0.04)*	0.000 *(-0.04)*	-0.007 *(-1.47)*	0.001 *(0.22)*	0.000 *(0.09)*	-0.006 *(-1.6)*	0.006 *(1.3)*
Team work (hours)	0.003 *(0.32)*	0.016 *(1.65)*	0.000 *(0.02)*	0.028 *(1.59)*	-0.008 *(-0.6)*	-0.012 *(-1.81)*	0.012 *(0.95)*	0.010 *(1.46)*
Marking (hours)	-0.008 *(-1.61)*	0.004 *(1.15)*	-0.015 *(-2.05)*	-0.008 *(-1.7)*	-0.004 *(-0.7)*	-0.003 *(-1.04)*	-0.001 *(-0.06)*	-0.006 *(-0.91)*
Student counselling (hours)	0.012 *(1.14)*	-0.002 *(-0.34)*	0.007 *(0.42)*	-0.008 *(-0.69)*	0.033 *(3.47)*	0.023 *(2.3)*	-0.004 *(-0.38)*	0.015 *(1.7)*
School management (hours)	0.009 *(1.21)*	0.002 *(0.27)*	0.007 *(0.94)*	0.001 *(0.08)*	0.018 *(2.67)*	0.013 *(3.38)*	0.015 *(1.76)*	0.011 *(1.12)*
Admin duties (hours)	-0.001 *(-0.17)*	-0.007 *(-0.92)*	-0.032 *(-1.97)*	-0.009 *(-1.13)*	-0.020 *(-1.71)*	-0.013 *(-1.83)*	0.009 *(0.67)*	-0.020 *(-1.58)*
Interaction with parents (hours)	0.023 *(2.36)*	0.009 *(1.54)*	0.027 *(1.13)*	0.012 *(0.7)*	0.006 *(0.31)*	0.025 *(2.2)*	0.035 *(1.98)*	0.008 *(0.56)*
Extra-curricular activities (hours)	0.001 *(0.23)*	0.006 *(0.99)*	-0.025 *(-1.48)*	-0.006 *(-0.53)*	0.008 *(0.96)*	0.019 *(3.07)*	0.007 *(1.05)*	-0.004 *(-0.54)*
Other (hours)	-0.006 *(-0.76)*	0.006 *(1.03)*	-0.003 *(-0.44)*	0.015 *(0.73)*	0.002 *(0.32)*	-0.004 *(-0.73)*	-0.007 *(-0.82)*	-0.004 *(-0.59)*

Table A3.6 continued

Variable	Singapore	Slovak Republic	Spain	Sweden	Abu Dhabi (UAE)	Alberta (Canada)	England (UK)	Flanders (BEL)
Constant	0.234	0.111	-0.066	0.092	0.140	0.155	0.066	0.169
	(4.5)	*(1.68)*	*(-0.91)*	*(0.95)*	*(1.58)*	*(1.78)*	*(0.67)*	*(2.07)*
Face-to-face teaching (hours)	-0.006	-0.006	0.000	0.002	0.000	-0.002	-0.001	0.003
	(-2.12)	*(-1.92)*	*(0.12)*	*(0.55)*	*(0.13)*	*(-0.81)*	*(-0.18)*	*(0.86)*
Planning (hours)	-0.001	-0.005	-0.009	-0.004	-0.011	-0.012	0.000	-0.001
	(-0.27)	*(-1.00)*	*(-1.87)*	*(-0.76)*	*(-1.7)*	*(-2.03)*	*(-0.05)*	*(-0.24)*
Team work (hours)	0.007	0.015	0.025	0.035	0.012	0.013	0.016	-0.012
	(0.9)	*(1.23)*	*(2.11)*	*(3.6)*	*(1.06)*	*(1.85)*	*(1.82)*	*(-0.95)*
Marking (hours)	-0.010	-0.002	0.001	-0.018	0.000	0.000	-0.023	-0.015
	(-3.19)	*(-0.36)*	*(0.16)*	*(-2.86)*	*(-0.03)*	*(0.03)*	*(-4.15)*	*(-2.44)*
Student counselling (hours)	0.014	-0.005	0.016	0.003	-0.016	-0.005	0.024	-0.008
	(1.42)	*(-0.46)*	*(0.99)*	*(0.53)*	*(-2.02)*	*(-0.92)*	*(2.33)*	*(-0.82)*
School management (hours)	0.022	0.021	0.028	0.026	-0.002	0.003	0.020	0.038
	(3.86)	*(3.44)*	*(4.94)*	*(2.04)*	*(-0.2)*	*(0.63)*	*(4.25)*	*(3.68)*
Admin duties (hours)	-0.011	-0.012	-0.043	-0.024	0.001	-0.004	-0.021	-0.036
	(-2.67)	*(-1.39)*	*(-4.04)*	*(-4.38)*	*(0.12)*	*(-0.38)*	*(-3.27)*	*(-4.02)*
Interaction with parents (hours)	-0.012	0.042	0.046	-0.013	-0.003	-0.003	0.028	0.009
	(-0.77)	*(2.15)*	*(2.08)*	*(-0.93)*	*(-0.29)*	*(-0.17)*	*(1.66)*	*(0.45)*
Extra-curricular activities (hours)	-0.007	-0.001	0.022	0.041	0.008	0.009	0.017	-0.001
	(-0.97)	*(-0.1)*	*(3.4)*	*(2.84)*	*(0.8)*	*(1.55)*	*(2.39)*	*(-0.13)*
Other (hours)	-0.009	0.002	-0.009	-0.022	0.001	-0.015	-0.005	-0.007
	(-1.86)	*(0.3)*	*(-1.41)*	*(-2.99)*	*(0.15)*	*(-1.75)*	*(-0.9)*	*(-1.00)*

Table A3.7: Chapter 3 – Model 2 (different activities)

Variable	Australia	Brazil	Bulgaria	Chile	Croatia	Czech Republic	Denmark	Estonia
Constant	-3.151	-3.070	-3.300	-3.218	-4.039	-2.508	-3.593	-2.863
	(-12.94)	*(-19.68)*	*(-11.57)*	*(-12.73)*	*(-20.65)*	*(-10.93)*	*(-10.21)*	*(-11.84)*
Cooperation with colleagues index	0.072	0.058	0.024	0.076	0.104	0.085	0.067	0.094
	(4.64)	*(7.51)*	*(1.77)*	*(7.01)*	*(7.74)*	*(6.68)*	*(3.92)*	*(5.59)*
Teacher's self-efficacy	0.099	0.115	0.135	0.089	0.154	0.118	0.150	0.082
	(5.94)	*(10.91)*	*(9.00)*	*(5.68)*	*(12.24)*	*(7.4)*	*(6.25)*	*(6.02)*
Teacher's age	0.004	0.002	-0.004	0.012	0.006	-0.001	0.006	0.004
	(0.97)	*(1.21)*	*(-1.07)*	*(3.3)*	*(2.38)*	*(-0.2)*	*(1.17)*	*(1.57)*
Experience at current school	0.009	0.000	-0.003	-0.002	0.000	0.002	0.015	-0.002
	(1.5)	*(0.18)*	*(-0.92)*	*(-0.44)*	*(-0.12)*	*(0.87)*	*(3.15)*	*(-0.69)*
Total teaching experience	-0.009	0.002	0.003	-0.004	-0.002	-0.006	-0.017	0.000
	(-2.09)	*(0.8)*	*(0.93)*	*(-0.9)*	*(-0.6)*	*(-1.69)*	*(-3.09)*	*(0.1)*
Face-to-face teaching (hours)	-0.007	-0.003	-0.002	-0.003	0.000	-0.016	-0.020	-0.010
	(-2.34)	*(-2.49)*	*(-0.61)*	*(-1.26)*	*(-0.05)*	*(-3.92)*	*(-3.97)*	*(-3.74)*
Planning (hours)	-0.009	0.003	0.000	0.010	0.001	-0.002	-0.002	-0.003
	(-1.72)	*(0.98)*	*(0.03)*	*(2.68)*	*(0.28)*	*(-0.38)*	*(-0.36)*	*(-0.6)*
Team work (hours)	0.002	0.001	-0.002	0.013	0.005	-0.027	-0.017	0.001
	(0.32)	*(0.28)*	*(-0.16)*	*(0.97)*	*(0.47)*	*(-2.02)*	*(-1.3)*	*(0.07)*
Marking (hours)	-0.010	0.000	-0.015	-0.002	-0.016	-0.011	-0.010	-0.024
	(-1.86)	*(-0.03)*	*(-2.11)*	*(-0.19)*	*(-2.73)*	*(-1.61)*	*(-1.29)*	*(-3.45)*
Student counselling (hours)	-0.001	-0.004	-0.013	-0.014	-0.032	0.009	-0.012	0.000
	(-0.11)	*(-1.22)*	*(-0.88)*	*(-1.23)*	*(-2.66)*	*(0.89)*	*(-1.00)*	*(-0.04)*
School management (hours)	0.007	0.009	0.009	0.012	0.007	0.013	-0.009	0.007
	(1.35)	*(2.59)*	*(1.08)*	*(1.32)*	*(0.36)*	*(1.93)*	*(-1.29)*	*(0.8)*
Admin duties (hours)	0.002	0.002	-0.008	-0.001	0.004	-0.001	-0.032	-0.016
	(0.25)	*(0.42)*	*(-0.91)*	*(-0.09)*	*(0.67)*	*(-0.14)*	*(-3.45)*	*(-1.41)*
Interaction with parents (hours)	-0.017	0.001	0.010	-0.012	-0.001	0.001	-0.002	-0.004
	(-0.88)	*(0.19)*	*(0.76)*	*(-0.61)*	*(-0.13)*	*(0.03)*	*(-0.2)*	*(-0.19)*
Extra-curricular activities (hours)	0.011	0.002	0.008	0.021	0.006	0.011	0.012	0.017
	(1.64)	*(0.6)*	*(1.03)*	*(2.52)*	*(0.91)*	*(1.33)*	*(1.18)*	*(2.44)*
Other (hours)	-0.012	-0.005	0.000	-0.019	0.002	-0.012	-0.004	0.005
	(-1.61)	*(-1.27)*	*(-0.05)*	*(-2.71)*	*(0.27)*	*(-1.94)*	*(-0.67)*	*(0.58)*
Female	0.047	-0.001	-0.042	0.145	0.194	0.111	0.010	0.053
	(0.75)	*(-0.02)*	*(-0.81)*	*(2.91)*	*(5.33)*	*(2.23)*	*(0.17)*	*(0.98)*
Full-time employment	0.006	0.039	-0.047	0.066	-0.026	-0.029	0.117	0.132
	(0.06)	*(1.25)*	*(-0.52)*	*(1.01)*	*(-0.46)*	*(-0.48)*	*(1.26)*	*(2.76)*
Is mentee	-0.086	-0.006	0.016	0.043	0.023	0.121	0.033	0.010
	(-1.41)	*(-0.08)*	*(0.23)*	*(0.38)*	*(0.47)*	*(1.72)*	*(0.45)*	*(0.15)*
Is mentor	0.360	0.077	0.268	-0.077	0.242	0.188	0.079	-0.079
	(4.75)	*(2.5)*	*(4.28)*	*(-0.65)*	*(3.11)*	*(2.22)*	*(0.49)*	*(-0.73)*
Permanent contract	-0.075	-0.046	-0.106	-0.028	-0.062	-0.061	0.039	0.005
	(-1.00)	*(-1.21)*	*(-1.49)*	*(-0.45)*	*(-0.96)*	*(-0.95)*	*(0.26)*	*(0.09)*

