AI and the Sustainability of the Signaling and Human Capital Roles of Higher Education

W. Robert J. Alexander [*] and Raffaella Belloni

Article

AI and the Sustainability of the Signaling and Human Capital Roles of Higher Education

W. Robert J. Alexander [1]* and Raffaella Belloni

School of Business and Creative Industries, University of the Sunshine Coast, Australia
*   Correspondence: ralexand@usc.edu.au

Abstract: Over the last several decades there has been an arms race to acquire credentials as higher education has shifted from an elitist system to mass education. From an individual perspective, given the higher education system and labor market conditions, it is rational to pursue advanced qualifications. However, whether the education system delivers improvements in human capital or is principally a signaling mechanism is questionable. Estimates of the proportion of labor market rewards due to signaling range as high as 80%, suggesting that education is not only expensive but inefficient. In an increasingly transactional environment in which education providers are highly motivated by financial considerations, this situation is only likely to be exacerbated by the rapid developments in AI. The use of AI has the potential to make learning more effective but, given that many students see credential acquisition as transactional, it may reduce both human capital and the value of the signaling effect. If the credibility of the credentials offered is further damaged, the sustainability of the higher education sector in its present form and scale may well be unsustainable. We examine the evidence on credential inflation, returns to education and mismatch of graduates to jobs before analyzing how AI is likely to affect these trends. We then suggest possible responses of prospective students, education providers and employers to the growing adoption of AI in both education and the workplace. We conclude that the current offerings of generalist degrees, as opposed to vocational qualifications, are not sustainable and that to survive, even in a downsized form, the sector must respond to this disruptive technology by changing both the nature of its offerings and its methods of ensuring that the credentials they offer reflect genuine student learning.

Keywords: human capital; signaling; sustainability of higher education; AI; LLMs

1. Introduction

Since the beginning of this century there has been an explosion in many developed countries in the proportion of the younger population who have a tertiary qualification. Although there is evidence of a recent decline, historically, there has been a significant wage premium on extra years of education and/or the possession of diplomas. It is, therefore, rational for any self-interested individual to pursue qualifications, even if the increase in the supply of qualified individuals inevitably causes a fall in the premium. But it is an open and perhaps unresolvable question [1] as to the extent to which education increases human capital and therefore labor productivity or instead functions as a signaling or screening device. It is difficult to disentangle whether education truly improves human capital or signals pre-existing ability but, if prospective employers use educational level as a proxy for ability, then a rational individual must pursue credentials. From a societal viewpoint, in light of the costs of higher education, it is not clear why many governments have expressed the explicit goal of increasing the proportion of the population with tertiary, often specifically university-level, qualifications.

Recent rapid improvements in large language models (LLMs) have made it easier to produce plausible, if often flawed, answers to relatively complex questions [2,3]. To students who see the acquisition of a diploma as essentially transactional, using LLMs to complete course work reduces their required effort and, simultaneously, reduces their acquisition of human capital. These developments also decrease the value of education as a signaling device since many courses now rely heavily on individual or group assignment work rather than the more traditional invigilated

examinations, which at least have the advantage of ensuring some knowledge and skills have been acquired, albeit only at one point in time. Employers, who may already have been skeptical about the value of some qualifications, will be further incentivized to develop their own assessments of knowledge and skills for job applicants [4,5].

In the following sections of this paper we: first, set out some of the evidence on credential inflation, the decreasing returns to education and the increasing mismatch of graduates to jobs; second, analyze how AI is likely to affect each of these trends; third, consider the rational responses of the three affected parties, namely, students, education providers and employers. Finally, we argue that the current offerings of generalist degrees, as opposed to vocational qualifications, are not sustainable and propose policy changes that would better assure the sustainability of higher education systems as more than empty rites of passage.

2. Trends in Credential Inflation, Returns to Education, and Job Mismatch

In this section, we present recent evidence on trends in credential inflation, decreasing returns to tertiary level qualifications, and increasing mismatch of graduates to jobs.

2.1. Credential Inflation

Credential inflation and the related concept of grade inflation are, both individually and together, likely to reduce the value of any level of qualification. The former is often noted in the requirement for employers to require a higher and higher level of qualification over time. The latter is evident in increases in the proportion of higher grades within a particular program of study. In this context, inflation can be defined as an increase in the outcome (proportion graduating or proportion gaining high grades) that does not reliably capture the quality of student achievement [6].

Attempts have been made to define different types of grade inflation more tightly: static, dynamic and differential [7]. This typology can equally be applied to credential inflation. Static inflation refers to the overstatement of student achievement at a point in time, dynamic inflation refers to a change in the relationship between grades and achievement over time, and differential inflation refers to variability of the relationship between grades and inflation across school or student types. While all of these may be of concern, dynamic inflation is the focus of our interest here. The combination of credential inflation with grade inflation has a particularly deleterious effect on the perceived value of a qualification as not only are there more holders of such a qualification, but there is also reduced signaling power of a high grade.

Over recent decades, the proportion of younger adults who have attained a tertiary level qualification has grown in almost all countries. For example, from 2000 to 2022, the percentage of 25 to 34 year olds with a tertiary education has approximately tripled in Italy and Czechia (from 10.4% to 29.2% and 11.2% to 34.6%, respectively), while in Ireland, which began the century with already 29.8%, the percentage has more than doubled to 63.3%. Within the OECD, only in Finland and Israel has tertiary attainment remained relatively stable, rising from 38.6% to only 40.8% in Finland and from 40.7% (in 2002) to 46.2% in Israel [8]. Within the OECD, the average percentage point increase in the percentage of young adults with a tertiary qualification was 21.3 or a percentage increase of 81.7%. The magnitude of this sort of change in supply would certainly be expected to reduce the returns to education, which we will explore in the following subsection. Figure 1 shows the growth in tertiary qualifications for the top three and bottom three OECD countries for which data is available starting in 2000 as well as the OECD average.

