Selection and Allocation Functions of Teacher Education Systems

Development and Test of a Model based on Open Systems Theory

vorgelegt von

Christoph König, Dipl.-Päd.

Regensburg, 2014

Inaugural-Dissertation zur Erlangung der Doktorwürde der Philosophischen Fakultät II (Psychologie, Pädagogik und Sportwissenschaft) der Universität Regensburg

Universität Regensburg

Erstgutachterin: Prof. Dr. Regina H. Mulder, Universität Regensburg

Zweitgutachter: Prof. Dr. W.H. Adriaan Hofman, Rijksuniversiteit Groningen

Abstract

Starting from several conceptual and methodological shortcomings of current research on the relation of teacher education and student achievement (Study 1), this thesis aims at developing an alternative, organizational perspective on teacher education based on Open Systems Theory. The resulting model of teacher education as an open system focuses on the selection and sorting of student teachers, as well as on the allocation of trained teachers to schools in the education system. This focus allows addressing the connection between student teachers, organizational features of teacher education, and the context of teacher education. At the same time it allows investigating the inherent selection and non-random allocation problem of research on the relation of teacher education and student achievement. The model includes characterizations of teacher education's selection and allocation functions as arrangement of structural elements governing the selection and allocation processes. Their characterization and the model are validated further by means of an interview study with experts in the fields of teacher education, education systems, and comparative education (Studies 2 and 3). The model is tested by means of two international comparative studies implementing a multigroup structural equation modeling approach. In case of the selection function, its structural arrangements in the teacher education systems of Singapore, Poland, and the USA are compared with regard to their impact on the motivational orientation of student teachers and their relation to their use of learning opportunities (Study 4). While the results show no differences in the relation between the motivational orientation of student teachers and their use of learning opportunities, there are differences in the motivational orientation of and the use of learning opportunities by student teachers across the three structural arrangements. In case of the allocation function, its structural arrangement in the teacher education systems of Singapore and Finland are compared with regard to associated differences in the degree of positive matching (Study 5). The results of this study show differences in the degree of positive matching across the two structural arrangements. Despite some methodological limitations, which are mainly due to characteristics and availability of adequate data, the results of both international comparisons allow deriving several policy recommendations. Eventually, these recommendations and the potential use of the model in further research are discussed.

Table of contents

List of tables

List of figures

CHAPTER 1

INTRODUCTION

1.1 Teachers and teacher education in the current global educational policy field

In a recent editorial of a special issue published by the *Comparative Education Review*, Paine and Zeichner (2012) state that teacher education has become one of the focal themes of debates in the global educational policy field. The Teacher Education and Development Study in Mathematics (TEDS-M), conducted by the International Association for the Evaluation of Educational Achievement, as well as the OECD Teaching and Learning International Survey (TALIS) aim at providing national and international policy makers with data and information for teacher education reforms. Additionally, international reports from other institutions compare teacher education and education systems in high performing countries and draw conclusions and recommendations for global reforms of teacher education (Barber & Mourshed, 2007). Similar to the situation when the first PISA cycles were conducted, it may be argued that with the rise of international comparative studies such as TALIS the OECD initiates a 'comparative turn' in policy and practice of teacher education, both on national and international levels (Grek, 2009; Paine & Zeichner, 2012). Grek, Lawn, Lingard and Varjo (2009) identify the shaping of policy through constant comparison of achievement data as the standard of the development and evaluation of education systems.

This standard of policy making becomes questionable when the current state of research on teachers and teacher education is considered. The global discourse is centered on the claim that teachers are crucial for student learning and achievement; differences in the quality of education systems, conceptualized as standardized mean and variance of student achievement, are attributed to differences in the quality of the teacher body (Staiger & Rockoff, 2010; Hanushek, 2011; Paine & Zeichner, 2012). In case of teacher education there is less consensus and evidence. Besides a small mean and considerable variance in teacher education effects on student

achievement, studies are unsuccessful at identifying the contribution of specific aspects of teacher education to these effects (Hattie, 2009; Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009). Moreover, it is unclear what causes differences in teacher education effectiveness with respect to the development of relevant knowledge, attitudes, and beliefs of student teachers (Boyd et al., 2009). The current state of research might not be considered a sufficient basis for informed discussions about national and international teacher education reforms.

1.2 The current state of research on teacher education

Zeichner (2005) goes beyond the aforementioned statement and claims that "[…] teacher education research has had very little influence on policymaking and on practice in teacher education programs […]" (p. 756). But what does the current state of research on teacher education look like? Zeichner (2006) distinguishes five overarching categories of contemporary research. The first category contains studies which describe current situations, practices, or contextual conditions of teacher education programs in different countries without relating these aspects to relevant outcomes. The second category involves conceptual or philosophical questions, and compares different approaches to teacher education. However, with its emphasis on societal factors and overarching questions related to the nature of teaching and the teaching profession, the focus might be too broad for policy relevant insights. The third category concentrates on how the learning and instruction of prospective teachers takes place, and focuses on the development of knowledge, attitudes, and beliefs of prospective teachers. Studies in this category also relate structural characteristics, such as field experiences, to this development. The fourth category includes studies on student teachers and teacher educators, linking their characteristics to teacher quality and student achievement in the education system. The fifth category contains studies investigating different teacher education programs and the policies in which they are embedded. This category includes questions about the effectiveness of different approaches to initial teacher training, for example, different pathways into teaching, as well as the impacts of various policies on the practice in these approaches.

It may be reasonable to assume that a combination of the last three categories of research on teacher education is most informative for a discourse about teacher education policy and practice. However, the connection between characteristics, learning, and subsequent teaching of prospective teachers with teacher education and its contextual conditions is not well established (Zeichner, 2005; Zeichner, 2006). For example, it remains unclear how structural characteristics of teacher education are related to the learning and development of prospective teachers. This is, on the one hand, due to a lack of clarity with which the respective characteristics are specified across studies (Zeichner & Conklin, 2008). On the other hand, respective studies are primarily comparisons of graduates from different teacher education programs with regard to outcomes such as student achievement (Zeichner, 2006). These comparisons focus primarily on graduates from either four-year or five-year programs, or traditional and alternative programs (Zeichner & Conklin, 2005). However, due to a lack of detail in the descriptions of the different programs it is not possible to determine which structural features are responsible for the differential

effectiveness of the programs. Moreover, the individual characteristics prospective teachers bring into teacher education are important determinants of their learning (Andrew, 1990; Feiman-Nemser & Remillard, 1996; Boyd et al., 2009). So far research has not succeeded in distinguishing the effect of the program from the effect of individual characteristics on the learning and instruction of prospective teachers, as well as in identifying characteristics which are predictive for study success (Zeichner, 2005). A further problem is that studies do not consider the context of teacher education. However, the influence of contextual conditions present in, for example, the teacher labor market or the institutional context, is important for the effectiveness of different single organizational characteristics and the teacher education program as a whole (Grossman & McDonald, 2008). In order to reach a connection between characteristics, learning, and subsequent teaching of prospective teachers with teacher education and its contextual conditions, Zeichner (2005, p. 743) suggests focusing on the following individual, organizational, and contextual characteristics:

> "[…] Among these are the individual attributes brought by prospective teachers to their teacher education programs; the specific features of these programs and their components and the institutions in which they are situated; the nature of instruction in teacher education programs, what prospective teachers learn in these programs; the schools in which teachers teach before, during, and after they complete their preparation; school district policies and practices; and state and federal policies […]."

Most policy decisions are based on studies investigating the effectiveness of graduates of different teacher education programs with regard to their impact on student achievement. The effectiveness of these graduates is then attributed to the effectiveness of teacher education. Although the meaningfulness of this outcome is questioned (Zeichner, 2005), this kind of research receives much attention. Buddin and Zamarro (2009) identify three different phases. The first, using cross-sectional, school-average student test scores and teacher characteristics found no effect of teacher education on student achievement (Hanushek, 1986). No prior achievement was included in these studies. The second introduced measures of prior achievement and other student background controls. Here, especially teacher test scores were found to be related to student achievement (Greenwald, Hedges, & Laine, 1996). In the current third phase, studies use panel data including controls for student, teacher, and sometimes school heterogeneity. These studies thus have a more detailed set of variables capturing relevant background factors influencing student achievement; in order to reduce bias due to unobserved student and teacher characteristics, the models include a variety of fixed effects. While these value-added models are a sophisticated methodological approach to investigate the relation between teacher education and student achievement, it does not allow drawing substantive conclusions with respect to the effectiveness or impact of teacher education aspects.

Estimates of the effect of teacher education on student achievement might be biased due to inappropriate theoretical conceptualizations of teacher education. As Yeh (2009) states, conceptualizations of teacher education in relevant studies focus mainly on distal aspects and are too narrow to adequately capture teacher education practice. Additionally, Harris and Sass (2011)

identify two methodological challenges which bias respective estimates: the ‘inherent selection problem’, that is, unobserved teacher characteristics influencing the amount of coursework a teacher obtains during initial training which, in turn, determines his instructional practice; and the non-random allocation of teachers to schools in the education system, where highly qualified teachers are systematically clustered in schools with students with higher socioeconomic status (Little & Bartlett, 2010). Research on selection effects is scarce; additionally, research does not provide explanations for the development of positive matching (Schalock, Schalock, & Ayres, 2006; Luschei & Carnoy, 2010). Furthermore, value-added models mostly utilize individual level data. Thus, besides knowledge, attitudes, and beliefs, even structural features of teacher education are considered individual background variables and then related to student achievement (Little & Bartlett, 2010). It becomes clear that studies implementing value-added models do not consider all of the characteristics suggested by Zeichner (2005).

And although recently some researchers have come to the conclusion that it takes a systemic view on teacher education in order to consider and connect all of these characteristics, research on teacher education is missing such a perspective (Zeichner, 2006; Grossman & McDonald, 2008). Besides recommendations to perceive teacher education as a coherent system of interrelated parts, there are few explicit models which allow for an investigation of specific parts of teacher education and their effects on, for example, competence development (Wang, Coleman, Coley, & Phelps, 2003; Maaz, Hausen, McElvany, & Baumert, 2006; Darling-Hammond & Rothman, 2011). As a consequence research does not provide explanations for the development, and thus a better understanding of the inherent selection problem and the non-random allocation of teachers to schools in the education system. Thus, the overarching question to be answered in this thesis is: *What does an organizational perspective on teacher education look like?*

1.3 Teacher education in the concept of teacher quality

In order to illustrate the usefulness of an organizational perspective on teacher education the concept of teacher quality provides an adequate framework. Teacher quality is a multidimensional construct consisting of three components (Goe & Strickler, 2008): (1) teacher qualifications and personal characteristics, (2) teacher practices, and (3) teacher effectiveness measured by standardized student test scores. Qualifications and personal characteristics of teachers influence their instructional practice, that is, their behavior in the classroom. These practices in turn influence student achievement as an indicator of teacher effectiveness. Goe and Strickler (2008) stress the difference between teacher quality and teaching quality: the former can be considered as all attributes teachers bring into the classroom, while the latter is what they actually do in the classroom. Teacher quality involves commitment for professional development, love of children, mastery of subject-didactics, a repertoire and understanding of multiple models of teaching and knowing when to use them, the ability to collaborate with colleagues, and a capacity of reflection over practice (Hopkins & Stern, 1996, in Hopkins, 2008).

Teacher education directly affects some of these attributes, and indirectly their instructional practice. Teacher education can be defined at two different levels. First, at the individual level, teacher education is subsumed under the term teacher qualifications and includes, for example, the amount of coursework obtained during initial teacher training, subject matter knowledge, pedagogical content knowledge, degrees, and credentials of teachers (Rice, 2003; Goe & Strickler, 2008). But the questions are these: do these teacher qualifications reflect, for example, the mastery of subject-didactics and the knowledge of alternative teaching models? Moreover, is it possible to draw any conclusions about the effectiveness of teacher education with respect to the development of relevant teacher characteristics? While the aforementioned teacher education variables may be differentiated according to the degree to which they reflect teacher quality, no answers can be given regarding the second question.

Answers to the second question can be provided by the second, organizational working definition of teacher education as structured learning opportunities provided to student teachers over a given period of time (Zeichner, 2006; Grossman & McDonald, 2008). This definition additionally includes structural characteristics governing the selection and sorting of teacher education candidates and student teachers, as well as the allocation of teachers to schools in the education system. The structural characteristics of a teacher education system influence and shape the personal characteristics of teachers (Morge, Toczek, & Cakroun, 2010). Structural aspects of teacher education and individual teacher characteristics are interrelated inputs which influence the behavior of teachers in the classroom. More precisely, the structural aspects influence what teachers bring into the classroom, which in turn influences what they actually do in the classroom. Eventually, teacher effectiveness is an indirect consequence of the structural and individual characteristics of teacher education.

1.4 Aim of the thesis

The overarching aim of this thesis is to reach a better understanding of teacher education policy and practice, with a specific focus on the selection and sorting of teacher education candidates and student teachers, and the allocation of teachers to schools in the education system. Both of these processes are not only important for teacher education and its effectiveness, but also for the quality and equity of education systems. With the focus on the selection and sorting of teacher education candidates and student teachers the connection between individual characteristics and structural features of teacher education is established, while simultaneously considering contextual conditions present in the teaching profession and the education system. Investigating the allocation of teachers to schools in the education system establishes a connection between structural features of teacher education and its context. Hence, this thesis addresses important shortcomings in research on teacher education (as explicated in section 1.2), and provides researchers with a theory-based conceptualization of teacher education as a system of interrelated components which illustrates the connection between student teachers, teacher education, and the context of teacher education. In order to reach the overarching aim there are two steps to be taken.

1.4.1 The development of an organizational perspective on teacher education

The first step involves the development of an organizational perspective on teacher education. This includes modeling teacher education not as an individual teacher attribute, but as a system of interrelated parts which is not independent of its context. Because of its impact on teacher education policy, the development of the organizational perspective on teacher education will be approached firstly from the point of view of current research on the relation between teacher education and student achievement. Respective studies will be reviewed with respect to their theoretical conceptualizations of teacher education and analyzed with regard to problems associated with the dominant conceptualization. Secondly, the identified problems will be discussed in light of the teacher quality concept and related to specific shortcomings of this concept. Both aspects depict the fundamental starting point of the development of the model of teacher education as an open system. Thus, the development of the model is closely connected to current theory and research. The model itself will be based on Open Systems Theory (Katz & Kahn, 1978), because this framework allows for authentic modeling of the core characteristics of teacher education as a system. In order to further validate the original model, it will be reviewed and discussed by experts in the fields of teacher education, education systems, and comparative education in an interview study. With this step Zeichner's (2006) plea for an organizational perspective on teacher education is addressed.

1.4.2 Investigating consequences of different approaches to teacher selection and allocation

The second step relates to the provision of insights into the effects of different approaches to selecting teacher education candidates and student teachers, and to allocating teachers to schools in the education system. This will be approached by testing the respective parts of the model in order to gain insight into two policy relevant aspects of teacher education. (a) The relation between the approach to selecting teacher education candidates and student teachers and their use of learning opportunities during initial teacher training. (b) The relation between approaches to allocate teachers to schools in the education system and positive matching, a distinctive manifestation of the non-random distribution of teachers in the education system. Besides addressing two prominent challenges in research on teacher education and its relation to student achievement, testing the model provides insights into the feasibility and utility of the model from which possibilities and necessities for its further development can be derived. With this step a connection between the learning of student teachers, structural features of teacher education, and the context of teacher education is established.

1.5 Structure of the thesis

The following five chapters of this thesis build on each other and include a series of complementary theoretical and empirical studies designed and conducted in order to reach the aforementioned overarching aim.

The second chapter, 'SETTING THE SCENE: *Why we need a different perspective on teacher education and student achievement',* contains a review of 49 studies investigating the relationship between teacher education and student achievement. The review focuses on the theoretical conceptualizations of teacher education, as well as the consideration of the inherent selection and non-random allocation problem. It addresses the following research questions: *(1) how do relevant studies conceptualize teacher education? (2) To what degree do relevant studies consider the inherent selection problem? (3) To what degree do relevant studies consider the non-random allocation of teachers?* The review depicts the starting point for the development of the model of teacher education as an open system by identifying relevant gaps in recent research on teacher education and its relation to student achievement.

The third chapter, 'TEACHER EDUCATION AS AN OPEN SYSTEM: *Development of an organizational perspective on teacher education',* is the core component of the first part of the thesis. It takes up the gaps in current research identified in the second chapter and directly addresses the overarching research question of the thesis: *what does an model of teacher education, taking into account the conceptual, complexity, inherent selection, and non-random allocation problems, look like?* In the first half of this chapter the core characteristics of teacher education as an open system are presented, as well as the characteristics of its selection and allocation functions and the consequences of the model for the concept of teacher education effectiveness. In the second half of the chapter an interview study with experts in the fields of teacher education and education systems is presented. The following research questions are addressed. *(1) Is the model of teacher education as an open system a valid representation of teacher education and its relation with the general education system? (2) Are the selection and allocation functions sufficiently and reasonably characterized by their dimensions and structural elements? (3) Does the operationalization of the structural elements reflect their theoretical meaning?* The study is conducted in order to further validate the core propositions and elements of the model of teacher education as an open system.

The fourth chapter, 'TESTING THE MODEL, PART I: *Comparing selection functions of teacher education systems: towards more certainty in sorting student teachers',* takes up the resulting model from the third chapter and investigates the relation between the selection function and the use of learning opportunities by student teachers. This chapter is the first part which addresses the second requirement for reaching the overarching aim of this thesis by addressing the following research questions. *(1) What is the relation between student teacher characteristics and their use of learning opportunities? (2) Does the configuration of teacher education's selection function moderate the relation between student teacher characteristics and their use of learning opportunities? (3) Are different configurations of teacher education selection functions associated with differences in the student teacher characteristics and their use of learning opportunities?* The multigroup structural equation modeling approach taken in this study illustrates a way to identify student teacher characteristics which are predictive of their use of learning opportunities, and additionally evaluate the effectiveness of different approaches to selecting candidates and student teachers with respect to these characteristics.

The fifth chapter, ‘TESTING THE MODEL, PART II: *Teacher allocation and positive matching: on the relation between teacher education’s allocation function and the non-random allocation of teachers’*, is the second part which addresses the second requirement for reaching the overarching aim of this thesis by addressing the following research questions. *(1) What is the relation between characteristics of the teaching staff and the average socioeconomic status of schools? (2) What is the relation between the average socioeconomic status of schools and teacher shortages? (3) Are different configurations of teacher education allocation functions associated with differences in the degree of positive matching?* By answering these research questions with a multilevel multigroup path analysis in a longitudinal framework, it is possible to identify differences in the degree of positive matching at two time points across different configurations of allocation functions.

The sixth chapter, ‘GENERAL DISCUSSION AND REFLECTIONS’, provides readers with an integrative discussion of the results of each of the studies. Moreover, this chapter illustrates the methodological limitations of the different studies, as well as policy implications and directions for future research which can be derived from the results of this thesis. An integrative assessment of the value added of this thesis to current theory and research on teacher education brings the thesis to a conclusion.

CHAPTER 2

SETTING THE SCENE: *Why we need a different perspective on teacher education and student achievement*

2.1 Aim and structure of the chapter

In this chapter I set the scene for this thesis and the model of teacher education as an open system by illustrating and identifying the current state and the major shortcomings of current research on the relation between teacher education and student achievement. The starting point and background of this literature review are the problems and challenges of this kind of research mentioned in the introduction. Its objective is to shed light on the reasons for the weak and inconclusive results and to develop the empirical framework for the development of the organizational perspective on teacher education. The studies included in this review are analyzed with a focus on their conceptualizations of teacher education, and the degree to which they consider the inherent selection problem, and the non-random allocation of teachers to schools in the education system. The review seeks to answer the following research questions: *(1) how do relevant studies conceptualize teacher education? (2) To what degree do relevant studies consider the inherent selection problem? (3) To what degree do relevant studies consider the non-random allocation of teachers?*

This chapter is structured as follows. In the next subsection I describe the selection of studies, that is, the search for relevant studies and the criteria by which I included them in the review. In the following sections I describe the results of the review. The summary of the chapter in the last subsection includes a brief discussion of the results.

2.2 Selection of Studies

This review focuses firstly on the theoretical conceptualizations of teacher education in studies on its impact on student achievement. This includes the complexity with which this relation is

modeled. It focuses secondly on their consideration of the inherent selection problem and the non-random allocation of teachers to schools. It adds to existing research by clarifying Yeh's (2009) statement about teacher conceptualizations in relevant studies being too narrow. Furthermore, it adds to existing research by identifying gaps which (a) may explain the inconclusive results regarding the impact of teacher education on student achievement, and (b) set the scene for a change in perspective of research on teacher education and student achievement.

I confined the time period for the search of relevant literature to the years from 2003 to 2012. This period was chosen because the last major reviews on conceptualizations of teacher education and its impact on student achievement were published in 2003 (for example, Wayne & Youngs, 2003; Rice, 2003). Furthermore, Hattie (2009) included meta-analyses which considered studies conducted prior to 2003. Hence, a comprehensive overview of current conceptualizations of teacher education in relevant research will be illustrated. In order to obtain relevant literature I searched the SSCI, PsycINFO, and ERIC databases. I used the following key words, which are oriented on the definitions of teacher education outlined in the introduction: student achievement and teacher education, teacher training, teacher preparation, teacher characteristics, teacher quality, and teacher credentials. The initial hits of the literature search and the databases are summarized in Table 1.

Table 1. Initial hits of the literature search

Keywords		Initial Hits (in Education)		
		SSCI	PsycINFO	ERIC
Student Achievement and				
	Teacher Education	986	13	338
	Teacher Training	132	10	183
	Teacher Preparation	109	12	154
	Teacher Characteristics	403	8	259
	Teacher Quality	560	19	478
	Teacher Credentials	20	0	21

Note. SSCI = Web Of Science.

After a scanning of the abstracts and the removal of duplicates, 113 studies were considered for inclusion. The full texts of these studies were read. In order to be included studies had to meet the following criteria: (1) articles were selected which included a standardized measure of student achievement explicitly as an indicator of teacher effectiveness. Studies which did not address student achievement as such an indicator were selected if the teacher education variables were related to the respective measure of student achievement. (2) Only empirical articles were selected which reported results regarding the effect of teacher education on student achievement. Empirical studies not reporting this result, as well as theoretical articles or narrative reviews, were not included. This exclusion was based on the aim of maintaining the focus on conceptualizations within the empirical relation between teacher education and student

achievement. (3) I selected only studies which aimed at the identification of teacher characteristics relevant for student achievement gains. Studies which addressed the variance in teacher effectiveness only were not included, because no direct relation between teacher education variables and student achievement was investigated. (4) The studies needed to address formal pre-service teacher education. Studies on professional development of teachers already in the profession were not included. The rationale behind this exclusion was based on the concept of teacher quality outlined in the introduction, where initial teacher training has a greater part in shaping teacher characteristics than the professional development of teachers. (5) Only studies were included which focused primarily on mainstream public education. Reason for this criterion was the higher policy-relevance of this part of the education system, as compared to private schools.

To ensure scientific quality I only included studies published in peer-reviewed journals; furthermore, studies were included based on the quality of the description of the conceptual framework and the respective teacher education variables, and the description of the respective method. The methodological approach taken, as well as the selected variables, had to be comprehensible.

Following this strategy I obtained 49 studies investigating the relationship of teacher education and student achievement, which were included in this review. By means of a structured form, and in order to answer the research questions, information was extracted from the studies regarding their conceptualizations of teacher education (i.e. the variables used to measure teacher education). Next, information was extracted from the studies regarding the complexity of the relation between teacher education and student achievement. Furthermore, information was extracted from the studies regarding the consideration of the inherent selection and non-random allocation problem. Lastly, the aims of the studies, their research questions, method and data, as well as their main findings were extracted (these information can be found in Appendix A, which gives an overview of the 49 studies listed in alphabetic order).

2.3 Results

This section is structured as follows: after a short overview of the context of the included studies, the second subsection presents their teacher education conceptualizations. The third subsection illustrates the prevalent complexity with which the studies model the relation between teacher education and student achievement. The third and fourth subsections present the considerations of the inherent selection problem and the non-random allocation of teachers to schools.

2.3.1 Overview of contexts and aims of the studies

The majority of studies (32) were conducted in the US. Two studies were conducted each in Sweden and Mexico. One study was conducted each in the UK, France, Germany, Pakistan, Peru, Guatemala, and Australia. Eight studies had an international comparative orientation and

analyzed TIMSS 1995, TIMSS-R 1999, TIMSS 2003, PISA 2000, PIRLS 2001, and PISA 2003 data.

Furthermore, the majority of studies (40) investigated achievement of elementary students. Twelve studies investigated achievement of middle school students, and five studies of high school students. Twenty-two studies used a representative school sample. Fourteen studies focused on urban schools, six studies compared urban and rural schools, and one study focused on rural schools. Six studies had a focus on high poverty schools.

The 49 studies had three distinct aims. The majority of studies examined (1) the impact of teacher quality variables on student achievement. (2) The development and test of instructional models, where teacher education, instructional practice, and student outcome variables were related to each other was the aim of two studies. Six studies focused on (3) the distribution of teacher quality across schools and its effect on student achievement. Further information about the data and designs of and methodological approach taken by the studies is included in Appendix A.

2.3.2 Research uses narrow sets of distal individual level indicators for teacher education

The model by Goe and Strickler (2008) shows, on the one hand, that teacher quality is a multidimensional and complex concept. Teacher education, on the other hand, plays a very specific role within this concept. One might expect that respective studies use a broad and specific set of indicators which adequately capture and reflect the role teacher education plays within the complex relation of teacher quality, teaching quality, and student achievement. However, this is not the case.

Relevant studies mostly use a narrow set of distal indicators to investigate the relation of teacher education and student achievement. Such a narrow set is characterized either by the use of a single variable indicating the general or subject-specific degree of the teachers (Akyüz & Berberoglu, 2010; Kaya & Rice, 2010; Munoz & Chang, 2007; Nye, Konstantopoulos, & Hedges, 2004; Slater, Davies, & Burgess, 2012), or a single variable indicating the certification status of the teachers (Neild, Farley-Ripple, & Byrnes, 2009), or a combination of degree and certification status of the teachers (Huang & Moon, 2009; Jepsen, 2005; Palardy & Rumberger, 2008). Other distal variables are the number of years of teacher training (Aslam & Kingdon, 2011; Connor, Son, Hindman, & Morrison, 2005) or a categorical variable indicating the educational level of the teachers (for example experienced, trained novice, untrained novice; Bressoux, Kramarz, & Prost, 2009).

With respect to degrees, it is a striking feature that this variable is not used in a consistent way. While some studies distinguish between teachers who have a Bachelor's degree and others (e.g. Jepsen, 2005), other studies use a binary distinction between teachers who have a Master's degree and others (e.g. Palardy & Rumberger, 2008). Hence, the meaning of a variable indicating the degree of the teachers shifts from study to study. The same is true for the certification status of the teachers. While some studies distinguish in a more global way between teachers who are certified and those who are not, other studies include several more refined certification

possibilities (for example, elementary, special, secondary, and no certification; Neild et al., 2009). While many researchers are aware of the limited usefulness of these distal variables, they still select them because of their policy-relevance or their availability in administrative datasets (e.g. Clotfelder, Ladd, & Vigdor, 2007; Munoz & Chang, 2007). There are two reasons why the use of such narrow sets of distal variables is problematic. First, these variables capture specific features of teacher education programs only implicitly. This is aggravated, for example, by the different contents and requirements of certification exams. They differ not only across countries, but also across states. A single variable indicating the certification status of a teacher, or a variable indicating if the teacher passed the certification exam on the first try (as included in the studies by Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008; Boyd et al., 2009; Goldhaber, 2007; Rockoff, Jacob, Kane, & Staiger, 2011) cannot reflect the different contents. This is a common explanation for insignificant effects of certification status on student achievement (Neild et al., 2009). The second reason is that these variables are only weak proxies for what is happening in the classroom. They say nothing about the amount and kinds of knowledge a given teacher has obtained during his initial teacher training. Moreover, they do not distinguish between what the teachers brings into the classroom, and what he does in the classroom. Thus, the difference between teacher quality and teaching quality, as stressed by Goe and Strickler (2008), is not considered.

In order to adequately capture what teachers bring into the classroom, i.e. their knowledge and personal attributes, more proximal measures of teacher education are necessary (Croninger, Rice, Rahbun, & Nishio, 2007). The most frequently used proxies for teacher knowledge are teacher test scores, which are derived from various sources: certification examinations (Boyd et al., 2008; Rockoff et al., 2011; Buddin & Zamarro, 2009; Clotfelder et al., 2006; 2007; 2010), the PRAXIS tests (Goldhaber, 2007; Sass, Hannaway, Xu, Figlio, & Feng, 2012), the Scholastic Assessment Test (SAT; Boyd et al., 2008; Kane, Rockoff, & Staiger, 2008; Harris & Sass, 2011; Rockoff et al., 2011), teacher evaluations by government authorities or educational boards in order to determine salary increases and career advancements (Leigh, 2010; Luschei, 2012; Santibanez, 2006), the grade point average during initial teacher training (GPA; Kukla-Acevedo, 2009), and purpose-built subject-specific knowledge tests in math, science and reading (Aslam & Kingdon, 2011; Baumert et al., 2010; Carlisle, Correnti, Phelps, & Zeng, 2009; Marshall & Sorto, 2012; Metzler & Woessmann, 2012; Rockoff et al., 2011). All of the studies with teacher test scores further include some or more distal variables indicating the general, subject-specific, or advanced degree, or the general or subject-specific certification status of the teachers. Moreover, Luschei and Chugdar (2011) include a variable called ‘readiness to teach’, available in the TIMSS 2003 database, where teachers rated their preparedness to teach math and science topics. They use this variable as a proxy for knowledge necessary to teach both subjects. It can be argued that this variable more adequately reflects such knowledge than certification status or degrees. Since the variables consist of self-reported responses to 16 (math) and 19 (science) items, their use as a proxy for relevant knowledge of teachers is limited. What is interesting is that only one study (Rockoff et al., 2011) considered personal characteristics (BIG-5 personality traits). Along with a

variety of test scores derived from a variety of knowledge tests, their study uses the richest set of variables indicating the education and quality of teachers.

Generally speaking, the variety of sources of teacher test scores limits the comparability of their effects on student achievement, and eventually a definitive assessment of the magnitude of the total effect. Similar to certification status, this is partly due to the different contents and requirements of the tests (Heck, 2007). The tests might differ in the degree to which they reliably measure knowledge and skills necessary for effective teaching. Therefore, they must be linked to the curriculum in a given subject (Baumert et al., 2010; Carlisle et al., 2009; Metzler & Woessmann, 2012). The linkage between the contents of a test and the intended curriculum might be used to assess the proximity with which a given teacher test is able to capture knowledge necessary for effective teaching. For example, the SAT scores and the subject-specific GPA of teachers are indicators for the academic achievement prior to and during initial teacher training. The linkage between SAT scores, GPA, and subject-specific curricula is low. Thus, it is reasonable to assume that these test scores do not fully reflect knowledge necessary for effective teaching. The same applies for scores on certification exams, which represent knowledge after initial teacher education. The case might be different for teacher evaluations and purpose-built subject-specific knowledge tests. On the one hand, teacher evaluations which include not only knowledge tests, but also scores from classroom observations might allow a direct assessment of the relation between knowledge and effective teaching. Leigh (2010), Luschei (2012), and Santibanez (2006) include teacher evaluations, which partially consist of scores from classroom observations by supervisors. Of these studies only Luschei (2012) directly assesses their relation to student achievement. On the other hand, purpose-built subject-specific knowledge tests might be better suited in order to directly capture knowledge specifically relevant for effective teaching, and overcome the limitations of certification test scores (Rockoff et al., 2011). However, only Baumert et al. (2010), Marshall and Sorto (2012), and Rockoff et al. (2011) explicitly distinguish between subject-specific content knowledge (CK) and pedagogical content knowledge (PCK). PCK comprises knowledge about different teaching methods and the learning of students, and is hypothesized to translate into effective teaching more directly. And indeed, Baumert et al. (2010) and Marshall and Sorto (2012) find substantial positive effects on student achievement.

2.3.3 Teacher education is directly related to student achievement in most studies

Another problematic aspect of the primarily distal and narrow sets of indicators is that most studies hypothesize that these distal indicators of teacher education have a direct effect on student achievement. However, Konold, Jablonski, Nottingham, Kessler, Byrd, Imig, Berry, and McNergney (2008, p. 310) state that "[…] there is little to be learned by examining the long jump between teacher characteristics and pupil learning. […]". Goe and Strickler (2008) explicitly distinguish between teacher and teaching quality. Only fourteen studies differentiate between teacher quality and teaching quality and model their association accordingly, but six studies investigate only the relation between teaching quality variables and student achievement without any links to teacher education or teacher characteristics (Akyüz & Berberoglu, 2010; Aslam &

Kingdon, 2011; Beese & Liang, 2010; Jepsen, 2005; Kaya & Rice, 2010; Myrberg, 2007). Although their theoretical frameworks explicitly state relations between teacher education, teacher characteristics, instructional practice, and student achievement, they do not estimate these relations. Hence, they give away potential, despite discussing the relevance of investigating the full complexity of the relation (for example, Jepsen, 2005). Three studies investigate the indirect relations between teacher quality, instructional practice, and student achievement, but do not differentiate between teacher education and teacher characteristics such as knowledge necessary for teaching (Konold et al., 2008; Marshall & Sorto, 2012; Palardy & Rumberger, 2008). Eventually, five studies take into account the distinction between what brings a teacher into a classroom and what he does in the classroom. These studies investigate the full association between teacher education, teacher characteristics, instructional practice, and student achievement (Baumert et al., 2010; Carlisle et al., 2009; Connor et al., 2005; Desimone & Long, 2010; Guo, Connor, Yang, Roehrig, & Morrison, 2012). The potential which is given away by studies not modeling the full association is related to two problems.

First, it hampers the identification of teacher characteristics which are predictive for student achievement. This problem is aggravated if only distal variables are used for measuring teacher education or teacher characteristics. Some authors consider especially the lack of non-cognitive teacher characteristics as a limiting factor with respect to the amount of explained variance in student achievement (Huang & Moon, 2009). The narrow focus might explain the blurry picture regarding teacher characteristics relevant for effective teaching, and in turn for student achievement (Rockoff et al., 2011). Disregarding the relation between teacher characteristics and instructional practice further hinders a clarification of the pedagogical mechanisms with which teacher knowledge translates into effective teaching (Croninger et al., 2007; Marshall & Sorto, 2012). Especially with respect to teacher knowledge, a relatively proximal indicator of teacher quality, information about teaching practice is necessary in order to be able to identify which kinds of knowledge are relevant specifically for effective teaching. Hence, the step between teacher characteristics, be it teacher knowledge or non-cognitive attributes, and instructional practice might be a way to identify teacher characteristics which are predictive of student achievement.

Second, the assessment of the effectiveness of different kinds or amounts of teacher education is limited when the relation between teacher education and teacher characteristics is not explicitly taken into consideration. For example, the amount of knowledge relevant for teaching a teacher brings into the classroom is a direct consequence of his initial teacher training (Baumert et al., 2010; Harris & Sass, 2011). The distal variables used to measure teacher education do not allow for an investigation of this relation. They might gain meaning, however, if they are included in conceptual models which explicitly state a relation between teacher education and teacher characteristics, and investigate differences in teacher characteristics and associated differences in instructional practices (Baumert et al., 2010). When they are not, which is the case for the majority of studies, teacher education effectiveness may not be adequately assessed. The question remains if individual-level variables measuring teacher education are suited in order to capture different features of initial training programs which are relevant for teacher knowledge building

and other teacher characteristics (Gansle, Hoell, & Burns, 2012). A promising way relies on an organizational perspective on teacher education.

Several studies conceptualize the degree or certification status of the teachers as an organizational property of schools (Andersson, Johansson, & Waldenström, 2011; Akiba, LeTendre, & Scribner, 2007; Beese & Liang, 2010; Croninger et al., 2007; Fuchs & Woessmann, 2007; Heck, 2007; Hogrebe & Tate, 2010; Myrberg, 2007; Rivkin, Hanushek, & Kain, 2005; Woessmann, 2003). This approach does not relate teacher education to differences in teacher characteristics. Hence, differences in teacher education effectiveness may not be adequately assessed. Only three studies take up the aforementioned organizational approach (Boyd et al., 2009; Gansle et al., 2012; Zhao, Valcke, Desoete, Verheaghe, & Xu, 2011). Gansle et al. (2012) include a categorical variable indicating the nature of the teacher training program (undergraduate, masters, practitioner, private practitioner), while Zhao et al. (2011) use a binary variable indicating the level of the teacher education institute, distinguishing between distance education or open university (low) and normal university or teacher training college (high). However, both studies do not consider differences in teacher characteristics, but relate the teacher education variable directly to student achievement. Boyd et al. (2009) do the same, but their conceptualization includes specific structural features of teacher education programs: the number of math and language courses required for entry or exit; a capstone project (for example, portfolios, a research paper) which students have to complete in order to graduate; the percentage of tenured personnel; the oversight of student teaching, including the quality and selection of the cooperating teachers, and the frequency with which the teachers are observed by a program supervisor. Furthermore, they included variables indicating the amount of learning opportunities used by the students and the congruence of training contents and the state-specific curriculum. In sum, an organizational perspective enabling researcher to assess the effectiveness of different teacher education programs is taken up only by a fraction of studies. Moreover, these studies do not consider the relation between teacher education and teacher characteristics, but relate teacher education directly to student achievement. Due to the aforementioned reasons this is problematic and further aggravated by the inherent selection problem, which is covered in the next section.

2.3.4 Studies consider the inherent selection problem implicitly

It is reasonable to assume that what teachers bring into the classroom depends on their use of learning opportunities during their initial teacher training. The amount of coursework a teacher obtained is further assumed to depend on personal characteristics, for example, his motivation (Harris & Sass, 2011). Thus, the inherent selection problem gains relevance because of its relation to important teacher characteristics such as knowledge relevant for teaching. Variation in relevant teacher characteristics consequently may be due to variation in their behavior during initial teacher training (Winters, Dixon, & Greene, 2012). Despite this relevance, few studies take into account the use of learning opportunities. Some studies implicitly include it in other indicators of teacher education, for example the college attended (Aaronson, Barrow, & Sander, 2007; Luschei, 2012; Zhao et al., 2011), or use teacher fixed effects to control for unobserved

heterogeneity of teachers in the relation between teacher education and student achievement (Jepsen, 2005; Leigh, 2010). Thirteen studies explicitly include variables related to the inherent selection problem. Barron's college ranking, a categorical variable indicating the selectivity of the teacher education institution (from least selective to most selective), is an organizational-level variable (Boyd et al., 2008; Clotfelder et al., 2006; 2007; 2010; Rockoff et al., 2011). Eleven studies include individual-level indicators: the SAT score of the teachers, a measure of pre-college ability (Boyd et al., 2008; Harris & Sass, 2011; Kane et al., 2008; Rockoff et al., 2011), the GPA of the teachers (Baumert et al., 2010; Kukla-Acedevo, 2009), or the number of course credits or relevant college courses a teacher obtained or completed during his initial teacher training, which are more proximal indicators of the use of learning opportunities (Boyd et al., 2008; Croninger et al., 2007; Desimone & Long, 2010; Harris & Sass, 2011; Kukla-Acedevo, 2009; Phillips, 2010; Winters et al., 2012).

