



Department
for Education

The 16 to 19 Bursary Fund impact evaluation - Interim Report

Research report

Revised June 2014

**Jack Britton, Haroon Chowdry and Lorraine
Dearden - Institute of Education & Institute
for Fiscal Studies**

Contents

List of figures	4
List of tables	5
Executive Summary	6
1. Introduction	11
2. Methodology	13
2.1. Evaluation design and methods	13
2.2. Data and policy definitions	16
2.3. Robustness of proposed evaluation methodology	20
3. Research Analysis Findings	25
3.1. Headline impacts on participation and attainment	25
3.2. Subgroup impacts on participation and attainment	30
4. Robustness Checks	33
4.1. DiD analysis on a placebo treatment year	33
4.2. Exclusion of 2006	34
5. Assessment of Value for Money	36
5.1. Assessment of cost effectiveness	37
5.2. Assessment of net social benefits	39
6. Conclusions	43
References	46
Appendix 1	47
Additional figures on pre-reform participation trends	47
Additional figures on pre-reform attainment trends	48
Appendix 2	49
Creation of socioeconomic index	49

Econometric specification	49
Subgroup Analysis	50
Bootstrapped Standard Errors	50

List of figures

Figure 1 Illustration of 'difference-in-differences' approach	14
Figure 2 Pre-reform common trends in male full-time Year 12 participation	21
Figure 3 Pre-reform common trends in male part-time Year 12 participation	22
Figure 4 Pre-reform common trends in male full-time Year 13 full-time participation	22
Figure 5 Pre-reform common trends in Male Level 2 threshold achievement	23
Figure 6 Pre-reform common trends in Male Level 3 threshold achievement	24
Figure 7 Pre-reform common trends in female full-time Year 12 participation	47
Figure 8 Pre-reform common trends in female full-time Year 13 full-time participation	47
Figure 9 Pre-reform common trends in Female Level 2 threshold achievement	48
Figure 10 Pre-reform common trends in Female Level 3 threshold achievement	48

List of tables

Table 1 Summary of estimated impacts on participation and attainment rates in 2011/12	8
Table 2 Number of pupil records analysed, by Year 11 cohort	16
Table 3 EMA eligibility	19
Table 4 EMA eligibility groups used in analysis	20
Table 5 Overall impacts of 16 to 19 Bursary Fund	25
Table 6 Impacts of 16 to 19 Bursary Fund: Males	26
Table 7 Impacts of 16 to 19 Bursary Fund: Females	27
Table 8 Subgroup analysis: males, females and overall	30
Table 9 Common Trends Assumption: 2011 as the placebo treatment year	34
Table 10 Common Trends Check: Excluding 2006	35
Table 11 Recap of impacts on participation and attainment	36
Table 12 Range of impacts across all EMA-eligible pupils	37
Table 13 Size of total EMA-eligible group	37
Table 14 Headcount impacts on participation and attainment	38
Table 15 Departmental savings due to 16 to 19 Bursary Fund (£m)	38
Table 16 Change in number of participants and achievers per £1m saved	39
Table 17 Scores used to create socioeconomic index	49

Executive Summary

The Institute of Education and Institute for Fiscal Studies were commissioned by the Department for Education (DfE) to undertake a statistical impact analysis of recent reforms to financial support for 16 to 19 year olds in education introduced by the Government.

The 16 to 19 Bursary Fund was implemented nationally in September 2011 as the replacement for the previous Education Maintenance Allowance (EMA), with transitional arrangements in 2011/12 for some of those still in education who had previously applied for EMA. Under EMA, students received up to £30 paid directly into their bank account every week, conditional on attendance in education or training. Eligibility and the size of payments (£10, £20 or £30) were determined by household income, and the policy cost £560m per year when it was abolished.

The policy intention was to provide more efficient targeted support for post-16 learning, by delegating £180m per year funding to education and training providers and allowing them to allocate it to the learners deemed most in need of additional support. Students can receive support through two routes. Young people in care, care leavers, and young people in receipt of income support, Employment Support Allowance or Disability Living Allowance can receive a bursary of £1,200 a year. Providers use additional funding at their discretion to make further awards to young people in ways that best fit the needs and circumstances of their students.

