

Participation in Higher Education: A Study to Determine Whether the Higher Education Initial Participation Rate Should be Disaggregated

Professor Brian Ramsden

Research Report
No 676

*Participation in Higher Education:
A Study to Determine Whether the Higher
Education Initial Participation Rate
Should be Disaggregated*

Professor Brian Ramsden

The views expressed in this report are the author's and do not necessarily reflect those of the Department for Education and Skills.

© Brian Ramsden 2005
ISBN 1 84478 565 3

Contents

Participation in Higher Education	Error! Bookmark not defined.
A. Executive summary	3
B. Introduction	5
C. Background.....	6
D. Methodology of this review	8
E. The Numerators – general issues.....	10
F. The Denominators – general issues.....	12
G. Alternatives to disaggregation of the HEIPR	13
Possible use of the API	13
Possible use of disaggregated numerators as an indicator.....	16
H. Disaggregation of the HEIPR by ethnicity	18
The numerators in relation to ethnicity	20
The denominators in relation to ethnicity	21
Options in relation to ethnicity.....	22
Recommendations in relation to ethnicity	23
I. Disaggregation of the HEIPR by Social Class	24
The numerators in relation to social class	25
The denominators in relation to social class	30
Options in relation to social class	31
Recommendations in relation to Social class	35
J. Disaggregation of the HEIPR by Disability	37
The numerators in relation to disability	37
The denominators in relation to disability	39
Options in relation to disability	41
Recommendations in relation to Disability	44
K. Disaggregation of the HEIPR by Region	45
The numerators in relation to regional disaggregation	45
The denominators in relation to regional disaggregation	45
Options in relation to regional disaggregation	46
Recommendation in relation to the disaggregation of the HEIPR by Region ..	47
L. Other issues arising from this project	48
Religion.....	48
The relationship of the HEIPR with other participation statistics	48
M. Conclusions and recommendations	50
General recommendations	50
Recommendations in respect of ethnicity	50
Recommendations in respect of Social class.....	51
Recommendations in respect of disability	52
Recommendations in respect of disaggregation by Region	52
Conclusion.....	52
N. Appendices	53
Appendix 1: Membership of the Steering Group for the project.....	53
Appendix 2: Individuals consulted during this project.....	54
Appendix 3: Issues concerning Religion	55
Appendix 4: Aggregated minority ethnic groups	58
Appendix 5: References.....	59

A. Executive summary

1. I was commissioned by the Department for Education and Skills (DfES) in October 2004 to undertake an option appraisal regarding the possible disaggregation of the Higher Education Initial Participation Rate (HEIPR) by ethnicity, social class and disability. I was subsequently asked to extend this remit and to explore the feasibility of disaggregating the total HEIPR by region.
2. The HEIPR is the measure of progress towards the Government's target (expressed as a Public Service Agreement (PSA) target for the DfES) as follows:

"By 2010, increase participation in Higher Education towards 50% of those aged 18 to 30."

3. The project has four components, relating to ethnicity, social class, disability and region (each to be considered independently). In each of these components, if the HEIPR is to be disaggregated, it is necessary to identify adequate definitions of both the numerators – i.e. higher education participants - and the denominators – i.e. national population estimates - and to ensure either that robust and comparable data is available at each individual year of age between 17 and 30, or that there is an acceptable method of estimation.
4. I have identified general issues relating to the statistics about Higher Education entrants, and also about the national population estimates which should constitute the denominators.
5. I have also considered issues relating to the disaggregation of the numerators and the denominators in respect of ethnicity, social class, disability and region.
6. In the process, I have found some impediments to a simple disaggregation of the HEIPR, arising from incompatibilities between the Higher Education statistics and the national population projections, and also from weaknesses in the coverage of each of these.
7. I have identified options for generating either a disaggregated HEIPR, or an alternative to that, in respect of ethnicity, social class, disability and regional disaggregation. I have been asked by the Steering Group to limit these options to those which are plainly viable, and have also been asked to comment on the cost implications and to indicate in relation to each option the extent to which it might be constructed retrospectively in order to generate time series data.
8. In relation to ethnicity, I have concluded that there is potentially a valid mechanism for disaggregating the HEIPR by ethnicity although some work would be needed in order to ensure that the numerators and denominators were comparable. I have emphasised that a simple split between "white" and "non-white" is far from ideal, and that if the Department intends to publish an ethnically disaggregated version of the HEIPR, it should do so if possible at a detailed level, including further disaggregation by

gender, while recognising that the robustness of the data will be questionable in respect of the smaller ethnic groups.

9. In relation to social class, I have questioned the conventional assumption that this can be meaningfully derived from the occupational category of either the student or his/her parents. And I have identified some alternative approaches.
10. In relation to disability, I have concluded that it would be unwise at present to consider any disaggregation of the HEIPR by disability, since there is no immediate prospect of aligning the numerators and the denominators.
11. In relation to regional disaggregation, I have concluded that there are bases for disaggregating the HEIPR.
12. I have also commented on:
 - ◆ The use of the Age Participation Index, about which I was specifically asked to report, and have suggested that, as presently defined, it provides no useful information about participation in Higher Education.
 - ◆ The availability of data about religion, since I was asked formally to add this to my remit.
13. And finally I have set out some recommendations to the Department for Education and Skills. My recommendations are summarised at page 50.

B. Introduction

14. I was commissioned by the Department for Education and Skills (DfES) to undertake an option appraisal regarding the possible disaggregation of the Higher Education Initial Participation Rate (HEIPR).
15. My terms of reference were:
 - ◆ To undertake an option assessment on behalf of the DfES in relation to the disaggregation of the Higher Education Initial Participation Rate by social class, ethnicity and disability. This work to provide views on:
 - ◆ disaggregating the HEIPR
 - ◆ providing other mechanisms for relating these areas to the HEIPR (e.g. surveys)
 - ◆ the pros and cons of monitoring social class participation using DfES' other existing measure – the Age Participation Index (API)
 - ◆ To liaise with the project manager and consult HE:AS on the options drawn up and then to present them to the other users, including UUK, HEFCE and possibly ONS and Ministers.
 - ◆ To prepare a draft report for consideration by HE:AS and, in the light of that discussion, provide a final report in Microsoft Word format on CD-ROM, together with a hard copy.
16. I was required to provide a progress report in early December 2004 and a final report by the end of February 2005.
17. A progress report was submitted to the Steering Group for consideration at its meeting on 10th December 2004, and it was subsequently agreed that a draft final report should be considered by the Steering Group on 9th February: there was also an opportunity to discuss the interim findings with other DfES officials. At that meeting, I was asked to conduct some rapid consultation with other stake-holders, with a view to providing a final draft to the Steering Group by 7th March, for comments by the Steering Group and perhaps also comments by Ministers, on the understanding that any remaining comments would lead to a finalised version of the report by the end of March 2005.
18. This is my final report, amended in the light of comments received in the course of consultation.

C. Background

19. In 2003, I undertook a quality review, commissioned by the DfES and ONS, into the methodology for calculating the Initial Entry Rate into Higher Education (the IER). That review led to the redefinition of the IER as the Higher Education Initial Participation Rate, (HEIPR) and its inclusion within the area of National Statistics. The HEIPR has now been published, together with a time series.
20. The HEIPR is significant in its own right as a measure of initial participation in Higher Education (HE). However it also has political significance as being the measure of how the Government is faring against its target of increasing participation in HE towards 50% of those aged 18-30 by the end of this decade.
21. In my final report of the IER quality review¹, I advised that the HEIPR should be disaggregated by gender, but I expressed doubts about the feasibility of disaggregating it by other variables. However, the Department subsequently indicated its wish to consider any options for monitoring participation by social class, ethnicity, disability and region, and that is the focus of this current project.
22. The calculation of the HEIPR is deceptively simple: in essence it is defined as follows:

“The Higher Education Initial Participation Rate is a measure of participation in Higher Education for the first time on the part of young people living in England, and is constructed by summing the percentages entering higher education for the first time at each age between seventeen and thirty years of age.

For the purposes of this definition, an entrant is included only if he/she remains engaged in the programme for at least six months.

Higher Education is defined according the definition which is applicable in the country in which the individual is studying, and includes all publicly funded higher education institutions and Further Education Colleges within the UK. At present students entering privately funded UK institutions including “corporate” universities are excluded, as are students pursuing their higher education experience outside the United Kingdom. The Initial Entry Rate therefore slightly undercounts the totality of higher education experience on the part of young English residents.”

23. It follows that the HEIPR, **and any robust disaggregated version of it**, must have as the numerators the number of initial English-domiciled² participants in higher education at each age between 17 and 30, and as the denominators the English-domiciled population again at each age. Plainly also, if the statistic is to be robust, the definitions applicable to the numerators and the denominators need to be either identical or sufficiently similar to enable valid comparison.

¹ Ramsden, 2004.

² My use of the term “English-domiciled” in this report as far as students are concerned follows the definition within the HESA student record, i.e. it is derived from the student's permanent or home address prior to entry to the programme of study.

24. It is this need to identify robust and comparable numerators and denominators which is the major challenge for the Department, in considering options for disaggregation.
25. If it were to be the case that robust and comparable numerators and denominators could not be found or generated by survey, then an alternative methodology might be envisaged for reporting on the components of the statistic.
26. The rest of this report looks at the individual elements which make up this project, and sets out options for meeting policy objectives.

D. Methodology of this review

27. In carrying out this review, I have sought input from a variety of sources. I have been helped in this by the Steering Group of the project (the members of which are listed in Appendix 1), both at our meetings, and by informal email contact.
28. In addition I have consulted colleagues in relevant departments and organisations about particular aspects of my remit. The list of those who have contributed to this report following consultation is given in Appendix 2. It should be noted that in some instances I have consulted individuals about a particular aspect of my work, rather than the entirety of it, and therefore it should not be assumed that individuals referred to were either aware of or in agreement with the conclusions which I draw in my report. I am grateful to all of those who have given their time to assist me in this project.
29. The project has four components, relating to ethnicity, social class, disability and region. In each of these components, if the HEIPR is to be disaggregated, it is necessary to identify adequate definitions of both the numerators and the denominators, and to ensure either that robust data is available at each individual year of age between 17 and 30, or that there is an acceptable method of estimation.
30. A starting position for this review, in the light of early discussions suggested that:
 - a) In relation to ethnicity, there is a common coding frame in both higher education and the population estimates derived from the census: the major challenge is to generate adequate inter-censal population figures as the denominators
 - b) In relation to social class, there is now a common coding frame in both higher education and the population estimates – the NS-SEC classification. However, in addition to the problem of generating adequate inter-censal figures for the denominators, there is a major issue in relation to the numerators, which is that the NS-SEC classification is currently only available in respect of a sub-set of entrants to HE – those who enter full-time undergraduate programmes through UCAS.
 - c) In relation to disability, there is no obvious common coding frame, and therefore it is necessary to consider what developments might take place in order to generate a basis for disaggregation of the HEIPR.
 - d) In relation to regional disaggregation, there is no fundamental difficulty, but a minor issue regarding the identification of the home region of students within Higher Education.
31. In the initial stages of this review, I sought to isolate the major problem issues in connection with the various numerators and the denominators separately. I also initially explored the extent to

which an alternative methodology might be adopted for identifying the relationship between the HEIPR and the specific areas of concern.

32. The issues which arose in this stage of the project were discussed by the Steering Group on 10th December 2004. The discussion was helpful in concentrating on the major issues of concern.
33. There is one general issue which I should raise here before entering the more detailed analysis: that is the particular vulnerability of a statistic which is reported on an annual basis. I shall comment later about another such statistic (the API) , but it is sufficient at this point to note that any statistic which is reported on a time series basis – and especially one which is reported annually – is subject to scrutiny at a level which may be greater than it can bear. For example, if a ratio in year x were to be 38% and in year $x+1$ were to be 35%, that might be regarded as a major change – but the significance of that change would depend on the basis for collecting and analysing the data, and it might simply be the case that 35% and 38% were within a margin of error.
34. I believe that it is important that the Department should not introduce analysis at a level which the data cannot support: to do so would simply serve to replicate mistakes which have been made in the past.
35. In the following paragraphs, I shall address the more specific issues about the numerators and denominators which might be used in the disaggregation of the HEIPR.

