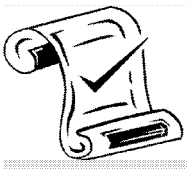


Network for Irish Educational Standards

Paper 9

New Metrics for Detecting Changes in Educational Standards

Brendan Guilfoyle

September 2011

NEW METRICS FOR DETECTING CHANGES IN EDUCATIONAL STANDARDS

BRENDAN GUILFOYLE

Abstract. The purpose of this paper is to develop a definition and quantification of educational standards. We do so initially from the theoretical point of view and therefore start by focusing on formal and deductive elements. Our key concept is that of the grade mapping: a map from any fixed taxonomy of abstract attributes constituting the standard, to a scale, namely the assessment grade.

We define declining standards to be a lowering in the taxonomy required to attain a given grade. We deduce that during times of declining standards not only does the mean grade increase but, for a hierarchical taxonomy, there is an accompanying decrease in skewness in the grade distribution. In a single-peaked grade distribution this can be seen as a migration of grades over the mean. In extreme cases, declining standards ultimately lead to a decrease in the standard deviation of the grades.

We apply these theoretical tools to the Irish educational system between the years 1992 to 2009. From the grade data at both second level (leaving certificate grades from 1992 to 2009) and third level (undergraduate University grades from 1998 to 2008) we demonstrate that they display the characteristic pattern of declining standards, namely, increasing mean and decreasing skewness.

Furthermore, we find that while the standard deviations of the leaving certificate grades have not shown an appreciable decrease, the standard deviations of the University grades have fallen over the period. The Irish Universities have undergone an intense inflationary period during which their ability to discriminate between students has been impaired.

Contents

Date: 28th September 2011.

1. Introduction

The purpose of this paper is to develop a fresh approach to the definition and quantification of educational standards in mass education environments. In order that the conclusions stand up to scrutiny and are as widely applicable as possible, we now put forward a theoretical structure of performance and assessment, seeking to make all underlying assumptions clear.

Our three fundamental assumptions are that:

(1) we are dealing with a mass education system,
(2) the results of assessment provides sufficient grade differentiation,
(3) assessment measures student performance against some hierarchical taxonomy of activity.

The first assumption reflects the fact that, while our initial analysis is theoretical and abstract, we test the conclusions drawn against empirical data collected from a mass education system and make robust statistical inferences. We are thus guided by metrics of output, namely grade distributions, rather than input. Care must therefore be taken in drawing direct conclusions of pedagogical practice.

The second assumption means that we consider systems in which the grade scale is sufficiently fine to allow for the analysis of second and third order effects. In practice this means that pass/fail systems are not suitable for our methods, while those with three or more grades are within the reach of our methods.

The final assumption means that the assessment seeks to measure student performance against some abstract framework of activity, perhaps articulated to the particular subject of the qualification. This is the *educational standard*. While the analysis may be extended to other activities, our interest will lie primarily in third level (tertiary) educational systems and the second level systems that feed them. Bloom's taxonomy of knowledge [2] and its variants [1] serve as the exemplar of such an abstract framework for cognitive activity at this level, and so we give some details of this specific model.

The key feature of this taxonomy is that it is *hierarchical*. By this we mean that the scale of cognitive activity has an ordering that is cumulative. Thus a lower activity must be mastered in order to advance to the next level of activity. Conversely, those who have mastered the higher levels find it easier to perform lower activities. It is important to note that our analysis will draw the same conclusions from any taxonomy with this property. Only a belief in a radical dissociativity of cognitive activity would lead one to reject this assumption and such a belief does not appear to be widespread among contemporary educationalists.

Consider then an educational system satisfying our three assumptions. That is, an educational system which produces a grade for a large number of students which measures their performance against a hierarchical taxonomy. This grade could be arrived at in any number of ways and could be, for example, an aggregation of a number of measurements. The resulting grade distribution reflects the attainment of the student cohort against the standard.

From an abstract point of view, assessment is then a mapping from an ordered set of abstract attributes to the grade distribution of the student population. Such a *grade mapping* is determined by numerous interlinked factors, including the nature of the material being assessed, the mode of assessment, the selection of students, as well as historical and institutional factors.

As educationalists, we aspire to make the grade mapping as consistent, fair and objective as possible. While comparison of mappings generated by different systems or institutions may be difficult because of the above-mentioned interlinked factors, distortions of the grade mapping of a given institution over time are a different matter. Indeed, consistency over time is one feature that we would like to observe in any grade mapping, as it engenders confidence that assessment is measuring the student against something fixed, i.e. the standard. On the other hand, large scale distortions over time require justification that student performance is really changing on the taxonomy and not just on the image of the grade mapping.

We define *declining standards* to be a change over time in a given educational system where the grade mapping gives higher grades to those at fixed levels in the taxonomy. Equivalently, decline means that a lower level in the taxonomy suffices to attain a fixed grade. In such a situation, the mean of the grade distribution would naturally increase and we see the first order effects of declining standards. Reversing the argument does not work directly, however, since grade increase on its own does not necessarily imply declining standards. It could result from a variety of factors, for example better teaching, higher selectivity of students etc.

Indeed, it is precisely this issue that is the main point of contention, if not controversy, in the debate about declining educational standards. When is grade increase a symptom of grade inflation (i.e. declining standards) and when is it a sign of higher student attainment? In the absence of other comparative metrics of performance, how can the former be distinguished from the latter?

In this paper we consider second and third order effects in the grade distribution and what we expect to see during times of declining standards in mass education systems. We argue that the hierarchical nature of the taxonomy implies that, during times of declining standards, in general those operating at a higher level in the taxonomy benefit more. That is, we expect to see a non-linear effect in which the grade distribution, aside from having an increasing mean, becomes negatively skewed.

Skewness is traditionally quantified by a single number γ which is zero if the distribution is symmetric, negative if the distribution has is peaking to the right and positive if the distribution has is peaking to the left. Thus by becoming more negatively skewed the grades are piling up in the higher awards, while the level of lower awards is thinning out.

Furthermore, an advanced decline in standards leads to a second order effect in the form of decreasing standard deviation. This is an artifact of the *ceiling effect* whereby the top grades cannot increase any further. Such a situation undermines the whole ethos of assessment as a measure of achievement in educational settings. These concepts are discussed in detail in the next section.

In section 3 we apply these theoretical tools to the Irish educational system between the years 1992 and 2009. What is fascinating about the Irish experience is that it has undergone radical changes relatively recently and a comprehensive set of empirical data is readily available to test theoretical constructs. Moreover, the population is big enough to be of interest, but small (and homogenous) enough to rule out large-scale internal dynamics that would lessen the validity of analysis.

From the grade data at both levels we find that the grade distributions display the characteristic pattern of declining standards: increasing grade mean and decreasing grade skewness. This is a feature of almost all subjects at second level and

across Honours Bachelor degrees in all seven Universities. Thus, this internal metric indicates a significant reduction of standards across the Irish education system during this period.

Perhaps more worryingly, University grades are found to have decreasing standard deviation - a hallmark of advanced decline in educational standards whereby assessment fails to distinguish between students. Thus, Irish Universities are slouching towards a pass/fail system with relative merit residing with institutional reputation rather than award level.

In the final section we discuss the implications of our analysis

2. Theoretical Analysis

2.1. **Assessment and Grade Distributions.** Consider the educational setting in which student learning is measured by assessment and a grade is awarded to indicate the level of attainment of the student. Such an assessment may be arrived at in any number of ways: summative written examination, continuous assessment, project work etc. It can also be an aggregation of different assessments, as in the final degree award at third level.

We assume that the grading system provides an ordered scale that is reflective of student achievement, usually given by letter or numeric grade. While such grading systems are by their nature discrete, the number of possible values of the grade will be left unspecified, although the analysis will assume a sufficient level of grade differentiation. Thus, for example, pass/fail systems would not be appropriate for the techniques we develop here.