Table A3.7 continued

Variable	Finland	France	Iceland	Israel	Italy	Japan	Korea	Latvia
Constant	-2.997 *(-13.11)*	-3.399 *(-11.71)*	-2.374 *(-6.85)*	-3.642 *(-17.26)*	-2.922 *(-11.31)*	-2.533 *(-10.41)*	-2.487 *(-8.36)*	-3.063 *(-9.53)*
Cooperation with colleagues index	0.064 *(4.44)*	0.096 *(5.33)*	0.101 *(5.13)*	0.083 *(5.42)*	0.078 *(6.53)*	0.068 *(5.01)*	0.018 *(1.54)*	0.081 *(4.47)*
Teacher's self-efficacy	0.133 *(11.08)*	0.122 *(7.71)*	0.048 *(2.44)*	0.116 *(9.39)*	0.130 *(9.3)*	0.123 *(9.13)*	0.089 *(9.41)*	0.104 *(4.58)*
Teacher's age	-0.001 *(-0.27)*	0.001 *(0.24)*	0.005 *(0.97)*	0.025 *(7.61)*	-0.008 *(-2.09)*	0.006 *(1.53)*	-0.004 *(-0.81)*	0.007 *(1.89)*
Experience at current school	0.002 *(0.63)*	0.012 *(3.33)*	-0.005 *(-0.9)*	-0.005 *(-1.36)*	0.011 *(3.46)*	0.004 *(0.9)*	-0.005 *(-1.47)*	-0.005 *(-1.72)*
Total teaching experience	-0.002 *(-0.34)*	-0.010 *(-1.86)*	-0.005 *(-0.86)*	-0.012 *(-3.08)*	-0.001 *(-0.27)*	-0.011 *(-2.68)*	-0.001 *(-0.29)*	0.001 *(0.24)*
Face-to-face teaching (hours)	0.002 *(0.67)*	0.007 *(1.24)*	-0.004 *(-0.71)*	-0.001 *(-0.27)*	-0.002 *(-0.36)*	-0.005 *(-1.55)*	-0.007 *(-2.41)*	-0.002 *(-1.00)*
Planning (hours)	-0.009 *(-1.61)*	-0.005 *(-1.32)*	-0.008 *(-1.58)*	0.003 *(0.53)*	-0.001 *(-0.17)*	0.000 *(-0.11)*	-0.004 *(-1.09)*	-0.002 *(-0.46)*
Team work (hours)	-0.006 *(-0.4)*	0.010 *(0.8)*	-0.030 *(-1.88)*	0.006 *(0.64)*	-0.004 *(-0.67)*	0.009 *(1.89)*	0.029 *(2.58)*	0.018 *(1.27)*
Marking (hours)	0.005 *(0.85)*	-0.018 *(-3.91)*	-0.002 *(-0.21)*	-0.014 *(-3.13)*	-0.014 *(-2.3)*	0.000 *(0.04)*	-0.009 *(-1.45)*	-0.022 *(-2.62)*
Student counselling (hours)	-0.009 *(-0.6)*	-0.003 *(-0.24)*	0.014 *(1.00)*	0.021 *(2.32)*	0.004 *(0.49)*	-0.013 *(-2.06)*	-0.003 *(-0.44)*	0.007 *(0.83)*
School management (hours)	0.023 *(1.71)*	0.054 *(4.02)*	-0.003 *(-0.37)*	0.013 *(3.48)*	0.019 *(2.89)*	-0.001 *(-0.4)*	-0.006 *(-0.8)*	0.009 *(0.78)*
Admin duties (hours)	-0.020 *(-1.41)*	-0.029 *(-2.00)*	0.002 *(0.19)*	-0.010 *(-0.89)*	-0.008 *(-0.7)*	-0.002 *(-0.74)*	-0.010 *(-2.3)*	-0.006 *(-0.89)*
Interaction with parents (hours)	-0.041 *(-1.62)*	-0.002 *(-0.15)*	0.020 *(0.81)*	-0.018 *(-1.11)*	-0.025 *(-1.67)*	0.001 *(0.1)*	0.003 *(0.34)*	-0.028 *(-1.59)*
Extra-curricular activities (hours)	-0.001 *(-0.09)*	0.002 *(0.27)*	-0.004 *(-0.26)*	0.000 *(-0.00)*	0.016 *(1.86)*	0.004 *(1.45)*	0.014 *(1.55)*	0.005 *(0.39)*
Other (hours)	0.008 *(0.91)*	-0.015 *(-1.7)*	-0.002 *(-0.22)*	-0.015 *(-2.61)*	-0.002 *(-0.39)*	0.002 *(0.57)*	-0.008 *(-1.36)*	-0.002 *(-0.3)*
Female	0.062 *(1.34)*	0.116 *(2.89)*	0.274 *(3.99)*	0.004 *(0.07)*	0.017 *(0.39)*	-0.051 *(-1.44)*	-0.043 *(-0.98)*	0.140 *(1.77)*
Full-time employment	0.128 *(1.43)*	-0.004 *(-0.07)*	0.075 *(0.79)*	0.042 *(0.65)*	-0.094 *(-1.15)*	-0.457 *(-4.4)*	0.209 *(1.32)*	-0.027 *(-0.41)*
Is mentee	0.032 *(0.29)*	0.027 *(0.28)*	0.085 *(1.02)*	0.037 *(0.88)*	0.032 *(0.42)*	-0.015 *(-0.29)*	0.109 *(2.39)*	0.084 *(0.99)*
Is mentor	0.056 *(0.52)*	0.175 *(1.38)*	0.357 *(2.42)*	-0.062 *(-1.31)*	0.029 *(0.27)*	0.207 *(5.3)*	0.213 *(4.72)*	0.051 *(0.48)*
Permanent contract	-0.172 *(-3.07)*	-0.277 *(-2.9)*	-0.113 *(-1.27)*	-0.084 *(-1.52)*	0.017 *(0.28)*	-0.106 *(-2.02)*	-0.154 *(-2.43)*	0.036 *(0.45)*

Table A3.7 continued

Variable	Malaysia	Mexico	Netherlands	Norway	Poland	Portugal	Romania	Serbia
Constant	-3.629 *(-6.78)*	-3.657 *(-16.48)*	-3.127 *(-7.44)*	-3.399 *(-11.68)*	-3.173 *(-13.73)*	-3.717 *(-16.54)*	-4.035 *(-13.00)*	-3.792 *(-20.16)*
Cooperation with colleagues index	0.055 *(5.05)*	0.071 *(7.65)*	0.092 *(3.84)*	0.113 *(6.55)*	0.047 *(2.71)*	0.065 *(4.89)*	0.058 *(4.23)*	0.094 *(9.82)*
Teacher's self-efficacy	0.168 *(12.92)*	0.133 *(9.56)*	0.183 *(9.46)*	0.126 *(6.16)*	0.133 *(8.71)*	0.128 *(9.19)*	0.137 *(6.84)*	0.151 *(12.56)*
Teacher's age	0.003 *(0.72)*	0.001 *(0.23)*	-0.004 *(-0.81)*	0.009 *(3.05)*	0.003 *(0.53)*	0.011 *(2.68)*	-0.007 *(-1.35)*	0.001 *(0.38)*
Experience at current school	0.006 *(1.8)*	0.008 *(2.68)*	0.008 *(1.86)*	0.006 *(1.56)*	0.003 *(0.89)*	0.017 *(4.08)*	0.006 *(1.89)*	0.003 *(1.12)*
Total teaching experience	-0.003 *(-0.74)*	0.001 *(0.3)*	-0.010 *(-1.68)*	-0.011 *(-2.47)*	-0.004 *(-0.69)*	-0.015 *(-3.18)*	0.009 *(1.67)*	0.004 *(1.28)*
Face-to-face teaching (hours)	-0.001 *(-0.51)*	-0.001 *(-0.77)*	-0.005 *(-0.87)*	-0.016 *(-3.00)*	0.000 *(-0.05)*	-0.004 *(-1.12)*	-0.004 *(-1.23)*	-0.004 *(-1.8)*
Planning (hours)	-0.006 *(-1.23)*	0.001 *(0.37)*	0.005 *(0.76)*	-0.004 *(-0.85)*	0.001 *(0.26)*	-0.001 *(-0.27)*	-0.004 *(-1.04)*	0.005 *(1.2)*
Team work (hours)	0.002 *(0.23)*	0.000 *(0.01)*	-0.020 *(-1.41)*	-0.005 *(-0.28)*	-0.015 *(-1.19)*	-0.016 *(-2.6)*	0.002 *(0.14)*	0.000 *(-0.06)*
Marking (hours)	-0.006 *(-1.27)*	0.004 *(1.07)*	-0.013 *(-1.69)*	-0.009 *(-2.19)*	-0.011 *(-1.96)*	-0.003 *(-0.99)*	-0.010 *(-1.13)*	-0.012 *(-2.1)*
Student counselling (hours)	-0.005 *(-0.54)*	-0.008 *(-1.64)*	0.006 *(0.35)*	-0.014 *(-1.35)*	0.024 *(2.62)*	0.020 *(2.2)*	0.000 *(0.02)*	0.006 *(0.74)*
School management (hours)	0.004 *(0.64)*	-0.004 *(-0.51)*	0.000 *(-0.04)*	-0.003 *(-0.4)*	0.008 *(1.28)*	0.007 *(1.95)*	0.006 *(0.71)*	0.001 *(0.11)*
Admin duties (hours)	0.000 *(0.08)*	-0.008 *(-1.03)*	-0.027 *(-1.45)*	-0.014 *(-1.95)*	-0.023 *(-2.16)*	-0.013 *(-1.86)*	0.004 *(0.28)*	-0.013 *(-1.26)*
Interaction with parents (hours)	0.011 *(1.36)*	0.007 *(1.27)*	0.010 *(0.42)*	-0.001 *(-0.1)*	0.014 *(0.75)*	0.020 *(1.96)*	0.003 *(0.16)*	-0.009 *(-0.61)*
Extra-curricular activities (hours)	0.001 *(0.11)*	-0.004 *(-0.71)*	-0.026 *(-1.8)*	-0.010 *(-0.9)*	0.004 *(0.5)*	0.011 *(1.94)*	0.009 *(1.36)*	-0.006 *(-1.01)*
Other (hours)	-0.006 *(-0.98)*	0.004 *(0.85)*	-0.005 *(-0.63)*	0.006 *(0.6)*	0.005 *(0.78)*	-0.002 *(-0.25)*	-0.008 *(-1.05)*	-0.006 *(-0.88)*
Female	-0.052 *(-1.12)*	-0.017 *(-0.42)*	0.013 *(0.22)*	0.128 *(2.42)*	0.096 *(2.3)*	0.046 *(1.06)*	0.132 *(2.51)*	0.088 *(2.76)*
Full-time employment	0.158 *(1.06)*	-0.012 *(-0.23)*	0.155 *(2.27)*	0.071 *(0.94)*	-0.019 *(-0.29)*	-0.099 *(-1.09)*	-0.075 *(-1.03)*	0.074 *(1.57)*
Is mentee	0.060 *(1.35)*	0.031 *(0.5)*	0.083 *(1.05)*	0.034 *(0.48)*	0.159 *(3.1)*	0.046 *(0.62)*	-0.011 *(-0.14)*	-0.022 *(-0.47)*
Is mentor	0.102 *(2.84)*	0.017 *(0.33)*	-0.066 *(-1.05)*	0.257 *(1.98)*	0.124 *(2.05)*	0.135 *(1.6)*	0.117 *(1.67)*	0.131 *(1.77)*
Permanent contract	-0.456 *(-1.03)*	0.025 *(0.49)*	-0.232 *(-2.7)*	-0.244 *(-2.9)*	-0.244 *(-3.46)*	-0.041 *(-0.75)*	-0.081 *(-1.68)*	-0.356 *(-7.61)*