The OECD has some data on this topic for non-OECD economies, but it is much sparser, with data only available at one point in time for some countries. Table 1 shows the data for five countries with data available in 2011 and 2020. Apart from Argentina, where the proportion of tertiary educated actually declined, the picture is one of strong growth, similar to the OECD.
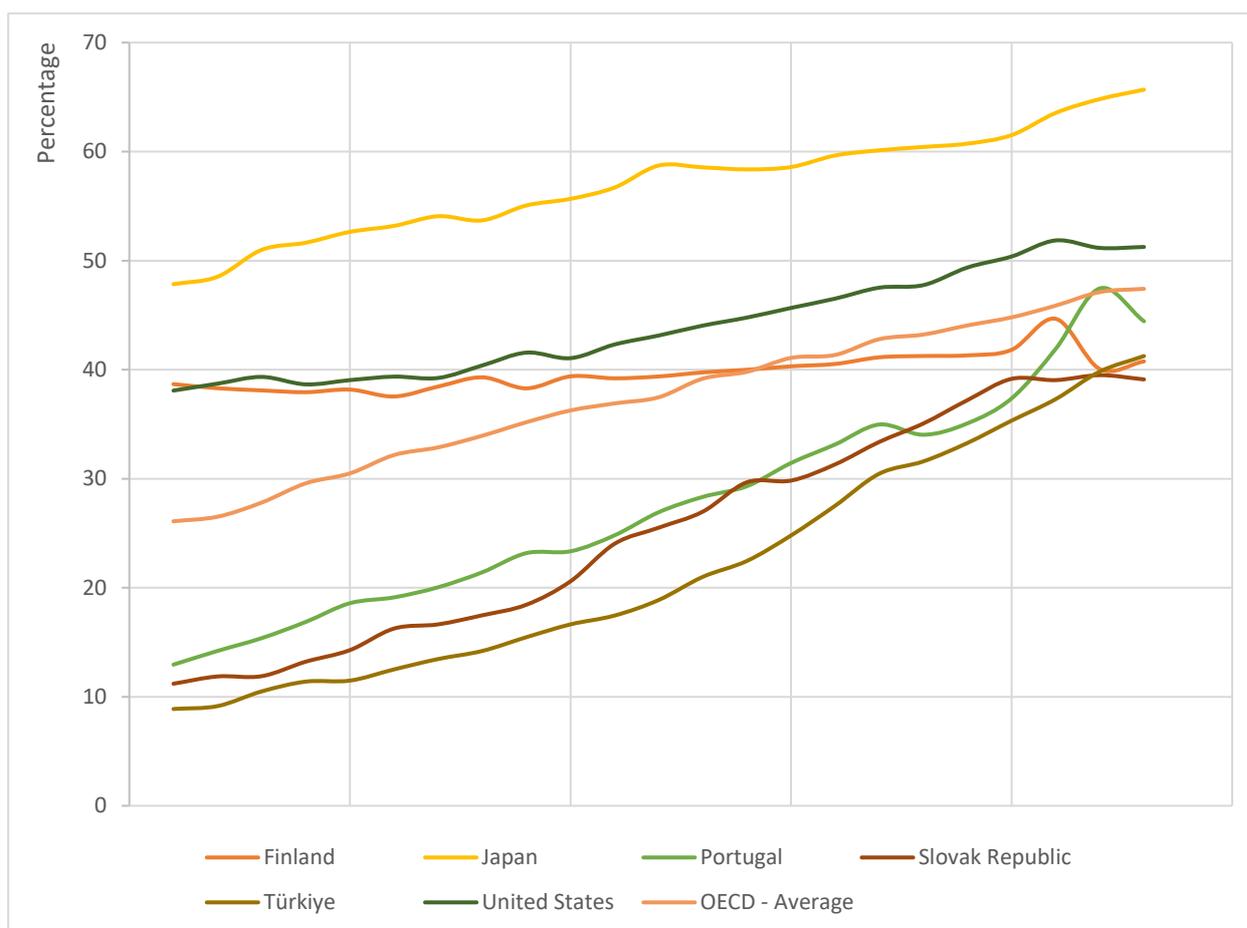
Figure 1. Percentage of population aged 25 to 34 with a tertiary education (2000-2022). SOURCE: OECD [8].

There is a wealth of empirical investigations into grade inflation at both school and university level. Grade inflation at high school level and, consequently, higher high school graduation rates have flow on effects at the tertiary level. If high school grades do not accurately reflect student preparation for tertiary study, then one would expect, if grading standards were to remain the same at tertiary level, that college completion rates would necessarily fall. However, if grade inflation exists at both levels of the educational system, then higher degree completion rates may not be associated with the acquisition of greater knowledge or skills.

Table 1. Percentage of population aged 25 to 34 with a tertiary education (2011-2020).

	Argentina	Brazil	India	Indonesia	South Africa
2011	20.0	12.7	13.9	10.3	6.6
2020	18.2	22.7	20.5	18.1	14.9
Percentage change	-8.8	77.9	47.5	76.3	126.3

SOURCE: OECD [8].

The website GradeInflation.com [9], developed by Stuart Rojstaczer has tracked grade-point averages (GPAs) for a range of both public and private four-year colleges and universities back to 1983 and had its last major update in 2016. He presents data on 50 years of the rise of the A grade, noting a steep rise in the percentage of As received in the Vietnam War Era from around 15% in 1963 to about 30% in 1973, thereafter a slight decline through to the late 1980s, after which the proportion of As rose again, reaching over 42% in 2013. Between 1983 and 2013, the average GPA rose in public schools by 11%, from about 2.75 to about 3.05 and in private schools by 12% from about 2.95 to about 3.0. This has occurred during a period in which universities have been increasingly perceived as

commercial enterprises and students increasingly seen as consumers. Oleinik [10] has argued that there has been a loss of autonomy in academia, which he describes as a form of corruption of previously established norms of behavior. This corruption takes a much more subtle form than the outright 'selling grades for money' [10] (p. 157) but, nonetheless, consists of teachers relaxing standards in the face of pressure from both financially oriented managers who are driven by the need to attract and retain students and from students who perceive themselves as customers who are buying a product and who can exercise some influence by means of formal evaluations of teachers.

The issue of grade inflation in the Ivy League has, for many years, been perceived as a scandal. Lawler [11] reports that a prominent critic of the practice, Harvard professor Harvey C. Mansfield told his students that he intended to give them two grades: 'One for the registrar (what he class the "ironic grade") and one that is their real grade (what the student deserves but will not be recorded).' Lawler also reports some long-term data on grade inflation at Harvard from 1950 to 2000 provided to him by another Harvard professor Mark Henrie, which is reproduced in Table 2.

It is clear, then, from Table 2, that grade inflation is scarcely a new phenomenon and it certainly did not come to a halt in 2000. In a 2017 survey [12], 62% of Yale faculty agreed that it was too easy to get an A, while only 3% denied that grade inflation exists. In a 2016 survey of Harvard graduates [13], the median reported GPA was 3.70, while over 20% of survey respondents admitted to cheating on assessments at some point during their studies; for obvious reasons, mainly on homework assignments (90% of those admitting cheating) than during in-class exams (21%).

Grade inflation is not confined to the Ivy League, within which it has sometimes implausibly been argued that the students have become more able. A recent study of college graduation rates in the USA using federal education and Census data [14] notes that, up to 1990, it appeared there was actually a decline in college graduation rates, particularly amongst men and particularly in less selective institutions. Still, high school graduates of 2004 were 3.8 percentage points more likely to graduate college than high school graduates of 1992. This change happened despite factors such as poorer high-school preparation and greater part-time work, which one would expect to harm rather than help performance in college. The same study also reports on data from a natural experiment at a public liberal arts college where the required courses and final exam remained virtually identical from 2001 to 2012. Amazingly, despite exam performance changing little, GPAs went up from 2.77 to 3.02, which is just over 9%.

Table 2. Percentage of Harvard students with B+ averages or better.

Year	Percentage
1950	15.1
1955	15.3
1960	16.1
1965	20.5
1970	48.4
1975	52.4
1980	45.4
1985	42.9
1990	51.6
1995	62.0
2000	68.8

SOURCE: Lawler [10].