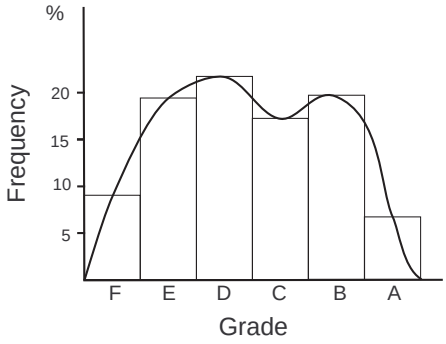


FIGURE 1

Grade distributions can be easily visualized as a frequency table over the grade scale, which is discrete. In Figure 1 we give an example of a commonly used grade scale for which there are six grades, F to A. Here and throughout we draw such distributions with lower grades on the left to higher grades on the right. The vertical axis indicates the frequency with which the grade was achieved as a percentage of the total student population. For illustrative purposes we draw a smooth or piecewise linear curve that interpolates the discrete data under consideration. All numerical calculations will of course be carried out with the discrete data.

Our primary interest is to investigate the grade distribution for a given educational system. We refrain from making any comments as to what such a distribution should look like, as different grading distributions may be appropriate for different educational settings. That said, the systems we consider mostly have a single-peaked grade distribution.

The first numerical quantifier of a distribution is the *mean* grade which gives us a single central value for the level of attainment. How clustered the grades are about the mean is measured by the second numerical quantifier, the *standard deviation*. Asymmetry about the mean is measured by the third quantifier, the *skewness* of the distribution. These quantifiers are mathematically defined in the appendix.

To visualize skewness, consider the three examples in Figure 2 below. As mentioned earlier, in each graph the grades awarded are on the horizontal axis with higher attainment to the right. These distributions are distinguished by their *skewness* - a key concept that will be discussed and quantified in subsequent sections.

Thus, in the left-most distribution the grades awarded are symmetric in the sense that the percentages of higher and lower grades is equal - we say the distribution is *symmetric*. In the middle distribution the frequency of higher grades awarded is more than that of lower grades and the distribution is said to be *negatively skewed*, while the opposite is the case in right-most distribution which is *positively skewed*.

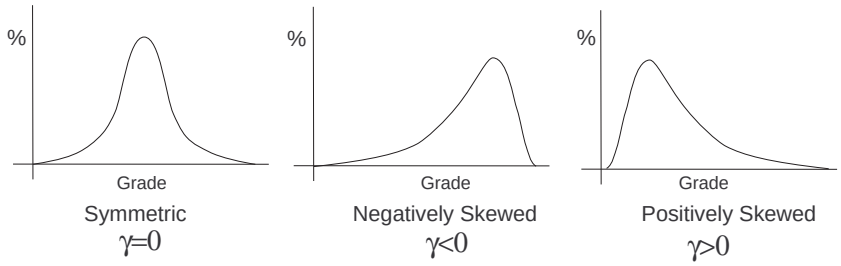


FIGURE 2

Specifying what grade distribution would be appropriate for a given educational system may be problematic. For our purposes we do not need to enter into this discussion, as we are concerned with changes in a given grade distribution (of any shape) over time. We then track the change of the grade mean, standard deviation and skewness and try to infer reasons for any large-scale movements.

In order to advance the analysis, we consider an abstract conceptual framework in which educational achievement can be benchmarked. Any hierarchical taxonomy would yield similar conclusions, but for illustrative purposes we consider the most widely cited framework.

2.2. **Bloom's Taxonomy of Knowledge.** A foundational classification of learning objectives in education is provided by the taxonomy of knowledge first presented in 1956 [2]. In the cognitive domain there are 6 levels summarized below.

While modifications and alternatives to this taxonomy have been put forward over the intervening decades [1], it remains the most widely used. From our point of view, the detailed categorization is not really important - rather it is the hierarchical nature of the levels.

The levels are hierarchical in the sense that they start from the simplest behavior to the most complex. The categories can be thought of as degrees of difficulties.

TABLE 1. Bloom's Taxonomy of Knowledge

Competence	**Skill**
Knowledge	Exhibit memory of previously-learned materials by recalling facts, terms, basic concepts and answers
Comprehension	Demonstrative understanding of facts and ideas by organizing, comparing, translating, interpreting, giving descriptions, and stating main ideas
Application	Using new knowledge. Solve problems in new situations by applying acquired knowledge, facts, techniques and rules in a different way
Analysis	Examine and break information into parts by identifying motives or causes. Make inferences and find evidence to support generalizations
Synthesis	Compile information together in a different way by combining elements in a new pattern or proposing alternative solutions
Evaluation	Present and defend opinions by making judgments about information, validity of ideas or quality of work based on a set of criteria

That is, the first one must be mastered before the next one can take place. Conversely, the lower levels of activity are easier for those who have already mastered the higher levels.

The taxonomy has played a significant role in educational theory and is one of the most cited works in the field. The philosophy, if not the detail, has been codified in many teaching methodologies, such as learning outcomes, and in international agreements which stratify educational awards, such as the Bologna Accords.

There is, however, no consensus as to how exactly to map Bloom's taxonomy onto a particular award level, nor how assessments can be designed to measure specific levels in the taxonomy. While educational systems strive for the highest levels within the taxonomy, it is not clear whether every individual is even capable of attaining all levels or that this is possible in mass education systems.

Now we consider temporal changes in educational standards, as defined by movement up or down the knowledge taxonomy, and how they impact on grade distributions.

2.3. **Temporal Changes in Educational Standards.** We now come to the key question of this paper. Suppose we have a particular educational system satisfying our three assumptions and the corresponding grade distribution. What effect would a change of educational standards have on the grade distribution?

The explicit definition of such a standard for a given system is bedeviled with difficulties, but our emphasis is on how we expect the grade distribution to change as the standard changes. In particular, suppose that the standard declines, that is, the level of attainment required to attain a particular grade moves down the taxonomy. How will this be reflected in the grade distribution?

Before considering this, let us compare with other possible temporal changes in the grade distribution and their causes. For example, take an educational system which normalizes grades by fixing the average at say 55%. A decision to raise or lower this average would see a translation of the whole distribution to the right or left, respectively. This could be effected without any change in either the standard deviation or the skewness of the distribution.

Or consider an educational system in which a significant amount of resources is released to assist the weakest students i.e. remedial teaching. This would affect the grade distribution by drawing the left tail into the grade distribution without affecting the right tail. This would increase the grade mean, while increasing the skewness (from negative to symmetric, for example).

2.4. **Increasing Mean, Decreasing Skewness.** Let us return to the central question: suppose that the standard declines, that is, the level of attainment required to attain a particulargrade moves down the taxonomy. How will this be reflected in the grade distribution?

It is obvious that we expect the mean grade to increase. However, there is a refinement of this shift: given a mixed ability cohort of students, those who have already attained a higher level in the taxonomy would be advantaged more by the downward shift in standards. That is, those with higher abilities would be better positioned to take advantage of the shift down the hierarchy.

Conversely, weaker students would not be in a position to take advantage of the downward shift as much as the more capable students. Such non-linear phenomena are familiar in a variety of competition models - the fitter are better able to take advantage of changing opportunities. Thus, in the presence of declining standards, we expect to see a proportionally greater effect higher up the hierarchy and its image under the grade mapping. Pictorially, this would be seen by a tendency of the grade distribution to become negatively skewed (middle distribution in Figure 2).

We are thus led to conclude that:

Conclusion 1: *In an educational system undergoing declining standards, aside from increasing mean, we expect to see a decrease in the skewness of the grade distribution.*

Mathematically, the skewness is measured by a number, which is proportional to the third moment of the distribution (see equation (5.3) in the Appendix). In a single peaked grade distribution, skewness is a measure of the relative frequencies to the left and right of the peak.

Thus a change in the skewness can be seen as a *migration of grades over the mean*, from left to right for decreasing skewness. This is induced by a decrease

in the number of mid-lower grades and a corresponding increase in the number of mid-upper grades. For example, with five grades: Fail D C B A, and mean lying at or near C, we would expect to see a decrease in D grades and a corresponding increase in B grades. Fails and A's would be similarly effected, but since they have lower frequencies the effect could be washed out by noise. Thus it is in the bulk grades, namely D's, C's and B's that one would expect to see the effect of decreasing skewness most clearly.

