Research in Comparative and International Education
Volume 6 Number 3 2011
www.wwwords.uk/RCIE

RESEARCH IN Comparative & International Education

The Use of Large-Scale Administrative Data Sets to Monitor Progression from Vocational Education and Training into Higher Education in the UK: possibilities and methodological challenges

GEOFF HAYWARD
School of Education, University of Leeds, United Kingdom
MICHAEL HOELSCHER
Institute of Sociology, University of Heidelberg, Germany

ABSTRACT This article describes two administrative data sources - UCAS applicant data and Higher Education Statistical Agency (HESA) data - and demonstrates how they can be utilised to monitor the progression of students from vocational and educational training (VET) programmes in to higher education (HE) in the UK. First the article provides a general discussion of the administrative data sets, their coverage and resulting limitations. Second, it illustrates how a combined HESA-UCAS data set can be utilised to provide a descriptive statistical analysis that demonstrates that including VET students within HE programmes widens participation for under-represented groups in the UK but shows that VET students are at a disadvantage compared with those progressing through academic pathways in terms of their probability of gaining access, which types of higher education institution (HEI) they gain access to, and an increased risk of dropping out during their first year of study. Third, multilevel modelling indicates that drop-out risks differ across HEIs and suggests that HEIs can take actions that will reduce the risk of VET students dropping out. Finally some conclusions about the utility of administrative data to answer substantive research questions are presented, with suggestions for improving the quality and access to such data for researchers.

Introduction

A central tenet of social policy in the UK under both New Labour and the current coalition governments is that economic efficiency and social justice run together. Potential workers are construed as actors striving to make themselves marketable in a more flexible labour market that policy makers should not interfere with, for example, through industrial policy.

> What all this means is not that the role of Government, of the collective, of the services of the State is redundant; but changed. The rule now is not to interfere with the necessary flexibility an employer requires to operate successfully in a highly fluid changing economic market. It is to equip the employee to survive, prosper and develop in such a market, to give them the flexibility to be able to choose a wide range of jobs and to fit family and work/life together. (Blair, 2007)

The state's role is thereby reconceptualised primarily as ensuring adequate opportunities for individuals to develop human capital rather than, say, regulating the labour market. Such a policy vision is predicated on a belief that:

1. The development of more diverse educational opportunities beyond the age of 16 will increase the participation rate of learners beyond compulsory schooling, a decisive pre-condition for

 http://dx.doi.org/10.2304/rcie.2011.6.3.316

increasing and widening participation in HE (see, for instance, Education and Employment Committee, 2001, p. 33); and

2. Any increase in educational participation and attainment will produce both individual and social returns on investment in further education and training.

Educational participation beyond the compulsory school age has increased in the UK constantly since 1945, with a massive increase in participation in full-time provision between 1985 and 1994 (Hayward et al, 2004; Hayward et al, 2005, 2006; Hayward, 2006; Pring et al, 2009). This increase can partly be attributed to the growing availability of Level 3 vocationally oriented qualifications aimed at 16-year-olds, particularly those with lower levels of prior school attainment that are highly correlated with demographic characteristics such as lower socio-economic status. Such qualifications are increasingly marketed as providing a means for progressing into HE, thus constituting an important component of attempts to widen participation for under-represented groups, especially those from poorer backgrounds, as a key means of both achieving social mobility and improving national economic performance.

However, the far-reaching discourse about HE access 'is the most troublesome item in talk about HE' (Watson, 2006, p. 93), because the research base in this area is fragmented to a degree that almost any conclusion can be drawn from it (Gorard et al, 2006). Linked to this debate about HE participation, vocational education and training (VET) policy has as one of its aims increasing access to HE via this route. However, there is little research that speaks to either the effectiveness or the efficiency of this policy, the redistributive potential of which lies not just in whether an individual participates in HE but also in which higher education institutions (HEIs) they access and in which subject areas (Chevalier & Conlon, 2003; Sloane & O'Leary, 2004; Walker & Zhu, 2005; Bratti et al, 2006; Powdthavee & Vignoles, 2007).

UK and international research has indicated that the connections made in policy discourse, between expanding post-compulsory participation via an increased vocational offer and participation in HE, are not necessarily realised in practice. For example, in the UK investigations into the educational value of some vocationally oriented qualifications, in terms of their currency for further progression, have concluded that they only offer a 'mirage of wider opportunities' (Pugsley, 2004, p. 28; see also Connor et al, 2006; Vickers & Bekhradnia, 2007; Wilde & Hoelscher, 2007). Given that the VET pathway is often construed as an alternative chance for those deemed 'unsuitable' for progression via the academic pathway, there is, therefore, a need to assess how good an alternative it is, in terms of where, what and why graduates of the VET system study in HE. The challenge is finding data sets that can be analysed to make such an assessment.

Assessing the Transition from VET to Higher Education

A variety of measures exist that purport to measure the effectiveness of VET systems. However, typically these are measures of stocks – the number or proportion of people enrolled on particular courses at particular levels; the number or proportion holding qualifications of a particular kind at a particular level – or various measures of individual rates of return to individuals who hold a particular level of qualification, deemed to be a proxy for the value of a qualification in the labour market. Such measures are useful, but they tell us little about the progression of individuals, or the factors that affect such progression (including the contribution of VET programmes), between different phases in their learning career - for example, from secondary to tertiary education or from VET programmes into the labour market, relative to other individuals who have followed other learning pathways. Without detailed information about such transitions, and about factors affecting the probability of success, it is difficult to provide individuals with sound career advice and guidance about which pathway to follow to realise their ambitions. Yet, arguably, this is how we should judge the value of VET programmes: the extent to which they provide access to valuable outcomes, such as good jobs or progression to higher-level education and training programmes that in turn lead to successful insertion into the labour market, as opposed to, say, 'warehousing' people in education and training programmes through times of economic downturn with reduced labour market opportunities (see DfE, 2011; Hayward & Williams, 2011).

