

How can major research findings about returns to qualifications illuminate the comparability of qualifications?

Jackie Greatorex

Research Division

Presented at the:

Journal of Vocational Educational and Training Conference, Worcester College, University of Oxford, 5-7 July 2013

British Educational Research Association Conference, University of Sussex, 3-5 September 2013

Abstract

Comparability of qualifications is a topic of national significance. When two qualifications lead to the same job it is important that they are of the same standard. This study investigates how major research findings about returns to qualifications illuminate the comparability of qualifications. Conventional comparability research by awarding organisations focuses on the demand of assessment tasks or the quality of the learners' work. Returns to qualifications fit within a broader view of comparability and provide different information. When the returns to qualifications are similar, the qualifications are comparable in terms of the economic value of the knowledge, skills, and personality attributes of the qualification holders. However, the qualifications do not necessarily assess the same knowledge, skills and competence. Returns to qualifications are subject to influencing factors; predominantly the supply of qualified people available for work and the demand for knowledge and skills. Other influences include grading standards, brand and government policy. Findings about the returns to qualifications can be synthesised with other evidence to determine which influencing factors are at work.

Introduction

The aim of this study is to consider how findings from returns to qualifications (returns) research can illuminate the comparability of qualifications from an economic perspective.

Returns measure how much more is earned by people with a particular qualification compared to people with similar characteristics who do not have the qualification. They are a statistical proxy for the productivity of people with a qualification, where productivity refers to the knowledge, skills, competencies and personality attributes a person uses in a job to provide goods and services of economic value (Machin and Vignoles, 2005, Sianesi, 2003). The definition can be adapted to apply to a group or type of qualifications, for example degrees.

Dickerson and Vignoles (2007) provide several examples of findings to returns to qualifications analyses including:

In aggregate, the returns to qualifications are quite similar for full-time men and women. The rate of returns to level 1 qualifications is negligible or zero; while at level 2 and above the returns are positive and significant, and quite substantial – around 13%- 16% for both level 2 and level 3 qualifications and rising to 23%-31% for level 4 and 5 qualifications (Dickerson and Vignoles, 2007, v).

Returns can be explained in economic terms. Qualifications are a positional good with potential benefits for qualification holders (van de Werfhorst, 2011). The supply of qualified people and the demand for people with particular qualifications can influence potential benefits of qualifications such as returns (Powdthavee and Vignoles, 2006, Page, 2007). For example, between 2004 to 2010 the returns to level 2 Apprenticeships decreased (Morse, 2012, Douglas et al., 2012). This may be because the number of adults (age 19 plus) starting Apprenticeships rose by 140% from 2006/07 to 2010/11 (Morse, 2012).

The paper proceeds as follows; an outline of the UK context, a summary of comparability research, a summary of returns research, an exploration of how returns research may be viewed as comparability research, the factors influencing returns and how awarding organisations may consider responding to such findings. Finally conclusions are drawn.

Both comparability and returns literature have many terms and abbreviations which are defined in the text. Both areas of research provide a vast body of knowledge and therefore key issues are briefly discussed.

Context

The UK has a complex qualifications system (Wolf, 2011, Leitch, 2006). There are hundreds of qualifications of several types at different levels in three qualifications frameworks (Isaacs, 2010), the Qualifications and Credit framework (QCF), the National Qualifications framework (NQF) and the Framework for Higher Education qualifications (FHEQ).

The QCF and NQF share a common levels framework; entry level, levels 1, 2, 3, 4, 5, 6, 7 and 8. Generally the qualifications in the QCF are vocational qualifications (VQ) and general (academic) qualifications (GQ) are in the NQF. The main GQ are General Certificate of Education (GCSE) and A Level. Most 16 year old learners sit GCSE examinations in an average of eight subjects at the end of compulsory schooling. Many 18 year olds sit A Level examinations in three or more subjects. VQ include Apprenticeships (work based government-funded training programmes) and BTEC (Business and Technology Education Council) qualifications in a variety of vocations. Awarding organisations develop qualifications (specifications and associated assessments). Note that some organisations have acronyms as a name.

Ofqual (regulator of qualifications in England and VQ in Northern Ireland, with the exception of the FHEQ qualifications) accredits qualifications in the QCF and NQF (Isaacs, 2010). The government funds schools and colleges to teach accredited qualifications (Isaacs, 2010). The NDAQ (National Database of Accredited Qualifications) shows 176 awarding organisations and the majority offer VQ.

Methods used in comparability research

Generally 'to have comparability' means to be equal, similar or alike. 'Comparable' usually means suitable for comparison or similar. Elliott (2011) explains that in comparability research the definition of 'comparability' and 'comparable' are operationalised in a variety of ways. In its broadest sense 'comparability of examination standards' concerns anything related to the comparison of one qualification (or group of qualifications) with another (Elliott, 2011). In this paper 'standard' can refer to both performance standards and the demand of the qualification. Elliott (2011) argues that definitions of comparability may or may not specify the context of the comparison. Newton (2008) uses a definition of 'comparability' which specifies the context of the comparison; he says that:

For two examinations to have comparable grading standards, students who score at equivalent grade boundary marks must be the same in terms of the character of their attainments. They must be equally good at knowing, understanding and being able to do X, where X is the set of knowledge, skill and understanding (KSU) that is common to both examined constructs (Newton, 2008, 4-5).

