Guidance prepared by Chris Taylor, Gareth Rees, Ceryn Evans and Caroline Wright (2015), Wales Institute for Social and Economic Research, Data and Methods (WISERD), Cardiff University. With support from Higher Education Funding Council for Wales (HEFCW) and the Economic and Social Research Council (ESRC) (award no. ES/K004247/1).
Foreword

Widening Access: securing inclusion, progression and success

Universities in Wales have been commended for their innovative practices and acting as agents of economic and social change in widening access.¹

Universities in Wales have been commended for their innovative practices and acting as agents of economic and social change in widening access.

HEFCW and Welsh universities are committed to delivering effective widening access outcomes for Wales and continuing to make a significant contribution to supporting equality of opportunity, promoting higher education and student success.

To measure progress in widening access, we funded the Wales Institute of Social and Economic Research, Data & Methods (WISERD), to enable it to extend ESRC-funded research, to evaluate aspects of widening access policy and practice. I welcome this report and I should like to extend my thanks to Professors Gareth Rees and Chris Taylor and their colleagues at WISERD for this valuable work. I should also like to thank the universities, colleges and Reaching Wider Partnerships for their time and contributions to this research.

The WISERD evaluation report is based on school, college and HE linked datasets. We recognise widening access is an all-age agenda and more work should be done in relation to mature students’ entry, participation and success in higher education. However, similar datasets are not available for mature higher education entrants and are, therefore, outside the scope of the report. The WISERD report findings and recommendations will inform our policy developments and this evaluation guidance should be read in conjunction with the evaluation report.

In addition to the WISERD report, this practical evaluation guidance document, developed by WISERD for us, is provided for institutions and Reaching Wider Partnerships. The evaluation guidance should inform further the practical implementation of widening access evaluation. This evaluation guidance is applicable to all-age widening access provision and should be a key mechanism in supporting critical assessment of widening access planning and delivery.

Dr David Blaney
HEFCW Chief Executive

¹www.heacademy.ac.uk - assets - nations - wales - Review of the Widening Access and Reaching Wider Strategies in Wales
The purpose of this guidance is to provide advice to universities and widening access practitioners about how to evaluate their widening access strategies and practices. It is important to note that this can include a wide range of initiatives ranging from university policies (such as contextualized admissions policy), to departmental activities with school subject teachers, to short-term interventions with targeted groups (such as peer mentoring or summer universities). Just as no initiative is the same then no two evaluations are likely to be the same. Consequently it is very difficult to be prescriptive about how such initiatives should be evaluated. Instead practitioners should consider the key characteristics of good evaluation when undertaking their own evaluation.

This guidance is supplementary to the HEFCE evaluation toolkit for practitioners (3rd Edition 2014). However, this guidance focuses more on the evaluation design rather than how to conduct an evaluation. In this guidance we set out four steps practitioners should consider when preparing to evaluate widening access activities:

1. Outline a policy logic model
2. Decide the aims of the evaluation
3. Design the evaluation
4. The use of administrative data
The first stage to evaluating widening access initiatives is to develop a policy logic model for the intervention. This involves identifying the characteristics of the intervention. This should be undertaken retrospectively even if the original decision and design of the initiative was not based on a policy logic model. An illustrative policy logic model is presented below:

The purpose of a policy logic model is to describe in detail the processes in which a problem is identified, what intervention will be implemented to address this problem, and then what impact the intervention is expected to have. A policy logic model can be applied to whole institution strategies or for individual discrete activities. A comprehensive evaluation of many university-wide widening access activities could develop several interconnected policy logic models. Ideally a policy logic model is developed prior to designing an intervention. However, even if this is undertaken retrospectively it is valuable in guiding and shaping the evaluation of an intervention.

It is useful to develop a policy logic model in two stages. The first stage is to complete a policy logic model prior to undertaking any evaluation or research. This will then include the intended impacts of the intervention. Then following...
the evaluation the policy logic should be updated to reflect the research findings. It is important to consider at this latter stage whether the initial assumptions (e.g. about conditions or problems) were correctly identified, whether the inputs were appropriately implemented and which outcomes can or cannot be evidenced. In developing a policy logic model we recommend that the following questions are considered (see Taylor and Rees for further information):

1. **HOW IS WIDENING ACCESS UNDERSTOOD?**

Previous research would suggest that there are four sets of inequalities that widening access activities can address. Distinguishing between these (or identifying whether an intervention is designed to address multiple issues) is important in understanding better what the intended outcomes are likely to be.