The challenge with providing evidence about the success, or otherwise, of such transitions is that it requires longitudinal data of sufficient quality to exert the statistical control needed to make,

ceteris paribus, valid claims about the effectiveness (or otherwise) of particular VET interventions to support transitions to successful outcomes. Such data are very expensive to generate and are, accordingly, rare. They do exist - for example, in England there are the various birth cohort studies, but the gaps between each cohort preclude drawing conclusions about issues of interest associated with changes in the labour market and the HE widening participation agenda in the UK, for instance, since the 1990s.

Other panel studies, such as the English Longitudinal Survey of Young People (LSYPE) [1], will provide useful information for one group of young people progressing to age 25. The main role of the study is to identify, and enable analysis and understanding of, the key factors affecting young people's progress in transition from the later years of compulsory education, through any subsequent education or training, to entry into the labour market or other outcomes. Data from the study will be used, among other things, to monitor the progress of the cohort group, evaluate the success or otherwise of policy aimed at this group and provide an evidence base for further policy development. Sample boosts took place for deprivation factors and for ethnicity. However, the data are for one cohort starting in 2003/04, so are of limited use at the moment in terms of examining labour market progression.

Yet other panel data sets are either too restricted in their age coverage (e.g. the Youth Cohort Study) or too limited in their sample size (e.g. the British Household Panel Survey) to be of much use in answering the questions of interest involved in measuring the effectiveness of VET systems in terms of supporting progression to HE.

Another alternative is to link administrative data sets and so follow individuals through different phases of their learning career. This was the approach adopted by the project 'Degrees of Success: the transition between VET and HE' funded by the Economic and Social Research Council (ESRC). In this article, we reflect on the challenges of using the administrative data sets employed by this project and report on some of the main findings to illustrate the potential utility of such administrative data sets in exploring the effectiveness of VET programmes as preparation for HE progression.

Background to the Project

The project 'Degrees of Success: the transition between VET and HE' analysed the transitional landscape of students progressing to HE from different educational backgrounds, with a focus on those holding vocational qualifications. The project described the distributional patterns and sought to explain them by looking at different factors and characteristics of students and institutions.[2] The empirical basis was threefold, consisting of: (a) an analysis of administrative large-scale databases; b) data from two surveys administered by the project to students in five HEIs in three subject areas; and c) interview data with admissions staff, lecturers and students.[3]

This article focuses on the methodological challenges of using administrative data sets. Such data are typically computerised records that are gathered for some administrative purpose, but contain information that can be used for other purposes as well. A classic example is birth records. These data are maintained as a matter of public record and have long been computerised to facilitate their use. In the case of education, administrative data are typically collected for funding and accountability purposes.

The main strength of administrative data is that no new data collection is required by the research team as another body has already gathered and entered the data, and they are apparently therefore lying there waiting to be analysed. Available databases are often relatively large – in the case of the Degrees of Success project, consisting of in excess of half a million records. That usually means it is possible to analyse population subgroups separately, and administrative data can often be useful in identifying problems specific to a particular subgroup or geographical area.

However, administrative data are not without serious problems and limitations. First, the data can be 'dirty', and careful scrutiny of all fields is needed to identify aberrant entries. Second, administrative data are often incomplete, and there is a consequential issue about how to deal with missing data. Third, it is critical to remember that the data were gathered for another purpose - for example, in the case of the data sets employed in the Degrees of Success study, they were gathered to collect information about course attendance for funding purposes or for managing entry to

HE. The fields of greatest interest - for example, the qualifications held by applicants to HE - may or may not be central to the primary record-keeping or payment purpose.

The Data Sets

No single data set yet exists within the UK that makes possible the tracking of students over different stages of a learning career from completion of compulsory secondary education to completion of an undergraduate degree across all qualification types. Therefore one has to look for different data sets and try to combine them in a meaningful way. The Degrees of Success project aimed to combine three different data sources: (a) the post-compulsory, but pre-HE, data included in the Individual Learner Record (ILR) from the Learning and Skills Council (LSC) (subsequently the Young People's Learning Agency [YPLA]) and the Pupil Level Annual School Census (PLASC) from the Department for Education (DfE); (b) the applicant/admission data from UCAS, the UK's HE clearing house; and (c) Higher Education Statistical Agency (HESA) data on students in HE. All three administrative data sets hold the full population of the specific groups, which should permit drawing quite reliable transition landscapes between different educational stages and pathways.

School and College Data Sets

However, the PLASC data set holds limited data on students enrolled in post-16 VET provision in schools. The restricted fields available at the time of the project were partly a function of the need to minimise the administrative burden on school managers when completing a return for each student on each field. New approaches to managing this data set in the future may yield better functionality for answering research questions about transitions from VET provision. Further, at the time of the project, plans were only in the process of being formulated to produce a unique pupil number that would enable an individual to be tracked across all data sets. Without such an identifier, matching of cases in the PLASC data set to other data sets needs to be done by fuzzy matching processes, which was beyond the scope of this project.