Ofqual uses a definition of 'comparability' that does not specify context:

The extent to which the same awards reached through different routes represent the same or equivalent levels of attainment (Ofqual, 2009).

Awarding organisations and Ofqual undertake comparability research to highlight similar, easy or difficult routes to jobs or further study. If results of such studies show a lack of comparability of standards then the results may be provided to appropriate authorities who determine what action is necessary to reduce any disparity. Broadly, there are three approaches to conducting such comparability research:

- comparing the demand of assessment tasks
- comparing the quality of learners' performance evidenced by their responses to assessment tasks
- comparing learners' achievement on multiple measures of attainment.

These approaches have been used to conduct research, particularly about GQs, for several years. Table 1 summarises the key requirements for conducting the studies, how a lack of comparability may be shown and examples of studies. Comparability research that combines well designed studies using different approaches provides comprehensive evidence about whether standards are comparable. For example, Greatorex et al. (2003), Arlett (2003) and Edwards and Adams (2003) compare the demand of the examinations offered by different awarding organisations as well as the performance of the learners who achieved particular grades in the same examinations. Further details of all the approaches are available in Elliott (2011) or Newton et al. (2007).

There are some circumstances in which comparability studies cannot be conducted as the requirements are not met, for example, if studies based on comparing the quality of work produced require physical examples of that work which might not be available. This can also happen when performance is assessed by a professional discussion, or after examination scripts have been destroyed after a period of time. A 'professional discussion' is a face to face discussion between a learner and an assessor. The assessor asks questions about how the learner went about particular tasks (such as work based activities), what they learnt, why they followed a particular course of action and what contingencies were in place. The learner answers the questions and provides supporting evidence to support their answers. For example, files, artefacts which may not be allowed out of the work-place. For further details see Scottish Qualifications Authority (2012), Scottish Qualifications Authority (undated-a), Scottish Qualifications Authority (undated-b) and Scottish Qualifications Authority (undated-c).

When the requirements for such comparability studies are not met, returns research methods can provide a means of highlighting where differences exist which can then be further explored.

Table 1 Key approaches to comparability studies and attendant requirements

Qualifications are compared using	Key requirements for the research	For example, a lack of comparability may occur when	Example(s) of a study
The demand of assessment tasks	<ul style="list-style-type: none"> • All (or a representative sample of) the assessment tasks • A robust sample of experts to judge demand • An appropriate scale(s) of demand 	The demand of assessment tasks is judged to be higher for one examination than another	Greatorex et al. (2012) QCA (2006)
The quality of learners' performance evidenced by their responses to	<ul style="list-style-type: none"> • All (or a representative sample of) the evidence of learners' performance such as responses to examination questions • A robust sample of experts 	The quality of learners' performance is judged to be greater for one examination than	Yim and Shaw (2009)

assessment tasks	<ul style="list-style-type: none"> to judge demand A measure of quality of learners' performance 	another	
Multiple measures of attainment	<ul style="list-style-type: none"> All (or a representative sample of) learners' marks /grades on multiple measures of attainment Each learners' marks/grade must be linked on the different measures Appropriate measures of attainment for example marks from the first and final years of the same degree, marks on two examinations for a qualification 	The learners' marks/grades are equivalent on one measure and different on another	<p>Bell and Dexter (2000)</p> <p>Murphy (2007)</p> <p>Newton (1997)</p>

Research about returns

Research about returns generally involves secondary statistical analysis of data obtained from surveys. The surveys usually have many participants and are longitudinal. Examples of frequently used surveys are named and described in Table 2. A minority of returns use administrative data, for example Patrignani and Conlon (2011) used Her Majesty's Revenue and Customs data about annual earnings.

Strengths and limitations

Before considering whether returns research may be used in comparability research it is important to consider the following issues.

The strength of the findings in a returns study depends on the available data. The statistical analyses control for the effects of variables to avoid influencing the effect of the variables under investigation. For example, the returns to a degree qualification could be due to earlier achievements such as A Level qualifications. The skills from the A Level qualifications helped people gain entrance to degree qualifications and made them valuable employees. The statistics control for this if the data is available. A statistical result could be found, however, the underlying reason for the result could depend on a factor that is not included in the data.

The results can be sensitive to the statistical model that is used. For example, contradictory results may be produced from alternative, but equally suitable, statistical models, often within the same report (Jenkins et al., 2007, Wolf, 2011, London Economics, 2011, Blundell et al., 2005). Therefore, exact figures for returns from different studies should only be compared when exactly the same statistical model and data were used (London Economics, 2011, McIntosh and Garrett, 2009).

The sample sizes, particularly with VQ can be small, and it is important to use only the results from a sizeable sample. For example Gambin et al. (2011) says that there small sample sizes when researching the returns to Apprenticeships.

The studies about returns are inconsistent regarding occupational groupings, sector groupings and categories of qualifications (Jenkins and Sabates, 2007). An 'occupation' is an activity that generates a person's regular source of livelihood. A 'sector' is part of the national economy, such as the public sector. There is some

overlap between occupations and sectors, for example people in the 'assembly line worker' occupation are likely to work in the Manufacturing sector. However, other occupations such as being a secretary or senior official occur in diverse sectors such as Manufacturing or Financial Services. The inconsistent groupings add to the complexity of using the research and comparing exact figures from various studies. However, general patterns in results can be synthesised from various surveys and statistical models.

