

IZA DP No. 9082

Is Ability Tracking (Really) Responsible for Educational Inequalities in Achievement? A Comparison between the Country States Bavaria and Hesse in Germany

Hartmut Esser
Ilona Relikowski

May 2015

Is Ability Tracking (Really) Responsible for Educational Inequalities in Achievement? A Comparison between the Country States Bavaria and Hesse in Germany

Hartmut Esser
*University of Mannheim
and IZA*

Ilona Relikowski
University of Bamberg

Discussion Paper No. 9082
May 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Is Ability Tracking (Really) Responsible for Educational Inequalities in Achievement? A Comparison between the Country States Bavaria and Hesse in Germany

It is still taken for granted that (early) ability tracking increases the impact of social origin on achievement in (lower) secondary education, but without gains in the overall level. This contribution addresses the question of whether this common conviction is really correct. The various deviations and inconsistencies obtained from analyses that use other approaches and data bases form the starting point. On the basis of a general theoretical model, the Model of Ability Tracking, we specify the preconditions for identifying the effects of ability tracking. These include considering the school level as well as cognitive abilities prior to ability tracking at the end of elementary school. Both conditions aren't included in common analyses using PISA data. As a consequence, effects of social origin have been systematically overestimated and those of cognitive abilities haven't been detected in the respective studies at all. Because PISA data are lacking information on cognitive abilities in the institutional sorting at the end of primary school and no other appropriate data set to compare educational systems is available, these assumptions will be tested with another data base: the BIKS-study. This study allows using the different levels of strictness of the institutional rules concerning ability tracking in the two country states Bavaria and Hesse in Germany. The results support the presumptions of the Model of Ability Tracking: If school effects on the one hand and cognitive abilities on the other hand were taken into account, all effects of a reinforcement of social origin disappear and increases in school effects of abilities on achievement are observed in Bavaria, the country state with an especially strict rule for ability tracking. Applying the misspecifications of the other approaches to these data, one again obtains their misleading findings, and they disappear by approaching the analyses to the specifications of the Model of Ability Tracking.

JEL Classification: I20, I24, I28

Keywords: educational systems, ability tracking, educational inequality, education and social origin, cognitive ability and educational achievement, school-effects, PISA-studies

Corresponding author:

Hartmut Esser
Universität Mannheim, MZES
68131 Mannheim
Germany
E-mail: hartmut-esser@t-online.de

1. The problem

There is hardly a country without social (and ethnic) educational inequalities. Yet, there are substantial differences between the countries in terms of these inequalities. From the very beginning, the various hypotheses on the reasons for these differences have pointed to the effects of the educational systems. Educational systems refer to the institutional rules and organizational design of a certain society's or region's educational establishments. Of major interest is the question on the effects of *differentiation* and a specific *institutional* sorting on the basis of cognitive abilities or educational achievement at the end of elementary school. The reason for this is the presumption that a homogenization of learning environments has certain advantages. These include a better adaptation of curricula and lessons and perhaps even a cumulative increase in the learning success due to mutual contacts between pupils having similar talents, predispositions, and interests. The counter argument states that less talented pupils could no longer benefit from the more favorable learning environment of an integrated school and would possibly perform worse due to an assignment to lower valued types of schools entailing stigmatizations and a negative self-perception. This applied particularly in view of the assumption that sorting according to achievement almost inevitably involves *stratification*, i.e. the systematically different distribution of pupils according to their social origin to rather more general-academic in contrast to more specific-vocational school types, which clearly differ in terms of their social composition, the quality of their equipment, and their societal evaluations. For quite some time scholarly and public debates have agreed on the point that differentiation according to school tracks was indeed associated with social (and ethnic) educational inequality and that the expected increases in achievement tended not to exist. The results of most contributions and summaries indeed point in this direction. However, there are several exceptions and inconsistencies, which cannot be explained solely by common variations in the approaches used, the data and statistical methods.

This contribution aims at clarifying these anomalies, contradictions and inconsistencies. It starts with a summary of the most important approaches and empirical results on the effects of differentiation of educational tracks according to achievement and social origin (chapter 2). In chapter 3 these approaches and empirical results will be systematized within the framework of a comprehensive theoretical model, the "Model of Ability Tracking". This model particularly clarifies the prerequisites for identifying and separating effects of differentiation as compared to stratification and integration of an educational system. In this way it provides the basis for

a more explicit and systematic theoretical specification of empirical analyses, also in comparison to previous approaches with their partly quite inconsistent results (chapter 4). Because the internationally comparative studies on which most analyses are based do not provide the required data for such a test, a different approach has to be chosen. This is done using data from BiKS, a panel study on educational pathways in elementary school and subsequent secondary school in Germany. These data allow for a comparison on the basis of the theoretical model between Bavaria and Hesse – two representative country states of the Federal Republic of Germany with different rules according to which differentiation is organized and which vary (substantially) in their degree of strictness. In chapter 5 we will describe the data used and the approach. The results of the analyses done on the basis of the Model of Ability Tracking and in comparison to the findings of previous approaches will be reported in chapter 6. In chapter 7 we will shortly evaluate the main results within the framework of (ongoing) scholarly and public debates.

2. Approaches and findings

Indeed there seems to be ample empirical evidence for the impression that differentiation according to achievement reinforces educational inequalities without increasing the performance level (cf., *inter alia*, Alba et al., 2011: 401ff.; Ammermüller, 2005; Bauer and Riphan, 2006; Betts, 2011: 368, 377; Horn, 2009, 2013; Hanushek and Wößmann, 2006; Hanushek and Wößmann, 2011: Chapter 4.4.4; Meier and Schütz, 2007: 19ff., 24; Müller and Kogan, 2010: 225ff.; Vandenberghe, 2006: Chapter IV; van de Werfhorst and Mijs, 2010: 415ff.; Wößmann, 2009: 27ff.; Wößmann et al., 2009: Table 7.A2). Virtually all contributions are based on the distinction between *two* levels: the individual level of children and families on the one hand and the level of countries and the respective educational systems on the other hand. The level of schools and school classes is not considered any further. In addition, mostly used PISA data are lacking information on prior abilities and performance development at the end of elementary school. We refer to this approach as Standard Approach and to the result obtained using this approach as *Standard Result*. There are, however, a couple of findings deviating from the Standard Result and even partly clear contradictions. They particularly occur when another theoretical and methodological approach than the Standard Approach is applied and insofar they do not represent merely the not uncommon, often idiosyncratic anomalies in replications.

Already the simple *OECD reports* on the PISA studies containing descriptive results for the meanwhile available data of five surveys conducted between the years 2000 and 2012 suggest that the relation between differentiation and stratification in terms of achievement is more complex than is often assumed. There was a convergence of the level of achievement and social permeability between the years 2000 and 2012, although there had been no (visible) substantial institutional changes, in particular not in the countries with integrated educational systems (OECD, 2014). In fact, Canada and Finland, for example, with their strict integrated educational systems remained also in the year 2012 the two countries displaying the (altogether) highest achievement levels and the highest social permeability. However, simple inferences are still difficult: In the year 2012 achievement levels in Germany and Switzerland, for instance, were substantially higher than in Sweden with its integrated educational system in the year 2000, whereby the achievement level in Sweden had decreased over the years without any discernible institutional changes in terms of institutional sorting.

Contributions with multivariate analyses also provide evidence for the hypothesis that other processes and conditions than differentiation as compared to integration are responsible for the cross-country differences. If controlled for, it becomes apparent that differences in inequalities on the basis of social origin between differentiated and integrated systems already exist *prior* to sorting and that there is subsequently no special own effect of sorting (cf. Waldinger, 2006 or Ruhose and Schwerdt, 2015 for a replication of the Standard Result considering effects of social origin for elementary schools using data of PIRLS; Merry, 2013: 234ff. for a similar comparison between the USA and Canada). Differences in early preschool attendance, i.e. previous performance development, obviously play a major role (Schlotter, 2011: Table 5). This corresponds to the (admittedly: again only descriptive) finding that pupils already perform better at elementary school in the country states of Bavaria and Baden-Württemberg in the Federal Republic of Germany with their particularly strict rules of differentiation (in 2009), especially as compared to the city states Berlin, Bremen and Hamburg (Böhme and Weirich, 2012: Table 5.1), and that this apparently merely continues in secondary school (Schipolowski and Böhme, 2010: Table 5.1; Esser, 2013: Tables 1 and 2), even in comparison with the specific conditions in (other) big cities (Böhme and Weirich, 2012: Figure 5.3).