The Individual Learner Record comprises two databases. One consists of the qualifications a learner is taking, and the second is a record of learners. The purpose of the data is to release funding to further education, sixth form, and tertiary colleges and a range of private training providers for the programmes they deliver. Potentially this is a useful data set, but it only covers a limited portion of the post-16 population – it excludes school students. Consequently, using this data to track the progress of students confounds two factors - the qualification taken, and the centre where learning took place, with no means of controlling for the absence of school-based students. Further, the data sets require considerable reworking to make them suitable for the type of analysis envisaged for the Degrees of Success project. Finally, matching of data from the ILR to UCAS and HESA data sets would again require fuzzy matching processes.[4]

UCAS Data

UCAS organises the applications of all applicants for full-time undergraduate study at HEIs in the United Kingdom.[5] The huge administrative data sets this process generates provide information about the gender, age, socio-economic status, ethnic background, and potential disabilities of applicants, as well as a restricted postcode of the applicants' home address, their UCAS tariff score (a numerical measure of attainment, but one which is only available for some qualification pathways), their choices (which subject they applied for, and at which type of institution), and, most importantly for this project, the applicants' prior qualifications and whether the applicant was accepted or not by an HEI. We compared data from a ten-year time span (1995-2004), so we were able to track trends over time.

The UCAS data are collected for administrative rather than research purposes. Therefore, although the data are a rich source of information, they are limited in different ways.[6] First, they only hold information about the application process. Although this includes some information about the applicants' prior qualifications, the data are not very detailed. Furthermore, although one knows whether they are accepted by a university or not, it is not clear from these data alone

(a) why they are (not) accepted [7], and (b) how successful they are later in their studies. Thus, nothing about individual educational pathways or learning careers over a longer period can be drawn on the basis of these data alone. Second, the data are, in the main, only for full-time HE students.[8] Although full-time students make up the majority of all students in UK HE, this might not be true for the special subgroup of students with a VET background.

An additional problem with time comparisons is the stability of particular fields in the data sets. Most of the variables in the data set stayed the same over the entire ten years; however, changes occurred in the coding of socio-economic status, a crucial variable, although UCAS provides information about how to transform the old codes into the new ones.[9] Changes were also made in the coding of prior qualifications, due to changes in the qualifications pupils could achieve. As far as one can see, these changes produce only minor changes over the time period. Another problem is the classification of qualifications. The data set we used is not perfect for providing full information about prior achievements.[10] Further limitations of the data are discussed when each of the variables are described below.

These limitations notwithstanding, the results produced below have a high reliability for the largest group of students - those under 21 entering full-time HE in England. One reason for this is that we have data for the whole population of applicants, not merely a sample.[11] The following analyses take only UK (-domiciled) applicants into account, although some cases have been dropped, as they had no information about choices (less than 1%). Table I gives an overview of the data sets for the three years 1995, 2003/04 and 2004/05.

Year	No. of cases	Women %	Age (mean, standard deviation)	SES[a] (mean, standard deviation)	Accepted %
1995	369,701	51.4	21.0 (5.88)	3.0 (1.89)	71.8
2003/04	406,165	54.3	20.6 (5.71)	3.1 (1.90)	82.2
2004/05	409,526	55.2	20.6 (5.62)	3.1 (1.90)	81.6

[a]Socio-economic status is classified in seven categories (1 = High, 7 = Low) and is not a metric variable. Nevertheless, mean and standard deviation (in brackets) are shown to give a first impression of the distribution.

Table I. Characteristics of UK higher education applicants: Overview.

HESA Data

HESA holds data on all individuals who have enrolled at an HEI, as well as on the type of HEI they are attending and details about completion/non-completion rates. There are fields for qualifications held but these are often only partially completed, and there are some socio-economic and demographic data. Importantly, the HESA data have a student identifier that can be mapped back to the identifier given to the student when they applied through UCAS, enabling the matching of the data sets. Thus, the more detailed background information provided by the UCAS data can be matched to data about progression into and success within HE.

We now present some results produced from analysing these data to illustrate their utility, before turning to some concluding remarks.

Applicants and their Characteristics by Educational Pathways

The UCAS data hold reasonably detailed information about prior qualifications. On the basis of this information, five educational backgrounds for HE applicants could be differentiated [12]:

- General academic;
- Vocational;
- Foundation and Access courses (FaA);
- Other;
- No qualifications.

Table II shows the changes in proportions applying for HE from the different qualification groups. There are distinct changes over time. The number of applicants holding a general academic

qualification, the largest group in UK HE, has increased by around 5% over the last ten years, but an even larger increase can be found for applicants holding vocational qualifications, rising from around 18% to 25% over the same time span. However, the biggest increase can be found for 'other' qualifications.[13] Interestingly, the share of people coming from a Foundation/Access (FaA) background is stable over time, despite the introduction of the new Foundation qualification during the period analysed.[14] The number of applicants with no reported qualification has decreased by half.

Type of qualification	1995	2003	2004
General academic	70.7	75.3	75.7
Vocational	17.8	25.8	24.9
Foundation/Access	7.8	8.6	8.9
Other	5.1	13.5	14.1
No qualification	6.0	3.5	3.2
Total[a]	107.3	126.7	126.8

[a]more than 100%, as applicants can hold multiple qualifications.

Table II. Type of qualifications held by applicants to UK higher education (%).

However, the increase in the share of applicants holding a certain qualification does not tell the whole story. The figures for each year in Table II add up to more than 100% because people can hold qualifications from more than one group. In 1995 only a small minority (7%) held different types of qualifications, but this had increased to 27% by 2003/04. The increase of around 12% in the overall number of applicants (see Table I) was accompanied by an increase in the diversity of the qualification types held.