The data and associated returns are historical experiences. The returns in previous years do not necessarily predict the returns in the future as the situation changes over time. As an example Hunt and McIntosh (2007) are cautious about the predicted impact on returns of lengthening compulsory education and other economic measures. However, others argue that returns inform decisions about which type of education or training yields the highest returns, for example Powdthavee and Vignoles (2006) and EEF (2012). EEF is a manufacturers' organisation, representing all aspects of the manufacturing sector in the UK.

Table 2 Survey names and descriptions

Survey name	Brief description of the survey
British Cohort Study (BCS)	<p>A survey of people born in England, Scotland and Wales in a week in 1970, started in 1970 and repeated approximately every five years.</p> <p>Data included details of the background of the mother, her pregnancy and labour and the first week of the baby's life. When participants were 10 and 16 years old, medical information was included. Data about the participants as 16 year olds included details of diet, leisure, behaviours and psychological characteristics. The remaining surveys covered education, employment, family formation, health, citizenship and values.</p>
British Household Panel Survey (BHPS)	<p>A nationally-representative random panel survey of UK households and began in 1991.</p> <p>Data included a range of social and economic indicators for households and individuals.</p>
General Household Survey (GHS)	<p>A representative survey of households in England, Scotland and Wales began in 1971 and was conducted yearly.</p> <p>Data included demographic information about household members, household and family information, household accommodation, housing tenure, consumer durables including vehicle ownership, migration, employment, pensions, education, health and use of health services, marriage, cohabitation, fertility history, and income.</p>
Individual Learner Record (ILR)	<p>A government survey of state funded FE in England began in 2003/04, and was collected yearly.</p> <p>Data included learner characteristics, funding, employment, route onto the course, outcomes and their final destination.</p>
Labour Force Survey (LFS)	<p>A representative sample survey of households in the UK began in 1973 and was repeated each quarter.</p> <p>Data included earnings, whether respondents were employed, hours worked, marital status, whether accommodation was rented/owned and the number of children in a household.</p>

Survey name	Brief description of the survey
National Child Development Survey (NCDS)	<p>A survey following people born in the same week of 1958 in England, Scotland and Wales, and repeated approximately every 5 years.</p> <p>The data included information on child development from birth to early adolescence, childcare, medical care, health, physical statistics, school readiness, home environment, educational progress, parental involvement, cognitive and social growth, family relationships, economic activity, income, training and housing.</p>
Annual School Census (ACS) Previously the Pupil Level Annual School Census (PLASC)	<p>A government census of state maintained primary, secondary and special schools as well as Academies in England and Wales, began in 1996.</p> <p>Data included pupil level data such as age, home postcode, mode of attendance, main ethnic group/background, special educational needs and free school meal eligibility. Data also included details about the school and its staff.</p>
Workplace Employment Relations Survey Information (WERS) Previously the Workplace Industrial Relations Survey (WIRS)	<p>A survey of a nationally representative sample of British workplaces, began in 1980 and was repeated every five years or so.</p> <p>Data included average gross hourly earnings, information on overtime hours, gender, full/part time working, recruitment and training, employee representation, payment systems, collective disputes and procedures, redundancies, grievance and discipline, equal opportunities, work-life balance and flexible working, health and safety, flexibility and performance, business strategy and workplace change.</p>
Youth Cohort Survey (YCS)	<p>A series of surveys of various cohorts (year groups of young people) following completion of compulsory education. This survey began in 1985 and was conducted at irregular intervals with a variety of sampling strategies.</p> <p>Data included young people's education and labour market experience, training, qualifications and demographic details.</p>

Compiled from The UK Data Service (ukdataservice.ac.uk), The Information Authority (theia.org.uk), the Economic and Social Research Council (esrc.ac.uk/research/survey/), the Department for Education (education.gov.uk/researchandstatistics/stats) and the literature referenced below.

Factors influencing returns

The terms 'comparable' and 'comparability' are used in various ways, including comparing between qualifications. Two or more qualifications are comparable in economic value if the returns to the qualifications are similar (Elliott, 2011). For instance, Jenkins et al. (2007) speak of the returns to National Vocational Qualification (NVQ) level 3 being comparable to the returns to other level 3 VQs, such as City & Guilds. Greatorex (2011) argues that when the returns are similar the qualifications are comparable in terms of:

- the productivity of the qualification holders
- the economic value of the knowledge, skills, competence and personality attributes of the qualification holders.

However, the qualifications do not necessarily assess the same knowledge, skills and competence. Returns do not focus on standards, and therefore, may not align with the results of conventional comparability studies.

As mentioned above, returns are generally explained in terms of economics; the supply of qualified people available for work and the demand for the knowledge and skills associated with particular qualifications (Powdthavee and Vignoles, 2006, Page, 2007). For instance, a decline in returns may be due to an increase in the supply of qualified people available for work. Table 3 and Table 4 each summarise the influencing factors which may have led to a range of results in returns studies.