Different results than those of the Standard Approach will also become apparent if one considers *subsequent* developments: Using data of the International Adult Literacy Survey

(IALS), Brunello and Checchi (2007: Table 14) find a reduction of the effects of social origin on performance in terms of differentiation two years after compulsory education for the then seventeen year old pupils. With regard to other aspects of educational success, like educational attainment, drop out, and tertiary education, differentiation still reinforces the effects of social origin. Dustmann et al. (2012: Tables 2, 3 and 5) prove that misallocation of pupils whose performance is close to the cut point of the requirements of institutional sorting in terms of differentiation has no long-term impact on the labor market success, neither with regard to income nor to (full-time) employment and even not with regard to occupational status. One has also to keep in mind that stratifying effects not necessarily have to be connected with sorting itself, but with other characteristics of the school types, which are not necessarily linked to sorting, like quality of the equipment and teachers' motivation (Burgess et al., 2014: Chapter 5). In addition, all these results are subject to the reservation that the available (cross-sectional) data do not allow for drawing causal conclusions and that the various methods for controlling unobserved heterogeneity often require special assumptions.

This is the reason why results of *experiments* on the reorganization of a differentiated system into an integrated one are particularly revealing (Betts, 2011: 357ff., 371ff.). There have been several natural experiments on the occasion of changing the educational systems from differentiation to integration in Great Britain, Sweden, and Finland. Overall, the effects are rather weak or completely missing, partly contrary to the expectations, or they do not relate to educational achievement, but to educational attainment, to military aptitude tests, or future income on the labor market. In addition, in case of natural experiments in single countries it always remains unclear, whether changes were caused by unobserved other processes of social change. The only experiment, which avoids this problem via randomization, was conducted in Kenia. For this experiment Duflo et al. (2011) report a significant improvement of the achievement level in the schools differentiated according to ability, even after controlling for various covariates and with an at least one-year sustainability. This strongly suggests that differentiation (strictly) according to ability may after all have the intended positive effects, because the decisive confounding variable was already structurally controlled through the special situation in Kenia where children displayed only small differences in terms of social origin. Without doubt, one can question whether this result may be transferred to western industrialized countries. However, this might have little to do with the fact, that the experiment involved African children and teachers and that social inequality was not particularly pronounced.

In addition, one would only expect specific effects of homogenization according to achievement, if differentiation *indeed* involved a sorting *strict* by abilities and were *not* confounded by other characteristics like social or ethnic origin (cf. Betts, 2011: 353ff.). Controlling for the stringency of sorting according to ability, including the binding force of recommendations limiting parents' freedom of choice or selectivity based on achievement in terms of recommendations for the higher tracks, effects of social origin both on attainment and on achievement clearly diminish or even disappear completely (cf., *inter alia*, Bol et al., 2014; Dronkers, 2015; Dronkers and Korthals, 2015; Korthals, 2012; Marks, 2005, 2014; Verwiebe and Riederer, 2013).

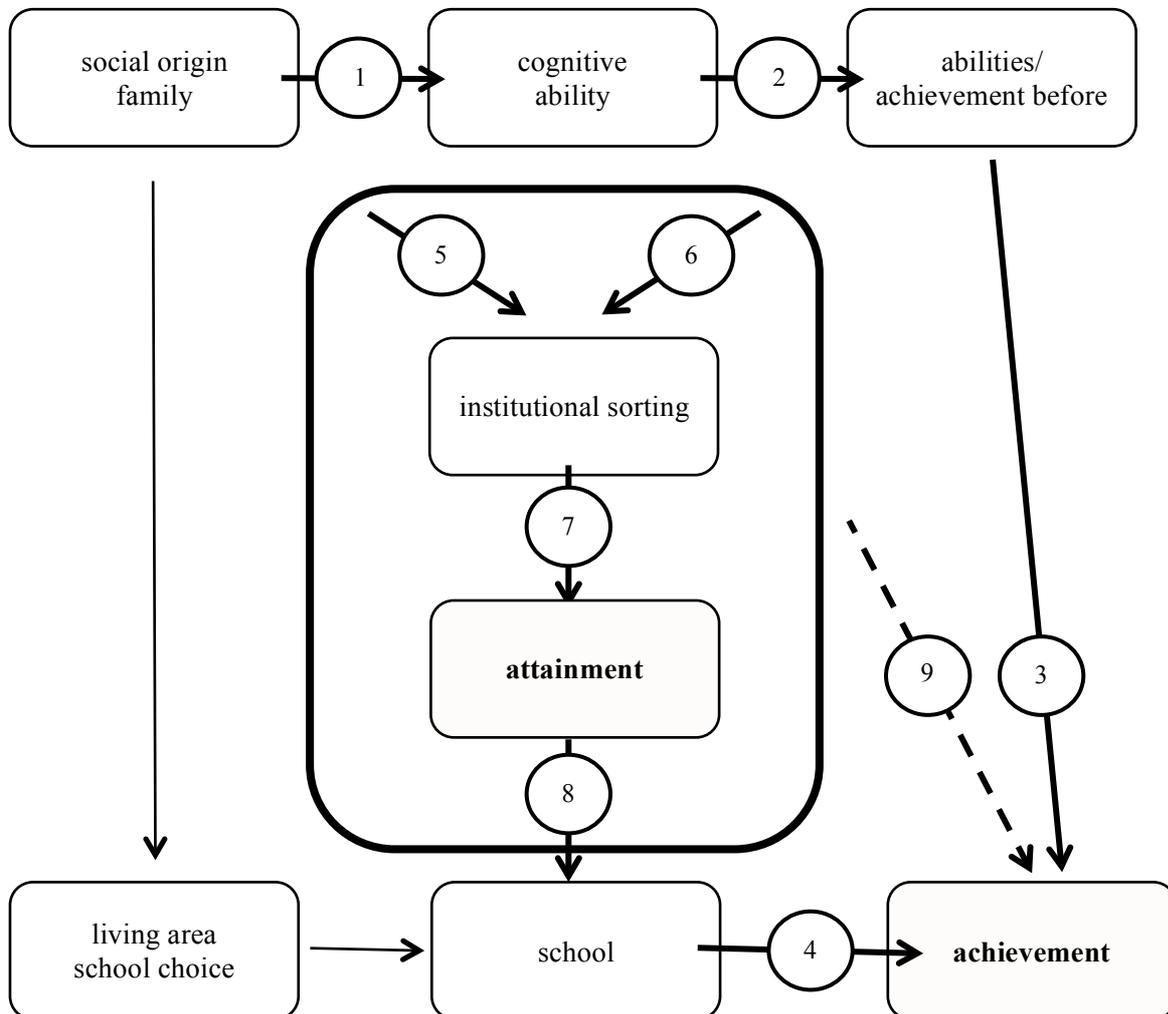
The most striking anomaly of the Standard Result appears, however, when one includes *schools and school classes* as an *own* level in the analysis in addition to the two levels of educational systems and children and their families of the Standard Approach, just as was done first by Dunne (2010) followed by Dronkers et al. (2011, 2012). Considering schools and school classes reveals clear effects of the proportion of higher status children within schools. Yet, at the same time one observes a *decrease* in the effect of *individual* social origin – and thus a *higher* social *permeability* for higher achievement in *differentiated* educational systems. In addition, the effect of social origin increases (slightly) with the proportion of higher class children in schools and school classes in a kind of Matthew-effect. This result has meanwhile been confirmed by other studies (Bol et al., 2014: 20ff.) and obviously applies also to children from families with migration background (Dronkers et al., 2011: Table I; Dronkers et al., 2012: Table I; Dronkers and Korthals, 2015; Teltemann, 2012: Chapter 6.2; Verwiebe and Riederer, 2013: Tables 3 and 4). Accordingly, differentiation does *not* exert its disadvantageous impact through the effects of *individual* social origin, but through social *segregation* within schools and school classes as a consequence of sorting. A combination of differentiation according to performance *and* social de-segregation of schools would thus possibly yield an optimum, especially for children from lower social strata. Stronger social barriers to performance development in secondary education in case of an integrated educational system as compared to a differentiated one would, however, contradict the Standard Result. To mark these differences and deviations, the approach considering *three* levels, including schools' composition, shall be referred to as *DVD Approach* and the quite counterintuitive finding shall be denoted *DVD-Puzzle* (as an acronym for the authors of the

respective contributions: Dronkers, Van der Velden and Dunne; cf. also Esser 2015a, 2015b, 2015c).