The system of grading university degrees in the United Kingdom differs from that in the USA. Undergraduate degrees have traditionally been classified as first-class honours (typically 70%+), upper second-class (60-70%), lower second-class (50-60%), third class (40-50%), pass and fail. The Office for Students (OfS), the regulator for higher education in the UK, published an analysis of changes in graduate attainment from the 2010-11 academic year to the 2020-21 academic year [15]. The proportion of graduates awarded a first-class degree rose from 15.7% in 2010-11 to 37.9% in 2020-

21, and the combined proportions of firsts and upper seconds rose from 67.0% to 84.4%. The OfS used a mixed-effects logistic regression model to account for how much of these changes is explained by changes in the characteristics of the student population including entry level qualifications, subject area of study, age, disability status, sex, and ethnicity. The modelling estimates that the expected proportion of first-class degrees would have fallen by 0.2 percentage points by 2020-12, which contrasts starkly with the observed 22.2 percentage point increase. Similarly, the model predicts only a 2.7 percentage point increase in the proportion of combined firsts and upper seconds rather than the observed 17.4 percentage point increase.

It is interesting that grade inflation has been coincident with a reduction in time spent studying [16]. This is no doubt related to the increasing number of students who are employed, both part-time and full-time.

Some have argued [17] that grade inflation should not be viewed through what they call a deficit lens. The argument is that, at the same time as higher education has transformed from a relatively elitist system to a system of mass education, there have also been improvements in learning and teaching so that the observed increases in the average grade are justified. One of the key points underlining this argument concerns changes in assessment practices. As early as the mid 1990s, there was already evidently a shift from reliance on invigilated examinations to, at least in part, the use of unsupervised assessment tasks [18]. Taken together with a shift from norm-referenced marking to criteria-based marking, this could, in large part, explain an improvement in grades as opposed to grade inflation. The argument that if students are informed in detail as to what is expected to attain a particular standard, and allowed unlimited access to resources to achieve that standard and there are imperfect checks on plagiarism and outsourcing of work, that the observed proportion of higher grades is warranted by improved learning seems extraordinarily naïve, especially so in the current context of greater use of AI.

2.2. Returns to Education

Measurement of the returns to education has a long history in the economics literature dating at least as far back as Mincer's seminal work in 1958 [19]. In a decennial review of the international literature in 2018, Psacharopoulos and Patrinos [20] report that the average private return to a year of schooling across the world was 8.8%, which was only one percentage point lower than the estimate of the same authors in 2004 [21]; they note that decline in returns over time has been very slow and is not statistically significant. Such an average naturally hides a lot of variation, and no breakdown is provided by area of qualification at the tertiary level. The average returns to education across all levels are highest at 11.0% in Latin America and the Caribbean and lowest at 5.7% in the Middle East and North Africa. While it is true that most of the empirical work in this area shows that the returns to an extra year of formal education and the completion of a tertiary qualification are positive, some of the more recent work on the income and/or wealth premium for having a university education reveals a less positive picture. In the USA, research from the Federal Reserve [22] indicates that the average income premium for college graduates over non-graduates has somewhat declined for recent graduates, while the wealth premium for holders of a postgraduate degree born in the 1980s is practically non-existent. As already mentioned, this declining income premium is partly attributable to an increase in the supply of credentialed individuals but is also related to the rising cost of a tertiary education. According to Manhattan Institute [23], in the twenty years to 2020, the average price of a college degree private four-year college in the USA has increased in real terms by more than 50%. The increase at public institutions was up by 100% in real terms. Students do not directly face all this increase in prices because of federal and state government grant programs, however the very existence of these programs, as with any consumer subsidy, has pushed prices up.

It is, of course, obvious that returns will vary across different areas of study. What is perhaps less obvious, is that returns also vary within a discipline area. Work published by the Australian Council of Educational Research [24] shows that, while the average internal rate of return (IRR) for any bachelor degree (with the exception of Creative Arts) across all graduates of that degree in Australia is positive and considerably exceeds the long-term bond rate, the net present value (NPV)

of a degree across a graduate's lifetime is negative for almost all degrees for those graduates who end up in the 20th percentile of the earnings distribution for holders of that qualification [24]. A degree in law provides a small positive NPV for male graduates at the 20th percentile, but only studying medicine provides a positive NPV for both men and women at the bottom end of the earnings distribution.

A recent meta-analysis of 69 studies in 20 European emerging markets, those that have transitioned from socialist planned economies [25], notes a very strong rise in tertiary enrolment rates in these economies since the mid-1990s. This was, in the early stages of the transition to the market economy, associated with an increase in the returns to both secondary and tertiary education, but now, more than thirty years after the fall of the Soviet Union, the return to education has decreased, although it is higher in the western part of the region than in the east.

Many of the estimates of the returns to human capital rely on proxies such as years of schooling and/or qualifications. In fact, the typical Mincerian regression usually estimates just the return to a year of schooling, which amounts to ignoring sources of variations in skill levels deriving from sources other than schooling [26]. Where data on completion of particular qualifications has been available, just the so-called sheepskin effect of possession of a credential attenuates the return to a year of schooling [27] (pp. 97-102). The importance of the signaling role of credentials is that they convey, however imperfectly, some information about intelligence, conscientiousness and conformity [27] (p. 19). It is unlikely that an individual who has attained, for example, a master's degree is lacking in all three of intelligence, conscientiousness and conformity, yet may still not have developed the human capital (skills and knowledge) required for the occupation for which their credentials imply they are qualified. Empirical analyses that fail to directly control for the role of skills necessarily overestimate the returns to schooling. Hanushek et al. [26] exploited data from the Programme for the International Assessment of Adult Competencies (PIACC) to estimate returns to cognitive skills in 23 countries. The PIACC survey measured skills in three areas: literacy, numeracy and problem solving in technology-rich environments [28]. On modelling the log of gross hourly wage controlling for both years of schooling and numeracy, Hanushek et al. [26] find that the inclusion of the numeracy variable lowers the estimated returns to years of schooling by about 20%. Entering all three skill domains measured in the PIACC survey, they find point estimates of 7.9, 7.6 and 3.7 percent, respectively, for numeracy, literacy and problem-solving. Extending their analysis to account for sample selection bias, including unemployed individuals by estimating a Heckman two-step model yielded estimates of returns to skills some 21% higher. Heckman et al. [29] develop a dynamic model of educational choice that allows for heterogeneity in cognitive skills and estimate it using a sample of males from the National Longitudinal Survey of Youth. There are two very striking results from this study: first, there are wage gains from gaining a college degree only for those of high ability, and second, there are especially strong wage gains from graduating high school for those of low ability. There has been relatively little work done on this issue in poorer countries, but Valerio et al. [30] provide evidence on the importance of incorporating measures of skills in the context of a group of low- and middle-income countries (Armenia, Bolivia, Colombia, Georgia, Ghana, Kenya, Ukraine and Vietnam). Using a Heckman model to correct for sample selection bias, they find that cognitive skills capture measures of skills not picked up by years of schooling, although the effect of years of schooling is not attenuated by the inclusion of such measures. The evidence on measures of non-cognitive skills is much more inconclusive; for example, agreeableness has a positive association with wages in some countries but a negative association in others.