This interim report introduces and tests the methodology being used and provides initial findings on the policy's impact on participation and attainment in the academic year 2011/12. A final report will further develop the analysis and provide estimates of the impact in 2012/13. The study is part of the Department for Education's wider research on the 16 to 19 Bursary Fund which includes a separately commissioned process evaluation investigating the characteristics of pupils receiving the Bursary, its administration and perceived impact.

The analysis in this evaluation report uses statistical techniques to estimate the impact of replacing EMA with the 16 to 19 Bursary Fund. The outcomes analysed are whether learners stayed in full-time (FT) participation in post-16 education and whether they had achieved the Level 2 or 3 attainment threshold by age 18. The estimated impacts in this research briefing should be interpreted as the changes in participation and attainment rates compared to a hypothetical no-reform scenario where the EMA had remained unchanged in 2011/12. They are not the impacts compared to a scenario without financial support for 16 to 19 learners. In 2011/12, learners who had successfully applied for 'full' EMA in 2010/11 at £30 per week were entitled to a £20 EMA payment under

transitional arrangements¹. Those previously claiming lower EMA payments of £20 or £10 per week were not entitled to payments but could apply for the new Bursaries. These transitional arrangements mean that in particular the impact on Year 13 participation may not in future reflect that observed here.

Measuring the impacts of the policy reform is complicated by the fact that it was implemented for all young people in England at the same time (i.e. there was no control group against which to measure the impact). It is therefore challenging to identify what participation and attainment rates would have been in 2011/12 had there been no reform to EMA. Reliably estimating this quantity is crucial as it is the baseline to which the actual levels of participation and attainment in 2011/12 should be compared in order to isolate the impact of the reform.

The analysis in this report is based on the administrative data records for approximately 5.2 million state school pupils who were in Year 11 between 2002/03 and 2010/11. The analysis tracks the post-16 education outcomes for lower-income pupils – whom it is believed would have been eligible for EMA in 2011/12 had it been retained – against the same outcomes for pupils whose family income was slightly too high for them to have been eligible for EMA in 2011/12 (had it been retained). The latter group is assumed to have been unaffected by the reform, since they would not have been eligible for EMA. It is also assumed that the change in their observed post-16 education outcomes in 2011/12 is a reliable guide to change in the education outcomes that would have been seen among lower-income pupils in 2011/12 had there been no policy reform, despite them having different levels in these outcomes, controlling for other factors. Pupils with family incomes just above the EMA eligibility income threshold are used as the basis for comparison in order to maximise the validity of this assumption, which is crucial to the ‘difference-in-difference’ (DiD) approach used here.

Without direct access to information on pupils’ family income – which is not recorded in any of the data sources – this analysis attempts to identify lower- and higher-income pupils on the basis of other socioeconomic characteristics such as Free School Meals eligibility and local neighbourhood characteristics including those measured at postcode level.² These characteristics are combined into a socioeconomic index which is used as a proxy of pupils’ household income rank and thus enables them to be placed into different income groups using annual nationally representative household income from the Family Resources Survey (FRS). More detail on this approach can be found in the ‘Methodology’ section.

¹ http://readingroom.ypla.gov.uk/ypla/ypla-16-19_transitional_arrangements_guide_for_2011-12-gn-aug11-v2.pdf

² The local neighbourhood characteristics are: IMD and IDACI scores, the ACORN socio-economic classification and the proportion of households that are owner-occupied.