When standards are mentioned, they are generally synonymous with the level of the qualification (Blanden et al., 2012, Page, 2007, Green et al., 2011). For example, returns increased with level (Garrett et al., 2010, Nikolaou and Theodossiou, 2006, Silles, 2007b). There is a paucity of research which defines standards as the demand of the assessment task or the quality of the learners' work. Furthermore, there is little research investigating whether recalibrating standards, such as setting tougher pass marks or grading standards, influences returns (Clark and See, 2011). Clark and See (2011) make two theoretical propositions:

- tougher grading standards result in fewer passes
- standards can rise without consequence until a tipping point at which the higher standard causes larger numbers of learners to drop out of the qualification.

Both these propositions theoretically lead to a reduction in the supply of qualified people and therefore an increase in returns. However their analysis showed that tougher grading standards did not impact on earnings. The lack of a statistical effect may be due to the small sample size, or variations in the cohort. However, in a different context, more lenient grading standards were linked to a decline in returns (Sulaiman, 2012), which fits the theoretical model.

Another factor which may influence the supply of qualified people, and subsequently returns, is the ability of the learners. 'Ability' is defined in various ways. The NCDS measured 'mathematical ability' and 'reading ability', which are often interpreted as 'innate ability' (Silles, 2007a). Furthermore, Blanden et al. (2010) use 'ability' to mean 'cognitive ability' and there is research which refers to 'ability' and offers no definition. Hereafter, 'ability' will be used to refer to any of these forms of ability. McIntosh (2007) argues that the number of people applying for Apprenticeships exceeds the number of places available, and adds that, if employers select the applicants in terms of ability that explains the high returns to Apprenticeships. The influence of learners' ability on returns is often referred to as the 'ability bias' (Jenkins et al., 2007, Dickson and Smith, 2011).

Table 3 A range of findings from returns research

		Different returns	Same returns
Key reasons for the result may include	The supply of qualified people available for work...	varied	was similar rose at the same rate that the demand decreased declined at the same rate that the demand increased
	The demand for the knowledge and skills...	varied	was similar rose at the same rate that the supply decreased declined at the same rate that the supply increased
	Learners' ability...	varied	was similar
	Grading standards...	varied	were equivalent in terms of economic value but the qualifications may not be rewarding similar knowledge and skills
	Employers...	valued the content of qualifications with a high return understood the content of qualifications with a high return placed less value on the content of qualifications with a low return did not understand the content of qualifications with a low return	equally valued the perceived content of two qualifications
	Occupational use of the knowledge and skills in qualifications...	was greater for some qualifications than others	was similar
	Users' trust in/value of a brand...	was higher than users' trust in / value of another brand of qualification.	of qualification was the same as their trust in / value of another brand of qualification
	Government policy...	varied	remained unchanged or ineffectual

Note that 'qualification' could be replaced with 'group of qualifications'.

The curriculum content of VQs influences returns (Dickerson and Vignoles, 2007, Greenwood et al., 2011, Jenkins et al., 2007). Low returns reflect that the curriculum content is of low value to employers (De Coulon and Vignoles, 2008) and variations in returns are probably due to the diverse content of the qualifications (Dickerson and Vignoles, 2007). Furthermore, there is a plethora of qualifications which vary in terms of curriculum content. Consequently it is difficult for employers to ascertain the true value of the VQs (Dickerson and Vignoles, 2007). The frequent changes in names and structure of the qualifications may contribute to employers', carers' and learners' lack of clarity about the content of qualifications (Wolf, 2011). If VQs have a high return in a variety of sectors they are arguably well understood by employers, and

their curriculum offers knowledge and skills which are in demand (Dickerson and Vignoles, 2007).

One of the factors influencing returns is whether the knowledge and skills from the qualification are used in the occupation. This is rarely accounted for in research, and infrequently investigated. However, Greenwood et al. (2011) find generally higher returns when the knowledge and skills from the qualification are used in the learners' subsequent occupation.

Table 4 Reasons for a range of changes to returns over time

		Decline in returns to a qualification	Increase in returns to a qualification
Key reasons for the result may include	The supply of qualified people available for work...	increased	decreased
	The demand for the knowledge and skills...	declined	increased
	Learners' ability...	declined	increased
	Grading standards...	became more lenient	became tougher
	Employers...	were not clear about the qualifications' content	valued the qualifications' content
	Occupational use of the knowledge and skills in qualifications...	decreased	increased
	Users' trust in/value of a brand...	decreased	increased
	Government policy...	changed	changed

Note that 'qualification' could be replaced with 'group of qualifications'.

Returns may be influenced by employers', carers' and learners' trust in or value of particular brands. Here 'brand' refers to awarding organisations (City & Guilds, OCR, and so on) or qualification types (BTEC, Cambridge Technical). Generally the research about returns shows that own-brand qualifications; BTEC, City & Guilds, Ordinary National Diploma (OND), Ordinary National Certificate (ONC) and Royal Society of Arts Examinations Board (RSA) resulted in higher returns than government initiated qualifications (for example NVQ) (Jenkins et al., 2007, McIntosh and Garrett, 2009). This is part of the wider market trend that own-brand qualifications flourish and there is lower demand for qualifications from government initiatives, such as NVQ, General National Vocational Qualification (GNVQ) and Advanced Vocational Certificate of Education (AVCE) (Stanton and Bailey, 2005). Brands are important as they can pre-date and outlive the government policies which change qualification names and the structure of qualifications, for example City & Guilds started in the nineteenth century (Wolf, 2011).