E. The Numerators – general issues

36. In relation to the overall HEIPR (and its gender-disaggregated form) the numerators at each age are derived from data provided by the institutions to the various organisations which are responsible for collecting information about higher education in the UK. The Higher Education Statistics Agency (HESA) collects information about all HE students in publicly-funded HE institutions³ throughout the UK, and in aggregate this amounts to approximately 92%⁴ of the data which forms the HEIPR. In order to produce robust statistics for the numerators, however, it is necessary also to include data collected by the Learning and Skills Council (LSC) in relation to students following HE programmes within Further Education Colleges in England: in order to gain a complete picture, it would be desirable also to include data collected by the Scottish Executive, the Higher Education Funding Council for Wales and the Department of Education and Learning, Northern Ireland (DELNI) in relation to the English-domiciled students entering programmes of study at HE level in FE colleges in those countries - although the numbers involved are small, especially in respect of Wales and Northern Ireland.⁵
37. In relation to some of the relevant items of data, there is not complete coverage across all categories of HE student: in particular, some data is available – and with some limitations – only for students entering Higher Education through the UCAS admissions system. This system provides application and acceptance processes only for a subset of students, which excludes all part-time entrants, all entrants at postgraduate level, and also some entrants to full-time undergraduate courses.
38. The full-time undergraduate students who are omitted from the UCAS entrant data are those who are directly admitted to a university. These students constitute, in any recent year, some 15% of full-time undergraduate entrants⁶; and analysis which I have commissioned during this project or undertaken in my own research suggests that they have significantly different characteristics from UCAS entrants, for example:
- ◆ The non-UCAS entrants to full-time first degree courses are concentrated in the post-1992 universities (especially in the major conurbations.)
 - ◆ They include a significantly higher proportion of minority ethnic groups than the undergraduate population as a whole.
 - ◆ They are, on average, older than UCAS entrants; and therefore their social class, as currently defined, is more likely to be

³ For the avoidance of doubt, this includes the Open University, which is the major part-time provider of HE courses in the UK.

⁴ My estimate - BAR

⁵ As regards Scotland, in 2000-2001, 428 English-domiciled students were in their first year of study at higher education level within FE colleges in Scotland – I have not thought it necessary to update this figure.

⁶ This estimate has been confirmed to me in the course of this project by both HESA and UCAS.

derived from their own occupations than that of their parents:
this is an issue which I shall address below (page 25).

- ◆ These entrants are also atypical in that they have a significantly lower average A-level points score⁷ on entry (on my own findings, approximately 13.5 as compared with 18.5)⁸
39. It is therefore in my view extremely unsafe to regard data provided and reported solely for UCAS entrants as being adequate to provide a basis for any general participation statistics concerning Higher Education students within the UK. To do so would be to admit a potentially significant bias into the statistics. I therefore recommend that the DfES should not use data available only for UCAS entrants as a basis for any general statistics about participation in Higher Education within England or the UK.
 40. To say this is not in any sense to denigrate the value of the application statistics collected by UCAS, which are a rich source of information within their limitations: it is simply to recognise a reality which is often overlooked in simple analysis of data about higher education, which is that UCAS entrants are not representative of all entrants to HE.

⁷ I am of course aware that the A-level points score has been replaced by the UCAS Tariff score, but within the timescale of this project I have not been able to pursue comparative figures under the new classification.

⁸ While the overall discrepancy between overall intake and UCAS intake is not in question, I would accept that the more detailed analysis of entrants contained in this paragraph requires further validation, and I hope that this will be undertaken.

F. The Denominators – general issues

41. In relation to the overall HEIPR (and its gender-disaggregated form) the denominators at each age are derived from the national population estimates calculated annually by the Office for National Statistics (ONS) having reference to the population projections developed by the Government Actuary's Department.⁹ These national population estimates are ultimately derived from the most recent census (currently 2001).
42. For the purposes of this report, I am assuming that these national population estimates may be taken as read, although I understand that there are suggestions that at certain age points there are concerns about the relationship of the overall population estimates with the statistics derived from the School Census, at some years. While these years are significant in terms of the movement into and out of schools, I do not feel that the discrepancies invalidate the national population statistics.
43. In order to assess the feasibility of disaggregating the HEIPR by social class, ethnicity, disability and region, we need to consider whether it is possible to obtain estimates annually of the numbers of English-domiciled people at each year of age between 17 and 30 falling into each disaggregation category.
44. In relation to the census year 2001, this is a comparatively simple exercise. However, in relation to the inter-censal years, it is necessary to find a basis for making a robust estimation.
45. There are in principle two bases for making such estimates:
 - ◆ A cohort projection approach based on the last census
 - ◆ An approach which uses survey data to adjust the estimated population.
46. I shall address each of these in relation to each of the categories which I am studying in the following paragraphs.

⁹ This is in fact an over-simplification: for example, the population estimates are in turn adjusted in order to generate a reference date equivalent to that which is used in HE data. However it is not felt to be necessary to re-visit the details of the calculation in this report, since they were described in detail in Ramsden, 2004.

G. Alternatives to disaggregation of the HEIPR

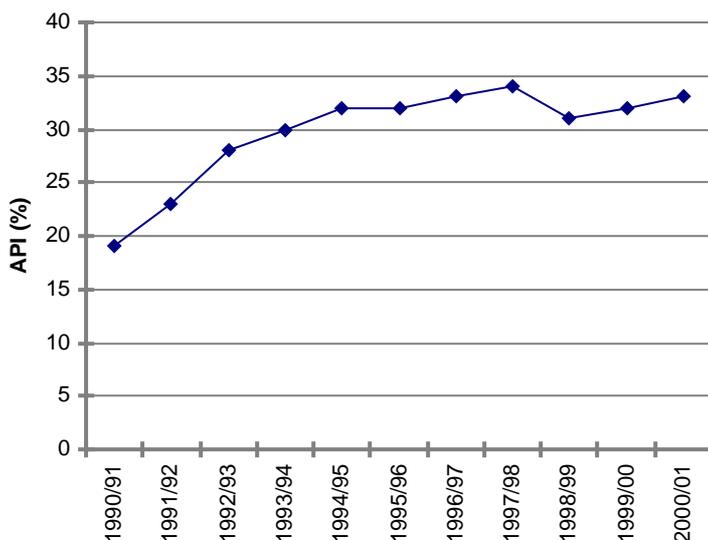
47. The previous paragraphs have discussed issues in relation to the numerators and denominators of the HEIPR, which would need to be addressed in order to achieve disaggregation in relation to the four factors which underlie this study.
48. This section of my report looks at alternative approaches, which might be adopted in the short to medium term, in order to provide relevant information, while not actually disaggregating the HEIPR.
49. Two such approaches are identified in this draft report:
 - ◆ Possible use of the API, as a basis for providing some disaggregated statistics
 - ◆ Possible disaggregation of the numerators in the HEIPR, without attempting disaggregation of the denominators

Possible use of the API

50. I have been asked to consider whether, and if so to what extent, the Age Participation Index (API) might be used as an alternative to the HEIPR in order to analyse participation by ethnicity, social class and disability.
51. I should note here that Government statistics about participation by social class have traditionally been drawn from the API data – and still are, at least in theory. It is necessary to consider two separate issues here, i.e. the basic API and its derivative version the API by social class

The basic API

52. The basic Great Britain API is defined as ‘the number of UK-domiciled young (aged under 21 years) initial entrants to full-time and sandwich undergraduate courses of higher education in Great Britain, expressed as a proportion of the averaged Great Britain 18 to 19 year old population’.
53. This basic definition itself raises questions, since the numerator is drawn from the UK and the denominator from Great Britain only, and since the numerator relates to all ages up to 20 while the denominator is limited to the “averaged” 18-19 year old population (and is therefore susceptible to the significant year-on-year movements in individual ages.) The latest version of the API is illustrated in the following chart.



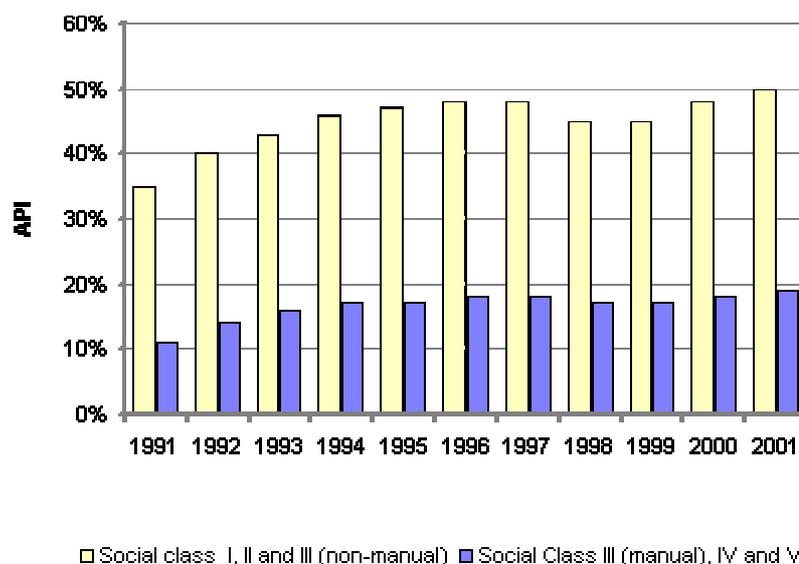
54. The overall GB API is disaggregated for England, Scotland and Wales (although the methodology for calculating these is, I understand, slightly different from the methodology for calculating the GB API).
55. In earlier years, two other participation measures were calculated: the Younger Mature Entry Index, which covered 21-24 year olds, and the Older Mature Participation Index covering 25-34 year olds, but these were not calculated after 1992. These also were limited to full-time undergraduates.
56. Although in many ways the API was intended to perform a similar role to that of the HEIPR, its construction is very different, and in some ways much simpler – or cruder, depending on your point of view.
57. The basic API appears to have the following weaknesses:
- ◆ It explicitly excludes entrants to part-time courses of Higher Education, as well as excluding students aged over 20: that is its major limitation as a participation measure.
 - ◆ While it recognises that not all students embarking on a course for the first time are new entrants to higher education, it makes no attempt to determine the actual number of new entrants in any year. In order to estimate the proportion of new entrants, factors are applied to reduce the numbers recorded in the index, as follows:
 - students in former UFC¹⁰ institutions - 0.977
 - teacher training students in former PCFC institutions - 1.00
 - non teacher training students in former PCFC institutions - 0.87
 - HE students in Further Education Colleges - 0.87

¹⁰ The Universities Funding Council (UFC) and the Polytechnics and Colleges Funding Council (PCFC) were disestablished in 1993 and their functions in England were taken over by the Higher Education Funding Council for England (HEFCE), following the 1992 Further and Higher Education Act.

- ◆ These factors have been unchanged for more than ten years, and therefore have not been adjusted in a decade which has seen very significant changes in the nature of higher education provision and the participation of students within HE.
 - ◆ There is no limitation on the actual or expected length of course, for inclusion in the API; this is perhaps understandable since it is limited to full-time undergraduates, but it follows that it is not in any way comparable with the HEIPR.
58. On the other hand, the strength of the API may be that it has a long history – although that strength may also be in a sense also a weakness, since its history pre-dates the expansion of higher education to include teacher education and nursing, for example, and uses some components which are plainly unsuitable in 2005. Although the time series based on the API is long, the changes which have taken place in the nature of higher education provision undermine comparisons.
59. In my view the Age Participation Index as currently defined has outlived its usefulness as a measure of participation in Higher Education. I recommend that, if the Department continues to provide information based on the API, it should develop appropriate and strong caveats which should accompany any such information.

The API by social class

60. The API is also disaggregated by “social class” and has been regularly published on this basis.
61. The following chart shows the API by social class as most recently reported:



62. The API by social class also suffers from some weaknesses, as follows:
- ◆ The methodology for calculating this statistic involves taking data about the proportions of UCAS entrants within each social

class and applying these to the totality of initial entrants as defined within the API. The dangers in doing this have been described above (page 10).