In sum, only a fraction of studies considers the inherent selection problem explicitly with proximal indicators of the teachers' course taking patterns during initial teacher training. Additionally, the organizational-level variables are treated as individual teacher attributes. Moreover, almost all of the aforementioned studies relate the amount of coursework directly to student achievement. Only one study specifies and tests the relation between these teacher education variables and teacher knowledge (Baumert et al., 2010). Desimone and Long (2010), although their conceptual model explicitly states this relation, do not test for an influence of teacher education on teacher knowledge. This overall picture is problematic due to several reasons.

First, treating organizational or structural features of teacher education as individual attributes allows no conclusions about specific features of teacher training programs and their effect on teacher characteristics. Although this is standard practice in the US literature (Clotfelder et al., 2007), conceptualizing the competitiveness of a teachers' college as an individual teacher attribute is a distal way to consider the inherent selection problem. If these organizational features are furthermore directly related to student achievement, the effectiveness of different teacher education programs cannot be assessed. It is not possible to investigate variation in relevant teacher characteristics due to variation in organizational features of teacher education programs. Second, the same is true for the direct relation of the other individual variables to student achievement. Despite the number of college courses or the course taking pattern of teachers closely captures the core characteristic of the inherent selection problem, it remains unknown how different course taking patterns translate into relevant teacher characteristics, for example their knowledge relevant for teaching. Hence, additional to the problems associated with relating teacher knowledge directly to student achievement, disregarding the complex chain from teacher education over teacher characteristics to teacher behavior in the classroom hinders the identification of teacher characteristics relevant for student achievement. Third, scores on academic ability tests or the grade point average have different notions. The SAT score is primarily a measure of an individual's pre-college ability, while the grade point average captures his performance during initial teacher training. Thus, both variables have different meanings with regard to the amount of coursework a teacher obtained. To adequately capture these meanings it

might be necessary to use the SAT score as a predictor of the GPA. This relation is only implicitly modeled in respective studies. Furthermore, the selectivity of a teacher training institution might influence the average SAT score of its student teachers, which in turn influences their course taking patterns and their GPA. What would be necessary to shed light on these relations is an organizational perspective on teacher education, which allows investigating the effect of specific features of teacher education, such as its selectivity, on individual attributes of student teachers (Boyd et al., 2009). Their influence on course taking patterns and, eventually, on relevant teacher knowledge would be a next necessary step. Hence, measures of teacher course taking patterns during initial teacher training implicitly describe and assume many important prior processes, without explicitly taking them into account or explaining them. Thus, their impact on teacher knowledge remains unclear. This in turn has consequences for the use of easily available, policy-relevant measures used by schools, boards, or other authorities when recruiting teachers. These consequences will be explained in the next section, in the context of the non-random allocation of teachers.

2.3.5 Studies consider the non-random allocation problem on a technical level

Consequences of the inherent selection problem with respect to teacher knowledge have an impact on aspects of the non-random allocation of teachers to schools and students. As Goldhaber (2007, p. 791) states, "[…] test performance [on certification exams] is not a silver bullet credential that can be used to predict teacher effectiveness […]". Partly due to variation in teachers' course taking patterns, certification exams are not only weak signals in the recruitment process, it is almost the only variable investigated with respect to the influence on the non-random allocation of teachers in the context of the relation between teacher education and student achievement.

The majority of studies included in this review consider the non-random allocation of teachers to students. They can be differentiated according to the level at which they include or directly investigate this problem. First, the most prevalent mechanism to account for a non-random allocation of teachers to students or schools can be found at the level of the dataset. All studies use either administrative or purpose-built datasets which match teacher registers or databases with achievement and other data of students. Moreover, studies use classroom, school, or country-level data and investigate the relation between teacher education and student achievement on these levels accordingly (Andersson et al., 2011; Akiba et al., 2007; Beese & Liang, 2010; Croninger et al., 2007; Fuchs & Woessmann, 2007; Heck, 2007; Hogrebe & Tate, 2010; Myrberg, 2007; Rivkin et al., 2005; Woessmann, 2003). Second, the next mechanism can be found at the level of the estimation of the respective models. Nineteen studies use a variety of student, grade, and school fixed effects, or a combination of these (Aaronson et al., 2007; Andersson et al., 2011; Aslam & Kingdon, 2011; Boyd et al., 2008; Boyd et al., 2009; Buddin & Zamarro, 2009; Clotfelder et al., 2006; 2007; 2010; Goldhaber, 2007; Harris & Sass, 2011; Jepsen, 2005; Kane et al., 2008; Kukla-Acevedo, 2009; Leigh, 2010; Phillips, 2010; Rockoff et al., 2011; Sass et al., 2012; Slater et al., 2012). Fixed effects, for example of students, teachers, or

schools, are means to control for unobserved heterogeneity in students, teachers, or schools. The third mechanism can be found at the level of research questions, i.e. the design. Ten studies specify and test a research question directly related the non-random allocation of teachers. Akiba et al. (2007) investigate the distribution of teacher quality in 46 countries, while Desimone and Long (2010), Heck (2007), Huang and Moon (2009), Luschei (2012), Metzler and Woessmann (2012), Nye et al. (2004), and Sass et al. (2012) look into differences in teacher quality between schools with low and high poverty or socioeconomic status or rural and urban schools. Two studies are concerned with the composition of the teacher body. Goldhaber (2007) investigates certification requirements and their effects on the composition of the teacher body. Winters et al. (2012) investigate the probability that a teacher gets into and remains in the classroom, based on his effectiveness. Furthermore, three studies apply an experimental design where teachers were randomly assigned to students (Bressoux et al., 2009; Konold et al., 2008; Nye et al., 2004).

Each approach has different consequences for the identification of the teacher education effect on student achievement, and the explanation of the development of the non-random allocation of teachers to students and schools. Without considering the non-random allocation estimates of the effect of teacher education on student achievement are biased by unobserved characteristics of students, their family background, and schools (Rivkin et al., 2005). As mentioned above, these unobserved characteristics are consequently controlled by fixed effects. However, Clotfelder et al. (2006) found that the relation between teacher qualifications and student achievement is due to a large extent to the sorting of teachers and students between and within schools, even with an estimation including student and school fixed effects. Thus, estimates of the effect of teacher education on student achievement are still biased. For reasons explained in the previous sections, this bias might be further aggravated by the use of distal teacher education variables. Despite numerous critiques (for example Rockoff et al., 2011; Metzler & Woessmann, 2012), the majority of studies rely on easily available, yet distal variables. The selection of such variables is often motivated and substantiated by their policy-relevance (Leigh, 2010), that is, they are selected because they are often used in the recruitment process and to determine the salary or career advancement of teachers. But exactly these measures, for reasons explained in the previous section, are weak signals for schools and principals (Goldhaber, 2007). Thus, since it remains questionable that these variables adequately capture what the teacher brings into, and further, what he is able to do in the classroom, researchers may not exactly determine the influence of the non-random allocation problem on the relation between teacher education and student achievement. The use of fixed effects or matched administrative databases is not suited for solving such theoretical problems.

Explanations of the development of the non-random allocation of teachers to schools are rare. Neither of the aforementioned approaches to the problem provides insights into mechanisms causing the distribution of teachers in the education system. For example, the results of almost all the studies with a research question dedicated to the non-random allocation, e.g. Heck (2007) and Luschei (2012), only allow a description of the respective distribution of teachers. They do not test for effects of structural features of the teacher labor market on this distribution. More specifically, they are ex-post adjustments to already accomplished facts, since the assignment of

teachers to students or schools already took place. In this regard, the distal teacher education variables prove to be problematic as well. A prominent example is the aforementioned certification status of teachers, a proxy for a certification system within a country or district. While only two studies investigate the composition of the teacher body (Winters et al., 2012; Goldhaber, 2007), only one relates the composition and distribution to certification measures (Goldhaber, 2007). In this study Goldhaber concludes that the effect of certification systems on the teacher distribution is not well understood. In light of these research characteristics the development of phenomena regarding teacher distributions, such as the non-random allocation of teachers and the positive matching of teachers and students, remain unclear and cannot be explained. Explanations for these phenomena require an organizational perspective on the non-random allocation problem, which allows relating relevant structural features of the teacher labor market to specific teacher distributions.

2.4 Summary of the Chapter

In this chapter I reviewed 49 studies, published between 2003 and 2012, investigating the relation between teacher education and student achievement. The primary focus was on the conceptualizations of teacher education within this relation, as well as on the considerations of the inherent selection problem and the non-random allocation of teachers. The research questions were: (1) how do relevant studies conceptualize teacher education? (2) To what degree do relevant studies consider the inherent selection problem? (3) To what degree do relevant studies consider the non-random allocation of teachers?

Based on the results of this review I have to confirm Yeh's (2009) observation and conclude that current conceptualizations of teacher education are characterized by the use of narrow sets of distal proxies for teacher education. Furthermore, these distal proxies are directly related to student achievement, without taking into account differences in teacher characteristics and their instructional practice. Thus, the difference between what a teacher brings into the classroom and what he is able to do in the classroom, which is stressed by Goe and Strickler (2008), may not be adequately addressed (Konold et al., 2008). These basic characteristics of conceptualizations of teacher education can be labeled the 'conceptual problem' and the 'complexity problem'. Both problems are related to and further aggravated by the way the studies consider the inherent selection and non-random allocation problems.

In case of the inherent selection problem studies include variables indicating the course taking pattern of teachers during their initial teacher training and relate them directly to student achievement. As explained above, these variables only implicitly assume many aspects of the use of learning opportunities of teachers without making them clear. Furthermore, organizational or structural aspects of teacher education, such as the selectivity of the teacher training institution, are treated as individual teacher attributes. The relation between such organizational or structural characteristics and the subsequent use of learning opportunities is not taken into account. Research on respective characteristics of teacher education programs is lacking (Zeichner & Conklin, 2005). Given the importance of the use of learning opportunities for teacher

characteristics such as knowledge, it cannot be fully explained *why* a given teacher brings a certain set of characteristics, such as knowledge and personality traits, into the classroom. Additionally, when studies neglect the relation between teacher characteristics and their instructional practice, means to test "[…] the proposition that teachers who have participated in teacher education are more likely to behave in ways that help pupils […]" are lacking (Konold et al., 2008, p. 309). Consequently, it hinders the identification of teacher characteristics relevant for effective teaching.

In case of the non-random allocation of teachers to schools, prominent means to account for this problem are various forms of fixed effects to control for unobserved heterogeneity in students, teachers, and schools, and administrative datasets matching teachers and students. These measures are primarily taken in order to reliably identify the effect of teacher education on student achievement. In light of the conceptual and complexity problems, however, it is questionable if such measures are sufficient, even when longitudinal data are analyzed. The lack of effects of teacher education on student achievement is frequently explained by differences in teacher education programs or teacher labor markets (Munoz & Chang, 2007). Respective variables, such as test scores on certification exams, are selected because of their relevance for policies concerning the teacher labor market. However, the significance of these variables is disputed (Rockoff et al., 2011). Consequently, some researchers argue that policies which shape the teacher body should be based on classroom observations, where the teaching quality of teachers can be assessed more directly (Goldhaber, 2007). But only few studies investigate such policies with regard to their possible influences on the teacher body and the distribution of teachers. Hence, researchers cannot explain *why* a given teacher brings his characteristics to the classroom in a specific school, and not in another. In other words, there are no explanations about specific manifestations of the non-random allocation problem, such as the positive matching of teachers and schools.

With a change in perspective on teacher education and student achievement the conceptual and complexity problems may be solved. This change involves (a) a slight alteration of the input variables in the model of teacher quality (Goe & Strickler, 2008), (b) modeling teacher education as a system of interrelated components embedded in its context, and (c) a change in the understanding of teacher education effectiveness (Morge et al., 2010). The alteration comprises an unfolding of the relation between teacher education, teacher characteristics (such as teacher knowledge and attitudes), and teacher behavior in the classroom. The model involves conceptualizing teacher education as an open system, with structural prerequisites for the selection of students, opportunities to learn for student teachers, and structural features responsible for the distribution of teachers into the education system. Teacher education effectiveness is then a question of how the teacher education system and its components account for the development of cognitive and non-cognitive characteristics of student teachers which are necessary for effective teaching. In other words, teacher education is responsible for the prerequisites for effective teacher behavior in the classroom. Based on these premises, in the next chapter I develop a model of teacher education as an open system.

CHAPTER 3

TEACHER EDUCATION AS AN OPEN SYSTEM: *Development of an organizational perspective on teacher education*[1]

3.1 Aim and structure of the chapter

In this chapter I address the problems identified in the literature review presented in the previous chapter and illustrate the development of a model of teacher education as an open system. This model takes into account that teacher education is embedded in multiple contexts, for example the higher education system and the general education system (Grossman & McDonald, 2008). Zeichner (2006) criticizes current research on the relation between teacher education and student achievement because it is lacking such an organizational perspective. Without such a perspective, even structural features of teacher education continue to be considered individual teacher characteristics, subsumed under the term teacher quality and directly related to student achievement (Wayne & Youngs, 2003; Little & Bartlett, 2010). Zeichner (2006) deems such a perspective necessary, because research cannot explain variation in effects on student achievement across teacher education programs (Boyd et al., 2009). This variation might not only be due to the provision of learning opportunities, but at the same time and for equal parts to a better selection of student teachers (Denzler & Wolter, 2009). Research on such selection effects is still scarce (Schalock et al., 2006). However, recent studies suggest a positive relation between selection procedures and competence development of student teachers (Blömeke, Felbrich, Müller, Kaiser, & Lehmann, 2008). Furthermore, variation in teacher education effects on student achievement might be biased by the non-random allocation of teachers to schools. As already mentioned, there might be teacher-school combinations which lead to better student achievement (Jackson, 2010). In light of the conceptual and complexity problems of current research on teacher education and student achievement, research cannot explain why there are manifestations of the non-random allocation of teachers to schools such as positive matching.

---

[1]Parts of this chapter and an updated version of the final model described in this chapter will be published as: König, C., & Mulder, R.H. (in press). A change in perspective – teacher education as an open system. *Frontline Learning Research*.

The overarching research question of this chapter is as follows: *what does a model of teacher education, taking into account the conceptual, complexity, inherent selection, and non-random allocation problems, look like?* The chapter is divided in two parts, which reflect the development of the model of teacher education as an open system. The first part deals with the general characteristics of the model, which is based on Open Systems Theory. This includes characteristics of its selection and allocation functions, the structural elements of both functions, and indicators of these elements. The second part deals with the validation and further refinement of this model. The basic characteristics of the model are used as input for an interview study with experts in the fields of teacher education, education systems, and comparative education. The aims of the interview study are twofold. First, I want to validate the model with its main propositions and elements. Second, based on the experts' knowledge and opinions, I want to derive decisions about changes in the model in order to make it testable and usable for further research.

3.2 Teacher education as an open system

The model of teacher education is based on Open Systems Theory (Katz & Kahn, 1978). Due to its focus on interactions between systems, and its specific consideration of the relation between the system and its participants, it adequately captures the central characteristics of teacher education systems. I conceptualize the teacher education system as initial teacher training programs for lower secondary education. As an open system it is defined as "congeries of interdependent flows and activities linking shifting coalitions of participants embedded in wider material-resource environments" (Scott & Davis, 2007, p. 34). This definition entails the transitions of student teachers into and out of the teacher education system (i.e. the interaction between systems), as well as the use of learning opportunities by student teachers (i.e. the relation between the teacher education system and its student teachers). Open Systems Theory does not consider student teachers elements of the teacher education system, but a part of its environment (Simon, 1976). This has consequences for learning opportunities and their use by student teachers, which will be described later in this chapter.

The core of the interaction between the teacher education system and the general education system, conceptualized as the lower secondary school system, is the provision of a sufficient number of qualified teachers. The success of this core function depends on characteristics of the teacher education candidates entering the teacher education system, further on the learning opportunities provided and their use by student teachers, and eventually on the assignment of the trained teachers to the schools in the general education system. For teacher education as an open system maintenance of the intake of student teachers as well as the output of trained teachers is essential (Scott & Davis, 2007). In this regard its selection and allocation functions play a key role (Musset, 2010). A function is in general defined as a contribution of one (or more) elements of a system to the effectiveness of the superior system (Parsons, 1951). More specifically, the selection function is defined as the selection and sorting of teacher education candidates and

student teachers (Van de Werfhorst & Mijs, 2010). The allocation function is defined as the assignment of newly trained teachers to schools (Parsons, 1951).

Both functions are arrangements of structural elements which establish the interaction with the general education system, and maintain the intake of student teachers and the output of trained teachers (Katz & Kahn, 1978; Wang et al., 2003). In case of the selection function these structural arrangements govern the selection and sorting of teacher education candidates and student teachers. In case of the allocation function they govern the assignment of trained teachers to schools. With the help of established routines and administrative conditions and regulations (for example admission regulations) they allow for candidates, student teachers, and trained teachers being screened out if they do not meet the requirements of teacher education, the teaching profession, or the requirements of a teaching position in a given school (Maaz et al., 2006). The specific characteristics of both functions will be described next.

3.3 Specific characteristics of the selection and allocation functions

3.3.1 The selection function

The arrangement of structural elements of the selection function governs the intake of teacher education candidates into the teacher education system. Furthermore, it maintains an adequate, or else, successful use of learning opportunities by student teachers within the teacher education system. The relevance of the selection function is based on the dependence of student teachers to successfully use the learning opportunities provided, in order to develop relevant characteristics such as knowledge, beliefs, and attitudes (Blömeke, Suhl, Kaiser, & Döhrmann, 2012). These learning opportunities require specific cognitive and personal characteristics. If these requirements are not met by teacher education candidates or student teachers, their use of learning opportunities becomes suboptimal and they may drop out of the teacher education system prior to graduation (Blömeke, 2009). Thus, the use of learning opportunities is inherently unstable (Katz & Kahn, 1978; Blömeke, Kaiser, & Lehmann, 2010). Furthermore, student teachers may not be intrinsically motivated to act according to the requirements of the learning opportunities. They join, stay, and leave depending on the relative advantage resulting out of their exchange with teacher education (Scott & Davis, 2007).

Hence, the arrangement of structural elements of the selection function includes control and socialization measures (Katz & Kahn, 1978; Scott & Davis, 2007). These measures provide information about teacher education candidates' cognitive and personal characteristics, as well as about the use of learning opportunities by student teachers. This monitoring serves as feedback for the teacher education system in order to make informed admission and progression decisions (Katz & Kahn, 1978). Socialization mechanisms initiate the transfer of professional role expectations and norms in order to integrate the student teacher into the teaching profession (Saks, Uggerslev, & Fassina, 2007). They serve as information for the student teacher and the teacher education system to assess his fitness for teaching.

Moreover, the selection function adapts to contextual conditions present in the general education system which influence the pool of potential candidates. For the teacher education system it is only possible to select candidates which (are able to) make themselves available, or else, which are able to decide to enter initial teacher training (Grodsky & Jackson, 2009). Thus, the arrangement of structural elements includes characteristics of the teaching profession, as well as of the general education system which influence the degree of freedom in educational decision making. The structural elements of the selection function, as well as their operationalization with respect to adequate indicators of these elements, will be described next.

3.3.2 Structural elements of the selection function and their operationalization

The structural elements of the selection function can be assigned to three distinct dimensions which are related to the success of the selection function in governing and maintaining the intake of, and the use of learning opportunities by student teachers: first, the accessibility of teacher education, which describes how easy it is for a potential candidate to decide for initial teacher training; second, the comprehensiveness, that is, the level of general information about candidates and student teachers. This information can include either the skills or experience needed for a successful use of learning opportunities, or the cultural fit of the candidate or student teacher to teacher education, or their long term potential for teaching (Baron, Hannan, & Burton, 1999). Third, the level of general integration of student teachers into teaching; this level is related to the degree to which the structural elements reduce the uncertainty of student teachers when they make their first teaching experiences. The elements of the selection function, their respective dimension, and their indicators are summarized in Table 2.

3.3.2.1 Elements and their indicators of the accessibility of teacher education

Rational choice models postulate that individuals analyze every educational alternative by weighing costs against benefits (Sicherman & Galor, 1990). They emphasize three core aspects relevant for accessibility: structure, finance, and status.

There are four respective structural elements. (1) The ***stratification of the education system*** describes the number of possible tracks in an education system; high stratification implies a high number of structural determined decision points which negatively influences the accessibility of teacher education. Consequently, the respective indicator is the number of school types or distinct educational programs available to 15 year old students, which varies widely from one country to another (OECD, 2010). (2) The ***stratification of tertiary education***, i.e. a high number of tertiary educational alternatives decreases the probability for a positive decision for teacher education, and decreases its accessibility (Becker & Hecken, 2009). The respective indicator is the number of school types or distinct educational programs available to 19 year olds. (3) The ***funding of teacher education*** includes either tuition fees charged from or financial aids available to students. The indicator describes the type of funding tertiary education, which is a combination of tuition fees charged and financial support available to students (OECD, 2012a). Individuals opt for

educational alternatives which secure or enhance their social status (Breen & Goldthorpe, 1997; Esser, 1999). (4) A high ***occupational status of teaching*** attracts a high number of potential candidates and implies a high accessibility. The indicator is taken from the Eurydice database, which suggests three different types of employment (Eurydice, 2012): public sector employee with contractual status, civil servant status, and career civil servant status, where teachers are appointed for life.

Table 2. The selection function: dimensions, structural elements, and respective indicators

Dimension	Structural Element	Indicator (Source)
Accessibility of Teacher Education	Stratification of the Education System	Number of school types available to 15 year olds (OECD, 2012a); continuous
	Stratification of Higher Education	Number of school types available to 19 year olds (OECD, 2012a); continuous
	Funding of Teacher Education	Type of combination of tuition fees and financial support available to students (OECD, 2012a); categorical
	Occupational Status of Teaching	Type of employment system (Eurydice, 2012); categorical
Level of General Information	Career Counseling for Teachers	Type of career counseling for teachers (own collection); categorical
	Admission Procedures Assessment Procedures	Structure of the selection process (OECD, 2012a); categorical
	Admission Criteria Assessment Criteria	Degree of central regulation of teacher education (Horn, 2009); categorical
Level of General Integration	Internal Support	Availability of guided teacher practica during initial teacher training (OECD; 2012a); binary
	Field Experiences	Minimum amount of professional training during initial teacher training (Eurydice, 2009); continuous

3.3.2.2 Elements and their indicators of the level of general information

Educational and selection decisions are characterized by an asymmetric distribution of information (Van der Velden & Wolbers, 2007). Lack of information increases the risk of maladaptation, opportunistic behavior and agency problems. Agency theories emphasize guidance and control mechanisms that increase the level of information.

There are five respective structural elements. (1) ***Career counseling for teachers*** provides career-relevant information leading to realistic perception of the requirements of a teaching career (Grodsky & Jackson, 2009). In some countries career counseling for teachers is seen as a complementary mechanism to admission procedures (Rothland & Terhart, 2011). The respective indicator describes the type of such career counseling: no counseling, non-mandatory counseling, and mandatory counseling. (2) ***Admission procedures***, and their (3) ***criteria*** which they are based

on, provide information about candidates and increase the level of information, and the likelihood of selecting candidates with a high probability of graduating. The criteria determine which individual characteristics are required for teacher education. (4) ***Assessment procedures***, and their (5) ***criteria***, which determine characteristics of student teachers necessary for teaching, are implemented to monitor individual progress and growth, either formative or summative. The indicators are combinations of admission and assessment procedures, and admission and assessment criteria. The first indicator captures the structure of the selection process and distinguishes three types: screening at entry to the profession, screening at entry into initial teacher training, and screening at both entry points. The level of information increases from type one to three. Regarding criteria, the concept of centralization is adopted, which is defined as the degree of central regulation of teacher education curricula and competences (Horn, 2009). There are three types (EU, 2009): regulation by teacher education institutes, regulation by centrally administered broad frameworks, and by central authorities. The level of information increases with a higher standardization.

3.3.2.3 Elements and their indicators of the level of general integration

Theories on organizational socialization emphasize socialization mechanisms complementary to control mechanisms that increase the level of integration. They support the teacher education system adapting to student teachers' needs in order to retain good students. Respective structural elements serve as means of information dissemination helping student teachers to take on new roles (Van Maanen & Schein, 1979).

There are two respective structural elements. Both elements reduce uncertainty of students about teaching. (1) ***Internal support*** gives access to structured forms of guidance by experienced teachers or sequenced in clearly defined courses. These structures reduce uncertainty of students about expectations and requirements when entering teacher education. Two types of teacher education systems can be differentiated: systems where such a teacher practicum is not available and systems where such a teacher practicum is required and mandatory (OECD, 2012a). (2) ***Field experiences*** describe opportunities for student teachers to make teaching experiences prior to entering the teaching profession. Combined with control mechanisms, the positively influence teacher education success rates. The indicator is continuous and describes the minimum amount of professional training during initial teacher training (Eurydice, 2009).

3.3.3 The allocation function

The arrangement of structural elements of the allocation function governs the assignment of trained teachers to schools in the general education system. Furthermore, it determines recruitment and hiring decisions of principals and trained teachers which eventually result in matches between teachers and schools (Lankford & Wyckoff, 2010). The relevance of the allocation is based on the unequal distribution of qualified teachers in education systems (Harris & Sass, 2011), and the fact that the quality of teacher-school matches is able to account for up to

40% of a teacher effect, and the effect of teacher-school matches on teacher turnover (Jackson, 2010; Maier & Youngs, 2009). The assignment of teachers to schools has long been based on the assumption of school equivalence: schools and teaching positions were assumed to be equal across districts and regions (Johnson & Kardos, 2008). Only recently research came to the conclusion that each school is a unique social system, with unique requirements with respect to the environment and students (Jackson, 2010). Another problematic aspect of the recruitment process is the reliance of principals and other authorities on weak signals of the knowledge and skills of teachers (Goldhaber, 2007). Similarly to the selection function, lack of information or an asymmetrical distribution of information may lead to maladaptation and, in case of the allocation function, to an increased teacher turnover (Van der Velden & Wolbers, 2007).

Similarly to the selection function, the arrangement of structural elements of the allocation function includes control and socialization measures (Scott & Davis, 2007). These measures provide information about the knowledge and skills of a trained teacher, as well as about his fit for a given teaching position in a school. This information serves as feedback for the schools in order to make informed recruitment decisions, and as feedback for the teacher education system regarding the adequacy of the knowledge and skills of the trained teachers. In other words, the teacher education system learns from this feedback if its teachers have acquired the necessary knowledge and skills for teaching. The socialization mechanisms initiate the transfer of school-specific role expectations and norms in order to integrate the trained teacher into the school (Saks et al., 2007). They further serve as information for the teacher and the school to assess his fitness for the specific teaching position.

The allocation function adapts to contextual conditions present in the general education system as well. These contextual conditions influence the number of available teachers (supply), and the number of available teaching positions (demand). Consequently, the structural arrangement includes characteristics of the teaching profession, as well as of the general education system which influence the decision of trained teachers to enter teaching at all, and to take a specific teaching position. The structural elements of the allocation function, as well as their operationalization with respect to adequate indicators of these elements, will be described next.

3.3.4 Structural elements of the allocation function and their operationalization

The structural elements of the allocation function can be assigned to three distinct dimensions which are related to the success of the allocation function in governing the assignment of teachers to schools: first, the capacity of the teacher labor market, which describes the number of available teachers and teaching positions or, in other words, teacher supply and demand; second, the level of specific information about trained teachers. This information can include either the skills or experience needed for the teaching position, or the cultural fit of the teacher to the school, or their long term potential for the teaching position (Baron et al., 1999). Third, the level of specific integration of trained teachers into the school; this level is related to the degree to which the structural elements reduce the uncertainty of trained teachers when they first start at a school or teaching position. The higher this level, the faster the teacher is able to act according to the

school's requirements. The elements of the allocation function, their respective dimension, and their indicators are summarized in Table 3.

Table 3. The allocation function: dimensions, structural elements, and respective indicators

Dimension	Structural Element	Indicator (Source)
Capacity of the Teacher Labor Market	Student Population	Student-teacher ratio (OECD, 2012a); continuous
	Salary Structure	Starting salary combined with the years required to top salary (Eurydice, 2012a); categorical
	Work Conditions	Type of continuing professional development (Eurydice, 2012); categorical
	Career Ladders	See Occupational Status
Level of Specific Information	Licensure/Certification	Requirements to enter the teaching profession (OECD, 2012a); categorical
	Probationary Periods	Type of recruitment process (OECD, 2012a); binary
	School Autonomy	Index of school responsibility for resource allocation (OECD, 2010; 2013); continuous
	Union Regulations	Percentage of schools where the principal reported that teacher union exert direct influence on staffing (OECD, 2010); continuous
Level of Specific Integration	Teacher Induction	Length of the induction period (Eurydice, 2012); continuous
	Teacher Mentoring	Type of regulation of teacher mentoring (Eurydice, 2012); categorical
	TE-School-Partnerships	Type of partnership between schools and teacher training institutions (EU, 2007); categorical

Note. TE = Teacher Education.

3.3.4.1 Elements and their indicators of the capacity of the teacher labor market

The capacity of the teacher labor market, i.e. the number of available teachers and teaching positions shapes the allocation process. Both aspects cover teacher supply and demand. Career mobility and segmentation theory suggest two aspects constituting the capacity of the teacher labor market (Doeringer & Piore, 1971; Sicherman & Galor, 1990): structure and incentives.

There are four respective structural elements. (1) Characteristics of the ***student population***, e.g. its size, affect the number of available teaching positions. An indicator for the student population is the student teacher ratio. It is, among others, one of the primary factors determining the size of the teacher body. It compares the number of students (in full-time equivalent) to the number of teachers (in full-time equivalent) at a given level of education (OECD, 2012a). A low student teacher ratio implies an increased capacity of the teacher labor market. The following

three elements are all related to teacher supply. (2) The ***salary structure*** influences the attractiveness of teaching, and thus the number of available teachers. High wages lower opportunity costs even for high ability teachers. Its indicator captures the starting salary combined with the number of years required to top salary (Eurydice, 2012; OECD, 2012a). A high starting salary combined with few years from start to top implies an increased capacity of the teacher labor market. (3) ***Work conditions*** can be considered nonpecuniary factors influencing the attractiveness of a teaching career. The respective indicator, type of continuing professional development, describes training activities including subject based and pedagogical training. Three types can be distinguished (Eurydice, 2012): a teacher labor market where CPD is optional; CPD is necessary for promotion; and CPD is a professional duty. The attractiveness of working conditions increases from type one to three, and thus increases the capacity of the teacher labor market. Furthermore, internal labor markets can be conceived of as incentives for teachers to apply for a given teaching position. (4) ***Career ladders*** structure the labor market, and simultaneously depict incentives for high performance and opportunities for professional development. Hence, they increase the attractiveness of a teaching career. The indicator is the same as for occupational status, since these aspects are included in the status of teaching.

3.3.4.2 Elements and their indicators of the level of specific information

An asymmetric distribution of information in the recruitment of trained teachers increases the risk of adverse selection and turnover (Van der Velden & Wolbers, 2007; Maier & Youngs, 2009). The reasons are differences in the knowledge and skills of trained teachers, even when they graduated from the same university. Signals about teachers' characteristics attenuate the lack of information (Stiglitz, 1975). Signaling theories emphasize control mechanisms and other structural features that determine the level of information.

There are four respective structural elements. Signals are provided by (1) ***licensure*** and ***certification*** of beginning teachers. The better these mechanisms are able to identify the knowledge and skills of a teacher, the higher the probability of a quick initial teacher-school match. These are one or more tests required to be passed before entering the teaching profession, thus exceed a simple graduation, and attest that a trained teacher is ready to teach (Goldhaber, 2011; OECD, 2012a). Three types of certification/licensure systems can be distinguished, where no additional certification after graduation is required, where certification is required to start teaching, and where certification is required to become a fully qualified teacher. The level of specific information increases from type one to three. Further specific information is provided by (2) ***probationary periods***. Here, tenured positions in schools are separated from probationary positions where teachers are monitored with respect to their performance and fit to the school (Staiger & Rockoff, 2010). Thus, the recruitment decision depends on the outcome of this monitoring. There are two distinct possibilities: a recruitment process where a probationary period is not required, and where a probationary period is required (OECD, 2012a). The level of specific information is higher in the second type. The degree of (3) ***school autonomy*** in recruitment determines the degree to which information is provided directly to schools. This

element includes job descriptions; high school autonomy makes job search more time consuming and requires teachers to be adequately informed (Mortensen, 1986). Job descriptions are a structural element adding to the level of specific information, thus facilitating matches between teachers and schools (Liu & Johnson, 2006). The index of school responsibility for resource allocation (REPRES) from the PISA studies is adopted as indicator (OECD, 2010; 2013). The level of specific information increases with a higher school autonomy. However, this kind of recruitment weakens in an institutional framework allowing less autonomy to choose applicants: autonomy can be overridden by (4) ***union regulations***. These regulations do not consider school specific needs regarding trained teachers and can be understood as constraints interfering with school based recruitment. The degree of union involvement is defined as percentage of students in schools where the principal reported that teacher unions exert a direct influence on staff decision making (OECD, 2010; 2013). An increasing degree of union involvement implies a lower level of specific information.

3.3.4.3 Elements and their indicators of the level of specific integration

To increase the level of specific integration the transfer of school-specific knowledge, values and norms is highlighted. First, the teacher learns the requirements of a role or teaching position (functional aspect); second, he integrates into the social structure of the school (inclusion aspect). Over time they get accustomed to the specific organizational characteristics and can adapt to them (Braxton, Hirsch, & McClendon, 2004). Theories on organizational socialization suggest respective socialization mechanisms.

There are three respective structural elements. (1) ***Teacher induction*** is a means to make the teacher acquainted to the specific characteristics at a given school. It includes a formalized system to support beginning teachers (Ingersoll & Strong, 2011). The length of this system varies from country to country (Eurydice, 2012). Longer induction periods imply a higher level of specific integration. (2) ***Teacher mentoring*** is personal guidance provided by a senior teacher at a school. It varies from single meetings to formalized programs involving frequent communications between beginning teacher and mentor. They depict means to assess the fit of teacher and school. There are three types (Eurydice, 2012): no regulations, regulations available and actual support measures in use. The level of specific integration increases from type one to three. Structural linkages between initial teacher training institutes and schools (Van der Velden & Wolbers, 2008) describe (3) ***partnerships between teacher education institutes and schools***. They not only facilitate recruitment, but also increase the level of specific integration. These partnerships can be differentiated into three types (EU, 2007): a type where schools and teacher education institutes are not connected at all (no connection), where only certain schools are connected to some teacher education institutes (partly connection), and where schools are an integral part of teacher education (full connection). The level of specific integration is highest in type three.

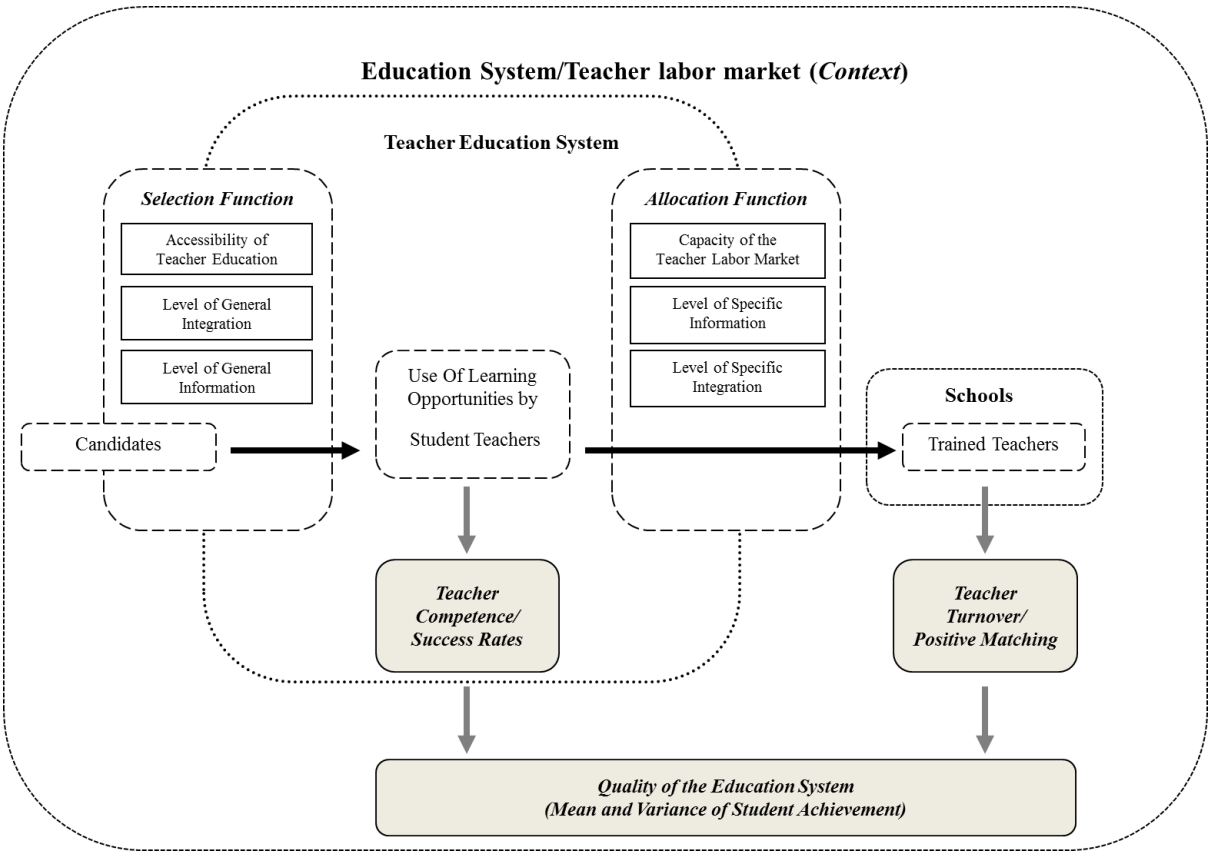

Figure 1. The original model of teacher education as an open system; rectangles depict the dimensions of the selection and allocation function, as well as contextual conditions in the education system/teacher labor market. The black arrows illustrate the transition of an individual through teacher education into schools, from teacher education candidate over student teacher to a trained teacher in a school. The gray arrows depict the consequence of the use of learning opportunities by student teachers on their competence and success rates as individual and organizational-level indicators of teacher education effectiveness. Teacher turnover and positive matching are individual and organizational manifestations of specific teacher distributions, and thus indicators of the effectiveness of the allocation function. They are both prerequisites for the indirect outcome of teacher education, namely the quality of the education system indicated by the mean and variance of student achievement.

