Schmollers Jahrbuch 128 (2008), 351 – 379
Duncker & Humblot, Berlin

Educational Effects of Alternative Secondary School Tracking Regimes in Germany

By Andrea M. Mühlenweg*

Abstract

This paper examines educational outcomes of pupils selected to secondary school types by different tracking regimes in a German state: Pupils are alternatively streamed after fourth grade or after sixth grade. Regression results indicate that, estimated on the mean, there are no negative effects of later tracking on educational outcomes in the middle of secondary school. Positive effects are observed for pupils with a less favorable family background. Quantile regressions reveal that effects of later tracking are positive for the lower quantiles but decrease monotonically over the conditional distribution of test scores, turning into negative effects for the upper quantiles.

Zusammenfassung

Dieses Papier vergleicht schulische Ergebnisse von Schülerinnen und Schülern im Bundesland Hessen, die zu unterschiedlichen Zeitpunkten den Sekundarschulformen zugewiesen werden: Alternativ zur klassischen Sekundarschulzuweisung nach der 4. Klasse, findet in sogenannten Förderstufen die Zuweisung erst nach der 6. Klasse statt. Die Regressionsergebnisse weisen darauf hin, dass es – im Durchschnitt – keine Effekte des Zeitpunkts der Zuweisung auf die Sekundarschulleistung zu geben scheint. Allerdings werden für Kinder mit einem benachteiligten familiären Hintergrund durchaus positive Effekte einer späteren Selektion beobachtet. Ergebnisse von Quantilsregressionen ver-

* Financial support from the Anglo-German Foundation within the project 'The Economics and Politics of Employment, Migration and Social Justice' of the initiative 'Creating Sustainable Growth in Europe' for parts of the study is gratefully acknowledged. I am very grateful to two anonymous referees for helpful comments. I thank Horst Entorf, Knut Gerlach, Szilvia Hamori, Uwe Jirjahn, Christian Pfeifer, Patrick Puhani, Sergiy Radyakin, Philip Savage as well as participants of the Ph.D. Seminars in Applied Labor Economics at the University of Hannover and the Darmstadt University of Technology for valuable comments. I also thank Hans-Peter Hafner from the Research Data Center (Forschungsdatenzentrum) of the Statistical Office of the state of *Hessen* for help with the administrative data for *Hessen.* Thanks are due to Ursula Dörger from the Ministry of Education in *Hessen* who kindly answered my questions concerning school regulations. All remaining errors in this study are my own.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

deutlichen, dass der geschätzte Effekt der Förderstufen monoton über die Quantile abnimmt.

JEL Classification: I21, I28

Received: October 12, 2007
Accepted: January 11, 2008

1. Introduction

Numerous European countries select pupils into more or less academic tracks at some point during their secondary education. The rationale behind educational tracking or streaming is to provide a homogeneous learning environment which is supposed to foster specific pupils' abilities and to improve educational outcomes. From a theoretical point of view, the educational setup with respect to tracking may be considered as the result of an optimization process. Thus, recent studies model optimal tracking time as being determined by a trade-off between negative and positive effects of early tracking: The negative effect stems from the assumption that the tracking decision is the more appropriate (with respect to actual, unobserved individual ability) the later tracking takes place. The counteracting positive effect is due to more able pupils benefiting from a more selective system (cf. Brunello / Giannini / Ariga, 2007; Ariga / Brunello / Iwahashi / Rocco, 2005).[1]

In Germany, pupils are generally tracked into three different types of secondary schools at a relatively early point of their educational careers (mostly at the age of ten). Track choice mainly depends on the decisions made by parents. Recently, researchers have argued that this early tracking regime is an important source of high educational inequality: For example, Dustmann (2004) states that early tracking enforces intergenerational immobility because of strong influences of parental views on the children's (early) educational decision. The study shows that parental education and occupational status have a significant impact on the children's secondary school choice and subsequent educational attainment in Germany. In addition, these parental influences yield to differences in the children's earnings later in life. These views are confirmed by recent studies mainly drawing on internationally standardized test score data for different countries: The cross-county comparisons by Hanushek / Wößmann (2006); Entorf / Lauk (2006); Ammermüller (2005); Schütz / Ursprung / Wößmann (2005)[2] and the Swiss cross-canton study by Bauer / Rip-

[1] Non-linear peer-effects are assumed in these models. Epple / Newton / Romano (2002) is a further study modelling implications of school tracking. However, this paper refers to the somewhat different context of ability tracking within public and private schools. Different selection mechanisms to school tracks are examined in Fernandez (1998).

[2] The empirical paper by Schütz / Ursprung / Wößmann (2005) also offers a theoretical model linking the timing of tracking to education inequality.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

hahn (2006) indicate that countries featuring tracking and especially early tracking systems are characterized by relatively high educational inequality and lower average performance. Pekkarinen (2005) shows that later tracking yields higher gender differences in education in favor of girls and decreases the subsequent gender wage gap.[3]

One special feature of the German educational system is that besides the traditional early tracking schools some later tracking schools exist, too: In so-called 'support stages' (*Förderstufe*) or 'orientation stages' (*Orientierungsstufe*) tracking is postponed for two years. The idea is that pupils are given more time to develop specific skills and interests and that teachers and parents receive improved information for the transition decisions to secondary schools. To date and to my knowledge, no empirical research has been undertaken to identify a causal effect of the 'support stages' on educational outcomes using appropriate statistical strategies.[4]

This study aims at examining educational effects of these special schools in one German state (Hessen) where this institution co-exists to the traditional school types. The empirical examination is based on two student-level data sets for the relevant federal state:[5] The PISA-E data (a national extension of the PISA data) provide information on pupils' test scores in ninth grade as well as information on their individual and family background and their tracking history. Additionally, an administrative data-set covering all students in the state of Hessen is used to examine the state's tracking practice in detail.

The central methodological problem when comparing educational outcomes by tracking regime is that tracking regime choice is endogenous to educational outcomes. Thus, estimates of the 'support stage effect' are likely to be biased in a simple regression framework. In brief, since the endogeneity bias can be considered to be an omitted variable bias, I examine how the estimated effect changes whilst a broad variety of background characteristics is controlled for based on the PISA-E data. This approach reduces the bias of

[3] While the focus of the present paper is on tracking of pupils to academic and vocational school types further empirical studies consider ability grouping within schools. Recent papers examining this version of tracking are for example, Zimmer (2003); Figlio / Page (2002), and Betts / Shkolnik (2000).

[4] An early study of the 'support stages' in Hessen is provided by Hopf (1979) and describes the development and organisation of the schools as well as experiences of parents, teachers and pupils in this school type. The study does not compare 'support stage' outcomes to outcomes of alternative school types using evaluation techniques. A similar approach is taken in the studies of 'orientation stages' in Bremen by Jürgens (1989) and Jürgens (1991). Henze / Sandfuchs / Zumhasch (1996) focuses on low ability pupils within 'orientation stages' in the state of Niedersachsen.

[5] Aside from the institution of the 'support stage', I consider Hessen to be a 'representative' (West) German state. According to administrative data, the proportions of pupils in the different traditional secondary school tracks are representative as compared to the pattern for Germany as a whole (cf. the figures presented in Puhani / Weber, 2007b).

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

the estimator and provides profound descriptive evidence on the effects of later tracking. Still, selection effects can not be ruled out completely and will be discussed carefully.

Since some of the above mentioned studies demonstrate (based on comparisons of different countries) that later tracking reduces educational inequality, this paper also focuses on inequality aspects of the tracking regime. First of all, regression results are presented for different sub-groups according to pupils' family background. Furthermore, quantile regressions demonstrate the difference of the later tracking effect for pupils at different quantiles of the conditional performance distribution.

This paper is organized as follows: Section 2 describes the German education system with an emphasis on the institutional framework of the state of Hessen. Section 3 provides descriptive evidence on tracking in Hessen. It is shown that pupils having attended later tracking schools perform worse (in terms of the secondary education level reached) than pupils who have been tracked early. However, these results are driven by the endogeneity of regime choice. The methodological framework for an analysis of track choice taking its endogeneity with respect to educational outcomes into account is introduced in Section 4 together with the results: Overall, there seem to be no negative effects of later tracking. However, sub-group analyses and quantile regression results reveal that 'support stages' seem to work in favor for children with a disadvantaged family background whilst there are negative effects on pupils on top of the conditional performance distribution. Thus, later tracking may, in fact, decrease education inequality but to the detriment of the top performers. Section 5 discusses the findings and presents conclusions.