- ◆ There is a further problem in that this numerator is then compared with a denominator which consists of the split in the *total economically active population at the time of the last census, unadjusted for change over years*. Plainly the numerators and denominators are inconsistent and so the calculation is invalid.
63. However, the calculation of this statistic ceased with the production of the 2001 data, although a series exists over several years. I am pleased to be advised that there are no current plans to re-issue it – not least because I believe that there is a strong possibility that it has significantly under-counted the participation in higher education of people from the lower socio-economic groups.
 64. If Government wishes to monitor “young participation” in higher education – which is of course an entirely different concept from participation up to the age of thirty which is the focus of the HEIPR – then there is a recently developed methodology which should be considered. That is set out in a report published in January 2005 by the Higher Education Funding Council for England¹¹, and which includes a proxy for social class – but one which could not be carried over into the 17 to 30 age range, since the denominator is based on enrolments in schools at the end of compulsory education.
 65. This methodology constructs Young Participation Rates (YPR) which can be applied at a detailed geographical level, since it is derived from local information. The relationship between the YPRs and the HEIPR is clear, and the author of the HEFCE report identifies the respects in which the two measures differ.
 66. If the Government wishes to monitor the social composition of the young entrant population into Higher education, I recommend that it should use a methodology similar to or derived from the HEFCE methodology for constructing Young Participation Rates, and should make clear that these measures relate solely to young entrants, and are not comparable with the HEIPR.

Possible use of disaggregated numerators as an indicator

67. There is an alternative approach to the reporting of information about HE participation which does not involve the generation of a disaggregated HEIPR, but which might be of some value and interest.
68. This is to report the overall HEIPR and to provide beneath it the proportion of the numerator which has specific characteristics.
69. So, for example, the HEIPR might be reported along the following lines:

¹¹ HEFCE, 2005

“The HEIPR in the year 200x-y is 46%, and within the new entrants to higher education:

- ◆ *x% are right-handed (an increase/decrease compared with $x_{-1}\%$ last year)*
- ◆ *y% have red hair (an increase/decrease compared with $y_{-1}\%$ last year)*
- ◆ *z% are wearing Rolex watches (an increase/decrease compared with $z_{-1}\%$ last year)*

70. This is quite clearly not a disaggregation of the HEIPR, since only the numerator is disaggregated, and no tracking is undertaken of the denominator. The statistic would therefore have little validity if the underlying population was subject to annual change.
71. However, it has in its favour that it would provide a time series of statistics, and there is no limit to the number of variables which could be included as bullet points (except the numerator definitions, which are generally consistent among the data providers).
72. I believe it also to be the case that the denominators are in fact less susceptible to annual change than the numerators: for example, whatever definition were to be used of “Social Class”, the annual change in the population at large is small, while the point of political interest – and the relevant statistic - is the intake into Higher Education, which may be more dynamic.
73. I shall illustrate this approach further in relation to options for reporting on disability.

H. Disaggregation of the HEIPR by ethnicity

74. I have been asked to consider the scope for disaggregating the HEIPR by ethnicity.
75. Ethnicity is not a characteristic which can be identified as clearly and certainly as, for example, gender. An individual's recorded ethnicity in any public records is either self-assessed or assessed by another member of the household: and so, for example, the ethnicity assigned to someone by the head of their household in the Census, might not tally with that own individual's assessment of his/her own ethnicity in Higher Education records. This factor potentially weakens the robustness of the indicator, perhaps especially in relation to mixed ethnic groups.
76. Also, over a period of time, an individual's own perception of his or her ethnicity may indeed change: research work is being undertaken in this area.¹² And, increasingly, mixed ethnicities are of increasing relevance (four such categories being recognised for the first time in the census definitions for 2001) and coding frames for Higher Education students have recently been amended to recognise this.
77. Some valuable work in relation to the disaggregation of the HEIPR by ethnicity has already been undertaken using the 2001 census data in order to construct the denominator, in the context of the DfES research report *"Why the Difference? A Closer Look at Higher Education Minority Ethnic Students and Graduates"*.¹³ That report includes a technical report setting out the methodology used, and many of the perceived implications of the work. I do not propose to repeat that analysis here, but should note that I am entirely in agreement both with the methodology and the conclusions in that work.
78. Ethnicity, unlike gender, is not simply a binary field. There is a tendency within the HE community to refer to "ethnic minorities" as though all of them had the same characteristics. This is plainly untrue, and trivialises important issues. The ethnic groups within the UK-domiciled population are complex and their participation within Higher Education varies very considerably – and the interaction of ethnic group and gender is also extremely important, because of different traditional patterns of participation. These are admirably summarised in *"Why the Difference?"* and it is appropriate to repeat a part of that summary here:
 - ◆ White, mixed ethnic, female Pakistani and especially Bangladeshi groups have the lowest female participation rates of any ethnic group (33-44 per cent range)
 - ◆ By contrast, female Black African, Asian other and Black other have the highest participation rates (over 70 per cent).

¹² See for example, research outline at Simpson, 2004.

¹³ Connors et al., 2004

- ◆ The lowest male participation rates are among Black Caribbean, White and mixed ethnic groups (34-36 per cent).
- ◆ while the highest male participation rates are among Black African, Indian and Asian other groups (over 70 per cent)

79. The full table as published in “Why the Difference?” is given below:

Disaggregated HEIPR, 2001, as published in What's the difference?

Ethnic group	Female			Male			Total		
	Est. pop.	HE entrants	HEIPR %	Est. pop.	HE entrants	HEIPR %	Est. pop.	HE entrants	HEIPR %
White	3,838,120	105,470	47	3,898,230	90,410	34	7,736,360	195,880	38
All minority ethnic groups	541,350	22,230	58	524,580	21,120	55	1,065,930	43,360	56
Black Caribbean	52,330	1,870	52	45,210	1,160	36	97,540	3,100	45
Black African	64,020	3,100	75	56,650	2,660	71	120,670	5,800	73
Black Other	11,480	610	72	10,320	440	56	21,800	1,050	64
Indian	131,670	6,470	72	129,630	6,390	70	261,310	12,900	71
Pakistani	102,460	3,330	44	102,020	4,090	54	204,480	7,420	49
Bangladeshi	44,300	1,030	33	39,000	1,220	43	83,300	2,310	39
Chinese	35,700	1,370	50	36,940	1,420	47	72,640	2,840	49
Asian Other	26,710	1,600	94	35,140	1,630	74	60,850	3,230	83
Mixed ethnic	73,700	2,580	44	69,680	2,040	35	143,350	4,610	40
All (known)	4,379,470	127,700	43	4,422,810	111,530	37	8,802,290	239,240	40

Source: HEFCE and DFES as quoted in DfES Research Report 552, table 4.1¹⁴

80. These significant differences in participation rates, as between ethnic groups by gender appear to point the way towards an approach to disaggregation of the HEIPR, i.e. that ideally it should be disaggregated at the most detailed level.
81. However, on closer examination, the table in fact exemplifies the danger of undertaking such detailed disaggregation. There is at least one point within the table where the data might be regarded as unconvincing: the results in respect of Asian other students show a very significant difference between the numbers in the female and male national populations, the male estimate being 32% higher than the female figure. It has been suggested that “new groups with high proportions of immigrants tend to have more young men than young women, until they are better established.” However, the very high participation rate – 94% - for women, suggests that either there is a problem in the data in either the numerators or the denominators, or that the numbers in a particular cell are too small to enable a valid and robust HEIPR to be calculated at this level of detail.¹⁵ (A further table from the “Why the Difference?” report is contained in Appendix 4: this shows an aggregated version of the entrant data, although it

¹⁴ It should be noted that this table is identified as having several caveats, relating to both the national statistics available and also the limitations of linking two datasets collected under different terms and reported by different individuals.

¹⁵ The *reductio ad absurdum* here would be to imagine a situation in which, within one ethnic group and gender, there was only one person aged, say, 28 and that that person was an entrant to Higher Education in year x. Since the HEIPR is a summation of participation rates **at each year of age, the HEIPR for that group in that year would inevitably exceed 100%. Quod est absurdum.**

should be noted that its definition is not compatible with the table above.)

82. Although I have sought to find out more about this issue, I have been unable to determine which of these two possibilities applies in this case; but whichever it may be, it argues for caution in manipulating small numbers – an issue to which I shall return.

The numerators in relation to ethnicity

83. The ethnicity of students is collected routinely for all students in both UK Higher Education institutions and FE colleges. While it is not compulsory for students to respond to questions about ethnicity, there is a high level of response, among both full-time and part-time students. (The response rate among full-time first degree students is approximately 95%, and among all new part-time students it is 84%.)
84. The definitions are consistent as between HESA and LSC collections, and also broadly within the Scottish FES.
85. They are also consistent with the definitions used for the denominator, i.e. census 2001 definitions.
86. The full coding frame for HESA ethnicity data is as follows:

10 White.
11 White - British.
12 White - Irish.
13 White Scottish.
14 Irish Traveller.
19 Other White background.
21 Black or Black British - Caribbean.
22 Black or Black British - African.
29 Other Black background.
31 Asian or Asian British - Indian.
32 Asian or Asian British - Pakistani.
33 Asian or Asian British - Bangladeshi.
34 Chinese.
39 Other Asian background.
41 Mixed - White and Black Caribbean.
42 Mixed - White and Black African.
43 Mixed - White and Asian.
49 Other Mixed background.
80 Other Ethnic background.
90 Not known.
98 Information refused.

87. A virtually identical coding frame exists for students in FE colleges in England, and with only appropriate variations in Scotland.
88. The major issue which arises is the extent to which the various ethnic groupings should – and can - be aggregated in relation to the HEIPR. I shall return to this issue when I consider options, below. A further issue to be considered is whether any ethnicity HEIPR should also be disaggregated by gender, in view of the different participation patterns of men and women in some ethnic groups.

The denominators in relation to ethnicity

89. While the numerators in relation to ethnicity cause little problem, there are several issues of concern in connection with the denominators.¹⁶
90. It is, for example, noteworthy that population estimates as published by ONS include students from overseas studying within the UK. It is obviously inappropriate that they should be included in the denominator, since the numerator is explicitly English-domiciled students. This issue is perhaps particularly significant in relation to the Chinese ethnic group, which numbers about 70,000 people in the age range 17 to 30.
91. It would appear to be the case that some adjustments would be necessary here in order to generate a robust HEIPR; since, for example, the numbers of Chinese nationals engaged in higher education are known, and since the majority of overseas students of Chinese ethnicity are understood to be Chinese nationals¹⁷, the population figures could be amended in order to eliminate Chinese nationals and other ethnic Chinese of non-England domicile from the denominator.
92. I recommend that this should be done, together with any other adjustments which might similarly distort the population denominators.
93. A further and more major issue is the estimation of the population by ethnic group for the inter-censal years. Considerable work has already been done on the possible use of projection models to project the population of ethnic minority groups.¹⁸
94. This work makes use of a wide variety of data including migration patterns, births and deaths, etc, as well as the integration of data from surveys (Labour Force Survey and the ONS Longitudinal Study (LS), notably.) The technique, which is described in detail by Rees¹⁹, is chiefly aimed at making estimates in relation to comparatively small local areas. However it can of course be used to estimate national populations, and is probably the most accurate methodology for doing so.
95. An alternative is to use a single survey as a proxy, and the Labour Force Survey is the obvious one here. It appears to be the case that the ethnicity data derived from the Labour Force Survey is, at the higher levels of analysis, reasonably comparable with that derived from the 2001 Census, at least at the level of all England, and that therefore to use trends identified in the LFS is likely to be an acceptable basis for adjusting the population denominators for the disaggregated HEIPR.

¹⁶ I exclude here the possibility of bias in the responses to the Census within particular ethnic groups, which is a broader issue.

¹⁷ Note the issues relating to students whose domicile is given as Hong Kong, Macao or Taiwan, and the Foreign Office status of these: the totality of overseas students in China, including Hong Kong, Taiwan and Macao in the most recent year exceeds 60,000.