3. The Model of Ability *Tracking*

It would be a hopeless attempt to disentangle the various inconsistencies and anomalies of the Standard Approach by investigating all possible variations in the approaches, data, and analysis strategies. Figure 1 outlines (therefore) a theoretical model, i.e. the Model of Ability Tracking (MoAbiT), which systematically specifies the conditions and processes being important for school achievement development. It is designed to discern the differences between the various approaches, identify possible misspecifications, classify empirical findings, and resolve, as far as possible, the different anomalies and inconsistencies.

Figure 1: The Model of Ability Tracking (MoAbiT)



Differences in educational achievement in general, i.e. in principle in all kinds of educational systems, constitute the central explanandum. Early development of cognitive abilities within the families and the emergence of differences in cognitive abilities according to social origin constitute the starting point for all other processes (path 1). Then follows the development of early school-related skills, possibly already at preschool, and the development of educational achievement up to the end of elementary school (path 2). The most important individual conditions for the achievement in secondary education (and subsequent education) are the cognitive abilities developed in family, preschool, and elementary school as well as the achievement before sorting (path 3). The causal mechanism is the higher efficiency of learning for a given input for children with higher cognitive abilities and prior achievement (path 3): Students with higher cognitive abilities will translate the same learning content in a

better learning outcome. The context of the school in secondary education, then, constitutes the current institutional learning environment for the acquisition of skills and subsequent achievement. This possibly involves certain *school effects*, which are generally connected with differences in incentives and opportunities for learning in terms of the various characteristics of schools and school classes (path 4).

Paths 1, 2, and 3 then describe the influence of the family and the prior institutional impact on current achievement, and path 4 depicts the current institutional influence thereupon. Together they constitute the basic model for explaining achievement and the effect of social origin in this process. In addition, other influences on children's achievement and the schools' characteristics may be important. Examples are the respective living environment, the assignment to school districts, and the (possibly also targeted) parents' choice of place of residence and school.

The model assumes (implicitly) an integrated educational system: There is no specific division of educational tracks through regulations of the educational system. In case of a differentiation, *institutional* sorting into different school types and educational tracks becomes important. This institutional sorting is programmatically, however not always actually, based on prior achievement at elementary school and the connected recommendations by the teaching staff: Depending on the specific rules of the educational system for differentiation, parents have the opportunity for an independent own decision on the further educational track (path 5). And depending on the concrete organization, the teaching staff can deviate from sorting strictly according to achievement. For example, they may have stereotyped expectations and efforts based on children's social origin with regard to their achievement, but above all, with regard to the assessment of the achievement in form of marks and recommendations (path 6). In addition to the indirect impact of the so-called *primary effects* via the development of cognitive skills and school-related achievement prior to institutional sorting (paths 1 and 2), one can specify two more effects of social origin for the generation of differences in institutional sorting based of children's social origin: *secondary effects* of parents' educational decisions, which may deviate from the actual achievement and teachers' recommendations; and independently from that, the here so-called *tertiary effects* of the

teaching staff's stereotyped expectations, efforts, evaluations and recommendations and their effects on institutional sorting.¹

The actual sorting is then the combined result of two different processes. *First*: the *development* of children's *achievement* until the end of elementary school resulting from early *primary* effects on the one hand (path 1) and from subsequent additional institutional influences from pre- and elementary school on the other hand (path 2). *Second*: independent from that, *deviations* in the sorting process strictly according to the observable achievement at this point either due to *secondary* effects of parents' decisions (path 5) or to *tertiary* effects of teachers' socially biased expectations, efforts, and evaluations (path 6). Sorting *strictly* according to abilities and achievement (with given primary effects) would then more likely occur in case of weaker secondary and tertiary effects. In addition, the weights of the effects may vary according to the conditions and rules of the educational systems: via restricting or allowing parents' freedom of choice in terms of secondary effects and/or via higher or lower selectivity regarding the achievement by the *organization* of schools and instructions in terms of tertiary effects.

The sorting process completed in this way then determines, according to the administrative rules the differences in children's allocation to the various school types and their participation (path 7). From this results in turn a different composition of schools than in integrated systems as regards achievement and pupils' social origin (path 8) – varying in each case (with given primary effects) according to secondary and tertiary effects. The bordered boxes in Figure 1 summarize the processes in a differentiated educational system with institutional sorting according to achievement, which differ from those of an integrated system and from the basic model.²

¹ The denotation of these three effects follows and extends the well-known distinction between primary and secondary effects by Raymond Boudon for explaining educational inequalities in terms of educational *attainment* and *sorting* (Boudon, 1974: 20-39; for further background information and various types of this distinction: Jackson, 2013: 3ff.). Primary and secondary effects relate to the process of institutional sorting (according to achievement) through *families'* influences. The extension to 'tertiary' effects refers to *additional* effects of social origin on the sorting process via the *school* context: Do *teachers'* attitudes vary according to children's social origin, possibly with consequences for children's achievement itself, but particularly for the evaluations in terms of marks and recommendations based on achievement?

² Paths 1 and 2 and 5 to 7 in Figure 1 correspond to the model by Erikson et al. (2005: 9732) for explaining inequalities in educational *attainment*. Their model does *not* consider subsequent *achievement* in secondary school (on this cf. Jackson, 2013: 18f.).

The *total* effect of *differentiation* on achievement as compared to integration, thus, consists in an *indirect* effect of institutional sorting and the various school effects (paths 5 to 6 and path 4) and a possibly remaining *direct* effect (path 9). Effects of differentiation are just part of the total effect of an educational system, like, for example, differences in the *input* of schools with regard to expenditures and quality or differences with regard to schools' *organization* and instruction including, for example, control of and incentives for the teaching staff. This makes it necessary to control for the peculiarities of input and organization as well as for other (macro) conditions of the receiving countries, like the socio-demographic and economic situation, cultural traditions, and general educational policy (cf. Esser, 2015a: Chapter 2).

4. Implications and theoretical expectations

The Model of Ability Tracking reveals the complexity of identifying the special effects of differentiation in comparison to integration. This is particularly true, if one aims at separating the *stratifying* effects from those of *differentiation* as institutional sorting according to cognitive abilities and prior achievement. This includes five implications.

First: Institutional sorting according to cognitive abilities and prior achievement *alone* has *no* effects. It initially involves only a decomposition of the total variance in terms of cognitive abilities in one part within schools or school types and another one between them. In order to refer to effects of differentiation, there have to be effects of the composition of schools or school types *in addition to* the mere decomposition of the distribution.

Second: Effects of differentiation (in addition to a possible direct effect) follow the indirect path of the effects of institutional sorting within schools and school classes and subsequent school effects, particularly the *then* given composition according to cognitive abilities and social origin. Institutional sorting assigns children after elementary school to different school types and also to spatially separated schools and single school classes. This generates *more* homogenized learning environments with regard to abilities and social origin as compared to integration (path 8). With this homogenization and spatial separation one can expect that peer contacts between children having similar abilities and social origin will be more frequent and more stable and that this results in a *higher* exposure to learning-related interactions, which causes additional effects going beyond the mere distribution (path 4). The additional (system) effect in question would appear empirically as an *interaction* effect between differentiation

and the composition of schools and school classes according to cognitive abilities and social origin.