2. Stylized Facts

2.1 Institutional Background

Traditionally, the German school system is characterised by early ability streaming of pupils. Table 1 provides an overview of the tracking systems in selected industrialised countries:[6] While many European countries track pupils to more or less academic secondary school types, Germany's regular tracking age of ten is rather early in international comparison. To be more specific, in Germany pupils are selected into three school types after four years of elementary school:[7] The most 'able' pupils are supposed to attend the Gymnasium,

[6] Besides explicitly streaming pupils to vocational and academic tracks, in some countries it is common to select pupils to different classes within comprehensive secondary schools according to ability (as it is the case in the United States). This version of tracking is not considered in Table 1.

[7] In the East German states Berlin and Brandenburg, primary school generally covers six grades.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

which is a nine- (or eight-) year higher-level secondary school and enables pupils to pursue further academic studies (for example at universities).[8] An alternative school track is offered by the Realschule as an intermediate level secondary school which generally lasts six years and prepares pupils for a rather vocational education. Finally, the Hauptschule, as the lowest level secondary school type, is supposed to be the most vocational and least academic track and lasts five years. In principle, it is possible to change tracks after the initial track decision. However, different curricula for the different school types complicate switching tracks, especially after sixth grade.[9]

Table 1

First age of selection in the education system

10	11	12	13	14	15	16
Austria Germany	Czech Republic Hungary Slovak Republic Turkey	Belgium Mexico Netherlands	Canada Luxembourg	Italy Korea	France Greece Ireland Japan Poland Portugal Switzerland	Australia Denmark Finland Iceland New Zealand Norway Spain Sweden U.K. U.S.

Source: OECD (2004, 262).

Besides the system of streaming pupils to the different secondary school types after fourth grade, later tracking school types also exist. These school types, which are called ‘support stages’ (Förderstufe) or ‘orientation stages’ (Orientierungsstufe), track pupils after sixth grade. Later tracking schools were mainly introduced in different regions at the end of the 1950s and in the 1970s:[10] Especially in the 1950s, educational experts developed the idea of

[8] Recently there has been a tendency to shorten the duration to eight years. In the East German states Sachsen and Thüringen, the higher secondary school generally takes eight years.

[9] Hardly any figures on switching tracks exist. Baumert / Trautwein / Artelt (2003) states that 14.4 % of German 15-year-old pupils in the PISA study claim to have switched from initial secondary school track to another track. Pischke (2007) explains that 7 % of pupils switched to higher level schools from lower or intermediate secondary schools in 1966. Recent evidence based on administrative data for Hessen is given in Puhani / Weber (2007b) demonstrating that track upgrades are more frequent than downgrades.

[10] For further information on the history of comprehensive secondary schooling see Hessisches Kultusministerium (1995) and Jürgens (1991).

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

so-called 'support stages'.[11] While the traditional elementary schools were to be maintained, the Förderstufe sought to combine grades five and six in an autonomous comprehensive school type which would be located at traditional German lower secondary or primary schools. In the states of Hessen and Niedersachsen, this school type was introduced on a larger scale alongside the traditional tracking system.[12] Reasons for introducing 'support stages' may have been rather theoretical ones (for example to foster equal educational opportunities) or practical ones: Schools in rural areas tended to introduce 'support stages' so that all fifth and sixth graders could be provided with local secondary education.[13]

All in all, discussions on the idea of prolonged comprehensive schooling generated a mixed system of institutions in Germany: The state of Hessen introduced the offer of 'support stages' (Förderstufe) in some schools coexisting with the traditional selective school types. Children in these 'support stage' schools are normally taught in comprehensive classes, while separate classes according to ability may exist for mathematics and the first foreign language (mostly English).

Concerning the regulations in the other German states, in most states, pupils are mainly still selected to different secondary school types after fourth grade. Furthermore, the states of Bremen and Niedersachsen used to have fully established comprehensive 'orientation stages' covering grades five and six but abolished them in 2005 and 2004 respectively. It is only in Berlin and Brandenburg that elementary school traditionally takes six instead of four years.

In addition, general comprehensive schools exist in Germany, too. Pupils in the former German Democratic Republic used to be taught in comprehensive schools (Einheitsschule) until tenth grade. In West Germany, comprehensive schools (Gesamtschule) were introduced as an 'experiment' in several schools in the 1960s and lead to grade 10 or 13 respectively. From 1973 to 1982 all German states introduced some experimental comprehensive schools. Pupils in comprehensive schools are taught in different ability groups (only) in some

[11] This idea was developed in the 'Rahmenplan zur Umgestaltung und Vereinheitlichung des allgemeinbildenden öffentlichen Schulwesens' of the Deutscher Ausschuß für das Erziehungs- und Bildungswesen in 1959.

[12] The first Förderstufe-type school was already introduced in 1955 in Hessen in the so-called Schuldorf Bergstraße. Whether a 'support stage' was introduced at a specific school was instigated by the school authority (Schulträger) and the respective school.

[13] A further discussion of the idea of prolonged comprehensive schooling emerged after the formation of the 'German Education Council' (Deutscher Bildungsrat) in 1965. In 1970, the council suggested that a comprehensive 'orientation stage' following the four years of elementary school should cover grades five and six. This is especially documented in the 'Strukturplan für das Bildungswesen' from 1970. In the following years, representatives of all German Länder in the Bund-Länder-Kommission discussed how to organise this new school type. However, the projected system of homogenous nation-wide 'orientation stages' could not be enforced.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

subjects (integrierte Gesamtschule) or are allocated to an internal track according to their proficiency similar to the traditional school tracks (kooperative Gesamtschule). Nowadays, the acceptance of comprehensive schools largely varies between the German states: While there is only one comprehensive school left in Bavaria (as a remnant of the nation-wide experiment), it is widely established in the state of Berlin, for example.

2.2 Principles of Tracking in Hessen

The present study focuses on the state of Hessen since traditional secondary schools and the two year comprehensive orientation stages co-exist in this state. As a further alternative, the institution of the Gesamtschule offers fully comprehensive education from grade 5 – 10. In principle, after fourth grade, parents decide on the secondary school type of their children based on children's abilities and previous school performance (according to § 77 of the school law of Hessen). Parents may opt for the 'support stage' or a comprehensive school (Gesamtschule) in order to give their children more time to assess their abilities and interests. Especially, parents wishing that their children attend the higher secondary track (Gymnasium) but are not sure that they will be able to cope with the demands of this school type may make them join a 'support stage' or a comprehensive school. The distance between a pupils' place of residence and the location of the respective school is a further determinant that is known to drive the decision to attend a 'support stage' school vs. a tracked secondary school in fifth grade.[14] Some regions in Hessen do not offer 'support stages' so that children hardly have the choice to attend this school type.[15] However, the school law states that if the desired school type is not offered in a pupil's region of residence the pupil has the right to attend this school type in another region (cf. ì 70, school law of Hessen).

If the 'support stage' is chosen after fourth grade a decision on the final secondary track must be reached after sixth grade. Again, the parents have the primary authority to decide on the school type. However, if the desired track is the highest secondary school, selection to this school type depends on the 'support stage' teachers' approval.

2.3 Data Sets and Descriptive Analysis

This section presents some descriptive evidence indicating the quantitative dimension of the different tracking regimes and the streaming of pupils to the different secondary school types in Hessen. Further descriptive illustrations

[14] This is illustrated in Hessisches Kultusministerium (1995, 36).

[15] For example the city of Darmstadt offers no 'support stages' but those located at generally comprehensive schools.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

refer to the incidences of track modification and grade repetition[16] after pupils have been tracked by one or the other regime. Due to the pre-selection of different groups of pupils into the tracking regimes it is important to keep in mind that the presented stylized facts do not provide insights into the causal educational effects of one tracking regime compared to the other.

The following descriptive statistics are mainly based on newly available individual level data provided by the local statistical office of the state of Hessen. The data set covers all pupils enrolled in general schools in Hessen in the school years 2002 / 2003 – 2005 / 2006. At time of writing this paper, besides the official statistical tables, there exist only two empirical studies drawing on this data base (Puhani / Weber, 2007a and Puhani / Weber, 2007b). One drawback of the data is that it does not provide a panel, i.e. pupils cannot be tracked using an individual identification number. Thus, even if several data waves exist, my analysis is based on a cross-section of observations. Little information is given on the prior development of the pupils (prior grade and school type) and this only refers to the previous year.

While the advantage of the data set is its large number of observations, a clear general disadvantage is the limited number of reported variables for each individual. Besides variables indicating region, school and class, individual information is given on gender, birth year and month, school entry year and month, and nationality. There are no outcome variables such as school marks or test scores. However, it is possible to identify the incidences of grade repetition and track modification (the correction of initial track choice) from one year to the following year.

