

Serial number 0149

Title How comparable are grades, between disciplines, universities and over time spans? To what extent do they reflect the performance, or are there other systematic influences on grading, which prevent a fair comparison?

Submitter Ms. Marita McGrory
Author, Prof. Volker Müller-Benedict

How comparable are grades, between disciplines, universities and over time spans? To what extent do they reflect the performance, or are there other systematic influences on grading, which prevent a fair comparison?

With the continued growth of university graduates, the function of their respective grades becomes more important, also in differentiating the graduates in preparation for the competitive labour markets. However, research in the US has detected various forms of "grade inflation" (HU 2005) (where there is an increase of better grades without the respective increase in performance). Consequently, also in the US, some universities prescribe proportions of A's and B's. Better grades need not necessarily be bad news, so long as their base is an improved student performance. As grades are the result of a grading process, a benchmark exists, may it be an explicit or written one, or may it be subjective or overt. This construct facilitates in reasoning as to how the benchmark should be, and what the result of the grade distribution should be.

Germany's Higher Educational institutions' (HEI) admission requirements are increasing, and the basis for admission are grades. The autonomy of German universities, since the 1990s, is increasing as is the competition to attract students, and here too, the mean grades are used for purposes of evaluation and quality management. Up until 1996, comparing grades was difficult in Germany, because grades were only recorded on paper. However, with this technical development comparisons have facilitated the German Education Council in a limited publishing of German HEI grades. Remarkable differences in mean grades between the HEIs were detected, as was a trend identifying grade inflation. And exactly that represents the core of the problem, the identification of grade inflation.

This is where our research programme enters the scene, and this author initiated researching grading. The research was involved staying for one and a half years in the dusty archives of seven universities to get dense and long-term time series of grades, in order to analyse grade developments over the last four decades. The data thus consists of these archival data from 1960 up to 1997, after which we continued with data from an electronic database from the German statistical bureau. This work will present descriptive results about the long-term trends of grades comparing disciplines, universities and types of exams, and will then continue with the trends and the differences between grades.

Thus, grades are steadily improving (they decline, bearing in mind that German grade 1 is very good, 2 is good etc.), but at the same time also the standard deviation is narrowing steadily, and this means that not a part of the students have improved their performance, but that the total distribution of grades is concentrated on a small span pivoting around the best grades. In some fields of study, the great majority of the students' grades are between 1 (very good) and 2 (good). Differences of performance between exams only, can be represented by position after the decimal point of grade 1. However, you can not subcategorize the grades in minor steps, therefore all students appear as a group with an homogeneous performance. This contradicts the function of grades, which is to map the differences of performances. So, here, we state a clear case of grade inflation.

In summary, the descriptive analysis identifies:

1. Over four decades grade inflation is proceeding in most disciplines.

2. Remarkable stable differences of mean grades exist between the disciplines.
3. Considerable differences of mean grades between universities (same discipline), are stable over one or more decades.
4. Stable considerable differences between state examinations and university exams (diploma, magister).
5. Mean grades are fluctuating with periods of a broad bandwidth of around 20 years.

In order to explain this, the first result was that grades are fluctuating in cycles around the inflation trend. The cycles are correlated with cycles in the academic labour market and/or with cycles in the frequency of the student cohorts. The second result that emerged was that if with a downward cycle the grades move in the direction 'better', the move is greater than if the opposite were the case, if the upward cycle is a movement of *disimprovement*, then the move is smaller. A possible explanation could be that examiners and academic teachers are in general the same persons and therefore, firstly, the mean grade of exams can serve as an indicator of their own teaching competencies. Thus an improvement of mean grades clearly is more welcome than a *disimprovement*. Secondly, the students use the mean grades to select courses and professors, also because better grades are preferred. Thirdly, it may be that a worsening of grades may request more explanation than an improvement. In sum a different elasticity could be expected.

It is difficult to derive a recommendation from these mechanisms, the onus is on the examiners morals'. They should be equally prepared to change grades in both directions better or worse. A worsening of the general level should be considered as equally important to an improvement in grades, otherwise grade inflation will be unavoidable. There are various explanations for the cyclical fluctuations of grades, and of grade inflation. This work will present these explanations with the aid of time series analysis. Unfortunately, a lot of possible sources of influence are short-termed or no data are available for the long time span. Therefore the author had to use the frequency of exams as a proxy of various different influences. Thus, as an indicator, the frequency of exams does not point to specific means and this should be established to minimize the changes of grades by influences independent of the performance.

References:

Hu, Shouping. (2005). *Beyond grade inflation. Grading problems in higher education*. Hoboken, NJ: Wiley Periodicals (ASHE-ERIC higher education report, 0884-0040, v. 30, no.6).