Third: Depending on the *strength* with which selectivity is based on achievement in institutional sorting, distribution according to achievement may be weakened and sorting according to other factors, like social origin, may be fortified: The stronger secondary and tertiary effects – with given primary effects – are, the less likely the homogenization in schools and school classes in terms of cognitive abilities, intended by institutional sorting, will occur. In addition, sorting could have a weaker impact through the school effect in question.

Fourth: School effects can relate to aspects of differentiation according to achievement and/or to those of stratification of schools and school classes in terms of the various *school types*: effects of homogenization of achievement levels on the one side, as compared to effects of the composition according to social origin, in terms of equipment, quality and schools' prestige as well as qualification and efforts of the teaching staff on the other side. Insofar as school types already differ with regard to the curricula, one would anyway expect differences in achievement: An exposure to more academic general contents as compared to more occupation specific contents will necessarily yield other learning results – even if everything else remains the same.

Fifth: Controlling for the mentioned indirect effect, one would expect *no* additional *direct* effect of differentiation (path 9 equals zero): All effects of differentiation are based on the homogenization of schools and school classes through sorting and the associated general school effects. If an additional direct effect would yet occur, it was due to other mechanisms, like anticipating the institutional sorting and the resulting presumably particularly early efforts in the families towards prior high achievement in elementary school or differences in the teaching staff's focus on achievement within schools.

The identification of the effects of educational systems and particularly the separation of the effects of differentiation from those of stratification, thus, necessarily requires three conditions: *First* the measurement of children's cognitive abilities and achievement prior to sorting, *second* the consideration of schools and school classes as an intervening level, and *third* that the effects of the composition of schools and school classes according to

achievement can be separated from those of stratification based on the composition according to social origin, school types, curricula, schools' quality, and teaching staff's efforts.

Table 1 summarizes the hypotheses on the (direct) effects of differentiation according to the MoAbiT that can be expected when these factors are controlled.

In terms of the *general* conditions for the acquisition of competencies in secondary school, the MoAbiT presumes positive effects of social origin due to the primary effects of the family, though they become less important in the course of school attendance and with the control for aspects of achievement (effect $SES \geq 0$). Positive effects would, however, occur in any case for cognitive abilities (effect $ABL +$). In terms of the effects of schools and school classes, there should be no effect of the composition according to social origin, if schools' quality and the teaching staff's efforts were controlled. In addition no interaction between individual origin and the respective school effects ($SSES$ and $SES \cdot SSES$ each = 0) should occur, because it would hardly change the composition according to achievement, and everything else would be controlled by inclusion of school's quality and the teaching staff's efforts in the analysis. Things would be different with regard to cognitive abilities. Here, one would expect positive school and interaction effects (both $SABL$ and $ABL \cdot SABL +$): The reinforcement of the effect of cognitive abilities in schools and school classes with higher proportions of talented children makes up the core of the MoAbiT's argument based on learning theories, which propose an increasing effect of the same exposure for higher abilities by a higher learning efficiency. Finally, the MoAbiT suggests an own positive effect for the *Gymnasium* as compared to the other school types – already because of its rather academic curriculum alone, but also because of the usually empirically observed better equipment and quality. The effect should, however, diminish and – finally – disappear completely, if these features were controlled ($GYM \geq 0$).

In terms of the *specific* effects of differentiation there should be *no* direct (additional) effect of educational systems (differentiation equal to 0) after the respective controls according to the MoAbiT. At most, indirect effects of homogenization and associated (additional) school effects mentioned above under points 2 and 3 could occur. In addition, neither changes in terms of the effects of individual social origin and homogenization of schools and school classes according to social origin are not expected ($SES \cdot Diff$ and $SSES \cdot Diff$ each = 0), nor changes in terms of individual cognitive abilities ($ABL \cdot Diff = 0$). One would expect,

however, an increase in the effects for the composition of schools and school classes according to cognitive abilities in case of a stronger orientation towards achievement in an educational system. The reason for this is that homogenization with regard to cognitive abilities within schools and school classes would then be more pronounced. In case of stronger secondary and tertiary effects which attenuate this homogenization one would expect this effect to decrease or disappear completely, so that one can assume a positive effect approaching zero in countries with a “liberal” institutional sorting ($SABL \cdot Diff \geq 0$).

Table 1: Specifications and theoretical expectations of the Model of Ability Tracking in comparison to empirical findings of other approaches to identify effects of differentiation (light grey-shaded): “---“ not considered by the approach

Conditions		Approaches			
Conditions/levels	Variables	Hypotheses MoAbiT	Findings OECD	Findings Standard	Findings DVD
specific effects: country/state					
educational system	Differentiation	≥ 0	≤ 0	≤ 0	-
·social origin	SES ·Diff	0	+	+	-
·Proportion social origin+ school	SSES ·Diff	0	---	---	+
·cognitive abilities	ABL ·Diff	0	---	---	---
·Proportion abilities+ school	SABL·Diff	≥ 0	---	---	---
general effects: Individuals					
social origin	SES	≥ 0	---	+	+
Ability	ABL	+	---	---	---
general effects: Schools/classes					
Proportion origin+ school	SSES	0	---	---	+
Origin·proportion origin+ school	SES·SSES	0	---	---	---
Proportion ability+ school	SABL	+	---	---	---
Ability·proportion ability+ school	ABL·SABL	+	---	---	---
School type/track	GYM	≥ 0	---	---	+

The hypotheses of the MoAbiT deviate partly significantly from so far available results (cf. chapter 2 above). However, nearly all approaches and analyses lack at least one of the preconditions for correctly estimating the respective effects. Regular *OECD reports* provide

little more than rough information on individual countries without any further systematic control of distributional effects, in particular with regard to prior achievement and social origin. Thus, they actually do not allow for drawing conclusions on effects of differences in the rules of institutional sorting according to achievement or social origin between the country states. The *Standard Approach* lacks the mediating level of schools and school classes. This is why the Standard Result can only refer to the *total* causal effect of differentiation and stratification, but not to the indirect effect of homogenization and school effects. In addition, they are based almost exclusively on PISA data, which lack information on cognitive abilities and prior achievement. The *DVD Approach* indeed considers school effects and thus makes an important step in the correct specification of empirical analyses. However, studies following this approach were also based on PISA data, so that the necessary information on children's achievement during sorting and the composition of schools and school classes according to prior achievement is missing.

For the comparison of the theoretical implications of the MoAbiT, Table 1 also contains typical empirical results of the three other approaches to analyzing the specific effects of differentiation (OECD reports, Standard Approach, and DVD Approach). With regard to all three approaches they only refer to social origin but not to cognitive abilities. Yet, the results of the DVD Approach additionally refer to the composition according to social origin, but also not to cognitive abilities. One immediately recognizes the incompleteness with respect to the preconditions for a correct specification as implied by the MoAbiT (cf. chapter 2 above and in more detail the overview in Esser, 2015a): OECD reports lack all multivariate controls, the Standard Approach lacks the school level and thus the decisive indirect effect of homogenization and school effects, and the DVD Approach, just like the others, lacks cognitive abilities and achievement prior to sorting.

5. Data base and operationalization

For the analyses, data from the study BiKS-8-14 were used, an educational panel conducted by the interdisciplinary DFG research group BiKS ('Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter' ['Educational Processes, Competence Development and Selection Decisions in Preschool- and School Age']) in Bavaria and Hesse.

The BiKS study concentrates on Bavaria and Hesse, as the two country states exhibit crucial differences in their educational systems. Following the typology by von Below (2011), Bavaria is to be classified as traditional-conservative, as the teaching content is clearly regulated by binding curricula and the structure is traditionally tripartite with little permeability between school tracks. The transition to secondary school is based on strict performance criteria as well as on a binding teacher recommendation. On the contrary, Hesse displays a reformed-liberal type, where parents are free to choose the preferred secondary school irrespective of prior school grades, comprehensive schools are widespread, changes between school types are more easily to perform, and the content of teaching is given more room for individual embodiment.