To illustrate these theoretical constructs consider the grade distributions for Leaving Certificate Examination Irish (Higher) in 1992 and 2009. The result is Figure 3:

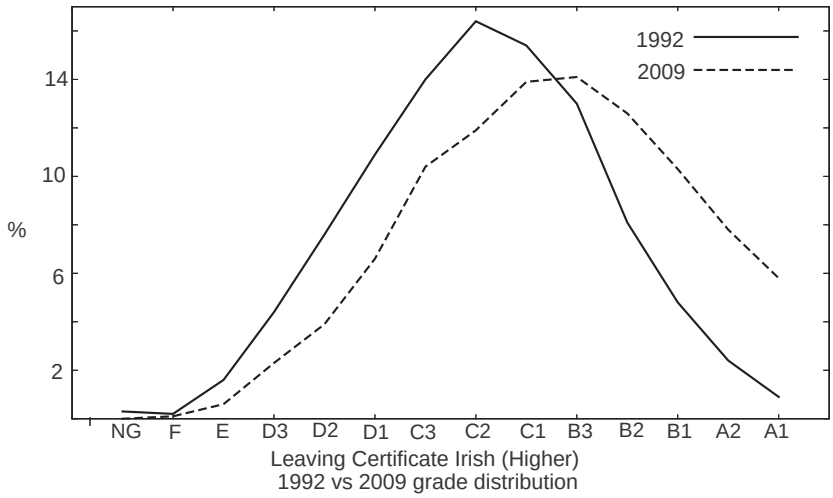


FIGURE 3

Here we clearly have a distortion to the right with increasing mean and decreasing skewness (remember skewness can be negative and get more so by decreasing). Note the biggest changes are the depletion in the mid-lower grades and a corresponding increase in the mid-upper translating up through the highest. This is migration of the grades over the mean and strong evidence of declining standards. We return to this example in the next section and put it in the context of other subjects.

2.5. **Decreasing Standard Deviation.** Of course, given the fact that students cannot go beyond the top of the award classification, there is less room for improvement in the very top students - what is referred to as the *ceiling effect*. As standards decline, the mean increases and skewness decreases, and there remains less and less of the grade scale to effectively differentiate between students.

Thus were the decline to be very severe or to take place for an extended period of time, we would also see a loss of spread in the grade distribution as the grades become more clumped together. Since the spread of the grades is measured by the standard deviation, we are led to the following:

Conclusion 2: *In an educational system undergoing severe or protracted declining standards, we expect to see a decrease in the standard deviation of the grade distribution.*

In summary, the evolution of grade distributions in time can be used as indicators of the kinds of changes taking place in the education system. The above two conclusions show us the finger-print of declining standards: increasing mean, decreasing skewness and, in extreme cases, decreasing standard deviation. In the following sections we quantify these concepts and study the data for the Irish educational system.

3. The Irish Educational System

3.1. **Leaving Certificate Examination 1992 to 2009.** The Leaving Certificate Examination (LCE) is the annual national examination administered to students in the Republic of Ireland upon completion of second level education. It covers a 2 year central syllabus in over 40 subjects and is currently taken by around 50,000 students.

The examinations are prepared, administered and graded by the State Examinations Commission which also publishes the statistics annually. Of the subjects involved, 3 are mandatory (Irish, English and Mathematics), while students usually choose 4 other subjects. Most subjects can be taken at one of two levels: Higher or Ordinary Level. Letter grades are awarded from No Grade to A1, in accordance with the percentage ranges given in Table 2. The table also gives the representative mark at the middle of each interval.

Table 2. Irish Leaving Certificate Grades

Grade	**Refinement**	**Percentage**	**Representative**
A	A1	90 – 100	95
	A2	85 – 90	87.5
B	B1	80 – 85	82.5
	B2	75 – 80	77.5
	B3	70 – 75	72.5
C	C1	65 – 70	67.5
	C2	60 – 65	62.5
	C3	55 – 60	57.5
D	D1	50 – 55	52.5
	D2	45 – 50	47.5
	D3	40 – 45	42.5
Fail	E	25 – 40	32.5
	F	10 – 25	17.5
	NG	0 – 10	5

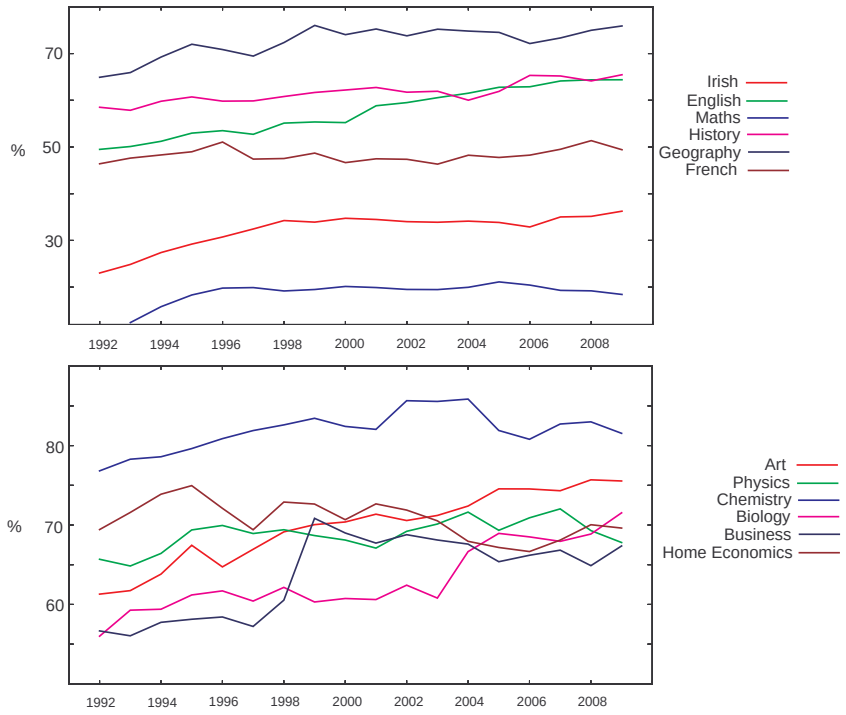
In [21] the grades for the LCE were studied over the period 1992 to 2006. The 24 most popular higher level subjects showed an increased rate of combined A and B grades in 2006 over 1992, with an average increase of 54.7%. Eighteen of the twenty ordinary level subjects showed an increase in A and B grades with an average increase of 101.2%. It was argued that the increase has been caused by declining standards. This was concluded on the basis of comparison with the PISA studies [5] and consideration of social and economic factors.

In this section we study the results obtained by second level students in the LCE in the period 1992 to 2009, as supplied by the State Examinations Commission

[26]. We consider the top 12 most popular subjects, both at Higher and Ordinary Levels, and compare the percentage of students obtaining various grades in each examination, with particular attention to their change over time.

One possible cause of mean grade change is change in the proportion of students taking higher and ordinary levels. Were a significant proportion of students to drop from higher to ordinary level, then we would expect an increase in the grades at both levels. At ordinary level because relatively stronger students would now be taking the easier exam and, at higher level because only the strongest students would be left taking that exam.