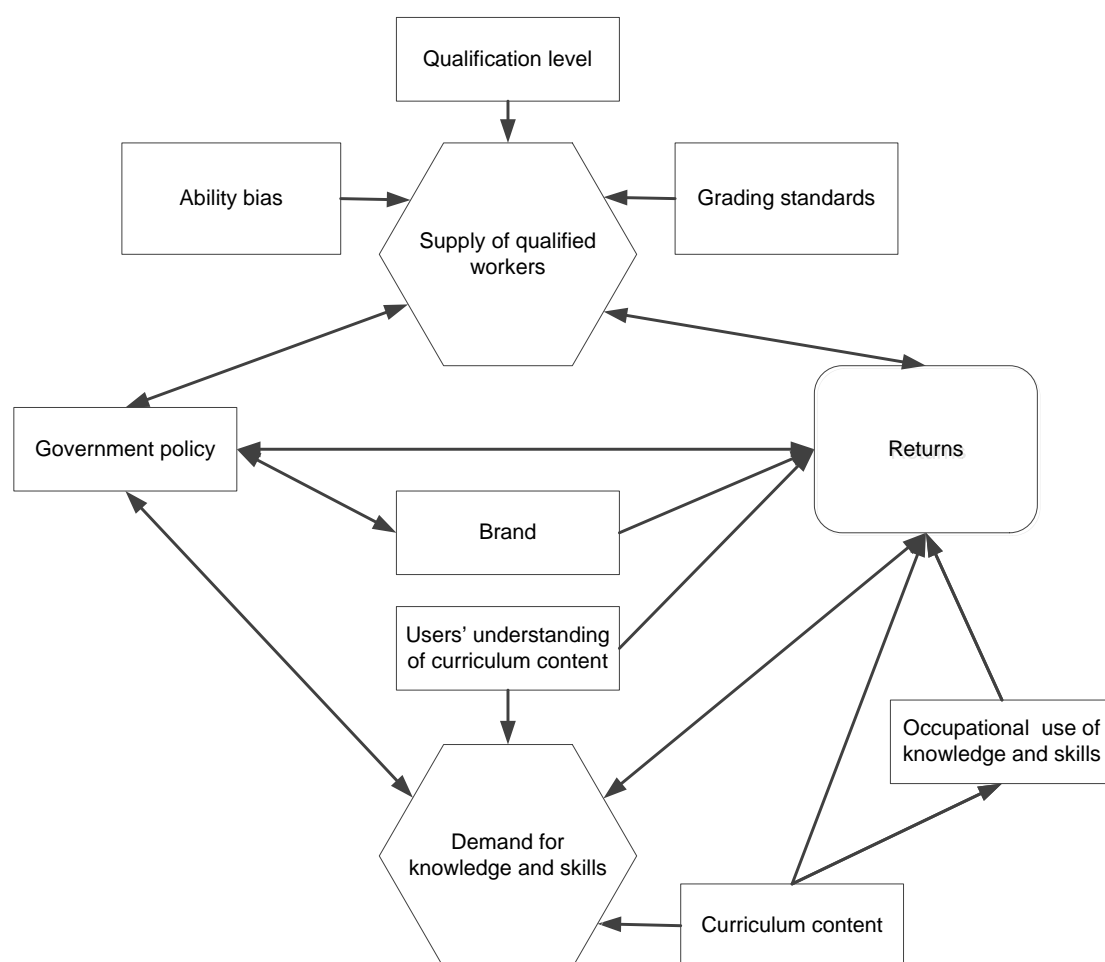
There can be a cyclical relationship between government policy and returns. Firstly, government policy can influence returns. For example, Hunt and McIntosh (2007) report several studies which show that increasing the school leaving age is associated with increased returns. Additionally, returns are used in government policy making. For example, the Department for Business Innovation and Skills (BIS) announced in December 2011 that they will focus on areas where Apprenticeships bring the greatest returns and the widest benefits (Morse, 2012).

There are ways of eliminating and confirming which factors are influencing returns depending on the available data and evidence. For instance the influence of supply, demand and applicants' ability can be quantified and included in analyses, or findings from trend analyses can be synthesised with findings from returns analyses (McIntosh, 2006). Qualitative research may help to answer questions which are not satisfactorily answered from the secondary analysis of longitudinal surveys such as:

- why low level VQs receive low returns (Dickerson and Vignoles, 2007)
- why particular qualifications yield a certain return (Greenwood et al., 2007).

The preceding discussion outlines relationships between factors influencing returns. For example, tougher grading standards can lead to a lower supply of qualified workers so returns increase. These relationships are summarised in Figure 1. The arrows indicate the direction of influence postulated in the literature, the major influences are represented in hexagons and the more minor influences are represented in boxes.