Table 1 summarises the estimated headline impacts of the reform on participation and attainment rates in 2011/12. The first row shows that the estimated effect of the implementation of the policy led to a 1.2 percentage point (ppt) fall in FT participation amongst Year 12 students who would otherwise have been eligible for the full EMA award. In other words, their participation rate in FT education in Year 12 would have been 1.2 ppts higher in 2011/12 had there been no reform to EMA. The impact among the wider group of pupils who would have been eligible for any level of EMA support is 1.1 percentage points which translates into a fall amongst the cohort as a whole of 0.7 ppts, to 83.9%. For the Year 13 transition cohort, there was a 1.8 ppts fall in FT participation among the poorest students who would have previously been eligible for the full EMA, and a 1.5 ppt fall among pupils who would previously have been eligible for any EMA. This translates into a reduction of 0.9 ppts across the entire cohort, to 69.7%. Overall our findings suggest that participation amongst all full time Year 12 and Year 13 students dropped by 8,100 individuals as a consequence of the policy change.

Table 1 Summary of estimated impacts on participation and attainment rates in 2011/12

	Impact on lowest-income pupils (who would have been eligible for maximum EMA support)	Impact across all pupils who would have been eligible for any EMA support	Impact across cohort as a whole
Y12 FT participation	-1.19ppts	-1.07ppts	-0.65ppts (83.9%)
Y13 FT participation	-1.75ppts	-1.50ppts	-0.88ppts (69.7%)
L2 by 18 attainment	-1.83ppts	-1.52ppts	-0.90ppts (82.8%)
L3 by 18 attainment	-0.09ppts*	-0.05ppts*	-0.03ppts* (48.0%)

The figures in the parentheses in column 3 give the true overall post-reform proportion. * Indicates not statistically significant.

The impacts on attainment are also the most negative among the poorest students: for those who would have been eligible for the full EMA award, there was a 1.8 ppts fall in the L2 achievement rate by age 18, leading to a 0.9 ppt fall across the whole Year 13 cohort, to 82.8%. No statistically significant impact is found on the L3 achievement by age 18.

The report estimates larger effects on males than on females, suggesting the reform led to a small increase in the gender gap both in terms of participation and achievement. However, it should be noted that whilst those differences are consistent, none of them are statistically significantly different. Further subgroup analysis suggests the reduction in participation was significantly larger amongst non-whites than amongst whites, and amongst EAL students against non-EAL students. These results hold for both males and females in both Y12 and Y13.

For males, the analysis finds that there is a positive overall effect on participation of SEN students in both Y12 and Y13, which is potentially attributable to increased eligibility for grants for SEN students under another component of the Bursary. However, it is important to note that the result does not hold for females.

Investigating Level 2 attainment by subgroup suggests negative overall effects and stronger negative effects for both SEN and EAL students, for both males and females. There is no statistically significant difference in Level 2 attainment for whites and non-whites.

For Level 3 attainment, meanwhile, the results of the subgroup analysis should be treated with caution, due to the unexpected signs and large magnitudes of the regression coefficients. For non-whites Level 3 attainment is estimated to have *improved* under the Bursary for both males and females. For SEN students, Level 3 attainment appears to have reduced, while there is no significant difference between Level 3 attainment for EAL and non-EAL students.

Importantly, it is likely that the overall impact estimates presented in this report are slightly conservative and may underestimate the true impacts of the policy reform. This is a consequence of the empirical strategy, in particular the use of proxy information on socioeconomic characteristics in the absence of actual income data to determine potential EMA eligibility. We feel we have adopted an approach which minimises the risks of this by including individuals with incomes up to £35,000 in our 'EMA eligible' group, reducing the extent to which our control group contains EMA eligible individuals. This means that our estimates of impact is on a group slightly larger than the true EMA eligible group³, but ensures that our control group is unlikely to contain individuals who would have been eligible for the EMA. It also means that our estimate of the overall impact on the whole cohort should be robust, as long as our modified DiD assumptions hold.

To perform a basic analysis of value for money, the headline impact estimates are combined with the departmental savings arising from the reform associated with the

³ For the remainder of the report, when we refer to the EMA eligible group we are referring to this slightly larger group and hence are likely to be underestimating the impact for those who are eligible for the EMA. Our overall estimate of income for the whole cohort will not suffer from this problem.

reduction in financial support payments and learner-driven funding to providers. This suggests that the reform led to 20 fewer FT participants and 12 fewer Level 2/3 achievers per £1m saved in 2011/12. In the 'worst-case' scenario, based on the largest plausible impacts, the reform led to 26 fewer FT participants and 17 fewer Level 2/3 achievers per £1m saved. In the 'best-case' scenario, based on the smallest plausible impacts, the reform led to 13 fewer FT participants and 6 fewer Level 2/3 achievers per £1m saved.