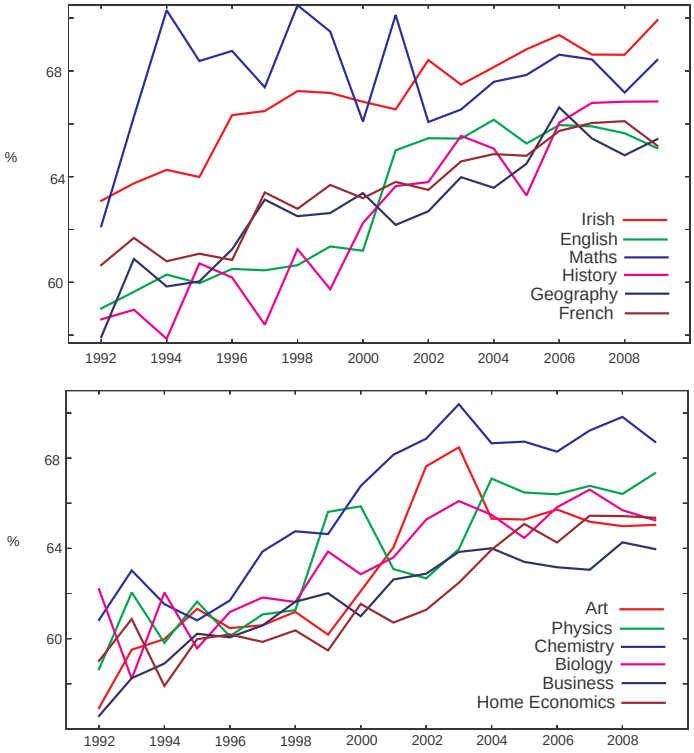
The data however shows the exact opposite trend: in most subjects the proportion taking the higher level has *increased* over the period. This has been noted previously in a study of increases in CAO points (for entrance to third level) [13].



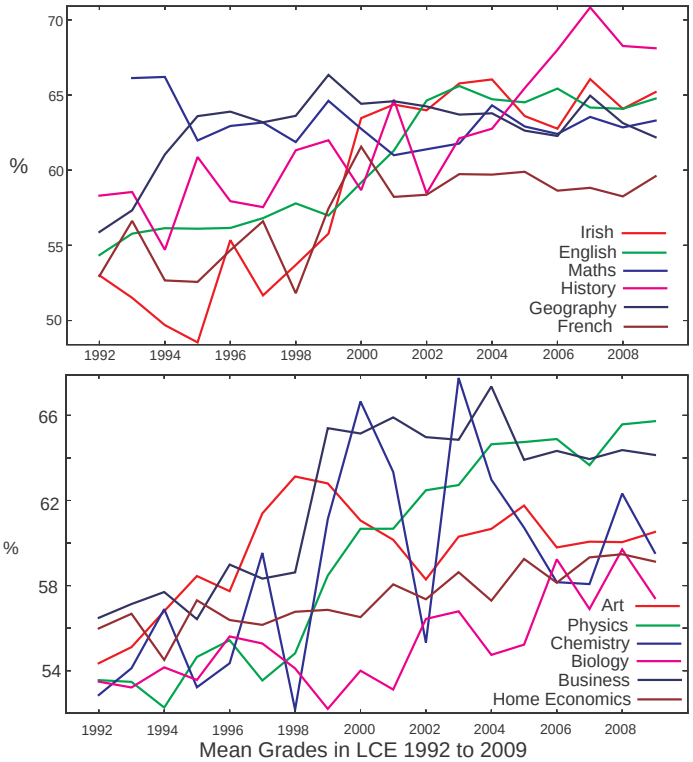
Percentage of students taking higher level in LCE 1992 to 2009

FIGURE 4

Given the fact that the numbers at higher level have increased and at lower level have decreased, we expect that, were standards maintained, the grades at both level should decrease. When we look at the data, however, we find the opposite: the mean grade has been increasing at both levels, as illustrated in Figure 5.



Mean Grades in LCE 1992 to 2009
Higher Level



Mean Grades in LCE 1992 to 2009
Ordinary Level

FIGURE 5

Clearly there has been a significant and consistent increase in the mean grades. This is entirely in keeping with the findings of a dramatic increase of higher awards found in [21]. To investigate these grade increases further, over the following four pages we consider the trends for skewness, D vs B grades and standard deviation for the grade data.

First consider the skewness of the grades, as graphed in Figure 6. From the data there has been a general decrease in skewness.

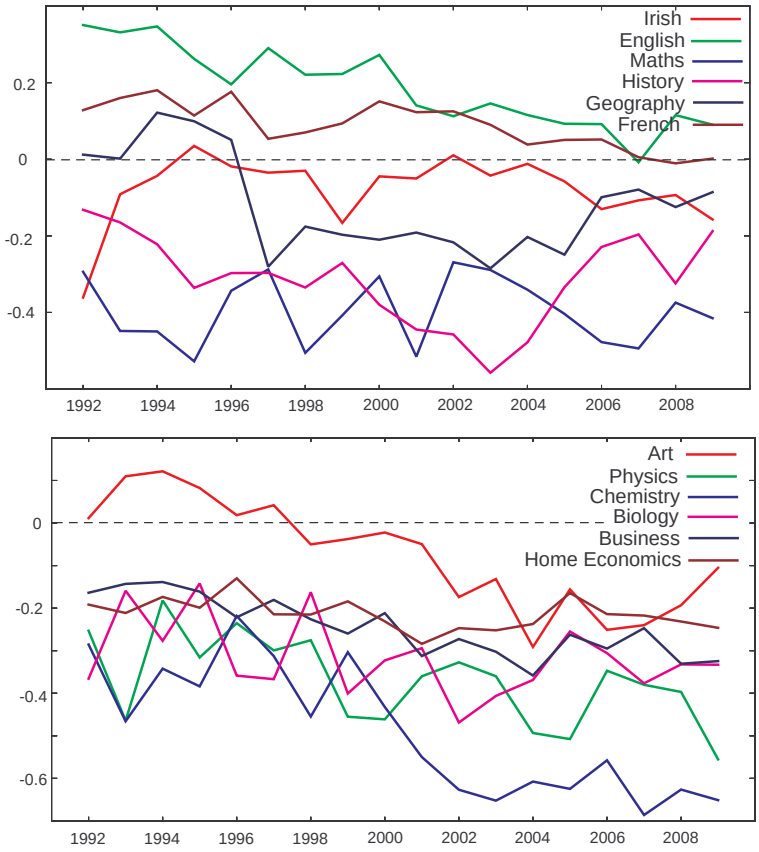
This can also be seen in the following manner. In Figures 7 and 8 we compare the proportion of those attaining a B and those attaining a D in all 12 subjects at both levels. Here, of course, we have amalgamated B1, B2 and B3 to a B and D1, D2 and D3 to a D. The fingerprint of declining standards once again appears: a decrease in the D's and a corresponding increase in B's. This is a graphical representation of migration of grades across the mean.

Note the almost exact mirroring of the D and B grades. This is an artifact of the fact that these are the most populous grades: one can only increase dramatically if the other declines. This lends further weight to our conclusion about skewness in single peaked distributions being most visible by migration close to the mean.

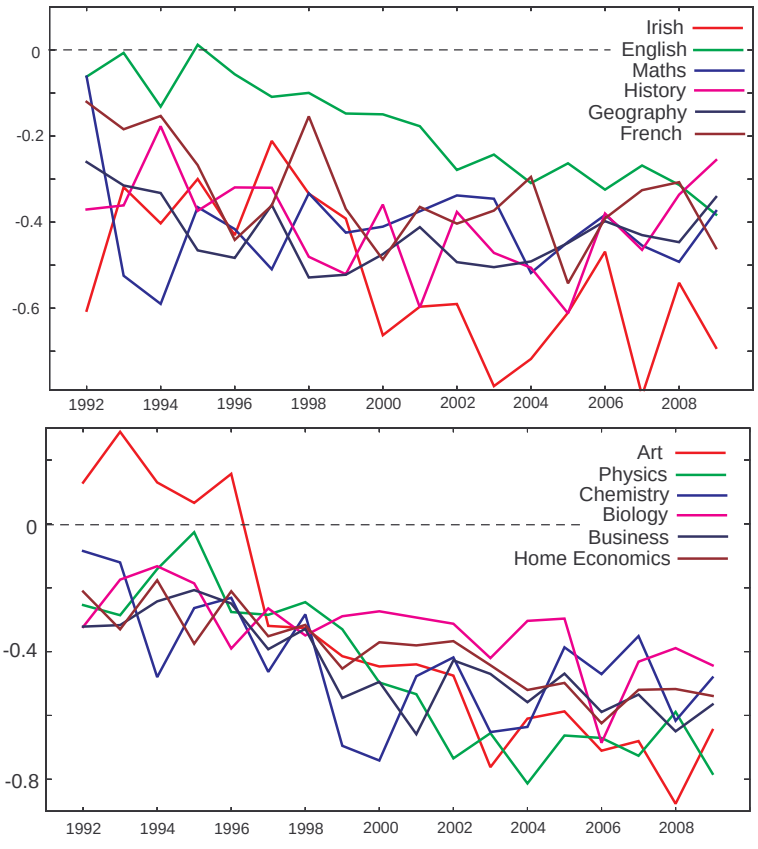
We finally turn to standard deviation. Here we find no immediately identifiable pattern and the results are shown in Figure 9.