¹⁸ see for example Haskey (ed.), 2002

¹⁹ Rees, 2002

Options in relation to ethnicity

96. In my view, no action needs to be taken in relation to data collection regarding the numerators, which are as robust as could reasonably be expected. The data collection agencies and the institutions have succeeded in generating more comprehensive statistical information about ethnicity than was originally envisaged when the data collection arrangements were put into place.
97. In relation to the denominator, options include the following:

Ethnicity Option 1: Cohort projection model

98. The first option is the cohort projection model which has been described above (paragraphs 98-99). This has the advantage of being probably the most robust basis for generating inter-censal statistics.
99. However, as I have been advised, a detailed methodology is not yet in place, although much work has been done in this area.
100. There may also be an issue in relation to transparency, in view of the complex nature of the cohort projection approach.
101. There is potentially a cost implication here, which is the cost of developing a methodology which can be implemented on a consistent basis from year to year, using all relevant data. I would regard this as being a medium cost implication.
102. The potential for retrospective time series comparisons exists, although it is limited by the change in data definitions in relation to ethnic groups, introduced for the 2001 Census.

Ethnicity Option 2: adjustment of denominator by LFS data

103. The second option for disaggregating the denominators in relation to ethnicity involves using the Labour Force Survey data as a proxy for calculating the annual change in the ethnicity of the population as a whole. This has the apparent advantage of simplicity, being based on one robust annual survey.
104. The simplicity of this approach follows through to its very low cost implications: if the confidence thresholds were thought to be satisfactory, a comparatively simple analysis from ONS of the LFS data might - at least in theory - be used in order to update the census population for the inter-censal years.
105. However, an examination of the data which I have commissioned during this study suggests that any resulting statistics would not be adequately robust except at a high level of aggregation, and this is confirmed in advice from the Office for National Statistics concerning confidence thresholds.
106. If the information were confined only to the question of the participation in Higher Education on the part of all ethnic minority groups compared with white ethnicity, by gender, some worthwhile information could be derived, although it would fall short of providing the level of detail which would be desirable.

107. Retrospective time series comparisons could be generated in respect of the high level data.

Recommendations in relation to ethnicity

108. In relation to the publication of the statistic, the groupings of ethnicity need to be agreed: there are potentially several options in relation to the aggregation of ethnic codes. While a high level aggregation is likely to be statistically most robust, it needs to be recognised that the participation patterns of ethnic groups among England-domiciled people are very different, as indicated above.
109. There is therefore, as we have seen, a need to balance the desirability of having detailed information with the imperative of having robust and annually comparable statistics.
110. **I recommend that** if the Government wishes to disaggregate the HEIPR by Ethnicity it should commission work to generate a basis for assessing the change each year in the population at individual year of age and by individual ethnic group, using the cohort projection approach.
111. In the interim, and until the work referred to above is complete, **I recommend that**, if the Government wishes to have a simple disaggregation of the HEIPR by ethnicity, then it should use the Labour Force Survey data to generate a headline participation rate for:
- ◆ English-domiciled minority ethnics as a single group, by gender
 - ◆ White English-domiciled as a single group, by gender
112. **I also recommend that** the population denominator should be adjusted to eliminate the counting of overseas students within the England population.

I. Disaggregation of the HEIPR by Social Class

113. I have been asked to consider options for disaggregating the Higher Education Participation Rate by "Social Class".
114. "Social class" is a complex concept which has long had some currency within the United Kingdom, but which is not identified as meaningful in many other Western countries.
115. Although the term continues to be loosely used within the UK, I think it is important that I should first state that I am not aware of any current objective measure of "Social Class", (nor am I aware that the concept of "social class" is one which would be recognised by the man in the street, or which is in any statistical sense susceptible to analysis now).
116. I have therefore considered alternatives, which are based on socio-economic factors – and especially economic factors – rather than on a simple "social" hierarchy, and this accords with the national statistics developments.
117. The first issue to be considered is whether there is a basis for comparing the numerators and denominators on a consistent basis.
118. At first sight, this appears to be the case, since both the Higher Education student statistics and the ONS National population statistics now adopt a new framework : the National Statistics - Socio-Economic Classification (NS-SEC).
119. This classification was introduced for the 2001 census, and took into account new work patterns in the UK, and the changes in education levels required for, and the status of, large numbers of occupations.
120. The National Statistics Socio-Economic Classification (NS-SEC) is an occupationally based classification designed to categorise the whole adult population. It replaces Social Class based on Occupation (SC, formerly) Registrar General's Social Class and Socio-Economic Groups (SEG). The information required to create the NS-SEC is occupation coded to the unit groups (OUG) of the Standard Occupational Classification 2000 (SOC2000) and details of employment status (whether an employer, self-employed or employee; whether a supervisor; number of employees at the workplace).
121. In place of the six categories included under the former social classification, it has seven employment categories, as follows:
 - Higher managerial and professional occupations
 - Lower managerial and professional occupations
 - Intermediate occupations
 - Small employers and own account workers
 - Lower supervisory and technical occupations
 - Semi-routine occupations
 - Routine occupations

122. The following paragraphs will consider the potential applicability of this classification to the numerators and denominators which would be calculated in disaggregation of the HEIPR, and will also consider alternatives.

The numerators in relation to social class

123. The NS-SEC classification has already been used in relation to Higher Education statistics, in the context of the Performance Indicators (PIs) published in September 2004,²⁰ and a basis for disaggregating higher and lower socio-economic groupings has therefore been piloted in the context of higher education students.
124. The basis used in order to calculate the PIs was that the first three employment categories should be regarded as "high" class and the other four should be regarded as "low" class.
125. It is important to note that the categorisation distinguishes only among those in employment: the further category of "unemployed" is not ranked. This is perhaps unimportant at a time of low unemployment, but could become significant if unemployment were to increase significantly and differentially. It could also be argued that the inclusion of "Small employers and own account workers" within the lower socio-economic grouping is inappropriate; by definition, this grouping consists of several different types of worker, including for example free-lance musicians, shop-keepers, IT consultants, self-employed farmers, etc.
126. The NS-SEC classification is available by age, for those students for whom it is returned (it is a compulsory field in respect of full-time entrants to undergraduate courses under the UCAS admissions process), and therefore in principle it could provide a basis for disaggregating the HEIPR.
127. However, there are several serious objections to the use of the NS-SEC field as it is currently compiled as the numerator for analysis in relation to the HEIPR.
128. Firstly, the NS-SEC classification is obtained, for most students in both Higher Education institutions and Further Education colleges, by analysis of application records held by UCAS, using the Standard Occupational Classification of either the student (in the case of students aged over 21) or the higher income earning parent in the case of young students - and so it involves an amalgamation of two very different components. The second of these is obviously the less robust, since it is not clear that a young student would necessarily know which of his or her parents (even assuming that two are available for assessment) would have the higher income. But irrespective of that, the amalgamation of these two very different fields is contentious in the extreme, and as one well respected and highly skilled user of HE data observed to me, in the course of my consultations about this project, "it seems to me that a statistic based on [such] a mixture would be almost meaningless".

²⁰ HESA, 2004

129. Even if this were not the case, the analysis of the occupational classification is inevitably subject to a margin of error: it has been suggested to me in advice given as part of this review that the margin of error is considerable. I am not able to estimate the extent of it, but some work undertaken by the Department in conjunction with HESA in 1999²¹ included the following observation:

Use of SOC in assigning occupations

In the paper Notes on social classifications and analysis by the HEFCE (PISG9816) it is stated that it is not always appreciated how difficult it is to assign occupations to the SOC. Paragraph 7 of this paper refers to the reliability/validity of allocation of cases to social classes, based on occupation in post-enumeration checks following the 1981 Census, to be estimated at around 87% [ref. SOC Volume 3, Social Classifications and Coding Methodology, OPCS, 1991]. The paper further quotes an investigation by UCCA showing that an accuracy of 75% was achieved [ref. UCCA Statistical Supplement to the 23rd Annual Report, UCCA, 1984-85]. The main source of this inaccuracy is seen by OPCS as the vagueness or inadequacy of many of the job titles and job descriptions given in free text style responses to questions about occupation and variability in the conclusions drawn by coders from such information.

130. The paragraph above may be unduly pessimistic, being based on a fairly old bit of research, and, in consultation about this project, the Chief Executive of UCAS has drawn my attention to the level of expertise of those who now undertake the coding. However, there are further problem areas.
131. As has been noted above, there is a large number of entrants to HE courses, including between 10 and 15% of full-time undergraduates and all part-time students, who do not enter through UCAS, and for whom this data is therefore not analysed on admission: and as has been noted, the characteristics of this group are significantly different from the UCAS entrants.
132. HESA encourages institutions if possible to collect and provide this data item for full-time undergraduate students not entering through the UCAS system, although the response rate is very low. However, in relation to students following Higher Education courses in Further Education Colleges in England, the LSC's Individual Learner Record is constructed on the assumption that "this field is not required for non-UCAS entry HE learners. In this case the null value would be four zeros."²²
133. It is clear that, even within the HESA institutional constituency, institutions are not generally inclined to collect information in this area when it is not a legal requirement to do so: a survey carried out in 1998 by HESA and the DfEE reported that 87% of institutions had no intention of collecting this information for students who were not included in UCAS statistics.²³
134. The same survey revealed some serious problems with even attempting to extend the data collection to other categories, e.g.

²¹ HESA, 1999

²² LSC, 2004

²³ HESA, 1999

“institutions that did attempt to mimic the information collected by UCAS (parental occupation for under 21’s and own occupation for others), found that even young students (particularly those studying part-time) tended to enter their own (temporary or part-time) occupation when completing registration forms rather than that of their parents.”²⁴

135. There is therefore some doubt as to whether the NS-SEC classification could be used as a reasonable basis for calculating the numerator of a disaggregated HEIPR, - *unless there were to be some major change to the régime of data collection*, an issue to which I shall return.
136. The question then arises: are there any alternatives which could be used in order to calculate the numerators of a reasonably robust disaggregated HEIPR.
137. Options which I have considered include
 - ◆ “low participation neighbourhoods” (as used in the Higher Education Performance Indicators²⁵)
 - ◆ the Index of Multiple Deprivation (or some components of it)
 - ◆ as yet undefined student finance and related statistics
138. I shall address each of these in turn.

Low participation neighbourhoods

139. Students from “Low Participation Neighbourhoods” have been the subject of analysis since the First Performance Indicators for HE institutions were published in 2001. They are based on the percentage who come from a neighbourhood (as denoted by its postcode) which is known to have a low proportion of 18 and 19-year-olds in higher education.
140. There are at least three problems involved in using this measure in order to disaggregate the HEIPR:
 - ◆ The measure is explicitly calculated using the proportion of 18 and 19-year olds in higher education: while in principle it might be re-calculated using a broader age range, other issues cloud its value.
 - ◆ There are understood to be some problems in relation to neighbourhoods in London and in Scotland, which would undermine any attempted relationship between neighbourhood participation and social class.

²⁴ op.cit., page 3

²⁵ Note that the geographical analysis which has been used in HEFCE’s recent report on Young Participation in Higher Education (HEFCE, 2005) is not susceptible to use in the analysis of the broader age ranges which constitute the HEIPR, since it is based on the 15/16 year-old cohort two to three years before entry to Higher Education: and that the concept of “low participation neighbourhoods” as used in performance indicators is not in any sense related to the analysis in that report.

- ◆ The concept of “low participation” – i.e. a relative rather than an absolute value – while being meaningful in the assessment of institutional performance - is unhelpful in the context of a statistic which will be considered on a time series basis, since, if “under-representation” in one neighbourhood were to be alleviated, that would by definition lead to “under-representation” in another.

141. I therefore do not regard the concept of low participation neighbourhoods as providing a basis for calculating the numerator.