The results based on the Hessen data are later on complemented by evidence based on the national PISA-E database covering about 2,300 ninth graders in the German state of Hessen. The PISA-E data are a national extension of the international PISA 2000 data including supplementary questions from pupils and parents questionnaires as well as test results from the standardized math, reading and science tests. No information is available from school questionnaires which are included in the PISA study. The main reason why I use PISA-E instead of PISA is that information on 'support stage' attendance in fifth grade is only available in the extension study. Compared to the Hessen data, the advantage of PISA-E is that it allows controlling for a variety of individual background characteristics. This is why the econometric part of this study (Section 3) focuses on the PISA-E data.

According to the administrative data, nearly 13 % of all the primary and secondary schools in Hessen offer 'support stages' (206 out of 1,642 schools as observed in school year 2005 / 2006). These schools are either located at

[16] In Germany, low performing pupils have to repeat a grade if they are not able to attain certain marks.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

elementary schools (22 %), fully comprehensive schools (45 %) or at further school environments offering different educational tracks. Table 2 considers the school track choice of pupils being streamed after fourth grade in 2003 / 04 and of those who opted for the 'support stage' in 2003 / 04 and are tracked after sixth grade (in 2005 / 06) as observed in the administrative data set. The corresponding numbers are calculated using two different waves of the data so that both groups under examination consist of pupils from approximately the same cohorts. Results from Table 2 indicate that most of the fifth graders have already been tracked to the 'classical' secondary school levels: The majority of them attend the higher secondary track (38 %), while the intermediate and lower secondary levels are less popular (14 % and 5 % respectively). Furthermore, 15 % of all fifth graders attend fully comprehensive schools and 28 % opt for the 'support stages'. The latter group of pupils is mostly streamed to secondary levels after sixth grade (except of those 2 % who decide to attend fully comprehensive schools): Pupils tracked in seventh grade mostly enter the intermediate (46 %) or even the lower level (32 %) schools. There are no feasible gender differences when tracking to the secondary levels takes place after fourth grade. However, for the pupils tracked after the 'support stage', girls tend to choose higher educational tracks compared to their male classmates.

Table 2

Track choice in the earlier and in the later tracking regime (Hessen data)

Selection after / into	4th grade (tracking of *all* pupils)			6th grade (tracking of *support stage* pupils)		
	all (%)	male (%)	female (%)	all (%)	male (%)	female (%)
lower secondary	4.64	5.13	4.14	32.09	35.49	28.42
intermediate secondary	14.40	14.39	14.42	46.37	44.66	48.22
higher secondary	37.74	36.16	39.37	19.15	17.24	21.21
fully comprehensive	15.27	15.59	14.95	2.38	2.61	2.14
support stage	27.94	28.73	27.13	–	–	–

Note: Sample of all pupils tracked after fourth grade of elementary school in 2003 / 2004 and after sixth grade of the 'support stage' in 2005 / 2006 respectively.

Source: Administrative data for *Hessen,* wave 2003 / 2004 and 2005 / 2006, own calculations.

Table 3 repeats the descriptive analysis of secondary school track choice based on the PISA-E data (since this data set is used for the econometric part of this paper). Note, that the PISA-E data refers to an older cohort who is mostly tracked to secondary schools in 1995 (end of fourth grade) or 1997 (end of sixth grade). Thus, there are slight differences in the track choice as compared to the more recent evidence provided by the administrative data set.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Especially, the proportion of pupils in lower secondary schools in fifth grade is somewhat higher (9 %) while the proportion of fifth graders in the highest track is lower (31 %). The proportion of pupils opting for the 'support stage' is only slightly lower than in the more recent data (24 %). After sixth grade, this earlier cohort of 'support stage' pupils more often enters the intermediate track (53 %) compared to today's 'support stage' pupils. All in all, the general picture as demonstrated based on the administrative data set remains valid.

Table 3

Track choice in the earlier and in the later tracking regime (PISA-E 2000)

Selection after / into	4th grade (tracking of *all* pupils)			6th grade (tracking of *support stage* pupils)		
	all (%)	male (%)	female (%)	all (%)	male (%)	female (%)
lower secondary	8.77	10.14	7.27	17.91	20.38	15.53
intermediate secondary	16.31	16.68	15.90	53.45	52.22	54.63
higher secondary	31.13	32.00	30.17	20.30	18.51	22.03
fully comprehensive	17.38	16.75	18.09	8.34	8.89	7.81
further / unspecified	1.92	1.48	2.40	0.00	0.00	0.00
support stage	24.48	22.94	26.17	–	–	–

Note: PISA-E information refers to the track chosen after fourth grade of elementary school (1994 / 1995, retrospective information) and to the 'support stage' pupils' final track choice (observed in ninth grade in 2000). The proportions differ from the proportions based on the administrative data base, because different cohorts of pupils are considered.

Source: PISA-E 2000, own calculations.

The administrative data source additionally allows providing descriptive evidence by nationality group which is presented in Table 4. The two major subgroups under analysis are 'native' pupils (as defined by pupils holding nationalities of German-speaking countries) and pupils holding another nationality ('non-natives'). Furthermore, I look at the two most frequent immigrant groups, which refer to pupils holding Turkish (about 6 % of the considered fifth graders) or Italian and Greek nationalities (1.6 % of the sample).[17] I do not consider further nationality groups because of the smaller sample sizes of these groups.

While 'native' pupils are most often tracked to the highest secondary schools after fourth grade (41 %) a relatively small proportion of 'non-native' fifth graders attend these schools (19 % of all 'non-natives', only 13 % of pupils from Turkey and 18 % of pupils from Italy / Greece). Most pupils with an immigrant background opt for the 'support stages' (34 % of all 'non-natives', 38 % and 32 % for pupils from Turkey and Italy / Greece respectively). This is

[17] The data at hand do not allow distinguishing between Greek and Italian nationals.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

consistent with the idea that these schools give them more time to integrate and learn the German language before having to decide on their educational (and professional) future.

Table 4

Track choice by nationality

Selection after / into	4th grade (tracking of *all* pupils)				6th grade (tracking of *support stage* pupils)			
	native	non-native	Turkey	Italy / Greece	native	non-native	Turkey	Italy / Greece
lower secondary	3.66	10.53	10.78	11.25	28.65	49.26	52.64	53.57
intermediate sec.	13.74	18.38	20.05	17.19	47.99	38.31	37.12	37.14
higher secondary	40.96	18.56	13.00	18.02	20.99	9.97	7.54	6.79
comprehensive	14.69	18.72	18.59	21.67	2.37	2.45	2.71	2.50
support stage	29.96	33.81	37.58	31.87	–	–	–	–

Note: Sample of all pupils tracked after fourth grade of elementary school in 2003 / 2004 and after sixth grade of the 'support stage' in 2005 / 2006 respectively.

Source: Administrative data for *Hessen,* wave 2003 / 2004 and 2005 / 2006, own calculations.

The educational decision after the 'support stages' differs between immigrants and natives as well: While the highest proportion of natives reaches the intermediate secondary track after attending the 'support stages' (48 %), immigrants are most often selected to the lowest secondary schools (49 % of all 'non-natives', even 53 % of pupils from Turkey and 54 % of pupils from Italy / Greece).

Table 5 and Table 6 aim at answering the question whether modification of the initial track choice and grade repetitions are unusual if pupils are tracked after six instead of four years of comprehensive schooling. As described above, one rationale behind the 'support stages' is that children are given more time to develop their abilities and skills and to obtain more information on their educational performance before deciding on the secondary track. If it is true that tracking after sixth grade is based on more reliable information on the pupils' abilities, one would expect that ex-post modification of the initially chosen track and grade repetitions are not frequent under the later tracking regime.