Table A3.7 continued

Variable	Singapore	Slovak Republic	Spain	Sweden	Abu Dhabi (UAE)	Alberta (Canada)	England (UK)	Flanders (BEL)
Constant	-2.939 *(-11.75)*	-3.014 *(-11.44)*	-3.263 *(-12.2)*	-3.305 *(-13.69)*	-4.459 *(-11.73)*	-2.638 *(-9.61)*	-3.202 *(-13.22)*	-2.574 *(-11.07)*
Cooperation with colleagues index	0.097 *(8.15)*	0.073 *(5.73)*	0.092 *(5.74)*	0.113 *(8.05)*	0.050 *(3.73)*	0.082 *(4.94)*	0.113 *(9.7)*	0.061 *(3.18)*
Teacher's self-efficacy	0.062 *(7.03)*	0.100 *(6.74)*	0.139 *(10.94)*	0.095 *(7.56)*	0.110 *(5.8)*	0.085 *(5.71)*	0.101 *(8.09)*	0.112 *(7.99)*
Teacher's age	0.015 *(3.88)*	0.011 *(2.8)*	0.002 *(0.44)*	0.009 *(3.17)*	0.018 *(3.32)*	0.005 *(1.12)*	-0.003 *(-1.03)*	0.005 *(1.56)*
Experience at current school	0.009 *(2.37)*	-0.002 *(-0.79)*	0.016 *(5.45)*	0.002 *(0.8)*	0.004 *(0.7)*	0.018 *(3.7)*	0.001 *(0.11)*	0.008 *(1.85)*
Total teaching experience	-0.001 *(-0.3)*	-0.009 *(-2.3)*	-0.015 *(-3.6)*	-0.010 *(-3.19)*	-0.006 *(-0.96)*	-0.009 *(-1.56)*	-0.001 *(-0.13)*	-0.011 *(-2.01)*
Face-to-face teaching (hours)	-0.007 *(-2.77)*	-0.002 *(-0.77)*	-0.004 *(-1.32)*	-0.004 *(-1.04)*	0.003 *(0.97)*	-0.002 *(-0.76)*	-0.003 *(-0.9)*	-0.003 *(-0.99)*
Planning (hours)	-0.001 *(-0.2)*	-0.006 *(-1.26)*	-0.005 *(-1.15)*	-0.003 *(-0.71)*	-0.010 *(-1.69)*	-0.006 *(-1.22)*	-0.002 *(-0.45)*	-0.007 *(-1.75)*
Team work (hours)	-0.003 *(-0.39)*	0.002 *(0.16)*	0.000 *(-0.02)*	0.016 *(1.66)*	0.016 *(1.69)*	0.003 *(0.37)*	0.002 *(0.31)*	-0.021 *(-1.71)*
Marking (hours)	-0.011 *(-3.68)*	0.000 *(-0.04)*	0.000 *(0.11)*	-0.020 *(-3.46)*	-0.005 *(-0.75)*	-0.005 *(-0.9)*	-0.024 *(-4.79)*	-0.020 *(-3.47)*
Student counselling (hours)	0.014 *(1.51)*	-0.018 *(-1.74)*	-0.001 *(-0.07)*	0.002 *(0.26)*	-0.016 *(-2.02)*	-0.006 *(-1.21)*	0.024 *(2.49)*	-0.016 *(-1.5)*
School management (hours)	0.011 *(1.91)*	0.017 *(2.9)*	0.022 *(4.17)*	0.013 *(1.36)*	-0.006 *(-0.77)*	-0.001 *(-0.11)*	0.012 *(2.46)*	0.029 *(2.65)*
Admin duties (hours)	-0.007 *(-1.77)*	-0.002 *(-0.25)*	-0.032 *(-3.74)*	-0.022 *(-4.44)*	-0.002 *(-0.3)*	-0.008 *(-0.8)*	-0.024 *(-3.87)*	-0.035 *(-3.94)*
Interaction with parents (hours)	-0.026 *(-1.81)*	0.019 *(0.98)*	0.032 *(1.82)*	-0.027 *(-2.26)*	-0.009 *(-0.89)*	-0.005 *(-0.31)*	0.018 *(1.03)*	0.009 *(0.44)*
Extra-curricular activities (hours)	-0.004 *(-0.61)*	0.004 *(0.43)*	0.012 *(1.98)*	0.038 *(2.79)*	0.000 *(0.04)*	0.005 *(0.85)*	0.003 *(0.55)*	-0.001 *(-0.15)*
Other (hours)	-0.012 *(-2.59)*	-0.001 *(-0.12)*	-0.013 *(-2.4)*	-0.016 *(-2.3)*	0.000 *(-0.03)*	-0.022 *(-2.66)*	-0.008 *(-1.46)*	-0.010 *(-1.47)*
Female	0.004 *(0.12)*	0.076 *(1.42)*	0.113 *(3.36)*	0.160 *(4.26)*	-0.095 *(-1.71)*	0.005 *(0.1)*	0.007 *(0.17)*	0.057 *(1.3)*
Full-time employment	0.119 *(1.62)*	-0.161 *(-2.22)*	-0.038 *(-0.56)*	0.076 *(1.6)*	0.098 *(0.65)*	-0.110 *(-1.39)*	0.107 *(1.38)*	0.168 *(2.86)*
Is mentee	-0.009 *(-0.23)*	0.015 *(0.23)*	-0.064 *(-0.89)*	0.243 *(3.00)*	-0.041 *(-0.71)*	0.119 *(2.03)*	-0.027 *(-0.6)*	0.088 *(1.29)*
Is mentor	0.308 *(7.55)*	0.041 *(0.55)*	-0.134 *(-1.24)*	0.386 *(3.55)*	0.026 *(0.54)*	0.173 *(2.3)*	0.160 *(2.77)*	0.178 *(2.44)*
Permanent contract	-0.274 *(-5.11)*	-0.178 *(-3.22)*	-0.131 *(-2.5)*	-0.243 *(-3.92)*	-0.174 *(-2.83)*	-0.341 *(-4.18)*	-0.072 *(-0.87)*	-0.285 *(-4.55)*

Table A3.8: Chapter 3 – Model 3 (different activities)