Applicants with different combinations of qualifications were classified into seven 'educational pathways':

- only general academic ('Academic');
- only vocational qualifications ('Vocational');
- only Foundation and Access courses ('FaA');
- only other ('Other');
- academic and vocational qualifications ('Ac + Voc');
- academic and FaA ('Ac + FaA'); and
- any other combination ('Other Combination').

Table III shows the proportions of applicants in each of these pathways, now adding up to 100%.

Qualification pathway	1995	2003	2004
Academic	63.4	50.8	51.3
Vocational	13.6	10.1	9.5
Foundation and Access Courses	6.2	3.9	4.2
Other	3.5	5.9	6.0
Academic and Vocational	4.2	14.1	13.7
Academic and Foundation and Access Courses	1.5	3.2	3.1
Other combination	1.6	8.5	8.9
No qualification	6.0	3.5	3.2
Total	100.0	100.0	100.0

Table III. The proportion (%) of students progressing to higher education in the UK through different qualification pathways.

The growing diversity of qualification types applicants hold is again clearly visible. The proportion in the first three groups in Table III decreased over time, while the proportion in the category 'Other' and the pathways followed by those taking combinations of different qualifications

increased. Nevertheless, the 'Academic' group is still the biggest, accounting for more than half of all applicants.

The increase in vocational qualifications (Table II) is mostly a result of an increase in their combination with academic qualifications. While the 'Vocational' group was much larger than the combined 'Ac + Voc' group in 1995, it decreased over time, and in 2003/04 the combined group is the larger group, having more than tripled its share of all applicants.[15]

In summary, the majority of those applying to HE (70.7% in 1995 and 75.7% in 2004) held general academic qualifications, but the proportion applying with vocational qualifications increased from 18% (1995) to 25% (2004). However, this growth was due to an increase in those combining vocational and general academic qualifications, up from 4% to 14% over this time period, while the proportion with vocational qualifications only decreased from 14% to 10%.

Do VET Pathways Widen Higher Education Participation?

This section presents results derived form the combined HESA-UCAS data sets. Unsurprisingly, those with 'Academic' qualifications mainly come from state school sixth forms, sixth form colleges, independent schools and grammar schools. Most applicants from the 'Vocational' pathways obtained their most recent qualification within the further education sector. Applicants to UK HE with a vocational background come from lower socio-economic groups, are more often male and older, are more likely to come from a non-white ethnic background, and are more often disabled, compared with those from the traditional academic route. Consequently, enabling learners from pre-HE VET routes to enter UK HE should widen participation in HE. However, a high proportion of those with combined academic and vocational qualifications gained their qualifications in state-school sixth forms, and they are more similar to the traditional general academic applicants from such schools. Consequently, the increasing proportion of students applying to HE with a combination of academic and vocational qualifications may be leading to some increase in widening participation, but this effect is not substantial.

Furthermore, those with vocational backgrounds are much more likely to apply for courses in post-92 HEIs (mainly the former polytechnics and HE colleges) [16] and further education (FE) colleges offering HE provision than those with purely academic backgrounds. Those who combine academic and vocational qualifications are less likely to apply to at least one pre-92 institution than those with just general academic qualifications, but more likely to do so than those with just vocational qualifications. However, the application rate to post-92 HEIs is similar for those with just vocational qualifications and those combining vocational and academic qualifications.

Students with a vocational background are therefore over-represented in 'less prestigious' HEIs, and this has potentially important consequences in terms of social mobility and future progression into the labour market. However, VET students are also, unsurprisingly, more likely to take degrees in applied fields such as creative arts and design and computer science and less likely to study subjects such as medicine, dentistry, law, languages, history and philosophy than those who have taken just general academic qualifications. It could just be that the more applied subjects chosen by VET students are more likely to be offered in post-1992 institutions, and this would account for the distribution of VET students across HEIs.

Controlling for the distribution of subjects across institutions, it becomes clear, however, that students with an academic background were much more likely to study at pre-1992 universities than their counterparts with vocational qualifications. Their over-representation was especially visible in those subjects in which they were normally under-represented. *If* someone with an academic qualification was studying 'agriculture', 'computer science' or 'creative arts' at all, *then* they were studying at a pre-1992 institution.

Controlling for socio-economic background, other demographic factors, and the type of educational institution attended prior to entry to HE, compared with traditional A-level students, those with VET qualifications have a much higher risk of not obtaining a place in HE and of dropping out after their first year (Table IV). But the picture is much more favourable for those combining academic and vocational qualifications: they are nearly as successful at entering and completing the first year of HE as those with only general academic qualifications.

The cross-tabulation shown in Table IV clearly demonstrates that having a VET background has a negative impact on the likelihood of continuing to study after one year in HE. Indeed, on this metric, the VET students seem to be doing worse than students entering HE in England via any other pathway. However, this result can be influenced by different factors. One is certainly prior attainment – those with higher prior attainment are more likely to stay on after one year (Smith & Naylor, 2001). Second, the HEI at which the student has chosen to study could influence the risk of dropping out. For example, the HEI may have a highly competitive atmosphere, or it might offer poor teaching and support for students. Third, it may be that the subjects chosen by VET students may have particularly high drop-out rates.

Progression route (*n*)	Staying on %	Dropping out %
Academic (150786)	92.9	7.1
Academic and Vocational (23235)	92.1	7.9
Vocational (15851)	86.4	13.6
Foundation and Access Courses (5716)	91.6	8.4
Other (8979)	87.4	12.6
Total (204567)	92.1	7.9

Table IV. Staying-on rates after one year of study by pathway - English students aged under 21.