Thus, Table 5 shows the proportions of pupils staying in the chosen track in fifth, sixth and seventh grade. As explained in Section 2, it is generally possible to modify the initially chosen track at any grade level, whilst track modification is somewhat complicated by different curricula at different school types. Note, that the data at hand are not available as a panel. Thus, it is principally not possible to observe individuals over time in order to determine whether the track modification behaviour of former 'support stage' pupils differs from other pupils. However, I use information on the shares of former 'support stage' pupils being in the respective school at a given grade level.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 5

Proportions of stayers in school tracks by previous 'support stage' attendance

	No incoming support stage pupils (0 %)			High share of incoming support stage pupils (> 80 %)		
All Track Types						
Stayers after ...	ratio	(s.d.)	observ.	ratio	(s.d.)	observ.
... 5th grade (2003 / 04)	0.98	(0.14)	15,938	–	–	–
... 6th grade (2004 / 05)	0.98	(0.13)	16,053	–	–	–
... 7th grade (2005 / 06)	0.98	(0.14)	15,937	0.96	(0.18)	13,877
Lower Secondary						
... 5th grade (2003 / 04)	0.97	(0.17)	1,640	–	–	–
... 6th grade (2004 / 05)	0.96	(0.19)	1,859	–	–	–
... 7th grade (2005 / 06)	0.98	(0.13)	1,975	0.99	(0.11)	4,561
Intermediate Secondary						
... 5th grade (2003 / 04)	0.95	(0.23)	3,539	–	–	–
... 6th grade (2004 / 05)	0.96	(0.21)	3,579	–	–	–
... 7th grade (2005 / 06)	0.95	(0.21)	3,620	0.96	(0.19)	6,455
Higher Secondary						
... 5th grade (2003 / 04)	0.99	(0.09)	10,759	–	–	–
... 6th grade (2004 / 05)	0.99	(0.08)	10,615	–	–	–
... 7th grade (2005 / 06)	0.99	(0.10)	10,342	0.94	(0.24)	2,861

Note: The 'proportions of stayers' indicate the number of pupils in the given school type divided by the number of pupils in the given school type who have already been in this school the year before. Only pupils in tracked school types moving from one grade to the following grade (for example from grade 5 to grade 6 in 2003 / 2004) are considered. The total number of pupils in a given grade is not equal to the total number of pupils in the previous grade times the proportion of stayers since grade retainees additionally lower the number of remaining pupils. Pupils dropping out of the school system or moving to another German state are not observed, grade retainees are not considered. Proportions are separately calculated for schools with no incoming 'support stage' pupils and schools with high shares of incoming 'support stage' pupils. The share of incoming pupils from the 'support stages' is calculated by the proportion of seventh graders in the respective school in 2004 / 2005 having attended 'support stages' in sixth grade. The proportions are very similar (and thus robust) if grade retainees are kept in the sample.

Source: Administrative data for *Hessen,* waves 2003 / 2004 to 2005 / 2006, own calculations.

Table 5 distinguishes between schools having no incoming pupils from 'support stages' in grade seven and those having high shares (80 % or more) of incoming 'support stage' pupils. Since the number of incoming 'support stage' pupils differs by school track, I additionally distinguish between school tracks.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 6

Proportions of retained pupils by share of incoming 'support stage' pupils

	No incoming support stage pupils (0 %)			High share of incoming support stage pupils (> 80 %)		
All Track Types						
Retainees in ...	retained	(s.d.)	observ.	retained	(s.d.)	observ.
... 5th grade (2003 / 04)	0.03	(0.17)	16,417	–	–	–
... 6th grade (2004 / 05)	0.03	(0.16)	16,480	–	–	–
... 7th grade (2005 / 06)	0.04	(0.20)	16,550	0.07	(0.25)	14,789
Lower Secondary						
... 5th grade (2003 / 04)	0.07	(0.26)	1,765	–	–	–
... 6th grade (2004 / 05)	0.06	(0.23)	1,973	–	–	–
... 7th grade (2005 / 06)	0.08	(0.27)	2,261	0.08	(0.28)	5,070
Intermediate Secondary						
... 5th grade (2003 / 04)	0.04	(0.20)	3,693	–	–	–
... 6th grade (2004 / 05)	0.04	(0.20)	3,736	–	–	–
... 7th grade (2005 / 06)	0.06	(0.24)	3,806	0.07	(0.26)	6,970
Higher Secondary						
... 5th grade (2003 / 04)	0.02	(0.13)	10,959	–	–	–
... 6th grade (2004 / 05)	0.01	(0.12)	10,771	–	–	–
... 7th grade (2005 / 06)	0.03	(0.17)	10,483	0.03	(0.16)	2,749

Note: The 'proportions of retained pupils' indicate the number of pupils attending the same grade as in the previous year divided by the number of pupils at the given grade level. Only pupils in tracked school types are considered. Pupils dropping out of the school system or moving to another German state are not observed. Retainees include pupils changing to another track if they are repeating the grade in this track. Proportions are separately calculated for schools with no incoming 'support stage' pupils and schools with high shares of incoming 'support stage' pupils. The share of incoming pupils from the 'support stages' is calculated by the proportion of seventh graders in the respective school in 2004 / 2005 having attended 'support stages' in sixth grade.

Source: Administrative data for *Hessen,* waves 2003 / 2004 to 2005 / 2006, own calculations.

Generally, for the schools not educating any former 'support stage' pupils, the proportion of pupils staying in the previously chosen school type when moving to the following grade after a given grade amounts to 98 % in grades five, six, and seven. The proportion of stayers is lower (96 %) in the seventh grade for schools primarily recruiting former 'support stage' pupils. The difference in the proportion of stayers between schools not educating any 'support stage' pupils and schools primarily educating 'support stage' pupils is especially high in the highest secondary school track: While 99 % of the se-

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

venth graders remain in the highest level school track in the schools without former 'support stage' pupils, only 94 % are stayers in the schools featuring a high proportion of former 'support stage' pupils. Even if one takes into account that the seventh graders in the first type of schools (no 'support stage' pupils) possibly already revised their initial track decision after grades five and six, the figure of six percent of track changers in the second type of schools (featuring a high share of 'support stagers') is comparably high.

All in all, a relatively high proportion of pupils in the higher secondary track decide to revise the track decision made after the 'support stages'. While a primary objective of the 'support stages' is the optimisation of school track choice through a longer period of observation and support in the comprehensive system, the changer rates following the tracking grade suggest that the 'support stage' based decisions may not be as appropriate as expected. However, it must be noted again that this descriptive evidence does not provide causal effects of the tracking regime in the statistical sense but merely looks at the educational performance of self-selected groups of pupils who have chosen one or the other tracking regime.

Table 6 additionally presents proportions of grade retainees (pupils who have to repeat a grade due to poor performance) following the same strategy as Table 5 above. A casual examination of the first set of rows in Table 6 gives the impression that the proportion of pupils not succeeding in the given grade is especially high for schools with high shares of incoming 'support stage' pupils. However, if the proportion of retained pupils is calculated by school track type (see the next sets of rows in Table 6) it is shown that the high proportion of retainees in schools receiving high shares of former 'support stage' pupils is due to the fact that these schools are mainly at the lower or intermediate secondary level. There are no feasible differences in the proportions of retained pupils if the comparison relates to schools of the same track type.

3. Econometric Strategies and Regression Results

3.1 Identification Strategy and Specifications for the Econometric Analysis

If the tracking regime were randomly assigned, the causal effect of 'support stage' attendance on educational outcomes could be estimated using a simple OLS regression framework. The corresponding regression equation is given by:

$$Y_i^t = \beta X_i + \gamma S_i + \varepsilon_i \,, \tag{1}$$

where Y_i^t is the educational outcome of individual i measured at time t (several years after the regime choice), X_i is a vector of explanatory variables, S_i refers to the tracking regime indicator, and ε_i is the error term. However, as stated

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

above, the prior choice of the tracking regime is endogenous to educational outcomes. One may assume that pupils choosing to attend the 'support stages' differ from the average pupil in (unobserved) characteristics which are also related to the schooling outcome so that $corr(S_i, \varepsilon_i) \neq 0$. For example, ambitious parents often decide that their children attend the 'support stages' if the children did not perform well enough in elementary school to suggest immediate tracking to the highest educational stream (cf. section 2.2). Thus, it can be expected that estimating the effect of 'support stage' attendance on later educational outcomes by OLS will yield (negatively) biased results.[18] Given the available data-sets, the feasible strategy to pin down the effect of 'support stage' attendance is as follows: Formally, I assume that the true model equation is

$$Y_i^t = \beta X_i + \gamma S_i + \delta U_i \,, \tag{2}$$

where U_i refers to a vector of non-controlled variables determining both the tracking regime choice after fourth grade and educational outcomes at a later point in time. The corresponding estimation equation is:

$$Y_i^t = \beta X_i + \gamma S_i + \delta U_i + u_i \,, \tag{3}$$

where $corr(S_i, u_i) = 0$. Thus, the underlying problem is taken to be an omitted variable problem where the error term in equation 1 contains both the influences of the characteristics (δU_i) and the error term of equation 3 (u_i). The feasible solution to this problem is to control for as many of the variables (U_i) causing the bias as possible using a relatively rich data set (the PISA-E data) on the pupils' individual and family background.

Table 7 gives an overview of the different specifications used in the regression analysis. Specification 1 simply includes the dummy variable of interest (indicating whether the pupils attended the 'support stage' regime) and a control dummy variable for attending the fully comprehensive system. In other

[18] One standard solution to such an endogeneity problem is to apply an instrumental variable strategy. The crux is whether it is possible to find a valid instrument which explains 'support stage' attendance but is not correlated to unobservable characteristics driving the outcome variable. In my opinion, it is not possible to find a valid instrument. One potential instrument that springs to mind is the density of 'support stages' in a region: Using this instrument it is assumed that pupils are more likely to decide to opt for the 'support stage' regime if there are many 'support stage' schools in their county of residence. However, the provision of 'support stages' cannot be considered as exogenous to educational outcomes: The local 'support stage' density is potentially driven by the same or similar characteristics of a region's residents as the individual decision to attend the 'support stage'. Conducting regressions on the local provision of 'support stages' using county data shows that the local 'support stage' density is significantly determined by observable regional variables which are also thought to be important determinants of educational outcomes (for example income and wealth variables).