In conclusion, there is ample evidence that academic standards in the Leaving Certificate Examination have declined over the period 1992 to 2009. This has manifested itself both by an increase in the mean grade and a decrease in the skewness of the grades awarded. The decline has occurred to a greater or lesser extent across almost all subjects.

This drift has not, however, impacted on the examinations ability to discriminate between levels of student attainment insofar as the standard deviation has not been appreciably decreasing. No doubt the number of grades available (14) has assisted in maintaining this spread, however in such situations it may be worthwhile to look at higher moments (e.g. kurtosis) to fill out the true picture of the grade dynamic.

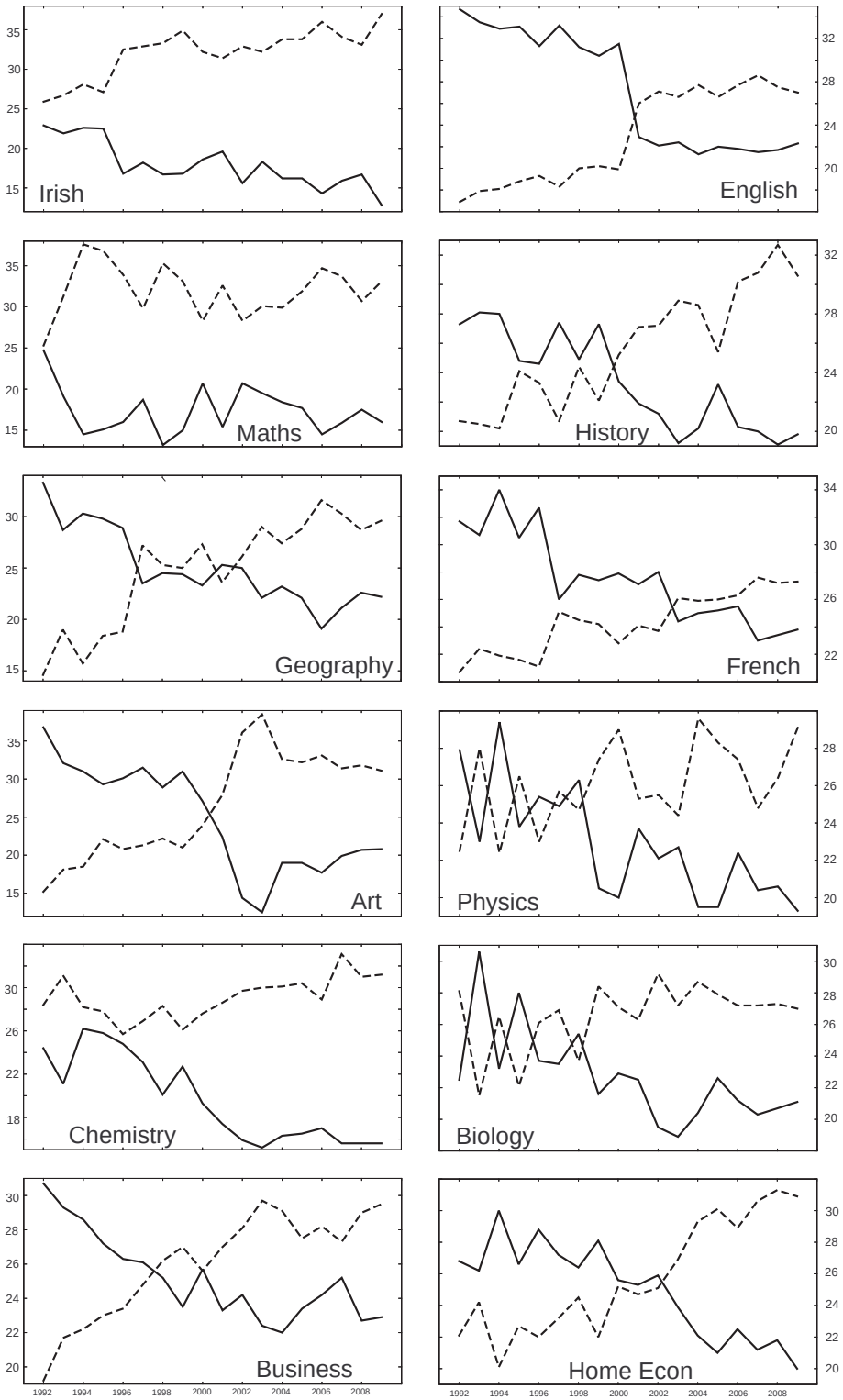


Grade skewness in LCE 1992 to 2009
Higher Level



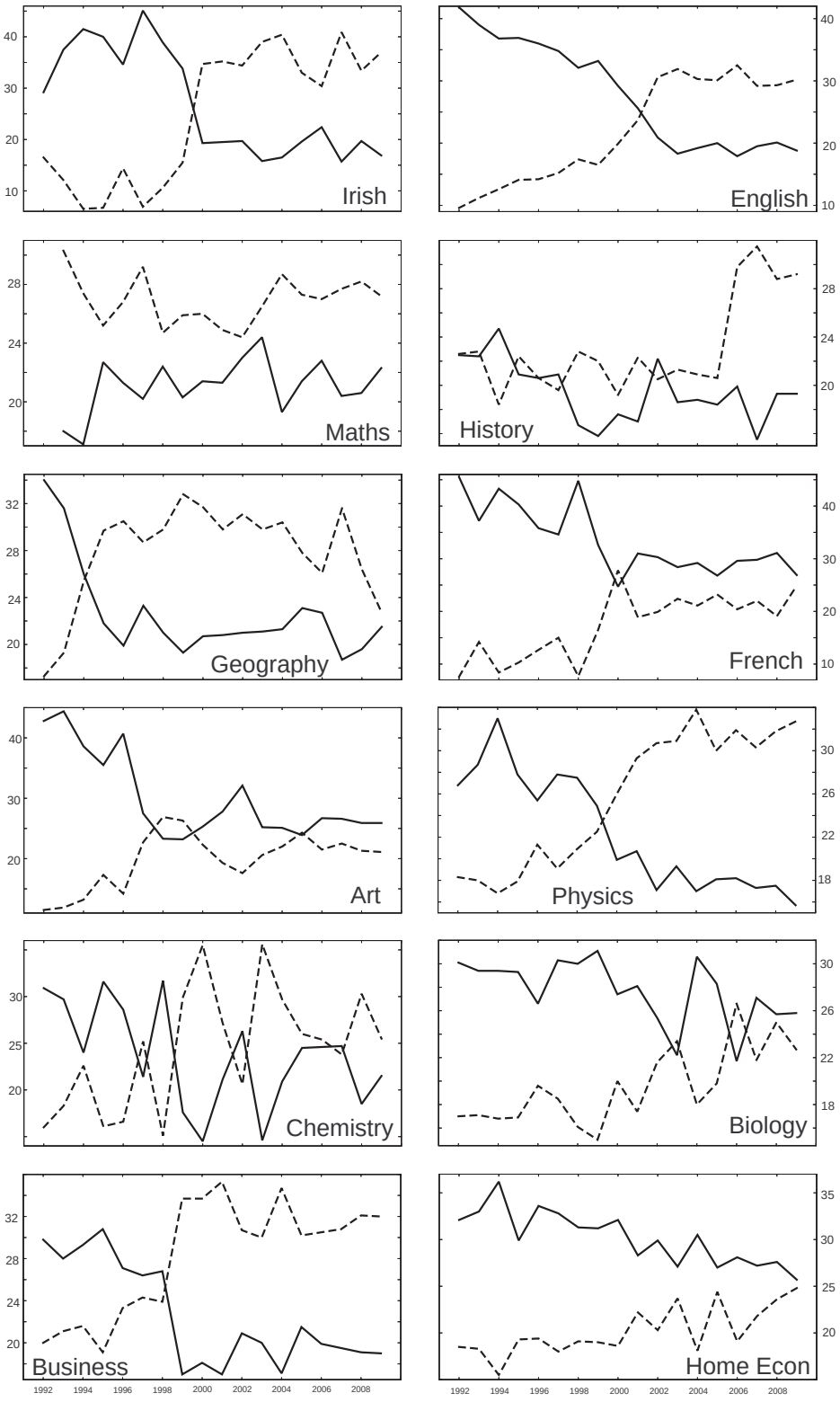
Grade skewness in LCE 1992 to 2009
Ordinary Level