- Educational prior achievement – Inequalities in educational achievement prior to entry to higher education
- Admission to higher education – Inequalities in admission to higher education (i.e. given levels of prior achievement)
- Higher education outcomes – Inequalities in higher education outcomes
- Knowledge of higher education – Inequalities in knowledge about higher education

2. **WHO IS WIDENING ACCESS FOR?**

The target group for any widening access activity is dependent upon the answer to the first question. The policy logic model should outline the rationale for identifying the target groups or individuals for the widening access activity. This is important in order to fully understand the likely impact of the activity. It is also useful to note here how the target groups or individuals will be identified in practice. This is important as it helps to say something about the probability that someone receives the intervention:

- Target groups or individuals are randomly selected
- Target groups are identified on the basis of their individual characteristics
- Target groups are identified in the basis of their membership of a particular group of individuals (such as by their area of residence or a school they attend),
- Participants are self-selecting
- Participants are referred or selected on an ad-hoc basis

3. **WIDENING ACCESS TO WHAT?**

The higher education sector is very complex and covers a wide range of subjects, courses and levels of qualifications. It is useful to know what the activity attempts to widen access to. This includes the following:

- The level of course/qualification entered at (e.g. HNC (NVQ Level 4) or First Degree (NVQ Level 6))
- The level of course/qualification expected to exit with (particularly for learners who will make incremental progress from one Level to the next)
- What mode of study is this for (e.g. part-time or full-time)
- What subject(s) is this for (e.g. university wide admissions or specific subject admissions)
- Which university(ies) is this for (e.g. from region, from Wales, UK-wide or specific university)
Following the development of a policy logic model the next step is to determine the aims of the evaluation. These should be determined on the basis of the content of the policy logic model, and in particular, whether the evaluation will evaluate some aspect of the intervention or attempt to evaluate the whole intervention. This might mean focusing on a particular set of inputs and activities rather than all the inputs and activities outlined in the policy logic model. Alternatively, it could mean focusing on particular outputs and outcomes rather than all anticipated outputs and outcomes. It is also useful at this stage to explicitly acknowledge what the evaluation will not be asking or addressing.

In practice, developing the aims of the evaluation is an iterative process as the evaluation is designed. But it is important that, if necessary, the aims of the evaluation (and what the evaluation does not aim to address) are amended to reflect the final evaluation design.

Generally, there are two kinds of evaluation:

1. Process evaluation – to understand better the implementation of an intervention

2. Outcome evaluation – to identify what impact the intervention has had
A process evaluation is largely concerned with evaluating the **rationale, inputs, activities and outputs** of an intervention. An outcome evaluation is largely concerned with evaluating the outcomes and broader impact of the intervention. The distinction between outputs and outcomes is an important one here. For example, a process evaluation may be able to say that 200 participants attended a widening access event (an output), but an outcome evaluation would want to look at whether those 200 participants experienced a change in their attitudes about going to university as a result of attending that event (an outcome). Process evaluations are also more likely to ask formative questions (i.e. to inform better practice or implementation) whereas outcome evaluations are more likely to ask summative questions (i.e. has something changed).

Although both kinds of evaluation can be done separately it is very useful to do both. The process evaluation will help understand why certain outcomes have been achieved (or otherwise), whereas the outcome evaluation is important to know whether certain processes actually have the desired outcome. In addition, there is often considerable overlap in a well-designed process evaluation and outcome evaluation. For example, the process evaluation could collect data that would be valuable for an outcome evaluation. Even if you are undertaking just a process evaluation, perhaps because the strategy or initiative is on-going, it is still useful to consider what the outcome evaluation might want to address and how this could benefit from how the process evaluation is designed. Finally, it is often assumed that a process evaluation is easier and less time-consuming to do than an outcome evaluation. This is not the case. As the initial policy logic model should demonstrate, there will often be multiple components to the implementation of an intervention – all of which need to be captured in a process evaluation. Conversely, the outcome evaluation is primarily concerned with whether there is a measurable change in some kind of outcome. Of course, just demonstrating a change in an outcome may also not be satisfactory, hence the need to consider doing both.
So, the third step is to design the evaluation. An evaluation design should be dependent upon the aims of the evaluation. However, as noted above, in practice determining the aims and the design of the evaluation is an iterative process. For example, the aims and design may have to be modified due to resource constraints or the availability of appropriate data. But throughout this process it is important that the policy logic model provides the basis for deciding these.

The HEFCE toolkit for evaluation design suggests using the RUFDATA framework for planning an evaluation. This includes the: Reasons and purposes, Uses, Foci, Data and evidence, Audience, Timing and Agency. These are useful questions. It also discusses different kinds of data and the ethics of conducting research. In particular, the diagram below (HEFCE Toolkit Figure 9, p.19) is helpful in considering the relationship between data sources and the ‘logic’ of an intervention:

<table>
<thead>
<tr>
<th>Input data</th>
<th>Participation data</th>
<th>Output data</th>
</tr>
</thead>
<tbody>
<tr>
<td>What leaners are like</td>
<td>What happened in the intervention</td>
<td>What happens next?</td>
</tr>
<tr>
<td>What schools do they come from?</td>
<td>What did they do?</td>
<td>Do they stay in school?</td>
</tr>
<tr>
<td>What areas do they come from?</td>
<td>How many times did they do it?</td>
<td>What grades do they achieve?</td>
</tr>
<tr>
<td>What are their predicted grades?</td>
<td>Who delivered it?</td>
<td>Do they apply to higher education?</td>
</tr>
<tr>
<td>What disadvantages do they have?</td>
<td>Did they value it?</td>
<td>Do they remain within / succeed in higher education?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Where do they work after leaving higher education?</td>
</tr>
</tbody>
</table>