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 7

Variables used in the different specifications

Variable	Explanation
Tracking Regime Indicators (Reference = Tracking after fourth grade):	
support stage	dummy variable for 'support stage' attendance in fifth grade
comprehensive school	dummy for comprehensive school attendance in fifth grade
Variables Added in Specification 2 (Individual Characteristics):	
gender	dummy for male gender
immigration	dummy indicating whether pupil or parents were born abroad
proxy for school entry age	dummy indicating whether pupil is born before the official school entry cut-off date of June (= theoretically entered school relatively young according to the official school entry rule)[a)]
Variables Added in Specification 3 (Family Characteristics):	
father	dummy indicating whether only a male guardian (mostly the father) lives with the child
mother	dummy indicating whether only a female guardian (mostly the mother) lives with the child
employment of mother	dummy indicating whether the mother is employed
employment of father	dummy indicating whether the father is employed
mother: no vocational education[b)]	dummy indicating whether mother does not hold a vocational degree
mother: tertiary education[b)]	dummy indicating whether mother holds a tertiary educational degree
father: no vocational education[c)]	dummy indicating whether mother does not hold a vocational degree
father: tertiary education[c)]	dummy indicating whether mother holds a tertiary educational degree
parental reading encouragement	parents often read to child before child learned to read
siblings	dummy indicating whether there are siblings of the child

Note: [a)] See the paper by Puhani / Weber (2007a) for the motivation of this variable. [b)] The reference category are mothers holding a vocational (upper secondary) degree. [c)] The reference category are fathers holding a vocational (upper secondary) degree. In addition to these variables dummy variables for missing information are included.

words: the regression results differentiate between effects of three options of tracking regimes (i.e. the earlier and the later tracking regime and the comprehensive system). Individual characteristics (gender, immigration background and a proxy for school entry age) are added in specification 2. Specification 3 additionally includes family background variables (i.e. indicating the presence

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

of parents at home, parental employment, education, and behavior and the presence of siblings). I assume that the endogeneity bias is reduced as one moves from specification 1 to specification 3. Especially, the variables added in specification 3 are mainly parental characteristics that influence the tracking regime choice as well as the children's educational outcomes. Ideally one would also directly control for initial ability of pupils, i.e. compare pupils who performed similarly before entering the different tracking systems. However, no appropriate performance measure is available in the data.[19]

A further issue is that in the PISA-E data there are missing observations for the variables of interest for some pupils. For each of the control variables up to five percent of the observations are missing. For parental education even 12 % (mother) and 16 % (father) of the observations are generally missing. Given that this might additionally bias the results, in the following regression analysis, I include dummy variables indicating missing observations.

In order to measure test results I use the averages of the plausible values of test scores which are given in PISA-E. For detailed information on the scaling of the PISA test results and test contents I refer the reader to the technical reports and documentaries (Adams / Wu, 2002 and especially the publication by Deutsches PISA Konsortium, 2003 for the German extension study). The plausible values correspond to the ones measured in the PISA-study but are standardized for each German state so that the mean score equals 100 and the standard deviation is 30 for each state. Thus, comparisons of test results across German states are not possible and analyses must be conducted at the single state's level.[20] For the sake of representativeness, all statistics are weighted using the sampling weights provided in the data-set. I also use clustered robust linear regressions with schools as clusters in order to avoid an underestimation of the standard errors since the variable of interest is on a higher aggregation level than the other variables.

I conduct several robustness checks in order to secure the validity of the results. Since pupils' observed and unobserved characteristics differ strongly between school types, one robustness check compares pupils who have been in 'support stages' to students who have been in the intermediate track in fifth grade (assuming that this is a more suited reference group).[21] Furthermore, in

[19] The only potential measure is the school level the pupil had been recommended to attend after fourth grade. For pupils attending the 'support stages' the indicated level might also be the one recommended after sixth grade and thus be an outcome of 'support stage' attendance. This is why I do not use this information.

[20] In the original PISA study scores are standardized to an international mean 500 and standard deviation 100 which allows international comparisons.

[21] I do not follow this approach as my main estimation strategy since I have no way to assure that pupils selecting (based on unobserved characteristics) to a specific track already in fifth grade are in fact comparable to pupils in the 'support stages'. However, restricting the sample to this group is still an insightful robustness check.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

order to secure that the findings are not driven by the parametric assumptions of the linear regression model, I use matching estimators as alternative estimation strategies. I will briefly mention the results from all these robustness checks when discussing the regression results.

So far, I only estimate the impact of later tracking at the mean of the conditional performance distribution. As mentioned in the introduction, from a theoretical point of view, there are counteracting effects of later tracking: While later tracking may result in a more appropriate tracking decision because of improved information concerning the children's proficiency, more proficient pupils may actually benefit from early tracking for example through positive peer effects. Thus, it is interesting to examine whether the later tracking effect differs for pupils with a different background and of different ability. Therefore, the presentation of regression results is complemented by sub-group analyses focusing on pupils' family background. Additionally, quantile regressions are conducted in order to directly consider pupils at different positions of the conditional distributions of test scores.

3.2 Regression Results

Table 8 shows the results of OLS regressions of test performance on tracking regime dummies and different sets of explanatory variables (as explained in Table 7).[22] Generally, all the estimated effects are negative if they are significant. This might indicate that the attendance of a comprehensive class in fifth grade reduces school performance in ninth grade but the negative coefficients might also be the result of a negative selection of pupils into the comprehensive regimes after fourth grade. Including individual control variables in specification 2 hardly changes the point estimates compared to specification 1. However, if parental background is considered in specification 3, the estimated coefficients decrease notably and become insignificant in most cases (except for the significance of the 'support stage' coefficient in the science regression and the coefficient on the comprehensive school indicator in the math regression).

The decrease in the absolute size of the negative coefficients as one moves from specification 2 to specification 3 reflects the 'negative selection' to the comprehensive school systems, i.e. pupils with a less favourable socio-economic background select to these systems.[23] This finding corresponds to a

[22] In addition to the presented regressions, I also conducted regressions where I allowed for a more flexible form by interacting the 'support stage' dummy and the explanatory variables. However, hardly any of the interaction coefficients proved to be significant in the full specification. Alternatively, I consider effects for some socio-economic sub-groups which will be discussed below.

[23] Section 2 demonstrated that especially pupils with an immigrant background select to these schools.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 8

Results of OLS regressions of PISA-E scores on 'support stage' attendance

			Maths			Reading			Science		
			all	female	male	all	female	male	all	female	male
SPECIFICATION	1	support stage	–5.90* (3.03)	–6.04 (4.52)	–5.28 (3.60)	–4.12 (2.84)	–3.23 2.94)	–5.63* (3.29)	–7.17* (3.80)	–6.11 (4.32)	–7.95* (4.07)
		comprehensive	*–6.65** (3.29)*	*–10.44* (4.03)*	*–2.91 (3.74)*	*–2.67 (2.66)*	*–4.67 (3.11)*	*–1.27 (3.12)*	*0.71 (2.84)*	*0.35 (3.10)*	*1.14 (3.82)*
	2	support stage	–5.38* (3.04)	–5.44 (4.50)	–5.95* (3.33)	–4.39* (2.63)	–3.37 (2.73)	–5.74* (3.06)	–8.48** (3.83)	–7.52* (4.42)	–9.54** (4.01)
		comprehensive	*–7.24** (3.16)*	*–10.89** (4.21)*	*–4.51 (3.31)*	*–3.28 (2.49)*	*–4.61 (2.97)*	*–2.26 (2.92)*	*0.59 (2.63)*	*0.34 (3.07)*	*0.03 (3.45)*
	3	support stage	–1.94 (2.47)	–0.32 (3.32)	–3.35 (2.91)	–1.08 (2.04)	0.98 (2.32)	–3.43 (3.48)	–5.25* (3.11)	–2.48 (3.44)	–7.98** (3.41)
		comprehensive	*–4.68* (2.63)*	*–7.06* (3.76)*	*–2.88 (2.94)*	*–0.96 (1.97)*	*–1.49 (2.45)*	*–0.91 (2.39)*	*2.45 (2.48)*	*2.25 (2.94)*	*0.91 (3.19)*
observations			1,222	548	674	2,306	1,074	1,232	1,262	577	685
# support stage			245	114	131	464	224	240	261	117	144
# comprehensive			208	96	112	386	190	196	196	90	106

Note: The reported coefficients (standard errors in parantheses) refer to the 'support stage' dummy and the dummy variable for attendance of a general comprehensive school in fifth grade. The different specifications are explained in Table 7. * Significant at the ten percent level. ** Significant at the five percent level.