Variable	Australia	Brazil	Bulgaria	Chile	Croatia	Czech Republic	Denmark	Estonia
Constant	-2.502 *(-4.94)*	-2.807 *(-12.52)*	-3.549 *(-7.71)*	-2.638 *(-7.64)*	-3.972 *(-12.9)*	-2.685 *(-8.68)*	-3.506 *(-6.94)*	-2.785 *(-7.99)*
Distributed Leadership Index	-0.050 *(-1.56)*	-0.038 *(-4.17)*	0.019 *(1.31)*	0.005 *(0.33)*	-0.005 *(-0.36)*	0.019 *(1.15)*	-0.031 *(-1.36)*	0.018 *(1.12)*
Instructional Leadership Index	-0.006 *(-0.34)*	0.028 *(2.78)*	0.019 *(0.71)*	-0.051 *(-2.39)*	-0.002 *(-0.11)*	0.000 *(-0.03)*	0.022 *(0.97)*	-0.026 *(-1.47)*
Teacher Administrative Support Personnel Ratio	0.013 *(1.13)*	0.006 *(0.97)*	-0.044 *(-0.68)*	-0.003 *(-0.19)*	-0.002 *(-0.55)*	0.005 *(0.4)*	0.010 *(0.69)*	0.003 *(0.38)*
Teacher cooperates with colleagues	0.078 *(5.08)*	0.062 *(8.31)*	0.026 *(1.86)*	0.077 *(7.39)*	0.106 *(7.79)*	0.083 *(6.39)*	0.069 *(3.93)*	0.093 *(5.65)*
Teacher Pedagogical Support Personnel Ratio	0.000 *(0.13)*	-0.004 *(-2.66)*	0.005 *(2.37)*	0.004 *(0.67)*	0.003 *(0.87)*	-0.004 *(-3.36)*	0.000 *(-0.08)*	-0.006 *(-1.23)*
Teacher's self-efficacy	0.097 *(6.03)*	0.112 *(10.58)*	0.133 *(8.76)*	0.091 *(5.82)*	0.155 *(12.32)*	0.121 *(7.5)*	0.151 *(6.27)*	0.083 *(6.09)*
Teacher's age	0.003 *(0.9)*	0.003 *(1.47)*	-0.004 *(-1.12)*	0.012 *(3.38)*	0.006 *(2.45)*	-0.001 *(-0.19)*	0.006 *(1.23)*	0.004 *(1.47)*
Experience at current school	0.008 *(1.39)*	0.000 *(0.02)*	-0.001 *(-0.52)*	-0.001 *(-0.34)*	-0.001 *(-0.29)*	0.002 *(0.84)*	0.014 *(3.32)*	-0.002 *(-0.75)*
Total teaching experience	-0.008 *(-1.92)*	0.002 *(0.87)*	0.003 *(1.02)*	-0.003 *(-0.83)*	-0.001 *(-0.56)*	-0.006 *(-1.62)*	-0.017 *(-3.05)*	0.000 *(0.16)*
Face-to-face teaching (hours)	-0.008 *(-2.46)*	-0.002 *(-2.37)*	-0.002 *(-0.6)*	-0.003 *(-1.24)*	0.000 *(-0.13)*	-0.016 *(-3.88)*	-0.021 *(-4.31)*	-0.010 *(-3.62)*
Planning (hours)	-0.008 *(-1.58)*	0.003 *(0.82)*	0.000 *(-0.03)*	0.009 *(2.4)*	0.001 *(0.31)*	-0.002 *(-0.37)*	-0.002 *(-0.36)*	-0.002 *(-0.56)*
Team work (hours)	0.002 *(0.23)*	0.002 *(0.58)*	-0.001 *(-0.06)*	0.015 *(1.17)*	0.005 *(0.56)*	-0.027 *(-1.94)*	-0.017 *(-1.33)*	0.001 *(0.12)*
Marking (hours)	-0.011 *(-2.05)*	-0.001 *(-0.22)*	-0.016 *(-2.27)*	-0.003 *(-0.36)*	-0.016 *(-2.71)*	-0.010 *(-1.48)*	-0.010 *(-1.24)*	-0.024 *(-3.42)*
Student counselling (hours)	-0.002 *(-0.21)*	-0.003 *(-1.04)*	-0.012 *(-0.8)*	-0.013 *(-1.16)*	-0.031 *(-2.64)*	0.010 *(0.98)*	-0.012 *(-0.9)*	-0.001 *(-0.09)*
School management (hours)	0.008 *(1.56)*	0.008 *(2.45)*	0.008 *(0.97)*	0.013 *(1.39)*	0.007 *(0.33)*	0.013 *(1.99)*	-0.007 *(-1.18)*	0.006 *(0.76)*
Admin duties (hours)	0.000 *(0.05)*	0.001 *(0.28)*	-0.007 *(-0.85)*	-0.001 *(-0.11)*	0.004 *(0.75)*	0.000 *(-0.04)*	-0.033 *(-3.46)*	-0.016 *(-1.44)*
Interaction with parents (hours)	-0.017 *(-0.86)*	0.002 *(0.32)*	0.010 *(0.81)*	-0.013 *(-0.65)*	-0.001 *(-0.09)*	-0.003 *(-0.16)*	-0.003 *(-0.27)*	-0.005 *(-0.27)*
Extra-curricular activities (hours)	0.010 *(1.51)*	0.003 *(0.71)*	0.008 *(1.02)*	0.021 *(2.37)*	0.006 *(0.92)*	0.011 *(1.39)*	0.012 *(1.22)*	0.018 *(2.4)*
Other (hours)	-0.012 *(-1.62)*	-0.005 *(-1.28)*	0.000 *(-0.05)*	-0.019 *(-2.91)*	0.002 *(0.33)*	-0.013 *(-2.06)*	-0.004 *(-0.61)*	0.005 *(0.61)*
Female	0.063 *(0.97)*	0.003 *(0.1)*	-0.038 *(-0.75)*	0.141 *(2.88)*	0.195 *(5.36)*	0.107 *(2.17)*	0.013 *(0.22)*	0.056 *(1.03)*
Full-time employment	0.010 *(0.1)*	0.043 *(1.44)*	-0.072 *(-0.83)*	0.062 *(0.95)*	-0.031 *(-0.54)*	-0.035 *(-0.58)*	0.116 *(1.26)*	0.130 *(2.72)*

Head teacher is female	-0.128 *(-1.93)*	-0.034 *(-0.76)*	-0.096 *(-1.52)*	-0.028 *(-0.42)*	-0.064 *(-1.5)*	-0.084 *(-1.72)*	-0.051 *(-0.61)*	-0.022 *(-0.43)*
Is mentee	-0.084 *(-1.38)*	-0.016 *(-0.25)*	0.012 *(0.17)*	0.047 *(0.4)*	0.020 *(0.42)*	0.117 *(1.71)*	0.035 *(0.48)*	0.027 *(0.42)*
Is mentor	0.346 *(4.56)*	0.066 *(2.15)*	0.256 *(4.13)*	-0.061 *(-0.55)*	0.239 *(3.2)*	0.185 *(2.21)*	0.075 *(0.47)*	-0.086 *(-0.8)*
Permanent contract	-0.073 *(-1.00)*	-0.074 *(-1.96)*	-0.121 *(-1.67)*	-0.033 *(-0.53)*	-0.061 *(-0.94)*	-0.061 *(-0.94)*	0.039 *(0.25)*	0.004 *(0.08)*

Table A3.8 continued

Variable	Finland	France	Iceland	Israel	Italy	Japan	Korea	Latvia
Constant	-3.167 *(-7.96)*	-3.125 *(-8.81)*	-2.334 *(-4.79)*	-3.294 *(-10.08)*	-2.758 *(-6.88)*	-2.241 *(-8.27)*	-2.876 *(-7.57)*	-3.273 *(-6.7)*
Distributed Leadership Index	0.011 *(0.43)*	-0.030 *(-2.82)*	0.012 *(0.59)*	-0.018 *(-0.66)*	0.028 *(0.93)*	0.012 *(1.01)*	0.026 *(1.6)*	0.005 *(0.22)*
Instructional Leadership Index	-0.007 *(-0.4)*	0.019 *(1.27)*	-0.003 *(-0.19)*	-0.016 *(-0.85)*	-0.016 *(-1.08)*	-0.033 *(-2.00)*	0.004 *(0.24)*	-0.011 *(-0.52)*
Teacher Administrative Support Personnel Ratio	0.006 *(0.97)*	0.001 *(0.07)*	0.009 *(0.68)*	0.013 *(0.88)*	-0.022 *(-3.18)*	-0.006 *(-0.81)*	-0.005 *(-0.31)*	0.019 *(1.45)*
Teacher cooperates with colleagues	0.064 *(4.38)*	0.093 *(5.16)*	0.102 *(5.36)*	0.083 *(5.32)*	0.081 *(6.74)*	0.069 *(5.03)*	0.019 *(1.57)*	0.079 *(4.34)*
Teacher Pedagogical Support Personnel Ratio	0.001 *(0.14)*	0.002 *(0.47)*	-0.026 *(-3.23)*	-0.004 *(-0.67)*	0.000 *(-0.54)*	-0.003 *(-1.15)*	0.000 *(0.14)*	0.016 *(2.17)*
Teacher's self-efficacy	0.133 *(11.15)*	0.123 *(7.85)*	0.051 *(2.66)*	0.116 *(9.5)*	0.127 *(8.9)*	0.124 *(9.28)*	0.090 *(9.65)*	0.109 *(4.78)*
Teacher's age	0.000 *(-0.13)*	0.001 *(0.17)*	0.005 *(1.08)*	0.022 *(6.83)*	-0.007 *(-1.89)*	0.006 *(1.47)*	-0.004 *(-0.78)*	0.008 *(1.94)*
Experience at current school	0.002 *(0.64)*	0.010 *(2.78)*	-0.005 *(-0.91)*	-0.005 *(-1.25)*	0.010 *(3.32)*	0.005 *(1.01)*	-0.004 *(-1.25)*	-0.004 *(-1.55)*
Total teaching experience	-0.002 *(-0.4)*	-0.009 *(-1.58)*	-0.006 *(-1.05)*	-0.011 *(-2.73)*	-0.001 *(-0.31)*	-0.011 *(-2.68)*	-0.002 *(-0.39)*	0.001 *(0.29)*
Face-to-face teaching (hours)	0.002 *(0.75)*	0.006 *(1.07)*	-0.004 *(-0.76)*	0.000 *(-0.05)*	-0.003 *(-0.56)*	-0.005 *(-1.44)*	-0.006 *(-2.33)*	-0.003 *(-1.14)*
Planning (hours)	-0.009 *(-1.64)*	-0.005 *(-1.29)*	-0.008 *(-1.5)*	0.003 *(0.5)*	0.000 *(0.01)*	0.000 *(-0.16)*	-0.004 *(-0.98)*	-0.001 *(-0.22)*
Team work (hours)	-0.007 *(-0.46)*	0.010 *(0.81)*	-0.035 *(-2.24)*	0.005 *(0.48)*	-0.005 *(-0.74)*	0.009 *(1.99)*	0.027 *(2.47)*	0.016 *(1.15)*
Marking (hours)	0.004 *(0.74)*	-0.019 *(-4.01)*	-0.002 *(-0.23)*	-0.014 *(-3.12)*	-0.015 *(-2.47)*	0.001 *(0.27)*	-0.009 *(-1.46)*	-0.022 *(-2.62)*
Student counselling (hours)	-0.010 *(-0.65)*	-0.002 *(-0.11)*	0.015 *(1.06)*	0.020 *(2.18)*	0.004 *(0.48)*	-0.013 *(-2.06)*	-0.002 *(-0.28)*	0.006 *(0.75)*
School management (hours)	0.024 *(1.79)*	0.053 *(4.01)*	0.000 *(-0.05)*	0.013 *(3.56)*	0.018 *(2.79)*	-0.002 *(-0.45)*	-0.008 *(-1.01)*	0.007 *(0.63)*
Admin duties (hours)	-0.021 *(-1.49)*	-0.029 *(-2.1)*	0.001 *(0.06)*	-0.009 *(-0.84)*	-0.008 *(-0.75)*	-0.003 *(-0.86)*	-0.011 *(-2.49)*	-0.005 *(-0.75)*
Interaction with parents (hours)	-0.041 *(-1.63)*	-0.001 *(-0.05)*	0.019 *(0.79)*	-0.017 *(-1.09)*	-0.028 *(-1.85)*	0.002 *(0.21)*	0.003 *(0.34)*	-0.025 *(-1.41)*