Source: HEFCE 2014:19

This is enormously valuable when designing a process evaluation, when the focus is on capturing information about the implementation of an intervention. But the design of an outcome evaluation (and some process evaluations) must follow stricter requirements in order to generate valid and generalisable conclusions.
The key to an outcome evaluation is to **reduce bias** in the selection or targeting of the intervention group, whoever they may be (e.g. who attends a widening access event or who is eligible for a contextual admissions offer). This is important because an outcome evaluation must be able to demonstrate that the impact of an intervention is the result of that intervention and not other factors associated with the selection of participants to the intervention.

The ‘gold standard’ for removing bias in any intervention is to randomly select participants to an intervention – in other words the chances of being selected for the intervention is the same for everyone. In practice, however, this is often impractical. But without randomly selecting participants there will always be some form of bias in the selection process. There are numerous ways that bias can be introduced to an evaluation, but the main type of bias is:

- **confounding bias** – where selection into an intervention is based on certain characteristics that could also be independently associated with the outcome.

To mitigate the impact of bias it is important that an outcome evaluation collects and analyses comparative data, something to compare the results of the intervention with. Generally there are three strategies for this:

- **Pre- and post-test evaluation design** – this compares the outcomes of those exposed to an intervention prior to the intervention with the same outcome after the intervention.
- **Intervention group and control group evaluation design** – this compares outcomes of two groups of people, those that were exposed to the intervention in whatever form that came (the ‘intervention’ group), with a similar group of people who were not exposed to this intervention (the ‘control’ group).
- **Pre- and post-test of an intervention group and control group evaluation design** – the combines both of the strategies outlined above in to one evaluation design.
The first of these is the easiest strategy to implement and is often used to measure changes in attitudes prior to an intervention with attitudes after receiving an intervention. This does provide some comparative element that helps in the evaluation. However, it is still the case that we do not know from such a design whether changes in outcomes, of say attitudes, would have changed anyway over time. The only way of controlling for this is to consider the outcomes of a different group who did not receive the intervention. It is for this reason that a simple pre- and post-test evaluation design is considered to be non-experimental.

An evaluation design that compares the outcomes of an intervention group with the outcomes of a control group is useful then in order to be able to say something about what would have happened without the intervention. However, because of the potential selection bias of the intervention group it is best practice to measure change over time for both groups using a pre- and post-test. In this scenario the outcome of interest is no longer just the post-intervention outcome. Instead we are interested in the difference in outcomes before and after the intervention compared with differences in outcomes of a control group over the same time period. This type of evaluation design allows the evaluator to use the 'difference-in-difference' statistical techniques, helping to minimise further any selection bias of the participants in the intervention. This is particularly helpful when it is difficult to find a group of non-participants with similar characteristics (the 'control' group) as those receiving the intervention.
It is useful to note here that the selection of participants to receive an intervention could be on the basis of their individual characteristics or their membership of a group entity (such as the school they attend or the area they live in). It is often easier to identify a control group when the selection of the intervention group is based on the individual characteristics of the participants. However, some widening access initiatives are targeted at groups of individuals where it is their membership of a group that determines whether they receive the intervention or not. The selection of a control group can be made on the same basis. That is, if a school or area is selected to participate in an intervention then it is reasonable to select the control group on the same basis – i.e. another school or another area. However, when it comes to ensuring that the intervention group and the control group are similar this is best done on the basis of their individual characteristics.

It is also useful to note that the selection of a control group does not necessarily require the recruitment of additional individuals to the evaluation. The control or comparison could be made using existing data. The key here is to create a ‘virtual’ control group based on known characteristics, ideally as similar as possible to the participants in the intervention and for whom the outcomes will also be known.

Just as there can be bias in the selection of participants in the intervention there can also be bias within the control group. This is known as:

- contagion bias – where non-participants in an intervention indirectly benefit from the intervention or are benefitting from other interventions not being evaluated.

It is just as important when selecting a control group to compare with the group receiving the intervention that any potential contagion bias is also reduced. One way to help with this is to ensure that the control group come from a different area or setting. For example, if some students in a college are in receipt of the intervention it is sensible to identify a control group of other students from a different, but similar, college. Choosing students from within the same college as the ones receiving the intervention could lead to contagion bias (i.e. the social networks within the college could mean that all students in the college benefit to some degree from the intervention).

One of the main limitations of any evaluation is that an intervention could produce positive outcomes because of the additional resource or attention the intervention attracts, rather than the design of the intervention per se. Because of this it is useful to consider evaluating two or more interventions and comparing their impact. This is a different form of the intervention-control evaluation design, and is instead described as an intervention-intervention evaluation design. Comparing two interventions can often be more useful in evaluating the efficacy of different approaches to the same problem.