Index of Multiple Deprivation

142. A second possibility is to use the whole or parts of the “Index of Multiple Deprivation”, a concept which has been developed and implemented by the Office of the Deputy Prime Minister, and which calculates a Multiple Deprivation Index at a very detailed geographical level. The Index was first calculated in the year 2000 and has subsequently been re-calculated for 2004 by the Department of Social Policy and Social Research at the University of Oxford.²⁶

143. The Index of Multiple Deprivation 2004 (IMD 2004) is a measure of multiple deprivation at the small area level. The model of multiple deprivation which underpins the IMD 2004 is based on the idea of distinct dimensions of deprivation which can be recognised and measured separately. These are experienced by individuals living in an area. People may be counted in one or more of the domains, depending on the number of types of deprivation that they experience. The overall IMD is conceptualised as a weighted area level aggregation of these specific dimensions of deprivation.

144. The IMD 2004 contains seven Domains of deprivation:

- ◆ Income deprivation,
- ◆ Employment deprivation,
- ◆ Health deprivation and disability,
- ◆ Education, skills and training deprivation,
- ◆ Barriers to Housing and Services,
- ◆ Living environment deprivation
- ◆ Crime.

145. Each Domain contains a number of indicators. The criteria for inclusion of these indicators are that they should be 'domain specific' and appropriate for the purpose (as direct as possible measures of that form of deprivation); measuring major features of that deprivation (not conditions just experienced by a very small number of people or areas); up-to-date; capable of being

²⁶ ODPM, 2004

updated on a regular basis; statistically robust; and available for the whole of England at a small area level in a consistent form.

146. This index as a whole is currently used by the Learning and Skills Council in defining circumstances in which a "deprivation uplift" should be applied in funding Further Education students. The LSC provides a database showing for each full (inward and outward) postcode the deprivation uplift associated with that student's address.
147. Since it is based on post-codes, it would be perfectly possible to analyse the students within higher education (including FE colleges) using this discriminator, and to relate them to the denominators – the population of England in the same geographical areas.
148. However, it has been suggested to me that the Index itself may be inappropriate as the basis for determining a proxy for social class, and that perhaps some of the individual components of the Index – where these are available – should be used instead.
149. It has also been suggested that the component of the index which reports on income is the most relevant, and that other components are largely co-linear with this.
150. In fact, a detailed analysis suggests that there is not such a clear co-linearity as might at first sight appear to be the case: an admittedly small random analysis of a selection of postcodes which I carried out as part of this project showed some significant differentiation of the rankings of IMD components.
151. However, that does not undermine the argument for considering one or more of the individual components of the IMD as a basis for generating a proxy for social class – or at least socio-economic grouping. For example, family income and educational factors might be combined to form an index, if this could be calculated with sufficient frequency.
152. I conclude that the use of the IMD or its components is a serious possibility as a proxy for socio-economic group.

New student finance and related statistics

153. Finally, as another possibility, I return to an issue which I raised in my report on the definition of HEIPR last year: that is the possible use of new student finance and related statistics as a basis for considering the disaggregation of the HEIPR.
154. In the White Paper, "The future of Higher Education"²⁷, HMG set out an intention to generate new benchmarking data in relation to widening opportunity . Specifically, the White Paper noted that:

"The current ways of measuring access relate to social class, postcode and state/private school. The Government favours moving towards more sensitive indicators, looking at a student's family income, their parents' levels of education, and the average results of the school or college they attended. Data on family income could become available as early as 2004, and on school performance by 2006. We expect all of the new indicators to be in place by 2007 at the latest. We shall keep the existing criteria in place until we are satisfied with the robustness of the new data".

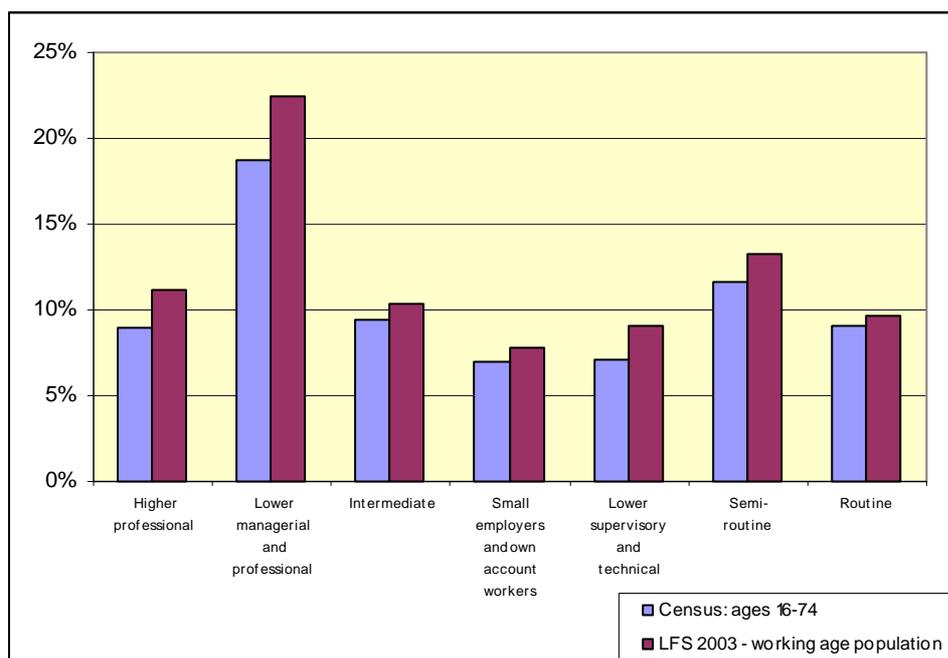
²⁷ Cm 5735

155. I have looked at this area briefly – but only briefly – because the reality is that the development and assessment of these measures is at a comparatively early stage.
156. My understanding at present is that:
- ◆ Parental Education data – which is supplied voluntarily - looks very limited so far: it may, in time, be a basis for reporting on the backgrounds of HE students.
 - ◆ Income data will only be available for those who apply for means tested support: while this will include data about such part-time students, the mechanisms for receiving and analysing the data are in their infancy, and it would be difficult at present to identify this data item as a basis for disaggregating a National Statistic
 - ◆ School performance data is not yet in place.
157. I conclude that it is too early to identify a basis for using these data items as a basis for disaggregating the HEIPR, but that the use of new measures of access to HE as a basis for disaggregating the HEIPR should be reviewed again after the full implementation of the collection of data.

The denominators in relation to social class

158. The issues in relation to the denominator include again the problem of generating data in respect of inter-censal years, which has been referred to above. This issue would arise if the NS-SEC classification were to be used as a basis for determining “social class”.
159. Again, it is conceivable that the Labour Force Survey can be used as a proxy, in order to report on annual changes. The LFS reports employment circumstances using the same basic coding frame as the census. Simply as an example of the comparability of data, the following chart shows the breakdown of the population under the census, and LFS, 2003 (Neither the years, nor the age range of the total population are comparable, and we should expect LFS to show higher values: however, the relativities are plainly very close.)²⁸

²⁸ ONS datasets used for this



160. The LFS data can be provided at single year of age.
161. Alternatively, if the Deprivation index were to be the basis of the Disaggregation of the HEIPR, there would be an issue in relation to the denominators, since these, as currently calculated are dependent on the census, and there is no basis – as far as I am aware – for estimating the deprivation index at a detailed level on an annual or otherwise frequent basis.
162. Finally, if the financial and other similar circumstances of students were to be used in the numerators, there would be a need to relate these to the population at large. The parental education level data (if it became sufficiently robust to use) could be replicated by data from the Labour Force Survey reporting on the qualifications of the workforce. The family income data - if that area were perceived to be a collectable item in relation to *all* initial entrants to Higher Education - could certainly be replicated through output from the Family Resources Survey²⁹.

Options in relation to social class

163. The options which I have identified are as follows:

Social class option 1: Require all institutions to return NS-SEC data (or occupational data) in respect of all students, both full-time and part-time

164. This is the logical option for extending the coverage of social class information from a large but limited group to all HE students.
165. It has the following advantages:

²⁹ DWP, 2004 (2)

- ◆ It generates a numerator which can, at least in theory, be directly related to a recognised denominator
- ◆ It uses a recognised definition of Socio-economic groups (although the reliability of NS-SEC in identifying “upper” and “lower” classes is as yet unproven).

166. Conversely, it has significant disadvantages:

- ◆ It would place a very large data capture burden on institutions. The arguments against increasing the data capture burden on institutions are set out fully (although not objectively and with limited evidence) in a Better Regulation Task Force report “Higher Education – easing the burden”.³⁰ Despite the considerable methodological and analytical shortcomings of that report, I doubt whether it is practical politics to argue against its basic tenets. (Nor, in fact, do I believe that the Higher Education institutions and related organisations which resisted the proposal some years ago would – or should - be willing to accept a similar imposition today).
- ◆ Even if NS-SEC were universally collected, there would still be inaccurate data in the numerator because of the doubtful quality of the conversion of occupation data into socio-economic data.

167. Although this is a theoretical option, I do not believe that it is a practical one. The data burden on institutions would in my view be totally disproportionate to the limited benefits which might be derived.

168. As regards cost, I have already identified this option as having an entirely disproportionate cost. As regards retrospective time series comparisons, these are plainly not achievable.

Social class option 2: undertake a survey of students in higher education in order to determine “social class” for missing categories

169. The second option would involve undertaking a survey of non-UCAS entrants to Higher Education on an annual or other regular basis, in order to determine the NS-SEC classification of the groups which are not covered by the administrative data collections (part-time and direct entrants among undergraduates, and all initial entrants as postgraduates, within the age range of the HEIPR).

170. This option would have the advantages of filling the data gap in a less resource intensive way than option 1, and also serving to provide a statistic which is based on a recognised definition of socio-economic groups.

171. Conversely, as a disadvantage, it would have a cost implication (dependent on the frequency of the survey): I would regard this as being a medium cost implication, and it is one which the

³⁰ Cabinet Office, 2002

Department could control, since it would commission the survey, and could determine both the coverage and the frequency.

172. Plainly, it would be impossible to produce retrospective time series comparisons.

Social class option 3: carry out a periodical survey of the new entrant population (including UCAS entrants)

173. A further option is to introduce a wider ranging survey, including not simply the non-UCAS entrants, but also those who enter through the UCAS system, in order to generate a consistent dataset of new entrants to HE.
174. An advantage of this option is that it could investigate the student's socio-economic group on a more objective basis than is possible through the conversion of occupations (either student's or parent's, depending on age) as is currently undertaken.
175. Under this option, the questions included in the survey could be generated to coincide with the national population statistics; and there is a range which might be included. For example, students might be asked to identify their own NS-SEC classification, which might then be related to the national population statistics; or alternatively the question might be more explicitly concerned with family income.
176. However, there is plainly an issue here about whether the desired outcome is a "disaggregation of the HEIPR" or a different measure of participation by social class. If the former is the case, then it follows that the survey questions should be explicitly related to data which can be derived in respect of the national population at each year of age, in order to generate the disaggregation of the HEIPR.
177. If this option were to be adopted in relation to social class, it might be linked also with the arrangements to collect information about disability: and indeed both of them might potentially be linked with another survey which is carried out frequently, i.e. the Student Income and Expenditure Survey; although care would need to be taken to ensure that samples were adequate to meet all the needs of such an omnibus survey.

Social class option 4: use a proxy for social class derived from the Index of Multiple Deprivation

178. I have discussed earlier the possibility of introducing a proxy for social class, and in principle I find this an attractive proposition.
179. A possible proxy might be derived from the Index of Multiple Deprivation, as described above, which is susceptible to detailed geographical analysis.
180. The advantages of this option would be that it would draw on detailed data which can be closely aligned with the domicile of each student: i.e. it can be derived objectively, without reference to the occupation of the student or the parent.
181. In many ways, I find myself inclined to suggest that this would be the best option: but I recognise that there is no formal basis for updating the data on an annual basis - the IMD calculations will be updated on a less than annual basis - and that therefore

realistically it might be found that statistics would need to be calculated annually using out-dated bases – which, as I have noted elsewhere, is one of the major criticisms of currently published statistics about social class. However, I believe that the underlying annual shifts in the denominators would be likely to be small.

182. In cost terms, this would be a very low cost option; and it would be susceptible to some retrospective time series comparisons.