Extra-curricular activities (hours)	-0.002 *(-0.14)*	0.000 *(0.05)*	-0.007 *(-0.41)*	0.000 *(-0.14)*	0.016 *(1.92)*	0.004 *(1.49)*	0.013 *(1.51)*	0.007 *(0.52)*
Other (hours)	0.006 *(0.71)*	-0.014 *(-1.65)*	0.000 *(-0.05)*	-0.013 *(-2.3)*	-0.001 *(-0.21)*	0.002 *(0.6)*	-0.007 *(-1.18)*	-0.002 *(-0.23)*
Female	0.062 *(1.34)*	0.111 *(2.8)*	0.274 *(3.99)*	-0.037 *(-0.67)*	0.020 *(0.46)*	-0.053 *(-1.49)*	-0.041 *(-0.96)*	0.143 *(1.85)*
Full-time employment	0.132 *(1.43)*	-0.010 *(-0.15)*	0.062 *(0.67)*	0.044 *(0.67)*	-0.067 *(-0.85)*	-0.471 *(-4.47)*	0.227 *(1.38)*	-0.020 *(-0.31)*
Head teacher is female	0.041 *(0.8)*	-0.103 *(-2.28)*	-0.157 *(-2.16)*	0.103 *(1.61)*	-0.031 *(-0.66)*	0.016 *(0.26)*	-0.053 *(-0.84)*	-0.030 *(-0.45)*
Is mentee	0.032 *(0.29)*	0.027 *(0.28)*	0.121 *(1.44)*	0.043 *(1.02)*	0.041 *(0.55)*	-0.010 *(-0.19)*	0.110 *(2.42)*	0.072 *(0.85)*
Is mentor	0.057 *(0.53)*	0.171 *(1.31)*	0.354 *(2.39)*	-0.050 *(-1.03)*	0.021 *(0.2)*	0.203 *(5.28)*	0.215 *(4.72)*	0.042 *(0.4)*
Permanent contract	-0.175 *(-3.13)*	-0.269 *(-2.91)*	-0.114 *(-1.28)*	-0.081 *(-1.49)*	0.002 *(0.03)*	-0.114 *(-2.17)*	-0.162 *(-2.51)*	0.042 *(0.51)*

Table A3.8 continued

Variable	Malaysia	Mexico	Netherlands	Norway	Poland	Portugal	Romania	Serbia
Constant	-3.692 *(-6.56)*	-3.719 *(-13.24)*	-3.435 *(-5.78)*	-4.254 *(-9.04)*	-2.986 *(-9.16)*	-3.261 *(-8.85)*	-4.317 *(-12.45)*	-3.785 *(-12.95)*
Distributed Leadership Index	0.018 *(1.1)*	0.017 *(1.87)*	0.026 *(1.49)*	0.074 *(2.2)*	-0.003 *(-0.21)*	-0.003 *(-0.18)*	0.025 *(2.17)*	0.000 *(-0.01)*
Instructional Leadership Index	0.008 *(0.47)*	-0.014 *(-1.16)*	-0.007 *(-0.36)*	-0.015 *(-0.83)*	-0.019 *(-1.17)*	-0.026 *(-1.94)*	-0.008 *(-0.59)*	-0.004 *(-0.23)*
Teacher Administrative Support Personnel Ratio	-0.016 *(-1.61)*	-0.001 *(-0.28)*	0.015 *(1.1)*	0.003 *(0.25)*	0.004 *(0.65)*	-0.010 *(-1.05)*	0.000 *(-0.00)*	0.002 *(0.37)*
Teacher cooperates with colleagues	0.055 *(5.12)*	0.071 *(7.58)*	0.095 *(4.11)*	0.113 *(6.72)*	0.048 *(2.73)*	0.066 *(4.91)*	0.057 *(4.25)*	0.093 *(9.89)*
Teacher Pedagogical Support Personnel Ratio	0.000 *(-0.57)*	0.003 *(1.98)*	0.000 *(0.02)*	0.003 *(0.69)*	0.007 *(3.16)*	-0.001 *(-0.69)*	0.001 *(0.99)*	0.001 *(1.18)*
Teacher's self-efficacy	0.166 *(13.13)*	0.134 *(9.63)*	0.183 *(9.74)*	0.126 *(6.19)*	0.133 *(8.93)*	0.129 *(9.25)*	0.138 *(7.04)*	0.151 *(12.64)*
Teacher's age	0.003 *(0.72)*	0.001 *(0.25)*	-0.004 *(-0.85)*	0.009 *(3.14)*	0.003 *(0.7)*	0.010 *(2.62)*	-0.007 *(-1.33)*	0.001 *(0.44)*
Experience at current school	0.006 *(1.63)*	0.008 *(2.58)*	0.007 *(1.64)*	0.006 *(1.49)*	0.003 *(0.69)*	0.016 *(3.98)*	0.006 *(2.07)*	0.003 *(1.15)*
Total teaching experience	-0.003 *(-0.69)*	0.001 *(0.24)*	-0.009 *(-1.55)*	-0.011 *(-2.43)*	-0.004 *(-0.73)*	-0.014 *(-3.16)*	0.009 *(1.55)*	0.004 *(1.22)*
Face-to-face teaching (hours)	0.000 *(-0.15)*	-0.002 *(-1.04)*	-0.006 *(-0.98)*	-0.016 *(-3.05)*	0.000 *(0.12)*	-0.004 *(-1.22)*	-0.005 *(-1.4)*	-0.004 *(-1.87)*
Planning (hours)	-0.006 *(-1.37)*	0.002 *(0.42)*	0.005 *(0.76)*	-0.004 *(-0.76)*	0.001 *(0.21)*	-0.001 *(-0.31)*	-0.004 *(-1.04)*	0.005 *(1.21)*
Team work (hours)	0.000 *(-0.05)*	0.000 *(0.03)*	-0.020 *(-1.35)*	-0.008 *(-0.47)*	-0.014 *(-1.08)*	-0.016 *(-2.55)*	0.001 *(0.06)*	-0.001 *(-0.08)*

Marking (hours)	-0.005 *(-1.09)*	0.004 *(1.08)*	-0.013 *(-1.74)*	-0.008 *(-1.93)*	-0.012 *(-2.15)*	-0.003 *(-1.02)*	-0.012 *(-1.29)*	-0.012 *(-2.12)*
Student counselling (hours)	-0.003 *(-0.36)*	-0.009 *(-1.69)*	0.008 *(0.49)*	-0.014 *(-1.35)*	0.025 *(2.76)*	0.020 *(2.26)*	0.001 *(0.11)*	0.006 *(0.73)*
School management (hours)	0.004 *(0.72)*	-0.003 *(-0.41)*	0.000 *(-0.04)*	-0.003 *(-0.49)*	0.010 *(1.52)*	0.007 *(2.03)*	0.007 *(0.8)*	0.001 *(0.12)*
Admin duties (hours)	0.000 *(0.04)*	-0.007 *(-1.01)*	-0.027 *(-1.46)*	-0.013 *(-1.93)*	-0.023 *(-2.15)*	-0.012 *(-1.75)*	0.002 *(0.2)*	-0.013 *(-1.22)*
Interaction with parents (hours)	0.011 *(1.3)*	0.008 *(1.25)*	0.009 *(0.38)*	0.002 *(0.17)*	0.011 *(0.63)*	0.020 *(1.99)*	0.003 *(0.19)*	-0.008 *(-0.59)*
Extra-curricular activities (hours)	0.001 *(0.11)*	-0.003 *(-0.65)*	-0.026 *(-1.83)*	-0.009 *(-0.76)*	0.003 *(0.42)*	0.011 *(1.98)*	0.009 *(1.35)*	-0.006 *(-0.96)*
Other (hours)	-0.005 *(-0.8)*	0.004 *(0.88)*	-0.004 *(-0.59)*	0.007 *(0.75)*	0.006 *(0.89)*	-0.002 *(-0.33)*	-0.009 *(-1.23)*	-0.006 *(-0.96)*
Female	-0.025 *(-0.53)*	-0.016 *(-0.41)*	0.017 *(0.29)*	0.128 *(2.4)*	0.114 *(2.71)*	0.043 *(1.01)*	0.132 *(2.52)*	0.090 *(2.87)*
Full-time employment	0.137 *(0.93)*	-0.006 *(-0.11)*	0.154 *(2.3)*	0.064 *(0.85)*	-0.030 *(-0.48)*	-0.112 *(-1.27)*	-0.075 *(-1.05)*	0.073 *(1.54)*
Head teacher is female	-0.141 *(-2.64)*	0.021 *(0.42)*	-0.049 *(-0.61)*	-0.070 *(-0.88)*	-0.093 *(-2.02)*	-0.034 *(-0.66)*	0.088 *(1.62)*	-0.016 *(-0.37)*
Is mentee	0.062 *(1.39)*	0.033 *(0.53)*	0.081 *(1.02)*	0.038 *(0.55)*	0.151 *(2.95)*	0.054 *(0.73)*	-0.017 *(-0.23)*	-0.023 *(-0.5)*
Is mentor	0.106 *(2.94)*	0.026 *(0.48)*	-0.064 *(-0.97)*	0.305 *(2.09)*	0.124 *(2.07)*	0.143 *(1.71)*	0.121 *(1.7)*	0.134 *(1.79)*
Permanent contract	-0.511 *(-1.18)*	0.010 *(0.19)*	-0.243 *(-2.77)*	-0.246 *(-2.85)*	-0.234 *(-3.39)*	-0.036 *(-0.66)*	-0.085 *(-1.79)*	-0.356 *(-7.64)*

Table A3.8 continued

Variable	Singapore	Slovak Republic	Spain	Sweden	Abu Dhabi (UAE)	Alberta (CAN)	England (UK)	Flanders (BEL)
Constant	-3.452 *(-10.27)*	-2.964 *(-7.81)*	-3.197 *(-10.52)*	-3.169 *(-8.38)*	-4.854 *(-8.34)*	-2.570 *(-7.67)*	-3.454 *(-9.8)*	-2.618 *(-7.96)*
Distributed Leadership Index	0.035 *(2.57)*	0.015 *(0.87)*	-0.013 *(-1.4)*	-0.002 *(-0.12)*	0.012 *(0.64)*	-0.007 *(-0.56)*	0.031 *(2.54)*	0.007 *(0.49)*
Instructional Leadership Index	0.007 *(0.56)*	-0.023 *(-1.39)*	0.010 *(0.98)*	-0.013 *(-0.93)*	0.032 *(1.36)*	0.005 *(0.34)*	-0.014 *(-1.00)*	0.005 *(0.29)*
Teacher Administrative Support Personnel Ratio	-0.001 *(-0.05)*	0.018 *(1.53)*	0.007 *(0.69)*	0.007 *(1.26)*	-0.020 *(-1.41)*	-0.018 *(-1.25)*	0.018 *(0.88)*	-0.003 *(-0.56)*
Teacher cooperates with colleagues	0.098 *(8.34)*	0.074 *(5.8)*	0.090 *(5.59)*	0.113 *(8.03)*	0.052 *(4.06)*	0.089 *(5.26)*	0.113 *(9.67)*	0.062 *(3.22)*
Teacher Pedagogical Support Personnel Ratio	0.004 *(2.94)*	-0.001 *(-0.62)*	0.000 *(-0.27)*	-0.002 *(-0.38)*	-0.003 *(-0.89)*	-0.004 *(-0.43)*	0.005 *(0.52)*	-0.001 *(-0.49)*
Teacher's self-efficacy	0.063 *(7.15)*	0.101 *(6.86)*	0.139 *(11.08)*	0.097 *(7.71)*	0.113 *(6.23)*	0.086 *(5.81)*	0.102 *(8.03)*	0.111 *(7.9)*