This has several benefits. First, this approach can be useful in considering the cost-benefits of an intervention, since the impacts of the two interventions can be compared against their relative costs. Secondly, this approach does not necessarily require having a control group (although including a control group will still have additional benefits).
As long as the pre- and post-outcome measures of both interventions are the same, and either the selection of participants is the same or can be controlled for statistically, then a direct comparison of the impact of both interventions is valid. However, it is important to note that the aim of this kind of evaluation is slightly different to that discussed above. This approach only can say whether intervention A has had more or less of an impact than intervention B. It cannot say whether this impact would have occurred with or without either intervention. To do this would still require the inclusion of a non-intervention control group.

It can be of particular importance when designing an evaluation of widening activities to include a longitudinal element, which involves conducting research over a period of time. This can be quantitative and/or qualitative. Since many widening access initiatives are not designed to have an immediate impact it may be necessary to consider the medium- to long-term processes and outcomes relating to them. This usually will require ‘tracking’ or ‘following’ individual participants over time, either directly (i.e. by interviewing or surveying them periodically over time) or indirectly (i.e. using existing administrative data to see what happens to them).

Finally, it should be acknowledged that applied evaluations can be ‘messy’ and will usually require the evaluators to make a number of compromises when designing an evaluation. Ultimately this is then about compromises to the kinds of questions that the evaluation wants to address. In such a situation it is useful to consider what the ‘least worst’ evaluation design would be, rather than starting with an “ideal” evaluation design and gradually reducing its appropriateness as and when compromises have to be made. Also as a result, it is not unusual for social research evaluations to provide only enough evidence to make partial or tentative conclusions. Critically a good evaluation recognises this and makes explicit the limitations of the evaluation when drawing conclusions from it.
The final step in this guidance is to consider the use of existing administrative data in the evaluation of widening access activities. All good quality evaluations often require considerable resource. Not only could an evaluation require the collection of data from an intervention group before, during and after the intervention, but it may also require the collection of data from a corresponding control group. This can often be difficult to justify when in all probability the cost of the evaluation has to come from the budget allocated to the intervention. As a rule of thumb, large-scale and relatively routine interventions should set aside 1% of the overall intervention budget for evaluation. However, more innovative and pilot interventions should set aside 10% of the intervention budget for evaluation. Ideally, the cost of evaluation should be considered when the intervention is being designed and implemented.

A significant way of reducing the cost of any evaluation is to use or improve the collection of routine widening access data, i.e. the data that is already being collected in the implementation of the initiative. For example, many widening access interventions already collect basic information from participants, such as their name, age, gender and ethnicity. It does not require much additional resource to ask a wider range of questions, such as what subjects they are studying, how likely it is that they will go to university, why they want to attend or participate in this intervention. The crucial thing here is to ensure that any routine information collected also asks participants for their permission to use this data for evaluation and research purposes.

The second main way of reducing the costs of evaluation is to use other administrative data collected for other purposes. This could be data collected by schools and/or colleges. Of most value to the evaluation of widening access initiatives is the use of the National Pupil Database (NPD), the Welsh Examination Database (WED), the Lifelong Learning Wales Record (LLWR) or Higher Education Statistics Agency (HESA) Student Records. Each of these datasets contains individual-level data on learners in different sectors of the education system. As recent research by WISERD has demonstrated, when combined together, by linking individuals across each of these datasets, they can be very powerful and enormously valuable in researching and evaluating widening access to higher education. In the coming years access to each individual dataset and linking individuals across each dataset will get increasingly easier with the adoption of the Unique Learner Number – a new method for tracking learners throughout their life as a student. Such datasets include basic information about an individual, such as their age, gender and ethnicity. They sometimes include useful socio-economic markers about an
individual, such as their eligibility for free school meals (up to the age of 15). They all contain important educational information; such as the courses and subjects a learner is enrolled on and measures of educational achievement.

A further enhancement to any evaluation that can be achieved through minimal cost is to be able to link routine data collected as part of the intervention with existing administrative datasets. Again, the adoption of the Unique Learner Number (ULN) in educational administrative data will make this increasingly easier. In order to be able to link individual data across to such educational data will just require the collection of the ULN. Alternatively the full name, gender and date of birth is usually sufficient to ensure the vast majority of individual data can be linked to other administrative datasets. However, it is very important that the appropriate informed consent is obtained from any participant who provides additional information to an intervention to link these records to records already held by a governmental department.

All requests to collect data for research purposes or to link to other administrative data must be approved by a regulated ethics committee. However, here is an example of explicit consent wording provided by the ESRC Administrative Data Liaison Service:

“In have read and understood the leaflet that provides further information on sharing my records collected by other organisations with the [Study name] for research purposes. I authorise the [Data controller name] or their equivalent to disclose to the organisation responsible for the [Study name] details they have routinely collected about me in the past, present and future relating to my [education and learning]. I authorise this information to be linked to the information held about me in the [Study name] for the purposes of their research studies into [main aim of research]. I allow research findings and datasets that contain information about me to be made publicly available, providing that they are anonymised and I cannot be identified. I understand that my consent will remain valid until it is withdrawn by me in the manner detailed on this form.”