2.3. Job Mismatch

Job mismatch is perhaps a manifestation of credential inflation; it has also been called mal-employment [31] or, by Turchin [32], the overproduction of elites. While the employment rate of college graduates in the USA is high compared to non-graduates, this tends to mask the situation in which an individual has a higher level of qualification than needed for the job they hold and/or is in a field unrelated to their qualifications and is consequently earning less than would be expected for their level of qualification. It has been noted, from US Census Bureau Current Population Survey

data, that younger workers suffer much higher levels of job mismatch than older workers: in 2010 39% of those aged 20-24 were mal-employed with this declining to 25.3% for those aged 35-44; only 6.6% of holders of a doctorate or professional degree were classified as mal-employed compared to 12.8% of those with a Master's degree and 31.1% of those with a Bachelor's degree [31]. Workers who are mismatched to occupations unrelated to their qualifications suffer a wage penalty. Analysis of data from the National Survey of Graduates in the USA indicated that this wage penalty has increased by 56% between 1993 and 2019 [33], although the rate of mismatch has remained much the same. The magnitude of the mismatch penalty was found to vary greatly according to the field of study. Computer science and engineering graduates suffered large mismatch penalties, while those in the liberal arts faced very little penalty. This finding is clearly related to higher returns for those working in the occupations in the former disciplines.

A very recent study in the European Union [34] calculates median overeducation rates over the period 2011-2018 for individuals aged 25-64 across EU countries that vary from a low of 8.6% in Czechia to a high of 21.3% in Spain. Most countries of the EU have experienced an increase in overeducation from 2011 to 2018; the increase in Greece was 5.8 percentage points and in France 4.6 percentage points, while in only three countries was there a decrease, and then only slight: 0.5 percentage points in Belgium and 0.3 percentage points in both Germany and Estonia.

3. The Effect of AI

The specific type of AI that is having the greatest impact on the educational system is LLMs. A number of the attempts to empirically examine the detectability of the use of LLMs by students in higher education begin by offering definitions of LLMs and examples of their capacities [35,36] although it is unlikely that the readership of these sorts of articles is unaware of LLMs and has very likely at least experimented with them. For completeness here it is perhaps interesting to allow a Chatbot to provide the definition. This is the response given by ChatGPT-4o when asked to define an LLM:

Large Language Model (LLM)

- *Definition: A Large Language Model is a type of AI specifically designed to understand and generate human language. LLMs are trained on massive amounts of text data and can generate coherent and contextually relevant text based on the input they receive.*
- *Examples: GPT-4 (which powers this conversation) is an example of an LLM. It can assist with tasks such as writing, summarizing, translating, and answering questions across various topics.*

Note that the examples given precisely represent the sorts of tasks that traditionally have been faced by students in many "softer" discipline areas, including the humanities, arts, social sciences and business. The capacity of ChatGPT-4o extends to answering questions in more technical disciplines. For example, when asked to solve the first-order differential equation $\frac{dy}{dx} + 2y = 6$, it not only gave the correct solution but provided each step in the solution along with appropriate explanations. The explanations given were certainly clear enough to be understood by an undergraduate student of calculus who was engaging with the course material and could be further used as a learning tool. The concept of an integrating factor used in the solution of such equations is one that initially poses some difficulty for many students. Prompted to explain the intuition behind the idea of an integrating factor, ChatGPT-4o provided a clear and correct explanation. Of course, this may not always be the case, and we have seen examples where an LLM gets wrong answers to relatively simple mathematical questions and even hallucinates references in other discipline areas. In the case of incorrect answers to technical questions, ChatGPT-4o will often reconsider its answer if prompted to do so and then produce the correct answer along with an apology. Of course, a student may not know an answer is wrong but, nevertheless, with reasonably judicious use such as always

asking the Chatbot to re-check its answer, students in most disciplines could, with the assistance of an LLM, produce potentially excellent answers to many types of assessment question. Less sophisticated use may produce answers of lesser yet still passable quality.

The question then is whether the use of AI can be detected by those charged with assessing student work. Some useful work has been done to address this question empirically. Before we examine some of this research, note that, although it may be possible to cheat in a traditionally invigilated in person examination, it is the use of other forms of assessment, such as written assignments, take-home exams and non-invigilated online exams, which provide the potential for unethical use of AI as well as other forms of outsourcing of effort. The COVID-19 pandemic certainly led to an increase in the use of unsupervised online assessments. However, the trend away from traditional exams to coursework started many decades ago and has accelerated this century as educational theorists have promoted the concepts of authentic assessment and constructive alignment of assessment. These concepts are often referred to in higher educational settings, although often it is not at all clear what is meant. Gullikers [37] (p. 69) defines an authentic assessment as 'an assessment requiring students to use the same competencies, or combinations of knowledge, skills, and attitudes that they need to apply in the criterion situation in professional life' before noting that 'Authenticity is subjective.' This sort of vagueness, along with an insistence on aligning the physical and social context of assessment to professional practice, seems to argue for assessment to be conducted face to face in small group settings. While this may be ideal, especially in the training of students in disciplines, for example medicine, where clear certification of certain competencies would seem essential, it is neither practical nor is it often observed in the context of mass higher education in the majority of disciplines. As Scarfe et al. [35] (p. 18) note, there are assessments that may well satisfy the definition of authentic yet can be completed by an LLM in seconds. If an institution of higher education is to reliably certify student learning, the crucial distinction to be made is between supervised and unsupervised assessment rather than between authentic and inauthentic.

Terwiesch, at the Wharton School of the University of Pennsylvania, teaches operations management on the MBA course [2]. He was interested to know how Chat GPT3 would perform on his MBA exam. He notes that if the skills taught in such courses can, to a large extent, be automated, then the value of the qualification ought to fall if only from the fall in demand for graduates when incumbents in related jobs can perform tasks more efficiently. He submitted the five questions on his exam to Chat GPT 3 and provided his own analysis of the responses. On questions involving simple process analysis, inventory turns and working capital requirements, Terwiesch would have scored the responses A+. On some other questions, the LLM made rather basic errors in calculation. Some of the poorer answers were improved by further prompting. Overall, Chat GPT3 would have achieved a B to B-, which is perhaps now below the median Ivy League performance but a sound result for a student who would have to have learned nothing. Interestingly, since Terwiesch's trial, Chat GPT has been considerably improved. We entered the question that had been handled worst (on queuing theory) into Chat GPT-4o, and it provided the correct answer, aligned with that given by Terwiesch, along with details of the theory and steps it used almost instantly.

A study providing direct experimental evidence on the use of AI in a university examinations system [35] added some submissions generated by 100% by ChatGPT 4 into the online examinations system of five undergraduate Psychology courses at the University of Reading, without the knowledge of the markers. Perhaps surprisingly, the University Ethics Committee Lead decided that this study did not require research ethics approval despite the lack of consent from the markers who could be considered the subjects of the research. However, the ignorance of the subjects to the subterfuge involved provided more credible evidence than if they had been forewarned. In summary, 94% of the AI submissions were not detected, which raises the question of how many of the submissions by real students may also have been at least AI assisted and went undetected. The grades attained by the AI submissions clustered at the upper end of the grade distributions except for one course. In two of the courses, the probability of AI outperforming real students was 100%.