Source: PISA-E 2000, own estimations.

situation where low performers at elementary school who are recommended to the lower level schools opt for the comprehensive system in order to get a 'second chance' to find out whether they still have the ability to attend the high (or intermediate) level track.

Furthermore, the low and mostly insignificant effects for specification 3 indicate that the choice of the tracking system does not matter at least for the math and reading outcomes of ninth graders. Even if the identification strategy does not allow for the identification of the true causal effect of the tracking regime, because of the negative selection into the comprehensive systems (as indicated by the change in coefficients between specification 2 and 3) there is no reason to believe that the presented coefficients suffer from a downward bias. Thus, it is reasonable to conclude that there is no negative effect of 'support stage' (or comprehensive school) attendance on fifth graders math (or science) and reading performance.

Generally, analysis by each gender yields similar findings as for the whole sample with the main conclusion that the 'support stage effect' drops down (mostly insignificant) if the full set of controls is included. However, there is one notable exception: For male pupils the negative science score effect does

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

not decrease as more controls are included. Still, the methodological framework of this paper does not allow identifying whether the persistent negative effect concerning the science score is due to education in the 'support stage' or due to a persistent selection bias caused by unobserved characteristics.

Tables 9 to 11 repeat the regressions for different sub-samples characterised by immigrant background and parental characteristics. Considering pupils with and without immigrant background, the following pattern emerges: For natives the 'support stage' effects generally decrease as the full set of controls is included where only the effect on the math scores remains significant at the ten percent level. For immigrants the effect is insignificant or becomes insignificant in all cases. Notably, the point estimate on immigrant pupils' maths and reading score effect even turns positive when using specification 3. If it is assumed that there is negative selection of pupils to the 'support stages' this finding suggests that there is a positive regime effect. Consequently, the results could be interpreted as indicating that immigrant pupils benefit from being educated in the later tracking regime. However, the estimated effects are not significant at the ten percent level.

Table 9

Regression results by immigration background

			Maths		Reading		Science	
			coefficients (s.e.)		coefficients (s.e.)		coefficients (s.e.)	
			native	immigrant	native	immigrant	native	immigrant
SPECIFICATION	1	support stage	–9.29** (3.13)	–0.56 (5.29)	–6.70** (3.18)	0.45 (4.07)	–7.63 (4.84)	–11.61** (5.01)
		comprehensive	*–9.12** (3.57)*	*–4.96 (4.81)*	*–3.75 (2.64)*	*–2.19 (3.91)*	*0.14 (3.20)*	*1.53 (3.49)*
	2	support stage	–9.28** (3.08)	0.20 (5.13)	–6.89** (3.11)	0.58 (3.89)	–7.52 (4.59)	–10.52** (4.92)
		comprehensive	*–8.92* (3.42)*	*–4.35 (4.94)*	*–3.74 (2.62)*	*–2.54 (3.89)*	*–0.18 (3.09)*	*1.46 (3.49)*
	3	support stage	–4.87* (2.76)	2.59 (3.83)	–3.11 (2.38)	4.19 (3.16)	–4.89 (3.56)	–5.13 (4.00)
		comprehensive	*–6.96** (2.91)*	*0.42 (4.11)*	*–2.27 (2.20)*	*3.20 (2.80)*	*0.47 (2.93)*	*6.67* (3.47)*
observations			802	420	1,562	744	866	396
# support stage			169	76	329	135	192	69
# comprehensive school			148	60	274	112	136	60

Note: 'Immigrant' refers to pupils who were born abroad or whose parents were born abroad (compare Table 7). The reported coefficients refer to the 'support stage' dummy and the dummy variable for attendance of a general comprehensive school in fifth grade. The different specifications are explained in Table 7. * Significant at the ten percent level. ** Significant at the five percent level.

Source: PISA-E 2000, own estimations.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Table 10

Reading regression results for different groups of immigrants

		First generation immigrants	Second generation immigrants	Foreign language spoken at home	First generation immigrants + foreign language spoken at home	Second generation immigrants + foreign language spoken at home
SPECIFICATION	1 support stage	6.52	–5.98	–1.89	4.29	–15.43**
		(4.18)	(6.97)	(3.80)	(3.98)	(6.10)
	comprehensive	*2.94*	*–7.50*	*–0.42*	*3.22*	*–5.96*
		(4.61)	*(5.01)*	*(4.09)*	*(4.71)*	*(6.18)*
	2 support stage	6.62	–6.18	–1.66	4.35	–15.67**
		(4.11)	(6.98)	(3.78)	(4.03)	(6.30)
	comprehensive	*2.45*	*–8.17*	*–0.49*	*2.79*	*–6.62*
		(4.68)	*(4.94)*	*(4.06)*	*(4.83)*	*(5.84)*
	3 support stage	10.22**	–2.71	1.77	7.10*	–7.96
		(3.64)	(4.35)	(3.36)	(3.86)	(5.42)
	comprehensive	*5.45*	*0.76*	*2.61*	*4.70*	*0.63*
		(4.07)	*(3.88)*	*(3.53)*	*(4.14)*	*(4.90)*
observations		386	358	539	334	176
# support stage		82	53	101	68	28
# comprehens. school		53	59	88	48	32

Note: Results are only presented for the reading sample, because sample sizes are even smaller for the science and mathematics test. The reported coefficients refer to the 'support stage' dummy and the dummy variable for attendance of a general comprehensive school in fifth grade. The different specifications are explained in Table 7. * Significant at the ten percent level. ** Significant at the five percent level.

Source: PISA-E 2000, own estimations.

Furthermore, it might be argued that the conclusion of a positive regime effect only holds if there is in fact negative selection of immigrant pupils to the 'support stages'. For the reading test, this assumption would not be valid if immigrant pupils with initially higher language skills (pupils who have spent longer time in Germany and use the German language at home) self-selected to the 'support stages'.[24] In order to take this objection into account, I estimate the 'support stage' effect separately for different groups of immigrants. The considered groups are: (1) pupils who were born abroad (mostly first genera-

[24] Generally , it might be argued that schools are more selective with respect to the early tracking decision if teachers and parents anticipate that pupils have a second chance of catching up to a higher level school track after attending the 'support stage'. However, it has been demonstrated in section 2.3 that a relatively high proportion of pupils in Hessen is streamed to the highest track in fifth grade. This is a first hint that the tracking policy in Hessen system is not specifically restrictive.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

tion immigrants), (2) pupils born in Germany whose parents were born abroad (second generation immigrants), (3) pupils who use a foreign language at home, (4) first generation immigrants who use a foreign language at home, and (5) second generation immigrants speaking a foreign-language at home. It is reasonable to assume that initial reading performance is better for second generation immigrants compared to first generation immigrants and especially compared to first generation immigrants speaking a foreign language at home.

Table 11

Reading regression results according to family background

		Both parents not working	Both parents low educated	Less favourable family background	Favourable family background
SPECIFICATION	1 support stage	6.20 (7.51)	13.95** (5.56)	0.57 (3.43)	–7.32** (3.16)
	comprehensive	*3.16* *(7.61)*	*11.30* *(7.60)*	*–1.56* *(3.55)*	*–3.51* *(2.65)*
	2 support stage	10.91 (7.50)	8.28 (5.72)	0.48 (3.33)	–7.32** (3.08)
	comprehensive	*5.38* *(7.43)*	*5.35* *(7.45)*	*–1.96* *(3.53)*	*–3.33* *(2.66)*
	3 support stage	18.82** (8.56)	9.62 (6.87)	4.30 (2.95)	–4.16* (2.44)
	comprehensive	*10.52* *(8.94)*	*4.98* *(8.44)*	*2.23* *(3.01)*	*–2.33* *(2.22)*
observations		123	104	846	1,407
# support stage		26	20	158	291
# comprehens. school		19	26	132	244

Note: Results are only presented for the reading sample, because sample sizes are even smaller for the science and mathematics test. The reported coefficients refer to the 'support stage' dummy and the dummy variable for attendance of a general comprehensive school in fifth grade. The different specifications are explained in Table 7. * Significant at the ten percent level. ** Significant at the five percent level.

Source: PISA-E 2000, own estimations.