In order to be able to link individual data across to such educational data will just require the collection of the ULN.
Generally it is accepted in the UK that individuals can make their own decision to consent from the age of 16 years. For any individual below 16 years of age it is necessary to seek consent from a parent, and good practice to also seek consent from the child (although your relevant ethics committee will provide further guidance on this).

When using administrative data for the purposes of evaluating widening access initiatives there are a number of known characteristics (or variables) that are important to include. These are:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Purpose</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Gender is known to be strongly associated with educational prior achievement, participation into higher education and progress through higher education</td>
<td>Readily available in all administrative data</td>
</tr>
<tr>
<td>Age</td>
<td>Age is usually a very important defining characteristic of the widening access target group</td>
<td>Date of birth is always useful to use since ‘age’ is dependent upon when the intervention or evaluation occurs. Knowing the date of birth of an individual means that age can always be derived when required. However, date of birth can be potentially revealing so should always be used with caution (e.g. year of birth should be used or an individual’s DoB should be kept separately from the rest of their data)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Ethnicity is strongly associated with educational prior achievement and access to higher education. However, this masks important variations between different ethnic groups. But in Wales the number of learners in minority ethnic groups can be relatively small which makes it difficult to generate reliable findings</td>
<td>Ethnicity is usually recorded in some way in all key administrative datasets. The more detailed the ethnic classification the better. However, the number of some ethnic groups can be very small, and could make an individual potentially identifiable. Given the ethnic composition of learners in wales is relatively small it may be appropriate to use a simple categorisation of White British and BME (Black and Minority Ethnic) group</td>
</tr>
<tr>
<td>Variable</td>
<td>Purpose</td>
<td>Data</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Socio-economic status (SES) (however measured) is strongly associated with educational prior achievement and access to higher education. In terms of progress at university our own analysis shows that SES is not associated with university outcomes after controlling for levels of prior attainment. SES is important in defining potential widening access target groups. The benefit of using the FSM indicator is that it is universally applied. Analysis that considers whether pupils have been eligible for free school meals over several years (i.e. two years or more) is a good predictor of the most disadvantaged children. However, recent analysis of the FSM indicator in Wales suggests that for every one pupil eligible for free school meals there are a further three pupils living in poverty. It is important to note that the majority of free school meals are taken up by children from single parent households. Because eligibility for free school meals is based on receiving a qualifying State benefit it is not a good indicator of in-work poverty or equivalised poverty (a measure of household income based on the number of people in the household). The other two measures of SES are more limited, either because of coverage (in the case of the SEC in HESA data) or because they only reflect the probability that a household has a certain SES (in the case of neighbourhood-level data). Despite the limitations of these three main sources of SES it is more important that some measure is included than no measure at all. But it is essential that the evaluator is fully aware of the limitations of their choice of measure and report findings accordingly. In future it is quite likely that the measure of SES from administrative sources will get better as a result of data linkage between different administrate datasets. The recently funded ESRC Administrative Data Research Network is embarking on a new research programme to identify and realise the potential of this. In the meantime primary data collection and secondary survey data provide the best means for obtaining detailed SES data. However, SES is a complex phenomenon and there is no straightforward way of capturing this. But crucially it is always useful to include at least three aspects of SES: income, occupational social class and highest educational qualification. Socio-economic status (SES) is not measured very well in administrative data. There are three main options: 1. Eligibility for free school meals (FSM) in the National Pupil Database — this is a proxy for SES and identifies pupils who have been registered to receive free school meals. Parents are able to claim free school meals if they receive a qualifying benefit from the Government (e.g. income support, Jobseekers Allowance, etc.). 2. National Statistics Socio-Economic Classification (NS-SEC) in HESA data (the SEC field in HESA Student Records). This data is obtained by UCAS from applicants who are invited (not compulsory) to provide the occupation of their parent/guardian who earns the most. This is then coded into one of seven socio-economic classifications: Higher managerial and professional; Lower managerial and professional; Intermediate; Small employers and own account workers; Lower supervisory and technical; Semi-routine; and Routine. Other outcomes include long-term unemployed or never worked and unknown classification. However, a significant number of applicants do not provide this information which means any analysis of such data has to be treated with caution (and the effect of any bias in response will vary depending on the population and outcome being studied). It is also important to note that this data is only available for applicants to higher education. 3. Other socio-economic data to which individuals can be linked. The main source of data for this is neighbourhood disadvantage as measured by the Wales Index of Multiple Deprivation (WIMD) or its equivalent in England. This is a measure of disadvantage based on multiple domains (including income and educational achievement) at the level of the Ward. An important and related classification used in Wales are Communities First areas/clusters. These were originally selected as the 100 most disadvantaged Wards in Wales (although these areas have been expanded over time to become geographical clusters). It is important to note that WIMD (or Communities First) are only measures of neighbourhood disadvantage, and reflects the distribution of disadvantage in an area. The more disadvantaged an area the more likely that households in that area are disadvantaged. However, it is not a measure of disadvantage at the household level and should not be treated as such. This is known as the ecological fallacy. However, if combined with socio-economic data at the household level it can be help produce a very detailed measure of socio-economic status. Such neighbourhood measures of disadvantage can be easily obtained by ‘looking up’ the postcode of an individual to the geographical unit at which the neighbourhood measure of SES is available (such as the Ward). The ONS provide Postcode Directories with this information historically and on a quarterly basis.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Use of Administrative Data
<table>
<thead>
<tr>
<th>Variable</th>
<th>Purpose</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postcode</td>
<td>As discussed above, the postcode of an individual can be very useful in being able to link to other existing SES data. It can also be used to explore the geographical relationships with access to higher education. In Wales this has been found to be particularly important, particularly in terms of the location and type of university being accessed.</td>
<td>Routinely recorded in most administrative data. However, accessing postcode information can be difficult because it can make an individual identifiable. Usually the actual postcode of an individual is not what is required and is only used to identify some other geographical unit (such as Ward or local authority). Hence one approach to mitigate the sensitive nature of the postcode data is to obtain the geographical information separately and then link this (without the postcode) to the other administrative data being used.</td>
</tr>
</tbody>
</table>
Prior educational attainment