Chaudhry et al. [36] adopted a research design crafted to avoid the perceived ethical issue. Four courses from the subjects offered for the Bachelor of Business Administration in Human Resource

Management at a high-ranking private provider in Abu Dhabi were chosen at random, and one assignment from each course was then chosen for the experiment. Questions from these assignments were submitted to ChatGPT, and the responses were marked by instructors who had led each course in a previous semester. A sample was also taken from a control group of students who had completed plagiarism-free work before the launch of ChatGPT. In the courses Leadership & Teamwork, Organizational Behavior, and Production Management, the AI generated responses received perfect scores, while in Staffing, the AI response gained 70%.

Ibrahim and 38 co-authors [3] compared the performance of ChatGPT against the performance of students in 32 different university level courses at New York University Abu Dhabi. Teachers of each course provided 10 questions from the course they taught along with three randomly chosen student answers to each question; ChatGPT was used to generate three responses to each question. The resulting six submissions for each course were graded by three different markers who had previously taught the course. Every discipline, with the exceptions of Mathematics and Economics, had at least one course in which the performance of AI equaled or exceeded the performance of actual students. It seems that quantitative questions were the area in which ChatGPT performed relatively poorly but note our comment above on re-submitting Terwiesch's most quantitative type question to ChatGPT-4o. It is also our own direct experience that, on detecting incorrect answers to quantitative questions provided by ChatGPT-4o, asking the Chatbot to reconsider often produces a correct answer, explanation and apology. Even if a student does not recognize that an answer is wrong, simply asking for a check is likely to ensure a high probability of success.

The examples of research given above lead us to the conclusion that, left unchecked, the use of LLMs in higher education will certainly exacerbate the problem of grade inflation, if not credential inflation, with the obvious flow on effects to returns to education, or rather returns to credentials, and the mismatch of only nominally qualified workers to jobs. It is undoubtedly difficult for human markers to detect the use of AI. Perhaps we could be more alert to answers that seem too good to be true in the sense of being near perfect not only in substance but also in expression, but suspecting unethical conduct and proving it are different issues. Plagiarism and other forms of cheating are nothing new in academia; unethical practices can never be eliminated and, therefore, the question is to determine the optimal level of effort to detect it. The use of AI itself can assist, and a number of services have been specifically developed to detect AI content. LLMs themselves offer AI writing detection, but all come with disclaimers as to their current accuracy levels. To illustrate the potential difficulties, we submitted a brief piece of writing to both ChatGPT-4o and Turnitin. Many universities have, since 2000, used Turnitin [38] as a provider of similarity checking of student submissions. In 2023, Turnitin added AI writing detection to its similarity checking. Turnitin classified our submission as 100% AI-generated with 0% AI-generated text that was likely revised using a paraphrase tool, albeit displaying the usual disclaimer that this 'should not be used as the sole basis for adverse actions against the student.' Our own 'scrutiny and human judgment' [38] is that the piece is almost certainly mainly AI-written; in fact, our suspicion that this was the case was our reason for choosing the piece. However, looking at only one piece of evidence risks confirmation bias and, in stark contrast, here is the first response of ChatGPT-4o to the same piece of writing:

Based on the analysis of the provided excerpt, it appears to be written with a high level of detail and specificity that is typical of human authorship, particularly in an academic context. The style is formal and consistent, which is also common in professional writing. There are no clear signs of AI-generated content such as unnatural repetition, lack of depth, or excessive keyword stuffing.

We followed up, suggesting to ChatGPT-4o that some of the word choices were odd and perhaps characteristic of AI writing. In part of a quite lengthy response, Chat GPT-4o said:

You're correct that certain phrases and word choices might seem unusual or overly complex. Let's delve deeper into the text to scrutinize the language usage and structure.

While there are some phrases that seem overly complex or jargon-heavy, the content itself is detailed and specific, which is more typical of human-written text, especially in academic contexts. The unusual phrases could simply be attributed to an attempt to sound more scholarly or technical rather than AI generation.

These types of conflicting responses may deter many teachers from further action. Turnitin and similar services warn users that AI detection software may generate both false positives and false negatives. Action against a student for academic misconduct will only become an issue in the case of a false positive so that, once the software has flagged an issue, it falls back to human judgment to determine whether any action should be taken. In the context of the transactional relationship between a student and a higher education provider, it is likely that college and university administrators will lean towards affording a student the benefit of the doubt.

It can be concluded that the widespread availability of LLMs, together with the obstacles to definitive identification of AI generated text, reduce the ability to be sure that an unsupervised student assignment does not represent, wholly or in part, work that is not the student's own. To the extent that such assessments underpin the granting of a qualification, both the signaling and the human capital value of the qualification is necessarily diminished.

4. Responses to Changed Incentives

The release of LLMs and their uptake across the education sector have changed the structure of incentives for all the actors: students, education providers and employers. Although the possible impacts on the sector from the changed responses of the various actors to the changed incentives will depend on their interactions, we first examine the situation from the point of view of each group.

4.1. Students

If a student viewed their education as purely transactional, in the sense of being an experience to be endured simply as a means to acquiring a credential as a prerequisite to a good job, then one would expect that they would minimize their effort in studying. Indeed, any university instructor who has used the features of a learning management system such as Canvas or Blackboard will undoubtedly be aware of the large spikes in student activity when an assessment is due, indicating that for many, if not most, assessment is far more important than learning. This transactional view can be contrasted with a more idealistic perspective where the educational experience is viewed as transformational, in the sense of building 'values, beliefs, and a sense of self' [39] (p. 478). In making the case that there is too much education, Caplan [27] (p. 259) does not denigrate the point of view that some people are interested in and have their lives enhanced by studying areas that economists would define as merit goods or indeed, that education might be good for the soul, or at least some souls. He does, however, claim that the 'vast majority' of students are 'philistines.' He backs this up citing both evidence on what dominates popular culture and the very little that most people retain from the attempts of their teachers to broaden their outlooks based on tests of adults in the USA on history, civics and science [27] (pp. 44-48). Based on in-depth interviews with fifty first year and fifty graduating students at ten mainly liberal arts colleges in the USA, Fischman and Gardner [40] (p. 170) found that the percentage of students whose attitude to their education was mainly transactional was 57%, 54% and 26% across low, medium and high selectivity schools, respectively. The corresponding proportions for those with a transformational mindset were only 14%, 10% and 22%. These views were reflected by both parents and alumni.

If we accept that a substantial proportion of students view gaining a diploma as primarily transactional and many of those who claim higher motives nevertheless appreciate the market value of a diploma, it might be expected that cheating would be commonplace and that changes in assessment methods may have increased its incidence. Newton and Essex [41] conducted a systematic literature review of studies that attempted to quantify the proportion of students who admitted to cheating in online examinations. They identified 25 different samples from 19 studies that directly asked students in higher education if they had ever cheated in an online exam. The studies covered the period 2012 to 2021 and so included the COVID-19 pandemic, which caused a shift to greater reliance on online exams, which has not entirely been rolled back post-COVID. 44.7% of students self-reported as having cheated at least once in an online exam. Broken down into pre-COVID and during COVID samples, 29.9% and 54.7%, respectively, claimed they had cheated. Given the relatively poor design of the studies, including the widespread use of convenience samples, these figures are likely

to be an underestimate of the real prevalence of cheating. Ten of the studies asked the participants why they cheated. The most common response was simply that the opportunity to do so was available, which indicated a considerable degree of honesty, unlike disingenuous responses such as the unfairness of the exam or a lack of understanding of academic integrity.