Further, the influence of these factors could also be different for students with different educational backgrounds. For example, it might be the case that vocational qualifications prepare students well for studying some subjects (resulting in a lower drop-out rate) but not for others (resulting in a higher drop-out rate). Or it may be that some HEIs offer teaching styles or other forms of support that are better adapted to meet the needs of students with 'non-academic' qualifications, thereby leading to lower drop-out rates in these institutions. To answer such questions, a multilevel modelling approach is needed.

Multilevel Modelling

Different approaches to examining the influence of institutions and subjects are possible using the data sets we have available for this analysis. One could take into account either the students nested in institutions, and the subjects they are studying as additional explanatory variables; or the students nested in the subjects, and the institutions as an additional explanatory variable. The analysis that follows adopts the first perspective - students nested in institutions - as this has proven to be the more influential factor.

Table V presents a random intercept model for the data employing logistic regression, as the dependent variable is binary (0 = stay on at the end of year 1, 1 = drop out), with second-order predictive quasi-likelihood approximation (PQL). Model 1 represents a baseline. The overall mean of the beta, -2.6, is highly significant, with the predicted probability of being in the non-continuation group being 0.074. Model 2 in Table V adds some basic variables to the baseline model: standardized UCAS tariff scores and different educational backgrounds. The higher the tariff scores for the prior attainment of a student, the lower their probability of dropping out: an increase of one standard deviation (=128 tariff points; Mean = 318) lowers the probability of dropping out by nearly 30%.

Coming through a non-traditional educational pathway (vocational education and training = VET; academic and vocational = AcaVet) has a very small and insignificant impact compared with those progressing through the academic pathway. The dummy variables for the two other educational pathways listed in Table IV - 'Foundation and Access courses' and 'other educational pathways' - are non-significant throughout the following analyses, with tiny beta values; consequently, they are excluded from further analyses.

In model 3, other socio-demographic variables are added. Higher socio-economic background (SEC), black or Asian ethnic background, and disability all increase the probability of staying on. Coming through clearing (a process whereby students who have not met the terms of the offer,

and who may therefore not be able to go to the HEI of their choice, are matched to one with spare places) and coming with a vocational prior qualification both increase the probability of drop-out. Now the effect of a VET background on drop-out, though still quite small, is significant.

	Model 1	Model 2	Model 3
Fixed effects			
Intercept	-2.60*	-2.71*	-2.59*
Tariff		-0.43*	-0.43*
VET		0.03	0.10*
AcaVet		0.02	0.04
SEC			-0.16*
Black			-0.59*
Asian			-0.46*
Other			-0.06*
Disabled			-0.19*
Clearing			0.39
Gender			-0.02*
Random effects			
Intercept	0.35*	0.22*	0.25*

Table V. Logistic regression models, dependent variable = staying on (0), dropping out (1), after one year of study. * = significant at $p = .05$.

One final logistic regression model, not shown here, included the 19 different subject fields. With the exception of veterinary science, students taking all subjects, when compared with those taking medicine and dentistry, have an increased risk of dropping out. The impact of the other variables remains similar to those in model 3. Controlling for a range of other variables suggests, therefore, that progressing into English HEIs with a VET background increases the risk of dropping out.

The models described so far have only included a fixed effect for the VET-background students. This means we assumed that the effect of coming through this pathway is the same in all institutions. However, it might be that some institutions are more responsive to VET students' needs, as suggested by Tinto's influential interactionalist theory of student retention in HE (Tinto, 1975, 2006, 2010). Table VI reports the results for a random slope model for VET qualifications. By freeing the VET parameter to have random effects, we allow this parameter to have a different influence across different HEIs. Compared with the fixed-effects model reported in Table V, it is clear that now the main fixed effect for vocational qualifications becomes insignificant.

	Model 5
Fixed Effects	
Intercept	-3.65*
Tariff	-0.42*
VET	-0.02
AcaVet	0.05
SEC	-0.16*
Black	-0.57*
Asian	-0.43*
Disabled	-0.20*
Clearing	0.38*
Random effects	
Intercept	0.24*
VET qualification	0.17*
Cov (HEI, VET)	0.17*

Table VI. Multilevel model, students nested in institutions, dependent variable = staying on (0), dropping out (1), after one year of study.
* = significant at $p = .05$.

However, the variance of the random effect is significant (Wald test: Joint c^2 (2 *df*) = 16.52, $p = .0003$). This suggests that the influence of VET background is significantly different across

institutions. The positive covariance of the random intercept and the random coefficient for VET background means that in institutions with higher dropping-out rates, VET students have an increased risk of dropping out. A further result, though not significant, is of interest: HEIs with a higher proportion of VET students seem to have a reduced risk of VET students dropping out. These results suggest that institutions that are better at retaining students in general reduce the risk of VET students dropping out and that effect is increased if such an HEI has a high proportion of VET students; such institutions may have additional safety nets that support VET students in particular. This suggests that the drop-out rate for VET students could be reduced if HEIs take appropriate supportive measures, thereby raising retention rates for VET students.

Concluding Remarks

The purpose of the article was threefold. The first part introduced two of the data sources of the 'Degrees of Success: the transition between VET and HE' project - the UCAS applicant and HESA data - and made general remarks about these administrative data sets, their coverage and resulting limitations. Some problems were discussed, but the point was made that since the combined data contain nearly all full-time applicants for a place in HE, they are, nevertheless, a highly valuable data source.