In order to judge the effectiveness or otherwise of the programme we need to look at the long term consequences of this increase in drop-out rates and reduction in qualification levels for the individual and for the Exchequer (for example in terms of reduced taxation and increased benefit payments over the affected cohorts lifetime. This will allow us to compare the costs (reduced taxation and increased benefit payments) and benefits (reduced cost of the 16 to19 Bursary Scheme compared to the EMA scheme) and to directly compare these benefits with other programmes to get a true sense of the effectiveness of the reforms. We will do this in our final report and we outline how we propose to do this in the report.

1. Introduction

This report is part of the 16 to 19 Bursary Fund evaluation and provides interim findings on the impacts of the policy on participation and attainment rates in 2011/12. The Institute of Education and Institute for Fiscal Studies were commissioned by the Department for Education (DfE) to undertake impact analysis of the 16 to 19 Bursary Fund as part of a wider evaluation. A second evaluation report, to be published in late 2014, will further refine and test the methodology and provide estimates of the impact of the policy on education participation and attainment in 2012/13 at which point a cohort will have gone through both Years 12 and 13 under the new system.

The 16 to 19 Bursary Fund was implemented nationally in September 2011 as the replacement for the previous Education Maintenance Allowance (EMA). Under EMA, students received up to £30 paid directly into their bank account every week, conditional on attendance in education or training. Eligibility and the size of payments (£10, £20 or £30) were determined by household income, and the policy cost £560m per year when it was abolished.

The Bursary Fund's intention was to provide more efficient targeted support for post-16 learning, by delegating £180m per year funding to education and training providers and allowing them to allocate it to the learners deemed most in need of additional support. Students can receive support through two routes. Eligible Young people can receive a bursary of £1,200 a year which include those who are:

- in care,
- care leavers,
- in receipt of income support
- disabled young people in receipt of Employment Support Allowance
- Disability Living Allowance

Providers can also use Discretionary Bursaries to make further awards, comprising the majority of the funding, to young people in ways that best fit the needs and circumstances of their students.

The EMA was closed to new applicants on 1 January 2011, and the new 16 to 19 Bursary Fund arrangements were in place from September 2011 onwards. Pupils who started Year 13 in September 2011 were subject to transitional arrangements, whereby those who had previously successfully applied for a full EMA award of £30 a week were to receive a reduced EMA of £20 a week, while those who had previously claimed a partial EMA award of £10 or £20 a week would no longer receive EMA but could apply for support from the 16 to 19 Bursary Fund. Meanwhile, pupils starting Year 12 in September 2011 faced the new bursary arrangements only.

The analysis in this research briefing uses statistical techniques to provide estimates of the impact of replacing EMA with the 16 to 19 Bursary Fund. The outcomes analysed are whether learners stayed in full-time (FT) participation in post-16 education and whether

they had achieved the Level 2 or 3 thresholds by age 18. The estimated impacts in this research briefing should be interpreted as the changes in participation and attainment rates compared to a hypothetical no-reform scenario where the EMA had been retained in 2011/12. They are not the impacts compared to a scenario of no 16 to 19 financial support.

Measuring the impacts of the policy reform in question is complicated by the fact that the reform was implemented across England at the same time. It is therefore challenging to identify what participation and attainment rates would have been in 2011/12 *had there been no reform to EMA*. Reliably measuring this quantity is crucial, as it is the baseline to which the actual levels of participation and attainment in 2011/12 should be compared, in order to isolate impacts that can be attributed to the policy reform itself.