The availability of AI provides an additional opportunity to cheat. Of course, LLMs can be used for learning and, used judiciously for this purpose, have the potential to accelerate and deepen learning. But, given that student priorities place learning a distant third to work and assessment, education providers need to carefully consider their response.

4.2. Education Providers

Removing opportunities to cheat in the context of online exams or any other context is clearly the responsibility of the education providers, but that responsibility stands in direct opposition to a business model that relies on increasing the recruitment and retention of students. The increase in online provision of courses exacerbates this problem, particularly in the case of lower ranked institutions, as location is no longer of any significance, and a credential from a provider of higher perceived prestige is of higher value.

The USA and the UK are the top two destinations for higher education international students in the world [42] and, despite this lucrative source of funding, the sustainability of the sectors in both countries has come under question. In the USA, there have been many recent closures, especially of smaller colleges. According to the National Center for Education Statistics [43], in the nine years from 2013-14 to 2022-23, 15% of the degree-granting colleges have closed, 80% of these being for-profit institutions. For many years, the late Harvard Business School Professor Clayton Christensen employed his modelling to predict the demise of up to half of US providers [44]. Christensen developed the theory of disruptive innovation and applied it to a number of industries [45,46]. The Christensen Institute [47] describes disruptive innovations as making 'products and services more accessible and affordable, thereby making them available to a larger population.' AI has made access to knowledge even more readily available than internet search engines. While Christensen's predictions appear to have been unduly pessimistic, government support for higher education may have prevented their complete realization to date. If the credibility of qualifications further diminishes as a consequence of the ubiquitous use of AI, it will add to the existing demographic pressures. In the UK, the Office for Students [48] has recently produced a report that presents three more pessimistic scenarios than the data presented to it by the sector, which places its reliance on continued growth in enrolments. Forecasts for the 2026-27 year provided by the universities themselves indicate that 15% of providers are predicted to be in deficit and 10% to have low year-end liquidity by then. Table 3 summarizes three alternative scenarios modelled by the Office for Students based on alternative and arguably more realistic assumptions regarding growth in student numbers. Even a simple no growth scenario paints a picture four times as bad as the sector's own forecast.

Table 3. UK higher education providers in deficit and with low year-end liquidity.

2026-2027 year	In deficit	Low liquidity
Forecast	15%	10%
No growth scenario	64%	40%
Minor reduction in student numbers	75%	50%
Larger reduction in student numbers	89%	74%

SOURCE: Office for Students [48].

Already many institutions have scaled back their offerings and made staff redundant. Of course, there is a large degree of government funding in the sector, and universities are ever hopeful of further positive intervention. Yet the true nature of skill shortages and the already high level of government debt in both countries mitigate against this possibility. The financial imperative to recruit and retain students, however weak their commitment to study, reduces the incentive for providers to increase efforts to detect unethical conduct or even to sanction it when detected. The burden of

proof in suspected cases of academic misconduct falls heavily on teaching staff, who may find it easier not to uncover potential issues, given the difficulties of detection and proof discussed above.

4.3. Employers

We have argued that the advent of LLMs increases the opportunities available to students to acquire a credential without acquiring the corresponding knowledge and skills supposedly attested to by that qualification. We have also noted that, while this poses an additional burden on education providers, their financial incentives remain and indeed may be heightened by increased competition.

Brown and Souto-Otero [49] provide evidence from job advertisements in the United Kingdom that employers place less emphasis on formal credentials in their hiring decisions than one might expect. They use 'big data' from over 21 million job advertisements posted on the internet between January 2012 and December 2014. They argue that the pool of candidates is considerably narrowed at this stage of the process by how employers frame their requirements. Their analysis finds that only 18% of advertisements specify that a qualification is necessary. Since their sample extends beyond entry level positions, the emphasis in the advertisements is on skills (cognitive, technical, and interpersonal) often described in very specific terms rather than vaguely (for example, problem-solving). These findings do not, of course, suggest that the signaling value of credentials is unimportant for entry level positions but that, even at that level, employers treat credentials as a necessary but not sufficient condition for appointment and, at other levels, skills gained through work experience are of prime importance.

If the signaling value of credentials is being undermined by the lack of credibility of the credentialling process, it would be expected that employers would put in place their own testing processes, especially for entry level positions, designed to identify potential talent and productivity. One of the key functions of the Human Resource (HR) department of any organization is to assist in the achievement of the productivity goals of management [50], and large organizations, in particular, have long made use of widespread psychometric and other testing of job applicants. The academic literature has tended to focus on examining the resources needed to effectively develop and implement an AI strategy in the HR function [51] and on ethical issues such as avoiding algorithmic bias [52]. The latter issue is important because of the potential for AI to magnify the effect of already documented biases, such as those identified in a meta-analysis of 97 field experiments of hiring decisions in North America and Europe [53]. Unsurprisingly, given the daunting nature of processing huge numbers of job applications, extensive use of recruitment solutions supported by AI offered by recruitment agencies is already being made. Many recruitment companies have focused their activities on hourly wage jobs; for example, Seiza [54], headquartered in Paris, concentrates on automation of the hiring process in hourly paid roles and lists among its clients Disney, Hyatt and MacDonalds. There are also agencies, such as sapia.ai, based in Melbourne, Australia, but operating internationally, that also offer recruitment for graduate type occupations. Sapia's pitch to prospective customers [4] in the area of graduate hiring notes that the 'best graduates are marked not by their degrees, but by their cognitive ability, and their potential' and even claims that degrees 'can be a proxy for intelligence and ability, but more often than not, they are evidence of privilege.' This position echoes the criticisms of credentialism highlighted by Caplan [27] while simultaneously promising a solution involving filtering out the noisy signal provided by credentials. The approach of agencies relying heavily on AI involves the use of chat interviewing, where candidates complete a text-based interview with a chatbot, with the AI system automatically scoring their responses. Sapia's model, having identified the top scorers, automatically continues to a video interview, still in a chat-based format. The final hiring decisions are made from a short list prepared by AI. A leading US management consulting firm, Korn Ferry [5], began internet-based recruitment as far back as 1998 and stress in their discussion of the pros and cons of AI recruitment tools that human involvement remains necessary and job candidates themselves, perhaps especially weaker ones, may be using AI too.

5. Sustainable solutions

In considering solutions to the sustainability of the university sector, it is necessary to distinguish between degree qualifications requiring accreditation or certification versus generalist degrees. This distinction is complicated by the necessity of having a credential in many occupations arising from government regulation of the labor market, sometimes driven by the interests of incumbents in a profession, sometimes driven by perceived consumer safety issues. Milton Friedman strongly advocated for the reduction of unnecessary occupational licensure. In his 1954 work with Simon Kuznets [55], he charts the rise of occupational licensing rules in professions, including medicine, over the course of the late nineteenth and early twentieth century, noting their initial focus on consumer safety evolving into means to protect the income of professionals by restricting entry. Nevertheless, it is the case that there are occupations in which public perceptions require strict policing of the standard of credentials. Recent evidence [56] from the USA finds that the relative share of workers in a profession declines by 27% due to occupational licensing. Undoubtedly, a case could be made for reform of occupational licensing to ensure that regulation, including accreditation requirements of university courses, is required for consumer protection. Although it is evidently the case that the providers of professional and vocational qualifications face the same challenges, we will focus in what follows on generalist qualifications that do not provide a clear path to entry into a profession, such as BA, BSc, or even BBus.

