Project Report for Data Literacy 2023/24
Grade Inflation in the German School System - Causes and Effects

Jonathan Schwab [* 1] **Lars Kasüschke** [* 2] **Niklas Munkes** [* 3] **Tom Freudenmann** [* 4]

Abstract

The reasons for the increase in Abitur grades have been a highly discussed topic in German society. Often, the reason for this is attributed to grade inflation. In this paper, we will show why this is not the case. Student competence has increased. One of the most important factors is the student-teacher ratio. We will show that this metric has a high correlation with the Abitur grades.

1. Introduction

The Abitur grades have constantly increased in the German school system over the past decades. Every year, when the Abitur takes place, the grades and the difficulty of the exercises are extensively discussed in the media and have been part of a fierce research discussion for decades. The central focus of the discourse revolves around the question of whether grade inflation occurs, signifying a rise in grades without a corresponding increase in competence or knowledge.

The discourse has predominantly centred around mathematics, since the difficulty of exercises is easiest to compare. The line is drawn between mathematicians arguing that specific exercises are easier than exercises in the past (Kühnel, 2015) (Jahnke et al., 2014) (Lemmermeyer et al., 2019) and studies claiming that grade inflation cannot be reliably proven since the competence of students has also increased (Schleithoff, 2015). In 2015, a data-driven approach was employed, involving the analysis of comprehensive data on the education system. The analysis was promising, but not yet enough to neglect the claim of a grade inflation (Grözinger & Baillet, 2015).

This paper expands on that work, disproving the claim that grade inflation is the main cause of the observed trend. An explanatory framework for the improvement of Abitur grades is provided by this data analysis, building upon past research. All analyzed data is taken from official federal resources such as the German Federal Statistics Office.[1]

The study undertakes an analysis of the quantifiable impacts stemming from the upward trend in Abitur grades within the educational system. Prognostications are offered concerning the trajectory of future grade developments and the implications for the German education system.

2. Methods

2.1. Mathematical Basics

The Pearson correlation coefficient (Rodgers & Nicewander, 1988), denoted as r, is a statistical measure used to assess the linear relationship between two sets of data, X and Y. It is computed as the ratio of the sample covariance of the X and Y to the product of their sample standard deviations:

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \cdot \sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \quad (1)$$

2.2. Datasets

This paper investigates both causes and effects of the phenomena discussed in section 1. Thus the analyzed datasets may be grouped by the information they represent into *cause* and *effect* datasets.

In the following, *causes* shall be defined as the social, demographic and/or political factors that may influence the german school system such as the number of students, teachers, or budget provided by the German government. In contrast, *effect* refers to the observable impact of these causes on any measure modelling the students' performance, e.g. average grades, PISA study results, or the rate of repeaters

---

[*]Equal contribution [1]Matrikelnummer 6765897, jonathan.schwab@student.uni-tuebingen.de, MSc Computer Science [2]Matrikelnummer 4247775, lars.kasueschke@student.uni-tuebingen.de, BSc Computer Science [3]Matrikelnummer 4269436, niklas.munkes@student.uni-tuebingen.de, MSc Media Informatics [4]Matrikelnummer 6631101, tom.freudenmann@student.uni-tuebingen.de, MSc Computer Science.
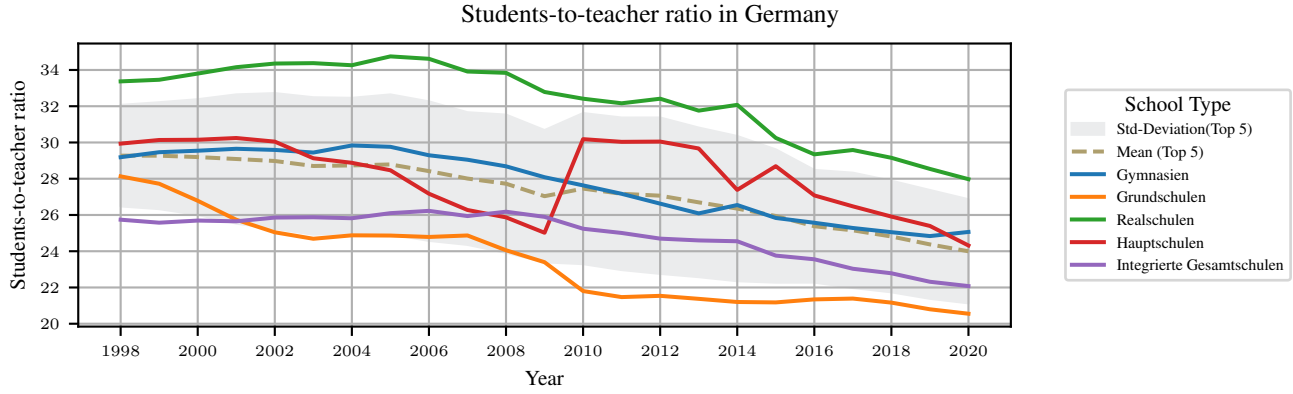
Figure 1. Students-to-teacher ratio of the five most common school types in Germany. The ratio of full- and part-time teachers is displayed for each school type and aggregated to their mean (━ ━) and standard deviation (▊).