3.4 Teacher education effectiveness in the open system framework

The effectiveness of teacher education is commonly determined by changes in student achievement. Teachers are ranked according to the actual achievement gains of their students, which are compared to their predicted achievement gains (Goe & Strickler, 2008). In other words, teacher education effectiveness is put on a level with teacher effectiveness. Apart from those two being very different concepts, this direct way of relating teacher education to student

achievement does not consider the full complexity of this relation (Konold et al., 2008). The open system framework allows for a more complex concept of teacher education effectiveness, which takes into account the inherent selection and non-random allocation problems.

The immediate outcomes of teacher education, and thus the primary criterion of teacher education effectiveness are teacher characteristics such as knowledge, attitudes, and beliefs (Blömeke et al., 2008). The development of these characteristics depends on student teachers and their use of the learning opportunities provided by the teacher education system. As already mentioned the selection function plays a prominent role with respect to the use of learning opportunities. It is assumed that the development of knowledge, attitudes and beliefs is facilitated when student teachers meet the requirements of the learning opportunities in the teacher education system (Tillema, 1994). Student teachers vary in the degree to which they possess such characteristics which make it likely that they successfully use the learning opportunities. In organizational terms the gap between student teacher characteristics and the requirements of the learning opportunities are called training costs (Glebbeek, Nieuwenhuysen, & Schakelaar, 1989). These training costs increase when student teachers are unsuccessful in their use of learning opportunities because they exceed the standard period of study, and are highest when they drop out of initial teacher training (Blömeke, Kaiser, & Lehmann, 2010). Hence, beside the mean and variability in teacher knowledge, attitudes, and beliefs as the primary indicator of teacher education effectiveness, there is also an organizational indicator. Low training costs imply a successful use of learning opportunities by student teachers which results in high organizational success rates (Gansemer-Topf & Schuh, 2006). An optimized selection function, which can be characterized by a high accessibility of teacher education, a high level of general information, and a high level of general integration, reduces variability in student teacher characteristics and further influences the success of the use of learning opportunities. This results in higher mean levels of knowledge of student teachers as well as higher success rates of the teacher education system.

This is the first integral part of the concept of teacher education effectiveness which takes into account aspects of the inherent selection problem (Harris & Sass, 2011). However, teacher education effectiveness cannot be adequately evaluated if the distribution of teachers in the school system in not taken into account. High success rates imply a high number of qualified teachers available for distribution in the school system. Thus, the allocation function depends on the success of the selection function and directly connects to these selection and sorting results. But higher numbers of qualified teachers do not automatically lead to higher student achievement. Given the unequal distribution of teachers in the school system, and the differential effectiveness of teacher-school combinations, the assignment of teachers to schools needs to be taken into account (Jackson, 2010; Little & Bartlett, 2010). The need for this consideration is illustrated by Boyd, Grossman, Hammerness, Lankford, Loeb, Ronfeldt, & Wyckoff (2012). They found that teacher education effects are diminished when teacher attrition is taken into account. Complementary to the high number of qualified teachers, which may be related to the mean student achievement in a given school system, the distribution of teachers is related to an equity-aspect of student achievement (Hofman, Hofman, & Gray, 2008). It does matter where

teachers bring their characteristics into the classroom. Positive matching, a prominent manifestation of a specific teacher distribution, restricts access to qualified teachers for low income or poverty students and negatively affects estimates of the effect of teacher education on student achievement. An optimal allocation function, which can be characterized by a high capacity of the teacher labor market, a high level of specific information, and a high level of specific integration, minimizes the association between school socioeconomic status and teacher shortages. It grants equal access to qualified teachers regardless of student background. This is the second integral part of the concept of teacher education effectiveness which takes into account the non-random allocation problem (Harris & Sass, 2011). In sum, the interaction between the teacher education system and the general education system is not finished with the provision of a sufficient number of qualified teachers, but requires an equal distribution of the teachers in the school system. Figure 1 illustrates the model of teacher education as an open system outlined in the previous sections. This version of the model was the theoretical input of the interview study with experts in the fields of teacher education, education systems, and comparative education. This second part of the model development will be described next.

3.5 An interview study with experts in the fields of teacher education/education systems

A shortened version of the outline of the model of teacher education as an open system was used as input for an interview study with experts in the fields of teacher education, education systems, and comparative education. Prior to testing the model and its propositions I wanted to make sure that the model is a valid representation of the relation between teacher education and the education system. Moreover, I wanted to make sure that the selection and allocation functions are sufficiently characterized by their dimensions and their elements, and that the elements are adequately operationalized. Consequently, the interviews had two distinct aims. First, I wanted to validate the model with respect to the general characteristics of teacher education as an open system (i.e. the overall model and its main propositions), the characteristics of the selection and allocation functions (their dimensions and elements), and the operationalization of the elements and the indicators of teacher education effectiveness. Second, based on the experts' evaluations of the validity of the model, as well as their knowledge and opinions, I wanted to identify necessary changes of the model in order to increase its feasibility and value for comparative research on teacher education. The research questions which I wanted to answer with this interview study were as follows. *(1) Is the model of teacher education as an open system a valid representation of teacher education and its relation with the general education system? (2) Are the selection and allocation functions sufficiently and reasonably characterized by their dimensions and structural elements? (3) Does the operationalization of the structural elements reflect their theoretical meaning?*

3.5.1 Sample

The sample consisted of eight experts working in eight different research institutions and universities in five different countries (Germany, Austria, Switzerland, Netherlands, and the USA). The international composition of the sample reflected the need for multiple perspectives on teacher education from different national and regional contexts (Blömeke & Paine, 2008). Their specific expertise ranged from international comparisons of education and teacher education systems over professionalism of teachers and teacher education effectiveness, to characteristics of teacher labor markets. These fields of expertise were necessary in order to be able to evaluate the general propositions of the model of teacher education as an open system (international comparisons of education and teacher education systems), the dimensions and structural elements of the selection and allocation functions (teacher education effectiveness, professionalism of teachers, and teacher labor markets) and the operationalization of the elements (international comparisons of education and teacher education systems). Formal criteria for participation were at least ten years of experience in their respective field and relevant publications in peer-reviewed journals or assistance in international comparative studies. Moreover, experts were selected based on the regional focus of their projects. The average experience of the experts in their fields of research was 18.25 years (*SD* = 8.83). One expert was a research assistant (PhD), and one expert was retired. Six experts were full professors at their respective universities. The experts conducted or assisted in international comparative studies in Western Europe, North and Latin America, Africa, and Asia.

3.5.2 Interview procedure

Prior to the interviews a handout was prepared containing the model of teacher education as an open system, the dimensions of the selection and allocation function including their respective structural elements, and the operationalization of the structural elements. This handout was sent to the experts after a date for the interview was confirmed, which was usually two weeks before the interview. Thus, the experts were given time to make themselves familiar with the model. The interviews were conducted in January and February 2013, either via Skype (4 interviews), telephone (3 interviews), or face to face at the experts' workplace (1 interview). The mean duration of the interviews was 43.25 minutes (*SD* = 19.23). They were digitally recorded and subsequently transcribed for analysis. A semi-structured guideline, which was based on the handout of the model, was used for the interviews. The interviews started with some information about the scope of the study. Each expert was asked for permission to record the interviews, and if there were any open questions regarding the model. First, the experts were asked about the validity of the relation between teacher education and the education system, if the role of the selection and allocation within this relation has been made clear, and if they saw any advantages or disadvantages with modeling teacher education as an open system. Next, experts were asked if both functions are sufficiently characterized by their respective dimensions, and if the structural elements are correctly assigned to the dimensions. Third, the experts were asked if they miss a

relevant structural element, and to rate and weigh the structural elements according to their relevance for their respective dimension. Fourth, the experts were asked if the indicators of the structural elements adequately reflect the theoretical meaning of the elements, or else, if the indicators sufficiently reflect the theoretical meaning of the elements. Next, the experts were asked if the combination of some specific structural elements into one indicator (see sections 3.3.2 and 3.3.4) is reasonable and justifiable. Lastly, the experts were asked if they wanted to add something which was not talked about during the interview. The semi-structured guideline was extended if an expert mentioned an aspect which needed further elaboration; the extended version was then used in the subsequent interviews. The statements of the experts were validated communicatively during the interviews. This means that if an expert made an explicit statement with respect to the validity of the model or to a potential change in a part of the model, his statement was repeated by the interviewer in order to make sure that it was understood correctly. After confirmation by the expert the interview was continued. The interviews were conducted until a saturation point was reached. This means that once no new aspects were raised in two consecutive interviews no new interviews were conducted. This was the case after the last interview.

3.5.3 Analysis

I used a direct approach to content analysis in which the categories for the initial coding were derived from the semi-structured guideline. More specifically, I used five categories which guided the initial analysis of the data: (1) the validity of the model, its general propositions, and the role of the selection and allocation functions; (2) advantages and disadvantages of modeling teacher education as an open system; (3) the dimensions of the selection and allocation functions; (4) the structural elements and their assignment to the dimensions; (5) the operationalization of the elements. New categories were introduced if themes were identified which could not be assigned to one of the existing categories. I used an iterative inductive strategy for categorizing the statements of the experts (Lichtman, 2013). The starting point was a single evaluative statement about either the validity of some part of the model, or directly about a potential change of the part of the model. This statement was initially coded and assigned to one of the aforementioned categories, depending on the part of the model it concerned. Thus, information about the validity of different parts of the model was grouped together. After the first coding of the interviews the interviews were read again and the initial codes were revisited and modified. Therefore I further refined the coded statements of the experts, eliminated redundant information, and checked the assignment of the coded statements to the categories again. Next, I reread the coded statements of the different categories, integrated them to an overarching statement of all experts and derived an overall evaluation of the respective part of the model. Based on these overall statements about the validity of the parts of the model, and the direct suggestions of changes to the model, I reached to decisions about which parts of the and how the model has to be changed in order to make it feasible and useful for further research.

3.5.4 Results

This section is structured according to the three research questions. In each section the overarching evaluative statements of the experts regarding the different parts of the model are presented. First, I describe the experts' evaluation of the validity of the overall model. Next, I describe the experts' evaluation of the characterization of the selection and allocation functions. Third, I describe the experts' evaluation of the operationalization of the structural elements of both functions. In every section I also describe and explain the respective decisions for changes in the model, based on the experts' evaluations.

3.5.4.1 Teacher education, student achievement, and the role of the functions

The experts took a quite similar view on the general conceptualization of teacher education as an open system and the role of the selection and allocation functions. They considered the overall interaction between teacher education as an open system and the general education system as a valid representation of the relation between the two systems, especially from an international comparative perspective. For example, Expert H mentioned that

> "[…] the advantage is to look at this as a system. At least in this area the research I have seen in the United States has not gone far in conceptualizing this. The problem with teacher sorting and teacher allocation, I mean there is a lot of good empirical work but not so much conceptual work like this. So just the fact that you are doing this conceptualization is great, but you are also thinking of it as a system. Incorporating all of the parts whereas a lot of the teacher-sorting research in the United States kind of just looks at that, you know, teacher preferences, the teacher-student match without really considering what the system is and how the system might vary across countries or across regions. So yes, I think in particular this will be useful for doing comparative research because it does include a lot of the elements that you would need to consider if you were doing comparative research […]" (p. 1, line 18).

This statement was further enhanced by Expert D who mentioned that, additionally to its usefulness in international comparative studies, the model can be expanded by the level of the individual teacher. In sum, from a practical research point of view the model offers fruitful areas for research on teacher education.

There was disagreement among the experts regarding the basic definition of teacher education being a part of the general education system, which was understood as the lower secondary school system. While Expert B suggested that teacher education is a part of the general education system in the Netherlands, Expert A stated that, by definition, the teacher education system cannot be a part of the lower secondary school system. They both mentioned that teacher education is also part of the higher education system. While the higher education system is not explicitly included in the model, the statements of the experts confirm Grossman and McDonalds (2008) statement that teacher education is at the nexus of multiple contexts.

With respect to theoretical aspects of the model two experts (Expert A and G) raised serious concerns about the scope of the functional chain from selecting teacher education candidates and student teachers over the allocation of trained teachers to student achievement. From the point of view of Expert G

> "[…] the model is in many respects an underspecified model and probably describes a relation which cannot be reasonably investigated in one single block. […] It might be better to keep the overall context in mind as a working hypothesis, and to select a part from which one is convinced to have found something theoretically new, and to look if there are respective data for that part […]" (p.1, line 39).

This is accompanied by a concern about the appropriateness of the indicator for the quality of the general education system being the mean and variance of student achievement. While the mean is considered appropriate, the variance of student achievement is not. Expert G suggested that variance in achievement is desirable above a given minimum level. Another theoretical concern was raised about the quality of the teacher education system, that is, what student teachers learn during initial teacher training. Three experts (Expert C, Expert D, and Expert F) raised the question if it is reasonable to relate structural and qualitative aspects of teacher education. However, after it was made clear that the model, in its current state, focuses on more quantitative aspects (provision of a sufficient number of qualified teachers) the experts agreed that the relation between the teacher education and the general education system is represented appropriately in the model.

With respect to the role of the functions within the relation between teacher education and the general education system the experts differed in their evaluations of specific aspects of the selection and allocation functions. As an overall evaluation of the role of both functions Expert A stated

> "[…] I have used these functions of education like selection function, allocation function, or skills production function in my own work, so I can see that you would like to also look at the allocation function, whether there is a good match between what is required in the labor market for teachers, that is, so in education, or whether the right kind of skills are being produced, and whether it relates to the selection function as well. I can see this, that you would like to address these three functions, and these have an effect on the quality of education, as you say. So if you have better teachers produced in higher education, or in teacher training, and if they are better allocated, if that works well that should also affect the quality of education. I think that is probably what I think the main contribution of your conceptual framework. That you really try to look at the role of teacher education and the different functions, so that is I think a good part […]" (p. 2, line 3).

In case of the selection functions all of the experts stated that the role has been made clear and is reasonably elaborated. As a representative example Expert B said that

> "[…] I see the logic of having clear selection, and having elements of the selection functions having an effect on the internal effectiveness of teacher education, and a more successful teacher education is in this you might expect a higher quality of those teachers which could have a positive effect on student achievement. That part I understand […]" (p. 2, line 23)

The case is different for the allocation function; more specifically in case of the indicator of the effectiveness of the allocation function. Expert A mentioned that allocation is a genuine economic aspect which is frequently transferred to education systems, but seldom with a similar meaning, and that allocation, in economic terms, is about matches between what the labor market requires and the students which are allocated to jobs for which they are trained. He suggested another indicator of the effectiveness of the allocation function, namely the percentage of teacher education graduates that end up in teaching. After a discussion about the strength of the relation between school socioeconomic status and teacher shortages being a specific characteristic of allocation in education systems, however, he stated

> "[…] Yeah, it might be, I could imagine that if the allocation function is working well, then you could say that on the one hand teacher shortages should be minimized, because, there should be more teacher shortages in a country where the allocation function is not well, so that could be one dependent variable. And another one would be that if the allocation function is well then you would see no main differences across sectors of education or schools in terms of teacher shortage. That could include your relation with SES, right, but it might also be regional, or any other variation across schools or across educational types in terms of shortages. So you can broaden, actually you can broaden the dependent variable […]" (p. 3, line 5).

Expert B pointed out the potential use of the model for research on the allocation function in the context of teacher education. He compared the level of control over the recruitment process in the Netherlands and Germany, and came to the conclusion that the variation in aspects of the recruitment process "[…] makes it indeed interesting to make some comparisons with the situation in other countries […]" (p.3, line 2).

Based on the experts' statements I decided to model the relation between teacher education and the education system in a more refined way. More specifically, teacher education is a subsystem of the general education system, but not a subsystem of the lower secondary education school system. At this level there are two different systems interacting with each other. A second decision concerned the indicators of effectiveness of the allocation function. I followed the suggestion of Expert A and included the percentage of teacher education graduates who end up teaching as a second indicator, thus broadening the scope of allocation effectiveness. However, I retained the association between school socioeconomic status and teacher shortages as primary indicator. This reflects the view that this indicator adequately reflects the peculiarities of allocation in the context of education systems. The last decision concerned the length of the chain between selecting teacher education candidates and student teachers, allocating trained teachers,

and the influence of both on student achievement in the education system. It is interesting that the concerns raised by the experts reflect the statement of Konold et al. (2008) that little is to be learned of taking the long jump between teacher education and student achievement. Given the system focus of the model it may not adequately reflect what the teacher does in the classroom, and how he is able to translate his knowledge into effective teaching. With respect to these underspecifications I dropped the mean and variance in student achievement as indicators for the quality of an education system and focus more specifically on the selection and allocation functions and their immediate outcomes. In light of the expected limited data availability the stronger focus enhances the feasibility of the model.

3.5.4.2 Characteristics of the functions – Their dimensions and structural elements

The focus of the second part of the interview was on the question if both functions are sufficiently and reasonably characterized by their dimensions and structural elements. This included the mapping of structural elements to their respective dimensions. With respect to the dimensions none of the experts raised serious concerns about the number or core propositions of the dimensions. On the contrary the experts deemed the dimensions as a reasonable characterization of the selection and allocation functions. For example, Expert E mentioned that

> "[…] I think it is valuable to have a model like this for the kind of work I do to even, you know, be able to consider a model, because not many researchers have attempted to do this. [...] So yes, I think it is valid […]" (p.2, lines 23/34).

Furthermore, Expert C, Expert D, and Expert H stated that in light of the complexity of the model it would not make sense to characterize the functions with more dimensions, especially since the included dimensions already sufficiently characterize both functions. Expert A raised the question if the integration function captures a separate socialization or reproduction function, which is frequently mentioned in sociological research. Given that Open Systems Theory explicitly states that socialization measures are important for either monitoring what student teachers do during initial teacher training, and integrating trained teachers into schools, I retained the two dimensions. After making clear that this dimension is a specific characteristic of the theoretical basis Expert A had no objections against retaining the dimensions.

The experts deemed the different dimensions as sufficiently characterized by their respective structural elements. Only two experts suggested additional structural elements. More specifically, Expert E mentioned structured mobility as a feature of systems were teachers rotate schools and are provided with bonuses when they teach at certain kinds of schools. Expert F mentioned alternative certification as an additional structural element countering teacher shortages. However, it might interfere with the level of specific information about trained teachers, because alternative certification programs have distinct prerequisites with respect to their learning opportunities and students.

With respect to a possible differential relevance of the structural elements Expert A, Expert C, and Expert H stated that in principle a weighing of the elements according to their relevance is unnecessary, because it depends on the specific contextual situation in a given country. For example, Expert A mentioned that

> "[…] I think they are all fine, as long as it is all related to teacher education, and they are, I think. So I think it is good, and I would not know, there is no specific weighing here in terms of which indicator is best. I would regard this as five different indicators of the same concept, and produce some composite score […]" (p. 10, line 11).

And Expert E noted further

> "[…] I think in decentralized systems like the United States, I think the teacher labor market conditions like students population, you know, I'm talking about student population from a supply side perspective, in other words, how teachers respect their preferences for the clients, the students they teach, and so forth. I think those variables are going to be much more relevant seeing them in a decentralized system where teachers are negotiating locally with schools, you know in there kind of weighing their options and saying, okay do I want to teach in school A with this kind of student population or do I want to teach in school B. So I think that is a very important, it would have a very strong weigh, if you had a weighing scheme, in a decentralized system. Whereas in a more centralized system teachers are not able to do this so much and maybe the student composition is not as important […]" (p. 3, line 16)

The only indication for a general differential relevance of the structural elements is a distinction between proximal and distal elements. More specifically, Expert F mentioned that it is possible to distinguish between elements directly related to, for example, the selection of teacher training candidates, and elements related to what happens before and after this initial selection.

With respect to the mapping of structural elements to their dimensions the experts noted that some of the dimensions are related across functions because some elements could be assigned to more than one dimension. Furthermore, they noted that these interrelations may be adequately covered if some of the dimensions are relabeled and combined. This relabeling and combination specifically affects the accessibility of teacher education and the capacity of the teacher labor market. In this regard Expert F mentioned

> "[…] and I also mean occupational status of teaching, because you could further elaborate, or to be precise, it is a question of the indicators. Because eventually you could say that career ladders are also an element of the selection function. […]" (p. 6, line 48).

The main critique regarding the mapping of the structural elements was concerned with these elements. Expert A took up the aforementioned thought and stated

> "[…] the capacity includes structural elements influencing the number of available newly trained teachers and the number of available teaching positions. Maybe here you could relate it more to, say, the capacity relates to factors affecting supply and demand. And this would make it clearer. […] But my main issue was actually that I cannot see any difference between occupational status of teaching, and starting salary, working conditions, and career ladders. […] I think it is really the same. So, and then it is very strange to put one part in the selection function and one part in the allocation function. I would say put them together, because it is, this is all about both the absolute and relative attractiveness of the teacher occupation, nothing else […]" (p. 4, line 33/52)

Expert B pointed out the relation between union regulations and the occupational status, since in some countries the salary, working conditions, and career options are negotiated between teacher unions and other authorities such as school boards or ministries.

> "[…] It interferes with school autonomy, well that is certainly true in the Dutch situation, but it also means that it is strongly related to, for example, the issue of the salary structure, the work conditions. Even career ladders, because part of these issues are negotiated by the school boards, and maybe that is also a distinction, it is not something that the Ministry Of Education does, it is the school boards themselves, that negotiate with unions on salary structure, work conditions, career ladders and so on. […] Of course the government plays a role, but it is at a distance, in the sense that our government and the parliament, they decide what is let's say the total lump sum that is available for education. So, in that sense you could argue that in the Dutch situation the union regulation belongs more to the first dimension […]" (p. 3, line 37)

Based on the experts' evaluations of the characterization of the both functions with their dimensions and structural elements, I decided to integrate the dimensions accessibility of teacher education and capacity of the teacher labor market into a single dimension labeled supply and demand of teachers. Furthermore, career ladders and occupational status of teaching were integrated into one structural element capturing the occupational status of teaching. The integration of both dimensions and their structural elements took into account that it is not easy to distinguish a differential effect of the elements in the context of selection or allocation. Moreover, all of the elements describe features of the immediate context of the teacher education system, and influence the supply and demand of teachers in some ways. I further decided not to include union regulations in the capacity dimensions because I considered it more important to cover the interference with school autonomy.

Lastly, I decided not to include alternative certification as a contextual condition. This decision is based on its low proximity to the core elements of the functions. However, I included the structured mobility element as another contextual condition and assigned the element to the capacity of the teacher labor market as well. The reason for this decision was that it is a peculiarity of East Asian teacher labor markets, and that it determines the context of the recruitment process. Overall these decisions made the model easier to understand and increased its meaningfulness.

3.5.4.3 The operationalization of the structural elements

The focus of the last part of the interview was on the question if the indicators of the structural elements adequately reflect their theoretical meaning. Furthermore, it was discussed if a combination of certain structural elements in single indicators is justified. The last question concerned the question how the structural elements influence their respective dimension, for example, how certification and licensure affect the level of specific information.

Overall, the experts stated that the indicators of the structural elements reflect the theoretical meaning of the elements well. An interesting point was raised by Expert E concerning the data availability and its relation to the selection of indicators.

> "[…] I mean at first glance I think they are well developed and they reflect the theoretical size of the model, let me see. I think what you have included is certainly relevant to be included in the model, it seems to capture pretty well the concepts that you are modeling. But I am just curious if you have gotten into data, what you are finding in terms of the actual data. […] I mean, like I said, at first glance I think it is really well done and I think they are relevant and capture the theoretical framework, but you know how once you start actually getting into the data you start realizing that, oh, there is that other variable that might work good. But yes, I think as a starting point they look good […]" (p. 4, line 21)

Other experts also pointed out that data availability might alter the selection of indicators for the structural elements, especially with regard to international comparisons. Expert F mentioned that it is hard to find indicators which are comparable across countries. Expert C pointed out that the indicators were selected in a pragmatic way, i.e. guided by data availability. Despite this pragmatism, the selection of indicators was evaluated as adequate and only few changes were suggested. These changes primarily concerned indicators which were not directly related to teacher education itself, especially indicators for the stratification of the education and higher education system. While it was acknowledged that these indicators capture the theoretical meaning of the elements well, an alternative indicator was suggested by Expert A.

> "[…] I would give you an alternative indicator. My first impression was actually, if I look at selection, I would say the most important difference across countries in the type of teacher training is whether it is university based or not. […] In my view the best indicator for selection would be the percentage of students in teacher education who get a MA degree, or who follow a MA degree course, or say relative to another type, or BA-University versus another type. I think then you would have something which is really different from stratification, because stratification is really something which directly affects students, not the quality of teachers; or the quality of teachers only very indirectly because it raises the skills of every student, so also teacher students. But that is not what you want. You want something that is directly related to teacher education. So I would choose something like that, which would also, I mean, would be my number one indicator for the selection function, the level or the kind of level where teacher

> training takes place, whether it is MA, or University, or professional college or whatever […]" (p. 9, line 26)

Furthermore, Expert A suggested that the funding of teacher education and its indicator is only loosely related to teacher education.

> "[…] So in that case it is really not a good indicator, because it says nothing about teacher education. So funding I would say is probably, I do not know of any country where there is a difference in tuition fees between teacher education and other forms of higher education […]" (p. 9, line 49)

Based on the operationalization of the structural elements, Expert A suggested dividing the level of specific information dimension of the allocation function into two separate dimensions.

> "[…] The other one is really interesting, so you have a number of different indicators about information specific, the union regulation, licensing, certification, job descriptions, school autonomy, probation periods. They all relate to some way or another to regulation and standardization. I am not sure whether I am totally correct but more or less you could say, you could argue well all of these things one way or another may have to do with regulation, whether the same, whether it is more or less standardization, whether it has the same standard nationwide, or there is a lot of autonomy for schools or not. […] So you actually have two sub-dimensions here. One is more about say having nationwide standards, which is about licensing, certification, maybe job descriptions. And one is about, at the same time you want to have autonomy in terms of, it is really about the skills and not just some rule that someone should teach. Which is I mean, unions negative, and school autonomy positive and probationary periods positive. And maybe you can link them together as such […]" (p. 10, line 26)

Another suggestion related to the indicators of the structural elements concerned the career ladders and occupational status of teaching. As already explained in the previous section, these elements were combined into a single structural element. Accordingly, it was suggested to combine the indicators into one single indicator, or else, to use only the indicator for occupational status of teaching. For example, Expert E mentioned

> "[…] Yeah, I haven't thought about that, it's an interesting idea. So for example if you are a civil servant than you are guaranteed to have a career ladder. So there's sort of this interaction between the contractual status and career options, whereas if you are contract teacher there is none. Yes, I think it makes sense. I think it's valid […]" (p. 5, line 18)

With respect to other indicators which combine two structural elements the experts raised no concerns. On the contrary, they stated that the combinations of certification and licensure, school autonomy and job descriptions, admission/assessment procedures and admission/assessment criteria, is not only a good idea but a requirement of the theoretical notions of the elements. In this regard, Expert D and Expert B stated

> "[…] I think they are well developed, these combinations you suggested. They even arise as a consistent result of what you have as input [the notion of the structural elements]. I always like that […]" (p. 10, line 17)

> "[…] I would say, school autonomy and job descriptions, yeah... because of our schools being that autonomous, they are completely responsible for job descriptions. But at the same time school autonomy and job descriptions, they are two different things. I would almost say that job descriptions are a small element within school autonomy, or autonomy […]" (p. 5, line 46)

A combination of structural elements which was not modeled prior to but came up during the interviews was combining teacher induction and teacher mentoring as aspects of school-teacher education partnerships: "Yes, I actually think that is a very good idea!" (Expert F; p. 7, line 31). It was suggested to drop school-teacher education partnerships because the two other structural elements already capture the theoretical meaning of a structural linkage which facilitates the transition from teacher education to teaching (Expert E).

> "[…] I think possibly. Although, I think of teacher mentoring to have a very informal uplift to it; I think a lot of teacher mentoring is informal. You know, just interaction between teachers across classrooms and so forth. But I mean if you are talking about formal mentoring programs and requirements, yes I believe so […]" (p. 5, line 49)

With respect to the influence of the structural elements on their respective dimensions, based on their operationalization, the overall statement of the expert was that it is difficult to assume an influence prior to the actual analysis. However, it became clear that this only concerned contextual conditions. For example, Expert F stated

> "[…] The teacher labor market [in Australia] is highly flexible, and they like it that way and deem our labor market highly constraining and highly restrictive. […] They would advocate the hypothesis that the accessibility is highest in System 1 and not in System 3 […]" (p.5, line 25)

In case of the core elements of the functions, the experts deemed their influence on their respective dimensions valid and reasonable. In this regard, Expert F suggested transforming the continuous indicator 'student-teacher ratio' into a categorical in order to better reflect the respective state of research.

> "[…] Now we know that if a certain limit is not reached, [changes in the student-teacher ratio] have no effect. For example, between 20 and 35 there is almost no difference in student achievement. Only if the groups are considerably smaller than 20 there is a difference. So, there are jumps. This is no continuous indicator, that is my point […]" (p. 7, line 41).

Based on the experts' evaluations I decided to accept the suggestion to drop both stratification elements and instead to include the level and length of initial teacher training. Both structural elements implicitly comprise the stratification aspects, but are at the same time more closely related to teacher education. Furthermore, I decided to drop the funding of teacher education and career counseling for teachers because of their expected low variance across countries and, in case of funding, its distal relation to teacher education. In light of the ambiguous influence of alternative certification on both accessibility and standardization, the experts' statements further confirmed the decision to not include this element in the model. While it certainly increases the accessibility of teacher education, at the same time it decreases the level of specific information about trained teachers. This decrease is explained by a lower degree of standardization. More alternative certification opportunities decrease the probability that teachers in a given country all possess the same kind of training, especially in light of the different requirements of the alternative certification programs.

Moreover, I decided to accept the suggestion to divide the level of specific information of the allocation function into two separate dimensions: the level of specific information about trained teachers, and the level of control over the recruitment process. With that differentiation I include the standardization and control aspects mentioned by the expert and reach a better reflection of the theoretical meaning of the certification, probationary periods, union regulations, and school autonomy. Moreover, the indicators of career ladders and occupational status were combined into a single indicator, according to the integration of both elements into one as explained in section 3.5.4.2. Furthermore, school-teacher education partnership was dropped since this element is already reflected by teacher mentoring and teacher induction. All other combinations of structural elements were retained.

Regarding the influence of the structural elements on their respective dimensions I decided to acknowledge that this influence is unclear with respect to contextual elements of the capacity of the teacher labor market. Hence, no hypotheses might be formulated regarding their specific influence on the dimensions, and in turn regarding the influence of the configurations of both functions on teacher education effectiveness. The case is different for the core elements of both functions. However, I stress the exploratory nature of the model testing in the following chapters, given the necessary pragmatism in selecting indicators for the structural elements of both functions in light of the limited data availability. Eventually, it might be possible that other indicators for the structural elements are chosen. The final model resulting from the experts' evaluations, suggestions, and my decisions is illustrated in Figure 2. The respective dimensions, structural elements, and indicators are summarized in Table 4.

Table 4. Dimensions and elements of the final organizational model of teacher education.

Function	Dimension	Structural Element	Indicator (Source)
Contextual Conditions (Supply and Demand)	*Capacity of the Teacher Labor Market*	*Length and Level of Teacher Education*	*Type and length of initial teacher training (OECD, 2012a); categorical*
		Student Population	*Student-teacher ratio (OECD, 2012a); categorical*
		Occupational Status	Type of employment system (Eurydice, 2012); categorical
		Salary Structure	Starting salary combined with the years required to top salary (Eurydice, 2012); categorical
		Work Conditions	Type of continuing professional development (Eurydice, 2012); categorical
		Structural Mobility	*Implementation of teacher rotation in the education system (own collection); binary (yes/no)*
Selection	Level of Information about Candidates and Students	Admission/Assessment Procedures	Structure of the selection process (OECD, 2012a); categorical
		Admission/Assessment Criteria	Degree of central regulation of teacher education (Horn, 2009); categorical
	Level of Integration of Students into Teaching	Internal Support	Availability of guided teacher practica during initial teacher training (OECD; 2012a); binary
		Field Experiences	Minimum amount of professional training during initial teacher training (Eurydice, 2009); continuous
Allocation	*Level of Information about Trained Teachers*	Licensure/Certification	Requirements to enter the teaching profession (OECD, 2012a); categorical
		Probationary Periods	Type of recruitment process (OECD, 2012a); binary (with/without)
	Level of Control over the Recruitment Process	School Autonomy	Index of school responsibility for resource allocation (OECD, 2010; 2013); continuous
		Union Regulations	Percentage of schools where the principal reported that teacher union exert direct influence on staffing (OECD, 2010); continuous
	Level of Integration of Teachers into Schools	Teacher Induction	Length of the induction period (Eurydice, 2012); continuous
		Teacher Mentoring	Type of regulation of teacher mentoring (Eurydice, 2012); categorical

Note. Changes in the final model compared to the original model are in italics.

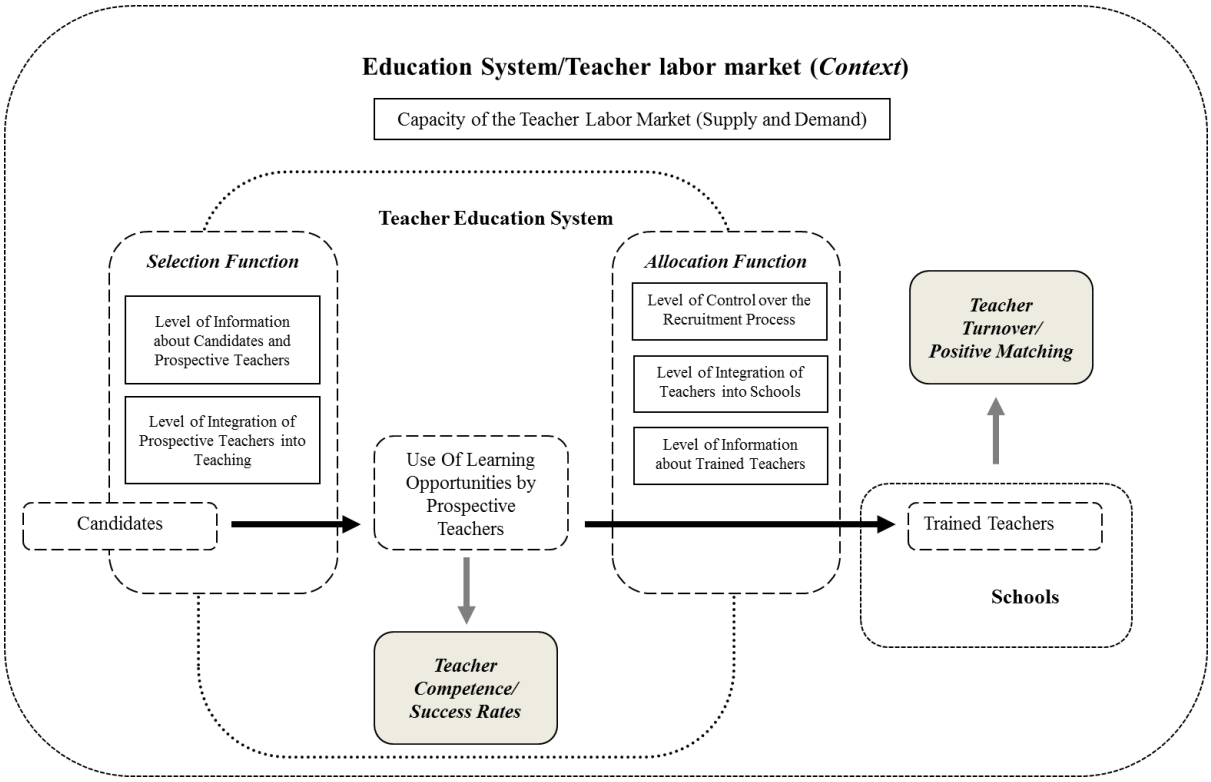

Figure 2. The final model of teacher education as an open system; rectangles depict the dimensions of the selection and allocation function, as well as contextual conditions in the education system/teacher labor market. The black arrows illustrate the transition of an individual through teacher education into schools, from teacher education candidate over student teacher to a trained teacher in a school. The gray arrows depict the consequence of the use of learning opportunities by student teachers on their competence and success rates as individual and organizational level indicators of teacher education effectiveness. Teacher turnover and positive matching are individual and organizational manifestations of specific teacher distributions, and thus indicators of the effectiveness of the allocation function.

3.6 Summary of the chapter

In this chapter I developed a model of teacher education based on Open Systems Theory following a two-stage strategy. First, the general outline of the organizational perspective on teacher education as an open system was developed. The overarching research question was: What does a model of teacher education, taking into account the conceptual, complexity, inherent selection, and non-random allocation problems, look like? In order to further develop the model, I conducted an interview study with experts in the fields of international comparisons of education and teacher education systems, professionalism of teachers, teacher education effectiveness, and characteristics of teacher labor markets. The research questions of the

interview study were: (1) is the model of teacher education as an open system a valid representation of teacher education and its relation with the general education system? (2) Are the selection and allocation functions sufficiently and reasonably characterized by their dimensions and structural elements? (3) Does the operationalization of the structural elements reflect their theoretical meaning?

Based on the experts' assessment of the validity and meaningfulness of the different parts of the model I conclude that the model of teacher education as an open system is an appropriate way to take into account the conceptual and complexity problem in research on the relation between teacher education and student achievement. The interaction between teacher education and the general education system is appropriately represented. However, as suggested by the experts, in light of a limited data availability and the complexity of the interaction of teachers, students, and the school context, it may be not possible to test the full complex chain between teacher education, teacher characteristics, teacher behavior and student achievement. Hence, although it still may not be possible to explain variation in effects on student achievement across teacher education programs, researchers are now in a position to explain variation in the development of knowledge, attitudes and beliefs of student teachers (Boyd et al., 2009; Denzler & Wolter, 2009). The model shifts the focus of teacher education effectiveness back into the teacher education system, which might be a more appropriate way to assess its effectiveness (Morge et al., 2010). Moreover, the model offers a framework which various studies on the relation between student teachers and their use of learning opportunities, on teacher characteristics and teacher behavior, and on teacher preferences in the recruitment process may be integrated in. As suggested by the experts I keep the full interrelation between teacher education, characteristics, their behavior and student achievement in mind, but focus on aspects of the model which also warrant attention given the current state of research. Nevertheless, the model might provide research on teacher education the organizational perspective Zeichner (2006) deemed necessary to be better able to explain teacher education policy and practice.

With respect to the inherent selection and non-random allocation problem the model offers new opportunities for research. Based on the experts' assessment of the characterization of the selection and allocation functions, their structural elements, and their operationalization I conclude that testing these parts of the model may provide research new insights and possible explanations of both problems. In general, the experts considered the characterization of both functions complete, thorough, and theoretically meaningful. The changes in the model which were based on the experts' suggestions and evaluations were important with respect to the integrity of the structural elements of both functions. In order to be able to explain the development of selection effects on the use of learning opportunities of student teachers, as well as the development of different teacher distributions, it is important for the model to include all possible relevant structural elements. It might still be possible that not all of the structural elements will be represented in respective studies, but that only reflects the specific situation in a given country, and in turn a specific configuration of the function in terms of information, control, and integration. It is important to note that in light of the pragmatism regarding the selection of the indicators, the tests of the selection and allocation parts of the model are

exploratory in nature, even though the experts deemed the indicators of the core structural elements as appropriate. Thus, in the next two chapters I test the model focusing on the inherent selection and non-random allocation problem. More specifically, I firstly investigate differences in the use of learning opportunities by student teachers across different configurations of selection functions. Secondly, I investigate differences in the degree of positive matching across different configurations of allocation functions.