The main aim of the project this article is reporting on was to investigate the transition landscapes and the success of students in HE who have a VET background. The second part of the article therefore gave a detailed description of key characteristics of the applicants progressing to HE along different educational pathways. The description took into account three domains of variables: socio-demographic characteristics; information about prior educational experiences; and some data about the application process. Looking at these characteristics, it becomes obvious that vocational routes open access to HE for non-traditional students. Applicants with a vocational background are from lower socio-economic classes, are more often male and from a non-white ethnic background, and are more often disabled, than those from the traditional general academic route.

The third and fourth parts of the article provided some results from an analysis of the combined UCAS-HESA data set. In particular, they demonstrated that students aiming to enter full-time UK HE from a purely vocational background were at a significant disadvantage compared with their counterparts following an academic pathway. Furthermore, they tended to attend less prestigious HEIs and were more at risk of dropping out during the first year of HE study. However, a multilevel modelling approach does suggest that HEIs can take actions that may reduce the risk of VET students dropping out and thereby widen participation.

Clearly, then, these large-scale administrative data sets do have utility for studying substantive questions of interest around the progression of VET students into HE. However, the challenges involved in working with these types of administrative data in terms of answering substantive research questions of interest should not be under-estimated. Upon reflection, three crucial questions need to be considered when working with such data:

1. Who 'owns' data collected using taxpayers' money, and who can legitimately use this information to answer research questions? The data collected by all the agencies involved are collected, to some extent, using public funding, yet the data are not readily publicly available. They can be requested for research purposes, but this involves paying not inconsiderable sums of money for the data to be extracted. Further data have to be requested, which means having an exhaustive knowledge of the variables and fields held in different data sets in order to specify accurately the data needed to answer a particular research question. Ongoing iterative negotiations were needed to ensure the best-quality data were obtained. There are good reasons for maintaining confidentiality around such data - for example, identifying individuals and institutions could lead to personal or organizational damage if the data are improperly used. However, other data sets are made available to researchers at no cost, with appropriate checks and balances to maintain confidentiality. If wider use is to be made of these potentially valuable administrative data sets for independently assessing the performance of the VET system, then, arguably, these data sets need to

be made publicly available at minimum, if any, cost. This suggests that a data warehouse should be developed with data being made available to researchers on request.

2. Who is missing from such administrative data sets, and why? It is clear that whole groups of applicants and HE students were missing from one or more of the data sets and that different information was being collected about the same learners because of the different uses that data were being put to administratively. Of particular importance for judging the effectiveness of VET provision for supporting progression to HE is the absence of part-time students from the UCAS data set. This reflects a historical tradition of direct entry for such students to an HEI without passing through the UCAS system, which is geared to support, in the main, full-time students. This means that important progression routes - for example, from apprenticeship programmes to part-time Foundation degrees - are all but invisible in current administrative data, even though such progression is currently a major policy concern. Clearly, steps will need to be taken to develop the administrative data sets to facilitate the tracking of learners through a highly heterogeneous set of pre-HE pathways.

3. What are the incentives and disincentives for organizations to provide the full returns needed to maintain comprehensive administrative databases that can be used for research purposes? An abiding source of frustration in working with these administrative data sets is the preponderance of missing data. This suggests that certain fields – notably those concerning qualifications held and background information about learners - are routinely not completed or only partially completed. In part, this is likely to be the result of the time needed to collect such information from students spread across a university campus. If the information is not essential to trigger the release of monies from a funding body, then the temptation must be to not complete that field but focus attention on the commercially more crucial ones. This is entirely rational for an administrator but completely frustrating for the researcher. Given the urgent policy need to evaluate the effectiveness and efficiency of VET provision, and given the need to utilise administrative data sets to do so, additional thought needs to be given as to how administrators can be incentivised to provide full returns.

A second reason for poor data around qualifications held is that it is difficult for bodies such as HESA and UCAS to keep their data collection systems up to date, with a rapidly evolving qualification system. This problem will be exacerbated by the fast-track procedures for registering new qualifications in the various UK Qualification and Credit Frameworks (QCF), by the intention to make such qualifications highly flexible in terms of their structure and composition in order to meet 'employer needs', and by the potential incorporation of company training programmes as qualifications to generate the release of government funding for such training within the English QCF. Such activity potentially further reduces the utility of vocational qualifications for supporting progression to HE, not least because little will be known about them by HE admissions staff and no data will be collected about their use in relation to HE progression. They will, in effect, be invisible, and so reduce the potential for those studying on VET programmes to progress into HE. The use of appropriately constructed administrative data sets would enable the detailed examination of the impact of such issues on the UK's HE widening participation agenda.

Finally, while students from a VET background contribute to a widening in access, the question remains as to whether their distribution across HEIs represents fair access for them. Of particular concern is the extent to which students from VET backgrounds find themselves primarily in less well-resourced HEIs. In addition, studying at a pre-1992 HEI brings extra economic benefits: students with an academic background have an advantage on the labour market. Clearly, as boundary objects (Star & Griesemer, 1989), vocational qualifications in the UK signify different attributes than academic qualifications do, confirm status hierarchies, and are interpreted, rightly or wrongly, as preparing students less well for HE study. HE remains a highly segmented market from the perspective of both VET applicants and HE admissions staff and tutors.

Notes

[1] See https://ilsype.education.gov.uk/workspaces/public/wiki/LSYPE for details.

[2] See Dunbar-Goddet & Ertl (2007) for more details about the project design and research questions.

[3] See Dunbar-Goddet & Ertl (2008) for an introduction and results of the first survey, and Hoelscher et al (2008) for some results from the quantitative analyses.