<table>
<thead>
<tr>
<th>Variable</th>
<th>Purpose</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>The most important predictor</td>
<td>the actual educational achievement of an individual without taking into account any background factors. This is known as a ‘raw’ measure of educational achievement. It is usually important to start with this approach before considering the other three. The reason why a ‘raw’ measure of educational attainment might be preferred is because the evaluation is interested in directly comparing the attainment of particular groups. This is important since the majority of university admissions is usually based on ‘raw’ measures of attainment.</td>
<td>There are three main sources of data on educational prior attainment available. 1. End of Key Stage 4 achievement (i.e. GCSE or equivalent qualifications) from the National Pupil Database. This is available for nearly all 15 year olds in the maintained education sector. Only those from independent schools and those in the maintained sector who are exempt from statutory assessments are not included. There are many ways Key Stage 4 achievement can be measured. Previous research shows that the a total points score of an individual’s best eight GCSE (or equivalent) qualifications is the strongest predictor of participation to higher education. Alternative measures commonly used include whether or not a pupil has achieved at least five ‘good’ GCSEs (i.e. with grades A*-C) including maths and English or Welsh (first language). 2. Post-compulsory educational achievement obtained from the Welsh Examinations Database (WED) (from school sixth forms) and Lifelong Learning Wales Record (LLWR) (from FE colleges). This measure is important since this is the basis on which admissions to university are determined. However, both datasets, and the LLWR dataset in particular, are difficult to use due to the complex nature of post-compulsory education and its qualifications. Although not recorded in the HESA Student Record these qualifications can usually be obtained from a university’s own student records. 3. UCAS Tariff obtainable from HESA Student Records. When a student enters higher education a measure of their post-compulsory prior attainment is recorded based on the UCAS Tariff associated with their previous qualifications. The total UCAS Tariff (TOTALTS in HESA Student Record) can then provide a proxy of prior attainment and is particularly helpful in taking into account the complex nature of post-compulsory qualifications. However, there are two important limitations of this. The first is that the total UCAS Tariff does not take into account the number of qualifications used to determine a student’s UCAS Tariff score. Therefore a student with six qualifications with low grades could have the same UCAS Tariff score as another student with three qualifications with high grades. For this reason the HESA Student Record contains a wide range of Tariff scores for different qualifications, including the number of qualifications the Tariff scores are based on. It is possible, therefore, to produce a more nuanced measure of prior attainment using these different measures. But the main limitation of using UCAS Tariff is that it is only available for university participants. So if an evaluation included non-participants it would be necessary to collect information on their prior attainment and calculate an equivalent Tariff score for them based on the UCAS formulae.</td>
</tr>
</tbody>
</table>
An important contextual factor to participation to higher education is the previous educational establishment a prospective applicant attended. Indeed, recent research shows that the secondary school a young person attended at age 15 can help determine the probability that they will go to university, even after controlling for their levels of prior attainment and socio-economic circumstances. Furthermore, undergraduate students from independent schools are less likely to get a ‘good’ degree than their equivalent state maintained counterparts (again after controlling for educational prior attainment). Therefore the school or previous educational establishment is important to widening access activities for both an important determinant of participation and as a ‘site’ for widening access activities or targeting.