The BiKS panel followed children starting in third grade elementary school (2006) throughout secondary school. Besides competence tests in different domains, in each wave, paper-and-pencil interviews were conducted with the students, and additional information was gathered through telephone interviews with the parents and via teacher questionnaires. With the children's transition to secondary school in 2007, the class mates were additionally recruited in order to attain an institutional sample within the new school context (Schmidt et al., 2009). For the newly recruited children in secondary school, relevant information on their prior school career, such as final elementary school grades and teachers' recommendation, were surveyed retrospectively.

When children changed over to secondary school, the overall sample consisted of $N = 2,984$ cases in grade 5. Due to the sampling strategy in elementary school (von Maurice et al., 2007), the Bavarian sample (BY) constitutes a significantly larger proportion of 67.69 %. In Bavaria, 23.02 % of the sample attended the *Hauptschule* (lower vocational track), 20.99 % the *Realschule* (higher vocational track), 52.62 % the *Gymnasium* (academic track), and 3.37 % were distributed among other school types (or school is unknown). The respective distribution in Hesse was: 2.18 % *Hauptschule*, 6.64 % *Realschule*, 52.18 % *Gymnasium*, 31.12 % *Gesamtschule* (comprehensive schools), and 7.88 % other school types (or unknown).

The analyses comprise all students who participated at least once in the annual competence assessments from grade 5 to 7 and did not change schools or drop out due to refusal during this observation period. However, competence tests had only been conducted with students in the institutional sample, which was restricted to the three traditional school types *Hauptschule*, *Realschule*, and *Gymnasium*. Children attending other school types were

followed up individually³ and did not participate in competence assessments. This limitation results in a bias of the analytical sample in Hesse with a proportion of 91.83 % attending the *Gymnasium* (Bavaria: 57.42 %) due to its strong orientation towards comprehensive schools. Albeit unfavorable, the intention of the presented theory test is that of revealing causal effects and not the prediction of exact point estimates. When correctly specifying the model, the causal associations should be unaffected by the distribution of case numbers.

Applying multiple imputations of missing values due to item-nonresponse⁴, the total analytical sample across the three panel waves under study consists of 1,614 students with overall 4,314 observations and a mean of 2.7 observations per child. Per school class, on average 9.8 children participated in grade 5, 8.8 children in grade 6, and 7.2 children in grade 7.

The dependent variable of achievement (ACH) is given by a measure of reading comprehension in order to obtain best possible comparability to the PISA results. The tests of reading comprehension were developed by the BiKS research group and consist of text samples with multiple-choice items to be answered by the students. As different tests were implemented in each wave, a common item design with a nonequivalent groups/anchor-item test design was applied. Following the approach by Pfof and Artelt (2013)⁵, the competence test scores are given by weighted likelihood estimates (WLE), which were T-standardized (M=50, SD=10) in grade 5. All tests show satisfactory reliabilities (grade 5 = 0.78, grade 6 = 0.77, grade 7 = 0.76).

The explanatory variables are operationalized as follows: The degree of differentiation is given by the state dummy: Bavaria (BY) as the case of high differentiation in reference to Hesse (HE) as the case of liberal differentiation⁶.

Students' social origin (SES) is measured by their parents' highest ISEI score (International Socio-Economic Index cf. Ganzeboom and Treiman, 1996). The social composition of the school context (SSES) is given by the aggregated means of parental ISEI scores of all participating children per school class (time-varying).

³ Besides, when certain sampling criteria were not met and in case of non-participation of secondary schools and school changes, children were followed up individually as well (for details see Schmidt et al., 2009).

⁴ Multiple imputation of $m = 20$ datasets based on a broad set of auxiliary variables was conducted by using the MICE system of chained equations implemented in Stata. 30 cycles of regression switching were carried out.

⁵ We kindly thank Dr. Maximilian Pfof for the provision of the extensively prepared test data.

⁶ With comprehensive schools not being part of the institutional sample, the overall proportion of Hesse in the analytical sample decreases to approximately 20 %.

To measure students' general cognitive abilities (ABL), the matrices subtest of the culture fair intelligence test (CFT 20-R, German version: Weiß, 2006) was administered in each of the three waves and is included in the analyses as sum scores. As with the social composition, the measure of the average ability level in the school context (SABL) is the aggregated mean of all participating children per school class.

The secondary school type is captured by a dummy variable for *Gymnasium* (GYM = 1, *Hauptschule/Realschule* = 0).

Considered controls are the students' gender (FEM 1 = female, 0 = male), migration background (MIG: both parents born in Germany = 0, one parent born abroad = 1, both parents born abroad = 2), and the duration of kindergarten attendance in years (KIGA). In all models, the survey wave is controlled for.

Table 2 provides a descriptive overview of the used variables both for the total analytical sample and separately by country state. For the multivariate analyses, central metric explanatory variables have been centered (SES, SSES, ABL, SABL).

Table 2: Descriptive statistics of used variables (BiKS-8-14, grades 5-7)

Variable	Measure	total		Bavaria (BY)		Hesse (HE)	
		%/mean	sd	%/mean	sd	%/mean	sd
dependent variable							
ACH	reading comprehension	54.02	.04	53.51	.05	56.04	.09
explanatory variables							
BY	Bavaria	80.13		--		--	
SES	parental ISEI (highest)	54.00	.06	53.25	.06	57.02	.12
SSES	mean parental ISEI (highest) in school class	54.04	.03	53.30	.04	57.00	.06
ABL	general cognitive abilities	11.45	.01	11.40	.01	11.67	.02
SABL	mean general cognitive abilities in school class	11.44	.00	11.39	.00	11.65	.01
GYM	attendance Gymnasium	64.26		57.42		91.83	
Controls							
FEM	female	52.36		52.59		51.46	
MIG	migration background						
	one parent migrant	7.44		7.23		8.29	
	both parents migrants	9.69		8.97		12.60	
KIGA	kindergarten duration	3.03	.00	3.02	.00	3.08	.01
no. of observations		4,314		3,457		857	
no. of cases (students)		1,614		1,296		318	

Note: For the time-varying variables ACH, SSES, ABL, and SABL averages across waves are displayed.

For the multivariate OLS-regression analysis using Stata, hybrid models are estimated (Allison, 2009), as they provide within estimates of level 1 variables and allow for including level 2 variables. This way, effects within and between students can be estimated in a single model and therefore combines advantages of fixed- and random-effects models (Schunck, 2013). For a robustness check, random-effects models were additionally estimated (not reported).

6. Results

According to the hypotheses of the MoAbiT *no* additional effects enhancing educational inequality in the case of differentiation according to achievement at the end of elementary school are expected. This is at the latest the case when prior cognitive abilities and the composition of schools and school classes in terms of these abilities are included in the analyses. It is also expected that results from other approaches that deviate from the hypotheses of the MoAbiT can be replicated by taking over their (mis-) specifications.

Differentiation and Stratification 1: The MoAbiT

Table 3 summarizes the results of the empirical test of the MoAbiT using data of the BiKS. Reference point of the comparison is Hesse with its institutionally (and also empirically) more “liberal” orientation towards a differentiation according to achievement as compared to Bavaria. In the simple comparison (yet already including gender, migration background⁷, kindergarten attendance, and development of achievement over the waves; model 0) one can at first observe a *lower* level in terms of educational achievement (for reading skills).⁸

⁷ The presentation in Table 3 is reduced to the effect of two immigrated parents.

⁸ This contradicts the usual descriptive results of partly significantly higher achievement as presented, for example, in PISA and IQB country reports (cf. Esser, 2015a , 2015b). This is, however, plausible because of the exclusion of comprehensive schools entailing low and intermediate performing students in Hesse.