Social class option 5: make use of new participation measures

183. The final option here is to make use of the new participation measures foreshadowed in the Higher Education White Paper and described above.
184. As I have indicated, I believe it is simply too early to make any judgement about these: we are currently in a pilot stage, at best, or a definitional stage in some respects.
185. There is no basis for assessing the benefits – or the costs – involved in using these indicators as a basis for disaggregating the HEIPR, and it would be wise to wait until the statistics have been developed and collected before making any such judgements.

Recommendations in relation to Social class

186. I have identified above some of the problems of identifying a measure of "social class" or "socio-economic group" which might be used in the disaggregation of the HEIPR. In particular, I have noted that the use of SOC data in relation to UCAS entrants only is likely to distort the data considerably. Therefore, given the currently available data in respect of HE students, it is simply not possible to disaggregate the HEIPR using the NS-SEC classification to relate the entrants to higher education to the population at large, and so, if there is a policy imperative to report on this area, some other methodology must be considered.
187. I should first explicitly say that I do not recommend that the data collection in relation to the student record should be extended by requiring institutions to provide data in relation to occupation grouping or socio-economic group for all new entrants to HE. The burden involved would in my view be grossly disproportionate to any benefit which might be achieved.
188. If HMG is of the view that it wishes to disaggregate the HEIPR by social class, I recommend that it considers three options:
- ◆ for the numerators, the existing administrative data reporting on the NS-SEC classification might be supplemented by periodic surveys, in order to ensure that there is an adequate representation of the socio-economic distribution of the new entrant population; the denominators should be refreshed annually by using the Labour Force Survey data to adjust the previous census population.
 - ◆ a survey be undertaken periodically of the new entrant population generally; the denominators should be refreshed annually by using the Labour Force Survey data to adjust the previous census population.
 - ◆ The use of the Index of Multiple Deprivation, or some components of it, as a proxy for socio-economic group. This is my favoured option.
189. However, I would regard this as a short-term expedient, and I recommend that the possible use of new measures of access to HE (such as parental educational background) as a basis for disaggregating the HEIPR should be reviewed again after the full implementation of the collection of data.
190. If the Government wishes to monitor the social composition of only the young entrant population into Higher education, I recommend that it should use a methodology similar to or derived from the HEFCE methodology for constructing Young Participation Rates, and should make clear that these measures relate solely to young entrants, and are not comparable to the HEIPR. This is, in my view, a realistic way of replacing the API, to which I have referred above.

191. It should however be noted that this is NOT a low cost option: the conversion of a long-term research project into an annual data collection activity will have some considerable resource implications.

J. Disaggregation of the HEIPR by Disability

192. This area is perhaps the most difficult of the four major disaggregations which are being addressed by this project.

193. The difficulty arises for four reasons:

- ◆ In common with the concept of social class, there is no generally recognised definition of disability, and also no general taxonomy of subsets of disability.
- ◆ In common with the concept of ethnicity, disability within the population statistics is essentially self-assessed
- ◆ Unlike both ethnicity and social class, the coding frames which are used in national statistics and higher education statistics are significantly at variance
- ◆ And even within the Higher Education constituency, there is no consistency of definition as between the HEIs and the Further Education Colleges which provide Higher Education courses.

The numerators in relation to disability

194. Within the HESA student record there are two fields which relate to disability.

195. The first is the student's self-assessment of any disability.

196. The coding frame (in relation to 2004-05 and subsequently) is shown below:

00	No known disability.
01	Dyslexia.
02	Blind/are partially sighted.
03	Deaf/have a hearing impairment.
04	Wheelchair user/have mobility difficulties.
05	Personal care support.
06	Mental health difficulties.
07	An unseen disability, e.g. diabetes, epilepsy, asthma.
08	Multiple disabilities.
09	A disability not listed above.
10	Autistic Spectrum Disorder.
98	Information not sought.
99	Not known.

197. Code 10 was not used in previous years, and therefore no analysis is available which would capture the use of this code retrospectively.

198. In the most recently collected HESA statistics, 5% of all new undergraduate students aged up to 30 declared themselves to be disabled.

199. Within the group which identified itself as disabled, the percentage attributable to each disability was as follows:

Dyslexia.	51.9%
Blind/are partially sighted.	2.7%
Deaf/have a hearing impairment.	3.7%
Wheelchair user/have mobility difficulties.	2.4%
Personal care support.	0.2%
Mental health difficulties.	3.7%
An unseen disability, e.g. diabetes, epilepsy, asthma.	19.1%
Multiple disabilities.	5.7%
Autistic Spectrum Disorder.	0.3%
A disability not listed above.	10.4%

200. Note that two disabilities, dyslexia and “an unseen disability” between them constitute 70% of the reported disabilities.
201. The second field which is available within HESA statistics reports on receipt of the Disabled Students Allowance (DSA).
202. The DSA is available to students (full-time and part-time) who can show that they have a disability, medical condition or specific learning difficulty which affects their ability to study. It is explicitly only available to cover costs incurred as a result of the disability, medical condition or learning difficulty. In general, students are not required to inform their university/college that they are eligible to apply for DSA (although they will need to do so if they are Open University students, since the OU administers the scheme for its own students.) The allowance is only payable following the completion of a needs assessment.
203. In 2002/03, the DSA field in the HESA student record was completed in respect of 98% of first year students aged 30 or less, and 2% of these were reported to be in receipt of the DSA.
204. The relationship between the 2% in receipt of DSA with the 5% self-assessed as disabled obviously presents some problems in identifying an adequate basis for disaggregating the numerators within the HEIPR.
205. However, a further problem arises from the fact that the LSC Individualised Learner Record (ILR), which provides data about students studying at HE level in Further Education Colleges in England, is not directly consistent with the HESA record. The most relevant field in the ILR asks for disability data analysed as follows:

01	visual impairment
02	hearing impairment
03	disability affecting mobility
04	other physical disability
05	other medical condition (for example epilepsy, asthma, diabetes)
06	emotional/behavioural difficulties
07	mental ill health
08	temporary disability after illness (for example post-viral) or accident
09	profound complex disabilities
90	multiple disabilities
97	other
98	no disability
99	not known/information not provided

206. While this coding frame has some components in common with the HESA coding frame set out above, it does not include Dyslexia, which constitutes over 50% of the disabilities reported by young people through the HESA student record. Instead, the ILR collects information about dyslexia in a separate field, "Learning difficulty", the coding frame for which is as follows:

01	moderate learning difficulty
02	severe learning difficulty
10	dyslexia
11	dyscalculia
19	other specific learning difficulty
90	multiple learning difficulties
97	other
98	no learning difficulty
99	not known/information not provided

207. It should be noted that if a student has dyslexia alongside another learning difficulty of equal importance, that students would be coded 90 – "other".

208. While it is not out of the question, I have a feeling that this issue, combined with other respects in which the two data collections are not compatible makes it difficult to envisage a robust aggregation of data into the numerators.

The denominators in relation to disability

209. It should initially be noted that there is no longer any concept of "registered disabled" .

210. The initial source of population statistics is, naturally, the Census, within which, in 2001, there was one question relevant to this area:

The term disability is used to refer to limiting long term illness or disability which restricts daily activities. It is calculated from a 'Yes' response to the question in the 2001 Census: "Do you have any long-term illness, health problem or disability which limits your activities or the work you can do?"

211. It does not appear to be the case that this definition can in any way be related to the potential numerators identified above.

212. The Disability Discrimination Act also includes a definition, which is summarised by the Disability Rights Commission (DRC) as follows:

The Disability Discrimination Act (DDA) protects disabled people. The Act sets out the circumstances in which a person is "disabled". It says you are disabled if you have:

- a mental or physical impairment*
- this has an adverse effect on your ability to carry out normal day-to-day activities*
- the adverse effect is substantial -the adverse effect is long-term (meaning it has lasted for 12 months, or is likely to last for more than 12 months or for the rest of your life).*

There are some special provisions, for example:

- if your disability has badly affected your ability to carry out normal day-to-day activities, but doesn't any more, it will still be counted as having that effect if it is likely to do so again*
- if you have a progressive condition such as HIV or multiple sclerosis or arthritis, and it will badly affect your ability to carry out normal day-to-day activities in the future, it will be treated as having a bad effect on you now*
- past disabilities are covered.*

What are "normal day-to-day activities"?

At least one of these areas must be badly affected:

- mobility*
- manual dexterity*
- physical co-ordination*
- continence*
- ability to lift, carry or move everyday objects*
- speech, hearing or eyesight*
- memory or ability to concentrate, learn or understand*
- understanding of the risk of physical danger.*

213. There are further definitions which are relevant, and the complexity of this area has led the Department of Work and Pensions to produce a "User's guide to disability estimates and definitions".³¹

214. This guide observes that:

- DWP's core estimates of the prevalence of disability are as follows:*
 - There are about 10 million adults in Britain covered by the Disability Discrimination Act (1995) from the Family Resources Survey 2002*
 - There are about 6.9 million people of working age with a long-term disability from the Labour Force Survey spring 2003.*
- (Long-Term disability is the measure used for the DWP's Public Service Agreement on employment and disability and consists of people who have a disability that affects the amount or type of work they could do (work limiting disability) and people with a disability likely to be covered by the DDA.*

³¹ DWP, 2004

215. The variation between the two different definitions set out above exemplifies the problem of determining a denominator for any disaggregated HEIPR in this area. It is however perhaps the case that the definition of Long-term disability used within the Labour Force Survey is the better for our purposes, since it relates to people of working age and includes the DDA definition.
216. However, it needs to be recognised that there is currently no clear (or even approximate) relationship between the definitions of the numerators and the potential denominators.
217. For the future, I understand the Office for National Statistics is working with other government departments towards a set of harmonised questions on disability. The focus is on data collected through surveys rather than administrative data and therefore would have more relevance to the denominators than to the numerators: and there is no reason to think that the new denominators would necessarily be more susceptible to comparison with the disability data collected by universities and colleges than is currently the case. The timescale is to propose a set of harmonised questions for agreement in spring 2006; and I recommend that the issue should be studied further then.
218. I was sorry that, despite several requests, the Disability Rights Commission was unable to provide comments on this draft report.

Options in relation to disability

219. Having considered the available data and having consulted the project Steering Group, I am of the view that there is no realistic likelihood of being able to identify a basis for disaggregating the HEIPR by disability through available datasets.
220. This is because the concept of "disability" can be – and is - defined in many different ways, and there seems to me to be no objective measure of disability which can realistically be applied to both the numerators and the denominators.
221. In the absence of such disaggregation, I suggest that there are only three viable options.

Disability option 1 – the numerator-only model

222. The first option would be to publish simply the proportion of the initial participants in Higher Education who are understood to be disabled: i.e. to adopt the "numerator only" approach described above. Since the HESA and LSC datasets are inconsistent in relation to Disability, a pragmatic approach would be to assume that the HESA dataset, which represents over 95% of the HEIPR students in England is representative.
223. On this basis, it would be possible to publish information along the following lines (the figures being representative, not actual)³²:

³² In fact the figures represent the proportion of students of known disability status among first year undergraduate entrants in 2002/03 in the UK, which are likely to correlate closely with the HEIPR figures when calculated.

Percentage of new entrants having a reported disability

	Column A	Column B	Column C	Column D	Column E
	2002/03	2001/02	2000/01	1999/ 2000	1994/95
Percentage reported to have a disability	5.4%	5.1%	4.5%	4.5%	4.2%
----- within which:					
Dyslexia	2.2%	1.9%	1.6%	1.5%	0.6%
Blind/Partially sighted	0.2%	0.2%	0.1%	0.2%	0.2%
Deaf/Hearing impairment	0.3%	0.3%	0.3%	0.3%	0.3%
Wheelchair user/Mobility difficulties	0.3%	0.3%	0.2%	0.2%	0.3%
Personal care support	0.0%	0.0%	0.0%	0.0%	0.0%
Mental health difficulties	0.3%	0.2%	0.2%	0.2%	0.1%
An unseen disability (for example, diabetes, epilepsy or asthma)	1.1%	1.2%	1.2%	1.3%	2.1%
Multiple disabilities	0.4%	0.4%	0.3%	0.3%	0.2%
Other disability	0.7%	0.6%	0.6%	0.6%	0.4%

224. This approach has the advantage that it provides straightforward and transparent data, within its definitional limitations.
225. While it has the disadvantage that it does not take into account changes in the underlying population of England, it might be reasonable to assume a comparatively slow rate of change in the national population disability statistics (however defined) within the 17-30 age range.
226. Whether or not this was the case, the presentation of Higher Education statistics on this basis would facilitate comparison with national statistics at a detailed level of specific disability.
227. It is also relevant to note that, within the sample figures quoted above, two categories (Dyslexia and "An unseen disability") show significant movement over the period under review, in opposite directions, and between them constitute over 60% of the reported disabilities among HE students – there may be some issue here about categorisation of disabilities in the past.
228. The cost implication of this option would be low, since the data can be derived from existing collections with no change.
229. As is exemplified above, it is possible to build a retrospective time series on this basis

Disability option 2 – a survey approach

230. The second option would be to carry out a periodical survey of students within HE-providing institutions in order to relate them to the nationally available statistical data.