There are a number of ways the previous educational establishment can be used to provide contextual data. The most important two are the characteristics of the school intake (e.g. the proportion of the pupil population eligible for free school meals) and the educational achievements of the school (e.g. the proportion of 15 year olds achieving at least 5 GCSEs with grades A*-C).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Purpose</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>School or previous educational establishment</td>
<td>An important contextual factor to participation to higher education is the previous educational establishment a student attended (if they applied through UCAS). (PREVINST in the HESA Student Record). For schools a UCAS school identifier is used (NB which is not the same the school identifier in the NPD). The HESA Student Records can also provide a derived variable for the type of establishment previously attended (e.g. UK state school, UK independent school, UK FE college or another UK HEI). Obviously the National Pupil Database (NPD) provides a school identifier for the school a pupil attends or attended. These official school identifiers are routinely used by government and the schools themselves. There are publicly available lists of schools and their contact details that include their official school identifier. The Welsh and UK Government regularly publish information about every state maintained school in Wales and England, including the characteristics of their intakes and their educational performance. In Wales the website mylocalschool.com provides easy access to this information on a school-by-school basis.</td>
<td>HESA Student Records provide the last educational establishment a student attended (if they applied through UCAS). (PREVINST in the HESA Student Record). For schools a UCAS school identifier is used (NB which is not the same the school identifier in the NPD). The HESA Student Records can also provide a derived variable for the type of establishment previously attended (e.g. UK state school, UK independent school, UK FE college or another UK HEI). Obviously the National Pupil Database (NPD) provides a school identifier for the school a pupil attends or attended. These official school identifiers are routinely used by government and the schools themselves. There are publicly available lists of schools and their contact details that include their official school identifier. The Welsh and UK Government regularly publish information about every state maintained school in Wales and England, including the characteristics of their intakes and their educational performance. In Wales the website mylocalschool.com provides easy access to this information on a school-by-school basis.</td>
</tr>
</tbody>
</table>
An example of how administrative data can be used in the evaluation of widening access activities can be found here:


This analysis examines the progress and outcomes of undergraduate students from ‘low participation neighbourhoods’ (i.e. Communities First and POLAR areas) using university administrative data. Two important findings emerged from this analysis. The first was that undergraduate students from ‘low participation neighbourhoods’ were just as likely to succeed at university as equivalently qualified undergraduate students from other areas. The second main finding was that many students from ‘low participation neighbourhoods’ had similar socio-economic characteristics as other students, highlighting the ecological fallacy of using area-based indicators for widening access policies and practices.

The analysis looked at six different measures of university progress (known as dependent variables), including whether students had withdrawn from university, whether they had to resit modules in Year 1, and their final degree classification. When comparing the outcomes of undergraduate students from ‘low participation neighbourhoods’ with undergraduate students from elsewhere the analysis had to consider differences in the composition of these two groups that may also be related to the outcomes of interest (i.e. to make the comparison between the two groups ‘fair’). One way of doing this is to undertake multivariate regression. This allows multiple variables to be considered simultaneously in order to assess how and to what extent they affect a given outcome. In other words when comparing the outcomes of the two groups of students we ‘control’ for differences in their composition. In this example the students’ prior attainment, age at entry, year of entry, gender, ethnicity, type of school last attended, socio-economic classification, disability and domicile were included in order to ensure the comparison between two groups of students was a ‘fair’ comparison.

The table below presents the result of binary logistic regression (a form of multivariate regression) for one of the outcome measures – whether an undergraduate student had to undertake at least one resit during their first year. The third column presents the odds ratio (or probability) of having a first year resit depending on the characteristics of the students (listed in the second column) against a comparison group of students (listed below the table). This shows, for example, that male students were more likely to have a resit in Year 1 than female students. The odds ratio for males is 1.73 – in other words, the average male was 73% more likely to have a resit than the average female after controlling for other characteristics included in the binary logistic regression. In contrast, the average undergraduate student from areas in Wales classified as Communities First (CF) and POLAR were only 9% (odds ratio of 1.09) more likely to have a resit than an equivalent undergraduate student from a non-low participation neighbourhood.

Undergraduate students from ‘low participation neighbourhoods’ were just as likely to succeed at university as equivalently qualified undergraduate students from other areas.
### An example of binary logistic regression using administrative data

**DV: Year 1 Resist(s)**

| Variable          | Reference groups: 2005/6 year of entry; Comprehensive Schools; Middle class, White students, non-Welsh speakers, females, domiciled in England, from non-LPN; * p < .05. ** p < .01.  
| Reference groups: 2005/6 year of entry; Comprehensive Schools; Middle class, White students, non-Welsh speakers, females, domiciled in England, from non-LPN; * p < .05. ** p < .01. From: Taylor et al. (2013:153) |  |
| Year of entry     | 2006/07 | 0.86**  
|                   | 2007/08 | 0.77**  
|                   | 2008/09 | 0.87**  
|                   | 2009/10 | 0.92  
|                   | 2010/11 | 0.00  
| Age               | Age on entry | 0.89**  
| Gender            | Male | 1.73**  
| Ethnicity         | Non-White | 1.88**  
|                   | Ethnicity not known | 1.20  
| Disability        | No known disability | 0.83**  
| School type       | Grammar | 1.22**  
|                   | Independent | 1.33**  
|                   | College | 1.06  
|                   | Unknown type | 0.83  
| Social class      | Working class | 1.00  
|                   | Not known | 1.08  
| Welsh language    | Welsh speaker | 1.00  
| Domicile          | UK (not Wales/England) | 1.94**  
|                   | Wales | 1.60**  
| LPN               | CF and POLAR | 1.09  
|                   | POLAR2 only | 1.13  
|                   | CF only | 1.13  
| Prior attainment  | Total UCAS Points | 1.00**  
| Constant          | | 2.70  