CHAPTER 4

TESTING THE MODEL, PART I: *Comparing selection functions of teacher education systems: towards more certainty in sorting student teachers*

4.1 Aim and structure of the chapter

In this chapter I address the inherent selection problem and investigate different configurations of selection functions with respect to differences in the relation between characteristics of student teachers and their use of learning opportunities provided by the teacher education system. As Schalock et al. (2006) state, little is known about how to select teacher education candidates. Consequently there is no definite consensus regarding selection criteria or methods in the context of teacher education (Blömeke, 2009). It is unclear which characteristics of teacher education candidates to look for because most studies relate student teacher characteristics directly to distal outcomes such as teacher knowledge, beliefs, and attitudes. Respective research does not consider the relation of student teacher characteristics and their use of learning opportunities (Casey & Childs, 2007). A notable exception is the Teacher Education and Development Study in Mathematics (TEDS-M; Tatto et al., 2012). This international comparative study explicitly focuses on the relation between learning opportunities and competence development. However, this study does not relate the use of learning opportunities to student teacher characteristics. The model outlined in the previous chapter provides an opportunity to consider both aspects of the inherent selection problem simultaneously. It might help to identify student teacher characteristics which are predictive of a successful use of learning opportunities during initial teacher training. Additionally, it may allow researchers to identify configurations of selection functions which are more effective in selecting and sorting the right candidates and student teachers.

Thus, in this chapter I aim at illustrating a potential solution to the inherent selection problem. By means of a multigroup analysis of TEDS-M data this study seeks to answer the following research questions. *First, what is the relation between student teacher characteristics and their use of learning opportunities? Second, does the configuration of teacher education's selection*

function moderate the relation between student teacher characteristics and their use of learning opportunities? Third, are different configurations of teacher education selection functions associated with differences in the student teacher characteristics and their use of learning opportunities? With these three questions it is possible to identify student teacher characteristics which are predictive of their use of learning opportunities, and additionally to evaluate the effectiveness of different approaches to selecting candidates and student teachers with respect to these characteristics.

This chapter is structured as follows. In the next subsection I describe the selection functions of three countries, namely Poland, Singapore, and the USA. The description is based on the characterization of the selection functions outlined in Chapter 3. In the following sections I describe the conceptual model to be tested and the methodological approach taken. The results are presented in the fourth subsection. A summary of the chapter is given in the last subsection.

4.2 Different configurations of selection functions – Poland, Singapore, and the USA

This section describes the structural arrangements of the selection functions of the lower secondary teacher education systems in Singapore, Poland, and the USA. The teacher education systems have been chosen due to the differences in the configurations of their selection functions. Information about the different dimensions of the selection function has been obtained from the TEDS-M institutional database and policy reports (Tatto et al., 2012; Ingvarson et al., 2013). The detail of the institutional database with respect to the admission and assessment procedures, as well as to the level of integration of student teachers allows for a more refined description of the selection function. Thus, the indicators of the structural elements and their respective dimensions differ from the indicators presented in Chapter 3.

The capacity of the teacher labor market is indicated by the occupational status of teaching in the respective country and the level of control over the supply of student teachers. Furthermore, I calculated the mean duration of initial teacher training as well as the credential earned (length and level of teacher education). Thus, with regard to contextual conditions I focus on aspects related to the supply of teachers. In Singapore the average length of initial teacher training is 3.00 years in the concurrent program, and 4.00 years in the consecutive program. In Poland the mean duration of the concurrent program is 3.58 years (no information about consecutive programs was provided). In the USA the mean duration of the concurrent program is 3.83 years, and the mean duration of the consecutive program is 5.44 years. Student teachers earn an ISCED 5A credential across programs in all countries. While length and level of initial teacher training is comparable across the three countries, there are differences in the teacher employment system. Singapore has implemented a career-based employment system, where teachers are employed as civil servants permanently. Working conditions are favourable (Ingvarson et al., 2013), and salaries are performance based and high. The USA has a position-based system implemented, where teachers are employed on a contractual basis. Working conditions are diverse and salaries are relatively low (compared to other occupations). Poland has a hybrid employment system with moderate salaries (no information about working conditions was available). Singapore is the only

country where national authorities match the number of available study places to the number of required teachers in the school system; in Poland and the USA teacher training institutions have more freedom in determining the available study places. In sum, supply and demand of teachers is well balanced in Singapore and less well balanced in Poland and the USA.

Policies for admission of student teachers are set by regional or national authorities or by the institutions with guidelines set by regional or national authorities in all countries. However, there are differences in the specific admission policies. Singapore requires potential student teachers to pass a national examination at the end of secondary education, and further assesses their suitability for teaching with a focus on the overall level of achievement in mathematics. Poland requires potential student teachers to pass a national examination and focuses more on the level of achievement at the end of secondary education. In the USA potential student teachers are required to pass an entry examination for admission to a specific training institution as well as an additional assessment of their suitability for teaching with a focus on the overall level of achievement in mathematics. However, the requirements vary across states and are limited in scope (Ingvarson et al., 2013). Policies for graduation of student teachers are set by state educational authorities and single teacher training institutions in the USA and Poland. Singapore does not have such policies. Regarding the requirements for successful graduation, all countries require a passing grade on all of the subjects and field experiences of the program. While Singapore and the USA require student teachers to successfully demonstrate teaching competence in the classroom, Poland requires students to write and defend a thesis. Poland is furthermore the only country which requires passing an institution specific exit examination. In sum, given the focus on the suitability for teaching of the admission and assessment policies in Singapore, the level of information about candidates and student teachers is highest. Poland and the USA have a more limited admission policy and focus more on graduation requirements. However, given that the USA requires student teachers to pass an examination administered by regional authorities, the level of information is higher than in Poland.

The level of integration of the three selection functions is indicated by the amounts of extended teaching practice and introductory field experiences (in days) provided by the teacher education systems. Extended teaching practice is continuous work in schools in order to enable student teachers to assume responsibility for teaching a whole class of students. Introductory field experiences are short term assignments in schools for preparatory purposes (Breese & Tatto, 2012). Singapore's teacher education system provides on average 56 days of extended teaching practice during initial teacher training, and no introductory field experience. Poland's teacher education system provides on average 48 days of extended teaching practice and 33 days of introductory field experiences. In the USA the teacher education system provides on average 92 days of extended teaching practice and on average 41 days of introductory field experiences during initial teacher training. Overall, the selection function of the USA teacher education system has a high level of integration of student teachers, while the selection function of the teacher education system in Poland has a moderate level of integration. The level of integration of student teachers of the selection function in Singapore's teacher education system is limited.

Table 5 summarizes the configurations of the selection functions of the three teacher education systems.

Table 5. Configurations of the selection functions – Singapore, Poland, and the USA

	Capacity of the Teacher Labor Market (Supply)			Level of Information	Level of Integration
Country	Occupational Status	Control over Supply	Length & Level of Teacher Education		
Singapore	High	High	3.00 / 4.00 (5A)	High	Limited
Poland	Moderate	Limited	3.85 (5A)	Limited	Moderate
USA	Limited	Limited	3.83 / 5.44 (5A)	Moderate	High

Note. Adapted from Tatto et al. (2012); Length: first number is the length of the concurrent, second number is the length of the consecutive teacher education program.

4.3 The selection function and the use of learning opportunities by student teachers

The differences in the configurations of the three selection functions provide an opportunity to test how the use of learning opportunities by student teachers varies across different approaches to select teacher education candidates and student teachers. More specifically, based on the theoretical framework outlined in Chapter 3 I assume that a selection function with high levels of information about student teachers and integration of student teachers avoids adverse selection of student teachers in terms of unfavourable characteristics. Following the use of learning opportunities model, one of the most important characteristics is motivation (Blömeke, 2009; Helmke, 2012). Motivation determines the quality of the learning activities constituting the use of learning opportunities, which in turn result in different knowledge and skill levels. As already mentioned, when the quality of the individual learning activities decreases, the use of the learning opportunities becomes suboptimal; this is likely the case when student teachers show unfavourable motivational characteristics. An important aspect is the motivational orientation of the student teacher towards the teaching profession (Malmberg, 2006). A common distinction between different motivational orientations of student teachers is between altruistic-pedagogical, subject-related, and extrinsic motives (Blömeke, Suhl, Kaiser, & Döhrmann, 2011). Intrinsic sources of motivation such as liking teaching (Younger, Brindley, Pedder, & Hagger, 2004) and having an impact on the lives of youth (Richardson & Watt, 2006) have been found to have a positive impact on learning activities of student teachers (Malmberg, 2008). At the same time there is concern with respect to extrinsic sources of motivation, such as job security and salaries (Moran, Kilpatrick, Abbott, Dallat, & McClune, 2001). Blömeke et al. (2011) found a negative relation between extrinsic motives and knowledge development of primary education student teachers. However, few studies directly relate the motivational orientations of student teachers to their use of learning opportunities.

Based on the theoretical framework and current research I hypothesize that an altruistic-pedagogical, subject-related, and extrinsic motivational orientation of student teachers has a positive impact on their use of learning opportunities. With respect to the use of learning opportunities I adopt the available data in the TEDS-M database and distinguish between subject-specific and pedagogical learning activities as operationalization of the use of learning opportunities. This part provides an answer to the first research question. Next, I investigate the second research question asking if this relation is moderated by the configuration of the selection function. The information about the different selection configurations is included as contextual information. This is similar to the configural approach to education systems taken by Hofman et al. (2008). This means that I estimate the model illustrated in Figure 3 (it already includes control variables which are explained in section 4.4.3) for each of the three teacher education systems and compare the strength of the relations between the constructs. I expect that the strength of the relation between motivational orientation and use of learning opportunities is equivalent across teacher education systems. Lastly, based on the theoretical framework I hypothesize that there are differences in the average motivational orientation of the student teachers and their learning activities across the teacher education systems. More specifically, I expect that a selection function characterized by high levels of information and integration is associated with student teachers showing higher altruistic-pedagogical and subject-related motivational orientations. Moreover, this selection function is expected to be associated with students with a lower extrinsic motivational orientation, and higher subject-specific and pedagogical learning activities.

4.4 Method

To test these hypotheses and to answer the research questions a multiple group analysis was conducted. This kind of analysis involves three steps: (1) tests of measurement invariance in order to determine if the constructs are comparable across countries, (2) tests of the equivalence of the structural paths in the conceptual model to determine if there are cross-country differences in the strength of the hypothesized relations, and (3) investigation of differences in the latent means across the three teacher education systems.

4.4.1 Database

I made use of the TEDS-M international database (Tatto et al., 2012). This international study was conducted in 2008 in 17 countries in order to investigate learning opportunities and mathematical content and pedagogical content knowledge of future primary and lower secondary teachers. All analyses were based on the TEDS-M dataset version 3.2 which was provided by the IEA data processing centre on April 4, 2012.

4.4.2 Sample

TEDS-M implemented a stratified multi-stage probability sampling design. Future teachers were randomly selected from a list of future teachers for each of the institutions in a country, which were randomly selected as well. The original sample of future lower secondary teachers participating in TEDS-M was 8332 in 16 countries (Tatto et al., 2012). In Poland 298, in Singapore 393, and in the USA 607 future lower secondary teachers participated. After excluding cases with missing values on all variables, the effective sample size of the present study was 298 (Poland), 392 (Singapore), and 500 (USA) future lower secondary teachers.

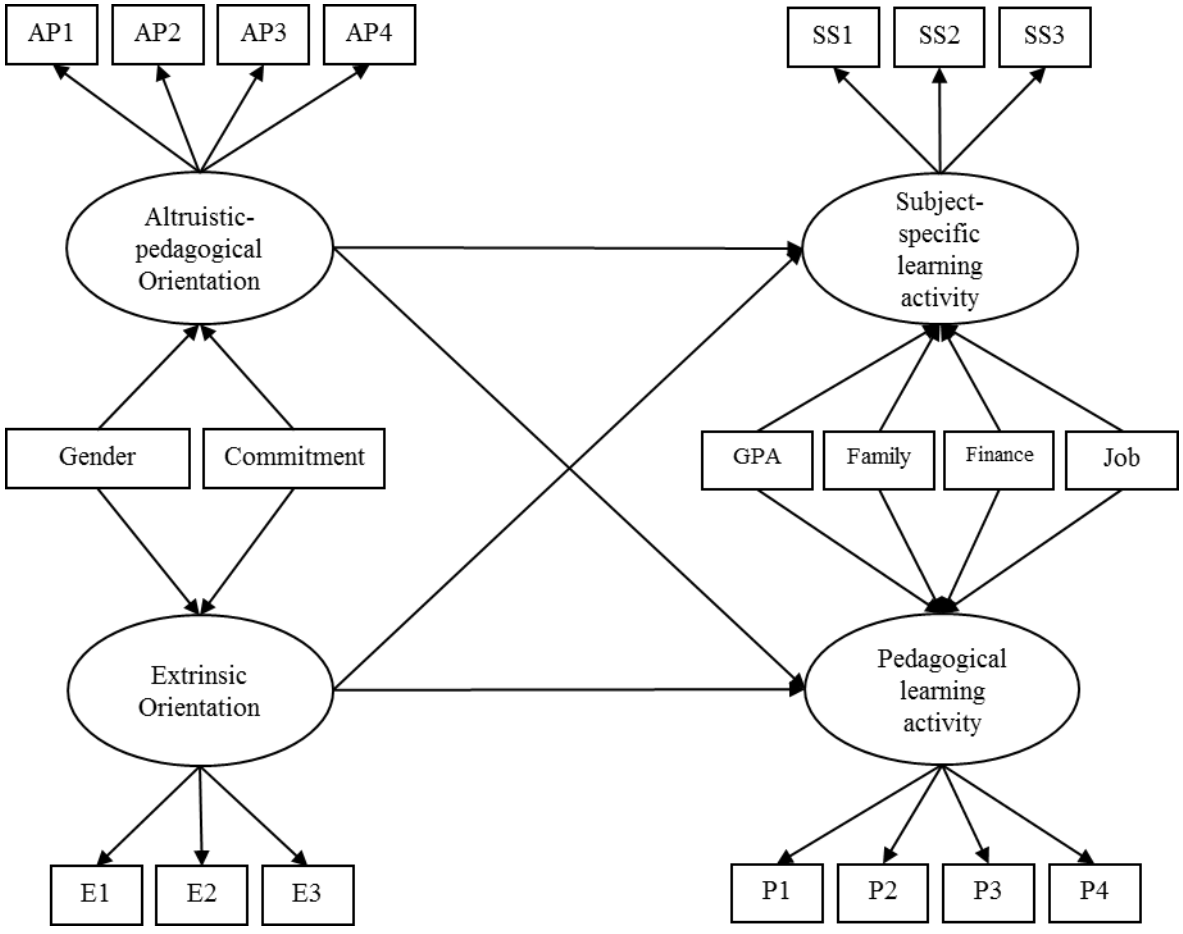

Figure 3. The hypothesized relations between motivational orientation and learning activities; altruistic-pedagogical and extrinsic motivational orientation are hypothesized to positively influence subject-specific and pedagogical learning activities of student teachers. The motivational orientations are controlled for gender and commitment differences. The learning activities are controlled for differences in mean achievement in secondary education, and differences in family-, finance- and job-related hindrances of student teachers. AP, E, SS, and P are indicators of the respective latent variables (see 4.4.3 for a description of the indicators).

4.4.3 Variables

In order to investigate differences in the relationship between student teachers' motivational orientation and their learning activities during initial teacher training across the countries, I used the following variables from the future lower secondary teacher questionnaire. The motivational orientation of student teachers was measured by three scales indicating their altruistic-pedagogical, subject-related, and extrinsic motivational orientation to become a teacher. The dimensions consisted of four, two, and three statements rated on a four-point Likert scale ranging from "not a reason" to "a major reason" (Blömeke et al., 2011). An indicator for altruistic-pedagogical motivational orientation was, for example, *"I believe that I have a talent for teaching"*. An indicator for subject-related motivational orientation was *"I love math"*, and an indicator for extrinsic motivational orientation was *"I am attracted by the availability of teaching positions"* (Brese, 2012). The reliability of the scales was satisfactory for the altruistic-pedagogical and extrinsic motivational orientation scales (Cronbach's α = .781 and .607, respectively). The reliability of the subject-related scale was unacceptable (Cronbach's α = .444). Blömeke et al. (2011) suggest that a reason for the low reliability is the low number of items. In light of the low reliability of the subject-related scale, and with respect to potential identification problems associated with latent variables with only two indicators, I decided not to include it in the model. The items were used as indicators for the latent variables altruistic-pedagogical and extrinsic motivational orientation of student teachers.

The use of learning opportunities by student teachers was measured by two scales indicating their learning activities in the mathematics education courses during initial teacher training, and their learning activities in the math education pedagogy courses during initial teacher training. The dimensions consisted of 15 and 26 statements rated on a four-point Likert scale ranging from "never" to "often" (Brese, 2012). An indicator for learning activities during mathematics courses was, for example, *"In the mathematics education courses that you have taken or are currently taking in your teacher preparation program, how frequently did you read about research on teaching and learning".* An indicator for learning activities during mathematics education pedagogy courses was, for example, *"In the mathematics education pedagogy courses that you have taken or are currently taking in your teacher preparation program, how frequently did you read about research on teaching and learning"* (Brese, 2012). The reliability of the scales was satisfactory (Cronbach's α = .831 and .944, respectively). The advantage of these indicators was their focus on what student teachers actually do in their teacher training courses. The high reliability of the scales and the high number of items justified the parcelling of the items. Three item parcels were formed for learning activities in mathematics courses, according to their thematic similarities outlined in the TEDS-M user guide (Brese, 2012). These parcels indicated the class participation, class reading, and problem solving activities of student teachers in their mathematics courses. Similarly, four parcels were formed for learning activities in mathematics education pedagogy courses: instructional practice, instructional planning, assessment uses, and assessment practice (Brese, 2012). These parcels were used as indicators for the latent variables subject-specific and pedagogical learning activities.

Furthermore, I included the gender of the student teachers as a control variable for their motivational orientations. The choice of this control is based on findings indicating gender differences in motivational orientations (Malmberg, 2006). This variable was dichotomous with two categories (1 = female, 2 = male). In the sample 63.8 % of the student teachers were female. Next, I included a variable measuring the commitment to teaching of the student teachers as another control for motivational orientation. This variable indicated if student teachers plan a career in teaching. The choice of this control variable is based on numerous studies finding a relation between commitment to teaching and the motivational orientation of student teachers (Malmberg, 2006). The variable had a four-point Likert scale ranging from "lifetime career" to "not seeking employment as a teacher".

With respect to controls for learning activities of student teachers, I included three dichotomous items indicating family-related, job-related, or financial circumstances hindering learning activities of student teachers (1 = yes, 2 = no). In the sample, 20.0% experienced family-related, 17.8% experienced money-related, and 27.2% experienced job-related hindering circumstances. Moreover, the overall level of achievement during secondary education was included as a control for learning activities. This variable had a five-point Likert scale ranging from "always at the top" to "generally below average". The choice of this control was based on the predictive value of this cognitive characteristic for study success (Blömeke, 2009). The means and correlations of the variables included in this study can be found in Appendix C. The conceptual model to be tested is illustrated in Figure 3.

4.4.4 Analysis

The multiple group approach to data analysis taken in this study consisted of three steps. In the first, measurement invariance of the four constructs was tested in order to ensure that they were comparable across countries. If measurement invariance holds, the differences in latent means can be meaningfully interpreted. The analysis is, in principle, an estimation of a series of models with specific constraints and an evaluation and comparison of the fit of these models. Therefore, I used the χ^2-test to assess model fit, where an insignificant result indicates good model fit. However, given its dependence on sample size (Kline, 2013) I furthermore inspected the absolute values of the correlation residuals. Absolute values below .10 indicate that the relation between two variables is adequately estimated. Moreover, the evaluation of model fit was based on the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA) with its 90% confidence interval, and the Standardized Root Mean Square Residual (SRMR). Acceptable model fit is indicated by a CFI higher than .90, a RMSEA lower than .08, and SRMR lower than .08 (Hu & Bentler, 1999; Marsh, Nagengast, & Morin, 2013). Good model fit is indicated by a CFI close to .95, a RMSEA lower than .06, and SRMR lower than .05 (Hu & Bentler, 1999; Steinmetz, Schmidt, Tina-Booh, Wieczorek, & Schwartz, 2009). Information about the goodness of fit indices used for model comparisons is included in the description of the different steps of the analysis.

The measurement invariance analysis of the present study followed the approach suggested by Thompson and Green (2013). To assess measurement invariance, the hypothesized measurement model was first specified and tested for each country separately. After evaluating the fit of these baseline models, the measurement model was tested in all countries simultaneously (the configural invariance model). Next, in order to establish metric invariance, the factor loadings of the indicators were constrained to be equal across the three countries. The metric model was compared to the configural model by means of a χ^2 difference test. In case of a non-significant difference in χ^2 the factor loadings are invariant across countries. Furthermore, I followed the suggestions of Chen (2007) to evaluate invariance of factor loadings. A change in CFI of lower than .010, lower than .015 in RMSEA, and lower than .030 in SRMR compared to the configural model is indicative of invariance of the factor loadings. In the last step, additionally to the factor loadings, the intercepts of the items were constrained to be equal across countries. Invariance of intercepts is achieved if the χ^2 difference between the scalar and metric models is insignificant, and the change in fit is lower than .010 in the CFI, lower than .015 in the RMSEA, and lower than .010 in the SRMR (Chen, 2007). While metric invariance is a prerequisite for equivalence tests of the structural paths of the model, scalar invariance is needed when differences in the latent means are to be compared.

The second step involved testing the equivalence of the structural part of the model. This analysis followed the approach taken by Levesque, Zuehlke, Stanek, & Ryan (2004). First, the structural model was estimated for each country separately. After evaluating the fit of these structural baseline models, the structural relations among the four constructs were tested for all countries simultaneously. Next, the structural paths were constrained to be equal across countries. The unconstrained and constrained models were then compared regarding their χ^2 values (Little, 1997). A significant change in the χ^2 values implies that there is at least one structural path which is not equivalent across the countries. To identify the non-equivalent paths, they were constrained one by one, that is, a series of models was estimated where only one structural path was constrained. Lastly, the structural paths were constrained according to the amount of change in χ^2, i.e. first the path with the least amount, then the path with the second highest change in χ^2. These steps were carried out until all structural paths producing no significant changes in χ^2 were included (Levesque et al., 2004).

The third step concerned the investigation of differences in the means of the latent constructs, following the reference-group approach described in Thompson and Green (2013). In the estimated models the intercept for the factor under investigation was fixed at zero in all countries. The other intercepts were freely estimated except in the reference country, where it was fixed to zero. A significant change in the goodness of fit indices when comparing the mean-restricted models to the scalar model implied that the respective factor mean was different across countries (Thompson & Green, 2013). Pairwise comparisons between the countries were based on χ^2 difference tests with two nested models. In the less constrained model the factor means of one country involved in the pairwise comparison were fixed to zero, while the factor means in the remaining countries were freely estimated (Model A). In the more constrained model, the factor means for both groups involved in the pair wise comparison were fixed to zero, and the

factor means for the remaining country were freely estimated (Model B). A significant change in χ^2 indicated differences in the factor means. In this kind of analysis, the parameter estimates for the factor means were differences in the means compared to the reference country.

All analyses were carried out in Mplus 6.12 (Muthén & Muthén 1998-2013). To account for non-normality of the data, I used the restricted Maximum Likelihood estimator (MLR). In combination with the routine for complex samples, this estimator additionally corrects standard errors for non-independence of observations. Weights and stratification information were incorporated to account for the complex sampling design and the associated unequal selection probabilities in order to obtain robust parameter estimates and standard errors (Blömeke et al., 2011; Brese, 2012).

4.5 Results

This section is structured as follows. First, the results of the measurement invariance analysis are presented. In the next sub-section, the results of the test of the equivalence of the full structural equation model are described. The last sub-section includes the results of the test for differences in the latent means of the latent constructs.

4.5.1 Testing construct comparability across the three countries

In order to determine if altruistic-pedagogical and extrinsic motivational orientations of student teachers are comparable across countries I firstly estimated the four factor (altruistic-pedagogical and extrinsic motivational orientation, subject-specific and pedagogical learning activities) measurement model separately for each country. After specifying nine residual covariances in the model for Singapore, and twelve residual covariances each in the models for Poland and the US, global and local model fit was good for each of these baseline models (see Table 6). All absolute correlation residuals were < .10 in all baseline models. Next, I estimated an unrestricted model where all measurement parameters (factor loadings and intercepts) were freely estimated. The fit of the configural model was good ($\chi^2(183) = 240.310$, $p < .01$, RMSEA = .028 with 90% CI [.017, .037], CFI = .985, SRMR = .043). All absolute correlation residuals were below .10 in all countries. The good fit of the configural model suggested a good fit of the hypothesized measurement model to the data in each of the three countries. Configural invariance is the necessary prerequisite for the test of the stricter invariance models.

In the second model the factor loadings were constrained to be equal across the three countries. The metric invariance model showed good fit as well ($\chi^2(203) = 297.886$, $p < .001$, RMSEA = .034 with 90% CI [.026, .042], CFI = .976, SRMR = .061). All absolute correlation residuals were < .10 in all countries. Changes in RMSEA, CFI, and SRMR between the configural and metric invariance models were marginal. These results suggest that the constructs altruistic-pedagogical and extrinsic motivational orientation, as well as subject-specific and pedagogical learning activities are understood similarly and thus are comparable across the three

countries. Furthermore, the requirements for comparing regression coefficients and for testing scalar invariance are met.

Table 6. Fit indices for the baseline and measurement invariance models.

Model	χ^2	df	RMSEA [90% CI]	CFI	SRMR	$\Delta\chi^2$ (df)	ΔRMSEA	*ΔCFI*	*ΔSRMR*
Baseline Models									
Poland	81.99*	62	.033 [.005, .051]	.982	.047				
Singapore	95.27*	63	.036 [.020, .050]	.984	.029				
USA	76.86	59	.025 [.000, .039]	.984	.051				
Invariance Models									
Configural	240.31*	183	.028 [.017, .037]	.985	.043				
Metric	297.89**	203	.034 [.026, .042]	.976	.061		+ .006	- .009	+ .018
Scalar	918.20**	223	.089 [.083, .095]	.822	.111		+ .055	- .094	+ .048
Partial Scalar	320.96**	209	.037 [.029, .045]	.971	.063		+ .003	- .005	+ .002
Structural Models									
Poland	204.30**	133	.043 [.031, .054]	.942	.052				
Singapore	285.01**	134	.054 [.045, .063]	.936	.041				
USA	171.79**	132	.025 [.012, .035]	.965	.049				
Unconstrained	669.54**	418	.039 [.034, .045]	.940	.056				
Constrained	674.22**	426	.039 [.033, .044]	.941	.058	6.63 (8)	+ .000	+ .001	+ .002

Notes. Model comparison and selection for the invariance models were based on differences in RMSEA, CFI and SRMR (Thompson & Green, 2013). Model comparison and selection for the structural and between-country models included χ^2 difference testing. * < .05, *** p < .001.

In the third model the intercepts were constrained to be equal across the three countries. The fully constrained model initially showed poor fit to the data ($\chi^2(223) = 918.204$, $p < .001$, RMSEA = .089 with 90% CI [.083, .095], CFI = .822, SRMR = .111). Five correlation residuals in Poland and four correlation residuals in Singapore had an absolute value > .10. Changes in RMSEA, CFI, and SRMR, compared to the metric model, clearly indicated non-invariance of the intercepts. Partial scalar invariance could be established after relaxing the constraints for seven intercepts. Model fit of the partial scalar invariance model was acceptable, and differences in RMSEA, CFI, and SRMR compared to the metric model were marginal ($\chi^2(209) = 320.958$, $p < .001$, RMSEA = .037 with 90% CI [.029, .045], CFI = .971, SRMR = .065). All absolute correlation residuals were < .10 in all countries. Thus, differences in the means of all the latent constructs can be meaningfully interpreted. This is important for the assessment of differences in the factor means across countries.

4.5.2 Influence of motivational orientation on learning activities equivalent across countries

Given that metric invariance could be established, it was possible to test the structural model depicted in Figure 3 in order to investigate the relation between the motivational orientations of student teachers and their learning activities. First, I estimated the structural model for each country separately to evaluate if the model fits the data in each of the three countries. Next, I estimated a model where the paths from both altruistic-pedagogical and extrinsic motivational orientation to subject-specific and pedagogical learning activities were unconstrained, and compared the resulting χ^2 with a model where these paths were constrained to be equal across countries. The paths from the control variables to the latent constructs were not constrained to be equal across countries.

Model fit was good in all of the three countries (see Table 6). However, the absolute correlation residuals between achievement level and the indicator "I believe I have a talent for teaching" had a value > .10. This means that the model overestimated the relation between the two variables. Next, the hypothesized model was tested simultaneously in all three countries. The paths between the latent constructs were freely estimated. The fit of this unconstrained model was good ($\chi^2(418) = 669.543$, $p < .001$, RMSEA = .039 with 90% CI [.034, .045], CFI = .940, SRMR = .056). Next, I constrained the paths from altruistic-pedagogical and extrinsic motivational orientation to subject-specific and pedagogical learning activities to be equal across countries. The fit of this fully constrained model was adequate as well ($\chi^2(426) = 674.223$, $p < .001$, RMSEA = .039 with 90% CI [.033, .044], CFI = .941, SRMR = .058). The change in χ^2 was insignificant ($\Delta\chi^2(8) = 6.6259$, $p = .58$) and the changes in RMSEA, CFI, and SRMR were marginal. This suggests that the strength of the influence of altruistic-pedagogical and extrinsic motivational orientation on subject-specific and pedagogical learning activities is equivalent across countries. Hence, the configuration of the selection function does not moderate the relation between motivational orientation and learning activities.

As expected, both altruistic-pedagogical and extrinsic motivational orientations positively influence subject-specific and pedagogical learning activities. However, the influence of the altruistic-pedagogical motivational orientation is stronger than that of the extrinsic motivational orientation (β = .314 (.053) and β = .263 (.055) compared to β = .068 (.027) and β = .092 (.034) on subject-specific and pedagogical learning activities, respectively). This confirms both hypotheses about the relation between motivational orientation and learning activities. The somewhat weaker relation between extrinsic motivational orientation and subject-specific and pedagogical learning activities further reflects the uncertainty with respect to the effect of extrinsic motives (Moran et al., 2001; Blömeke et al., 2011). Moreover, commitment to teaching significantly influences the motivational orientation of student teachers in all countries. Students who plan to stay in teaching show a higher altruistic-pedagogical and extrinsic motivational orientation than students who do not. An exception is the USA where there is no relation between commitment to teaching and extrinsic motivational orientation.

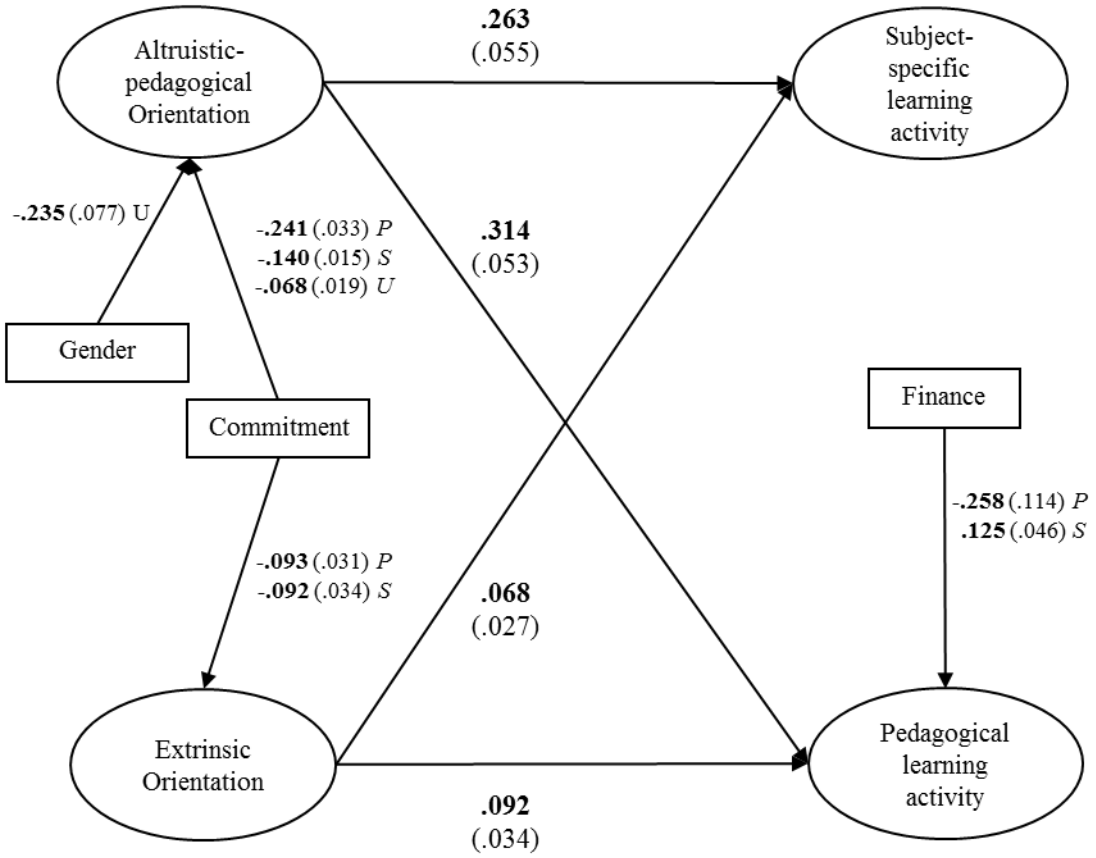

Figure 4. The estimated relations between motivational orientation and learning activities. Fit of the final model: $\chi^2(426) = 674.223$, $p < .001$; four absolute correlation residuals > .10; RMSEA = .039 with 90% CI [.033, .044], CFI = .941, SRMR = .058. R^2 for altruistic-pedagogical and extrinsic motivation, and subject-specific and pedagogical learning activity: Poland (*P*) .094 (.057), .311 (.067), .258 (.087), .162 (.054); Singapore (*S*) .029 (.022), .136 (.023), .294 (.063), .158 (.036); US (*U*) .023 (.036), .112 (.060), .091 (.037), .063 (.023). Unstandardized estimates for the structural coefficients are shown; all estimates significant at $p < .01$ except from finance to pedagogical learning activity ($p < .05$). One single estimate per structural path indicates equivalence across the countries. Standard errors are shown in parentheses. Measurement part and insignificant paths not illustrated.

The results furthermore show gender differences in the altruistic-pedagogical motivational orientation of student teachers in the US and Poland. Female student teachers show a higher altruistic-pedagogical motivational orientation (γ = -.235 (.077) for the US, γ = -.279 (.086) in Poland). Neither family-, money-, job-related hindrances nor GPA have a significant influence on subject-specific or pedagogical learning activities. The only exception are money-related hindrances which have a negative influence in Poland, and a positive influence in Singapore on pedagogical learning activities (γ = -.258 (.114) for Poland, γ = .125 (.046) in Singapore). Hence,

while student teachers without money-related hindrances show a higher use of pedagogical learning activities in Singapore, students with money-related hindrances show a higher use of pedagogical learning activities in Poland. This result might be explained by the pressure financial hindrances exert on students with regard to their completion of studies. For example, Heineck, Kifmann, and Lorenz (2006) showed that tuition fees are related to the duration of studies in Germany. Students experiencing financial pressure might be urged to finish their studies timely in order to earn money. The final model is illustrated in Figure 4.

Table 7. Fit indices for the comparison of mean differences across the countries.

Model	χ^2	df	RMSEA [90% CI]	CFI	SRMR	$\Delta\chi^2$ (df)
Comparison Model						
Partial Scalar	320.96**	209	.037 [.029, .045]	.971	.063	
Constrained Models (Factor Means)						
Altruistic-pedagogical	586.89**	211	.067 [.061, .073]	.904	.097	40.62(2) **
Extrinsic	462.70**	211	.055 [.049, .062]	.936	.080	177.28(2) **
Subject-specific	321.26**	211	.036 [.028, .044]	.972	.063	2.22(2)
Pedagogical	477.34**	211	.056 [.050, .063]	.932	.110	527.25(2) **
Pairwise Comparison Models (USA/SIN)						
Model A	320.96**	209	.037 [.029, .045]	.971	.063	
Model B	561.78**	213	.064 [.058, .071]	.911	.086	64.94(4) **

Note. ** p < .001.

4.5.3 Differences in motivational orientations and learning activities across countries

Up to this point it has been shown that the motivational orientation of student teachers influences their learning activities during initial teacher training. Furthermore, I can tentatively conclude that an altruistic-pedagogical motivational orientation is a better predictor of student teachers' use of learning opportunities than extrinsic motivational orientation. What remains to be clarified is the question if there are differences in the motivational orientations of student teachers across the three countries. In other words, it is investigated how effective the different configurations of selection functions are in selecting student teachers with favourable characteristics.

Table 7 summarizes the results of the cross-country differences in student teacher motivational orientation and learning activities. Compared to Poland, student teachers in the USA and Singapore show a higher altruistic-pedagogical motivational orientation as well as a higher extrinsic motivational orientation. The estimate of the difference in means on altruistic-pedagogical motivational orientation is + .831 (.113) for the US and + .296 (.086) for Singapore. The estimate for the difference in means on extrinsic motivational orientation is + .712 (.135) for the US and + .363 (.063) for Singapore. While there are no differences in subject-specific

learning activities across the countries, student teachers in the US and Singapore show higher pedagogical learning activities compared to Poland (+ .635 (.066) for the USA and + .349 (.052) for Singapore, respectively). All differences are significant at $p < .001$.

For a more differentiated picture I further compared the motivational orientations and learning activities of student teachers between Singapore and the USA. Student teachers in the USA showed higher altruistic-pedagogical and extrinsic motivational orientations than student teachers in Singapore (mean difference + .536 (.073), $p < .001$ and + .349 (.127), $p < .01$ respectively). Furthermore, student teachers showed higher pedagogical learning activities in the US (+ .287 (.044), $p < .001$).

Table 8. Estimated and standardized mean differences in the latent constructs.

	Mean Difference (Standard Error) Standardized Mean Difference			
Latent Construct	Poland (Reference country)	Singapore (compared to Poland)	US (compared to Poland)	US (compared to Singapore)
Altruistic-pedagogical Orientation	.000	+ .296 (.086) .50	+ .831 (.113) 1.43	+ .536 (.073) 1.15
Extrinsic Orientation	.000	+ .363 (.063) 1.05	+ .712 (.135) 2.06	+ .349 (.127) .54
Pedagogical learning activity	.000	+ .349 (.052) .65	+ .635 (.066) 1.18	+ .287 (.044) .72

Note. The first row of each cell contains the estimated mean differences (standard errors in parentheses). The second row of each cell contains the standardized mean differences.