Teacher's age	0.015 *(3.92)*	0.011 *(2.8)*	0.002 *(0.42)*	0.009 *(3.15)*	0.018 *(3.36)*	0.005 *(1.16)*	-0.004 *(-1.05)*	0.005 *(1.6)*
Experience at current school	0.010 *(2.55)*	-0.003 *(-1.01)*	0.015 *(5.35)*	0.002 *(0.75)*	0.002 *(0.29)*	0.017 *(3.29)*	0.001 *(0.33)*	0.008 *(1.73)*
Total teaching experience	-0.002 *(-0.37)*	-0.009 *(-2.24)*	-0.015 *(-3.41)*	-0.010 *(-3.22)*	-0.006 *(-0.91)*	-0.008 *(-1.48)*	-0.001 *(-0.25)*	-0.011 *(-1.95)*
Face-to-face teaching (hours)	-0.007 *(-2.83)*	-0.002 *(-0.67)*	-0.005 *(-1.51)*	-0.004 *(-1.04)*	0.003 *(0.98)*	-0.002 *(-0.57)*	-0.002 *(-0.7)*	-0.003 *(-0.95)*
Planning (hours)	0.000 *(0.01)*	-0.005 *(-1.21)*	-0.005 *(-1.13)*	-0.003 *(-0.74)*	-0.010 *(-1.63)*	-0.006 *(-1.14)*	-0.002 *(-0.33)*	-0.007 *(-1.71)*
Team work (hours)	-0.004 *(-0.53)*	0.002 *(0.14)*	-0.001 *(-0.13)*	0.017 *(1.81)*	0.014 *(1.45)*	0.004 *(0.45)*	0.004 *(0.54)*	-0.020 *(-1.65)*
Marking (hours)	-0.012 *(-3.8)*	-0.002 *(-0.33)*	0.000 *(0.07)*	-0.020 *(-3.42)*	-0.004 *(-0.63)*	-0.005 *(-0.95)*	-0.024 *(-4.73)*	-0.020 *(-3.67)*
Student counselling (hours)	0.015 *(1.67)*	-0.016 *(-1.63)*	0.000 *(0.01)*	0.002 *(0.29)*	-0.013 *(-1.84)*	-0.006 *(-1.3)*	0.023 *(2.44)*	-0.016 *(-1.39)*
School management (hours)	0.010 *(1.75)*	0.017 *(2.99)*	0.022 *(4.11)*	0.013 *(1.45)*	-0.004 *(-0.51)*	-0.002 *(-0.3)*	0.012 *(2.57)*	0.029 *(2.66)*
Admin duties (hours)	-0.007 *(-1.77)*	-0.002 *(-0.2)*	-0.032 *(-3.68)*	-0.022 *(-4.51)*	-0.006 *(-0.7)*	-0.007 *(-0.73)*	-0.024 *(-3.89)*	-0.034 *(-3.75)*
Interaction with parents (hours)	-0.025 *(-1.7)*	0.016 *(0.84)*	0.030 *(1.72)*	-0.027 *(-2.3)*	-0.009 *(-0.9)*	-0.004 *(-0.23)*	0.018 *(1.01)*	0.011 *(0.56)*
Extra-curricular activities (hours)	-0.004 *(-0.64)*	0.004 *(0.46)*	0.012 *(2.02)*	0.038 *(2.74)*	0.001 *(0.15)*	0.006 *(0.96)*	0.002 *(0.31)*	-0.001 *(-0.13)*
Other (hours)	-0.012 *(-2.53)*	0.000 *(-0.07)*	-0.013 *(-2.34)*	-0.016 *(-2.28)*	-0.002 *(-0.32)*	-0.020 *(-2.44)*	-0.008 *(-1.45)*	-0.010 *(-1.56)*
Female	0.004 *(0.09)*	0.082 *(1.54)*	0.111 *(3.34)*	0.158 *(4.2)*	-0.089 *(-1.41)*	0.004 *(0.08)*	0.010 *(0.24)*	0.064 *(1.46)*
Full-time employment	0.114 *(1.56)*	-0.163 *(-2.22)*	-0.027 *(-0.4)*	0.074 *(1.54)*	0.077 *(0.52)*	-0.100 *(-1.27)*	0.098 *(1.27)*	0.163 *(2.8)*
Head teacher is female	-0.072 *(-1.68)*	-0.057 *(-1.16)*	-0.035 *(-0.76)*	-0.029 *(-0.5)*	-0.067 *(-0.8)*	-0.079 *(-1.2)*	-0.078 *(-1.39)*	-0.114 *(-2.07)*
Is mentee	-0.010 *(-0.24)*	0.022 *(0.33)*	-0.067 *(-0.93)*	0.248 *(3.04)*	-0.034 *(-0.62)*	0.115 *(1.99)*	-0.029 *(-0.65)*	0.095 *(1.42)*
Is mentor	0.308 *(7.51)*	0.040 *(0.52)*	-0.135 *(-1.22)*	0.384 *(3.49)*	0.027 *(0.57)*	0.142 *(1.91)*	0.163 *(2.8)*	0.162 *(2.19)*
Permanent contract	-0.269 *(-4.91)*	-0.179 *(-3.25)*	-0.136 *(-2.56)*	-0.252 *(-4.01)*	-0.153 *(-2.64)*	-0.337 *(-4.26)*	-0.082 *(-0.96)*	-0.302 *(-4.66)*

Table A3.9: Chapter 3 – model 4 (different activities)

Variable	Australia	Brazil	Bulgaria	Chile	Croatia	Czech Republic	Denmark	Estonia
Constant	-3.087 *(-12.65)*	-2.825 *(-20.1)*	-3.375 *(-12.56)*	-2.931 *(-10.00)*	-4.016 *(-21.62)*	-2.541 *(-11.66)*	-3.244 *(-9.22)*	-2.807 *(-12.36)*
Cooperation with colleagues index	0.083 *(6.3)*	0.062 *(8.18)*	0.041 *(3.07)*	0.079 *(6.97)*	0.101 *(7.47)*	0.082 *(5.93)*	0.048 *(2.66)*	0.086 *(6.4)*
Teacher's self-efficacy	0.085 *(5.71)*	0.098 *(9.75)*	0.129 *(8.14)*	0.077 *(4.34)*	0.162 *(12.07)*	0.121 *(7.34)*	0.136 *(6.36)*	0.079 *(6.24)*
Teacher's age	0.007 *(1.63)*	0.005 *(2.71)*	0.000 *(-0.02)*	0.010 *(2.09)*	0.006 *(2.37)*	0.001 *(0.36)*	0.006 *(1.3)*	0.004 *(1.78)*
Experience at current school	0.006 *(1.07)*	-0.001 *(-0.34)*	-0.003 *(-1.12)*	0.002 *(0.41)*	0.000 *(0.01)*	0.003 *(1.06)*	0.007 *(1.56)*	-0.002 *(-0.86)*
Total teaching experience	-0.011 *(-2.89)*	0.001 *(0.56)*	0.002 *(0.6)*	-0.004 *(-0.79)*	-0.001 *(-0.21)*	-0.007 *(-2.01)*	-0.011 *(-2.26)*	-0.001 *(-0.34)*
Face-to-face teaching (hours)	-0.007 *(-2.55)*	-0.002 *(-2.17)*	0.000 *(0.14)*	-0.002 *(-0.67)*	-0.001 *(-0.4)*	-0.013 *(-3.32)*	-0.022 *(-4.37)*	-0.009 *(-3.51)*
Planning (hours)	-0.004 *(-0.71)*	0.000 *(0.00)*	0.007 *(1.7)*	0.009 *(2.22)*	0.000 *(0.16)*	-0.002 *(-0.43)*	0.004 *(0.69)*	-0.001 *(-0.27)*
Team work (hours)	0.001 *(0.18)*	0.006 *(1.72)*	0.008 *(0.77)*	0.011 *(0.79)*	0.002 *(0.17)*	-0.023 *(-1.6)*	0.000 *(-0.03)*	0.003 *(0.3)*
Marking (hours)	-0.010 *(-1.99)*	-0.003 *(-0.96)*	-0.021 *(-3.23)*	-0.005 *(-0.61)*	-0.012 *(-1.91)*	-0.013 *(-1.88)*	-0.007 *(-0.94)*	-0.019 *(-2.92)*
Student counselling (hours)	-0.005 *(-0.6)*	-0.004 *(-1.14)*	-0.017 *(-1.33)*	-0.008 *(-0.78)*	-0.023 *(-2.06)*	0.011 *(1.08)*	-0.010 *(-0.9)*	0.004 *(0.44)*
School management (hours)	0.009 *(1.91)*	0.007 *(2.14)*	0.009 *(1.17)*	0.009 *(1.05)*	0.001 *(0.03)*	0.015 *(2.3)*	0.003 *(0.31)*	0.007 *(1.02)*
Admin duties (hours)	-0.002 *(-0.31)*	0.001 *(0.24)*	-0.010 *(-1.12)*	0.005 *(0.44)*	0.001 *(0.11)*	0.002 *(0.33)*	-0.022 *(-1.92)*	-0.013 *(-1.52)*
Interaction with parents (hours)	-0.025 *(-1.35)*	0.008 *(1.19)*	0.007 *(0.56)*	-0.015 *(-0.88)*	0.001 *(0.08)*	0.005 *(0.28)*	-0.006 *(-0.62)*	0.012 *(0.66)*
Extra-curricular activities (hours)	0.005 *(0.65)*	0.000 *(-0.09)*	0.007 *(0.96)*	0.019 *(1.99)*	0.009 *(1.47)*	0.013 *(2.00)*	0.004 *(0.49)*	0.015 *(2.1)*
Other (hours)	-0.009 *(-1.14)*	-0.002 *(-0.65)*	0.000 *(-0.04)*	-0.015 *(-1.99)*	-0.001 *(-0.16)*	-0.009 *(-1.34)*	-0.003 *(-0.42)*	0.003 *(0.41)*
Female	0.033 *(0.45)*	0.013 *(0.46)*	-0.038 *(-0.75)*	0.116 *(2.14)*	0.179 *(4.85)*	0.146 *(3.03)*	0.059 *(1.17)*	0.063 *(1.25)*
Full-time employment	-0.068 *(-0.63)*	0.033 *(1.23)*	-0.063 *(-0.69)*	0.110 *(1.63)*	-0.066 *(-1.12)*	-0.021 *(-0.38)*	0.166 *(1.94)*	0.104 *(2.32)*
Is mentee	-0.027 *(-0.37)*	0.001 *(0.02)*	0.061 *(0.81)*	0.115 *(0.98)*	0.010 *(0.2)*	0.131 *(2.00)*	-0.027 *(-0.37)*	0.062 *(0.95)*
Is mentor	0.312 *(4.56)*	0.067 *(2.13)*	0.202 *(2.92)*	0.047 *(0.42)*	0.261 *(3.48)*	0.142 *(1.53)*	0.029 *(0.19)*	-0.030 *(-0.27)*
Permanent contract	0.045 *(0.47)*	-0.112 *(-2.99)*	-0.196 *(-2.66)*	-0.135 *(-1.79)*	-0.037 *(-0.57)*	-0.129 *(-2.09)*	0.106 *(0.8)*	-0.041 *(-0.74)*