Figure 1 Factors influencing returns



Many of the influencing factors in Figure 1 are integral to awarding organisations' activities. These include:

- writing specifications for qualifications which must meet particular criteria to be allocated to a qualification level in the NQF or QCF
- setting grade boundaries for examinations or tests that are part of VQs, for example, the Angoff procedure is often used to set grade boundaries in VQs (Ofqual, 2012, Novakovic, 2008). In the Angoff procedure a group of subject

experts work together to define a borderline pass. The subject experts then consider each question and predict the percentage of learners that will answer the question correctly. An average of the percentages given across items and across subject experts is used to work out the pass mark for the examination.

- engaging with government (OCR, 2012, Edexcel, 2008, Rogers, 2009, Frank, 2011).
- offering own-brand qualifications which tend to achieve higher returns than qualifications from government initiatives (Jenkins et al., 2007, McIntosh and Garrett, 2009).
- contracting external verifiers to check assessors' judgements and other quality issues in centres (schools, colleges and work-places) including the users' understanding of curriculum content which underpins the centres' work (Greatorex, 2005, Stasz, 2011).
- engaging with curriculum content, which is why City & Guilds funds research about vocational teaching, learning and curriculum (Lucas, 2012).

Therefore returns research findings offer comparability evidence on a broad range of issues and might inform the improvement and development of qualifications and associated services, once the reason for the finding is established. Different returns do not always necessitate remedial action by an awarding organisation or Ofqual. For example, if there were different returns to qualifications in administration and hairdressing. However, others differences could indicate serious flaws which need to be remediated. For example, if there were differences between the returns for two qualifications which are meant to be the same, but are offered by different awarding organisation, then some action may be needed.

Conclusions

Returns research can be viewed as a particular form of comparability research, when 'comparability' is defined in its broadest sense. When the returns are similar the qualifications are comparable in terms of:

- the productivity of the qualification holders
- the economic value of the knowledge, skills, competence and personality attributes of the qualification holders.

However, the qualifications do not necessarily assess the same knowledge, skills and competence, and so the focus of the comparisons is different from conventional comparability studies.

Nevertheless returns research is multifaceted and can inform several key issues including government policy in education and skills and relative earnings. Such information is important for a variety of organisations including awarding organisations. For instance, generally the returns to own-brand qualifications were higher than returns from government initiated qualifications (Jenkins et al., 2007, McIntosh and Garrett, 2009). Moreover, Stanton and Bailey (2005) claim that in 2005 there were only own brand VQ in the NQF. This suggests that non-government organisations are better placed to develop fit for purpose qualifications which meet educational and economic needs, and that own-brand qualifications are trusted by employers and learners. Arguably this supports the Cambridge Assessment view that the government should give awarding organisations a free hand to work with stakeholders to develop qualifications (Lebus, 2012). Furthermore, returns research provides learners with information about relative earnings which, along with other information, may inform decisions about further study, whilst being mindful that returns are historical and are not a prediction of future earnings. Greenwood et al.

(2011) report positive returns to a range of VQ in STEM (Science, Technology, Engineering and Mathematics) subjects compared to holding the same qualification in non-STEM subjects. Such comparative information could be useful to students as they make decisions about further study.

This type of research provides additional information to be considered alongside comparability research related to inherent aspects of qualifications, such as standards of content and performance. Given the range of qualifications on offer in today's educational and assessment system, such a research method is a useful addition to the toolbox.

References

- Arlett, S. (2003). A Comparability Study in VCE Health and Social Care, Units 3, 4 and 6: A review of the examination requirements and a report on the cross moderation exercise. *A study based on the Summer 2002 examination and organised by AQA on behalf of the Joint Council for General Qualifications.*
- Bell, J. F. & Dexter, T. (2000). Using multilevel models to assess the comparability of examinations. *Fifth International Conference on Social Science Methodology.*
- Blanden, J., Buscha, F., Sturgis, P. & Urwin, P. (2010). Measuring the returns to lifelong learning *Centre for the Economics of Education Discussion Papers.* London: Centre for the Economics of Education.
- Blanden, J., Buscha, F., Sturgis, P. & Urwin, P. (2012). Measuring the earnings returns to lifelong learning in the UK. *Economics of Education Review*, 31(4), 501-514.
- Blundell, R., Dearden, L. & Sianesi, B. (2005). Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(3), 473-512.
- Clark, D. & See, E. (2011). The impact of tougher education standards: Evidence from Florida. *Economics of Education Review*, 30(6), 1123-1135.
- De Coulon, A. & Vignoles, A. (2008). An analysis of the benefit of NVQ2 qualifications acquired at age 26-34 *Centre for the Economics of Education* London: Centre for the Economics of Education.
- Dickerson, A. & Vignoles, A. (2007). The distribution and returns to qualifications in the Sector Skills Councils. *Sector Skills Development Agency: Research Series.* Department of Economics University of Sheffield; Institute of Education University of London.
- Dickson, M. & Smith, S. (2011). What determines the return to education: an extra year or a hurdle cleared? *IZA Discussion Paper.* Bonn: Institute of Labor in Bonn (IZA), University of Bristol.
- Douglas, R., Malinowski, M., J., P., Tlupova, D., Wright-Anderson, P. & Gray, P. (2012). Adult Apprenticeships. Estimating economic benefits from apprenticeships - Technical paper. London: Department for Business, Innovation and Skills, Skills Funding Agency and National Apprenticeship Service.
- Edexcel. (2008). Response to the QCF consultation. Available: http://www.edexcel.com/Policies/Documents/2008_4_Response_to_the_QCF_consultation.pdf [Accessed 4 July 2013].
- Edwards, E. & Adams, R. (2003). A Comparability Study in GCE Advanced Level Geography (Including the Scottish Advanced Higher Grade Examinations): A review of the examination requirements and a report on the cross moderation exercise.