and school-leavers.

The first cause dataset is the *Fachreport Schuljahr 2020/21* of the Statistische Bundesamt presenting the number of teachers from 1992 until 2020. The dataset groups them by their contract type, federal state, and school type. For the analysis, this paper merges the teacher counts with two student datasets, which are published in the *Genesis* database provided by the Statistische Bundesamt. These contain the number of children as different groupings and aggregations. The first (Table $21111 - 0002$) contains the number of children per grade and school type for the years 1998 to 2022. In contrast, the second table (Table $21111 - 0010$) provides the absolute amount of children, leavers, and beginners in each federal state from 1997 to 2022. Therefore, the analysis of the merged dataset can only be conducted separately for school types and federal states.

Additionally, this paper considers the budget per child (Table $21711 - 0011$) as a possible cause which is provided in the *Genesis* database of the Statistische Bundesamt. The dataset contains the budget per child for the years 2010 to 2022 and is grouped by federal states. In contrast to the demographic and societal causes above, the budget can be regarded as a primarily political factor. To adjust for inflation, the budget is multiplied with the *Verbraucherpreisindex* relative to 2022 provided by the Statistische Bundesamt.

Moreover, the effects on students' performance are the basis for the analysis of the German school system, since they indicate whether the grade inflation exists or not. One of the few publicly available datasets containing grades is the average Abitur grades per federal state. The grades are published every year in a separate report by the Kultusminister Konferenz. Each file contains the count of children per written grade and federal state. In addition, the grades are given in 0.1 steps, with 4.0 as the worst and 1.0 as the best grade. The amount of children who failed with a grade worse than 4.0 is aggregated in an additional column.

Although this is a great model for the performance of children attending grammar schools, a general performance measure for all school types is required to translate the results to all school types. Accordingly, this paper uses the number of repeaters (Table $21111 - 0014$) derived from the *Genesis* Database of the Statistische Bundesamt. There, the absolute count of repeaters by federal state, school type, and year is provided for the years 1998 to 2022.

2.3. Exploratory Data Analysis

Having introduced all used datasets, this paragraph aims to investigate potential patterns through an exploratory data analysis of the potential causes and effects.

Firstly, regard the demographic effects on the number of children attending school and teachers employed by school type and federal state. The exploratory data analysis has shown that the number of schoolchildren decreases steadily from 1998 to 2014. Instead, it increased from 2019 to 2022 because more children started their education and fewer left school. Furthermore, more children graduate from grammar schools with university entrance qualifications. This demographic effect is combined with an increasing number of teachers across all German school types and federal states. Although, the percentage of part-time teachers is increasing, the number of full-time teachers is decreasing until 2020.

Given the hypothesis that having more teachers per student increases the quality of teaching, the datasets can be merged. As already explained, this merge can only been done separately for school types and federal states. Furthermore, the student-to-teacher ratio is calculated over all full- and part-time teachers, since they represent the majority ($\sim 90\%$) of the distribution. In contrast, the teachers who are employed on an hourly basis are excluded due to their small impact on the teaching quality and sparse representation in the data. The results (Figure 1) show that from 1998 to

2020, the ratio decreased for the five most common school types. As a result, the average decreases from 29 to 24 children per teacher. Together with the hypothesis, it follows that the quality of teaching should increase, and thus the performance measures should increase.

Besides the demographic measures, the budgets for the school system differ between the federal states. The analysis of the adjusted budget to inflation per child has shown that it steadily increases for all federal states. Although, this may be caused by the increasing number of teachers and the goals of digitalization of schools in the last years (Cone et al., 2022).