[4] Subsequent to the project, the PLASC and ILR data have been merged with HESA data.

[5] There are some minor additional routes for direct entry, but the numbers are very small.

[6] See Gorard et al (2006, Appendix A) for a more detailed analysis of problems with current analyses in HE.

[7] See Wilde & Hoelscher (2007) for a more detailed analysis.

[8] Even a small minority of full-time students can bypass UCAS through direct applications for HE courses or through internal progression from further education (FE) level to HE level at FE colleges. Additionally, not all HEIs are members of UCAS. HESA data for 2004 show that 59% of all first-year students are full-time students, another 6% are on sandwich courses, and 36 % are on part-time study. For students studying for a degree, the figure for part-time study reduces to 13%, and for full-time it rises to 78% (National Statistics reports a figure of 90%; see http://www.statistics.gov.uk/cci/nugget_print.asp?ID=9). It appears that the figures have been quite stable since the 1980s. A comparison over time by the Department for Education and Skills (DfES) shows that in 1970/71, 26.1% of the students in HE were part time, 34.3% were in 1980/1, and 33.9% were in 1990/1. The figures are not directly comparable to current data, as they changed from headcounts to enrolments (see http://www.statistics.gov.uk/STATBASE/xsdataset.asp?vlnk=189 for further information).

[9] The provision of socio-economic status (SES) information by the applicant is voluntary and therefore not that reliable for comparisons.

[10] Other problems with qualifications in surveys are discussed in Jenkins & Sabates (2007).

[11] As we use the whole population, traditional significance tests are not applicable, and the respective figures are not given below. There is, however, an ongoing debate as to whether one can nevertheless view a whole population as a sample – for example, the population of one year of applicants as a sample for applicants in different years (see Behnke, 2005). This depends on the research questions one poses. We will address this problem in another methodological paper.

[12] The full list of qualifications and the coding of the five educational backgrounds are given in Appendix B of Hoelscher et al (2008).

[13] The UK and overseas degrees account for just over 20% of all 'Other' qualifications in 2003 and 2004.

[14] Around 44 % of these applicants have completed the newly introduced Foundation courses (not available in 1995). Foundation courses are one-year university preparation courses suitable for mature students who may not have formal qualifications, and also for students of any age without the entry qualifications for specific degree programmes, especially overseas students who have studied a non-British curriculum.

[15] However, the changes might be, at least in part, a result of changes in the coding of data by UCAS for the different years.

[16] The 1992 Further and Higher Education Act ended the binary divide between universities and polytechnics. The former group are known as pre-1992 HEIs; the polytechnics and those more recently obtaining degree-awarding powers, such as colleges of higher education (mainly set up as teacher training institutions), are termed post-1992 HEIs.

References

Behnke, J. (2005) Lassen sich Signifikanztests auf Vollerhebungen anwenden? Einige essayistische Anmerkungen, *Politische Vierteljahresschrift*, 46(1), O1-O15.

Blair, T. (2007) *Our Nation's Future – the role of work*. http://www.number10.gov.uk/output/Page11405.asp

Bratti, M., Naylor, R. & Smith, J. (2006) Different Returns to Different Degrees? Evidence from the British Cohort Study 1970, University of Warwick, December 2006. http://www2.warwick.ac.uk/fac/soc/economics/staff/academic/naylor/publications/bcs70.pdf

Chevalier, A. & Conlon, G. (2003) Does it Pay to Attend a Prestigious University? Discussion Paper No. 33. London: Centre for the Economics of Education London School of Economics.

Connor, H., Sinclair, E. & Banerji, N. (2006) *Progressing to Higher Education: vocational qualifications and admissions*. Ormskirk: Action on Access, together with Aimhigher.

Department for Education (DfE) (2011) Review of Vocational Education – the Wolf Report. https://www.education.gov.uk/publications/standard/publicationDetail/Page1/DFE-00031-2011

Dunbar-Goddet, H. & Ertl, H. (2007) Degrees of Success: research on the transition from vocational education and training to higher education. Oxford: Department of Education, University of Oxford.

Dunbar-Goddet, H., & Ertl, H. (2008) Investigating the Learning Experience of HE Students with a Vocational Pathway: findings from a questionnaire survey. Oxford: Department of Education, University of Oxford.

Education and Employment Committee (2001) *Higher Education: access*. Fourth Report. London: The Stationery Office.

Gorard, S., Smith, E., May, H., Thomas, L., Adnett, N. & Slack, K. (2006) *Review of Widening Participation Research: addressing the barriers to participation in higher education*. A report to HEFCE by the University of York, Higher Education Academy and Institute for Access Studies (Review). York: University of York, Higher Education Academy and Institute for Access Studies.

Hayward, G. (2006) *Participation, Progression and Success in Vocational Learning: a quantitative analysis of system performance*. London: Learning & Skills Research Centre.

Hayward, G., Hodgson, A., Johnson, J., et al (2005) *The Nuffield Review of 14-19 Education & Training. Annual report 2004-05*. Oxford: Department of Educational Studies, University of Oxford.

Hayward, G., Hodgson, A., Johnson, J., et al (2006) *The Nuffield Review of 14-19 Education & Training. Annual report 2005-06*. Oxford: Department of Educational Studies, University of Oxford.

Hayward, G., Hodgson, A., Keep, E., et al (2004) *Nuffield 14-19 review annual report*. Oxford: Department of Educational Studies, University of Oxford.