Table A3.9 continued

Variable	Finland	France	Iceland	Israel	Italy	Japan	Korea	Latvia
Constant	-2.862 *(-13.22)*	-3.438 *(-10.55)*	-2.568 *(-7.5)*	-3.570 *(-14.95)*	-2.841 *(-8.88)*	-2.071 *(-7.43)*	-2.426 *(-6.08)*	-2.853 *(-9.17)*
Cooperation with colleagues index	0.055 *(3.99)*	0.107 *(5.8)*	0.091 *(4.52)*	0.069 *(3.38)*	0.077 *(5.39)*	0.073 *(5.23)*	0.029 *(2.23)*	0.081 *(4.52)*
Teacher's self-efficacy	0.132 *(12.64)*	0.128 *(8.11)*	0.062 *(3.32)*	0.123 *(9.66)*	0.132 *(8.87)*	0.112 *(8.23)*	0.094 *(9.26)*	0.091 *(4.19)*
Teacher's age	0.001 *(0.23)*	0.000 *(0.03)*	0.008 *(1.62)*	0.014 *(3.67)*	-0.011 *(-2.72)*	0.005 *(1.35)*	-0.005 *(-0.92)*	0.005 *(1.26)*
Experience at current school	0.001 *(0.36)*	0.009 *(2.21)*	-0.004 *(-0.81)*	-0.003 *(-0.83)*	0.006 *(1.94)*	-0.002 *(-0.53)*	-0.014 *(-2.93)*	-0.001 *(-0.41)*
Total teaching experience	-0.004 *(-0.97)*	-0.009 *(-1.63)*	-0.007 *(-1.37)*	-0.005 *(-1.15)*	0.004 *(0.92)*	-0.010 *(-2.5)*	0.000 *(-0.04)*	0.001 *(0.21)*
Face-to-face teaching (hours)	0.005 *(1.63)*	0.013 *(2.5)*	-0.005 *(-0.86)*	0.002 *(0.64)*	-0.004 *(-0.64)*	-0.003 *(-0.81)*	-0.005 *(-1.7)*	-0.004 *(-1.76)*
Planning (hours)	-0.013 *(-2.32)*	-0.006 *(-1.63)*	-0.009 *(-1.68)*	0.002 *(0.38)*	0.000 *(0.01)*	-0.002 *(-0.53)*	-0.005 *(-1.27)*	0.000 *(-0.09)*
Team work (hours)	0.005 *(0.33)*	0.014 *(1.09)*	-0.017 *(-1.12)*	0.015 *(1.43)*	0.000 *(-0.05)*	0.008 *(1.73)*	0.024 *(1.87)*	0.003 *(0.3)*
Marking (hours)	0.003 *(0.45)*	-0.019 *(-4.01)*	-0.005 *(-0.56)*	-0.014 *(-2.89)*	-0.014 *(-2.21)*	0.003 *(0.62)*	-0.008 *(-1.13)*	-0.017 *(-2.09)*
Student counselling (hours)	-0.008 *(-0.51)*	-0.002 *(-0.1)*	0.005 *(0.35)*	0.002 *(0.18)*	0.002 *(0.31)*	-0.008 *(-1.31)*	-0.003 *(-0.51)*	0.007 *(0.87)*
School management (hours)	0.015 *(1.15)*	0.048 *(3.14)*	-0.001 *(-0.2)*	0.011 *(2.53)*	0.019 *(2.28)*	-0.002 *(-0.53)*	-0.010 *(-0.99)*	0.008 *(0.76)*
Admin duties (hours)	-0.016 *(-1.02)*	-0.025 *(-1.76)*	0.004 *(0.3)*	-0.001 *(-0.05)*	0.005 *(0.44)*	-0.002 *(-0.83)*	-0.011 *(-2.05)*	-0.004 *(-0.57)*
Interaction with parents (hours)	-0.024 *(-0.9)*	0.008 *(0.44)*	0.022 *(1.12)*	0.004 *(0.22)*	-0.044 *(-2.64)*	0.010 *(0.82)*	0.009 *(0.89)*	-0.026 *(-1.73)*
Extra-curricular activities (hours)	-0.001 *(-0.06)*	-0.001 *(-0.14)*	-0.011 *(-0.68)*	-0.001 *(-0.18)*	0.009 *(0.91)*	0.005 *(1.83)*	0.005 *(0.42)*	0.021 *(1.89)*
Other (hours)	0.000 *(0.00)*	-0.015 *(-1.89)*	-0.007 *(-0.72)*	-0.006 *(-1.26)*	-0.006 *(-0.74)*	0.001 *(0.24)*	-0.012 *(-1.87)*	0.004 *(0.43)*
Female	0.086 *(2.18)*	0.166 *(3.67)*	0.187 *(2.89)*	-0.046 *(-0.79)*	-0.006 *(-0.13)*	-0.052 *(-1.41)*	-0.007 *(-0.15)*	0.129 *(1.74)*
Full-time employment	0.027 *(0.3)*	0.007 *(0.11)*	0.048 *(0.49)*	0.022 *(0.33)*	-0.046 *(-0.57)*	-0.596 *(-5.85)*	0.280 *(1.03)*	-0.042 *(-0.67)*
Is mentee	0.120 *(1.28)*	0.056 *(0.61)*	0.035 *(0.39)*	-0.004 *(-0.08)*	-0.010 *(-0.12)*	0.049 *(0.97)*	0.110 *(2.18)*	0.020 *(0.24)*
Is mentor	0.086 *(0.82)*	0.203 *(1.7)*	0.373 *(2.62)*	-0.019 *(-0.35)*	0.127 *(1.00)*	0.146 *(3.54)*	0.182 *(3.54)*	0.179 *(1.82)*
Permanent contract	-0.155 *(-3.1)*	-0.234 *(-2.67)*	-0.066 *(-0.68)*	-0.106 *(-1.82)*	-0.025 *(-0.37)*	-0.126 *(-2.39)*	-0.193 *(-2.62)*	0.010 *(0.12)*

Table A3.9 continued

Variable	Malaysia	Mexico	Netherlands	Norway	Poland	Portugal	Romania	Serbia
Constant	-3.429 *(-6.29)*	-3.473 *(-14.41)*	-2.690 *(-5.99)*	-3.187 *(-11.77)*	-3.093 *(-13.07)*	-3.703 *(-14.95)*	-4.023 *(-13.06)*	-3.850 *(-20.9)*
Cooperation with colleagues index	0.056 *(4.78)*	0.057 *(6.12)*	0.069 *(2.41)*	0.089 *(5.32)*	0.053 *(3.02)*	0.066 *(5.73)*	0.044 *(3.34)*	0.096 *(9.17)*
Teacher's self-efficacy	0.153 *(10.96)*	0.131 *(9.2)*	0.196 *(7.5)*	0.118 *(5.28)*	0.123 *(7.81)*	0.126 *(8.35)*	0.144 *(7.67)*	0.150 *(11.76)*
Teacher's age	0.000 *(-0.05)*	-0.001 *(-0.18)*	-0.003 *(-0.63)*	0.013 *(4.17)*	0.004 *(0.77)*	0.010 *(2.16)*	-0.005 *(-1.21)*	0.002 *(0.75)*
Experience at current school	0.007 *(1.73)*	0.006 *(1.67)*	0.006 *(1.43)*	0.003 *(0.81)*	0.003 *(0.71)*	0.012 *(2.96)*	0.005 *(1.71)*	0.000 *(0.03)*
Total teaching experience	-0.001 *(-0.17)*	-0.002 *(-0.55)*	-0.011 *(-2.4)*	-0.013 *(-2.98)*	-0.005 *(-0.81)*	-0.010 *(-2.36)*	0.008 *(1.55)*	0.005 *(1.65)*
Face-to-face teaching (hours)	0.000 *(0.18)*	-0.001 *(-0.61)*	-0.005 *(-0.92)*	-0.020 *(-3.58)*	-0.002 *(-0.33)*	-0.006 *(-1.96)*	-0.005 *(-1.5)*	-0.001 *(-0.7)*
Planning (hours)	-0.004 *(-0.89)*	-0.003 *(-0.74)*	-0.002 *(-0.29)*	-0.005 *(-1.03)*	-0.002 *(-0.24)*	-0.004 *(-0.99)*	-0.003 *(-0.7)*	0.001 *(0.31)*
Team work (hours)	0.006 *(0.68)*	0.006 *(0.68)*	-0.026 *(-1.76)*	-0.005 *(-0.37)*	-0.017 *(-1.41)*	-0.010 *(-1.55)*	0.000 *(-0.00)*	0.002 *(0.33)*
Marking (hours)	-0.005 *(-1.01)*	0.004 *(1.05)*	-0.021 *(-2.96)*	-0.011 *(-2.49)*	-0.008 *(-1.48)*	-0.005 *(-2.00)*	-0.017 *(-1.84)*	-0.008 *(-1.27)*
Student counselling (hours)	-0.006 *(-0.58)*	-0.005 *(-1.09)*	0.016 *(1.2)*	-0.015 *(-1.32)*	0.019 *(2.18)*	0.018 *(2.33)*	0.011 *(1.05)*	0.008 *(1.03)*
School management (hours)	0.006 *(1.05)*	0.001 *(0.21)*	-0.003 *(-0.38)*	-0.003 *(-0.54)*	0.013 *(1.92)*	0.009 *(2.3)*	0.010 *(1.29)*	0.010 *(0.94)*
Admin duties (hours)	-0.001 *(-0.15)*	-0.008 *(-1.24)*	-0.013 *(-0.65)*	-0.012 *(-1.47)*	-0.028 *(-2.21)*	-0.006 *(-0.86)*	0.009 *(0.66)*	-0.013 *(-1.4)*
Interaction with parents (hours)	0.002 *(0.19)*	0.008 *(1.26)*	0.018 *(0.65)*	-0.004 *(-0.27)*	0.042 *(1.99)*	0.018 *(2.00)*	0.004 *(0.21)*	-0.009 *(-0.59)*
Extra-curricular activities (hours)	-0.002 *(-0.43)*	-0.002 *(-0.53)*	-0.023 *(-2.19)*	-0.011 *(-0.85)*	-0.001 *(-0.11)*	0.005 *(0.89)*	0.009 *(1.31)*	-0.005 *(-0.91)*
Other (hours)	-0.001 *(-0.18)*	0.007 *(1.69)*	-0.006 *(-0.88)*	0.013 *(1.39)*	0.012 *(1.91)*	-0.006 *(-1.04)*	-0.018 *(-2.51)*	-0.006 *(-0.94)*
Female	0.000 *(0.01)*	-0.008 *(-0.18)*	-0.027 *(-0.39)*	0.153 *(3.04)*	0.094 *(2.02)*	0.110 *(2.8)*	0.156 *(3.07)*	0.063 *(2.04)*
Full-time employment	0.100 *(0.7)*	0.016 *(0.32)*	0.174 *(2.18)*	0.079 *(1.21)*	0.045 *(0.67)*	-0.014 *(-0.17)*	-0.088 *(-1.15)*	0.020 *(0.44)*
Is mentee	0.072 *(1.48)*	0.005 *(0.09)*	0.000 *(0.00)*	0.041 *(0.68)*	0.129 *(2.32)*	0.019 *(0.26)*	-0.035 *(-0.54)*	-0.019 *(-0.38)*
Is mentor	0.095 *(2.44)*	0.046 *(0.78)*	-0.081 *(-1.05)*	0.269 *(1.85)*	0.129 *(2.04)*	0.173 *(2.2)*	0.110 *(1.77)*	0.142 *(2.00)*
Permanent contract	-0.336 *(-0.74)*	0.054 *(1.11)*	-0.349 *(-3.8)*	-0.211 *(-2.23)*	-0.266 *(-3.78)*	-0.041 *(-0.81)*	-0.137 *(-2.91)*	-0.379 *(-8.44)*