Now that the basic effects that may influence the students' performance have been identified, it is possible to study the performance measures. As the analysis of the students datasets has shown, more children are attending grammar schools in Germany. Thus, the average Abitur grade of the children is a great measure of the performance of many children. Figure 2 shows that the average grades are increasing in all federal states. Furthermore, a linear regression can be employed to represent their mean. Importantly, the regression is calculated on the data before 2021 because of the COVID-19 pandemic beginning in 2020. In 2022, the grades significantly increased compared to the years before the pandemic. This could indicate that the pandemic has had novel consequences for the educational system. Due to the lack of data following the pandemic, this paper will solely focus on the linear trend until 2020. Furthermore, an additional analysis of the relative number of failed students has shown that the failure rate has no repetitive or linear pattern. Therefore, the provided results in Figure 2 are only valid for children graduating with a grade of at least 4.0.
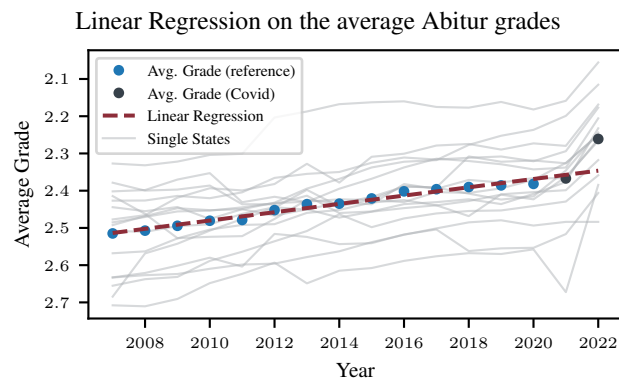

Figure 2. Average Abitur grades before (●) and after the COVID-19 pandemic (●) with a linear regression line (▬▬) of the years 2007 to 2020. In the background the figure contains average grades foreach federal state (▬).

Moreover, all children attending other schools have no direct impact on the results of the Abitur grades. Therefore, the number of repeaters per federal state, school type, grade,

and school year is analyzed. To enhance the relevance of the results, the relative ratio of repeaters is calculated by dividing the absolute counts by the absolute number of schoolchildren. This results in an aggregation for the federal states per year and in one for the school types per year. As a result, the number of repeaters has decreased for all educational institutions and federal states from 1998 to 2020. Hence, the trend equals the expected result, after analyzing the Abitur grades.

To summarize, the exploratory findings indicate an increasing number of students and teachers, resulting in a decreasing ratio of students to teachers and a rise in the budget per child. The possible outcomes include a linear increase in Abitur grades in grammar schools and a shrinking proportion of repeaters in general.

Correlation and Relationships   The relationships can be explored by combining the datasets, plotting the interesting variables, and calculating the correlation coefficients between them. Since the student-to-teacher ratio aggregates the number of students and teachers, this section will correlate the other variables against it.

Therefore, the first correlation exists between the average Abitur grade in Germany and the student-to-teacher ratio. Initially, the average grades across all federal states are calculated and then compared to the student-to-teacher coefficient for German grammar schools. As shown in Figure 3 the relationship between both is nearly linear. In addition, the result contains neither clusters nor outliers. Hence, a smaller student-to-teacher ratio results in better grades. Additionally, the Pearson correlation value is 0.98, indicating a strong positive correlation. This emphasizes the initial hypothesis and increases the importance of a good care factor between teachers and children.
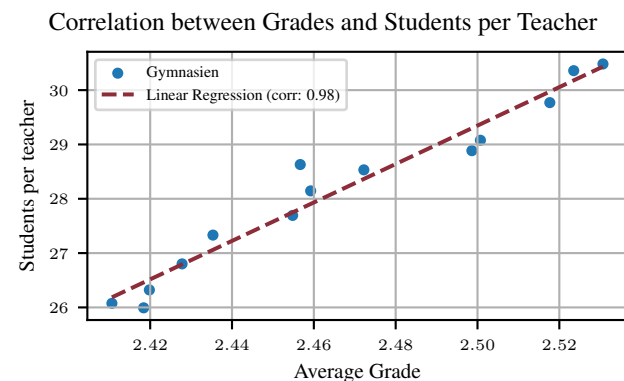

Figure 3. Linear regression on the students-to-teacher ratio by average Abitur grade. The resulting regression line (▬▬) is calculated over the aggregated average overall grammar schools (●) in Germany.

As already explained, these results correspond only to gram-

mar schools and are not necessarily representable for other schools or single federal states. Thus, the student-to-teacher ratio is also compared to the repeaters and the budgets per child in each federal state. Therefore, their Pearson correlation coefficients are calculated for each federal state across all years.