FIGURE 6

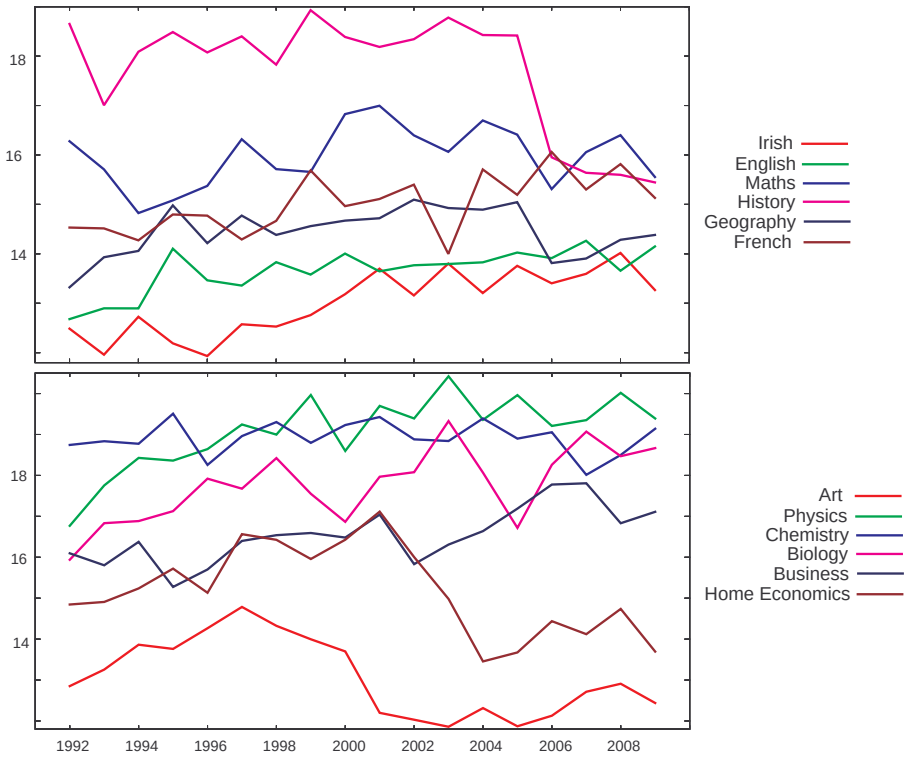


LCE grades: D (solid) and B (dotted) 1992 to 2009
Higher Level by subject

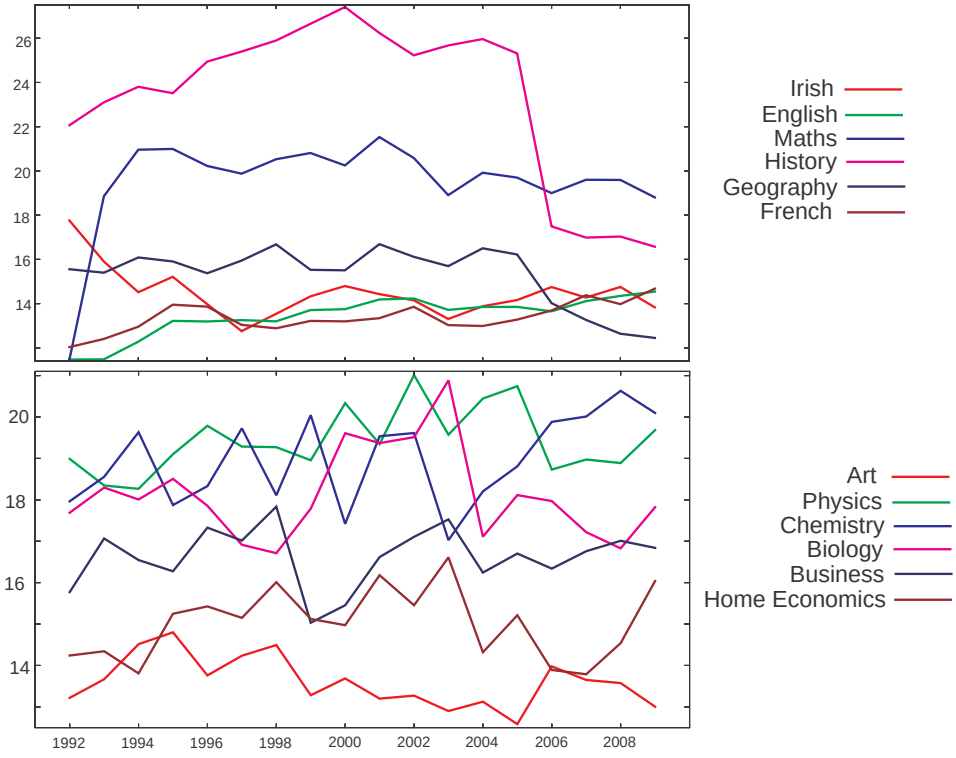
FIGURE 7



LCE grades: D (solid) and B (dotted) 1992 to 2009
Ordinary Level by subject
FIGURE 8



Grade standard deviation in LCE 1992 to 2009
Higher Level



Grade standard deviation in LCE 1992 to 2009
Ordinary Level

FIGURE 9

3.2. **University Undergraduates 1998 to 2008.** We now consider the grade distributions of the Irish University system. In earlier work, a very significant increase in higher grade award levels was identified in Irish Universities [17] and Institutes of Technology (IoTs) [16] [22] over the period 1994 to 2004 for the Universities and 1994 to 2008 for IoT's.

In particular, with regard to the University undergraduates, it was found that there had been a 76% increase in first class and a 21.4% increase in upper second awards over the period. It was argued that ample evidence exists that the average level of ability and motivation among University students declined during the period under consideration and that the inevitable conclusion was that the increase could only be attributed to declining standards.

In this section we focus on the more recent decade 1998 to 2008 and consider not just higher awards, but the whole available grade distribution data for the Universities. The seven institutions under consideration are:

- University College Dublin (UCD)
- University College Cork (UCC)
- National University of Ireland Galway (NUIG)
- Trinity College Dublin (TCD)
- National University of Ireland Maynooth (NUIM)
- Dublin City University (DCU)
- University of Limerick (UL)

The data we use is the grades for Honours Bachelor Degrees obtained by full-time students, as provided by the Irish Higher Education Authority [9]. These are classified into four grades: pass, second class honours (grade 2), second class honours (grade 1) and first class honours. We abbreviate these to pass, 2.2, 2.1 and 1, respectively. Some courses also award undifferentiated second class honours or third class honours, but the numbers involved are small and so have been removed for our analysis. One large course, namely UCD Arts, has a pass/fail grade and has also been removed.