- EEF. (2012). *Careers Guidance for Young People* [Online]. Parliament. Available: <http://www.publications.parliament.uk/pa/cm201213/cmselect/cmeduc/writev/632/m39.htm> [Accessed 22 November 2012].
- Elliott, G. (2011). A guide to comparability terms and methods. *Research Matters: A Cambridge Assessment Publication* [Online], Special Issue 2. Available: <http://www.cambridgeassessment.org.uk/images/109991-research-matters-special-issue-on-comparability.pdf> [Accessed 12 June 2013].
- Frank, M. (2011). The Wolf Report Report and Analysis from ASDAN. Available: http://www.asdan.org.uk/ASDAN_ESF_Launch [Accessed 23 June 2013].
- Gambin, L., Hasluck, C., Hogarth, T., Ryan, P. & Elias, P. (2011). Options study for the long-term evaluation of apprenticeships. *Department for Business Innovation and Skills Research Papers*. London: Warwick Institute for Employment Research.
- Garrett, R., Campbell, M. & Mason, G. (2010). The Value of Skills: An Evidence Review. *UKCES Evidence Reports*. London: UK Commission for Employment and Skills, National Institute for Economic and Social Research.
- Greatorex, J. (2005). Assessing the evidence: Different types of NVQ evidence and their impact on reliability and fairness. *Journal of Vocational Education and Training*, 57(2), 149-164.
- Greatorex, J. (2011). Comparing different types of qualifications: An alternative comparator. *Research Matters: A Cambridge Assessment Publication* [Online], Special Issue 2. Available: <http://www.cambridgeassessment.org.uk/images/109991-research-matters-special-issue-on-comparability.pdf> [Accessed 26 June 2013].
- Greatorex, J., Hamnett, L. & Bell, J. F. (2003). A Comparability Study in GCE A level Chemistry: A review of the examination requirements and a report on the cross moderation exercise.
- Greatorex, J., Shaw, S., Hodson, P. & Ireland, J. (2012). Using scales of cognitive demand in a validation study of Cambridge International A and AS Level Economics *Research Matters: A Cambridge Assessment Publication* [Online], 15. Available: http://www.cambridgeassessment.org.uk/ca/digitalAssets/205996_ResearchMatters15_-_01_2013.pdf [Accessed 22 May 2013].
- Green, A., Mason, G. & Unwin, L. (2011). Education and inequality: introduction. *National Institute of Economic review*, 215(1), 1-5.
- Greenwood, C., Harrison, M. & Vignoles, A. (2011). The labour market value of STEM qualifications and occupations London: Institute of Education.
- Greenwood, C., Jenkins, A. & Vignoles, A. (2007). The returns to qualifications in England: updating the evidence base on level 2 and level 3 vocational qualifications. *Centre for the Economics of Education Discussion Paper*. London: London School of Economics and Political Science.
- Hunt, E. & McIntosh, S. (2007). Raising the Participation Age: An Assessment of the Economic Benefits. *Department for Children, Schools and Families Research Report*. Department for Children, Schools and Families, Department of Economics, University of Sheffield.
- Isaacs, T. (2010). Educational assessment in England. *Assessment in Education: Principles, Policy & Practice*, 17(3), 315-334.
- Jenkins, A., Greenwood, C. & Vignoles, A. (2007). The returns to qualifications in England: Updating the evidence base on level 2 and level 3 vocational qualifications. London: Centre for the Economics of Education.
- Jenkins, A. & Sabates, R. (2007). The classification of qualifications in social surveys. *Working papers*. London: Institute of Education, University of London.
- Lebus. (2012). Speech to staff about the challenges faced by examination boards. Available: <http://www.cambridgeassessment.org.uk/images/113238-simon->