Figure 4 presents a visualization of the Pearson correlation coefficients, analyzing the relationship between the number of children per teacher and the average number of repeaters, as well as the educational budget per child. To visually represent the data across various federal states, a heatmap is generated. Therefore, the Pearson correlation coefficients for each state is normalized to the used color map scale. Consequently, each state is assigned a color representing the correlation coefficient. It is evident that there exists a strong positive correlation between the student-to-teacher ratio and the rate of repeaters in the most new federal states, while the correlation in the new federal states is significantly weaker. Moreover the correlation between the student-to-teacher ratio and the inflation adjusted budget per child tends to be positive for the new federal states and negative for the old federal states.
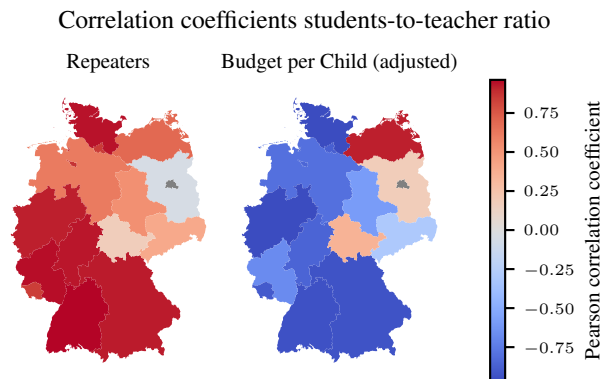
Correlation coefficients students-to-teacher ratio

Figure 4. Pearson correlation coefficients between the student-to-teacher ratio and the relative repeater count (left) and the inflation-adjusted average budget per child (right). Red indicates positive, gray neutral, and blue negative correlations between the variables.

3. Results

The most intriguing discovery in our analysis is the robust correlation observed between the student-teacher ratio and Abitur grades. While it may seem intuitive, the data provides concrete evidence of the strength of the correlation. The association between student performance and the presence of teaching personnel is a familiar topic in research, often examined in the realm of university performance (Dickson, 1984). However, its application to this specific problem is novel and the interpretation in this context is important. Since the student-teacher ratio got smaller over the current years and the grades went on a steep increase, This obser-

vation and the strong correlation underline the necessity of having enough teaching personnel available.

Suggesting that schools simply require additional funding to hire more teachers might seem like a straightforward solution. The data of Thüringen, Mecklenburg-Vorpommern, and Brandenburg shows, that this is not a solution. For them, the correlation between budget and students per teacher is negative. This seems odd, but there is an explanation. These states have had an increase in the number of students over the past decade (Bra)(?) (Mec). If certain schools get more children, they might want to employ more teachers. This does not happen, since they don't have enough teachers in these states available (Sekretariat der Ständigen Konferenz, 2023). This leads to schools getting more money, but the number of teachers staying the same or even decreasing because people retire. This leads to a negative correlation.

The same anomaly can be observed with the repeaters. Here, we think a different phenomenon is accountable for this. Schools in Thüringen and Brandenburg rely more and more on Teilzeitkräfte. This means that the overall proportion of teachers increases, but since they only account for about half the teaching time, the number of children failing a class still increases.

For every other of the 16 federal states, there is a very strong positive correlation, not only between Abitur grades but also the number of repeaters. This can be interpreted as meaning that the availability of teachers not only increases grades but is especially beneficial for the challenged students. But money doesn't necessarily help here. There needs to be enough teachers available to employ. From this analysis, we conclude that making sure that many teachers are available is one of the most important challenges for the education system. The prognosis of the Kultusministerkonferenz (Sekretariat der Ständigen Konferenz, 2023) shows, there are still more open positions than teachers that can fill them. Fortunatly, they predict that this gap will gradullay close in the coming decade. This means that we can expect a further increase in grades in the future.

It is important to note that having enough teachers is not the only factor at play here. We have shown that it is one of the most important ones. While the German education system faces several challenges, our demonstration illustrates that it has effectively addressed certain issues over the past decade and is poised to continue resolving them in the next. The increasing grades are a result of an increase in the competence of the students, facilitated by an improvement in the education system, especially a decrease in the student-teacher ratio.

4. Conclusion

We have introduced a new approach to explaining the increasing Abitur grades. There is a very strong correlation between the student-teacher ratio and the Abitur grades. Additionally, we also found a negative correlation between this ratio and the repeater number. This means that the number of teachers not only has a positive impact on the grades but also on the more challenged students. Improving the budget does not necessarily help. It is essential for the German education system to have enough teachers available.