Briefly, the analysis in this report compares the post-16 education outcomes for lower-income pupils – who would have been eligible for EMA in 2011/12 had it been retained – against the same outcomes for pupils whose family income was slightly too high for them to have been eligible for EMA in 2011/12 (had it been retained). The latter group is assumed to have been unaffected by the reform, since they would not have been eligible for EMA. Furthermore, it is assumed that the change in their observed post-16 education outcomes is a reliable guide to the change in education outcomes that would have been seen among lower-income pupils in 2011/12 had there been no policy reform, despite them having different *levels* in these outcomes controlling for other factors. Pupils with family incomes just above the EMA eligibility income threshold are used as the basis for comparison in order to maximise the validity of this assumption, which is crucial to the ‘difference-in-difference’ (DiD) approach used here.

This research briefing is structured as follows. The ‘Methodology’ section sets out in detail the challenges involved in reliably measuring the impacts of the reform on participation and attainment rates, and the approach proposed by this research to deal with those challenges. It describes the data sources that this research uses and the information on outcomes and pupil characteristics that will frame the impact analysis. It also tests the likely validity of the proposed empirical approach and identifies the particular outcomes for which the proposed approach is most likely to provide reliable impact estimates. The ‘Research Analysis Findings’ section then provides the estimated impacts of the 16 to 19 Bursaries Fund on participation and attainment rates in 2011/12, both across the cohort as a whole and for specific groups of pupils. The common trends assumption – which is key to the validity of the overall analysis – is then tested formally in the “Robustness Checks” section. Finally, the ‘Conclusions’ section draws together all of the findings and underlines the implications of this research.

2. Methodology

This section describes the empirical challenges and approaches involved in robustly measuring the impact of the 16 to 19 Bursaries on post-16 participation and attainment rates. It also provides evidence on the likely suitability of the proposed approach and identifies the outcomes examined in the impact analysis.

2.1. Evaluation design and methods

This report attempts to measure the impact of the policy reform as the difference between the actual participation and attainment outcomes that were observed in 2011/12, and the participation and attainment outcomes that *would have* been observed in 2011/12 had there been no policy reform. The latter cannot be measured directly as it relates to a hypothetical – often referred to as ‘counterfactual’ – scenario. The evaluation strategy must therefore use statistical techniques to best approximate the participation and attainment outcomes that would have been observed in the counterfactual scenario. A major challenge here is that the policy reform in question was implemented across England at the start of the 2011/12 academic year. This means there are no easily identified areas or groups of learners still eligible for the scheme’s predecessor (EMA) which could serve as potential comparators.⁴

In brief, the analysis in this report attempts to uncover the impact of policy reform in 2011/12 by comparing the change in post-16 education outcomes in 2011/12 amongst pupils who *would have been potentially eligible* (on income grounds) for EMA had it been kept, with the change in post-16 education outcomes in 2011/12 amongst pupils who *would not have been potentially eligible* (on income grounds) for EMA had it been kept. The assumption motivating this approach is that the latter group were not affected by the policy reform and therefore provide a reliable guide to the change in participation and attainment outcomes that would be expected to prevail had there been no policy reform.