- [lebus-speech-to-staff-about-the-challenges-facing-exam-boards.pdf](#)
[Accessed 3 July 2013].
- Leitch, S. (2006). Prosperity for all in the global economy - world class skills. Final Report. London: HM Treasury.
- London Economics (2011). Returns to Intermediate and Low Level Vocational Qualifications. *BIS Research Paper*. London: London Economics.
- Lucas, B. (2012). A theory of vocational pedagogy. *12th Learning and Skills Research Network National Research Event*. Pearson College, London: NIACE.
- Machin, S. & Vignoles, A. (2005). *What's the good of education? The economics of education in the UK*, Princeton, New Jersey, Woodstock Oxfordshire, Princeton University Press.
- McIntosh, S. (2006). Further analysis of the returns to academic and vocational qualifications. *Oxford Bulletin of Economics and Statistics*, 68(2), 225-251.
- McIntosh, S. (2007). A cost-benefit analysis of apprenticeships and other vocational qualifications. *Research Report* University of Sheffield.
- McIntosh, S. & Garrett, R. (2009). The economic value of intermediate vocational education and qualifications. *UKCES Evidence Report*. Wath-upon-Deane, London: University of Sheffield.
- Morse, A. (2012). Adult Apprenticeships. London: The Stationery Office.
- Murphy, R. (2007). Common test methods. In: Newton, P., Baird, J., Goldstein, H., Patrick, H. & Tymms, P. (eds.) *Techniques for monitoring comparability of examination standards*. London: QCA.
- Newton, P., Baird, J., Goldstein, H., Patrick, H. & Tymms, P. (2007). *Techniques for monitoring the comparability of examination standards*, Qualifications and Curriculum Authority London.
- Newton, P. E. (1997). Measuring comparability of standards between subjects: Why our statistical techniques do not make the grade. *British Educational Research Journal*, 23(4), 433-449.
- Newton, P. E. Exploring tacit assumptions about comparability. 34th annual conference of the International Association for Educational Assessment, September, 2008 Cambridge, UK. Ofqual, 7-12.
- Nikolaou, A. & Theodossiou, I. (2006). Returns to qualifications and occupation for males and females: evidence from the British Workplace Employee Relations Survey (WERS) 1998. *Applied Economics Letters*, 13(10), 665-673.
- Novakovic, N. (2008). Thinking on the edge: the influence of discussion and statistical data on awarders' perceptions of borderline candidates in an Angoff awarding meeting. *International Journal of Training Research*, 6(1), 74-102.
- OCR. (2012). *Written evidence submitted by Oxford Cambridge and RSA Examinations (OCR)* [Online]. Parliament. Available: <http://www.publications.parliament.uk/pa/cm201213/cmselect/cmeduc/632/632vw47.htm> [Accessed 23 June 2013].
- Ofqual. (2009). *Glossary* [Online]. The National Archives. Available: <http://webarchive.nationalarchives.gov.uk/+www.ofqual.gov.uk/501.aspx#C> [Accessed 20 June 2013].
- Ofqual (2012). Findings from the review of Level 4 Retail Investment Adviser Qualifications. Ofqual.
- Page, R. (2007). The Impact of Gaining an NVQ Level 2: Will the Leitch Review Recommendations Address the Low Returns? *Local Economy*, 22(2), 138-147.
- Powdthavee, N. & Vignoles, A. (2006). Using Rate of Return Analyses to Understand Sector Skill Needs. CEE DP 70. *Centre for the Economics of Education*. London.
- QCA (2006). Comparability study of assessment practice. Personal licence holder qualifications. London: QCA.

- Rogers, G. (2009). Response to the Children's, Schools and Families Select Committee 'New Inquiry into School Accountability' on behalf of EDEXCEL Ltd.
- Scottish Qualifications Authority. (2012). *Evidence gathering form* [Online]. Available: http://www.understandingstandards.org.uk/Subjects/SVQ/BusinessAdministration/images/Highlighted_professional_discussion.pdf [Accessed 12 June 2013].
- Scottish Qualifications Authority. (undated-a). Case study 2 — Professional Discussion. *Guide to Assessment* [Online]. Available: http://www.sqa.org.uk/files_ccc/GuideToAssessment_CaseStudy2_ProfessionalDiscussion.pdf [Accessed 12 June 2013].
- Scottish Qualifications Authority. (undated-b). *Professional Discussion* [Online]. Scottish Qualifications Authority. Available: http://www.understandingstandards.org.uk/Subjects/SVQ/BusinessAdministration/Evidence/professional_discussion [Accessed 10 June 2013].
- Scottish Qualifications Authority. (undated-c). *Recording an Audit Trail* [Online]. Scottish Qualifications Authority. Available: http://www.understandingstandards.org.uk/Subjects/SVQ/BusinessAdministration/Processes/audit_trail [Accessed 12 June 2013].
- Sianesi, B. (2003). Returns to education: a non-technical summary of CEE work and policy discussion. *Institute for Fiscal Studies and Centre for the Economics of Education* [Online]. Available: http://194.66.89.8/docs/cee_summ.pdf [Accessed 27 May 2011].
- Silles, M. A. (2007a). Adult education and earnings: evidence from Britain. *Bulletin of Economic Research*, 59(4), 313-326.
- Silles, M. A. (2007b). The returns to education for the United Kingdom. *Journal of Applied Economics*, 10(1), 391-413.
- Stanton, G. & Bailey, B. (2005). In search of VET. *SKOPE Research Paper*. Coventry: ESRC funded Centre on Skills, Knowledge and Organisational Performance.
- Stasz, C. (2011). The purposes and validity of vocational qualifications. *SKOPE Research Paper*. Cardiff: RAND Corporation.
- Sulaiman, M. (2012). *Social protection and human capital accumulation in developing countries*. The London School of Economics and Political Science (LSE).
- van de Werfhorst, H. G. (2011). Skills, positional good or social closure? The role of education across structural–institutional labour market settings. *Journal of Education and Work*, 24(5), 521-548.
- Wolf, A. (2011). Review of Vocational Education - The Wolf Report. Cheshire: King's College London.
- Yim, L. & Shaw, S. (2009). A comparability study using a rank-ordering methodology at syllabus level between examination boards. *International Association for Educational Assessment*. Brisbane, Australia.