Hayward, G. & Williams, R. (2011) Joining the Big Society: am I bothered? *London Review of Education*, 9(2). 175-189. http://dx.doi.org/10.1080/14748460.2011.585877

Hoelscher, M., Hayward, G., Ertl, H. & Dunbar-Goddet, H. (2008) The Transition from Vocational Education and Training to Higher Education: a successful pathway? *Research Papers in Education*, 23(2), 139-151. http://dx.doi.org/10.1080/02671520802048679

Jenkins, A. & Sabates, R. (2007) The Classification of Qualifications in Social Surveys. Working paper 2007/2. London: Institute of Education.

Powdthavee, N. & Vignoles, A. (2007) Succeeding in Higher Education: a widening participation issue. http://www.powdthavee.co.uk/resources/dropout+paper+27+March.pdf

Pring, R., Hayward, G., Hodgson, A., et al (2009) *Education for All: the future of education and training for 14-19 year olds*. London: Routledge.

Pugsley, L. (2004) *The University Challenge: higher education markets and social stratification*. Aldershot: Ashgate.

Sloane, P.J. & O'Leary, N.C. (2004) The Return to a University Education in Great Britain. Forschungsinstitut zur Zukunft der Arbeit (IZA) [Institute for the Study of Labor] Discussion Paper No. 1199. Bonn: IZA.

Smith, J.P. & Naylor, R.A. (2001) Dropping out of University: a statistical analysis of the probability of withdrawal for UK university students, *Journal of the Royal Statistical Society A*, 164, part 2, 89-405.

Star S.L. & Griesemer J.R. (1989) Institutional Ecology, 'Translations' and Boundary Objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39, *Social Studies of Science*, 19(3), 387-420. http://dx.doi.org/10.1177/030631289019003001

Tinto, V. (1975) Dropouts from Higher Education: a theoretical synthesis of recent research, *Review of Educational Research*, 45, 89-125.

Tinto, V. (2006) Research and Practice of Student Retention: what next? *Journal of College Student Retention*, 8(1), 1-19.

Tinto, V. (2010) From Theory to Action: exploring the institutional conditions for student retention, in J.C. Smart (Ed.) *Higher Education: handbook of theory and research 25*, pp. 51-89. Berlin: Springer. http://dx.doi.org/10.2190/4YNU-4TMB-22DJ-AN4W

Vickers, P. & Bekhradnia, B. (2007) Vocational A Levels and University Entry: is there parity of esteem? Higher Education Policy Institute working paper. Oxford: Higher Education Policy Institute.

Walker, I. & Zhu, Y. (2005) The College Wage Premium, Overeducation and the Expansion of Higher Education in the UK. Forschungsinstitut zur Zukunft der Arbeit (IZA) [Institute for the Study of Labor] Discussion Paper 1627. ftp://repec.iza.org/RePEc/Discussionpaper/dp1627.pdf

Watson, D. (2006) New Labour and Higher Education, *Perspectives: Policy and Practice in Higher Education*, 10, 92-96. http://dx.doi.org/10.1080/13603100601002936

Wilde, S. & Hoelscher, M. (2007) *Missed Opportunities? Non-placed Applicants (NPA) in the UCAS Data.* Project completed for UCAS by Department of Education, University of Oxford. Oxford/Cheltenham: UCAS.

GEOFF HAYWARD is Head of School and Professor of Education at the University of Leeds. Prior to that he was Director of Research, Associate Director of SKOPE and Reader in Education at the Department of Education, University of Oxford. He was a visiting professor in the Department of Sociology and Economics at the Autonomous University of Barcelona and Economics and Business Studies at the University of Paderborn. Geoff has wide experience in leading and managing large research projects, including a systems analysis of VET provision for the Learning and Skills Development Agency, and the Nuffield 14-19 Review undertaking, in particular, analysis of systems performance and management. With Dr Hubert Ertl he co-directed the ESRC-funded Degrees of Success project looking at progression from VET to higher education in the UK. He helped devise the evaluation strategy for the new Diploma qualifications and provided advice to the then DfES on the 14-19 reform strategy. He is regularly invited to present at international conferences on VET system performance. He has worked for many different governmental and non-governmental organizations and bodies as a consultant/adviser and is a governor of Oxford and Cherwell Valley College. *Correspondence*: Professor Geoff Hayward, School of Education, University of, Leeds, Leeds LS2 9JT, United Kingdom (g.f.hayward@leeds.ac.uk).

MICHAEL HOELSCHER is a university lecturer in sociology at the University of Heidelberg and senior research fellow at the University of Oxford. He wrote his PhD thesis (Wirtschaftskulturen in der erweiterten EU [Economic Cultures in the Enlarged EU], 2006) at the Free University in Berlin, Germany. He worked as a researcher at the HoF Wittenberg and at the Department of Education, Oxford, where he was part of the 'Degrees of Success' project. His main fields of interest are higher education, cultural globalisation, European integration, and comparative quantitative methods. Main publications include *Cultural Policy and Governance in a New Metropolitan Age* (2012, Sage; as co-editor, with H. Anheier & Y.R. Isar); 'Transitions from VET to HE: a successful pathway?' in *Research Papers in Education* (2008) (with G. Hayward, H. Ertl, & H. Dunbar-Goddet); and 'Universities and Higher Learning', a chapter to be published in the *Encyclopedia of Global Studies*, ed. H. Anheier & M. Juergensmeyer (Sage, 2012). *Correspondence*: Dr Michael Hoelscher, Max Weber Institute of Sociology, University of Heidelberg, Bergheimer Straße 58, D-69115 Heidelberg, Germany (michael.hoelscher@soziologie.uni-heidelberg.de).