Table 3: Empirical results based on the MoAbiT (BiKS; OLS estimations; hybrid regression according to Allison (2009)⁹; multiple imputation of missing data; waves controlled; light grey: specific effects of differentiation; dark grey: empirical support of the MoAbiT)

	MoAbiT	0	1	2	3	4	5	6	7	8
BY/HE	≥ 0	-2.63	-2.41	-1.77	-0.67	-0.46	-0.59	-0.64	-0.46	0.79
SES ·BY	0		0.04	-0.14	-0.03	-0.04	-0.04	-0.04	-0.04	-0.04
SSES ·BY	0		0.30	0.30	-0.15	-0.15	0.00	0.00	-0.02	-0.12
ABL ·BY	0		1.00	1.00	1.01	0.99	-0.47	-0.49	-0.45	-0.43
SABL ·BY	≥ 0		2.34	2.35	2.35	2.26	2.09	2.19	2.06	1.80
SES	≥ 0			0.18	0.07	0.09	0.09	0.08	0.08	0.08
SSES	0				0.45	0.44	0.20	0.18	0.22	0.21
SES·SSES	0					-0.005	-0.003	-0.003	-0.004	-0.004
ABL	+						2.16	1.92	1.97	1.98
SABL	+							1.11	1.38	0.34
ABL·SABL	+								0.31	0.26
GYM	≥ 0									4.83
FEM		2.67	1.94	1.99	1.97	1.95	1.86	1.88	1.90	1.81
MIG		-3.99	-2.41	-1.67	-1.36	-1.17	-1.08	-1.15	-1.15	-1.94
KIGA		0.63	-0.27	-0.26	-0.22	-0.21	-0.22	-0.25	-0.25	-0.25
R ²		0.06	0.13	0.21	0.21	0.21	0.30	0.30	0.32	0.34
N		4314	4314	4314	4314	4314	4314	4314	4314	4314

The heart of the test of the MoAbiT consists of the findings on the four *specific* additional or system effects of a greater differentiation according to abilities and social origin, both as an individual characteristic and as a school effect (SES·BY, SSES·BY, ABL·BY, SABL·BY) as well as the then remaining country effect (BY/HE).¹⁰

⁹ Analyses applying random effects models (with random intercept and random slope) yield hardly differing estimates.

¹⁰ The coefficient for comparing Bavaria and Hesse indicates in model 0 the difference in achievement between Hesse and Bavaria after controlling for the three covariates gender (FEM), migration background (MIG), and kindergarten attendance (KIGA). With the specification of the interaction effects (models 1 to 9) the values for the main effects for SES, SSES, ABL and SABL have to be interpreted as ‘conditional effects’: it is the value for Hesse being the reference category of the comparison.

Model 1 in Table 3 describes the four (system-) effects without any further control.¹¹ All effects are significant and positive. Model 2 then controls for the effect of social origin (SES) as the first of the general conditions. Although system effects (SES·BY, SSES·BY, ABL·BY, SABL·BY) decrease, they remain significant and positive – with one exception: the system effect of individual social origin (SES·BY) not only disappears, it even becomes negative. Models 3 and 4 control for school effects of social origin, including the interaction with individual social origin (SSES and SES·SSES). As usual, there is a reinforcing school effect of the composition according to social origin. However, this effect is decreasing with an increasing proportion of higher class children and all system effects of social origin disappear. In model 5 cognitive abilities at the end of elementary school (ABL) are controlled and in model 6 the corresponding school effects and their interaction with individual cognitive abilities (SABL and ABL·SABL) are included. All these conditions have a substantial effect on achievement. Now the systems-effect of differentiation with individual ability turn to a significant negative sign, which contradicts the hypotheses of the MoAbiT.¹²

In model 7, the interaction of individual ability and school ability is introduced. The significant positive effect denotes the (Matthew-) effect of a reinforcement of a better exposure with a higher efficiency by the combination of a higher individual ability with a higher school-ability. All this does not cause any changes in the system effects. Especially the system effect of the composition of schools and school classes according to cognitive abilities clearly persists (SABL·BY).¹³

Things also remain the same when controlling for school types (GYM) in model 8. Here a clear additional positive effect occurs, without changing the system effect of the reinforcement of the achievement-related school effect. There are, however, some changes regarding the general school effects: The interaction between social origin and social composition of schools and school classes is now no longer significant. In addition, the

¹¹ That is not surprising, because no main effect of SES, SABL, ABL, and SABL are controlled here. Model 1 serves as a more or less formal reference for an upper limit of possible *specific* system effects, to be controlled stepwise by the respective *general* (main) effects.

¹² With control of ABL and SABL in models 5 and 6 the effect of individual ability is *decreasing* with the stricter differentiation: ABL·BY becomes negative. That would mean, that children with lower talents are (relatively!) better off in systems with a stricter ability tracking. In the following models this effect becomes insignificant (and fits with the MoAbiT with its prediction for no effect), but remains in its direction and strength.

¹³ Considering grades and recommendations at the end of elementary school does not change the results (not reported). Although in this case explained variances overall increase, the effects of the different variables do not change. Obviously, grades, recommendations, and the measures for cognitive abilities function as equivalent and complementary indicators for the increase in learning efficiency for a given exposure.

composition according to cognitive abilities now exerts its effect solely through the interaction with individual cognitive abilities and thus through the mutual reinforcement of individual abilities and a stimulating learning environment as expected by the MoAbiT. The finding on the *Gymnasium* also points to further influences of the school types themselves and independent from sorting according to achievement. The available data do not allow for further clarifying the basis for these additional effects: differences in curricula, schools', or teachers' efforts?

Thus, three central findings can be reported. *First*: All *stratifying* consequences of a stricter differentiation *disappear*. Although general effects of social origin and the corresponding composition of schools and school classes persist, they apply to *all* regimes of educational systems. *Second*: From the specific system effects of a stricter differentiation, only that of the homogenization of schools and school classes with regard to cognitive abilities persists (SABL·BY). This means that there are indeed particular effects enhancing children's achievement in differentiated systems which involve a *strict(er)* adherence to sorting according to cognitive abilities and achievement in elementary school by, for example, keeping secondary and tertiary effects as low as possible and sorting in separated schools and schoolclasses. *Third*: There are *no* additional direct (residual) effects of differentiation as compared to integration. All cross-country differences are explained by the general and specific conditions and processes and the interaction with differentiation and thus by the *indirect* path of (stringent) sorting and school effects.

Differentiation and Stratification 2: Other Approaches

Table 4 depicts the findings regarding the replication of the three other approaches using data of the BiKS study. The comparison is arranged in such a way that the approaches gradually approximate the requirements of the MoAbiT, thus allowing determining whether the results converge to the MoAbiT or not.

Table 4: Reanalysis OECD-, Standard- and DVD-Approach (BiKS; OLS estimations; hybrid regression according to Allison (2009); multiple imputation of missing data; gender, migration background, kindergarten attendance and waves controlled; white-shaded: contrary to the MoAbiT; dark grey-shaded: corresponding to the MoAbiT)

	MoAbiT		OECD Reports		Standard Approach		DVD Approach	
	Hypotheses	Result BiKS	Results so far	Result BiKS	Results so far	Result BiKS	Results so far	Result BiKS
BY/HE	≥ 0	0.79	≤ 0	-13.30	≤ 0	-3.30	-	10.8
SES ·BY	0	-0.04	+	0.20	+	0.03	-	-0.03
SSES ·BY	0	-0.12					+	-0.13
ABL ·BY	0	-0.43						
SABL ·BY	≥ 0	1.80						
SES	≥ 0	0.08				0.18	+	0.08
SSES	0	0.21					+	0.28
SES·SSES	0	-0.004					+	-0.005
ABL	+	1.98						
SABL	+	0.34						
ABL·SABL	+	0.26						
GYM	+	4.83					+	9.01
R ²		0.34		0.13		0.14		0.29
N		4314		4314		4314		4314

OECD Reports basically present only the mean values for the achievement level and the respective ‘social gradient’, i.e. the correlation between achievement and social origin. They are reported separately for the single countries and depicted in graphs of the combination between achievement level and social impermeability (cf. chapter 2 above). Replicating this approach with data of the BiKS study using a regression model of conditional effects for Hesse and Bavaria and the system effect of social origin yields the usual result as indicated in the OECD Reports regarding countries with differentiated educational systems: a *greater* impact of social origin (SES·BY: 0.20) without a higher level of achievement in Bavaria.¹⁴

The *Standard Approach* differs from the simple descriptions of OECD reports in that it considers partly extensive lists of control variables, including individual social origin and the

¹⁴ See however footnote 8 for the finding of a negative effect for Bavaria in the analysis.

interaction effect between social origin and differentiation. Again a negative effect for Bavaria appears, but different from most contributions using the Standard Approach *no* reinforcement of the effect of social origin (SES·BY: 0.03, not significant) is found in the reanalysis using data of the BiKS study. Therefore, one can interpret the negative effect of the comparison between Bavaria and Hesse (-3.30, significant) as main effect. This also corresponds to the usual finding of the Standard Approach indicating that differentiation rather leads to a decrease in the achievement level. However, the analysis also yields the usual (clear) effect of individual social origin (SES: 0.18, significant). This, of course, applies generally and thus equally to both countries and educational regimes.