To help interpret the magnitude of the differences in the means of the latent constructs I further calculated standardized mean differences by dividing the differences in factor means by the respective disturbance in the reference country (Thompson & Green, 2013). The standardized mean differences are summarized in Table 8. If it is assumed that the factor means are normally distributed, the standardized mean differences show that the average student in the USA and Singapore would be on the 98th and 85th percentiles among students in Poland with regard to extrinsic motivational orientation. Moreover, the average student in the USA and Singapore would be on the 92th and 69th percentile among students in Poland with regard to altruistic-pedagogical motivational orientation. In case of pedagogical learning activities, the average student in the USA and Singapore would be on the 88th and 74th percentile among students in Poland. Furthermore, the average student in the USA would be among the 71th, 87th, and 76th percentile among students in Singapore on extrinsic and altruistic-pedagogical motivational orientation, and pedagogical learning activities. With the exception of the difference in extrinsic and altruistic-pedagogical orientation across the USA and Poland, which can be considered large effect sizes, the other effect sizes are small to moderate (Thompson & Green, 2013).

The differences in the motivational orientations of the student teachers can be explained by the differences in the level of information of the respective selection function in the USA and Singapore. Both teacher education systems require student teachers to pass an assessment of their suitability for teaching, which includes written applications and interviews focusing on their motivation to teach. And while Singapore has a more comprehensive level of information, it seems that it is more important to have mechanisms implemented which provide information specifically about the suitability of teacher education candidates. The large differences in the extrinsic motivation of the student teachers can be explained by the occupational status of teachers. There seem to be cultural differences in the perception of this status. While Singapore has a career-based system, and thus a highly attractive teaching profession, the USA has a position-based system. However, it seems that this system is considered highly attractive by student teachers in the USA. The differences in the learning activities may be consequences of the motivational orientations of the student teachers. In sum, I tentatively conclude that the selection function of the teacher education system in the USA is most effective in selecting student teachers with favourable motivational orientations, followed by the selection function in Singapore. It seems that the level of information about teacher education candidates and student teachers is important for a provision of necessary information about the suitability for teaching of student teachers.

4.6 Summary of the chapter

In this chapter I tested the model of teacher education as an open system with a focus on the inherent selection problem, i.e. the relation between the selection function and the association of student teachers motivational orientation and their learning activities during initial teacher training. By means of a multigroup analysis of future secondary education teachers of the TEDS-M database I sought to answer the following research questions: what is the relation between student teacher characteristics and their use of learning opportunities? Does the configuration of teacher education's selection function moderate the relation between student teacher characteristics and their use of learning opportunities? Are different configurations of teacher education selection functions associated with differences in the student teacher motivational orientations and their learning activities? The answer to each of these questions can be considered a step towards the identification of student characteristics relevant for their use of learning opportunities, and further for their study success and development of knowledge, attitudes, and beliefs. Each step is a necessary prerequisite for more certainty in selecting teacher education candidates and student teachers, an aspect where research lacks insights and consensus (Schalock et al., 2006).

In line with theory and previous research I found a positive relation between altruistic-pedagogical and extrinsic motivational orientations of student teachers and their subject-specific and pedagogical learning activities during initial teacher training. The strength of the relation differed for the two motivational orientations; I found a stronger association with learning activities for altruistic-pedagogical orientation than for extrinsic motivation. Altruistic-

pedagogically motivated student teachers show higher subject-specific and pedagogical learning activities than extrinsically motivated student teachers. Hence, I am able to confirm Blömeke (2009) who found a high predictive validity of motivational criteria and consequently suggested motivation to be used as a criterion for selection procedures in the context of teacher education. The more refined measure of motivation used in this study allows refining this suggestion. Based on the results of the analysis I suggest that student teachers are selected with respect to their altruistic-pedagogical motivational orientation. The relevance of this motivational orientation is further strengthened by the finding that its relation with learning activities is equivalent across the three countries included in the analysis. It is not moderated by the differential configuration of the selection functions, which furthermore are embedded in diverse cultural settings.

Given the relevance of motivational orientations of student teachers for their use of learning activities, the effectiveness of the selection functions is evaluated with respect to the average motivational orientation of the student teachers in their respective teacher education systems. The question is how successful the selection functions are in selecting teacher education candidates and student teachers with an altruistic-pedagogical motivational orientation. In this context I have to conclude that the selection function of the teacher education system in the USA is most effective because student teachers show the highest altruistic-pedagogical motivational orientation, followed by the selection function of the teacher education system in Singapore. This finding is in line with the theoretical framework and can be explained by the fact that both functions have structural elements implemented which provide information about the motivational orientations of the student teachers. This information is in turn used in order to make adequate admission and selection decisions. It seems that the more comprehensive level of information in Singapore does not give its selection function an advantage with respect to the altruistic-pedagogical motivational orientation. The advantage becomes clearer when one considers the higher attractiveness of the teaching profession in Singapore. The status of teaching is positively related to the size of the candidate pool (Schwille & Dembele, 2007). Although such a contextual condition is desirable, it is more likely to admit student teachers with primarily extrinsic motivational orientation. The lower extrinsic motivational orientation of student teachers in Singapore, compared to the USA, despite having a more attractive teaching profession might be indicative of an adequate adaption of the selection mechanisms to contextual conditions. Thus, when the extrinsic motivational orientation is included in the assessment of the effectiveness of the selection functions, one might interpret the difference in extrinsic motivational orientation between Singapore and the USA as an increased effectiveness of the selection function of the teacher education system in Singapore. The low extrinsic motivational orientation of student teachers in the teacher education system in Poland may be explained by the low attractiveness of the teaching profession in this country.

In sum, the test of the first part of the model of teacher education as an open system illustrates a way to make unobserved characteristics influencing the use of learning opportunities of student teachers observable. With regard to the selection and sorting teacher education candidates and student teachers, I present a way to identify relevant characteristics of student teachers which can be used as criteria for admission and selection decisions. Thus, I covered individual and

organizational aspects of the inherent selection problem. In the next chapter, I move on to the second part of the model of teacher education as an open system. In this part I investigate the relation between the allocation function of teacher education and the degree of positive matching.

CHAPTER 5

TESTING THE MODEL, PART II: *Teacher allocation and positive matching: on the relation between teacher education's allocation function and the non-random allocation of teachers*

5.1 Aim and structure of the chapter

In this chapter I address a specific manifestation of the non-random allocation problem, namely positive matching of teachers and schools in the education system. I investigate different configurations of allocation functions with respect to differences in the degree of positive matching. It describes a kind of teacher distribution where high ability teachers are clustered in schools with higher socioeconomic status. Significant changes in student achievement could be obtained if better teachers can be hired and retained in schools with lower socioeconomic status and lower achieving students (Little & Bartlett, 2010). However, while positive matching of teachers and schools in the education system is sufficiently demonstrated, only few studies relate the policies to attract, select, and retain teachers to the overall distribution of the teacher workforce. Respective research investigates features relevant for the allocation of teachers in isolation and does not connect them directly to teacher distributions (Luschei & Carnoy, 2010). For example, certification is more often related to teacher quality than teacher distributions. Moreover, it is considered a barrier to raise teacher quality which does not predict teacher performance (Angrist & Guryan, 2008; D'Agostino & Powers, 2009). Liu and Johnson (2006) suggest that a school-based hiring process is not sufficient for adequate recruitment of teachers. Balter and Duncombe (2008) investigate recruitment practices in New York and show that a wide variety of practices is in use, and that districts using only few practices hired less qualified teachers. In case of teacher retention, research on probationary periods remains inconclusive (Loeb & Myung, 2010). Hence, research lacks explanations for the development of teacher distributions in general and positive matching in particular. The model of teacher education as an open system provides an opportunity to relate structural features of the allocation function to

positive matching. Hence, it may allow researchers to identify configurations of allocation functions which are associated with lower degrees of positive matching.

Thus, in this chapter I aim at taking first steps towards explanations for the positive matching of teachers and schools. By means of a multilevel multigroup path analysis of PISA 2009 and PISA 2012 data this study seeks to answer the following research questions. *First, what is the relation between characteristics of the teaching staff and the average socioeconomic status of schools? Second, what is the relation between the average socioeconomic status of schools and teacher shortages? Third, are different configurations of teacher education allocation functions associated with differences in the degree of positive matching?* By answering these research questions with a multilevel multigroup path analysis, it is possible to identify differences in the degree of positive matching at two time points across different allocation functions. Hence, one is able to relate the degree of positive matching directly to differences in approaches to assigning teachers to schools.

This chapter is structured as follows. In the next subsection I describe the allocation functions of two countries, namely Finland and Singapore. The description is based on the characterization of the allocation functions outlined in Chapter 3. In the following sections I describe the hypothetical model to be tested as well as the method. The results are presented in the fourth subsection, and the last subsection provides the reader with a summary of the chapter.

5.2 Different configurations of allocation functions – Finland and Singapore

This section describes the structural arrangements of the allocation functions of the lower secondary teacher education systems in Finland and Singapore. The teacher education systems have been chosen due to the differences in the configurations of their allocation functions. Furthermore, with Finland and Singapore I investigate two reference countries which have a high relevance for global educational policy. Thus, it reflects the reference shift of global educational policy from Europe to Asia (Sellar & Lingard, 2013). Information about the different dimensions of the allocation function has been obtained from the TEDS-M policy report, the PISA 2012 report on educational institutions, and a report on high achieving teacher education systems of the Stanford Centre for Opportunity Policy in Education (Ingvarson et al., 2013; OECD, 2013; Darling-Hammond & Rothman, 2011).

In Finland and Singapore graduation from teacher education institutions is sufficient for gaining entry into the teaching profession (OECD, 2012a; Ingvarson et al., 2013). There are neither examinations required to become certified, nor examinations required to become a fully qualified teacher. Moreover, the contents and curricula of teacher education, as well as the graduation requirements, are centrally coordinated in Finland (Sahlberg, 2011). Similarly, in Singapore the National Institute of Education has set a competence framework the teachers have to fulfil in order to successfully graduate from initial teacher training (Choo & Darling-Hammond, 2011). However, there are differences across the two countries with regard to probationary periods. In Finland there are no accountability systems or measures for dismissal, except when there is a violation of the ethical rules of teaching (Sahlberg, 2011). Singapore has

implemented probationary periods which are not used as a certification requirement. There are assessments during this probationary period which are used for the evaluation of teacher performance (Ingvarson et al., 2013). Based on the theoretical framework outlined in chapter 3, I consider the level of information about trained teachers as moderate in Finland and Singapore. It is gradually higher in Singapore due to the probationary periods implemented as another mechanism to gather relevant information about the trained teachers for confirmation decisions (Choo & Darling-Hammond, 2011).

In Finland the principal and school, in coordination with the school board, make the decisions about the recruitment of teachers (Sahlberg, 2011). Consequently, 41% of the schools report a high level of autonomy with respect to teacher hiring (OECD, 2013). At the same time there are many union regulations. Teacher unions play a prominent part in the recruitment of teachers; almost the whole teaching force is unionized (Sahlberg, 2011). The situation is different in Singapore. In this country, teachers are recruited centrally by the Ministry of Education from the top third of each cohort of graduates (Choo & Darling-Hammond, 2011). Only 8% of the schools report a high level of autonomy with respect to teacher hiring (OECD, 2013). There are no union regulations; only 4% of the schools report a union responsibility about teacher hiring. The placement of the teacher is based on the needs of the schools. Two years after the initial placement teachers can request a transfer to a school of their choice. Furthermore, there is a yearly placement in which teachers who requested a transfer are assigned to schools based on their staff requirements (Choo & Darling-Hammond, 2011). Based on the theoretical framework (Chapter 3), I consider the level of control over the recruitment process as high in Finland, and low in Singapore.

Table 9. Configurations of the allocation functions – Finland and Singapore

Country	Level of Information about Trained Teachers	Level of Control over the Recruitment Process	Level of Integration of Trained Teachers into Schools
Finland	Moderate	High	Low
Singapore	Moderate/High	Low	High

With respect to teacher induction and mentoring, Finland has a variable system. The responsibility for induction and mentoring lies within the school (Sahlberg, 2011). Hence, there are some schools which have implemented a sophisticated induction and mentoring program for beginning teachers, whereas other schools have no such measures developed and implemented. Moreover, induction and mentoring activities are in some schools assigned to the principal, and in other schools assigned to experienced teachers (Sahlberg, 2011). The situation is different in Singapore. Beginning teachers are induced and mentored by experienced teachers for two years (Choo & Darling-Hammond, 2011). These experienced teachers are specifically trained by the National Institute of Education. During the induction and mentoring period, which serves as an

extended practicum, beginning teachers take further courses offered by the Ministry of Education (Choo & Darling-Hammond, 2011). Furthermore, they have a lighter workload. Based on the theoretical framework (Chapter 3), I consider the level of integration of trained teachers into schools as low in Finland, and high Singapore. Table 9 summarizes the configurations of the allocation functions of the two teacher education systems.

5.3 The allocation function and the degree of positive matching

The differences in the configurations of the two allocation functions make it possible to investigate if the degree of positive matching varies across the two countries. More specifically, based on the theoretical framework outlined in Chapter 3, I assume that an allocation function with high levels of information about trained teachers, a high level of control over the recruitment process, and a high level of integration of trained teachers into schools minimizes the degree of positive matching. According to the country descriptions in the previous section it becomes clear that neither of the two configurations fulfil the requirements of an optimal allocation function. Given the characteristics of the both functions, I hypothesize that the allocation function of Singapore is associated with a lower degree of positive matching than the function of Finland.

The positive matching problem, one prominent manifestation of the non-random allocation of teachers in the school system, is basically a consequence of the shift from the assumption of school equivalence to the characterization of schools as being unique social systems with very specific characteristics (Johnson & Kardos, 2008). It describes a kind of teacher distribution where high ability teachers are clustered in schools with higher socioeconomic status (Little & Bartlett, 2010). Luschei and Carnoy (2010, p. 180), investigate the teacher distribution in Uruguay and conclude that "[…] teachers with attributes associated with higher student outcomes are systematically concentrated in schools with higher socio-educational context […]". Moreover, effective schools are able to hire teachers with higher abilities (Loeb, Kalogrides, & Beteille, 2012). Students with low socioeconomic status are twice as likely to be assigned to newly trained teachers, compared to students with a higher socioeconomic status (Peske & Haycock, 2006). Furthermore, they are more likely to be taught by uncertified, out of field, or low ability teachers (Darling-Hammond, 2004; Jerald & Ingersoll, 2002; Shen, Mansberger, & Yen, 2004). This is problematic, since newly trained and especially high ability teachers are more likely to quit when teaching low achieving students, given they enter teaching in the first place (Boyd, Lankford, Loeb, & Wyckoff, 2005; Guarino, Santibanez, & Daley, 2006).

Based on the theoretical framework outlined in Chapter 3 and the state of current research I model the positive matching problem as follows. I hypothesize that the quantity of the teaching staff predicts the school average socioeconomic status. This is the first component of positive matching. The respective hypothesis is based on the findings by Luschei and Carnoy (2010), and Little and Bartlett (2010). The second component describes the relation between school average socioeconomic status and the shortage of qualified teachers. I hypothesize that the school average socioeconomic status is negatively related to teacher shortages at a school. This

hypothesis is based on the findings by Peske and Haycock (2006), Shen et al. (2004), and Guarino et al. (2006). The degree of positive matching varies with the strength of the hypothesized relations. The two components of the positive matching problem are illustrated within the dashed box in Figure 5. An optimal configuration of the allocation function governs the allocation of teachers to schools in a way that the quantity of the teaching staff and teacher shortages are independent of the average socioeconomic status of the schools. As mentioned before, I hypothesize that the allocation function of Singapore is associated with a lower degree of positive matching than the function of Finland.

5.4 Method

To test these hypotheses and to answer the research questions a multilevel multigroup path analysis was conducted. This kind of analysis involves three steps: (1) the specification of the hypothesized multilevel model, (2) an evaluation of the plausibility of the model and the level specific model fit, and (3) an investigation of the equivalence of the focal structural paths between the quantity of the teaching staff, school socioeconomic status, and teacher shortages.

5.4.1 Database

I made use of the PISA international databases 2009 and 2012 (OECD, 2012b; OECD, 2013). This international study was conducted in 75 countries in 2009 and 65 countries in 2012, and investigated the performance of 15-year-olds on reading, mathematics, and science. The analyses were based on the PISA 2009 and PISA 2012 dataset versions which were provided by the OECD on December 16, 2012 and November 29, 2013.

5.4.2 Sample

In order to reach a representative sample of 15 year old students, the PISA international study implemented a two-stage stratified sampling design (OECD, 2012b). In the first stage schools having 15 year old students were sampled by a probability proportional to size sampling (PPS). In the second stage 35 students were sampled from each selected school. The original sample of PISA 2009 consisted of 470000 students in 18641 schools and the original sample of PISA 2012 consisted of 510000 students in 18139 schools.

Based on the criteria described in section 5.2 Finland and Singapore were selected as focal countries for this study. The original samples of Finland consisted of 5810 students in 203 schools in PISA 2009, and of 8829 students in 311 schools in PISA 2012. The original samples of Singapore consisted of 5283 students in 171 schools in PISA 2009, and of 5546 students in 172 schools in PISA 2012. After excluding cases with missing values on all variables, the effective sample size of the present study was 4591/3051 students in 156/100 schools in PISA 2009 (Finland/Singapore), and 7477/5297 students in 263/164 schools in PISA 2012 (Finland/Singapore).

5.4.3 Variables

In order to investigate differences in the degree of positive matching across the countries, I used the following variables from the student and school questionnaires of PISA 2009 and PISA 2012. Means, standard deviations, and correlations of the variables can be found in Appendix D.

The *PISA index of economic, social, and cultural status* (ESCS) was used as an indicator of school socioeconomic status (School SES). This index was derived from three indices: home possessions, the highest occupational status of the parents, and the highest educational level of the parents using principal component analysis (OECD, 2012b; OECD, 2013). For the PISA 2009 data, the internal consistency of ESCS was .58 and .68 (Finland and Singapore; OECD, 2011). For the PISA 2012 data, its internal consistency was .61 and .73 (Finland and Singapore).

The *PISA index on teacher shortage* was used as an indicator of the shortage of teachers at a school. This index was derived from four items measuring potential factors hindering instruction at schools rated by the school principals. Higher scores on this index are indicative for a higher shortage of teachers (OECD, 2012b; 2013). For the PISA 2009 data its internal consistency was .52 and .82 (OECD, 2012b), while for the PISA 2012 data its internal consistency was .71 and .76 (Finland and Singapore).

The *quantity of the teaching staff* was used as an indicator for the teaching staff at a given school. This is a simple index calculated in PISA 2009 and PISA 2012 as dividing the number of teachers with an ISCED 5A qualification by the number of total teachers at a school (OECD, 2012b; 2013).

In order to obtain accurate estimates of the regression coefficients between the quantity of the teaching staff, school socioeconomic status, and teacher shortages, I further included the following control variables. For each of the aforementioned variables school size, school type, school competition, and the quality of the school's educational resources were included as controls. The *PISA index of school size* indicates the total enrolment at a school, summing up the total number of boys and girls at a school (OECD, 2012b; 2013). The *PISA index on school type* is a categorical variable with three categories: public, government-dependent private and government-independent schools (OECD, 2012b; 2013). Similarly, the *PISA index on school competition* is a categorical variable with three categories: two or more schools, one other school, and no other school. It indicates the degree to which a school competes with other schools regarding student intake, teaching staff, and other educational resources. The *PISA index on the school's educational resources* was derived from seven items indicating the principal's perception of potential factors hindering instruction at school, related to equipment, computers, and software (OECD, 2012b; 2013). For the PISA 2009 data its internal consistency was .84 and .82 (OECD, 2012b), while for the PISA 2012 data its internal consistency was .80 and .84 (Finland and Singapore).

I further included two variables indicating *school policies with respect to the admission of students*. These variables were derived by a principal component analysis of seven items indicating the reasons why a student is admitted to the respective school. For the PISA 2009 data, the internal consistency of this scale was .64 for Finland and Singapore. For the PISA 2012

data, the internal consistency of this scale was .61 for Finland and .63 for Singapore. The components indicated achievement related or ascriptive reasons (e.g. residence, family endorsement) for admission. Higher factor scores indicated a higher importance of the admission reasons. Additionally, I included three variables indicating *school policies with respect to student transfer*. These variables were derived by a principal component analysis of six items indicating the reasons why a student is transferred to another school. For the PISA 2009 data, the internal consistency of this scale was .61 and .52 (Finland and Singapore). For the PISA 2012 data, the internal consistency of this scale was .68 and .65 (Finland and Singapore). The components indicated achievement related, behavior related (e.g. behavioral problems), or ascriptive reasons (e.g. family requests) for admission. Higher factor scores indicated a higher importance of the reasons for student transfer.

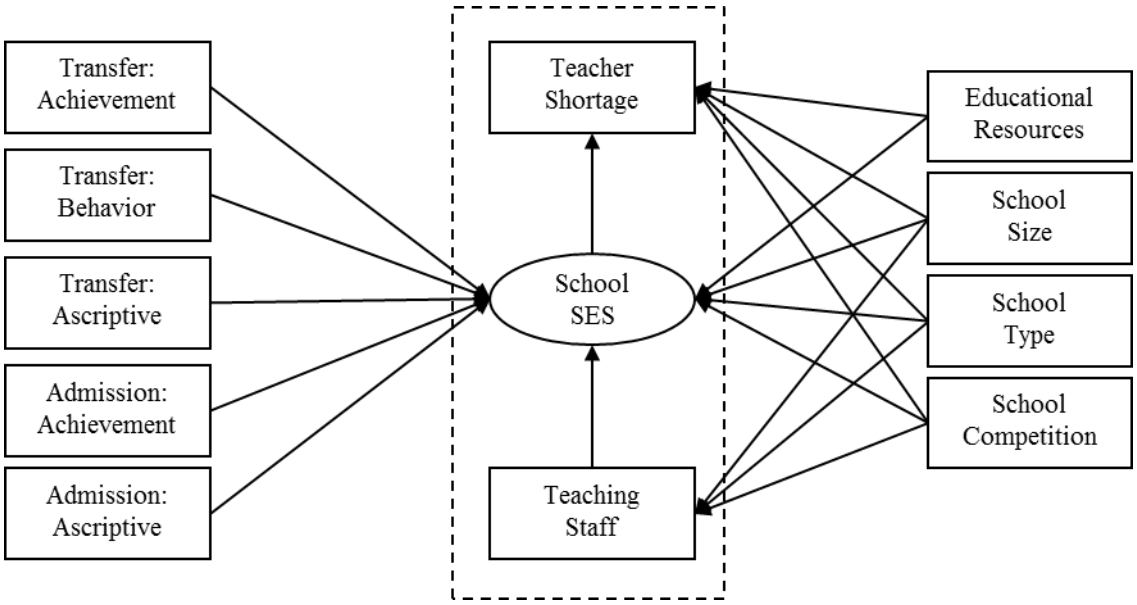

Figure 5. The hypothesized positive matching model; student transfer and admission policies of schools are hypothesized to influence their average socioeconomic status. The schools' educational resources, size, type, and competition with other schools are hypothesized to influence the quantity of the teaching staff, its socioeconomic status, and degree of teacher shortage. These are control variables for the two component relations of the positive matching problem (illustrated in the dashed box): the quantity of the teaching staff at a school predicts its average socioeconomic status (i.e. its intake), which in turn influences the degree of teacher shortage at a school. The individual level of the model is not shown in this picture since it has only one variable (student socioeconomic status).

5.4.4 Analysis

In order to investigate if there are differences in the degree of positive matching across the configurations of the allocation functions of the teacher education systems in Finland and

Singapore a multilevel multigroup approach was taken. The configurations of the allocation functions were included as contextual information characterizing the two groups. The methodological approach taken in this study involved three steps.

The first step consisted of the specification of the multilevel model. The socioeconomic status was specified to operate on the within (student) and between (school) levels of the multilevel model. This is similar to a single level model where the average of the student socioeconomic status is used as a school level variable. The multilevel model has the advantage that it allows the incorporation of student level weights and the consideration of the within school variance of the socioeconomic status. All other variables were specified to operate on the school level given their focus on school characteristics. The intra-class correlation was calculated for socioeconomic status in order to determine if a multilevel model was appropriate. The correlation ranged from .133 to .253 across countries and time points and indicated that this was the case (Heck & Thomas, 2009).

The second step involved the evaluation of the plausibility of the specified relations in the model. More specifically, I evaluated the level specific model fit, in order to assess the plausibility of the specified relations on the school level (following Stapleton, 2013). Therefore, I firstly estimated models were all possible covariances among the variables at the school level were fixed to zero (the independence models; *ind*). Next, I estimated models where the hypothesized relations were specified on school level as shown in Figure 5 (the partially saturated models; *ps*). The χ^2-values and the degrees of freedom of the independence and partially saturated models were then used to calculate the level specific CFI and RMSEA of the school level model as follows (*J* = number of schools; Ryu & West, 2009; Stapleton, 2013):

$$\boldsymbol{CFI}school-level = 1 - \frac{\max[\boldsymbol{\chi 2}ps - \boldsymbol{df}ps, 0]}{\max[\boldsymbol{\chi 2}ind - \boldsymbol{df}ind, 0]}$$

$$\boldsymbol{RMSEA}school-level = \sqrt{\frac{\boldsymbol{\chi 2}ps - \boldsymbol{df}ps}{\boldsymbol{df}ps(\boldsymbol{J})}}$$

The level specific SRMR is given by Mplus as default. In order to determine if the hypothesized model was plausible I used the fit indices and associated cut-off criteria described in chapter 4.4.4. The models described in this section were estimated for each country separately with the 2009 and 2012 data (the χ^2-values and the degrees of freedom of the independence models can be found in Appendix D).

The third step of the analysis provided the answers to the research questions. After making sure that the hypothesized model is plausible in each of the two countries, it was estimated simultaneously in both countries. I first estimated a model where the structural paths between the focal constructs were freely estimated. Next, I estimated a model where these paths were constrained to be equal across the countries. The unconstrained and constrained models were then compared regarding their χ^2-values (Little, 1997). A significant change in χ^2-values implied

that there was at least one path different across the countries. In order to identify which path was different I conducted Wald tests for the null hypothesis that the focal path is equal across the countries (Wang & Wang, 2012). A significant Wald test indicates that the path under investigation is not equal across the countries but different, and thus is moderated by group membership (i.e. the configuration of the allocation function). Lastly, I estimated the final models where the paths were constrained according to the results of the Wald tests. The unconstrained, constrained, and final models were estimated for 2009 and 2012, respectively.

All analyses were carried out in Mplus 6.12 (Muthén & Muthén, 1998-2013). To account for non-normality of the data, I used the restricted Maximum Likelihood estimator (MLR). In order to obtain robust parameter estimates and standard errors, student and school level weights were incorporated to account for the complex sampling design and the associated unequal selection probabilities of students and schools.

5.5 Results

This section is structured as follows. First, the results of the evaluation of the level specific model fit and the final model of the year 2009 are presented. Second, the evaluation of the level specific model fit and the final model of the year 2012 are presented. Thus, I illustrate the situation regarding the relation between the allocation function and positive matching in Finland and Singapore for these years. The last section contains robustness checks of the final model of 2012.

5.5.1 Differences in the degree of positive matching in 2009

Evaluation of the level specific model fit indicated that the theoretical school level model was plausible for both countries. It has to be noted that in order to establish the good model fit (see Table 10) I had to specify correlations between the quality of the school's educational resources and the quantity of the teaching staff, as well as between teacher shortages and the quantity of the teaching staff. However, given that a common omitted cause for these variables is reasonable to assume, I deemed both modifications as in line with the hypothesized model.

The fit of the multigroup models were good as well. The χ^2- tests were all insignificant, and the level specific RMSEA, CFI, and SRMR were all above or below the respective cut-off values (see Table 10). This indicated that the theoretical school level model was plausible as well, and that an interpretation of the parameter estimates was warranted.

Table 10. Fit indices of the partially saturated and comparison models 2009.

Model	χ^2	df	RMSEA_b	CFI_b	SRMR_b	$\Delta\chi^2$ (df)
Partially saturated models						
Finland	23.621	18	.044	.966	.039	
Singapore	20.652	19	.029	.991	.052	
Comparison (multigroup) models						
Unconstrained model	41.798	36	.041	.986	.045	
Constrained model	51.938	38	.037	.968	.050	17.169 (2)***
Final model	47.496	37	.033	.969	.048	10.253 (1)***

Note. RMSEA, CFI, and SRMR are specific to the school-level part of the model.
*** $p < .001$

With respect to the control variables the results showed that achievement related admission policies positively influenced the school average socioeconomic status in Singapore. Schools reporting that they have implemented these admission policies showed a higher average socioeconomic status. School size had a significant influence on teacher shortages in Finland. An increase in school size leads to an increase in teacher shortages. School competition had a significant influence on school average socioeconomic status and teacher shortages in Finland. A higher competition between schools regarding students and resources was associated with an increase in school average socioeconomic status. At the same time it led to a decrease in teacher shortages. This means that schools with no competition with other schools have a disadvantage over schools with one and two or more competing schools when it comes to socioeconomic status and teacher shortages. The quality of the schools' educational resources had a significant influence on school average socioeconomic status and teacher shortages in Singapore. Higher qualities of these resources led to an increase in school average socioeconomic status and a decrease in teacher shortages (see Figure 6).

All the significant relations between the control variables and school average socioeconomic status and teacher shortages had the expected direction. It is interesting, however, that student transfer policies had no influence on either school average socioeconomic status or teacher shortages in each of the two countries.

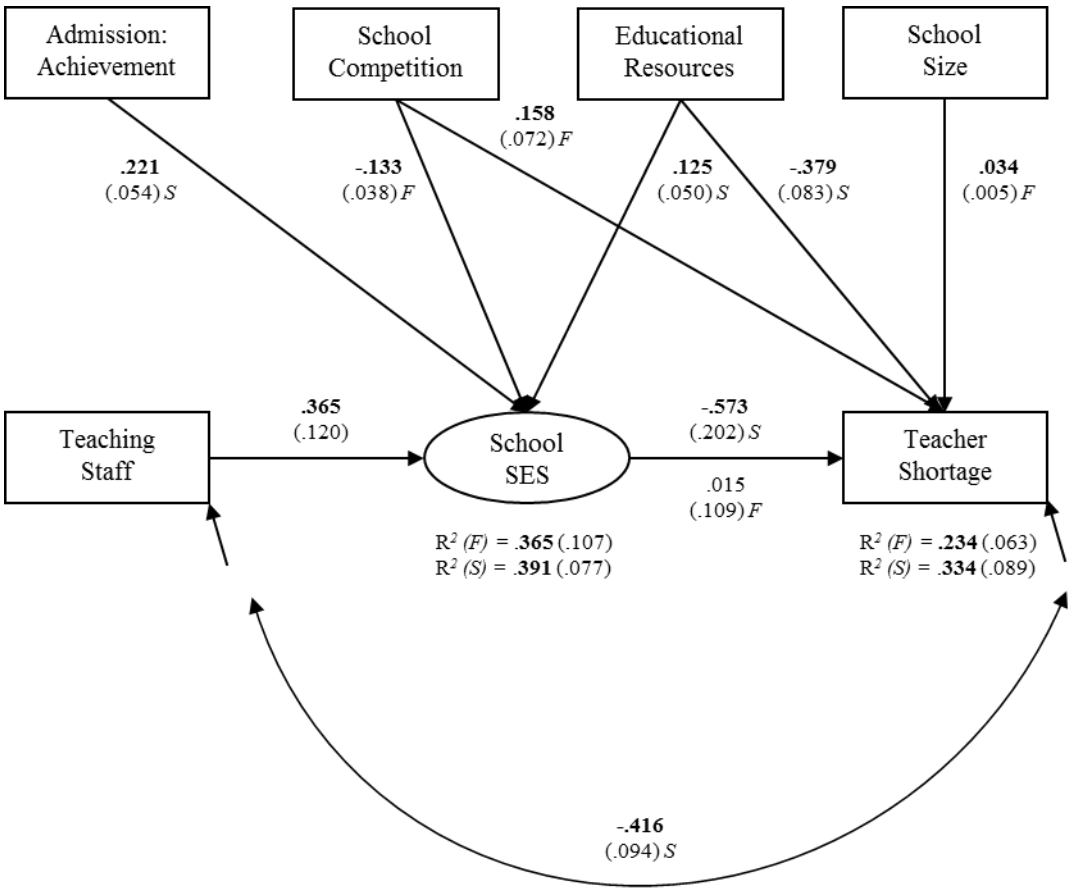

Figure 6. The final 2009 positive matching model. Model fit: $\chi^2(37) = 47.496$, p = .12, RMSEA = .033, CFI = .969, SRMR = .048. Unstandardized estimates are shown (standard errors in parentheses); all estimates significant at p < .01 except from school competition to teacher shortage (p < .05). A single estimate per structural path without a letter behind it indicates equivalence across the countries. A single estimate per structural path with a letter behind it indicates non-equivalence across the countries, and an insignificant estimate in the omitted country. *F* = Finland; *S* = Singapore. For better legibility the insignificant paths are omitted from the figure.

Regarding differences in the degree of positive matching across the two countries, the results showed an interesting pattern. The quantity of the teaching staff of the school significantly predicted the school average socioeconomic status in both countries (Wald test statistic 2.783 with df = 1, p = .096). An increase of one percent in the proportion of teachers with an ISCED 5A qualification led to an increase of .365 (.120) in the school average socioeconomic status in Finland and Singapore. Thus, I conclude that the contribution of this component to the degree of positive matching was equivalent across the two countries. The situation is different with respect to the second component of positive matching. The school average socioeconomic status significantly predicted teacher shortages only in Singapore (Wald test statistic 3.623 with df = 1,

p = .05). The lower the average school socioeconomic status was the higher was the schools shortages in qualified teachers. The differences between the two countries indicated that the configuration of the allocation function moderated the second component of positive matching. In sum, I have to conclude that in 2009 neither of the two configurations of allocation functions in Finland and Singapore governed the allocation of teachers to schools in a way that the quantity of the teaching staff is independent of the school average socioeconomic status.

5.5.2 Differences in the degree of positive matching in 2012

In order to gain a more comprehensive insight into the relation between the allocation function and positive matching of teachers and schools in the education system, I further investigated the situation in 2012. Given the insignificant relation between student transfer policies and the components of positive matching I dropped these policies from the 2012 model. The increase in the explained variance of school average socioeconomic status (see Figure 7) as well as the evaluation of the level specific model fit justified this decision. It indicated that the theoretical school level model was plausible for both countries in 2012 as well. In order to establish the good model fit of the partially saturated models, no modifications were necessary (see Table 11). The fit of the multigroup models were good as well. The χ^2- tests were all insignificant, and the level specific RMSEA, CFI, and SRMR were all above or below the respective cut-off values. This indicated that the theoretical school level model was plausible in 2012 as well, and that an interpretation of the parameter estimates was warranted.

Similarly to 2009, the results showed that achievement related admission policies positively influenced the school average socioeconomic status in Singapore. Schools reporting that they have implemented these admission policies showed, similarly to 2009, a higher average socioeconomic status. School size had a significant influence on the quantity of the teaching staff and school average socioeconomic status in Finland. An increasing school size was associated with an increase in the quantity of the school staff, and a decrease in the school average socioeconomic status. Similarly to 2009, school competition had a significant influence on school average socioeconomic status in Finland. An increased competition between schools regarding students and resources increased the school average socioeconomic status. This is comparable to the situation in 2009. The quality of the schools' educational resources had a significant influence on teacher shortages in Singapore. Higher quality of these resources led to a decrease in teacher shortages. Contrary to 2009, school type significantly influenced all components of positive matching in 2012. Public schools had a higher quantity of the teaching staff in Singapore. Furthermore, public schools had a higher average socioeconomic status than government-dependent and government-independent schools in Finland and Singapore. Lastly, while public schools had higher teacher shortages in Finland, they experienced lower teacher shortages in Singapore, compared to government-dependent and government-independent private schools (see Figure 7).

Table 11. Fit indices of the partially saturated and comparison models 2012.

Model	χ^2	df	RMSEA_b	CFI_b	SRMR_b	$\Delta\chi^2$ (df)
Partially saturated models						
Finland	9.292	6	.045	.975	.044	
Singapore	5.250	6	.027	1.000	.026	
Comparison (multigroup) models						
Unconstrained model	16.769	12	.036	.982	.036	
Constrained model	24.435	14	.041	.962	.043	7.7916 (2)*
Final model	17.025	13	.026	.985	.037	.4181 (1)

Note. RMSEA, CFI, and SRMR are specific to the school-level part of the model. * $p < .05$

Regarding differences in the degree of positive matching across the two countries in 2012, the results again showed an interesting pattern. Contrary to 2009, in 2012 the quantity of the teaching staff of the school significantly predicted the school average socioeconomic status in Singapore, but not in Finland (Wald test statistic 20.297 with df = 1, p = .000). An increase of one percent in the proportion of teachers with an ISCED 5A qualification led to an increase in the score for the school average socioeconomic status of 3.053 (.680) in Singapore. Thus, the differences between the two countries indicated that the configuration of the allocation function moderated the first component of positive matching. Compared to 2009, the situation in 2012 is different with respect to the second component of positive matching. Here, teacher shortages were independent of school average socioeconomic status in both countries (Wald test statistic .432 with df = 1, p = .51). The average school socioeconomic status did not predict shortages in qualified teachers in both countries. In sum, I have to conclude that in 2012 only the configuration of the allocation function in Finland governed the allocation of teachers to schools in a way that the quantity of the teaching staff is completely independent of the school average socioeconomic status. Thus, while the situation in 2012 is comparable to the situation in 2009 for Singapore, Finland manages to provide an effective allocation of teachers to schools. Except for a shift from the second to the first component of positive matching, the overall picture of the relation between the allocation function and the degree of positive matching remained the same in Singapore.