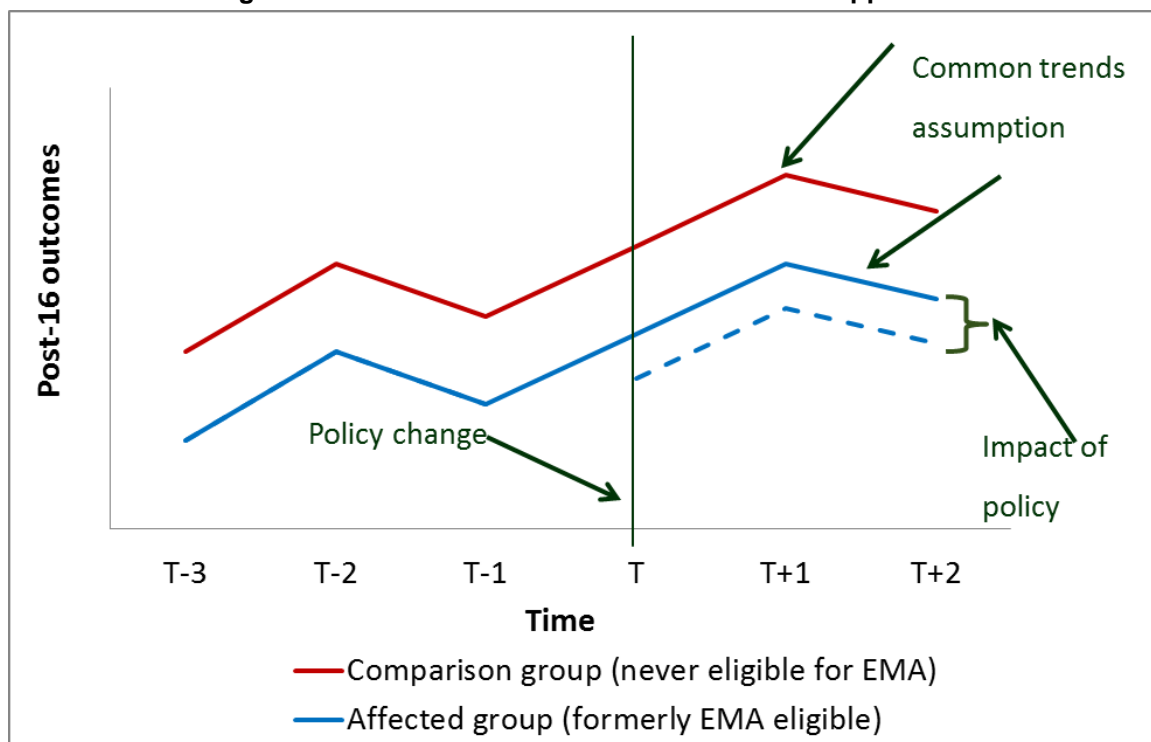
Since EMA eligibility was determined on the basis of family income, a potential comparator would have to be pupils from higher-income backgrounds who would not have been eligible for EMA even if it had been left unchanged. However, it is well-known that pupils from higher-income backgrounds generally have higher post-16 participation and attainment rates. This could confound any comparison and lead to misleading estimates of the impact of the policy reform: the difference in outcomes using higher-income pupils as a comparator is likely to be negative – suggesting a fall in participation and attainment – but this could reflect other factors, such as prior attainment, rather than the actual reform itself.

⁴ Pupils in Year 14 in 2011/12 were still eligible for EMA. However, such pupils would have been a poor comparator because of their age difference (compared with pupils in Year 12 and Year 13) and because there are few of them, thereby preventing precise statistical analysis.

A further issue arises when attempting to make comparisons over time, between education outcomes in 2011/12, when the 16 to 19 Bursary Fund was in operation, and education outcomes in previous years. In particular, other factors could have occurred which might cause a rise or fall in participation or attainment rates over this period; examples might be the broader economic environment, the state of the youth labour market or reforms to the higher education finance regime.⁵ It is clearly necessary to strip out the effects of these factors, where they occur, in order to be confident that any changes in participation and attainment only reflect the introduction of the 16 to 19 Bursary Fund.

In an attempt to circumvent this issue, the empirical analysis in this report compares the *trends over time* in participation and attainment outcomes for lower-income pupils against the *trends over time* in the same outcomes for higher-income pupils. Specifically, the analysis tracks successive cohorts of young people and examines the *change* in lower-income pupils' outcomes in 2011/12 (the first year of the 16 to 19 Bursaries); this is then compared against the *change* in higher-income pupils' outcomes in 2011/12. The *difference* between these two changes is then estimated as the impact of the policy reform; as a result, this methodology is referred to as a 'difference-in-differences' (DiD) approach.⁶

Figure 1 Illustration of 'difference-in-differences' approach



⁵ The cohort in Year 13 in 2011/12 would have been the first cohort to face the new higher education finance regime, involving maximum tuition fees of £9,000 per year.

⁶ This approach has been used before to evaluate the quantitative impact of education programmes and reforms. See, for example, Kendall et al. (2005) and Tanner et al. (2011).