231. This approach would have the advantage of generating high quality data which could be compared with national population estimates. As one colleague who was consulted during this project observed:

"I doubt that it would ever be possible to collect disability data through any administrative sources to the same degree of quality as a survey. We would need trained interviewers enrolling learners at colleges/universities. Participation in surveys is optional, so they suffer from unit non-response - people choosing not to participate in the survey. However, people who participate in surveys tend to answer most of the questions, so they get less item non-response. In administrative collections, people don't have the option to not participate. However, it is not compulsory for a learner to give information for every data item."

232. On the other hand, there are disadvantages in terms of costs.

233. The approach would be highly resource-intensive, since it would involve a survey of HE students across the UK, and, in view of the comparatively small proportion of students aged 17 to 30 who are disabled, the survey would need to be directed to a large number of entrants.

234. If this option were to be adopted it would be important to ensure that an appropriate definition was being adopted. As noted above (page 41) the definition used within the Labour Force Survey is perhaps the most appropriate current definition, but it is understood that new definitions will be agreed in Spring 2006 for subsequent surveys.

235. It should also be noted that, obviously, retrospective time series comparisons would be impossible.

Disability Option 3 – change in data collection arrangements

236. The final option would be to add an additional field into the data collected by HESA and the other data collection organisations, which would be more closely aligned to the national population figures which are represented in the denominators.

237. The option has the advantage of simplicity.

238. On the other hand, it has the disadvantage of imposing a further burden on institutions by requiring them to collect an item of data which is not of any immediate use to the institutions themselves (unlike the two current fields, both of which are of relevance to institutions as providers of services.)

239. In cost terms, it is probably an intermediate option, but note that the costs would be incurred primarily by the institutions, and not by the central bodies which would benefit from the provision of the data.

240. Plainly, retrospective time series would be impossible to generate.

Recommendations in relation to Disability

241. In relation to disability I recommend that

- ◆ the HEIPR should not be disaggregated by disability for the time being
- ◆ consideration be given to the publication of an annual series of statistics based solely on the initial entrants to Higher Education, and showing the proportion falling into each disability category (as exemplified in the table on page 42.)
- ◆ further consideration should be given in 2006 to the possibility of carrying out a survey of students in Higher Education providing institutions or of amending data collection arrangements in order to capture data about HE entrants which may be directly aligned with the national population survey definitions which are due to be agreed in that year.

K. Disaggregation of the HEIPR by Region

- 242. I was asked, as a supplement to the initial brief for this project, to consider the feasibility of disaggregating the HEIPR by Region.
- 243. In this context “region” implies Government Office Region (GOR) in England.

The numerators in relation to regional disaggregation

- 244. There are, I believe no issues in connection with the regional disaggregation of numerators of the HEIPR.
- 245. The Government Office Region of the student’s home address can be derived robustly from the postcode field within the HESA record and the records of students in Further Education Colleges, and can be attributed to the individual age on entry.

The denominators in relation to regional disaggregation

- 246. It is in the denominators that there is potentially a problem, although it is one which I believe can be surmounted.
- 247. As I observed in my earlier report about the definition of the HEIPR the problem here is that the population estimates which provide the denominators at each age use the term-time address of students rather than the home address. If the statistic were to be disaggregated by region, therefore, the relationships of the numerators and the existing denominators would only be valid if the inflow and outflow to each region (and indeed at each age) were approximately the same, as is the case in relation to migration between England and the other countries of the United Kingdom.
- 248. That this is not the case is well-known. Some regions (notably London) are net importers of students while others (e.g. the Eastern Region and the South-East) are net exporters. An illustration of the differences is to be found (in relation to the academic year 1998/99) in the following table³³:

³³ Ramsden, 2001

Comparison of (UK domiciled) HE student numbers and population

Region	Percentage of HE students	Percentage of population	<i>Percentage point difference</i>
East Midlands	7.3%	7.0%	0.2%
Eastern	6.0%	9.0%	-3.0%
London	16.4%	12.1%	4.3%
North East	4.5%	4.4%	0.1%
North West	11.2%	11.7%	-0.5%
Northern Ireland	2.5%	2.8%	-0.3%
Scotland	10.2%	8.7%	1.5%
South East	11.3%	13.5%	-2.2%
South West	7.1%	8.3%	-1.2%
Wales	5.7%	5.0%	0.7%
West Midlands	8.2%	9.0%	-0.8%
Yorkshire & Humberside	9.8%	8.5%	1.3%
Grand Total	100.0%	100.0%	

249. In absolute terms, the scale of this issue is exemplified by the following table which shows the number of people at critical ages who were reported in the 2001 census as living away from home:

Students living away from home at relevant ages (2001 Census) ³⁴	
18	37257
19	98816
20	113510
21	95933
22	52093
23	20425
24	8864

250. Plainly, any option for the disaggregation of the HEIPR by Region needs to have regard to this issue, and to adjust for it.

Options in relation to regional disaggregation

251. There are, it seems to me, few options to discuss here, except the methodology for adjusting the population denominators. There are two theoretical ways to overcome this issue. The first is to arrange for the generation of an ONS population estimate which adjusts student addresses to their home address, rather than the term-time address where this is different. The second would be simply to use the known home and study regions of the entrants to higher education and to adjust the national population estimates accordingly.

252. I understand from the Office for National Statistics that the first of these two options is not practicable. In order to achieve it, I am advised that ONS would need an extract from the 2001 Census

³⁴ National Statistics

with students at their home address rather than their term-time address, and then to

- a) remove those students at their term-time address who have a different home address
 - b) add these students back in at their home address. While basic information on students at their home address (where this is not their term-time address) exists, this is thought to be an undercount as it depends on the other people at that address including the person.
253. I therefore suggest that, in order to generate an appropriate denominator of any regionally aggregated version of the HEIPR, the population estimates by year of age should be adjusted according to the numbers of new entrants at each age domiciled in one region and studying in another region.
254. I would however caution against any assumption that the disaggregation by Region could cascade into a disaggregation by significantly smaller geographical units, since population numbers in the numerators at particular years of age might lead to unreliable results.

Recommendation in relation to the disaggregation of the HEIPR by Region

255. I recommend that the HEIPR should be disaggregated by Region of England, while noting that it should not be assumed that the methodology can be applied at a significantly lower level of geographical disaggregation.

L. Other issues arising from this project

256. Is relevant here to mention one related issue which has arisen, and on which I have been asked to report - that is data about Religion; and one other issue which I feel that it appropriate to review which is the general issue of the relationship between the HEIPR and other published statistics.

Religion

257. The Steering Group of this project at its second meeting on 10th December 2004 minuted that "as there has been interest in the participation of Muslims....." I should explore the scope for any measure by religion.
258. In relation to Higher Education students, religion is collected only for students resident in Northern Ireland, since this is a legal requirement in relation to equal opportunity statistics in Northern Ireland.
259. In relation to the population at large, religion is collected under several different definitions by the Census, the Labour Force Survey, and also by the Home Office Citizenship Survey.
260. Definitions and explanatory notes extracted from the ONS website are contained in Appendix 3.
261. In the absence of any data concerning the religion of English-domiciled students, it is clearly impossible to derive a disaggregation of the HEIPR by religion, and I am sure that this is not in fact sought.
262. I should however also note that there is not, within the Higher Education data, any basis for determining participation by religion in England.

The relationship of the HEIPR with other participation statistics

263. The Higher Education Initial Participation Rate is explicitly a measure of participation among the English-domiciled population, and relates to people aged up to 30. It is a valid measure of how the Government is faring against its target of increasing participation in HE towards 50% of those aged 18-30 by the end of this decade.
264. It is entirely different from some other participation figures. It is not, for example, in any sense comparable with the Scottish Executive's published participation rate, nor with the OECD participation rate published in Education at a Glance, since each of these is dependent on its own methodology, and neither is comparable with the HEIPR.
265. The HEIPR is also different from the Young Participation Rates used in the recent HEFCE publication to which I have referred earlier³⁵, since they are only applicable to very young entrants,

³⁵ HEFCE, 2005

being based on the population as at the end of compulsory education, and assume no significant mobility in the intervening years from (generally) 16 to 18.

266. It should also be noted that there are some aspects of participation which are captured at institutional level, and reported in the annual performance indicators for Higher Education. For example, social class, ethnicity and disability all feature within these indicators. The methodology associated with these, however, is not transferable to national participation statistics, since they are not associated with a denominator to enable tracking by time: they are quite simply a basis for comparing institutions at one point in time. A time series exists and is published annually by HESA, but that does not provide the basis for disaggregation of the HEIPR.
267. The purpose of this report is to identify the extent to which the HEIPR may be disaggregated according to social class, ethnicity and disability: however, I think it appropriate to emphasise that the monitoring of change in these areas may to some extent be accomplished by other mechanisms than simply the disaggregation of the HEIPR, and indeed it may prove to be the case that other mechanisms may be more meaningful.

M. Conclusions and recommendations

268. There are two general issues which I would like to flag before returning to the specific outcomes of my study.
269. The first issue is quite simply whether it is wise to consider using the HEIPR as a basis for assessing participation in Higher Education by specific and comparatively small sub-groups of the population. The HEIPR was devised as an instrument for monitoring overall participation in Higher Education among the population of England aged 30 and below. It has been (I think correctly) decided that the HEIPR should be disaggregated by gender. In this report I have identified that disaggregation by Government Office Region of England would be comparatively straightforward. However, I have identified some difficulties in relation to the other three aspects of disaggregation which I have been asked to consider. I do wonder whether we are confusing specific policy issues with more general questions about participation.
270. The second issue is about the basis on which information concerning higher education is collected. There is in place a very high quality system for collecting administrative information about higher education. It should not necessarily be assumed that an extension of the HESA/LSC administrative data collections would be the best way of providing statistical estimates which might inform policy decisions in areas where some subjectivity in relation to input and analysis is needed: surveys are intrinsically more receptive to subtle issues than administrative data collections: and the extension of administrative data collection into more complex areas will inevitably lead to increasing burdens on institutions, and potentially greater intrusion into the privacy of individuals.
271. For convenience I summarise in the following paragraphs my recommendations arising from this project

General recommendations

272. I recommend that the DfES should not use data available only for UCAS entrants as a basis for any general statistics about participation in Higher Education within England or the UK.
273. As regards the Age Participation Index, I recommend that, if the Department continues to provide information based on the API, it should develop appropriate and strong caveats which should accompany any such information. I support the Department's decision to stop the calculation and publication of the Age Participation Index by social class.

Recommendations in respect of ethnicity

274. I recommend that if the Government wishes to disaggregate the HEIPR by Ethnicity it should commission work to generate a basis for assessing the change each year in the population at individual

year of age and by individual ethnic group, using the cohort projection approach.

275. In the interim, and until the work referred to above is complete, I recommend that, if the Government wishes to have a simple disaggregation of the HEIPR by ethnicity, then it should use the Labour Force Survey data to generate a headline participation rate for:

- ◆ English-domiciled minority ethnics as a single group, by gender
- ◆ White English-domiciled as a single group, by gender

276. I also recommend that the population denominator should be adjusted to eliminate the counting of overseas students within the England population.