The respective reading score results by immigrant sub-group are presented in Table 10. Results for the maths and science scores (not shown) are very similar to the presented evidence but mostly insignificant due to very reduced sample sizes.[25] Generally, most of the findings considered are insignificant which might be due to limited sample sizes when considering sub-groups. However, looking at the point estimates, familiar patterns emerge for all sub-

[25] All results are presented in the discussion paper version of this paper and available from the author.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

groups and subjects: If the 'support stage' effect is negative in the initial specification (without control variables) it decreases in absolute size or turns insignificant or positive in the full specification. For some sub-groups (first generation immigrants and first generation immigrants using a foreign language at home for reading) the 'support stage' effect is positive even if no control variables are included. In these cases, the positive effect becomes more pronounced (and is significant for the reading score) if the full set of control variables is included. Interestingly, the positive 'support stage' effect in reading is especially high for first generation immigrants and first generation immigrants using a foreign language at home who might be considered to be a 'negative selection' (as concerns their initial reading skills) among the group of immigrant pupils. Since the positive effect becomes more pronounced as additional control variables are included, this is indicative of a negative selection bias being reduced. Summing up, I interpret these robust and consistent findings as supportive for the conclusion that 'support stages' are beneficial for the reading performance of immigrants.

Sub-group results by parental background are presented in Table 11. The considered groups are: (1) Children whose both parents are not employed, (2) children whose both parents do not hold a vocational degree, (3) children with a general 'disadvantaged' family background (children having either an immigrant background or having low educated or unemployed parents) and (4) children with an 'advantaged' family background (children having no immigrant background, no unemployed parent and no lowly educated parent). Since sample sizes drop to very small numbers for most of the sub-groups, I only present the results for the reading sample which is the largest sample. As a matter of fact, due to the limited sample size most of the sub-group results for the mathematics and science samples are insignificant (not shown here) but the general pattern emerging from these samples corresponds to the findings from the reading sample. The numbers of observations are already very limited for the reading regressions as can be deduced from Table 11. However, the results provide some interesting insights: First of all, and similar to Table 10 the 'support stage' effects are generally positive for the full specification when groups with a 'disadvantaged' family background are considered (in the first three columns of Table 11). The strong positive effect for children whose parents are not employed turn significant in the full specification (the effects in the second and third column are only significant on the 16 % and 15 % level respectively). However, if children with a favourable family background are examined, the point estimates are (again) negative and turn insignificant in the full specification. Thus, it seems that later tracking exerts different effects on different groups of children. If it is true that children with a less favourable family background benefit from the 'support stages' while this institution does not harm pupils with an advantaged family background, as it is suggested by these results, 'support stages' might reduce education inequality.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

These findings are also confirmed by several robustness checks based on matching estimators instead of regression analysis. First of all, based on nearest neighbor matching (based on all variables as in specification 3), I estimated the 'support stage' effect for the entire sample of ninth graders, where pupils in any school type apart from the 'support stage' in fifth grade are in the control group.[26] The point estimates for maths, science and reading are insignificant and rather low (ranging from -2.19 for science to 1.02 for maths). Sub-group analyses reveal the same pattern as for the regression analysis: Based on the reading sample, pronounced and significant positive effects are found for parents whose children are not employed (27.88 score points), first generation immigrants (10.26 points) and children with a general disadvantaged family background (7.31 points). Results for all other sub-groups by family background are not significant.

Furthermore, I re-estimated the effects only considering pupils in the intermediate track in fifth grade in the treatment group. Additionally and as a further check, I included not only pupils in the 'support stage' in fifth grade but all comprehensive types in fifth grade in the treatment group. Again, results for the sub-groups are either insignificant or positive and the patterns observed in the regression analysis are confirmed.

Distributional considerations are directly addressed using quantile regressions (Table 12). Figure 1 – Figure 3 show the estimated 'support stage' effects for different quantiles of the conditional test score distributions together with the mean regression results and its confidence bounds.[27] An interesting pattern emerges for all test scores: While there are significant positive 'support stage' effects for the lower quantiles, the effect decreases nearly monotonically and turns to a significant positive effect for the upper quantiles. For the 10 %-quantile for example the positive effect ranges between 5.35 scores for science and 6.65 for the reading score; this is equivalent to about one-fifth of the PISA-E standard deviation in the sample for Hessen. Looking at the 90 %-quantile, the effect is also sizeable and ranges between –4.58 (science) and –4.14 (reading) which corresponds to about 15 % of a standard deviation.

Thus, the quantile regression results suggest that 'support stages' work in favor of children with a disadvantaged education background whilst there are negative effects on pupils on top of the conditional performance distribution. Therefore, 'support stages' might reduce education inequality to the detriment of pupils on top of the (conditional) performance distribution. These findings are consistent with results from studies comparing tracking systems for differ-

[26] Matching results are estimated based on STATA's 'attnd' estimation calculating the average treatment effect on the treated using nearest neighbor matching. The results of all robustness checks are available from the author upon request.

[27] Standard errors are obtained via bootstrapping using STATAs 'sqreg'.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ent countries concluding that later tracking reduces education inequality (compare Section 1). Additionally, the theoretical literature on tracking provides explanations for the fact that tracking exerts differential impacts on pupils of different abilities: For example non-linear peer effects imply that high ability pupils specifically benefit from early segregation.

Table 12

Quantile regression results

Quantiles	Maths	Reading	Science
0.10	6.23**	6.65**	5.35**
	(2.11)	(2.24)	(2.36)
0.20	3.94*	2.03	2.92
	(2.21)	(2.32)	(2.11)
0.30	3.06	2.68	2.32
	(2.14)	(2.09)	(1.64)
0.40	2.67	2.48	1.74
	(1.78)	(1.92)	(2.03)
0.50	1.32	1.48	1.97
	(1.63)	(1.78)	(1.78)
0.60	–0.94	–0.71	0.35
	(1.65)	(1.79)	(1.65)
0.70	–2.57	–1.58	–1.39
	(1.86)	(1.83)	(1.94)
0.80	–3.73**	-3.68	-3.72*
	(1.77)	(2.05)	(1.94)
0.90	–4.25**	-4.14**	-4.58**
	(2.13)	(1.96)	(2.00)
0.99	–8.45**	–9.79**	–9.98**
	(2.17)	(2.36)	(2.32)

Note: The reported coefficients refer to the 'support stage' effect in the regressions using all control variables. Numbers in parentheses are the bootstrapped standard errors. The effects are also illustrated in Figure 1 – Figure 3. * Significant at the ten percent level. ** Significant at the five percent level.

Source: PISA-E 2000, own estimations.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