Both approaches do not consider schools and school classes as an own level of school effects based on the respective composition. If they were included according to the *DVD Approach*, the findings on system effects change as compared to the Standard Result. They also change in the reanalysis using the BiKS data, but not in the usual way. Now, the unexpected decrease in the effect of social origin of the DVD Puzzle does not occur for Bavaria (SES·BY: -0.03, not significant; cf. chapter 2 above). Moreover, unlike in the DVD Puzzle there is no increase in the effect of the composition of schools and school classes according to social origin (SSES: -0.15, not significant). So there is *no* stratifying effect of differentiation *at all* when applying the DVD Approach to the BiKS data. However, including the level of schools and school classes indeed revises the impression one gets from OECD reports and from the Standard Approach: Achievement is now clearly *higher* in Bavaria with its *stricter* sorting (BY/HE with a significant value of 10.8 and missing interaction effects). Effects with regard to general conditions remain the same as found in previous analyses applying the DVD Approach: Achievement increases with higher social origin (SES: 0.08, significant), with an increasing proportion of children from upper classes in schools and school classes (SSES: 0.28, significant), and even increases once more in *Gymnasium* (GYM: 9.01, significant). This again applies everywhere and also corresponds to the finding according to the MoAbiT.

It is obvious that the results obtained by replicating the three approaches using the BiKS data clearly differ in terms of the core aspects from those of the MoAbiT. But it is also obvious that with an increasing approximation of the approaches to the specifications of the MoAbiT, the results indeed converge: According to the OECD Reports the still usually assumed strong connection between differentiation and stratification seems indeed to exist. However, this connection already disappears in the replication with BiKS data using the Standard Approach,

but at the latest when applying the DVD Approach and including also the school effects of social origin.

Yet, when comparing the results obtained by the DVD Approach with those of the fully specified MoAbiT, it becomes apparent that general school effects (effects of the composition according to social origin and school types) are more pronounced in the DVD Approach than in the MoAbiT. The reason for this is obvious: Unlike the DVD Approach, the MoAbiT approach includes cognitive abilities and prior achievement as the core conditions for all effects. This leads to a further decrease in the stratifying effects. And then one can also detect what lies behind the strong positive direct effect for Bavaria found with the DVD Approach: the *reinforcement* of the effect of homogenization of schools and school classes in terms of cognitive abilities (SABL·BY: 1.67, significant). However, this will only be recognizable if the model is fully specified and if cognitive abilities are empirically considered with regard to individual and school effects. It also has to be mentioned here, that independent from whether the DVD Approach or the MoAbiT is applied, all of the remaining stratifying effects are *negative*, although they do not exceed the significance threshold. It can be summarized that a *stricter* differentiation rather tends to cause an *increase* in achievement and, if at all, even a higher social permeability for better performance. At least it should be noted that – different from what is usually assumed – institutional sorting according to cognitive abilities by no means inevitably entails an increase of stratification in educational achievement, which is not existent in general.

7. A short conclusion

The central finding of analyzing the effects of ability tracking with data of the BiKS study for the country states Bavaria and Hesse in Germany with their differing educational regimes can thus be summarized as follows: Unlike OECD Reports and the Standard Results mostly indicate and provided that the relevant conditions and processes are theoretically correctly specified and empirically controlled, there is *no* empirical evidence for the reinforcement of stratifications in terms of achievement in secondary school due to differentiation of the educational system. In fact, an actually additional *positive* effect of homogenization of schools and school classes according to cognitive abilities emerges, just like it is assumed by the advocates of ability tracking. This finding is supported by the replication of the various

(incomplete) approaches: They yield, in essence, the same misleading conclusions in terms of the effects of differentiation, which disappear when approximating a correct specification.

Not all questions could be answered with the data of the BiKS study, like particularly the question of which special conditions of the educational systems are *precisely* responsible for the specific effects. To answer this question it would have been necessary to refer the comparison to more countries and to higher variances especially in terms of contextual variables. Moreover, all variations only refer to the still rather strongly differentiated educational system in Germany. The ‘National Educational Panel Study’ (NEPS) would allow for extending the data base for Germany, thus making a *national* comparison of all 16 country states possible. The actually required *international* comparison cannot be conducted in the foreseeable future: Although the PISA study is currently the only suitable internationally comparative study which contains the necessary performance measurements in secondary schools, it currently does not contain the decisive conditions of institutional sorting after elementary school – ability or achievement before sorting. And that this will change is at the moment unforeseeable.

Finally, it also reveals that effects of differentiation are not particularly strong overall. This also applies to the positive additional effect of a strict(er) differentiation according to achievement through the homogenization of schools and school classes. And it almost appears as if the debate was nothing but a storm in a teacup which is artificially stimulated by ideologies concerning educational and social policies that have actually no consequences. In any case it should be noted that educational achievement in secondary school *primarily* depends on children’s cognitive abilities *and* on the corresponding *composition of schools and school classes* and that there are almost no effects of social origin as a consequence of differentiation according to achievement, if this was statistically controlled. This result of a ‘decline of the social’ for explaining educational inequality has meanwhile become increasingly obvious (cf. for current overviews and findings Marks, 2014, 2015). The reason why another impression had gained support for such a long time may have something to do with the fact, that a sufficiently systematic theoretical basis had been missing which would have sounded a note of caution. That is to say, that the crucial condition for both educational achievement and institutional sorting is children’s cognitive abilities which they have by birth and further develop within families and during *prior* elementary school attendance.

References

- Alba R, Sloan J and Sperling J (2011) The integration imperative: The children of low-status immigrants in the schools of wealthy societies. *Annual Review of Sociology* 37: 395-415.
- Allison PD (2009) *Fixed effects regression models*. Thousand Oaks, CA: Sage Publications.
- Ammermüller A (2005) Educational opportunities and the role of institutions. *ZEW Discussion Papers* No. 05-44.
- Bauer P and Riphahn RT (2006) Timing of school tracking as a determinant of intergenerational transmission of education. *Economic Letters* 91(1): 90-97.
- Betts JR (2011) The economics of tracking in education. In: Hanushek EA, Machin S and Wößmann L (eds) *Handbook of the Economics of Education. Volume 3*. Amsterdam: North Holland, pp. 341-381.
- Böhme K and Weirich S (2012) Der Ländervergleich im Fach Deutsch. In: Stanat P, Pant HA, Böhme K and Richter D (eds) *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann, pp. 103-116.
- Bol T, Witschge J, Van de Werfhorst HG and Dronkers J (2014) Curricular tracking and central examinations: Counterbalancing the impact of social background on student achievement in 36 countries. *Social Forces* 92(4): 1545-1572.
- Boudon R (1974) *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.
- Burgess SM, Dickson M and Macmillan L (2014) Selective schooling systems increase inequality. *IZA Discussion Paper* No. 8505.
- Brunello G and Checchi D (2007) Does school tracking affect equality of opportunity? New international evidence. *Economic Policy* 22 (52): 781-861.
- Dronkers J (2015) In wiens voordeel werkt selectie aan het begin van het voortgezet onderwijs? Een nieuwe benadering van een oude vraag. *Mens & Maatschappij* 90(1): 5-24.
- Dronkers J, Van der Velden R and Dunne A (2011) The effects of educational systems, school-composition, track-level, parental background and immigrants' origins on the achievement of 15-years old native and immigrant students. A reanalysis of PISA 2006. *ROA-RM-2011/6*. Maastricht: ROA Research Centre for Education and the Labour Market.
- Dronkers J, Van der Velden R and Dunne A (2012) Why are migrant students better off in certain types of educational systems or schools than in others? *European Educational Research Journal* 11(1): 11-44.