From: Taylor et al. (2013:153)
Useful references

Administrative Data Liaison Service (ADLS) www.adls.ac.uk. This website provides an overview on all administrative data in the UK, including the National Pupil Database (NPD) and the HESA Student Record. It also provides links to other relevant sources of details about these datasets.


### Illustrative examples of evaluating two different kinds of widening access initiatives

<table>
<thead>
<tr>
<th>SUMMER UNIVERSITY</th>
<th>CONTEXTUAL OFFERS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Context</strong></td>
<td><strong>Context</strong></td>
</tr>
<tr>
<td>Address inequalities in knowledge of higher education amongst Year 10 pupils.</td>
<td>Address inequalities in prior educational attainment of recent school leavers.</td>
</tr>
<tr>
<td><strong>Main aims</strong></td>
<td><strong>Main aims</strong></td>
</tr>
<tr>
<td>Better understanding of HE and how to access it. Increased motivation to do well in GCSEs and to further educational career.</td>
<td>Greater rate of participation amongst applicants who have experienced prior educational inequalities.</td>
</tr>
<tr>
<td><strong>Rationale</strong></td>
<td><strong>Rationale</strong></td>
</tr>
<tr>
<td>Some 14/15-year olds do not have access to the necessary social and cultural capital required to make positive decisions about their educational future, including what HE is for, what benefits it brings and how to secure a place in HE.</td>
<td>Some applicants have not had the same educational opportunities as other applicants and so may not achieve the same entry requirements despite having the potential ability to achieve well at university.</td>
</tr>
<tr>
<td><strong>Inputs &amp; activities</strong></td>
<td><strong>Inputs &amp; activities</strong></td>
</tr>
<tr>
<td>3-week visit to university during summer vacation by FSM pupils from local schools to be introduced to the range of subjects/research taught, university life, admissions information, tutoring for core GCSEs, etc., plus 1-to-1 mentoring during Year 11.</td>
<td>New university policy that (a) identifies potential WA applicants and (b) permits admissions tutors to lower typical entry requirements by two grades.</td>
</tr>
<tr>
<td><strong>Outputs</strong></td>
<td><strong>Outputs</strong></td>
</tr>
<tr>
<td>Number of pupils attending the event. Improved understanding about barriers and access to HE knowledge for practitioners.</td>
<td>Number of: applicants flagged by admissions tutors; offers made; applicants who take up offer; places confirmed (and by department).</td>
</tr>
<tr>
<td>SUMMER UNIVERSITY</td>
<td>CONTEXTUAL OFFERS</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><strong>Process evaluation</strong></td>
<td>Interviews with admissions tutors asking them about their interpretation and implementation of the new university admissions policy. Focus groups with academic teaching staff about the policy and how they think students will progress on their course. Interviews with WA students during their first year about their transition to HE and their academic progress.</td>
</tr>
<tr>
<td>Pre-event survey asking pupils about their background, current understanding of HE and probability that they will go to HE. Focus groups and interviews with pupils during visit. Post-event survey asking about the visit, what they liked etc. and how it might benefit them. Follow-up survey with participants one year after event asking about future plans.</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome evaluation</strong></td>
<td>Using university administrative records analyse the academic progress of WA students who received lower entry requirements with other students, controlling for a range of background factors (such as prior attainment, SES, ethnicity, gender, type of school previously attended, etc.). Examine range of outcomes, such as withdrawal in Year 1, average marks in Years 1, 2 and 3, and degree outcomes.</td>
</tr>
<tr>
<td>Similar survey of Year 11 FSM pupils in schools in a different locality and corresponding follow up survey of these pupils a year later (control group). Statistical comparison of responses to survey in Year 12 comparing control group with intervention group, and controlling for attitudes in Year 11 and other background factors.</td>
<td></td>
</tr>
<tr>
<td><strong>Longitudinal / extended evaluation</strong></td>
<td>Using HESA data compare the progress of the intervention group of students with equivalent students at similar universities, controlling for other background characteristics. Examine destinations of WA students 6 months after graduating.</td>
</tr>
<tr>
<td>Link pupils in evaluation to NPD and HESA records in order to see whether the intervention group are more likely to go to university than the control group.</td>
<td></td>
</tr>
</tbody>
</table>