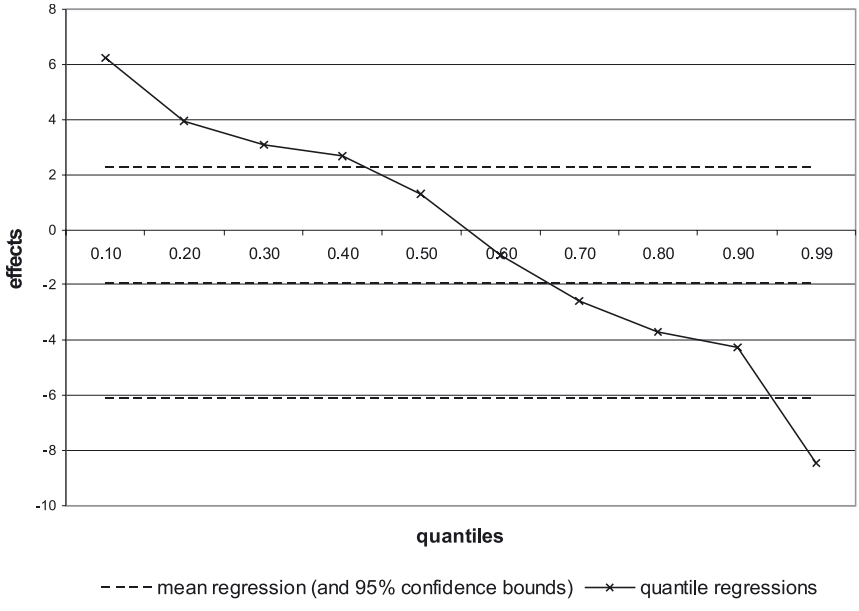

Figure 1: 'Support stage' effects on PISA-E maths scores by quantiles

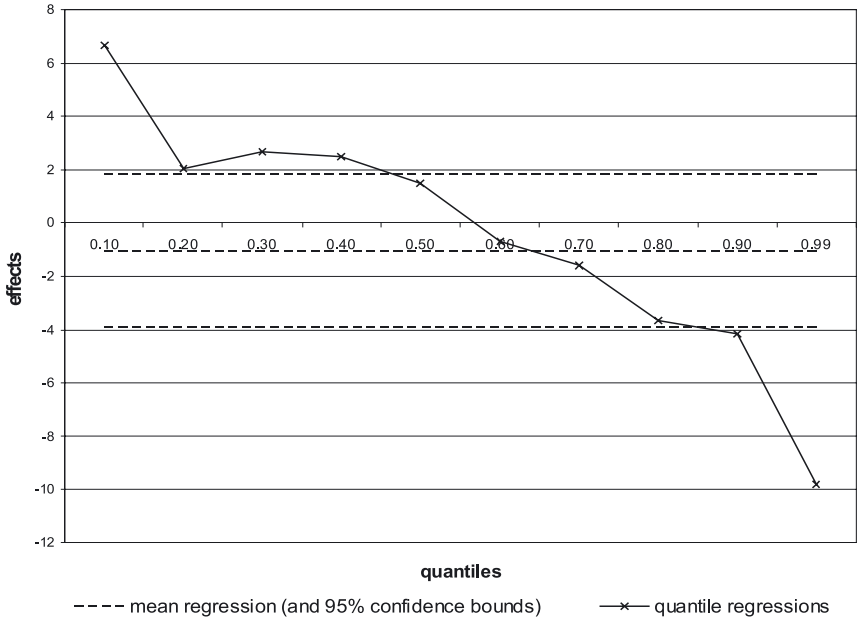

Figure 2: 'Support stage' effects on PISA-E reading scores by quantiles

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

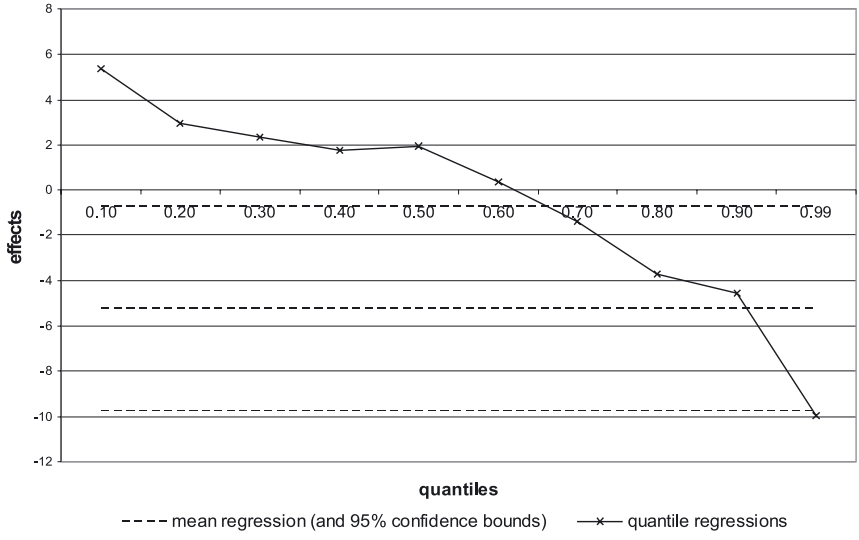

Figure 3: 'Support stage' effects on PISA-E science scores by quantiles

4. Conclusions

The optimal tracking system is an issue of controversial discussion among educationalists and social scientists. This paper considered an alternative tracking regime which allows streaming pupils to secondary school types after six instead of four years in the German state of Hessen. It has been argued that pre-selection into the alternative tracking regime (the 'support stages') is not random. It seems that especially lower performers are selected to the later tracking regime. Thus, it is not surprising, that children attending the 'support stages' are more often tracked to the lower secondary school types later, as can be seen from the descriptive statistics.

In an attempt to reduce the endogeneity bias in estimating the regime choice effect, I controlled for a variety of individual and family characteristics such as parental education, employment and behavior. Overall, the estimated negative coefficients on the 'support stage' or comprehensive school indicators drop in absolute size as one controls for family background (and turn insignificant in most cases): I conclude that there seems to be no general negative effect of 'support stage' (or comprehensive school) attendance on educational outcomes of ninth graders when estimated at the mean. However, sub-group analyses reveal that later tracking exerts positive effects on pupils with a less favourable family background. The sub-group results are complemented by quantile regressions demonstrating that the 'support stage' effects decrease nearly monotonically over the conditional performance distributions. Thus,

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

pupils at the lower quantiles benefit from later tracking in the sense that their PISA-E mathematics, reading and science score increase by one-fifth of a standard deviation. Unfortunately, the results also suggest that education inequality decreases to the detriment of 'top performers'.

Recently, policy-makers in Germany discuss the modification of the tracking system. Whether another system is considered to be beneficial depends from the objectives behind such a reform. If the major objective is to improve the educational situation of 'disadvantaged' pupils and to reduce education inequality, evidence from this paper suggests that delaying the timing of tracking is favourable. However, one needs to bear in mind that such a reform might negatively impact the 'top performers'. Thus, it cannot be unambiguously concluded from this study that a change of the schooling system towards later tracking is the panacea for improving educational outcomes in Germany.

References

Adams, R. / *Wu,* M. (2002): PISA 2000. Technical Report, Paris.

Ammermüller, A. (2005): Educational Opportunities and the Role of Institutions, ZEW Discussion Paper No. 05 – 44.

Ariga, K. / *Brunello,* G. / *Iwahashi,* R. / *Rocco,* L. (2005): Why Is the Timing of School Tracking So Heterogeneous?, IZA Discussion Paper No. 1854.

Bauer, Ph. / *Riphahn,* R. T. (2006): Timing of School Tracking as Determinant of Intergenerational Transmission of Education, Economic Letters 91, 90 – 97.

Baumert, J. / *Trautwein,* U. / *Artelt,* C. (2003): Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens, in: Deutsches PISA-Konsortium (eds.): PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland, Opladen.

Betts, J. R. / *Shkolnik,* J. L. (2000): The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools, Economics of Education Review 19, 1 – 15.

Brunello, G. / *Giannini,* M. / *Ariga,* K. (2007): The Optimal Timing of School Tracking, in: P. E. Peterson / L. Wößmann (eds.): Schools and the Equal Opportunity Problem, Cambridge.

Deutsches PISA Konsortium (2003): Ergänzungen zum Skalenhandbuch, Berlin.

Dustmann, C. (2004): Parental Background, Secondary School Track Choice and Wages, Oxford Economic Papers 56, 209 – 230.

Entorf, H. / *Lauk,* M. (2006): Peer Effects, Social Multipliers and Migrants at School: An International Comparison, IZA Discussion Paper No. 2182.

Epple, D. / *Newton,* E. / *Romano,* R. (2002): Ability Tracking, School Competition, and the Distribution of Educational Benefits, Journal of Public Economics 83, 1 – 48.

Fernandez, R. (1998): Education and Borrowing Constraints: Tests vs. Prices, Starr Center for Applied Economics Economic Research Reports No. 98 – 17.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Figlio, D. N. / *Page,* M. E. (2002): School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?, Journal of Urban Economics 51, 497 – 514.

Hanushek, E. A. / *Wößmann,* L. (2006): Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries, The Economic Journal 116, 63 – 76.

Henze, G. / *Sandfuchs,* U. / *Zumhasch,* C. (1996): Fördern in der Orientierungsstufe: Eine empirische Untersuchung zur Förderung lernschwacher Schüler in Kleingruppen, Bad Heilbrunn.

Hessisches Kultusministerium (1995): Empfehlungen für die Arbeit in der Förderstufe, Wiesbaden.

Hopf, V. (1979): Orientierungsstufe als Förderstufe, Frankfurt am Main.

Jürgens, E. (1991): 20 Jahre Orientierungsstufe. Beiträge zu einer umstrittenen Schulform, Sankt Augustin.

Jürgens, E. (1989): Die Orientierungsstufe im Urteil von Eltern und Lehrern, Frankfurt am Main.

OECD (2004): Learning for Tomorrow's World – First Results from PISA 2003, Paris.

Pekkarinen, T. (2005): Gender Differences in Educational Attainment: Evidence on the Role of the Tracking Age from a Finnish Quasi-Experiment, IZA Discussion Paper No. 1897.

Pischke, J.-S. (2007): The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years, The Economic Journal 117, 1216 – 1242.

Puhani, P. A. / *Weber,* A. M. (2007a): Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Educational Effects of Age of School Entry in Germany, Empirical Economics 32, 359 – 386.

Puhani, P. A. / *Weber,* A. M. (2007b): Persistence of the School Entry Age Effect in a System of Flexible Tracking, IZA Discussion Paper No. 2965.

Schütz, G. / *Ursprung,* H. W. / *Wößmann,* L. (2005): Education Policy and Equality of Opportunity, IZA Discussion Paper No. 1906.

Zimmer, R. (2003): A New Twist in the Educational Tracking Debate, Economics of Education Review 22, 307 – 315.

DOI https://doi.org/10.3790/schm.128.3.351 | Generated on 2025-01-22 13:08:10
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/