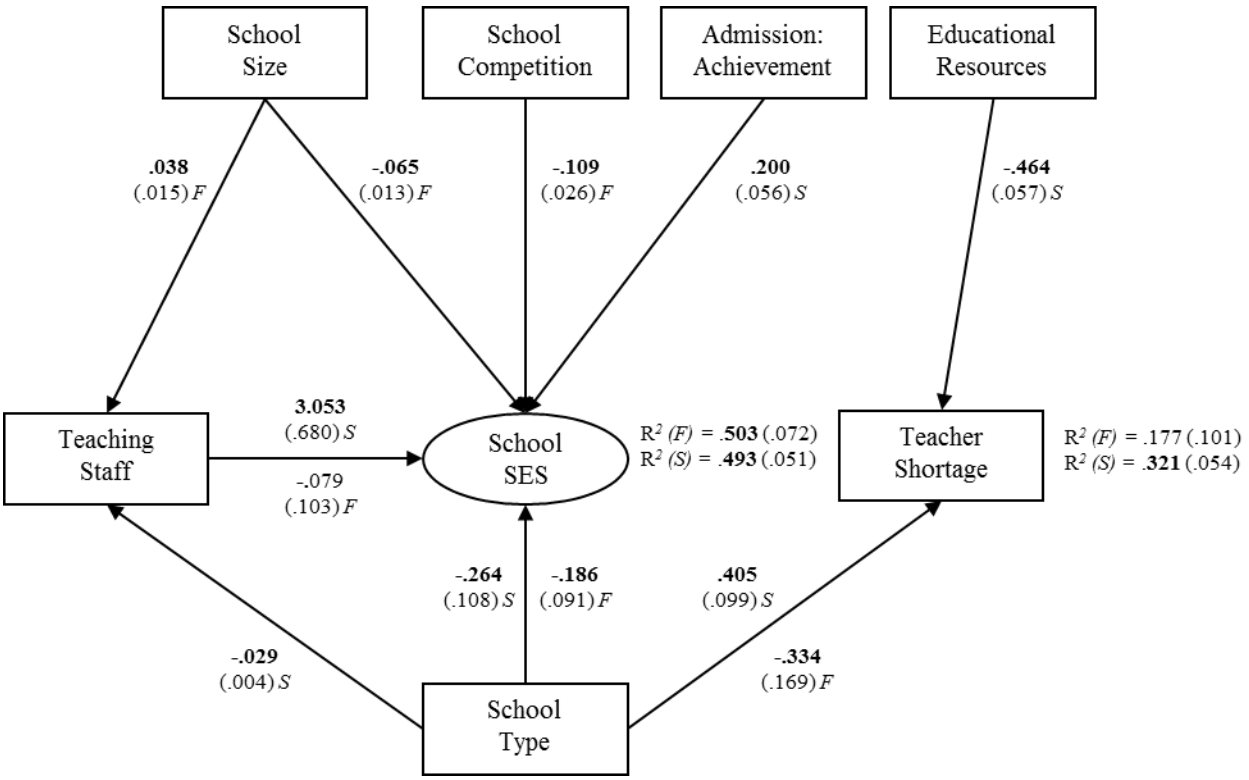

Figure 7. The final 2012 positive matching model. Model fit: $\chi^2(13) = 17.025$, p = .19, RMSEA = .026, CFI = .985, SRMR = .037. Unstandardized estimates are shown (standard errors in parentheses); all structural paths significant at p < .01 except from school type to school SES and teacher shortage (p < .05). A single estimate per structural path without a letter behind it indicates equivalence across the countries. A single estimate per structural path with a letter behind it indicates non-equivalence across the countries, and an insignificant estimate in the omitted country. *F* = Finland; *S* = Singapore. For better legibility the insignificant paths are omitted from the figure.

5.5.3 Robustness of the final model

Since dropping student transfer policies from the 2012 model was based on empirical reasons (insignificant influences of all student transfer policies on the quantity of the teaching staff, school average socioeconomic status, and teacher shortages), I investigated the plausibility of the hypothesized relations of this model in further countries. For this robustness check I randomly selected three more countries, namely Chile, Canada, and the Netherlands. The 2012 model showed good fit to the data in all of these countries (see Table 12). It has to be noted that in order to establish the good fit in Chile and the Netherlands I had to specify a direct effect from the quality of the school's educational resources on the quantity of the teaching staff in the Chilean model, and a covariance between teacher shortages and the quantity of the teaching staff in the Netherlands model. However, given that a relation and a common omitted cause for these

variables is reasonable to assume, I deemed both modifications as in line with the theoretical model. Moreover, the amount of explained variance in the teacher shortages and the school socioeconomic status was moderate to good (R^2 = .31 and .62 for Chile, R^2 = .12 and .59 for the Netherlands; and R^2 = .16 and .38 for Canada). In sum, the results of the robustness check of the final 2012 model showed that the hypothesized relations were plausible not only in Finland and Singapore, but in three other cultural contexts as well.

Table 12. Fit indices of the partially saturated models of the robustness check.

Model	χ^2	df	RMSEA_b	CFI_b	SRMR_b	$\Delta\chi^2$ (df)
Partially saturated models						
Chile	6.308	5	.036	.996	.020	
Canada	8.317	6	.030	.989	.028	
Netherlands	3.726	5	.053	1.000	.034	

Note. RMSEA, CFI, and SRMR are specific to the school-level part of the model.

5.6 Summary of the chapter

In this chapter I tested the model of teacher education as an open system with a focus on the non-random allocation of teachers to schools in the education system, i.e. the relation between the allocation function and the degree of positive matching. By means of a multilevel multigroup path analysis of the PISA 2009 and 2012 databases I sought to answer the following research questions: what is the relation between characteristics of the teaching staff and the average socioeconomic status of schools? What is the relation between the average socioeconomic status of schools and teacher shortages? Are different configurations of teacher education allocation functions associated with differences in the degree of positive matching? With the answers to each of these questions I identified the degree of positive matching in two countries at two different time points, and could show that there are differences in the degree of positive matching associated with different approaches to assigning teachers to schools.

More specifically, I could show that the relation between the quantity of the teaching staff, school average socioeconomic status, and teacher shortages was more developed in 2009. Schools with a higher average socioeconomic status were associated with a higher quantity of the teaching staff in both countries, and with lower teacher shortages in Singapore. In 2012, teacher shortages were independent of the school average socioeconomic status in both countries, but schools with a higher socioeconomic status were associated with more qualified teachers in Singapore. In Finland, school socioeconomic status was independent of the quantity of the teaching staff. Based on the overall results regarding the relation between quantity of the teaching staff, school average socioeconomic status, and teacher shortages, the results confirmed the hypotheses. Furthermore, I confirmed a finding by Akiba et al. (2007) stating that even countries with a high mean level of student achievement do not provide equal access to teachers.

The overall degree of positive matching decreased from 2009 to 2012 in both countries. However, with regard to the effectiveness of the different configurations of the allocation functions, the respective hypothesis was not confirmed. I assumed that the allocation function of the teacher education system in Singapore would be associated with a lower degree of positive matching. This hypothesis was based on its level of information about trained teachers, level of control over the recruitment process, and the level of integration of teachers into schools, which was closer to the theoretical optimal configuration outlined in Chapter 3. However, based on the relational pattern of the components of positive matching I have to conclude that the allocation function of Finland is more effective in allocating teachers independent of school average socioeconomic status. In 2009, the quantity of the teaching staff was related to school average socioeconomic status, whereas teacher shortages were not. In 2012, school average socioeconomic status was completely independent from the teaching staff and teacher shortages.

This result is in so far surprising as the allocation function of the teacher education system in Finland is from a theoretical point of view less optimal configured in terms of information and integration than the allocation function in Singapore. In this regard, the level of control over the recruitment process, although deemed low in a theoretical sense, has to be reconsidered as high because the centrally governed initial placement of teachers to schools is based on the schools' staffing needs (Choo & Darling-Hammond, 2011). But despite this characteristic of the allocation process the degree of positive matching is higher compared to Finland, where control over recruitment is fully decentralized. However, I already mentioned that in Singapore teachers are able to request a transfer to another school two years after their initial placement. Thus, the individual decision making of teachers in the context of the allocation functions needs to be considered. From an individual point of view, the configurations can be understood as structural backgrounds within which decisions for or against a specific school are made. It might be possible that the allocation function of the teacher education system in Singapore provides a degree of freedom in these transfer decisions for teachers, which possibly overrides the results of the initial placement of teachers to schools. Teacher mobility including entry decisions, decisions for specific teaching positions, and transfer decisions after initial hiring seems to be a promising element to be added to the model of teacher education as an open system. Moreover, from a policy perspective, further studies could shed light on the question if structured teacher mobility can counterbalance the possible negative effect of individual decision making on the equality of teacher distributions.

When I consider the level of integration of teachers into schools in combination with the level of control over the recruitment process the results become even more surprising. Induction and mentoring phases and measures are means to make the teacher acquainted to the specific situation at a given school. School and teacher can use this phase in order to collect information about their fit to each other. The possibility of teachers requesting transfers to another school may be conceived of as a means to dissolve teacher-school matches without an adequate fit between teacher and school. Hence, the centrally governed placement of teachers to schools, combined with demand-based transfer decisions might imply an equal distribution of teachers in the education system. However, this is not the case. It might be that teachers request their

transfers to other schools ‘for the wrong reasons’, that is, not because they do not fit in the school, but because they do not like to work with lower achieving students or students with a lower socioeconomic status. The decentralized allocation process within a less defined and systematic induction and mentoring system seems to be more effective in this regard. However, Liu and Johnson (2006) state that a school-based recruitment is not enough. And indeed, in Finland there are incentives for teachers to work in more rural or other disadvantaged schools (Sahlberg, 2011). Thus, despite not having a structured system of induction and mentoring, Finland implements other measures which have their effects in places. Moreover, because recruitment is school-based in Finland it might be that individual preferences regarding recruitment are already considered in the initial placement of teachers to schools, because the teachers apply directly to their school of choice (Eurydice, 2013). The individual decision making of teachers in the allocation process might then not override or alter the initial distribution of teachers. However, these explanations are speculative. In order to determine the role of individual decision making of teachers in the relation between the allocation function and the degree of positive matching, the model of teacher education as an open system needs to be extended by an individual level.

Of the three dimensions of the allocation function the level of information seems to play only a minor role. However, the differences in this dimension are only gradual across the two countries. In each of the two countries there are no other certification requirements than graduation from a teacher education institution. The relative unimportance of the level of information for both allocation functions might be explained by the standardization of the teacher education systems of both countries. As already mentioned, in Singapore the Ministry of Education sets standards which have to be met by each student teacher in order to be eligible to teach (Choo & Darling-Hammond, 2011). Thus, it might be that the variance in teacher knowledge and skills is low, and although the allocation of teachers is governed centrally, the schools do not need additional information about the knowledge and skills of the teachers. The probationary period might also provide sufficient information about the performance of a teacher. Finland implemented an integrative approach to initial teacher training where field experiences in training schools are a prominent feature (Sahlberg, 2011). Additionally, in the school-based recruitment process it is possible to acquire sufficient information about the specific knowledge and skills of the teachers in order to make informed recruitment decisions. Thus, further signals such as certification or licensure are not required, especially since these measures are considered only noisy signals for the knowledge and skills of teachers (Staiger & Rockoff, 2010). However, it has to be noted that certification systems might provide necessary information in other configurations of allocation functions. In other words, I do not propose to generalize their low importance in the contexts of Singapore and Finland to the allocation functions of other countries. It might be assumed that in countries with multiple ways to enter the teaching profession, that is, with various alternative certification opportunities, schools depend on signals provided by certification systems. In this regard Singapore and Finland have very homogenous teacher education systems (Sahlberg, 2011; Choo & Darling-Hammond, 2011).

Another aspect to consider is the school average socioeconomic status itself. It constitutes of the socioeconomic status of the students at a school, and depends in part of parental school choice. With the quality of the school's educational resources and school competition I included factors potentially influencing parental school choice. Although the explained variance in school average socioeconomic status is relatively high in both of the models, it is still reasonable to assume that there are important factors omitted, and that this might bias the results. The limitations in scope of the PISA school and student datasets did not allow including further control variables related to parental school choice. Including such factors is important for reliable estimates of the relation between the allocation function and positive matching, because parental school choice is no direct element of the allocation function of teacher education systems.

In sum the test of the second part of the model of teacher education as an open system illustrates a way to investigate the relation between teacher education's allocation function and the non-random allocation of teachers to schools in the education system. The results indicate that an allocation function whose configuration differs from the theoretical optimal configuration is still effective (a) if its elements are aligned to each other, and (b) if it adequately adapts to the teacher education system and the specific characteristics of the school system. Hence, the effectiveness of the allocation function is more a question of its contextual alignment than of a mere comparison with a theoretical optimum. Such a comparison would be further constrained by the difficulty of finding a 'baseline value' of an allocation function without any structural elements governing the allocation process. I will further elaborate on this thought in the next chapter, the general discussion.

CHAPTER 6

GENERAL DISCUSSION AND REFLECTIONS

The overarching aim of this thesis was to reach a better understanding of teacher education policy and practice, with a specific focus on the selection and sorting of teacher education candidates and student teachers, and the allocation of teachers to schools in the education system. With this focus a connection between student teachers, organizational features of teacher education, teachers, and contextual conditions of teacher education was established. In Chapter 2 it was shown that research on the relation between teacher education and student achievement does not take into account these connections and interrelations. Thus, despite its relevance for policy makers, this kind of research lacks a perspective enabling researchers to connect the aforementioned aspects and thus facilitating explanations of teacher education policy and practice (Zeichner, 2006). In Chapter 3 the first step was taken and an alternative, organizational perspective on teacher education was developed. The core of this development was the organizational model of teacher education as a system of coherent and complementary components. The model addresses the embedding of teacher education in multiple contexts, as well as the connection between student teachers and organizational or structural features of the teacher education system. Both aspects are considered important for new insights and fruitful policy recommendations by various authors (Zeichner, 2006; Barber & Mourshed, 2007; Zeichner & Conklin, 2008; Grossman & McDonald, 2008; Darling-Hammond & Rothman, 2011). The model of teacher education as an open system was the basis for the second step consisting of investigations of the inherent selection and non-random allocation problems; both problems are challenges in teacher education research but lack explanations (Harris & Sass, 2011; Little & Bartlett, 2010). The respective parts of the model were tested. In Chapter 4 structural features of teacher education, namely the structural arrangements governing the selection and sorting of student teachers, were connected to the learning of student teachers and to contextual conditions in the education system and the teaching profession. In Chapter 5

structural features of teacher education were related to the allocation of teachers to schools in the education system, thus connecting teacher education to its context as well. Both studies can be considered building blocks for a better understanding of selection and allocation practices and for explanations of the inherent selection and non-random allocation problem.

The aim of this last chapter is to provide a comprehensive discussion of the main findings of the studies conducted in this thesis. It illustrates how their results can be interpreted in light of the model of teacher education as an open system, and how the results contribute to a better understanding of teacher education policy and practice. This discussion further critically addresses limitations of the methodological approaches taken in the studies. Based on both results and limitations it outlines areas for further research and implications for teacher education policy and practice. A brief conclusion summarizes the overall value of this thesis.

6.1 Main findings

The insights provided by this thesis are related to three overarching aspects. First, it takes up the suggestion by Darling-Hammond and Rothman (2011) to consider teacher education as a system of interrelated components, that is, complementary policies aimed at the development of student teachers and teachers in general. With Open Systems Theory the model was based on an adequate theoretical fundament and thus offers a connection between theory and research. It goes beyond existing descriptive structural models which lack such a theoretical basis (for example the model by Wang et al., 2003). The characterization of teacher education as an open system explicitly includes two important aspects: the relation between system and student teachers, as well as the connection between system and context. These characteristics are addressed in the explication of the role of the selection and allocation functions for teacher education effectiveness.

The investigation of the role of the selection function in the context of a teacher education system is the second aspect. It illustrates how organizational features of teacher education have an impact on the learning and instruction of student teachers. Connecting the selection function to the relation between characteristics of student teachers and their use of learning opportunities not only addresses procedural and structural aspects of the inherent selection problem. It can further be considered as a way to disentangle what Kennedy (1998) called "enrollment influences" and "learning influences": the development of knowledge, attitudes, and beliefs not only depend on characteristics of the teacher education system, but also on characteristics student teachers bring into the system (Zeichner & Conklin, 2005). The test of the respective part of the model showed that these characteristics influence their learning, and that the structural arrangement of the selection function in turn affects what characteristics student teachers bring into the teacher education system.

The third aspect is the investigation of the role of the allocation function in the context of the teacher education system. The test of the respective part of the model illustrated how the allocation function has an impact on "[…] the ultimate goal of ensuring that each school in each jurisdiction is filled with highly effective teachers […]" (Darling-Hammond & Rothman, 2011, p. 2). It showed that different approaches to allocating teachers to schools entail different degrees

of positive matching, i.e. different distributions of teachers in the education system. Hence, the effectiveness of a teacher education system with regard to the development of teachers may be diminished by an unequal distribution of teachers. Furthermore, the allocation function is an example not only for the fact that the teacher education system depends on its context, but also that the teacher education system has an impact on its context. In other words, the connection between system and its context is bidirectional.

6.1.1 The model: teacher education as an open system

Zeichner (2005) argues that for new insights into teacher education practice a connection between theory and research is necessary. Modeling teacher education as an open system provides research with such a connection. This model establishes the relation between student teachers, organizational characteristics of teacher education, and contextual conditions of teacher education. Especially the distinction between system and individual, or else, teacher education and teacher characteristics offers a means to identify what makes some teacher education systems more effective than others. Individualized conceptualizations of teacher education, which are used in value-added models investigating the relation between teacher education and student achievement in the education system, lack such means (Zeichner, 2006; Boyd et al., 2009).

Barber and Mourshed (2007) and Darling-Hammond and Rothman (2011), among others, showed that countries with high performing education systems rely on well developed teacher education systems. These systems are furthermore embedded in a highly attractive teaching profession (Schwille & Dembele, 2007). The model of teacher education as an open system developed in this thesis allows explaining and refining what is meant by a ‘well developed teacher education system’, because it illustrates the potential interrelations between the different components. Viewing teacher education a complementary system of multiple components rather than an individual teacher attributes further raises the awareness of the scope of policy changes. When changing certain teacher education practices it may be necessary to consider the impact of these changes on other parts of the system (Darling-Hammond & Rothman, 2011). The model may also be used by policy makers and researchers for an evaluation of the developmental status of teacher education systems in other countries. Eventually, typologies of teacher education systems can be developed according to their developmental status, i.e. to the degree to which the different elements of the system are complementary and adapted to country-specific contextual conditions. These advantages of the model were also pointed out by the experts in the interview study.

When talking about the overarching use of the model for research and policy, it has to be noted that the model is not fully specified yet. Given the complexity of the relation between student teachers, organizational and contextual features of teacher education, teacher behavior, and student achievement (Zeichner, 2006; Konold et al., 2008), the focus of the model was on the selection and allocation functions of teacher education systems. Both functions are considered important aspects in the aforementioned relation (Harris & Sass, 2011). The experts involved in the interview study indicated that both functions, as well as the contextual conditions of the

teacher education system, are sufficiently and validly characterized in terms of structural elements and dimensions. Thus, it can be argued that the model in its current state is a valid starting point for a further elaboration of the model in future studies.

A particularly important aspect of the model to be extended is the interplay between system and student teacher. The learning opportunities are structured in ways that either facilitate or constrain individual development. At the same time, the individual development serves as feedback and input for the development of the system. More specifically, the model can be extended to include characteristics of the learning opportunities provided to student teachers, for example their realism and authenticity. It may be argued that teacher education systems with authentic and realistic learning opportunities, e.g. frequent classroom practice, teaching hours, and guidance by experienced teachers, are better aligned with the requirements of the teaching profession and thus are able to produce highly qualified and effective teachers. This may include changes in the way student teachers perceive teaching, the teaching profession, and the nature of the subjects they are going to teach (Morge et al., 2010). In other words, besides knowledge building, individual development entails conceptual change.

Moreover, monitoring the performance of student teachers in such authentic learning environments may facilitate the identification of personal characteristics required for effective teaching behavior. This information can then in turn be used for refining admission and assessment procedures in order to make valid selection decisions. It may render the distinction between personal characteristics relevant for study success and personal characteristics relevant for effective teaching obsolete (Blömeke, 2009), because the requirements of teacher education reflect more closely the requirements of the teaching profession. This illustrates the aforementioned interplay of individual and organizational development.

6.1.2 The impact and interplay of the selection and allocation functions

First and foremost the two studies investigating the relations between the two functions and their immediate outcomes illustrate how the model of teacher education as an open system can be translated into a methodologically viable and feasible model. This further illustrates the usefulness of the model for further research.

The relation between motivational characteristics of student teachers and their use of learning opportunities depicts the connection between student teachers and organizational or structural features of teacher education. The results show that individual development, that is, their learning depends on the characteristics they bring into teacher education. Moreover, which characteristics they bring into teacher education depends on the structural arrangement of the selection function of the teacher education system. It can be argued that development of student teachers consists of a combination of enrollment and learning influences (Kennedy, 1998; Zeichner & Conklin, 2005). However, for a more refined characterization of the learning influences the model needs to be extended with regard to the learning opportunities, as mentioned in the previous section, and by inclusion of performance measures of student teachers.

From an organizational perspective the relevance of motivational characteristics for the use of learning opportunities has important implications for the admission of teacher education candidates into the teacher education system. First, building admission procedures on criteria including altruistic-pedagogical motivational aspects increases the amount and the validity of the information provided by the procedures. Hence, the teacher education system may be more likely to avoid higher amounts of ex-ante training costs, because candidates without respective characteristics are not admitted into initial teacher training. Second, with the development of more refined admission procedures teacher education systems are more likely to avoid training costs during initial training of student teachers, because the admitted student teachers are more likely to finish their studies and less likely to drop out prior to graduation. Third, with more student teachers graduating from initial teacher training there are more potential new teachers available for distribution into the education system, thus increasing the likeliness that the teacher education system can provide a sufficient number of teachers available for allocation into the education system.

The relation between the allocation function and the distribution of teachers in the education system depicts the connection between the teacher education system and its context. With a relatively detailed characterization of the structural arrangements of the allocation functions of the teacher education systems in Finland and Singapore it was found that the differences in these configurations were associated with the degree of positive matching in the education systems of these countries. This study is among the first which relates not only single elements, for example certification of teachers, to teacher distributions, but configurations of elements relevant for the allocation of teachers. And although it was not possible to relate the elements of these configurations directly to the degree of positive matching, the differential relevance of these elements can be discussed. However, this discussion has to consider the interplay between the selection and allocation functions, as well as their relation to the quality and equity of the education system.

The high performance of the education systems of Finland and Singapore in 2009 and 2012 might be in part explained by the results of the studies. Taking up the perspective on teacher education as a system of interrelated parts, it can be seen that in both countries the teacher education system is embedded in a highly attractive teaching profession (Sahlberg, 2011; Choo & Darling-Hammond, 2011). Next, both teacher education systems have a focus on a selection at the entry into the teacher education system. This focus is necessary because of the large pool of teacher education candidates (as shown in Chapter 4). In combination with a highly standardized teacher education system, both systems are more likely to produce a sufficient number of highly qualified teachers which are available for allocation into the education system; training costs of student teachers are low. Thus, the effectiveness of the selection function leads to an increased effectiveness of the teacher education system. The allocation function connects to this effectiveness and distributes the trained teachers in case of Finland completely, and in case of Singapore partly independent from the socioeconomic status of the schools. In light of the low attrition rates in both countries (3% in Singapore and 1% in Finland) the teachers are able to contribute to their students' learning in a way that results in a high performance of the students

(Boyd et al., 2012). Thus, from an organizational perspective it can be concluded that especially in Finland the distribution of teachers does not override the effectiveness of the teacher education system. The absence of a certification system at entry into the teaching profession carries no weight for the amount of information available to the schools about the aptitude of the teachers because of the high standardization and standard of the teacher education system.

In sum, the results of both studies and underlines the importance of a coordination of both functions with teacher education and its context. The pattern of interdependencies and related outcomes which emerges from these studies is further in line with the main propositions of the theoretical model of teacher education as an open system. Thus, the model of teacher education as an open system provides a better understanding of teacher education policy and practice than the currently dominant individualized approach utilized by most studies. The methodological limitations of the studies, policy implications, and directions for further research will be discussed in the next sections.

6.2 Methodological limitations

In order to assess the scope and generalizability of the results there are three overarching methodological limitations to consider. First, as indicated in the previous section, the subject-centred approach of including the configurations of the selection and allocation functions as contextual information about group membership constrains the assessment of the relevance of single structural elements of the functions. Such identification requires a possibility to relate the structural elements directly to the learning of student teachers or the degree of positive matching. The limited availability of data prevented such a direct approach of investigating the impact of both functions. However, the indirect approach taken in this thesis allowed investigating the coordination of the functions with teacher education and its immediate context, the education system and the teacher labor market. As was explicated in the previous section, the coordination between and configuration of structural elements, functions, and context seem to be the more important aspect of teacher education policy and practice than the effect of a single, isolated structural element.

Additionally, in light of the dependency of the effects of the selection and allocation function on country specific aspects of the teacher education and education system, it may be argued that a generalizability of the results is neither warranted nor necessary. This addresses the lack of a baseline function without any structural elements, as well as configurations that do not include all structural elements as theoretically specified. It is important to note that there might be no single best way to select or allocate student teachers and newly trained teachers. As indicated by the concept of equifinality, several configurations of different structural elements of the functions might be leading to similar effects. It is not likely that teacher education systems in different countries consist of exactly the same or all theoretically specified structural elements. The primary question is how exactly a given configuration is coordinated with contextual conditions of the education system and the teacher labor market. Hence, the results are not to be transferred to other countries, even if they have similar configurations of their selection and allocation

functions, without a thorough consideration and analysis of teacher education and its context in a given country. In sum, in light of the context-dependencies of the function as well as the limited data availability it is argued that the multiple group approach to the cross-country comparisons taken in this thesis offer an adequate balance between generalizability and scope.

Second, limitations with regard to data availability are the reason for the discrepancies between the operationalization of the structural elements described in Chapter 3 and the characterization of the functions in the following chapters. For example, the empirical operationalization or description of internal support, in case of the selection function, and teacher mentoring, in case of the allocation function, are not congruent with their theoretical operationalization. The pragmatism of the initial operationalization of the structural elements was also mentioned by the experts in the interview study. The experts stated that there was a high probability that the eventual operationalization of the structural elements might not reflect the initial operationalization; however, they did not suggest that this was a problematic aspect. It might even be argued that with the more qualitative descriptions of the selection and allocation functions used in the two empirical studies it was possible to draw a more detailed picture of both functions, compared with the initial operationalization of their structural elements.

Third, data availability had also consequences for the quality of the data used for the two studies testing the theoretical model of teacher education as an open system. For a more detailed investigation of the relation between individual and organizational development a longitudinal approach might be required, which is based on a richer set of variables. The data provided by TEDS-M, however, do not allow for such a longitudinal investigation because this study was based on a cross-sectional design. Moreover, although steps were taken to include a number of relevant variables, the results might be blurred due to omitted variable bias. This is also reflected by the low amount of explained variance in the respective latent constructs of the model (from a minimum of 9% to a maximum of 31 % across constructs and countries). The model specification can be explained by the limited scope of the available TEDS-M data. Furthermore, although achievement data would have been available it was not included in the models, because data regarding the content and pedagogical content knowledge of the student teachers was not comparable across the three countries. This was due to a reduced coverage of the population in Poland and the USA (Tatto et al., 2012); in the former country, only concurrent programs were included and in the latter only public institutions were covered.

With regard to data quality, the PISA databases are limited as well. While the hypothesized model tested in the context of the allocation function uses a richer set of variables than the model tested in the context of the selection function, the reliability of some measures is low. Especially the internal consistency of the PISA index of economic, social, and cultural status and the PISA index on teacher shortages are low for Finland in PISA 2009 and 2012, and for Singapore in PISA 2009. Moreover, internal consistencies are consistently low for the indices on school admission and student transfer policies. The low internal consistencies may explain the insignificant results with respect to these variables, and therefore may be a justification for dropping them from the 2012 model.

Nevertheless, in case of both tests steps were taken to carefully select an adequate database, and further to include variables depicting the core aspects of the relation between characteristics of student teachers and their use of learning opportunities, and depicting the core aspects of positive matching. These limitations can be conceived of as fruitful starting points for further research.

6.3 Implications for policy and practice

The results of the studies and its methodological limitations entail the following implications for policy and practice. These implications concern the general approach to developing and implementing policy changes, as well as specific recommendations for the selection and sorting of teacher education candidates and student teachers and the allocation of teachers to schools in the education system.

The model of teacher education as an open system developed in this thesis points out the dependencies of the different components of the teacher education system, as well as its dependence on its context. Policy makers have to be aware that changes in one part of the system may have unintended impacts on other parts of the system. The development and implementation of reforms may be more likely to succeed if their potential impacts on other parts of the teacher education system are considered in advance. Darling-Hammond and Rothman (2011) outline a promising approach to developing and implementing policy changes. These changes may firstly address the most urgent need, and secondly address the complementary elements of the system which are affected by the change.

Moreover, the success of any teacher education reform is a question of its adaptation to country-specific characteristics of the teacher education system, the education system, and the teacher labor market. What works in Finland may not quite work in Germany, what does not work in Japan might work in the USA. Hence, respective reforms should be built around these country-specific characteristics, and not be based on international benchmarks and comparisons. The ‘comparative turn’ in its current state, where policy makers look at top performing and reference countries and selectively adapt certain practices to their own country, might be misleading (Bulle, 2011; Grek, 2009; Paine & Zeichner, 2012). Policy makers may be more likely to succeed in developing and implementing teacher education reforms if they withstand the pressure of the OECD and its international assessments to inform policy and practice by other countries’ performances (Paine & Zeichner, 2012).

The results of the investigation of the relation between selection function and the learning of student teachers points out a specific recommendation for teacher education policy. As mentioned before, an important characteristic of well developed teacher education systems is their embeddedness in an attractive teaching profession (Barber & Mourshed, 2007; Sahlberg, 2011; Darling-Hammond & Rothman, 2011). The recommendation concerns the problem of the role of admission procedures at entry into initial teacher training when there are teacher shortages. Research suggests that admission procedures and criteria adapt to the size of the candidate pool (Blömeke, 2006). While some authors argue that in case of teacher shortages admission

requirements should be attenuated or removed completely, other authors state that an increased attractiveness of teaching is the appropriate answer. Rothstein (2012) showed that changing the quality of the teaching force through selection is only successful if teacher evaluation systems and increased teacher salaries are introduced simultaneously.

Based on the results of this study it is recommended to address the attractiveness and status of the teaching profession in order to increase the candidate pool, from which the teacher education system is able to select its student teachers. While removing admission procedures may increase the pool of candidates and the number of student teachers, it is more likely that training costs increase for the teacher education system. Moreover, it may be possible that the variance in teacher knowledge might further increase, which then requires specific certification procedures at the entry into the teaching profession, which in turn have to be based on standards set by a central authority in order to make sure that each teacher meets some minimum qualification. It can be seen that removing admission procedures completely entails a number of further policy changes required to, in a sense, compensate for the removal. Addressing the size of the candidate pool with changes in the attractiveness of teaching might be the more fruitful way to react to teacher shortages. At the same time this policy change takes into account the connection between teacher education and its context.

The results regarding the relation between the allocation function and the degree of positive matching point out the role of the decision making of teachers during and after the recruitment process. The crucial aspect seems to be a way to avoid that individual decision making of teachers after the initial placement overrides the quality of the initial match between teacher and school, as provided by the allocation function. Such forms of structured mobility require additional monitoring procedures which assess the quality of the initial matches with regard to the degree of positive matching, as well as with respect to the fit of teachers and schools.

Probationary periods may be implemented or extended to include such monitoring procedures which provide additional information about the performance of a given teacher at a given school. After a designated period of time the quality of the match between teacher and school can be reassessed, and decisions about transferring the teacher to another school can be made. Just as there are teacher-school combinations which are more effective than others (Jackson, 2010), this delayed form of allocation further takes into account the possibility that there are teacher-school combinations which are ineffective. In other words, transfer decisions of teachers may not always override the quality of the initial matches, but also improve the distribution of teachers with respect to their equity. The value of probationary periods as a monitoring device may further be independent of the general approach to teacher allocation being centralized or decentralized, because in either approach information is provided directly to the schools.

Connecting the results of both studies with results of other studies another tentative recommendation can be made. This recommendation concerns the interplay between the selection and allocation function. As could be shown the success of the allocation function in distributing teachers independent of schools' socioeconomic status depends on the effectiveness of the teacher education system in producing sufficient numbers of teachers. Given that this effectiveness depends on enrolment and learning influences the focus of the selection function is

on providing as much valid and reliable information about the aptitude of student teachers as possible. Depending on the standardization of the teacher education system, the need for the allocation function to provide further information about the trained teachers is less pronounced. The focus of the allocation function is then on the specific way of distributing the teachers to the schools in the education system. In other words, the allocation function does not need further standardization or selection measures such as certification systems, which may also function as a barrier for teachers preventing them from entering the teaching profession (Angrist & Guryan, 2008). The amount of information available to schools for recruitment can also be influenced by the specific level of control over the recruitment process, as well as by probationary periods serving as measures for the integration of teachers into schools. Hence, it is recommended to make a clear distinction of the purpose of both functions while simultaneously taking into account their contextual dependencies.

6.4 Directions for further research

The results of the studies and their limitations further offer various directions for future research. These can be broadly categorized as follows: (a) extensions of the model; (b) investigating interrelations of the functions not covered in this thesis; and (c) further investigations that include different designs or methodological approaches.

The model can be extended by an individual level containing individual characteristics and behavior of student teachers, trained teachers, and teacher educators. Teacher educators and other staff working in a teacher education system are the individuals who convey the learning opportunities to the student teachers, and who implement policy changes (Hökkä & Eteläpelto, 2013). Thus, a fruitful extension of the model of teacher education as an open system may be a distinction between structural components, members of the system (teacher educators and other staff), and student teachers. Moreover, individual decision making during the recruitment process may be a necessary step in order to reach a better understanding of the development of manifestations of unequal teacher distributions such as positive matching. More specifically, this extension is required in order to investigate the role individual decision making plays in overriding the initial matches between teachers and schools provided by the arrangement of the structural elements of the allocation function.

Another extension of the allocation part of the model is related to the specification of the positive matching problem. As already mentioned, the average socioeconomic status of a school constitutes itself primarily of the socioeconomic status of its students. Parental school choice plays a very prominent role in this regard. An inclusion of factors influencing parental school choice might be required to get a more complete picture of the average socioeconomic status of a school, and further to reach a better understanding of the development of positive matching.

Furthermore, the model of teacher education as an open system is relatively underspecified with regard to characteristics of the learning opportunities. In order to be able to better explain the development of relevant knowledge, attitudes and beliefs it may be necessary to model the interplay between student teachers and learning opportunities in more detail. With a connection

of this interplay with performance related outcomes it is furthermore possible to better explain the differential effectiveness of different teacher education systems and teacher education programs. Moreover, with the characterization of the selection function provided in this thesis, it may further be possible to adequately clarify the inherent selection problem and the role of the selection function in competence development of student teachers.

Further investigating the coordination and interplay between the selection and allocation functions, and their relation with the quality of the education system is a promising approach to identify the impact of teacher education on the quality of education systems. An interesting question, which was tentatively answered in the previous sections, is how the coordination of both functions leads to better student achievement in the education system. Moreover, it may be possible to identify complementary configurations of selection and allocation functions which lead to better student achievement in the education system. The aforementioned further specification of the opportunities to learn of the teacher education system may lead to more detailed explanations of the connection between teacher education effectiveness and student achievement. It has to be noted that this area of research requires a much more detailed and refined database than is currently available.

A longitudinal approach to the investigation of the relation between the selection function and the use of learning opportunities by student teachers might further enhance the meaningfulness of the results provided by this thesis. It is clear the model tested in the respective study is relatively underspecified with regard to student teacher characteristics other than the motivational orientation. Similarly, the model is relatively underspecified with regard to variables capturing the learning opportunities provided by the teacher education system. A longitudinal investigation of a more fully specified model may facilitate the identification of student teacher characteristics which not only are predictive of study success, but also are predictive of success in teaching.

The initial operationalization of the structural elements of both functions as described in Chapter 3 can be developed further in order to construct composite indices of the different dimensions of the selection and allocation functions. This enables researchers to investigate the relation of the configurations of both functions with teacher education effectiveness and positive matching in a more direct way, because the composite scores may be used as instrumental variables in multilevel models. For example, the relation between relevant student (background) characteristics and achievement constitutes the lowest level. The second level comprises the classroom/teacher and their characteristics. Teacher characteristics are hypothesized to influence their behavior in the classroom; the influence of teacher characteristics on student achievement is mediated by this behavior. This is modeled as cross-level interactions. Teacher education variables can be included in two different ways: either on (1) teacher level as antecedents of teacher characteristics, or (2) as a property of the teacher education system (contextual variables at country level in the case of cross-country comparisons). In the first case, teacher education variables are included on level two, as direct antecedents of teacher characteristics (e.g. knowledge, attitudes, and beliefs), and indirect antecedents of teacher behavior. In the second case, teacher education is conceptualized on school level, cross-classified with other relevant school characteristics. Selection, qualification, and allocation aspects of the teacher education

system influence the teacher characteristics, modeled by cross-level interactions. This direct approach also facilitates longitudinal investigations of the relation between the selection and allocation functions and teacher education effectiveness and positive matching. However, data currently available may not be sufficient for these kinds of analyses.

The multiple group approach taken in this study can be further used in order to test the propositions and relations for different levels of teacher education. Thus it is possible to investigate differential effects of teacher education variables across different educational levels, as discussed by Huang and Moon (2009). For example, the importance of obtaining a degree for student achievement differs across elementary, middle, and high school levels (Phillips, 2010; Harris & Sass, 2011). The differential relevance is explained by the generalist/specialist distinction between elementary and secondary teacher education. Different educational levels or institutional settings can easily be integrated into the structural equation modeling approach by specifying the multilevel model for each educational level (either elementary, middle, or high school level or elementary and secondary level). The structural relationships can then be compared across levels or settings. Any difference in coefficients across the groups informs us about the differential relevance of teacher education aspects. Another potential use of the theoretical model of teacher education as an open system is that it can be used as a frame to investigate teacher education programs as well as single teacher education institutions.

Lastly, the results discussed in the previous sections suggest that the investigation of teacher education policy and practice is more a question of coordination and adaption to contextual conditions than a question of comparing certain configurations of functions with a theoretical optimal configuration. Hence, any future studies may formulate respective hypotheses on the basis of the complementarity of the teacher education elements under investigation, and of their coordination with contextual conditions present in the education system and the teacher labor market. The focus of this approach is on the question if a certain configuration works under specific contextual conditions, and if certain changes in configurations may work under specific contextual conditions. This takes into account the argument of the previous section which stated that in light of the context dependency of teacher education practice, generalizability may not be of the importance usually attributed to meaningful results.

6.5 Conclusion

Eventually, there remains one question to be answered: what is the value added by this thesis to the field of teacher education research, policy, and practice? The model of teacher education as an open system addresses a gap in current research, both in a theoretical and empirical sense. The model offers a theoretical basis in which researchers can integrate their studies. It can be conceived of the connection between theory and research Zeichner (2005) claimed to be necessary for a better understanding of teacher education practice. With its focus on the relation between student teachers, organizational features of teacher education, and characteristics of its context the model establishes a connection between important elements; it furthermore offers a

theoretical foundation of the claim that it is necessary to view teacher education as a system of interrelated components (Darling-Hammond & Rothman, 2011).

The studies conducted in this thesis further confirmed that view. Teacher education effectiveness is not the consequence of single policies or structural features, but a result of the interplay between context, system, and individual student teacher. The empirical value of the model and the studies can further be enhanced by the various directions for future studies derived from this thesis. The model may also be used by policy makers all around the world in order to carefully develop and implement changes to teacher education systems. The studies in this thesis show that respective changes need to be designed around country-specific contextual conditions. They serve as a critique of the current practice of adapting policies working in reference countries in order to retain legitimacy of their own teacher education systems.