Recommendations in respect of Social class

277. I recommend that the data collection in relation to the student record should not be extended by requiring institutions to provide data in relation to occupation grouping or socio-economic group for all new entrants to HE.

278. If HMG is of the view that it wishes to disaggregate the HEIPR by social class, I recommend that it considers three options:

- ◆ for the numerators, the existing administrative data reporting on the NS-SEC classification might be supplemented by periodic surveys, in order to ensure that there is an adequate representation of the socio-economic distribution of the new entrant population; the denominators should be refreshed annually by using the Labour Force Survey data to adjust the previous census population.
- ◆ a survey be undertaken periodically of the new entrant population generally; the denominators should be refreshed annually by using the Labour Force Survey data to adjust the previous census population.
- ◆ The use of the Index of Multiple Deprivation, or some components of it, as a proxy for socio-economic group. This is my favoured option..

279. However, I would regard this as a short-term expedient, and I recommend that the possible use of new measures of access to HE (such as parental educational background) as a basis for disaggregating the HEIPR should be reviewed again after the full implementation of the collection of data.

280. If the Government wishes to monitor the social composition of the young entrant population into Higher education, I recommend that it should use a methodology similar to or derived from the HEFCE methodology for constructing Young Participation Rates, and should make clear that these measures relate solely to young entrants, and are not comparable to the HEIPR. (It should also be

noted that this is NOT a low cost option: the conversion of a long-term research project into an annual data collection activity will have some considerable resource implications.)

Recommendations in respect of disability

281. I recommend that the HEIPR should not be disaggregated by disability for the time being.
282. I recommend that consideration be given to the publication of an annual series of statistics based solely on the initial entrants to Higher Education, and showing the proportion falling into each disability category (as exemplified in the table on page 42.)
283. I recommend that further consideration should be given in 2006 to the possibility of carrying out a survey of students in Higher Education providing institutions or of amending data collection arrangements in order to capture data about HE entrants which may be directly aligned with the national population survey definitions which are due to be agreed in that year.

Recommendations in respect of disaggregation by Region

284. I recommend that the HEIPR should be disaggregated by Region of England, while noting that it should not be assumed that the methodology can be applied at a significantly lower level of geographical disaggregation.

Conclusion

285. I have been pleased to carry out this review on behalf of the Department of Education and Skills.
286. The level of and measurement of participation in Higher Education are major issues of concern at the moment, and understandably the Government wishes to find the best method of measuring participation at as fine a level of detail as is realistic. I have suggested in this paper that there are both opportunities and limitations in respect of the presentation of the Higher Education Initial Participation Rate at a disaggregated level: I hope that this report will assist in the formulation of improved strategies for the publication of meaningful information.

Brian Ramsden
May 2005

N. Appendices

Appendix 1: Membership of the Steering Group for the project

Dr Stella Mascarenhas-Keyes, Chair and Project Manager

Gregory Boone
Ian Mitchell
Ron Allen
Steve Hamilton
Adam Hatton

Appendix 2: Individuals consulted during this project

In addition to the members of the Steering Group, I have consulted and been assisted by many people in the course of this short project, and I have attempted to list most of them here. I should emphasise that, in several instances, the individuals were consulted only about a specific aspect of this work, and therefore it should not be implied that they are in agreement with or even aware of the overall thrust of my findings.

Mr Andrew Battarbee	DfES
Mr James Halse	DfES
Ms Karen Hancock	DfES
Ms Anna Heyworth	DfES
Ms Jackie Jackson	DfES
Mr Andrew Longton	DfES
Ms Stephanie Robson	DfES
Mr Mark Gittoes	HEFCE
Mr John Thompson	HEFCE
Ms Carole Barrington	HESA
Professor Robin Sibson	HESA
Mr Jonathan Waller	HESA
Ms Jane Wild	HESA
Mr Colin Stronach	LSC
Ms Ruth Fulton	ONS
Mr Anthony McClaren	UCAS
Ms Liz Peters	UCAS
Professor Philip Rees	University of Leeds

Some other individuals and organisations were invited to comment on my draft report but did not submit comments.

Appendix 3: Issues concerning Religion

As noted in the text, this is not an issue which is technically covered by the terms of reference of this study, but in view of the Steering Group's minute of 10th December, I have included this brief summary of available data, which is derived from the ONS report "Focus on Religion"³⁶:

There are a number of key sources for religion data which have been used in the "Focus on Religion" report. The exact question wording from each of these sources is detailed below:

2001 England and Wales Census

What is your religion?

- None
- Christian (including Church of England, Catholic, Protestant and all other Christian denominations)
- Buddhist
- Hindu
- Jewish
- Muslim
- Sikh
- Any other religion (please write in)

2001 Scotland Census

What religion, religious denomination or body do you belong to?

- None
- Church of Scotland
- Roman Catholic
- Other Christian (please write in)
- Buddhist
- Jewish
- Muslim
- Sikh
- Another religion (please write in)

What religion, religious denomination or body were you brought up in?

- None
- Church of Scotland
- Roman Catholic
- Other Christian (please write in)
- Buddhist
- Jewish
- Muslim
- Sikh
- Another religion (please write in)

³⁶ ONS, 2004

2001 Northern Ireland Census

Do you regard yourself as belonging to any particular religion?

- Yes
- No

If yes

What religion, religious denomination or body do you belong to?

- Roman Catholic
- Presbyterian Church in Ireland
- Church of Ireland
- Methodist Church in Ireland
- Other, please write in

If no

What religion, religious denomination or body were you brought up in?

- Roman Catholic
- Presbyterian Church in Ireland
- Church of Ireland
- Methodist Church in Ireland
- Other, please write in

Labour Force Survey (2002 onwards)– Great Britain

What is your religion, even if you are not currently practising

- Christian
- Buddhist
- Hindu
- Jewish
- Muslim
- Sikh
- Any other religion
- Or no religion at all

Labour Force Survey (2002 onwards)– Northern Ireland

What is your religious denomination?

- Catholic
- Presbyterian
- Church of Ireland
- Methodist
- Other Protestant
- Other religion
- No denomination

Home Office Citizenship Survey

What is your religion?

- No religion

Christian (including Church of England, Catholic, Protestant, and all other Christian denominations)
 Buddhist
 Hindu
 Jewish
 Muslim
 Sikh
 Atheist/agnostic
 Any other religion

Religion data for Great Britain

The way in which people answer questions on religion is very sensitive to the exact question wording. This is particularly true for people who have a loose affiliation with a religion. Slight differences in question wording can produce large differences in the proportion of people who say they are Christians or have no religion, although the proportion of people from other religions tends to be more stable.

Since the Labour Force Survey asked the same question on religion across Great Britain, data from Scotland can be easily compared with data from England and Wales. The proportion of people who said they were Christian was very similar in Scotland to the proportion that answered this way in England and Wales.

The 2001 Census in Scotland asked two questions on religion: current religion and religion of upbringing. Neither of these was the same as the Census question asked in England and Wales.

The following table compares responses to the LFS religion question with responses from Census 2001:

Religion	Percentages				
	England & Wales		Scotland		
	Census ¹	LFS ²	(current) ³	(upbringing) ⁴	LFS ²
Christian	77.7	79.9	68.9	79.2	81.9
Buddhist	0.3	0.3	0.1	0.1	0.2
Hindu	1.2	1.1	0.1	0.1	0.2
Jewish	0.5	0.5	0.1	0.2	0.2
Muslim	3.2	3.0	0.9	0.9	0.8
Sikh	0.7	0.6	0.1	0.1	0.0
Any other religion	0.3	0.8	0.6	0.2	0.6
No religion	16.1	13.8	29.1	19.1	16.1
All population excluding 'not stated'	48,031,258	148,540	4,783,950	4,639,149	15,223

Notes to table:

1 2001 Census question in England and Wales asked respondents: 'What is your religion?'

2 Religion question on Labour Force Survey asked respondents: 'What is your religion even if you are not currently practising?' Data are for 2003/2004.

3 2001 Census question in Scotland asked respondents, firstly: 'What religion, religious denomination or body do you belong to?'

4 2001 Census question in Scotland asked respondents, secondly: 'What religion, religious denomination or body were you brought up in?'

Appendix 4: Aggregated minority ethnic groups

The following is an additional table from DfES Research Report 552, "Why the Difference?" which sets out a tentative HEIPR for some aggregated groupings by ethnicity.

Table 4.2: Higher Education Initial Participation Rates (HEIPRs) for English domiciled first- time entrants (full- and part-time) to HE courses (in universities and colleges), by broad ethnic/gender group, 2001/02

Ethnic group	Male	Female	All
White	34	41	38
All Minority ethnic	55	58	56
- Asian or Asian British	62	59	60
- Black or Black British	55	66	61
- Chinese or Other Ethnic	47	50	49
- Mixed Ethnic	35	44	40
All (with known ethnicity)	37	43	40

Note (from "Why the Difference") : The student coverage here is different from Table 4.1. It is based on number of entrants to HE (not undergraduate study only, though most will be) in 2002/02, domiciled in England but studying anywhere in UK.

Appendix 5: References

Connors, H., Tyers, C., Modood, T., and Hillage, J., 2004: Research Report 552, "Why the Difference? A Closer Look at Higher Education Minority Ethnic Students and Graduates", DfES Research Report 552, 2004.

Cabinet Office, 2002: "Higher Education – easing the burden", report of the Better Regulation Task Force, available at <http://www.brtf.gov.uk/reports/highedentry.asp>

Cm 5735: "The future of Higher Education", DfES 2003

DWP, 2004: "User's guide to disability estimates and definitions", Department of Work and Pensions, March 2004, available at <http://www.dwp.gov.uk/asd/asd5/ih2003-2004/IH128userguide.pdf>

DWP, 2004 (2): "Family Resources Survey, 2002/03", available at http://www.dwp.gov.uk/asd/frs/2002_03/pdfonly/frs_2002_03_report.pdf

Haskey, J. (ed) 2002: "Population projections by ethnic group: a feasibility study", ONS, 2002

HEFCE, 2005: "Young participation in higher education", HEFCE, January 2005, available at http://www.hefce.ac.uk/pubs/hefce/2005/05_03/

HESA, 1999: "Collection of Social Class Data for all HE Students: Feasibility Study": unpublished report.

HESA, 2004: Higher Education Statistics Agency, Performance indicators of Higher Education Institutions, <http://www.hesa.ac.uk/pi/default.htm>

LSC, 2004: Learning and Skills Council, Individualised Learner Record Specification 2004/05, available at <http://www.lsc.gov.uk/National/Partners/Data/DataCollection/ILR/ILRSpecification/ILRSpec200405.htm>

ONS, 2004: "Focus on Religion", Office for National Statistics, 2004, available at <http://www.statistics.gov.uk/focuson/religion/default.asp>

ODPM, 2004: Office of the Deputy Prime Minister, The English Indices of Deprivation 2004: Summary (revised), http://www.odpm.gov.uk/stellent/groups/odpm_urbanpolicy/documents/downloadable/odpm_urbpol_029247.pdf

Ramsden, B. 2001: "Patterns of Higher Education Institutions in the UK", published by Universities UK and SCOP, 2001

Ramsden, B. 2004: "Review of the Initial Entry Rate into Higher Education", Office for National Statistics, National Statistics Quality Review Series, report number 24: available at http://www.statistics.gov.uk/methods_quality/quality_review/downloads/IER_Report.pdf

Rees, P. 2002: "New models for projecting UK populations by ethnic group at national and subnational scales" in Population projections by ethnic group: a feasibility study, edited by John Haskey, ONS, 2002

Simpson, 2004: Ludi Simpson, research outline, "Ethnic identity and change", at <http://www.celsius.lshtm.ac.uk/projects/simpson.html>

Copies of this publication can be obtained from:

DfES Publications
P.O. Box 5050
Sherwood Park
Annesley
Nottingham
NG15 0DJ

Tel: 0845 60 222 60
Fax: 0845 60 333 60
Minicom: 0845 60 555 60
Online: www.dfespublications.gov.uk

© Brian Ramsden 2005

Produced by the Department for Education and Skills

ISBN 1 84478 565 3
Ref No: RR676
www.dfes.go.uk/research