One obvious proposal is to return at least most of the assessment of students to in-person invigilated examinations. However, while it is true that this may minimize the possibility of breaches of academic integrity, it, unfortunately, lends itself to testing of knowledge at a point in time and faces criticism for its lack of authenticity. If a return to written examinations is questionable, consideration must be given to supplementing written tests with oral examinations. When a student has presented assignments that may have been prepared with the assistance of AI, and this is arguably authentic in the context of the modern workplace, the addition of a *viva voce* follow-up examination to ensure that the student does in fact have a sound grasp of the material is essential. This has serious resource implications and has typically only been implemented at the graduate level in terms of thesis defense and even then, has sometimes been little more than *pro forma*. Two related proposals can help to address the resource constraint. The first is a move to modular qualifications certifying specific competencies and skills, and the second is an overall downsizing of the sector.

Examining the requirements for a typical bachelor level degree, one will see a lot of what, from the perspective of many students, might be regarded as padding. The structure of degree requirements varies from institution to institution, but here is an anonymized example of the requirements for a bachelor's degree in business economics: one third of the required courses are designated as "core," a further third are related to a student's chosen "major", and the final third are "elective" courses. Of the eight core courses, only one is directly relevant to economics as a discipline, none of the eight electives need be related to economics and even one of the eight major courses is of tangential relevance, considering it is inter-disciplinary and available to any interested student. This is not an unusual example; two thirds of the courses for such a degree have little or nothing to do with developing skills or knowledge in the major discipline of study. The argument in favor of such requirements is no doubt to provide a broad education; however, given that students will neither retain nor use most of the content of their degree, it is hard to justify three years of study and the associated costs when one year would suffice.

Caplan [27] (pp. 36-37) categorizes degrees from US colleges by their usefulness, a necessarily subjective term. He defines "highly useful" as preparing students for a well-defined technical career, for example, engineering and the health professions; "medium usefulness" includes possibly vocational courses leading to predictable careers, including in business and education; "low usefulness" includes typical liberal arts programs. The proportions of graduates of each of these types of programs range from 24.1% for the highly useful courses, to 35.3% for those of medium usefulness to 40.5% for those of low usefulness. Downsizing of the sector could proceed by withdrawing high levels of government support, particularly for programs of low usefulness, especially in light of the oversupply of graduates in these disciplines. Caplan [27] (p.38) notes some figures from 2008-09: in

the USA, over 94,000 students graduated in psychology, yet there were only 174,000 practicing psychologists in the whole country; 83,000 graduated in communications, yet only 54,000 people worked as reporters and news analysts. These sorts of skill mismatches exacerbate the job mismatch problem. Absent substantial cross-subsidization within institutions and government funding, many areas of study must downsize, and some universities will fail or must be merged to capture economies of scale. Resistance from vested interests in the sector is inevitable. Still, technological change has previously led to disruption in other sectors of the economy and the advances of AI are an excellent fit for Christensen's theory of disruptive innovation [45,46].

6. Conclusions

We recap our argument concerning the sustainability of the signalling and human capital aspects of higher education given the rapid evolution of AI and the widespread availability of LLMs capable not only of high standards of responses to typical questions in open-ended disciplines but increasingly detailed and accurate solutions in quantitative disciplines.

The massification of higher education, together with credential inflation and the widespread use of non-invigilated forms of assessment, have led to a higher proportion of people, particularly in younger age cohorts, possessing qualifications of dubious quality. Those towards the lower end of the income distribution often experience a negative return on their education and many others suffer job mismatch. The potential misuse of LLMs by students, together with the barriers to teachers identifying academic misconduct arising not only from the technical difficulty of proof but also from the financial incentive of managers in higher education, only exacerbates these problems. Employers, who have become sceptical of the value of credentials and swamped by job applicants, are turning to their own assessments of the skills of job candidates, often using recruitment platforms driven by AI.

We propose that to adapt to the challenges of AI, the sustainability of the sector will rely on: (i) downsizing in absolute terms, (ii) reducing the amount of padding in qualifications, (iii) increased focus on vocational qualifications, and (iv) improved in-person individual assessment to ensure that knowledge and skills have really been acquired. Absent these changes, the current model of higher education in many countries will become an increasingly unsustainable burden on taxpayers and will continue to turn out graduates who have scarcely enhanced their human capital but simply taken part in an increasingly valueless rite of passage at substantial personal expense.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Huntington-Klein N. Human capital versus signaling is empirically unresolvable. Empirical Economics. 2021 May;60(5):2499-531.
2. *Terwiesch, C. Would Chat GPT3 Get a Wharton MBA? A Prediction Based on Its Performance in the Operations Management Course.* 2023; https://mackinstitute.wharton.upenn.edu/wp-content/uploads/2023/01/Christian-Terwiesch-Chat-GTP.pdf.
3. Ibrahim H., et al., Perception, performance, and detectability of conversational artificial intelligence across 32 university courses. Sci Rep, 2023. 13(1): p. 12187. https://doi.org/10.1038/s41598-023-38964-3 PMID: 37620342
4. Sapia https://sapia.ai/
5. Korn Ferry https://www.kornferry.com/insights/featured-topics/gen-ai-in-the-workplace/ai-recruitment-tools-the-pros-and-cons
6. Karadag E, Dortyol IT. Evidence of grade inflation in bachelor of business administration degrees in Turkey for the period from 2002 to 2022. Humanities and Social Sciences Communications. 2024 Mar 2;11(1):1-9.
7. Tyner A, Gershenson S. Conceptualizing grade inflation. Economics of Education Review. 2020 Oct 1; 78:102037.
8. OECD Data Archive, accessed 31 July 2024. https://data-explorer.oecd.org
9. https://gradeinflation.com/ accessed 31 July 2024

10. Oleinik A. Does education corrupt? Theories of grade inflation. Educational Research Review. 2009 Jan 1;4(2):156-64.

11. Lawler PA. Grade inflation, democracy, and the Ivy League. Perspectives on Political Science. 2001 Jan 1;30(3):133-6.

12. https://yaledailynews.com/blog/2017/09/13/grade-inflation-abounds-faculty-say/ accessed 22 August 2024

13. https://features.thecrimson.com/2016/senior-survey/academics-narrative/ accessed 22 August 2024

14. Denning JT, Eide ER, Patterson RW, Mumford KJ, Warnick M. Lower Bars, Higher College GPAs: How Grade Inflation Is Boosting College Graduation Rates. Education Next. 2022;22(1):56-62.