Paine and Zeichner (2012) identified a convergence of teacher education practice around the globe. However, convergence without an adequate theoretical foundation is most likely to fail. In light of the ‘comparative turn’ in teacher education policy and practice, the theoretical model of teacher education as an open system might provide not only researchers, but also policy makers with such a foundation. It enables them to compare their teacher education system with others based on more relevant aspects than performance, namely coordination and coherence of the different elements of a teacher education system and its context.

References

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

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. Educational Researcher, 36, 369-387.

Akyüz, G., & Berberoglu, G. (2010). Teacher and classroom characteristics and their relations to mathematics achievement of the students in the TIMSS. *New Horizons in Education, 58,* 77-95.

Andersson, C., Johansson, P., & Waldenström, N. (2011). Do you want your child to have a certified teachers? *Economics of Education Review, 30,* 65-78.

Andrew, M. (1990). Differences between graduates of four-year and five-year teacher preparation programs. *Journal of Teacher Education, 41*, 45-51.

Angrist, J., & Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review, 27,* 483-503.

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

Balter, D., & Duncombe, W. (2008). Recruiting highly qualified teachers. Do district recruitment practices matter? *Public Finance Review, 36,* 33-62.

Barber, M., & Mourshed, M. (2007). *How the world's best-performing school systems came out on top (The McKinsey Report).* McKinsey & Company. http://www.mckinsey.com/clientservice/socialsector/resources/pdf/Worlds_School_Systems_Final.pdf (Accessed 28/07/10)

Baron, J.N., Hannan, M.T., & Burton, M.D. (1999). Building the iron cage: determinants of managerial intensity in the early years of organizations. *American Sociological Review, 64,* 527-547.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47,* 133–180.

Becker, R., & Hecken, A. (2009). Higher education or vocational training? An empirical test of the rational action model of educational choices suggested by Breen and Goldthorpe and Esser. *Acta Sociologica, 52,* 25-45.

Beese, J., & Liang, X. (2010). Do resources matter? PISA science achievement comparisons between students in the United States, Canada and Finland. *Improving Schools, 13,* 266-279.

Blömeke, S. (2006). Struktur der Lehrerausbildung im internationalen Vergleich. Ergebnisse einer Untersuchung zu acht Ländern. *Zeitschrift für Pädagogik, 52,* 393-416.

Blömeke, S. (2009). Ausbildungs- und Berufserfolg im Lehramtsstudium im Vergleich zum Diplom-Studium – Zu prognostischen Validität kognitiver und psycho-motivationaler Auswahlkriterien. *Zeitschrift für Erziehungswissenschaft, 12,* 82-110.

Blömeke, S., & Paine, L. (2008). Getting the fish out of the water: considering benefits and problems of doing research on teacher education at an international level. *Teaching and Teacher Education, 24,* 2027-2037.

Blömeke, S., Felbrich, A., Müller, C., Kaiser, G., & Lehmann, R. (2008). Effectiveness of teacher education. State of research, measurement issues and consequences for future studies. *ZDM – The international Journal on Mathematics Education, 40,* 719-734.

Blömeke, S., Kaiser, G., & Lehmann, H. (2010). *TEDS-M 2008: Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich.* Weinheim: Beltz.

Blömeke, S., Suhl, U., & Kaiser, G. (2011). Teacher education effectiveness: Quality and equity of future primary teachers' mathematics and mathematics pedagogical content knowledge. *Journal of Teacher Education, 62,* 154-171.

Blömeke, S., Suhl, U., Kaiser, G., & Döhrmann, M. (2012). Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries. *Teaching & Teacher Education, 28,* 44-55.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review, 95,* 166-171.

Boyd, D.J., Lankford, H., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis and Management, 27,* 793-818.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31,* 416-440.

Boyd, D., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2012). Recruiting effective math teachers: Evidence from New York City. *American Educational Research Journal, 49,* 1008-1047.

Braxton, J., Hirschy, A., & McClendon, S. (2004). *Understanding and reducing college student departure. ASHE-ERIC Higher Education Research Report Series*. San Francisco: Jossey-Bass.

Breen, R., & Goldthorpe, J. (1997). Explaining educational differentials. Towards a formal rational action theory. *Rationality and Society, 9,* 275-305.

Brese, F. (2012). *TEDS-M 2008 user guide for the international database – Supplement 1*. Hamburg: IEA.

Bressoux, P., Kramarz, F., & Prost, C. (2009). Teachers' training, class size, and student outcomes: Learning from administrative forecasting mistakes. *Economic Journal, 119,* 540-561.

Buddin, R., & Zamarro, G. (2009). Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics, 66,* 103-115.

Bulle, N. (2011). Comparing OECD educational models through the prism of PISA. *Comparative Education, 47,* 503-511.

Casey, C., & Childs, R. (2007). Teacher education program admission criteria and what beginning teachers need to know to be successful teachers. *Canadian Journal of Educational Administration and Policy, 67,* 1-24.

Carlisle, J. F., Correnti, R., Phelps, G., & Zeng, J. (2009). Exploration of the contribution of teachers' knowledge about reading to their students' improvement in reading. *Reading and Writing: An Interdisciplinary Journal, 22,* 457-486.

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464-504.

Choo, T., & Darling-Hammond, L. (2011). Creating effective teachers and leaders in Singapore. In L. Darling-Hammond & R. Rothman (eds.), *Teacher and leader effectiveness in high performing education systems* (pp. 33-42). Washington, DC: Alliance for Excellent Education.

Clotfelder, C.T., Ladd, H.F., & Vigdor, J.L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources, 41,* 778-820.

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

Clotfelder, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school. A cross-subject analysis with student fixed effects. *Journal of Human Resources, 45*, 655-681.

Connor, C.M., Son, S.H., Hindman, A.H., & Morrison, F.J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology, 43,* 343-375.

Croninger, R.G., Rice, J.K., Rahbun, A., & Nishio, M. (2007). Teacher qualifications and early learning: effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review, 26,* 312-324.

D'Agostino, J., & Powers, S. (2009). Predicting teacher performance with test scores and grade point average: a meta-analysis. *American Educational Research Journal, 46,* 146-182.

Darling-Hammond, L. (2004). Inequality and the right to learn: access to qualified teachers in California's public schools. *Teachers College Record, 106,* 1936-1966.

Darling-Hammond, L., & Rothman, R. (2011). Lessons learned from Finland, Ontario, and Singapore. In L. Darling-Hammond & R. Rothman (eds.), *Teacher and leader effectiveness in high performing education systems* (pp. 1-12). Washington, DC: Alliance for Excellent Education.

Denzler, S., & Wolter, S. (2009). Sorting into teacher education: how the institutional setting matters. *Cambridge Journal of Education, 39,* 423-441.

Desimone, L., & Long, D. (2010). Teacher effects and the achievement gap: Do teacher and teaching quality influence the achievement gap between black and white and high- and low-SES-students in the early grades? *Teachers College Record, 112,* 3024 – 3073.

Doeringer, P.B., & Piore, M.J. (1971). *Internal labor markets and manpower analysis.* Heath: Lexington Books.

Esser, H. (1999). *Soziologie: Spezielle Grundlagen. Band 1: Situationslogik und Handeln*. Frankfurt am Main: Campus

EU (2007). *Relationships between teacher education institutions and schools.* Brussels: European Commission.

EU (2009). *Teacher education curricula in the EU.* Jyväskylä: Finnish Institute for Educational Research.

Eurydice (2009). *Key data on education in Europe 2009.* Brussels: EACEA.

Eurydice (2012). *Key data on education in Europe 2012*. Brussels: EACEA.

Eurydice (2013). *Key data on teachers and school leaders*. Brussels: EACEA.

Feiman-Nemser, S., & Remillard, J. (1996). Perspectives on learning to teach. In F. Murray (ed.), *The teacher educator's handbook* (pp. 63-91). San Francisco: Jossey-Bass.

Fuchs, T., & Woessmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics, 32,* 433-464.

Gansemer-Topf, A., & Schuh, J. (2006). Institutional selectivity and institutional expenditures. *Research in Higher Education, 47,* 613-142.

Gansle, K. A., Noell, G. H., & Burns, J. M. (2012). Do student achievement outcomes differ across teacher preparation programs? An analysis of teacher education in Louisiana. *Journal of Teacher Education, 63,* 304-317.

Glebbeek, A., Nieuwenhuysen, W., & Schakelaar, R. (1989). The labour market position of dutch sociologists: An investigation guided by a theoretical model. *Journal of Social Sciences, 25,* 57-74.

Goe, L., & Strickler, L. (2008). *Teacher quality and student achievement: Making the most of recent research*. Washington: National Comprehensive Center for Teacher Quality.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources, 42,* 765-794.

Goldhaber, D. (2011). Licensure: Exploring the value of this gateway to the teacher workforce. In E.A. Hanushek, S. Machin, & L. Woessmann (eds.), *Handbook of the economics of education (Volume 3)* (pp. 315-339). San Diego: North Holland.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66,* 361–396.

Grek, S. (2009). Governing by numbers: the PISA effect in Europe, *Journal of Education Policy, 24,* 23-37.

Grek, S., Lawn, M., Lingard, B., & Varjo, J. (2009). North by northwest: quality assurance and evaluation processes in European education. Journal of Educational Policy, 24, 121-133.

Grodsky, E., & Jackson, E. (2009). Social stratification in higher education. *Teachers College Record, 111,* 2347-2384.

Grossman, P., & McDonald, M (2008). Back to the future: directions for research in teaching and teacher education. *American Educational Research Journal, 45,* 184-205.

Guarino, C., Santibanez, L., & Daley, G. (2006). Teacher recruitment and retention: a review of the recent empirical literature. *Review of Educational Research, 76,* 173-208.

Guo, Y., Connor, C.M., Yang, Y.Y., Roehrig, A.D., & Morrison, F.J. (2012). The effects of teacher qualification, teacher self-efficacy, and classroom practices on fifth graders' literacy outcomes. *Elementary School Journal, 113,* 3-24.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in the public schools. *Journal of Economic Literature, 24,* 1141–1178.

Hanushek, E.A. (2011). The economic value of high teacher quality. *Economics of Education Review, 30,* 466-479.

Harris, D.N., & Sass, T.R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95,* 798-812.

Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Heck, R.H. (2007). Examining the relationship between teacher quality as an organizational property of schools and students' achievement and growth rates. *Educational Administration Quarterly, 43,* 399-432.

Heck, R. H., & Thomas, S. L. (2009). *An introduction to multilevel modeling techniques* (2nd ed.). New York: Routledge.

Heineck, M., Kifmann, M., & Lorenz, N. (2006). A duration analysis of the effects of tuition fees for long-term students in Germany. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik), 226,* 82-109.

Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Kallmeyer.

Hofman, R.H., Hofman, W.H.A., & Gray, J.M. (2008). Comparing key dimensions of schooling: towards a typology of European school systems. *Comparative Education, 44,* 93-110.

Hogrebe, M.C., & Tate, W.F. (2010). School composition and context factors that moderate and predict 10th-grade science proficiency. *Teachers College Record, 112,* 1096-1136.

Hopkins, D. (2008). *A teacher's guide to classroom research.* Maidenhead: McGraw-Hill.

Hökkä, P., & Eteläpelto, A. (2013). Seeking new perspectives on the development of teacher education: A study of the Finnish context. *Journal of Teacher Education, 65,* 39-52.

Horn, D. (2009). Age of selection counts: A cross-country analysis of educational institutions. *Educational Research and Evaluation, 15,* 343-366.

Hu, L.-T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Huang, F.L., & Moon, T.R. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment Evaluation and Accountability, 21,* 209-234.

Ingersoll, R.M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research, 81,* 201-233.

Ingvarson, L., Tatto, M., Peck, R., Schwille, J., Rowley, G., & Senk, S. (2013). *An analysis of teacher education context, structure, and quality-assurance arrangements in TEDS-M countries*. Amsterdam: IEA.

Jackson, C. K. (2010). *Match quality, worker productivity, and worker mobility: Direct evidence from teachers*. NBER Working Paper 15990. http://www.nber.org/papers/w15990. (Accessed 13/05/12)

Jepsen, C. (2005). Teacher characteristics and student achievement: Evidence from teacher surveys. *Journal of Urban Economics, 57,* 302-319.

Jerald, C., & Ingersoll, R. (2002). *All talk, no action: putting an end to out-of-field teaching*. Washington, DC: The Education Trust.

Johnson, S.M., & Kardos, S.M. (2008). The next generation of teachers: who enters, who stays, and why. In M. Cochran-Smith, S. Feiman-Nemser & D.J. McIntyre (eds.), *Handbook of research on teacher education* (pp. 445-467). New York: Routledge.

Kane, T.J., Rockoff, J.E., & Staiger, D.O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27,* 615-631.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations.* New York: Wiley.

Kaya, S., & Rice, D. (2010). Multilevel effects of student and classroom factors on elementary science achievement in five countries. *International Journal of Science Education, 32,* 1337-1363.

Kennedy, M. (1998). *Learning to teach writing: Does teacher education make a difference?* New York: Teachers College Press.

Kline, R.B. (2013). Assessing statistical aspects of test fairness with structural equation modelling. *Educational Research and Evaluation, 19,* 204-222.

Konold, T., Jablonski, B., Nottingham, A., Kessler, L., Byrd, S., Imig, S., Berry, R., & McNergney, R. (2008). Adding value to public schools – Investigating teacher education, teaching, and pupil learning. *Journal of Teacher Education, 59,* 300-312.

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review, 28,* 49-57.

Lankford, H., & Wyckoff, J. (2010). Teacher labor markets: an overview. In D.J. Brewer & P.J. McEwan (eds.), *Economics of Education* (pp. 235-242). London: Elsevier.

Leigh, A., & Ryan, C. (2008). How and why has teacher quality changed in Australia? *Australian Economic Review, 41,* 141-159.

Levesque, C., Zuehlke, A.N., Stanek, L.R., & Ryan, R.M. (2004). Autonomy and competence in German and American university students: A comparative study based on self-determination theory. *Journal of Educational Psychology, 96,* 68-84.

Little, T.D. (1997). Mean and covariance structure (MACS) analysis of cross-cultural data: practical and theoretical issues. *Multivariate Behavioral Research, 32,* 53-76.

Luschei, T. F. (2012). In search of good teachers: Patterns of teacher quality in two Mexican states. *Comparative Education Review, 56,* 69-97.

Luschei, T., & Carnoy, M. (2010). Educational production and the distribution of teachers in Uruguay. *International Journal of Educational Development, 30,* 169-181.

Luschei, T. F., & Chudgar, A. (2011). Teachers, student achievement and national income: A cross-national examination of relationships and interactions. *Prospects: Quarterly Review of Comparative Education, 41*, 507-533.

Lichtman, M. (2013). Qualitative research in education (3rd ed.). Los Angeles: Sage.

Little, J., & Bartlett, L. (2010). The teacher workforce and problems of educational equity. *Review of Research in Education, 34*, 285-328.

Liu, E., & Johnson, S. (2006). New teachers' experiences of hiring: late, rushed, and information-poor. *Educational Administration Quarterly, 42*, 324-360.

Loeb, S., & Myung, J. (2010). Economic approaches to teacher recruitment and retention. In P.Peterson, E. Baker, & B. McGaw, *International Encyclopedia of Education* (3rd ed.) (pp 473-480). Oxford: Elsevier.

Loeb, S., Kalogrides, D., & Beteille, T. (2012). Effective schools: teacher hiring, assignment, development, and retention. *Education Finance and Policy, 7,* 269-304.

Maaz, K., Hausen, C., McElvany, N., & Baumert, J. (2006). Keyword: Transitions in the educational system. *Zeitschrift für Erziehungswissenschaft, 9,* 299-327.

Maier, A., & Youngs, P. (2009). Teacher preparation programs and teacher labor markets. *Journal of Teacher Education, 60,* 393-407.

Malmberg, L.-E. (2006). Goal-orientation and teacher intrinsic motivation among teacher applicants and student teachers. *Journal of Teaching and Teacher Education, 22*, 58-76.

Malmberg, L.-E. (2008). Student teachers' achievement goal orientations during teacher studies: antecedents, correlates and outcomes. *Learning and Instruction, 18,* 438-452.

Marsh, H.W., Nagengast, B., & Morin, A. (2013). Measurement invariance of Big-Five factors over the life span: ESEM of gender, age, plasticity, maturity, and La Dolce Vita effects. *Developmental Psychology, 46,* 1194-1218.

Marshall, J. H., & Sorto, A. M. (2012). The effects of teacher mathematics knowledge and pedagogy on student achievement in rural Guatemala. *International Review of Education, 58(2),* 173-197.

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

Moran, A., Kilpatrick, R., Abbott, L., Dallat, J., & McClune, B. (2001). Training to teach: motivating factors and implications for recruitment. *Evaluation & Research in Education, 15*, 17-32.

Morge, L., Toczek, M-C., & Chakroun, N. (2010). A training programme on managing science class interactions: Its impact on teachers' practises and on their pupils achievement. *Teaching and Teacher Education, 26,* 415-426.

Mortensen, D.T. (1986). Job search and labor market analysis. In O. Ashenfelter & R. Layard (eds.), *Handbook of Labor Economics Volume 2* (pp. 849-919). London: Elsevier.

Munoz, M.A., & Chang, F.C. (2007). The elusive relationship between teacher characteristics and student academic growth: A longitudinal multilevel model for change. *Journal of Personnel Evaluation in Education, 20,* 18-18.

Musset, P. (2010). *Initial teacher education and continuing training policies in a comparative perspective: Current practices in OECD countries and a literature review on potential effects.* OECD Working Papers No. 48. Paris: OECD Publishing.

Muthén, L.K., & Muthén, B.O. (1998-2013). *Mplus user's guide* (6th ed*)*. Los Angeles, CA: Muthén & Muthén.

Myrberg, E. (2007). The effect of formal teacher education on reading achievement of 3rd-grade students in public and independent schools in Sweden. *Educational Studies, 33,* 145-162.

Neild, R.C., Farley-Ripple, E.N., & Byrnes, V. (2009). The effect of teacher certification on middle grades achievement in an urban district. *Educational Policy, 23,* 732-760.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26,* 237–257.

OECD (2010). *PISA2009 results: What makes a school successful?* Paris: OECD Publishing.

OECD (2012a). *Education at a glance 2012: OECD indicators.* Paris: OECD Publishing.

OECD (2012b). *PISA 2009 Technical Report*. Paris: OECD Publishing.

OCED (2013). *PISA 2012 Results: What makes schools successful? Resources, policies and practices (Volume IV)*. Paris: OECD Publishing.

Paine, L., & Zeichner, K. (2012). The local and the global in reforming teaching and teacher education. *Comparative Education Review, 56,* 569-583.

Palardy, G.J., & Rumberger, R.W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30,* 111-140.

Parsons, T. (1951). *The social system.* London: Routledge.

Peske, H., & Haycock, K. (2006). *Teaching inequality: how poor and minority students are shortchanged on teacher quality. A report and recommendations by the Education Trust.* Washington, DC: The Education Trust.

Phillips, K. (2010). What does 'highly qualified' mean for student achievement? Evaluating the relationships between teacher quality indicators and at-risk students' mathematics and reading achievement gains in first grade. *Elementary School Journal, 110,* 464-493.

Rice, J.K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes.* Washington, D.C.: Economic Policy Institute.

Richardson, P., & Watt, H. (2006). Who chooses teaching and why? Profiling characteristics and motivations across three Australian universities. *Asia-Pacific Journal of Teacher Education, 34,* 27-56.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73,* 417-458.

Rockoff, J.E., Jacob, B.A., Kane, T.J., & Staiger, D.O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy, 6,* 43-74.

Rothland, M., & Terhart, E. (2011). Eignungsabklärung angehender Lehrerinnen und Lehrer. *Zeitschrift für Pädagogik, 57,* 635-638.

Rothstein, J. (2012). *Teacher quality policy when supply matters.* NBER Working Paper 18419. National Bureau of Economic Research. http://www.nber.org/papers/w1841. (Accessed 12/01/13)

Ryu, E., & West, S. (2009). Level-specific evaluation of model fit in multilevel structural equation modeling. *Structural Equation Modeling, 16,* 583-601.

Sahlberg, P. (2011). Developing effective teachers and school leaders: The case of Finland. In L. Darling-Hammond & R. Rothman (eds.), *Teacher and leader effectiveness in high performing education systems* (pp. 13-22). Washington, DC: Alliance for Excellent Education.

Saks, A., Uggerslev, K., & Fassina, N. (2007). Socialization tactics and newcomer adjustment: A meta-analytic review and test of a model. *Journal of Vocational Behavior, 70,* 413-446.

Santibanez, L. (2006). Why we should care if teachers get A's: Teacher test scores and student achievement in Mexico. *Economics of Education Review, 25,* 510-520.

Sass, T.R., Hannaway, J., Xu, Z.Y., Figlio, D.N., & Feng, L. (2012). Value-added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics, 72,* 104-122.

Schalock, D., Schalock, M., & Ayres, R. (2006). Scaling up research in teacher education. New demands on theory, measurement, and design. *Journal of Teacher Education, 57,* 102-119.

Schwille, J., & Dembele, M. (2007). *Global perspective on teacher learning: improving policy and practice.* Paris: UNESCO International Institute for Educational Planning.

Scott, W.R., & Davis, G.F. (2007). *Organizations and organizing. Rational, natural, and open system perspectives.* Upper Sadle River: Pearson.

Sellar, S., & Lingard, B. (2013). Looking east: Shanghai, PISA 2009 and the reconstitution of the global education policy field. *Comparative Education, 49,* 464-485.

Shen, J., Mansberger, N., & Yang, H. (2004). Teacher quality and students placed at risk: results from the Baccalaureate and Beyond longitudinal study 1993-97. *Educational Horizons, 4,* 1-6.

Sicherman N., & Galor O., (1990). A theory of career mobility. *Journal of Political Economy, 98,* 169-192.

Simon, H.A. (1976). *Administrative behavior: a study of decision making processes in administrative organizations.* New York: Free Press.

Slater, H., Davies, N.M., & Burgess, S. (2012). Do teachers matter? Measuring the variation in teacher effectiveness in England. *Oxford Bulletin of Economics and Statistics, 74,* 629-645.

Staiger, D.O., & Rockoff, J.E. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives, 24,* 97-118.

Stapleton, L. (2013). Multilevel structural equation modeling with complex sample data. In G.R. Hancock & R.O. Mueller (eds.), *Structural equation modeling: a second course* (pp. 345-383). Charlotte, CA: Information Age Publishing.

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S.H. (2009). Testing measurement invariance using multigroup CFA: differences between educational groups in human values measurement. *Quality and Quantity, 43,* 599-616.

Stiglitz, J.E. (1975). The theory of screening, education and the distribution of income. *American Economic Review, 65,* 283-300.

Tatto, M.T., Schwille, J., Senk, S.L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M., & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam: IEA.

Thompson, M.S., & Green, S.B. (2013). Evaluating between-group differences in latent variable means. In G.R. Hancock & R.O. Mueller (eds.), *Structural equation modeling: A second course* (pp. 119-169). Charlotte, CA: Information Age Publishing.

Tillema, H. H. (1994). Training and professional expertise: bridging the gap between new information and pre-existing beliefs of teachers. *Teaching and Teacher Education, 10*, 601-615.

Van de Werfhorst, H.G., & Mijs, J.J.B. (2010). Achievement inequality and the institutional structure of educational systems: a comparative perspective. *Annual Review of Sociology, 36,* 407-428.

Van der Velden, R., & Wolbers, M.H.J. (2007). How much does education matter and why? *European Sociological Review, 23,* 65-80.

Van der Velden, R., & Wolbers, M.H.J. (2008). A framework for monitoring transition systems. *OECD Education Working Papers, 20.* Paris: OECD Publishing.

Van Maanen, J., & Schein, E.H. (1979). Toward a theory of organizational socialization. *Research in Organizational Behavior, 1,* 209–264.

Wang, A., Coleman, A., Coley, R., & Phelps, R. (2003). *Preparing teachers around the world.* Princeton: Educational Testing Service.

Wang, J., & Wang, X. (2012). *Structural equation modeling. Applications using Mplus*. Chichester: Wiley.

Wayne, A.J., & Youngs, P. (2003). Teacher characteristics and student achievement. *Review of Educational Research, 73,* 89-122.

Winters, M.A., Dixon, B.L., & Greene, J.P. (2012). Observed characteristics and teacher quality: impacts of sample selection on a value added model. *Economics of Education Review, 31,* 19-32.

Woessmann, L. (2003). Schooling resources, educational institutions, and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics, 65,* 117-170.

Yeh, S.S. (2009). The cost-effectiveness of raising teacher quality. *Educational Research Review, 4,* 220-232.

Younger, M., Brindley, S., Pedder, D., & Hagger, H. (2004). Starting points: student teachers' reasons for becoming teaching and their preconceptions of what this will mean. *European Journal of Teacher Education, 27,* 245-264.

Zeichner, K. (2005). A research agenda for teacher education. In M. Cochran Smith, & K. Zeichner (eds.), *Studying teacher education* (pp. 737-761). Mahwah: Lawrence Erlbaum.

Zeichner, K. (2006). Studying teacher education programs: enriching and enlarging the inquiry. In C.F. Conrad, & R.C. Serlin (eds.), *The Sage handbook for research in education* (pp. 79-95). Thousand Oaks: Sage.

Zeichner, K., & Conklin, H. (2005). Teacher education programs. In M. Cochran Smith, & K. Zeichner, (Eds.), *Studying teacher education* (p. 645-736). Mahwah: Lawrence Erlbaum.

Zeichner, K., & Conklin, H. (2008). Teacher education programs as sites for teacher preparation. In M. Cochran-Smith, S. Feiman-Nemser, D. McIntyre, & K. Demers (eds.), *Handbook of research on teacher education (3rd ed)* (pp. 269-289). New York: Routledge.

Zhao, N.N., Valcke, M., Desoete, A., Verhaeghe, J., & Xu, K. (2011). A multilevel analysis on predicting mathematics performance in Chinese primary schools: Implications for practice. *Asia-Pacific Education Researcher, 20,* 503-520.

Appendix A

Studies included in the literature review (Chapter 2).

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Aaronson et al. (2007) [USA]	Estimation of importance of teachers and relating measures of teacher effectiveness to observable teacher characteristics DV: student test scores	VAM including lagged test scores (Administrative data of 88 Chicago public high schools; 1996-98, 52957 students, 783 teachers)	College attended (dummy coded; 8 alternatives: US News 1-5 and else, local, missing) College major (dummy coded; 4 alternatives: education, math, science, else) Advanced degree (binary, MA or PhD) Certification status (dummy-coded; 6 alternatives: bilingual, child, elementary, high school, special, substitute) (Student level)	Implicitly in quality of college attended	School fixed effects Dataset	Education background characteristics of teachers loosely, if at all, related to student achievement (less than 1% explained variance)
Akiba et al. (2007) [INTERNATIONAL]	Assessment of national levels of teacher quality and gaps in access to qualified teachers DV: math achievement (IRT scaled)	Correlation and multiple regression analyses (TIMSS2003 data; 46 countries)	Certification (binary), Math major (binary), Math education major (binary; all yes/no); Percentage of students taught of qualified teachers (overall measure: full certification, math and education major); (Student level)	---	Cross-country analysis	All teacher education variables (except math major) significantly related to student achievement; Unstandardized coefficients: Certification 2.71 (1.16) M. Ed. Major 1.15 (0.59) Overall 2.04 (0.56)
Akyüz & Berberoglu (2010) [INTERNATIONAL]	Investigation of the impact of teacher and classroom characteristics on student achievement DV: math achievement (IRT scaled)	MLM (HLM) with two levels (student and classroom) (TIMSS-R 1999; 10 countries, 1642 classrooms, 38109 students)	Highest degree (binary; MA or PhD/BA or less) included in teacher characteristics (Classroom level)	---	Cross-country analysis	Highest degree without impact except in Slovakia (negative; unstandardized coefficient -39.08 (19.89)); *Relation between teacher education and classroom practice not investigated!*
Andersson et al. (2011) [SWEDEN]	Estimation of the compositional effect of certification on student achievement DV: student GPA	VAM with an instrumental variable specification including school fixed effects (Administrative data of 583698 grade 9 students; 1993-2004)	Percentage of non-certified teachers (School level)	---	School fixed effects	Share of non-certified teachers associated with a decrease in school average student GPA (standardized effect -0.018 SD)

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Aslam & Kindgon (2011) [PAKISTAN]	Examination of the impact of teacher characteristics and behavior on student achievement DV: math and language grades (standardized)	VAM with a student-level within-subject estimation including student fixed effects (Administrative, school based data of Punjab district; 50 schools, 100 teachers, 1410 grade 8 students; 2002/03)	Advanced degree (binary; MA/Mphil/PhD or else) Subject-specific degree (binary; yes/no) Years of teacher training Teacher scores ELA test (range 1-5) (Student level)	---	Student fixed effects Dataset	Degree without impact on student achievement; teacher test scores have a positive impact (standardized effect 0.09 SD)
Baumert et al. (2010) [GERMANY]	Investigate the influence of teacher content knowledge and pedagogical content knowledge on instruction and student achievement DV: mathematics achievement (latent variables)	MLM (SEM) mediation model with two levels (students, classrooms) and two time points, estimated separately for CK and PCK; prior knowledge included (COACTIV; 181 teachers, 194 classes, 4353 grade 10 students (80/114 academic/non-academic track classes); 2003)	Content knowledge Pedagogical content knowledge (both latent constructs) Type of teacher training (three certification types) (Classroom level)	GPA (rating, 1-6)	Prior math knowledge Mental ability Parental education Social status Immigrant status Sampling	Academic track teachers had higher levels of CK and PCK (persistent over career) Substantial effect of PCK on student achievement (effect size d_{class}= 0.328 (.10)), fully mediated by the cognitive structure of learning opportunities (i.e. instructional practices)
Beese & Liang (2010) [INTERNATIONAL]	Examine the degree of variability in science teacher shortage and instructional resources between countries and how these school factors interact with student characteristics for an explanation of differences in student achievement DV: science achievement (IRT scaled)	MLM (HLM) with two levels (students and schools) (PISA2003 database; 869 Canadian, 166 US, 155 Finnish schools with 22646, 5611, and 4714 students)	Teacher qualifications included in schools reporting the ratio of part-time and full-time teachers and lack of qualified teachers (percentages) (School level)	---	Cross-country analysis	Finland: insignificant influence of teacher qualifications; US, Canada: significant impact of part-/full-time ratio (-0.13, 0.48 points per unit) and significant negative influence of lack of science teachers (-2.69, -5.65 points per unit) *Relation between teacher education and classroom practice not investigated!*
Boyd et al. (2008) [USA]	Examine changes in teacher quality distribution, observed teacher qualifications on student achievement DV: math/ELA test scores	VAGM including student, grade, and time fixed effects (Administrative data, 4th and 5th grade, 578630 observations)	Certification exam passed (binary, yes/no) Certification test score Path (categorical) Area of certification (categorical) Barron's college ranking (categorical) SAT score (math/verbal) (Student level)	Barron's college rank (categorical) SAT scores	Student fixed effects Grade fixed effects Time fixed effects Dataset	Certification 0.042 SD SAT scores: Math 0.043 SD Read 0.034 SD Competitive college 0.014 SD

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Boyd et al. (2009) [USA]	Examine how the preparation of first year teachers influences student achievement DV: math/ELA test scores	VAM including school fixed, random, and OLS specifications; (1) distribution of value-added of teachers, (2) impact of program characteristics, (3) impact of teachers' training experiences (Administrative data and documents, 31 programs across 18 institutions, 65-80000 grade 4-8 students, 8773 teachers; 2000-2006)	Path (categorical) Exit examination passed (binary yes/no) (Teacher level) Nr. Of math/ELA courses Capstone project (binary, yes/no) Oversight of student teaching (binary, yes/no) Tenure track (binary, yes/no) (Program level)	---	School fixed effects Program level analysis in one district	Program characteristics significantly related to student achievement across cohorts and subjects; Variation in institutional effectiveness (difference 0.07 SD); Variation in program effects (difference across programs 0.18/0.10 SD in math/ELA)
Bressoux et al. (2009) [FRANCE]	Examine the impact of teacher characteristics on student outcomes DV: math/reading test scores	VAM including random class effects in a quasi-experimental design (Administrative data; 12 departements, 4001 grade 1-5 students, 197 teachers; 1991/92)	Education level (categorical; experienced, trained novice, untrained novice) (Student level)	---	Quasi-experimental	Significant positive impact of education level on student achievement (0.25 SD) in math
Buddin & Zamarro (2009) [USA]	Examine the relation between teacher quality and student achievement DV: CAT/6 scores in math/reading	VAM and VAGM with student and teacher fixed effects; FGLS regression on teacher fixed effects (indirect approach) (Administrative data; Grades 2-5, 332538 and 16412 teachers (level model); 325521 students and 13047 teachers (gain model); 2000 observations in the FGLS regression)	Advanced degree (binary; MA/other) Teacher test scores (CBEST, CSET, RICA) (Student level)	Test scores	Student fixed effects Dataset	VAM and VAGM show similar results: Advanced degrees without impact, teacher test scores show little association with student achievement in reading (significant negative effect in math)
Carlisle et al. (2009) [USA]	Examine the extent to which teachers' knowledge about reading is related to differences in student achievement DV: ITBS reading	MLM (HLM) with two levels (student and classroom) (Evaluation of Reading First in Michigan; 112 elementary schools with an average of 357 students in grade 1-3; 2003/04 and 2004/05; Teacher data from Teacher's Quest; 977 teachers, 2004/05)	Teacher LRC test scores (metric and categorized Certification (binary, permanent yes/no) Advanced degree (binary, yes/no) (Classroom level)	---	Dataset	No significant influence of teacher knowledge on student achievement; First grade: negative influence of not having a master's degree (-2.55 (0.80); Second and third grade: negative influence of not having permanent certification (-2.33 (1.12); -2.07 (0.94) Weak, inconsistent results across grades

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Clotfelder et al. (2006) [USA]	Examine the impact of non-random matching of students and teachers on the relation of teacher characteristics and student achievement DV: student test scores math/reading	VAM with student and school fixed effects (Administrative data, 60791 grade 5 students, 3223 teachers; 2000/01)	Advanced degree (binary) Certification status (binary) Certification test scores Barron's college ranking (categorical) (Student level)	Barron's college ranking	Student fixed effects School fixed effects Dataset	Certification test scores positive impact on math achievement (0.02 SD); Positive influence of certification status on reading Negative impact of advanced degrees on math and reading
Clotfelder et al. (2007) [USA]	Examine the relation of teacher characteristics and student achievement DV: student test scores math/reading	VAM and VAGM with student and school fixed effects (Administrative data; 1.8 million observations in level model, 1.0 million observations in gain model; 1995-2004)	Advanced degree (binary) Licensure status (categorical) Certification status (binary) Certification test scores Barron's college ranking (categorical) (Student level)	Barron's college ranking	Student fixed effects School fixed effects Dataset	Negative impact of advanced degree on student achievement; negative effect for students with teachers without regular license; Certified teachers more effective; Non-linear effect of teacher test scores relative to average scoring teachers in math (difference 0.13/0.074 SD gains/level model)
Clotfelder et al. (2010) [USA]	Examine the relation of teacher characteristics and student achievement in high school DV: EOC test scores across 5 subjects	VACSM including student fixed effects (Administrative data with four cohorts of grade 10 students, 857548 observations; 1999/2000, 2000/01, 2001/02, 2002/03)	Licensure status (categorical) Certification status (subject-specific, categorical) Certification test scores Barron's college ranking (categorical) (Student level)	Barron's college ranking	Student fixed effects Dataset	Quality of undergraduate institution predictive of student achievement Subject-specific test scores with positive influence (math 0.0472 SD) Negative effect of non- and lateral licensed teachers Positive certification effect Overall credential effect 0.23 SD (difference between teachers with weak and strong credentials)
Connor et al. (2005) [USA]	Examine the relation of teacher qualifications to classroom practice and student achievement DV: Reading achievement (Letter-word identification, word attack, vocabulary)	SEM with ML estimation testing an ecological instructional model, controlling for prior achievement (Study of Early Child Care and Youth Development; 787 children, 4 classrooms)	Number of years of education Elementary teaching credential (dropped due to skewness) (Student level)	---	Dataset	Teacher education directly affects student outcomes (higher vocabulary, lower letter word scores; 0.061; -0.088) Inclusion of SES variables: indirect effect of teacher education on vocabulary through teacher warmth; negative impact on letter-word identification

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Croninger et al. (2007) [USA]	Analyze the relation between teacher qualifications and student achievement DV: Cognitive assessment in reading/math (IRT scaled)	MLM (HLM) with three levels (students, classrooms, schools) (Early Childhood Longitudinal Study, 10980 students in 2148 elementary schools; 1998/99 and follow up assessment in 2000)	Certification status (binary) Advanced degree (binary) Elementary degree (binary) Course ratio math/reading (Classroom level) Proportions with certification, advanced/elementary degrees; average course ratios (School level)	Course ratios	Dataset	Elementary degree (0.08 SD in reading) Average course ratios (0.05 SD in reading, 0.03 SD in math; contextual effect) Negative effect of proportion of teachers with advanced degree (-0.07 SD in math)
Desimone & Long (2010) [USA]	Examine the extent to which teacher quality and teaching quality influence math achievement growth and the achievement gap between white/black and low/high SES students DV: Cognitive assessment math (IRT scaled)	MLM (SEM) growth curve model with three levels (level 1: achievement, level 2: SES, level 3: teacher quality, instruction, controls) (Early Childhood Longitudinal Study, 10.980 students in 2.148 elementary schools; 1998 – 2000)	Degree (binary; less/more than BA) Level of teacher certification (categorical; no, high, emergency, alternative) number of college-level mathematics courses (Teacher level)	Number of college courses	Dataset	No consistent relationship between teacher quality and achievement growth; Math achievement growth slower if teacher has less than BA (b = -0.70), and faster if teacher has a high or alternative certification (b = 0.30/0.60); similar results in subgroup analyses
Fuchs & Wößmann (2007) [INTERNATIONAL]	Estimation of international education production functions DV: Student achievement in math/science/reading (IRT scaled)	International educational production function with WLS estimation (PISA2000 data; 31 countries, 174.227 students grade 7-11)	Share of teachers with masters in pedagogy, subject-specific masters, certification (percentages) (Student level)	---	Cross-country analysis	Pedagogy masters 8.338***/4.283* (science/reading); Certification 11.178***/10.484***, 6.741*** (math/science/reading); Subject-specific masters 11.847***/10.101***/17.583*** (math/science/reading)