Table A3.9 continued

Variable	Singapore	Slovak Republic	Spain	Sweden	Abu Dhabi (UAE)	Alberta (Canada)	England (UK)	Flanders (BEL)
Constant	-2.923 *(-12.03)*	-2.927 *(-11.2)*	-3.094 *(-10.51)*	-3.160 *(-11.41)*	-4.033 *(-13.51)*	-2.599 *(-8.79)*	-3.026 *(-11.51)*	-2.475 *(-9.14)*
Cooperation with colleagues index	0.094 *(7.53)*	0.061 *(4.85)*	0.103 *(5.85)*	0.123 *(8.12)*	0.067 *(7.16)*	0.086 *(4.94)*	0.116 *(9.7)*	0.068 *(3.46)*
Teacher's self-efficacy	0.068 *(8.07)*	0.106 *(7.15)*	0.126 *(9.38)*	0.094 *(6.83)*	0.097 *(6.36)*	0.070 *(4.54)*	0.105 *(7.16)*	0.104 *(7.35)*
Teacher's age	0.014 *(3.61)*	0.012 *(3.21)*	0.002 *(0.54)*	0.009 *(3.41)*	0.012 *(2.64)*	0.002 *(0.36)*	-0.005 *(-1.21)*	0.009 *(3.04)*
Experience at current school	0.011 *(2.72)*	-0.002 *(-0.63)*	0.006 *(1.81)*	0.003 *(0.91)*	0.002 *(0.38)*	0.011 *(2.44)*	-0.002 *(-0.38)*	0.003 *(0.82)*
Total teaching experience	-0.003 *(-0.6)*	-0.009 *(-2.26)*	-0.011 *(-2.22)*	-0.011 *(-3.9)*	0.001 *(0.24)*	-0.003 *(-0.53)*	0.003 *(0.56)*	-0.012 *(-2.8)*
Face-to-face teaching (hours)	-0.006 *(-2.34)*	-0.004 *(-1.47)*	-0.009 *(-2.45)*	-0.003 *(-0.69)*	-0.001 *(-0.47)*	-0.001 *(-0.3)*	-0.002 *(-0.59)*	-0.003 *(-0.87)*
Planning (hours)	0.000 *(0.18)*	-0.005 *(-1.07)*	-0.005 *(-1.13)*	-0.004 *(-0.75)*	-0.006 *(-1.57)*	-0.008 *(-1.5)*	-0.002 *(-0.52)*	-0.007 *(-1.63)*
Team work (hours)	-0.005 *(-0.73)*	0.004 *(0.38)*	-0.001 *(-0.07)*	0.011 *(1.11)*	0.010 *(1.18)*	0.000 *(0.04)*	-0.002 *(-0.24)*	-0.010 *(-0.81)*
Marking (hours)	-0.009 *(-3.25)*	-0.009 *(-1.29)*	-0.003 *(-0.9)*	-0.025 *(-4.24)*	-0.006 *(-0.95)*	-0.002 *(-0.38)*	-0.022 *(-4.31)*	-0.022 *(-3.97)*
Student counselling (hours)	0.014 *(1.54)*	-0.017 *(-1.98)*	-0.002 *(-0.14)*	0.005 *(0.71)*	0.002 *(0.37)*	-0.004 *(-0.66)*	0.020 *(2.3)*	-0.024 *(-2.09)*
School management (hours)	0.008 *(1.24)*	0.021 *(3.31)*	0.022 *(4.04)*	0.010 *(0.92)*	0.003 *(0.4)*	0.001 *(0.27)*	0.015 *(2.67)*	0.025 *(2.34)*
Admin duties (hours)	-0.006 *(-1.53)*	-0.002 *(-0.29)*	-0.025 *(-2.93)*	-0.019 *(-3.25)*	-0.004 *(-0.68)*	-0.005 *(-0.57)*	-0.016 *(-2.51)*	-0.033 *(-3.42)*
Interaction with parents (hours)	-0.019 *(-1.22)*	0.009 *(0.5)*	0.033 *(1.93)*	-0.017 *(-1.62)*	-0.001 *(-0.18)*	-0.001 *(-0.05)*	0.000 *(-0.01)*	0.028 *(1.49)*
Extra-curricular activities (hours)	-0.006 *(-0.96)*	0.004 *(0.49)*	0.007 *(1.6)*	0.020 *(1.39)*	0.000 *(-0.01)*	0.010 *(1.39)*	-0.003 *(-0.5)*	-0.009 *(-0.97)*
Other (hours)	-0.010 *(-2.27)*	-0.002 *(-0.33)*	-0.011 *(-1.81)*	-0.017 *(-2.31)*	-0.006 *(-0.9)*	-0.019 *(-2.31)*	-0.007 *(-1.32)*	-0.011 *(-1.61)*
Female	-0.009 *(-0.21)*	0.108 *(2.15)*	0.133 *(3.36)*	0.161 *(4.15)*	0.069 *(1.02)*	0.031 *(0.54)*	0.000 *(0.00)*	0.076 *(1.61)*
Full-time employment	0.143 *(1.91)*	-0.141 *(-1.91)*	0.131 *(1.77)*	0.080 *(1.7)*	-0.056 *(-0.41)*	-0.064 *(-0.85)*	0.100 *(1.05)*	0.170 *(2.78)*
Is mentee	0.028 *(0.69)*	0.051 *(0.77)*	-0.107 *(-1.45)*	0.333 *(4.13)*	-0.002 *(-0.04)*	0.098 *(1.63)*	-0.042 *(-0.84)*	0.069 *(1.18)*
Is mentor	0.307 *(7.7)*	0.132 *(1.74)*	-0.111 *(-1.18)*	0.327 *(3.05)*	0.076 *(1.96)*	0.100 *(1.28)*	0.195 *(3.17)*	0.144 *(1.79)*
Permanent contract	-0.254 *(-4.43)*	-0.178 *(-3.14)*	-0.136 *(-2.31)*	-0.252 *(-3.85)*	-0.104 *(-2.00)*	-0.399 *(-4.59)*	-0.104 *(-1.11)*	-0.280 *(-3.93)*

Table A3.10: Chapter 3 – Model 3b (England)

Variable	Estimate
Total workload	-0.005 *(-3.01)*
Full-time employment	0.153 *(2.12)*
Permanent contract	-0.101 *(-1.24)*
Female	-0.023 *(-0.55)*
Is mentor	0.118 *(1.81)*
Is mentee	0.008 *(0.18)*
Teacher's self-efficacy	0.114 *(11.3)*
Cooperation with colleagues index	0.117 *(9.92)*
Head teacher is female	-0.070 *(-1.19)*
Instructional Leadership Index	-0.014 *(-1.08)*
Distributed Leadership Index	0.025 *(1.94)*
Teacher pedagogical support staff ratio	-0.000 *(-0.05)*
Head teacher's age	-0.003 *(-0.64)*
Teacher administrative staff ratio	0.001 *(0.07)*
School size	0.000 *(1.59)*
Proportion of pupils receiving free school meals	-0.002 *(-0.66)*
Average Key Stage 2 score	0.027 *(0.83)*
Proportion of pupils receiving 5 A*-C grades in their GCSE exam	0.003 *(1.34)*
Independent school	0.193 *(2.11)*
Maintained School	0.022 *(0.36)*
Constant	-3.399 *(-3.77)*

Table A3.11: Chapter 3 – Model 4b (England)

Variable	Estimate
Total workload	-0.001 *(-0.69)*
Full-time employment	0.078 *(1.09)*
Permanent contract	0.035 *(0.54)*
Female	0.00 *(0.00)*
Is mentor	0.108 *(1.80)*
Is mentee	-0.014 *(-0.29)*
Teacher's self-efficacy	0.078 *(6.98)*
Cooperation with colleagues index	0.041 *(3.54)*
Well-behaved pupils	0.383 *(6.56)*
Effective school management	0.306 *(3.67)*
School managers give clear vision and direction	0.155 *(2.31)*
Insufficient autonomy	-0.300 *(-5.46)*
Teachers are underpaid	-0.042 *(-1.13)*
Own pay is fair given performance	0.073 *(1.83)*
Unmanageable workload	-0.323 *(-6.98)*
No additional pressure from accountability system	0.092 *(1.33)*
Supportive parents	0.201 *(2.91)*
No additional workload from accountability system	0.059 *(1.05)*
Scope to progress as a teacher	0.279 *(4.41)*
Scope to progress to leadership role	0.080 *(1.63)*
Scope to progress to higher pay.	0.10 *(1.84)*
Constant	-2.324 *(-14.33)*