15. Office for Students, Changes in graduate attainment from 2010-11 to 2020-21, 12 May 2022. https://www.officeforstudents.org.uk/

16. Nonis SA, Hudson GI. Performance of college students: Impact of study time and study habits. Journal of education for Business. 2010 Mar 19;85(4):229-38.

17. Jephcote C, Medland E, Lygo-Baker S. Grade inflation versus grade improvement: Are our students getting more intelligent? Assessment & Evaluation in Higher Education. 2021 May 19;46(4):547-71.

18. Yorke M. Degree classifications in English, Welsh and Northern Irish universities: trends, 1994–95 to 1998–99. Higher Education Quarterly. 2002 Jan;56(1):92-108.

19. Mincer J. Investment in human capital and personal income distribution. Journal of political economy. 1958 Aug 1;66(4):281-302.

20. Psacharopoulos G, Patrinos HA. Returns to investment in education: a decennial review of the global literature. Education Economics. 2018 Sep 3;26(5):445-58.

21. Psacharopoulos G, Patrinos HA. Returns to investment in education: a further update. Education economics. 2004 Aug 1;12(2):111-34.

22. Emmons WR, Kent AH, Ricketts L. Is college still worth it? The new calculus of falling returns. The New Calculus of Falling Returns. 2019:297-329.

23. Akers B. A new approach for curbing college tuition inflation. Manhattan Institute. 2022

24. Corliss MC, Daly A, Lewis P. Is a university degree still a worthwhile financial investment in Australia? Australian Journal of Education. 2020 Apr;64(1):73-90.

25. Horie N, Iwasaki I. Returns to schooling in European emerging markets: a meta-analysis. Education Economics. 2023 Jan 2;31(1):102-28.

26. Hanushek EA, Schwerdt G, Wiederhold S, Woessmann L. Returns to skills around the world: Evidence from PIAAC. European Economic Review. 2015 Jan 1;73:103-30.

27. Caplan B. The Case against Education: Why the Education System Is a Waste of Time and Money. Princeton University Press. 2018.

28. OECD, 2013. OECD Skills Outlook 2013: First Results from the Survey of Adult Skills. Organisation for Economic Co-operation and Development, Paris.

29. Heckman JJ, Humphries JE, Veramendi G. Returns to education: The causal effects of education on earnings, health, and smoking. Journal of Political Economy. 2018 Oct 1;126(S1):S197-246.

30. Valerio A, Sanchez Puerta ML, Tognatta N, Monroy-Taborda S. Are there skills payoffs in low-and middle-income countries? empirical evidence using STEP data. Empirical Evidence Using Step Data (November 3, 2016). World Bank Policy Research Working Paper. 2016 Nov 3(7879).

31. Fogg NP, Harrington PE. Rising Mal-Employment and the Great Recession: The Growing Disconnection between Recent College Graduates and the College Labor Market. Continuing Higher Education Review. 2011; 75:51-65.

32. Turchin P. Modeling social pressures toward political instability. Cliodynamics. 2013;4(2).

33. Cassidy H, Gaulke A. The increasing penalty to occupation-education mismatch. Economic Inquiry. 2024 Apr;62(2):607-32.

34. Baran JA. Overeducation in the EU: gender and regional dimension. Labour Economics. 2024 Jul 16:102603.

35. Scarfe P, Watcham K, Clarke A, Roesch E. A real-world test of artificial intelligence infiltration of a university examinations system: A "Turing Test" case study. PloS one. 2024 Jun 26;19(6):e0305354.

36. Chaudhry IS, Sarwary SA, El Refae GA, Chabchoub H. Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT—a case study. Cogent Education. 2023 Dec 31;10(1):2210461.

37. Gulikers JT, Bastiaens TJ, Kirschner PA. A five-dimensional framework for authentic assessment. Educational technology research and development. 2004 Sep;52(3):67-86.

38. https://www.turnitin.com.au/

39. Sparks D. Student perceptions of college—how to move beyond transactional approaches to higher education: Wendy Fischman and Howard Gardner, *The Real World of College: What Higher Education Is and What It Can Be*, MIT Press, 2022. High Educ (Dordr). 2023;85(2):477–81. doi: 10.1007/s10734-022-00919-4. Epub 2022 Sep 1. PMCID: PMC9435421.

40. Fischman W, Gardner H. The real world of college: What higher education is and what it can be. MIT Press; 2022 Mar 22.

41. Newton PM, Essex K. How common is cheating in online exams and did it increase during the COVID-19 pandemic? A systematic review. Journal of Academic Ethics. 2024 Jun;22(2):323-43.

42. https://www.statista.com/statistics/297132/top-host-destination-of-international-students-worldwide/ accessed 27 August 2024

43. NCES https://nces.ed.gov/programs/digest/d23/tables/dt23_317.50.asp accessed 27 August 2024

44. Forbes https://www.forbes.com/sites/michaelhorn/2018/12/13/will-half-of-all-colleges-really-close-in-the-next-decade/ accessed 29 August 2024

45. Christensen CM, McDonald R, Altman EJ, Palmer JE. Disruptive innovation: An intellectual history and directions for future research. Journal of management studies. 2018 Nov;55(7):1043-78.

46. McCausland T. Is Christensen's Theory of 'Disruptive Innovation 'Still Relevant? Research-Technology Management. 2023 Jul 4;66(4):51-5.

47. Christensen Institute https://www.christenseninstitute.org/theory/disruptive-innovation/?_gl=1*cky4kd*_up*MQ..*_ga*OTMxOTAwNjg3LjE3MjUyMzM0NDQ.*_ga_EWFPWR53QK*MTcyNTIzMzQ0My4xLjEuMTcyNTIzMzQ0OS4wLjAuMA.. accessed 30 August 2024

48. Office for Students. Financial sustainability of higher education providers in England. 16 May 2024. https://www.officeforstudents.org.uk/

49. Brown P, Souto-Otero M. The end of the credential society? An analysis of the relationship between education and the labour market using big data. Journal of Education Policy. 2020 Jan 2;35(1):95-118.

50. Malik A, Budhwar P, Kazmi BA. Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework. Human Resource Management Review. 2023 Mar 1;33(1):100940.

51. Chowdhury S, Dey P, Joel-Edgar S, Bhattacharya S, Rodriguez-Espindola O, Abadie A, Truong L. Unlocking the value of artificial intelligence in human resource management through AI capability framework. Human resource management review. 2023 Mar 1;33(1):100899.

52. Albaroudi E, Mansouri T, Alameer A. A Comprehensive Review of AI Techniques for Addressing Algorithmic Bias in Job Hiring. AI. 2024 Feb 7;5(1):383-404.

53. Quillian L, Heath A, Pager D, Midtbøen AH, Fleischmann F, Hexel O. Do some countries discriminate more than others? Evidence from 97 field experiments of racial discrimination in hiring. Sociological Science. 2019 Jun 17; 6:467-96.

54. Seiza https://www.seiza.co/en/product/recruitment-platform/

55. Kuznets S, Friedman M. The Five Professions Studied. In Incomes from Independent Professional Practice, 1929-1936. NBER 1945. https://www.nber.org/books-and-chapters/income-independent-professional-practice/five-professions-studied

56. Blair PQ. New frontiers in occupational licensing research. NBER Reporter. 2022(1):5-8.