- Dronkers J and Korthals R (2015) Systems' tracking, schools' entrance requirements and educational performance by migrants' children. In: Hadjar A and Gross C (eds) *Education systems and educational inequalities*. Bristol: Policy Press (forthcoming).
- Duflo E, Dupas P and Kremer M (2011) Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review* 101(5):1739-1774.
- Dunne A (2010) *Dividing lines. Examining the relative importance of between- and within-school differentiation during lower secondary education*. PhD Thesis, European University Institute, Florence.
- Dustmann C, Frattini T and Lanzara G (2012) Educational achievement of second generation immigrants: An international comparison. *Economic Policy* 27(69): 143-185.
- Erikson R, Goldthorpe JH, Jackson M, Yaish M and Cox DR (2005) On class differentials in educational attainment. *Proceedings of the National Academy of Sciences of the United States of America* 102(27): 9730-9733.
- Esser H (2013) Ethnische Bildungsungleichheit und Bildungssysteme: Der Blick in die Bundesländer. In: McElvany N, Gebauer MM, Bos W and Holtappels HG (eds) *Jahrbuch der Schulentwicklung. Band 17: Daten, Beispiele und Perspektiven*. Weinheim: Beltz Juventa, pp. 80-101.
- Esser H (2015a) Bildungssysteme und ethnische Bildungsungleichheit. In: Diehl C, Hunkler C and Kristen C (eds) *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten*. Wiesbaden: Springer VS (forthcoming).
- Esser H (2015b) Sorting and (much) more. The standard-approach and the DVD-puzzle on effects of educational systems on educational inequalities. In: Hadjar A and Gross C (eds) *Education Systems and Educational Inequalities*. Bristol: Policy Press (forthcoming).
- Esser H (2015c) Educational systems and educational inequality. The model ability tracking and empirical findings. In: Blossfeld H-P (ed) *Differentiation in secondary education and its short- and longer-term effects on inequality of educational opportunities* (forthcoming).
- Ganzeboom HBG and Treiman DJ (1996) Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research* 25: 201-239.
- Hanushek EA and Wößmann L (2006) Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal* 116: C63-C76.
- Hanushek EA and Wößmann L (2011) The economics of international differences in educational achievement. In: Hanushek EA, Machin S and Wößmann L (eds) *Handbook of the Economics of Education. Volume 3*. Amsterdam: North-Holland, pp. 89-200.
- Horn D (2009) Age of selection counts: a cross-country analysis of educational institutions. *Educational Research and Evaluation* 15(4): 343-366.

- Horn D (2013) Diverging performances: the detrimental effects of early educational selection on equality of opportunity in Hungary. *Research in Social Stratification and Mobility* 32: 25-43.
- Jackson M (2013) Introduction: How is inequality of educational opportunity generated? The case for primary and secondary effects. In: Jackson M (ed) *Determined to succeed? Performance versus choice in educational attainment*. Stanford, CA: Stanford University Press, pp. 1-55.
- Korthals R (2012) Selection and tracking in secondary education: A cross country analysis of student performance and educational opportunities. *ROA-RM-2012/14*. Maastricht: ROA Research Centre for Education and the Labour Market.
- Marks GN (2005) Cross-national differences and accounting for social class inequalities in education. *International Sociology* 20(4): 483-505.
- Marks GN (2014) *Education, social background and cognitive ability. The decline of the social*. London: Routledge.
- Marks GN (2015) Are school-SES effects statistical artefacts? Evidence from longitudinal population data. *Oxford Review of Education* 14(1): 122-144.
- Meier V and Schütz G (2007) The economics of tracking and non-tracking. *Ifo Working Paper* No. 50.
- Merry JJ (2013) Tracing the U.S. deficit in PISA reading skills to early childhood: Evidence from the United States and Canada. *Sociology of Education* 86(3): 234-252.
- Müller W and Kogan I (2010) Education. In: Immerfall S and Therborn G (eds) *Handbook of European Societies: Social Transformations in the 21st Century*. New York: Springer, pp. 217-289.
- OECD (2014) *PISA 2012 Results: What students know and can do. Student performance in mathematics, reading and science. Volume I (Revised edition, February 2014)*. OECD Publishing.
- Pfost M and Artelt C (2013) Reading literacy development in secondary school and the effect of differential institutional learning environments. In: Pfost M, Artelt C and Weinert S (eds) *The development of reading literacy from early childhood to adolescence. Empirical findings from the Bamberg BiKS longitudinal studies*. Bamberg: University of Bamberg Press, pp. 229-277.
- Ruhose J and Schwerdt G (2015) Does Early Educational Tracking Increase Migrant-Native Achievement Gaps? Differences-In-Differences Evidence Across Countries. IZA Discussion Paper 8903.
- Schipolowski S and Böhme K (2010) Die sprachlichen Kompetenzen in den Ländern. In: Köller O, Knigge M and Tesch B (eds) *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann, pp. 87-96.
- Schlotter M (2011) The effect of preschool attendance on secondary school track choice in Germany. Evidence from siblings. *Ifo Working Paper* No. 106.

- Schmidt S, Schmitt M and Smidt W (2009) Die BiKS-Studie. Methodenbericht zur zweiten Projektphase. *PsyDok (Online)*, 2534. Available at: <http://psydok.sulb.uni-saarland.de/volltexte/2009/2534/> (accessed 25 April 2015).
- Schunck R (2013) Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *The Stata Journal* 13(1): 65-76.
- Teltemann J (2012) *Gleichheit oder Leistung? Wie Bildungssysteme, Wohlfahrtsstaatsregime und Integrationspolitiken die Schulleistungen junger Migranten beeinflussen*. PhD Thesis, University of Bremen, Germany.
- Vandenberge V (2006) Achievement effectiveness and equity: the role of tracking, grade repetition and inter-school segregation. *Applied Economics Letters* 13(11): 685-693.
- Van de Werfhorst HG and Mijs JJB (2010) Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology* 36: 407-428.
- Verwiebe R and Riederer B (2013) Die Lesekompetenzen von Jugendlichen mit Migrationshintergrund in westlichen Gesellschaften. Eine Mehrebenenanalyse auf Basis der PISA-Studie von 2000 bis 2009. *Zeitschrift für Soziologie* 42(3): 201-221.
- von Below S (2011) Bildungssysteme im historischen und internationalen Vergleich. In: Becker R (ed) *Lehrbuch der Bildungssoziologie, 2. Aufl.*, Wiesbaden: VS Verlag für Sozialwissenschaften, p. 137-162.
- von Maurice J, Artelt C, Blossfeld HP, Faust G, Roßbach HG and Weinert S (2007) Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter: Überblick über die Erhebungen in den Längsschnitten BiKS-3-8 und BiKS-8-12 in den ersten beiden Projektjahren. *PsyDok (Online)*, 1008. Available at: <http://psydok.sulb.uni-saarland.de/volltexte/2007/1008/> (accessed April 25 2015).
- Waldinger F (2006) Does tracking affect the importance of family background on students' test scores? Unpublished manuscript. London: LSE.
- Weiß RH (2006) *Grundintelligenztest Skala 2 - Revision - (CFT 20-R) mit Wortschatztest und Zahlenfolgentest - Revision (WS/ZF-R)*. Göttingen: Hogrefe.
- Wößmann L (2009) International evidence on school tracking: A review. *CESifo DICE Report - Journal for Institutional Comparisons* 7(1): 26-34.
- Wößmann L, Lüdemann E, Schütz G and West MR (2009) *School accountability, autonomy and choice around the world*. Celtenham: Edward Elgar Publishing.