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Gansle et al. (2012) [USA]	Prediction of student achievement by teacher characteristics and other factors DV: iLEAP, LEAP of the ITBS in ELA, reading, math, science and social studies)	MLM (HLM) with three levels (students, teacher/classrooms, schools) (Administrative data; 162500-237000 grade 4-9 students, 5100-7300 teachers, 1050-1250 schools; 2007-2008)	Teacher preparation program (categorical; undergraduate 1, undergraduate 2, masters alternate 1 and 2, non-masters certification only, practitioner 1 and 2, private practitioner 1 and 2) (Classroom/teacher level)	Implicitly included in program characteristics	Dataset	Graduates from private practitioner 2, Masters alternate certification 1, and university practitioner program did better than the average new teacher in social studies and math; Mixed results for University Practitioner Program 1; Graduates from Undergraduate 1 and 2, Non-Masters Certification1 and Private Practitioner 1 were similar to average new teachers; Graduates from Private Practitioner 1 in reading and Non-Masters 1in ELA did worse than average new teachers
Goldhaber (2007) [USA]	Investigation of the relationship between teacher licensure and student achievement DV: NCDPI test scores (math/reading)	VAM variants including student and school fixed effects (Administrative data; 24237 teachers, 722166 students; 1994-2004)	Advanced degree (binary; MA or not) Certification (binary; fully or other) Licensure test passed (binary; pass/fail) Scores on PRAXIS I/II tests (normalized) (Student level)	Only sample selection: subsample of novice teachers	Student fixed effects School fixed effects	Teachers graduating from approved programs outperform teachers who do not (0.01 SD); NBPTS certified teachers outperform non-certified teachers (0.01-0.04 SD); Licensure 0.06 SD (math); Teacher scoring in top quintiles more effective (math 0.035, read 0.022 SD)
Guo et al. (2012) [USA]	Examine the relations between teacher characteristics and student literacy outcomes DV: WJ-R reading achievement (Letter-word identification, word attack, vocabulary)	SEM with FIML estimation testing an ecological instructional model controlling for prior achievement (G3 reading) (Study of Early Child Care and Youth Development;1043 grade 5 students and teachers; Phase III 2000-2005)	Master's degree (binary; yes/no) (Student level)	---	Dataset	No direct or indirect effects of teacher education on literacy outcomes or teacher self-efficacy (indirect paths tested)

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Harris & Sass (2011) [USA]	Examination of the relationship between teacher productivity and teacher training DV: Student test scores grade 3-10 (math/reading)	VAGM with student and teacher/school spell effects; regression of within-school teacher effects on pre-service education variables (Administrative data; grades 3-10, elementary, middle and high schools; 216.893/240.317 students; 1.125/616/325 elem., middle, high school math teachers, 1.307/419/368 elem., middle, high school reading teachers; 1999-2005)	Advanced degree (binary, MA or not) (Student level) Number of credits earned in various types of education, SAT total score (Teacher level)	Teacher fixed effects Course credits SAT total score	Student fixed effects School fixed effects	Advanced degree without impact on student achievement; Majors and SAT scores exert no influence on teacher productivity; neither do any coursework type; exception: classroom management positive influence (high school, reading), statistics credits (middle school, math)
Heck (2007) [USA]	Examination of the relation of school-level differences in teacher quality and student achievement DV: SAT-9 test scores (math/reading)	SEM multilevel growth model with maximum likelihood estimation, including interactions of student composition and teacher variables and cross-level interactions (Administrative data; 197 elementary schools, 14.082 5th grade students; 2004-2006)	Share of fully qualified and fully certified teachers (percentages) (School level)	---	School level analysis	Mean teacher quality significantly related to math and reading achievement levels (γ = 3.798 reading, γ = 2.783 math); Mean teacher quality only related to achievement growth in math (γ = 2.291); Increasing teacher quality reduces achievement gap
Hogrebe & Tate (2010) [USA]	Investigation of the relationship between science proficiency and school composition and context factors DV: SciMAP index scores (school level)	Multiple regression of 10th grade science proficiency on school environment, course related and teacher related variables (Administrative data; 423 high schools; 2002)	Share of teachers with advanced degree and regular certification (percentages) (School level)	---	School level analysis	Increases in teachers with MA degrees, teachers with regular certification associated with increases in average SciMAP scores (β = 0.142 and 0.174 for degree and certification); reduction of achievement gap
Huang & Moon (2009) [USA]	Investigate teacher characteristics and their relation with student achievement gains DV: SAT-10 test scores (reading)	MLM (HLM) with three levels (students, classrooms, schools) and 6 specifications (Administrative and survey data; 2210 2nd grade students, 195 teachers, 53 elementary Title I schools; 2006-2007)	Degree (binary; MA or higher) Certification status (binary; yes/no) (Teacher/Classroom level)	---	Dataset	No significant impact of teacher education variables on student achievement

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Jepsen (2005) [USA]	Investigation of non-traditional teacher characteristics and their influence on student achievement DV: CBTS-4 scores (math/reading)	VAGM with classroom (teacher + peer effects combined) and school fixed effects; regression of classroom effects on teacher and student characteristics (GLS/OLS; indirect) (Prospect study; two cohorts: 18837 and 18639 1st and 3rd grade students; fall and spring 1991)	Degree (binary; less/more than BA) Certification status (binary; fully/other) (Classroom level)	Teacher fixed effects included in classroom fixed effects	School fixed effects Dataset	No significant effect of teacher education variables on achievement in either cohort
Kane et al. (2008) [USA]	Investigate the effectiveness of newly hired teachers (relationship between certification and student achievement) DV: Student test scores (math/reading)	VAM with school and school-grade-year fixed effects (Administrative data; grades 4-8 (elementary and middle schools), 623482 students, 18856 teachers in math, 607563 students, 19083 teachers in reading; 1999-2005)	Certification status (categorical; TF, TFA, International recruits, not certified) SAT scores math/verbal (Student level)	SAT scores	School fixed effects Dataset	Considerable variation in teacher effectiveness, but certification status is largely irrelevant; SAT score shows no differential effects; Differences between certification groups irrelevant, but high variance within certification groups
Kaya & Rice (2010) [INTERNATIONAL]	Examine the effects of individual student and classroom factors on student achievement within and across five countries DV: Science achievement (IRT scaled)	MLM (HLM) with two levels (students, classroom) (TIMSS2003 data; 5 countries, 120-171 schools, 2665-6122 students; between country analysis: 913 teachers, 24333 students; 2003)	Degree (binary; MA-PhD or not) Subject-specific major (binary; science/math or not) (Classroom level)	---	Cross-country analysis	No significant relations between teachers' degree and subject-specific majors and science achievement (cross-country analysis did not include teacher education variables)
Konold et al. (2008) [USA]	Examination of teacher value added to student learning (investigate one school of education program) DV: Student test scores (math: data representation and interpretation)	SEM in a multigroup framework including teacher education, teacher behavior and student outcomes (controlling for prior achievement) (Project data; 2 academic years; random assignment of 680 students grade 6, 7 and 8 to 43 teacher or 47 art/science students;	Students with and without formal teacher education (BA/MA and postgraduate teacher education program) (Program level)	---	Experimental design	Teaching behavior significantly influences student achievement (standardized regression weight 0.36; stays significant across teacher students groups; n.s. for arts/science students); Teacher behavior factor mean for teachers without formal training 6.4 units below that for teachers with formal training

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Kukla-Acevedo (2009) [USA]	Examine the relation between teacher preparation and student achievement DV: KCCT test scores (math)	VAGM with fixed effects and between effects specification including interactions of teacher education variables and experience (Administrative data; 3812 elementary students; 2000-2003)	Number of content hours Number of education hours, Overall GPA, Math GPA Math-education GPA (Student level)	Amount of coursework Teacher fixed effects	Student fixed effects School fixed effects Dataset	Declining effect of teacher education; Positive interaction math GPA/experience; Only overall GPA consistently predicts student achievement (effect size range 0.34-0.84 SD); Positive effect of math/math education GPA and hours only significant in the pooled sample
Leigh (2010) [AUSTRALIA]	Estimation of teacher productivity using changes in student test scores DV: DETA literacy/numeracy test scores (math/reading)	VAM with teacher, grade by year and student fixed effects (1); (2) regression of teacher fixed effects on teacher characteristics (indirect) (Administrative data; 90000 grade 3, 5, and 7 primary students (3 cohorts), 10000 teachers; 2001-2004)	Advanced degree (binary; MA/further or not) DETA rating (categorical; outstanding, quality, satisfactory, eligible for temporary employment applicants) (Student level)	Teacher fixed effects	Student fixed effects Dataset	Approx. 1% variance in student achievement explained by teacher characteristics; Negative influence of advanced degree; Teachers with higher DETA ranking show higher student achievement
Luschei & Chudgar (2011) [INTERNATIONAL]	Examine the relation of teacher characteristics and student achievement DV: math and science achievement (IRT scaled)	Multiple regression (OLS) separate for 25 countries (TIMSS2003 data; 25 countries, 836-8025 grade 4 science and math teachers; 2003)	Teacher education (dichotomous; less than first degree, first degree, higher than first degree) Teacher readiness math Teacher readiness science (Indexes) (Student level)	---	Cross-country analysis	Overall lack of significance; First degree positively related to math student achievement in Slovenia; More than first degree negatively related to math student achievement in Hungary Readiness significant positive influence in Iran, Philippines, and US) Less than first degree negatively associated with science achievement in Moldavia and Singapore, and positively related in Norway; more than first degree significant negatively related to science achievement in US; teacher readiness positive influence in Taipei and Scotland
Luschei (2012) [MEXICO]	Identification of teacher attributes related to student achievement and description of the distribution of those attributes DV: student test scores (CM administered test)	VAM on grade level with school-by-grade any year fixed effects (Administrative data; Aguascalientes, Sonora; 3722 and 5177 teachers; two cohorts; 2003-2004)	Teacher test scores [Sonora] (Percentage correct answers) Teacher test points [Agua] (points, 0-28) Training evaluation score (federal, state; points, 0-12) Percentage of teachers with level of education (Normal basica, Normal licenciatura, Normal superior) (Grade level)	Implicitly included in training evaluations	Grade level analysis Dataset	Fixed effect specification with lagged achievement: Aguascalientes: Teacher test score points 0.04 SD; Teacher federal training score 0.06 SD; Sonora: Teacher test score (percentage) 0.09 SD; Teacher federal training score 0.05 SD; Teacher state training score 0.06 SD Quality teachers (high test scores) are concentrated in low-poverty municipalities and urban schools (exception Sonora)

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Marshall & Sorto (2012) [GUATEMALA]	Analyze the impact of teachers mathematics knowledge on student achievement DV: Overall and content area scores on PRONERE tests (IRT scaled)	MLM (HLM) with two levels (students and teachers) (PRONERE data including survey and classroom observations; 699 students, 65 classrooms; 2000-2001	Common content knowledge (primary; percent correct) Common content knowledge (middle, percent correct) Specialized content knowledge (percent correct) Mathematics knowledge for teaching (percentage correct on all common and specialized items) (Teacher level)	---	Dataset	CCK 0.06 and 0.08 SD SCK 0.06 SD MKT 0.07 SD (whole test) MKT 0.07 SD (problem solving) MKT 0.10 SD (fraction subtest); Teacher process variables positively related to student achievement (0.10-0.15 SD); *influence of teacher education on teacher processes not investigated!*
Metzler & Woessmann (2012) [PERU]	Estimate the impact of teachers' academic skills on student achievement DV: National evaluation of students (math/reading)	VAM with a correlated random effects specification (cross sectional) (Administrative data; 12000 grade 6 students in 900 randomly sampled schools; same teacher one classroom (STOC) sample; 2004)	Degree (binary; yes/no) Teacher test scores (subject-specific tests) (Student level)	---	Dataset Sample restriction	Significant association between teacher test score and student achievement: 0.13 SD/0.17 SD (math/read); Degree n.s.; STOC sample: Increase of teacher subject knowledge in math increases math test scores by 6.4% SD (read n.s.); robust across several specifications
Munoz & Chang (2007) [USA]	Evaluate effects of teacher characteristics on student achievement DV: Predictive Assessment Series (PAS) in reading, percent correct	MLM (HLM) growth model with two levels and three time points (PAS/Time and teacher) (Administrative data; 1487 grade 9 (high school) students and 58 grade 9 teachers; 2005/2006)	Level of education (binary; BA or MA/above)	---	Dataset	Teacher education did not predict either initial status of student achievement or achievement growth
Myrberg (2007) [SWEDEN]	Investigate the influence of teacher competence on reading achievement DV: reading achievement (IRT scaled)	MLM (SEM) with two levels (student and school/teacher) (PIRLS2001 data; 292 elementary schools, 717 classes, 1067 teachers, 10632 students; 2001)	Certification status (dummy coded; 9 alternatives) (Teacher level)	---	Dataset	Main effect of certification on reading achievement (25.6, t = 4.40); effect remains when entering student background (certification 22.60, t = 4.39); Teacher education effect significant across school types

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Neild et al. (2009) [USA]	Examine the relationship between student academic growth and certification status of teachers DV: CTBS test scores (math/science)	MLM (HLM) specified as three level mixed model (students – teachers – schools) controlling for prior achievement (fall achievement score) (Administrative data; grade 5 through 8 (middle schools); 539 teachers, 22853 students (math), 495 teachers, 21989 students (science); 2002/03)	Certification status (categorical; elementary, special, secondary, non certified) (Teacher level)	---	Dataset	Moderate effect size (0.09-0.12 SD) of elementary certification (math); Secondary certification effect size in science 0.20; Not being certified and special education certification have lower student gains than teachers with elementary certification
Nye et al. (2004) [USA]	Estimation of teacher effects on achievement status DV: Average SAT score (math/reading)	MLM (HLM) with three levels (within classroom, school/teacher; across schools); regression of teacher effect on teacher education (indirect); (STAR project; 79 elementary schools, 5766 kindergarten, 6377 grade 1, 5968 grade 2, 5903 grade 3 students)	Teacher education (dichotomous; graduate/advanced degree or other) (Teacher level)	---	Experimental design	Not more than 5% of variance in teacher effects explained; At grade 3, teacher education had significant estimated effects for reading and math of 0.06 and 0.09 SD, on achievement gains only
Palardy & Rumberger (2008) [USA]	Investigation of the importance of teacher background qualifications, attitudes, and instructional practices to achievement gains DV: NCES achievement gains (reading/math)	MLM (HLM) with three levels and two time points (students, classrooms, schools) (ECLS study; 3496 elementary students, 887 classrooms, 253 schools; longitudinal sample 1998-2003)	Certification status (binary; regular/advanced) Advanced degree (binary; yes/no) (Classroom level)	---	Dataset	Teacher effect size 0.30 SD (read), 0.25 SD (math); Full certification 0.09 SD (read); Relative effect of teacher background: instructional practices have higher effect than on student achievement *Influence of teacher education on instructional practices not tested!*

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Phillips (2010) [USA]	Investigates the relationship between teacher quality and student achievement DV: cognitive assessment in math/reading (IRT scaled)	VAGM with school fixed effects (ECLS-K study; 435 schools, 1126 teachers, 4021 students (math), 431 schools, 1078 teachers, 3897 students (reading); 1999/2000; wave 3/4)	Certification status (dichotomous; full/other) Subject-specific degree (dichotomous; elementary) Undergraduate education (dummy-coded: elementary, education related, not education related, no BA) Graduate education (dummy-coded: elementary, education related, not education related, no graduate degree) Course taking patterns (Amount of college courses taken; ranging from 0-6) (Student level)	Course taking pattern	School fixed effects Dataset	Limited associations of teacher education variables and student achievement growth: Full certification -0.16 SD (math); Teacher course taking 0.14 SD (math; but only in the at risk sample); Graduate degree in elementary education +0.11 SD (reading); Results driven by at-risk-students sub-sample
Rivkin et al. (2005) [USA]	Disentangling the impact of schools and teachers in influencing achievement DV: TAAS test scores (math/reading)	VAGM with student, school, school by grade, and school by year fixed effects (Administrative data - UTD Texas school project; three cohorts, each cohort approx. 200000 students in 3000 elementary and middle schools; grades 3-7)	Proportion with graduate degree (percentage) (School level)	---	Student fixed effects School fixed effects	Little or no evidence that degree raises quality of teaching (i.e. student achievement)
Rockoff et al. (2011) [USA]	Investigation of non-traditional teacher characteristics on teacher effectiveness DV: Standardized test scores (math)	VAM with school zip code and grade level fixed effects (1); teacher level regression with school zip code fixed effects (indirect) (Administrative data and online survey; elementary teachers (grade 2-4): 418 respondents, 184 non-respondents, 4275 not invited to the survey; 82977 student-teacher observations; 2006/2007)	Certification status (dummy coded) Licensure test scores Advanced degree (dichotomous; yes/no) College major (dummy coded) SAT verbal/math scores Barron's college ranking (categorical) Cognitive ability Math knowledge for teaching, (Student level)	SAT scores Barron's college rank	School/zip fixed effects Dataset	Few traditional credentials with significant effects; Non-traditionals: Math knowledge and personal efficacy significant at 10 percent level; Factor analysis: Cognitive and non-cognitive factors significantly related to student achievement (+0.024/0.025 SD)

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Santibanez (2006) [MEXICO]	Exploration of the relationship between teacher test scores / characteristics and student achievement DV: Average (class-level) student achievement	VAM with variables aggregated on classroom level following a first-differences specification eliminating all fixed or time-invariant variables (Administrative data of the CM program; grades 3-9, 25000 elementary and 10000 secondary school teachers; 1996-2001)	Degree (percentages; no degree, BA (Pasante), Normal Basica, Normal licenciatura, MA/PhD) Teacher test scores (percentage correct) (Classroom level)	---	First differences spec. Dataset	Weak positive relationship between teacher test scores and student achievement at primary level (+0.03 SD); stronger relationship at secondary level (+0.25 SD; score in content/methodology section +0.18/0.21 SD); Advanced degree with a negative influence on student achievement at primary level; non-certified teachers have a negative influence at secondary level)
Sass et al. (2012) [USA]	Comparison of teacher productivity in high/low poverty elementary schools DV: FCAT-NRT (FL) and end of grade exams (NC) (math/reading)	VAM using Blinder-Oaxaca decomposition with student fixed and teacher-school type effects (Administrative data;of Florida, North Carolina; 9000/8000 elementary school teachers in Florida/North Carolina; 2000-2005)	Share of teachers with advanced degree (percentages) Certification status (percentages; NBPTS) Licensure (percentages; regular) PRAXIS scores (School level)	---	Student fixed effects Sample restrictions School level analysis	Differences in observed teacher characteristics account only for modest proportion of the variance of teacher value added (25%); Certification status has a significant influence; Differences in teacher quality across school types primarily due to unmeasured characteristics
Slater et al. (2012) [UK]	Estimation of the effect of individual teachers on student outcomes and the variability of teacher quality DV: GCSE and Keystage 3 scores (math/science/english)	VAM with prior achievement (Keystage 3) including student, teacher, and school fixed effects; regression of teacher effects on teacher observables (indirect) (Administrative data; 7305 students, 740 teachers, 33 schools; 1999-2002)	Degree (dummy coded; first class, second class) Subject-specific major (dummy coded; math, English, math, science, social science) (Student level)	---	Point-in-time student fixed effects Dataset	Variability in teacher effects with an SD of 0.610 GCSE points (effect size 0.23 SD); Teacher education variables do not play a significant role in explaining teacher effectiveness

Authors (Year) [Country]	Purpose of study	Method (Data)	Conceptualization of teacher education	Selection problem	Non-random allocation	Main findings
Winters, Dixon, & Greene (2012) [USA]	Examination of the relation between teacher characteristics and student proficiency gains DV: FCAT test scores (math/reading)	VAM in a linear probability specification including a teacher-school spell fixed effect (1) and a measure indicating the likelihood for a teacher to be found in the sample; (2) regression of spell effect on time-invariant teacher characteristics and school fixed effects (Administrative data; grades 4 and 5 (elementary); (1) 176359 observations (math), 178700 observations (reading); (2) 4862 math, 6571 reading observations; 2000-2004)	Subject-specific certification status (dummy coded; English, math, professional education, science, elementary education, special education, English as second language, other) Number of course credits (Reading, reading pedagogy, math; math pedagogy, behavior management, research, internship, curriculum and assessment, other, other pedagogy, other education) Advanced degree (binary; yes/no) (Teacher level)	IMR estimation to account for sample selection Number of course credits	Teacher/school spell fixed effect Multiple teacher observations	Math: Advanced degree and certification status has little or no relation to student achievement; pedagogic course credits positive influence, behavior/management and curriculum/assessment negative; Reading: Advanced degree without influence, certification to teach English as second language negative influence; course credits similar to math results
Woessmann (2003) [INTERNATIONAL]	Estimation of the impact of institutional arrangements of different schooling systems on student performance DV: math and science achievement scores (IRT scaled)	VAM with clustering robust linear regression and WLS estimation (TIMSS1995 data; 266545 middle school students, 39 countries; 1995)	Degree (Three indicator variables: secondary, BA, MA) (Student level)	---	Cross-country analysis	Teacher level of education positively related to student performance in science and math (15.682/10.571/25.576 Secondary/BA/MA in math; 24.243/12.378/32.106 Secondary/BA/MA in science)
Zhao et al. (2011) [CHINA]	Development and test of a model to explain math learning performance including school, class, and student level variables DV: performance test math (IRT scaled)	MLM (HLM) with three levels (students, classrooms, schools); (Administrative data; multistage stratification sampling, 5 provinces, 197 teachers, 253 classrooms, 10959 elementary students grades 1-6)	Type of teacher education institute of the teacher (dichotomous; 0 = lower level, 1 = higher level) (Classroom level)	---	Sampling Dataset	Teachers' graduation school level significantly related to student achievement gains; after inclusion of student level variables teacher education no longer significant

Appendix B

Interview guide and category system for the interview study (Chapter 3).

	Primary Question	Aspects	Type	Category	Control question
0	**Before we start with the actual interview, do you have any questions regarding the model or its components?**				
1a	The first part of the interview deals with the model and its primary components. In the model, teacher education is conceptualized as a part of the education system. **Is this a valid conceptualization of the relation of teacher education and the education system?**	Effectiveness Mechanisms Core function	K/O	Val-I	
1b	**Has the role of the selection and allocation functions within this relation been made clear?**		K/O	Val-I	Role reasonable?
1c	Both functions are described by three primary functional dimensions. **Are both functions sufficiently characterized by these dimensions?**	Selection Allocation	K/O	Val-I	Further dimensions?
2a	Now we are already in the second part of the interview which deals with the structural elements of both functions. The structural elements are a result of a theory review. Based on their theoretical meaning, the elements have been assigned to one of the six functional dimensions. **Is the mapping of structural elements and dimensions valid?**	Accessibility/Capacity Information (gen/spec) Integration (gen/spec)	K/O	Val-I	Changes?
2b	If we look at each dimension separately: **How do you rate the relevance of the respective structural elements for the selection of student teachers?** **How do you rate the relevance of the respective structural elements for the allocation of beginning teachers?**	Accessibility/Capacity Information (gen/spec) Integration (gen/spec)	O	Rel-E	
2c	**If you had to weigh the structural elements according to their relevance for the respective dimension, which orders would result?**	Accessibility/Capacity Information (gen/spec) Integration (gen/spec)	O	Rel-E	
2d	**Do you know structural elements (teacher education and context) which are empirically found to be relevant for the selection and allocation of student teachers and beginning teachers, but are not included in the model?**	Selection Allocation	K	Rel-E	Source? Mapping?
3a	The third and last part of the interview deals with the operationalization of the structural elements. I developed for each structural element an indicator to be able to transform the theoretical model into a measurement model. **To what extent do the indicators of the structural elements of both functions reflect their theoretical meaning?**	Selection Allocation	K/O	Val-II (AS)	Characteristics (+/-) Theoretical meaning preserved? To what extent are structural aspects captured?
3b	Some indicators combine two structural elements. **Is a combination of two elements by a single indicator justified?**	Status + Career Ladders A + A Procedures/Criteria Certification + Licensure School Autonomy + Job Descr.	K/O	Val-II (AS)	Are the theoretical meanings preserved despite the combination?
3c	**Do you know any empirically validated indicators measuring some of the structural elements?**		K	Val-II (AS)	Source? Which elements?

Note. Val-I = Validity I (General propositions); Rel-E = Reliability of the characterization of the functions; Val-II = Operationalization and alternative structural elements (completeness); Type: K = Knowledge, O = Opinion

Appendix C1

Country-specific descriptive statistics of the test of the first part of the model (Chapter 4).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Altruistic-pedagogical orientation																				
AP1 (MFA011G)	1.00																			
AP2 (MFA011E)	.453	1.00																		
AP3 (MFA011H)	.530	.506	1.00																	
AP4 (MFA011D)	.438	.666	.479	1.00																
Extrinsic orientation																				
E1 (MFA011B)	.124	.107	.026	.139	1.00															
E2 (MFA011F)	.246	.117	.116	.189	.159	1.00														
E3 (MFA011I)	.148	.130	.237	.182	.250	.212	1.00													
Subject-specific learning activity																				
SS1 (PART)	.214	.316	.326	.243	.107	.116	.053	1.00												
SS2 (READ)	.315	.241	.280	.225	.096	.197	.102	.403	1.00											
SS3 (SOLV)	.149	.084	.136	.047	.030	.104	-.038	.271	.341	1.00										
Pedagogical learning activity																				
P1 (PRAC)	.108	.213	.171	.089	.185	-.030	.008	.456	.358	.433	1.00									
P2 (PLAN)	.235	.213	.271	.102	.121	.108	.012	.564	.379	.448	.708	1.00								
P3 (AUSE)	.180	.176	.178	.065	.156	.169	.047	.481	.297	.313	.691	.660	1.00							
P4 (APRAC)	.245	.199	.178	.151	.151	.245	.003	.484	.351	.358	.510	.644	.601	1.00						
Controls																				
Sex	-.145	-.332	-.235	-.240	-.079	-.099	-.010	-.169	-.057	-.192	-.172	-.162	-.079	-.103	1.00					
GPA	-.027	-.081	-.156	-.222	-.028	-.103	.003	-.065	-.069	-.133	-.085	-.066	-.083	-.034	.049	1.00				
Family	-.103	-.036	-.031	.056	.079	.100	.020	-.043	-.058	-.008	.013	-.072	-.054	-.014	.026	-.171	1.00			
Money	-.002	.007	.011	.052	.078	.031	.039	-.043	-.050	-.007	-.183	-.103	-.108	-.096	.098	-.059	.006	1.00		
Job	-.112	.045	.018	.087	.049	.132	.073	.020	-.073	.035	.002	.001	.025	-.050	-.060	-.084	.165	.319	1.00	
Commitment	-.251	-.483	-.365	-.441	-.182	.014	-.178	-.227	-.188	-.170	-.232	-.140	-.095	-.089	.278	.135	.046	-.032	-.072	1.00
Mean	2.055	2.883	2.521	2.568	1.373	1.189	2.189	2.452	1.875	2.701	2.599	2.479	2.299	2.507	1.258	2.560	1.819	1.831	1.709	2.580
Standard Deviation	.871	.981	.984	.986	.618	.477	1.059	.556	.679	.632	.659	.644	.773	.671	.437	.874	.384	.375	.453	1.123

Note. Descriptives for Poland; $N = 290$.

Appendix C1

(continued)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Altruistic-pedagogical orientation																				
AP1 (MFA011G)	1.00																			
AP2 (MFA011E)	.495	1.00																		
AP3 (MFA011H)	.411	.385	1.00																	
AP4 (MFA011D)	.377	.476	.316	1.00																
Extrinsic orientation																				
E1 (MFA011B)	.030	-.011	.144	.062	1.00															
E2 (MFA011F)	.050	.059	.165	.170	.412	1.00														
E3 (MFA011I)	.116	.068	.237	.206	.477	.412	1.00													
Subject-specific learning activity																				
SS1 (PART)	.173	.120	.174	.127	.037	.077	-.008	1.00												
SS2 (READ)	.141	.124	.178	.142	.147	.080	.171	.287	1.00											
SS3 (SOLV)	.092	.086	.038	.140	.029	.021	.041	.225	.357	1.00										
Pedagogical learning activity																				
P1 (PRAC)	.193	.138	.155	.218	.075	.101	.072	.318	.430	.439	1.00									
P2 (PLAN)	.195	.131	.126	.149	.047	.096	.066	.326	.427	.344	.661	1.00								
P3 (AUSE)	.158	.142	.213	.100	.045	.104	.057	.410	.375	.370	.620	.582	1.00							
P4 (APRAC)	.126	.145	.170	.186	.116	.095	.112	.277	.398	.355	.566	.577	.538	1.00						
Controls																				
Sex	-.011	.013	.002	.013	.024	.069	-.108	.135	-.021	-.018	-.138	-.134	-.083	-.093	1.00					
GPA	-.060	-.148	.038	-.148	-.125	-.086	-.017	-.065	-.117	-.122	-.109	-.045	-.031	-.137	-.136	1.00				
Family	.053	.057	.051	.057	.004	.042	.037	-.106	.023	.000	.009	-.033	-.014	-.012	-.064	-.031	1.00			
Money	-.056	-.008	-.068	-.008	-.012	.060	-.042	.070	-.049	-.097	-.031	.025	-.001	.008	.051	-.026	.098	1.00		
Job	.111	.023	-.034	.023	-.013	.050	-.048	-.118	.108	.009	-.011	.016	-.015	-.046	.007	-.035	.155	-.007	1.00	
Commitment	-.229	-.237	-.243	-.237	-.114	-.051	-.161	-.077	-.122	-.115	-.106	-.127	-.140	-.093	-.082	.080	-.011	.010	-.014	1.00
Mean	3.093	3.221	2.796	2.861	1.811	2.018	2.195	2.973	2.173	2.779	2.857	2.820	2.783	2.876	1.510	2.497	1.795	1.998	1.979	2.225
Standard Deviation	.801	.810	.882	.831	.783	.848	.963	.407	.582	.500	.506	.489	.575	.493	.500	.959	.404	.045	.141	1.165

Note. Descriptives for Singapore; $N = 387$.

Appendix C1

(continued)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Altruistic-pedagogical orientation																				
AP1 (MFA011G)	1.00																			
AP2 (MFA011E)	.249	1.00																		
AP3 (MFA011H)	.410	.311	1.00																	
AP4 (MFA011D)	.104	.376	.311	1.00																
Extrinsic orientation																				
E1 (MFA011B)	.069	.054	-.028	.047	1.00															
E2 (MFA011F)	-.072	-.033	.121	.108	.273	1.00														
E3 (MFA011I)	-.026	-.063	.056	-.007	.439	.264	1.00													
Subject-specific learning activity																				
SS1 (PART)	.178	.090	.114	.119	.135	.136	.029	1.00												
SS2 (READ)	.166	.027	.102	-.038	.097	.038	-.020	.471	1.00											
SS3 (SOLV)	.069	-.065	.059	.033	-.037	.142	-.021	.175	.282	1.00										
Pedagogical learning activity																				
P1 (PRAC)	.185	.061	.026	.126	.105	.188	-.049	.409	.512	.462	1.00									
P2 (PLAN)	.265	.094	.101	.116	.132	.150	-.038	.443	.424	.227	.710	1.00								
P3 (AUSE)	.213	.073	.063	.083	.094	.165	-.041	.355	.401	.303	.680	.731	1.00							
P4 (APRAC)	.160	-.031	.041	.032	.011	.113	-.082	.376	.454	.392	.689	.713	.684	1.00						
Controls																				
Sex	-.161	-.160	-.244	-.113	.066	.028	.044	-.023	-.067	.178	-.029	-.132	-.107	-.064	1.00					
GPA	.041	.056	.024	-.059	-.068	-.044	.083	.032	-.100	-.060	-.055	-.007	-.046	-.068	.098	1.00				
Family	.097	.032	.126	.073	-.076	-.027	-.020	.000	-.005	-.004	.050	.042	.018	.052	.066	.009	1.00			
Money	-.025	.003	.004	.070	-.023	.006	-.134	.003	-.088	-.082	.043	.063	.065	.050	-.033	-.008	.322	1.00		
Job	-.028	.016	-.053	.051	-.014	.072	-.100	.010	-.016	-.002	.092	.089	.062	.074	-.071	-.023	.131	.330	1.00	
Commitment	-.101	-.087	-.105	-.127	.163	-.024	.014	.033	-.006	-.184	.025	.053	-.041	.031	.035	-.065	-.086	-.010	-.056	1.00
Mean	3.514	3.705	3.080	3.485	2.133	1.366	2.621	3.049	2.762	2.658	3.054	3.096	2.897	3.155	1.207	2.300	1.728	1.627	1.442	1.509
Standard Deviation	.636	.582	.894	.670	.932	.618	.989	.548	.714	.752	.629	.612	.763	.636	.405	.996	.445	.493	.497	.840

Note. Descriptives for the US; N = 486.

Appendix C2

Unstandardized parameter estimates of the measurement part of the final model (Chapter 4).

	Factor Loadings				Residual Variances		
Observed Variable	Altruistic-pedagogical orientation	Extrinsic orientation	Subject-specific learning activity	Pedagogical learning activity	Poland	Singapore	US
AP1 (MFA011G)	1.000				.436	.425	.298
AP2 (MFA011E)	1.066				.506	.448	..222
AP3 (MFA011H)	1.316				.413	.428	.583
AP4 (MFA011D)	1.112				.546	.434	.325
E1 (MFA011B)		1.000			.240	.230	.440
E2 (MFA011F)		.478			.202	.577	.308
E3 (MFA011I)		.925			.957	.588	.625
SS1 (PART)			1.000		.165	.128	.177
SS2 (READ)			1.387		.303	.216	.263
SS3 (SOLV)			.973		.286	.181	.482
P1 (PRAC)				1.000	.152	.087	.125
P2 (PLAN)				.992	.109	.089	.089
P3 (AUSE)				1.139	.246	.143	.186
P4 (APRAC)				.942	.188	.116	.132

Note. All factor loadings significant at $p < .01$. Factor loadings of the first indicators of each latent construct fixed to 1 for model identification purposes.

Appendix D1

Country-specific descriptive statistics of the 2009 final model (Chapter 5).

	1	2	3	4	5	6	7	8	9	10	11	12
Components of postive matching												
QUAL	1.00											
TCSHORT	-.405	1.00										
ESCS	.136	.001	1.00									
School-related controls												
SCEDU	-.047	-.063	.096	1.00								
COMP	-.078	.071	-.540	.016	1.00							
SIZE	-.063	.437	.280	.003	-.262	1.00						
TYPE	-.013	-.005	.066	.024	-.210	-.026	1.00					
ADAC	.066	.057	.386	-.042	-.445	.452	.365	1.00				
ADAS	.095	-.091	.065	-.117	-.207	-.277	.074	-.175	1.00			
TRAS	.121	.199	-.072	-.053	.051	.033	-.124	-.129	.222	1.00		
TRBE	-.106	-.029	-.096	-.007	.031	-.155	.103	-.001	-.089	-.022	1.00	
TRAC	-.024	-.114	.008	-.010	-.031	-.063	.412	.194	-.026	-.238	.353	1.00
Mean	-.002	.043	.355	.070	-.084	1.813	.002	.086	-.130	-.026	-.012	.003
Standard Deviation	.158	.682	.292	.748	.951	9.960	.207	.735	.868	.769	.479	.482

Note. Descriptives for Finland; N = 4591, J = 156.

	1	2	3	4	5	6	7	8	9	10	11	12
Components of postive matching												
QUAL	1.00											
TCSHORT	-.376	1.00										
ESCS	.307	-.428	1.00									
School-related controls												
SCEDU	.281	-.509	.300	1.00								
COMP	-.077	-.020	.249	-.060	1.00							
SIZE	.200	-.145	.093	.147	.087	1.00						
TYPE	-.324	.093	.189	-.147	.221	-.495	1.00					
ADAC	-.128	-.148	.371	-.108	.296	-.077	.452	1.00				
ADAS	-.076	.100	-.070	-.248	.153	-.291	.285	.394	1.00			
TRAS	-.076	.062	-.020	-.068	.038	-.266	.211	.123	.248	1.00		
TRBE	.018	-.029	.186	-.075	-.082	-.077	.166	.068	-.180	-.107	1.00	
TRAC	-.022	.077	-.040	.115	-.095	-.038	-.114	-.140	-.058	-.020	-.316	1.00
Mean	.003	.011	-.390	-.005	-.004	.024	-.022	-.012	-.011	-.001	-.001	.012
Standard Deviation	.170	.840	.476	.399	.916	3.633	.614	.679	.917	.632	.356	.690

Note. Descriptives for Singapore; N = 3051, J = 100.

Appendix D2

Country-specific descriptive statistics of the 2012 final model (Chapter 5).

	1	2	3	4	5	6	7	8	9
Components of postive matching									
QUAL	1.00								
TCSHORT	-.096	1.00							
ESCS	.182	-.085	1.00						
School-related controls									
SCEDU	.082	-.079	-.044	1.00					
COMP	-.118	.130	-.566	-.104	1.00				
SIZE	.408	-.209	.632	.051	-.508	1.00			
TYPE	-.022	-.034	-.245	-.035	.271	-.147	1.00		
ADAC	.182	.055	.070	.157	-.199	.157	-.181	1.00	
ADAS	-.205	-.072	.122	.114	-.207	.114	-.001	-.005	1.00
Mean	-.031	-.020	.261	.036	.030	-.437	.004	.050	.008
Standard Deviation	.173	.686	.288	.711	.862	2.123	.164	.774	.811

Note. Descriptives for Finland; N = 7477, J = 263.

	1	2	3	4	5	6	7	8	9
Components of postive matching									
QUAL	1.00								
TCSHORT	-.168	1.00							
ESCS	.436	-.210	1.00						
School-related controls									
SCEDU	.056	-.443	.079	1.00					
COMP	.016	.082	.159	.023	1.00				
SIZE	-.013	.075	.328	.106	.312	1.00			
TYPE	-.415	.314	-.359	.070	.089	.357	1.00		
ADAC	-.017	-.080	.349	.047	.071	.247	-.070	1.00	
ADAS	-.089	.045	-.110	-.122	-.008	-.091	-.104	.186	1.00
Mean	.001	-.007	-.238	-.013	-.004	-.265	-.039	-.009	.000
Standard Deviation	.045	.834	.466	.834	.259	4.865	.672	.596	.748

Note. Descriptives for Singapore; N = 5297, J = 164.

Appendix D3

χ^2-values and the degrees of freedom of the independence models (Chapter 5).

Model	χ^2	df	Scaling Correction Factor
2009			
Independence Model	478.036*	132	1.6250
2012			
Independence Model	349.397*	72	2.1967

Note. * = p < .001

Acknowledgements

First of all I want to thank Regina Mulder most sincerely for her constant support, guidance, and constructive feedback throughout the years. I further want to thank the experts who participated in the interview study for their helpful and insightful comments and suggestions. Special thanks go to all of my colleagues, especially Franziska, Veronika, Gerhard, Birgit, Susanne, Sebastian, Martin, and Willi, for the lively discussions and the laughs we had at work, conferences, and various fall schools.

I owe many thanks to my parents and my sister Verena, not only for believing in me, but also for the generous funding. Without you this work would not have been possible!

Patrizia, your help was beyond words! Thank you so much for your patience, your endurance, and for keeping me going when the going got rough.

Last but not least, thank you Amy. You might not be aware, but somehow you started all of this.

I dedicate this work to my grandparents. Sadly it took too long.