The numbers who fail are not published by the HEA and it is not clear that these are even provided by the Universities. While it would be desirable to include this data to complete the picture, it is unlikely, given its political nature, that this dark matter of the Irish University system will be publicly available in the future. Our frequency rates are therefore calculated as a percentage of total non-failing students. Another omission is the grade data for 2003, as this data is not available through the HEA.

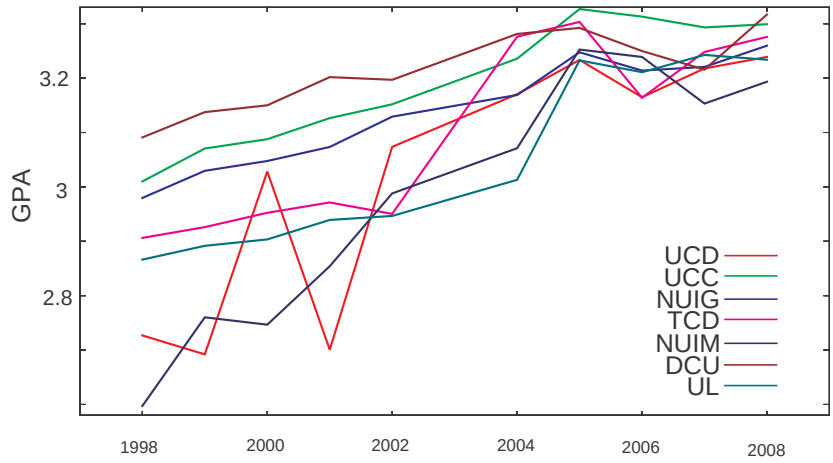
As in the case of second level, for statistical purposes the grade was converted into a numerical score, in this case the Grade Point Average. This was done as per Table 3.

TABLE 3. Irish University Grades

Grade	**GPA Range**	**Representative**
Pass	2 – 2.47	2.235
2.2	2.48 – 3.07	2.775
2.1	3.05 – 3.67	3.375
1	3.68 – 4.2	3.94

Note that the specific numerical value we assign to a particular grade is relatively unimportant - we are interested in how the frequency of these grades change over time for a given institution. A word of caution is in order about comparing different institutions, as their different grade frequencies may be caused by any number of factors: quality of students, courses on offer etc.

Using these values, the mean, standard deviation and skewness can be computed using equations (5.1), (5.2) and (5.3). For the mean the results are graphed in Figure 10.



Universities' Grade Mean 1998 to 2008

FIGURE 10

Over the time period, the mean GPA has been inexorably increasing in every institution, in agreement with the increase in higher grades documented in [17]. It is noteworthy that the increase seems to have slackened somewhat in recent years, a fact also noted in a recent report to the University council of Trinity College Dublin [3]. This would appear to be the ceiling affect: after a period of high inflation, the grades become so cramped at the top of the scale that the increase slows down.

Turning to the skewness, we find the trend over time illustrated by Figure 11.

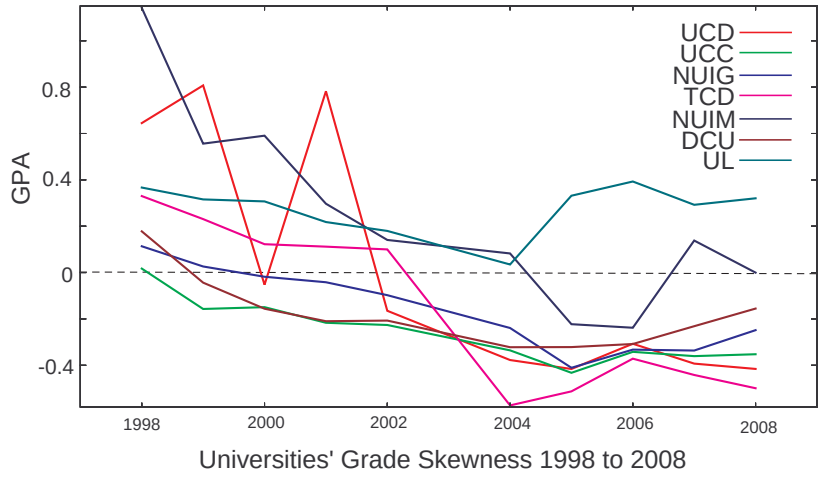


FIGURE 11

The skewness of the grades has been decreasing steadily over time, generally tipping from being positively skewed to being negatively skewed. In line with our theoretical discussion of section 2, this gives strong evidence that the increase in grades has been brought about by a reduction in standards.

Finally, let us turn to the standard deviation. The resulting changes in time are illustrated in Figure 12.

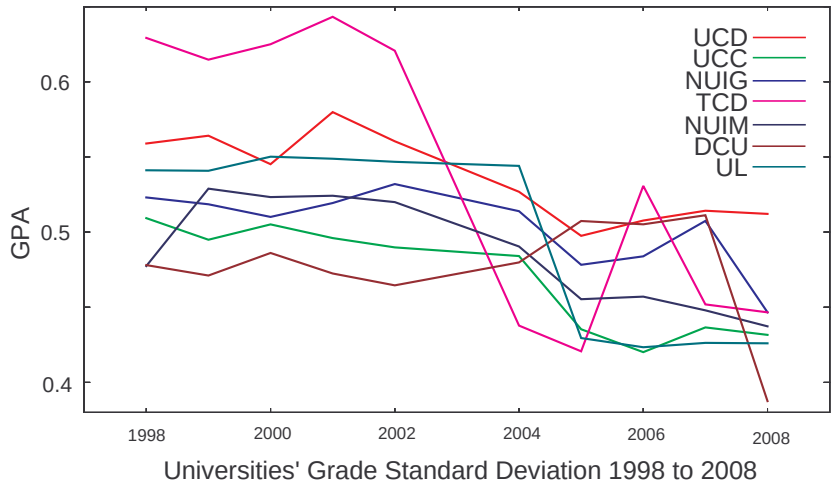


FIGURE 12

We see a steady decrease in the standard deviation of the grade distribution. This is evidence that the decline in standards has reached such a level that, while the ceiling effect is slowing the increase in grades, the ability of examinations to distinguish between students has been undermined. No doubt the small number of grades available on the scale has contributed to this phenomenon - in contrast to the Leaving Certificate Examination.

In summary, as for the Leaving Certificate Examination, there is strong evidence that standards have declined over the period. In addition, and more worryingly, the University examinations have seen a decrease in their ability to differentiate student performance.

In the final section, we will discuss the socio-political context in which all of these changes have been taking place.

4. Discussion

This work originates in the investigation of grade inflation in the Irish education system by the author and collaborators [16] - [22]. Educational decline, manifested through grade inflation, is the subject of much debate in other countries [11] [12] [15] [24] [25], but has only recently come to the fore in Ireland. For example, there is a strong body of evidence which shows that the introduction of student evaluation of teachers contributes to grade inflation [6] [8] [23].

In previous sections we have attempted to address the deficiency regarding theoretical constructs. In particular, given an educational setting and associated grade distribution, how is one to interpret this relationship and how it might change over time? Comparisons between institutions, educational awards or metrics are problematic and can lead to conflicting conclusions.

We have proposed a method for detecting declining standards through analysis of the internal relationship between parts of the grade distribution. Thus we propose a fingerprint for declining educational standards which can be detected without reference to external metrics, namely increasing grade mean and decreasing grade skewness. In advanced cases, decreasing standard deviation may also become evident.

As we have seen, this fingerprint appears in the grade data for the Irish educational system. The inevitable conclusion is that there has been a pervasive decline in educational standards, as measured by a lowering in the knowledge taxonomy.

One can pose the question, how is the lowering on the taxonomy actually occurring? At second level the issue continually high-lighted is predictability. Whether it is the use of an x rather than a t in a formula [14] or the absence of an old reliable question [4], the State Examinations Commission is routinely forced to publicly defend any examination that does not follow a predetermined pattern [10]. This predictability is further exploited by private schools specializing in examination tuition [7].