Grade inflation in the Abitur grades has not been scientifically proven so far. What has been proven is that student competence is increasing. It is important to acknowledge that the student-teacher ratio is not the only factor at play in improving education and, thus, student competence. Multiple factors are at play; some are already known through research, and some still need to be discovered. We have shown that the student-teacher ratio is a crucial one.

References

Zahlen - schülerinnen und schüler. https://mbjs.brandenburg.de/sixcms/detail.php/bb1.c.763755.de. Accessed: 2023-01-25.

Statistik. https://www.regierung-mv.de/Landesregierung/bm/Ministerium/Statistik/. Accessed: 2023-01-25.

Cone, L., Brøgger, K., Berghmans, M., Decuypere, M., Förschler, A., Grimaldi, E., Hartong, S., Hillman, T., Ideland, M., Landri, P., Van De Oudeweetering, K., Player-Koro, C., Bergviken Rensfeldt, A., Rönnberg, L., Taglietti, D., and Vanermen, L. Pandemic Acceleration: Covid-19 and the emergency digitalization of European education. *European Educational Research Journal*, 21(5):845–868, September 2022. ISSN 1474-9041, 1474-9041. doi: 10.1177/14749041211041793. URL http://journals.sagepub.com/doi/10.1177/14749041211041793.

Dickson, V. A. An economic model of faculty grading practices. *The Journal of Economic Education*, 15(3):197–203, 1984. doi: 10.1080/00220485.1984.10845072.

Grözinger, G. and Baillet, F. Gibt es auch beim abitur eine noteninflation? zur entwicklung der abiturnoten als hochschulzugangsberechtigung – eine darstellung und analyse aus soziologischer perspektive. *Bildung und Erziehung*, 68(4):473–494, 2015. doi: 10.7788/bue-2015-0407.

Jahnke, T., Klein, H. P., Kühnel, W., Sonar, T., and Spindler, M. Die hamburger abituraufgaben im fach mathematik. entwicklung von 2005 bis 2013. *Mitteilungen der Deutschen Mathematiker-Vereinigung*, 22(2):115–122, 2014. doi: doi:10.1515/dmvm-2014-0046. URL https://doi.org/10.1515/dmvm-2014-0046.

Kühnel, W. Modellierungskompetenz und problemlösekompetenz im hamburger zentralabitur zur mathematik. *Mathematische Semesterberichte*, 62:69–82, 2015.

Kultusminister Konferenz. Abiturnoten im Ländervergleich, 2024. URL https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/abiturnoten.html.

Lemmermeyer, F., Kühnel, W., Spindler, M., and Klein, H. P. Zentralabitur 2019: Weitere absenkung der mathematischen anforderungen zentralabitur 2019: The lowering of mathematical standards continous. *Journal für Didaktik der Naturwissenschaften und der Mathematik (F)*, 3:92–98, 2019.

Rodgers, J. L. and Nicewander, W. A. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, 42(1):59–66, 1988. ISSN 0003-1305. doi: 10.2307/2685263. URL https://www.jstor.org/stable/2685263. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].

Schleithoff, F. Noteninflation im deutschen schulsystem – macht das abitur hochschulreif? / grade inflation in the german school system. *ORDO*, 66(1):3–26, 2015. doi: doi:10.1515/ordo-2015-0103. URL https://doi.org/10.1515/ordo-2015-0103.

Sekretariat der Ständigen Konferenz, d. K. d. L. i. d. B. D. Statistische verÖffentlichungen der kultusministerkonferenz. 2023. URL https://www.bildungsserver.de/Lehrerbedarf-und-Lehrerbedarfsprognosen-in-den-Bu html#Modellrechnung_und_Beschluesse_der_Kultusministerkonferenz_KMK_.

Statistische Bundesamt. Allgemeinbildende Schulen - Fachserie 11 Reihe 1 - Schuljahr 2020/2021 (Letzte Ausgabe - berichtsweise eingestellt), March 2022. URL https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Publikationen/_publikationen-innen-schulen-allgemeinbildende.html.

Statistische Bundesamt. Statistisches Bundesamt Deutschland - GENESIS-Online, November 2023. URL https://www-genesis.destatis.de/genesis/online?language=de&sequenz=tabellen&selectionname=21111*#abreadcrumb.