Predictability leads to a lowering in the taxonomy as it encourages rote-learning rather than understanding (dropping from level 2 to level 1 on Bloom's taxonomy). Ironically this has spawned a rote-learning culture at second level that exacerbates the unpreparedness of students entering higher education. Thus, even as the Universities and IoT's have been flooded with a large number of academically weaker students, the skill sets of the better students have been driven down the knowledge taxonomy.

At third level, predictability of examinations can also be an issue. In fact, the skewness it induces can easily been seen. If a lecturer gives a hint, say, that a topic will appear on an end-of-term examination, this information will only benefit a student to the degree to which they are able to take advantage of it. That is, the best students will pick up on it immediately and make a note, the average student may know that a hint has been given, but be unclear as to exactly what it

refers to, while the weak student, if they are even present, will have little awareness as to what has transpired. Thus negative skewness will be introduced into the examination grades.

Should the errant lecturer go so far as to show the students the test beforehand, not only will the mean jump, but most of the students will be squeezed into the top grades and will be well-nigh indistinguishable. The standard deviation will also have decreased.

Another possible mechanism for lowering standards is altering the relative weighting of harder and easier questions on marking schemes. If one reduces the proportion of marks awarded for a more difficult part of a question, it is not hard to show that it introduces negative skewness into the distribution with the usual consequences of dragging mid-lower grades over the mean.

There are many such mechanisms for declining standards which can cause the effects we are seeing in the grade distributions. The challenge is in finding methods to remove these mechanisms and to halt the decline in educational standards.

5. Appendix: Mathematical methods

The statistical methods used in this paper are entirely standard, but we briefly summarize them for the sake of convenience.

Suppose we have a frequency table for a distribution with n grades:

Table 4. Frequency of Grades

Grade	x_1	x_2	...	x_n
Frequency	f_1	f_2	...	f_n

Furthermore, assume that the frequencies are expressed as a percentage (as is the case in our work). Mathematically, this means that $\sum_{i=1}^{n} f_i = 1$. Such data is usually analysed by considering the *mean* $\bar{x}$ of the data, which is given by the equation

$$\bar{x} = \sum_{i=1}^{n} f_i x_i \tag{5.1}$$

This will have a value between 0% and 100%.

The dispersion of the data about the mean is measured by the *standard deviation* σ:

$$\sigma = \left(\sum_{i=1}^{n} f_i (x_i - \bar{x})^2 \right)^{\frac{1}{2}} \tag{5.2}$$

This is a positive number that measures how clustered or spread out the data is about the mean.

Finally, the *skewness* γ measures the asymmetry about the mean and is obtained from:

$$\gamma = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^3}{(\sum_{i=1}^{n} f_i (x_i - \bar{x})^2)^{\frac{3}{2}}} \tag{5.3}$$

This can have any positive or negative value, zero representing symmetric, negative meaning negatively skewed and positive meaning negatively skewed as per Figure 2.

The three numbers $\bar{x}$, σ and γ are independent in the sense that there could exist grade distributions in which they can take on any values. Thus an increase or decrease in one does not in itself imply any specific changes in the other two quantities. In this paper we have considered how one would expect these numbers to change for a grade distribution in an educational system undergoing declining standards. Conversely, if such movements appear in the data, as is the case for the Irish educational system, then there is a strong indication that standards have indeed been declining.

References

[1] L.W. Anderson, D.R. Krathwohl, P.W. Airasian, K.A. Cruikshank, R.E. Mayer, P.R. Pintrich, J. Raths, and M.C. Wittrock (eds), *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*, Allyn and Bacon, 2000.

[2] B.S. Bloom (ed), *Taxonomy of educational objectives: The classification of educational goals*, Susan Fauer Company, 1956.

[3] P. Callaghan, *Grade Inflation in Irish Universities?*, Report for the University Council, Trinity College Dublin, November 2009.

[4] G. Cooney, *Tough paper – but hell hath no fury like a maths student scorned*, Irish Independent, June 11, 2011.

[5] J. Cosgrove, G. Shiel, P. Archer, and R. Perkins, *Comparisons of Performance in Ireland PISA 2000 to PISA 2009: A Preliminary Report to the Department of Education and Skills*, Educational Research Centre, St Patricks College, Dublin 2010.

[6] C.F. Eiszler *College students' evaluations of teaching and grade inflation*, Research in Higher Education **43**, 483–501, 2002.

[7] S. Flynn, *Almost 50% of Leaving Cert pupils take grinds*, Irish Times, September 21, 2011.

[8] M.-L. Germain and T.A. Scandura, *Grade inflation and student individual differences as systematic bias in faculty evaluations*, Journal of Instructional Psychology **32**, 58–67 2005.

[9] Higher Education Authority, *Undergraduate graduates of HEA institutions: honours Bachelor degrees obtained by full-time and part-time students, by institution, award, and field of study*. Retrieved September 2011, from www.hea.ie

[10] L. Holden, *Furore forces exams commission to defend 'extremely challenging' paper*, June 11, 2011.

[11] R. Kamber and M. Biggs, *Grade conflation: a question of credibility*, The Chronicle of Higher Education, April 12, 2002.

[12] M.S. Kolevzon, *Grade inflation in higher education: A comparative study*, Research in Higher Education **15**, 195–212, 1981.

[13] E. Mac Aogáin, *The points system and grading of the leaving certificate examination*, Irish J. Educ. **36**, 3–24, 2005.

[14] N. Murray, *T proves to be problem in maths paper*, Irish Examiner, June 11, 2011.

[15] B. Nagle, *A proposal for dealing with grade inflation: the relative performance index* , Journal of Education for Business **74**, 40–43, 1998.

[16] M. OGrady and B. Guilfoyle,*Evidence of grade inflation 1994 - 2004 in the Institute of Technology sector in Ireland*, Paper 1, Network for Irish Educational Standards. Retrieved March, 2007, from www.stopgradeinflation.ie

[17] M. OGrady and B. Guilfoyle,*Grade inflation in Irish Universities (1994 to 2004)*, Paper 2, Network for Irish Educational Standards. Retrieved March, 2007, from www.stopgradeinflation.ie

[18] M. OGrady, *Regulation changes contributing to grade inflation: the NCEA/HETAC case study*, Paper 3, Network for Irish Educational Standards. Retrieved March, 2007, from www.stopgradeinflation.ie

[19] M. OGrady and S. Quinn, *The causes of grade inflation: an exploration of social and institutional pressures and policy choices*, Paper 4, Network for Irish Educational Standards. Retrieved March, 2007, from www.stopgradeinflation.ie

[20] M. OGrady and S. Quinn, *The consequences of grade inflation*, Paper 5, Network for Irish Educational Standards. Retrieved March, 2007, from www.stopgradeinflation.ie

[21] M. OGrady, *Grade inflation in the Leaving Certificate Examination 1992 - 2006*, Paper 7, Network for Irish Educational Standards. Retrieved January, 2009, from www.stopgradeinflation.ie

[22] M. OGrady, *Evidence of grade inflation 2005 - 2008 in the Institute of Technology sector in Ireland*, Paper 8, Network for Irish Educational Standards. Retrieved September, 2011, from www.stopgradeinflation.ie

[23] R.E. Redding, *Students' evaluations of teaching fuel grade inflation*, American Psychologist **53**, 1227–1228, 1998.

[24] S. Rojstaczer and C. Healy, *Grading in American Colleges and Universities*, Teachers College Record, retrieved April 2010, from www.tcrecord.org

[25] J.K. Shoemaker, M. DeVos, A. Loustau and A.B. O'Connor, *Are we a gift shop? A perspective on grade inflation*, Journal of Nursing Education **38(9)**, 394–99, 1990.

[26] State Examinations Commission, *State examinations statistics*. Retrieved June 2010, from www.examinations.ie

Brendan Guilfoyle, Department of Computing and Mathematics, Institute of Technology, Tralee, Clash, Tralee, Co. Kerry, Ireland.

E-mail address: brendan.guilfoyle@ittralee.ie