Figure 1 illustrates the DiD approach schematically. The crucial assumption which underpins the DiD approach is that of common underlying trends after the policy reform: while the levels of outcomes are different, the *trends* in outcomes between different groups follow a parallel trajectory and would have continued to follow a parallel trajectory had there been no policy reform. In other words, the assumption is that there is no convergence or divergence between the education outcomes of the affected and comparison groups; if so, then any convergence or divergence that takes place after the policy reform takes place can be attributed to the policy reform itself.

The common trends assumption cannot be assessed using data following the policy reform (e.g. in 2011/12) because that would involve outcomes which may be affected by the reform itself. Instead, the validity of this approach is assessed by examining whether the affected and comparison groups exhibited common trends in their outcomes *before* the policy reform. If this is likely to be true then, while it does not *prove* conclusively that the two groups' outcomes would have continued to follow parallel trends after the policy reform (had the reform not happened), it does provide some confidence in the common trends assumption.

While the above description provides a basic flavour of how a DiD approach works; the empirical analysis presented in this report goes a step further and uses a *modified* version of a DiD model to provide greater statistical robustness. As the data sources on which the analysis is based cover many time periods, it is possible to estimate what the long-run underlying linear trend in education outcomes is for each income group as well as the effects of other observed characteristics. The DiD approach used in this analysis 'strips out' these underlying trends for each income group, as well as the effects of background characteristics, using statistical regression modelling and then assesses the common trends assumption. Appendix 2 gives further details of the modelling approach. We show that by doing this we do well for all outcomes of interest except part-time education participation. We use this modified DiD approach to examine the impact of the reforms on full-time participation in Y12 and Y13 as well as the proportion obtaining L2 and L3 qualifications.⁷

Since these trends are estimated directly, they are known – hence the assumption of parallel trajectories is no longer necessary. Instead, the assumption is now that the trends follow some linear path⁸ – not a curve – but that these linear paths can go in different directions for different income groups. Hence the education outcomes for

⁷ Technically, this approach of estimating separate trends for each group can be thought as running a before-after analysis for each group, and then examining the differences between each affected group's before-after estimates compared against the comparison group's before-after estimate.

⁸ This is because a linear trend in education outcomes is estimated for each income group. The analysis experimented with estimating quadratic trends, but this did not lead to qualitatively different results. It did, however, provide less precision in the impact estimates because of the greater demand placed by the models on the data.

different groups can converge or diverge: for example, it is not a problem if the participation rate of low-income groups is seen to ‘catch up’ with that of higher-income groups before the policy reform. What matters is whether it can be believed that the low-income group would have continued to catch up *at the same rate* after the date of the policy reform, had the reform not happened. This assumption will be plausible if the trends estimated for each income group before the policy reform appear to follow a linear path.

This section of the report will provide graphical evidence on whether the parallel trajectories assumption holds for our outcomes for different income groups once we strip out the effects of linear trends and background characteristics. The impact analysis will then focus on those outcomes for which this appears to hold, since the empirical approach is the most valid for them.

Before doing so, it is necessary to describe in more detail the data that is used in the empirical analysis and the methods for constructing the different income groups. This is done in the next section.

2.2. Data and policy definitions

This analysis uses National Pupil Database (NPD) records on state school pupils who were in Year 11 from 2002/03 to 2010/11. Table 2 Provides a breakdown of the number of pupil records analysed, by the academic year (cohort) in which pupils were in Year 11.

Table 2 Number of pupil records analysed, by Year 11 cohort

Year 11 cohort	Number of records
2002/03	563,721
2003/04	585,769
2004/05	582,285
2005/06	591,499
2006/07	598,949
2007/08	595,974
2008/09	577,260
2009/10	576,180
2010/11	564,584
Total	5,236,221

For each pupil, the following information is available on their demographic and socio-economic characteristics from the School Census:

- Gender;

- Ethnicity;
- Free School Meals (FSM) eligibility;
- English as an Additional Language (EAL) status;
- Special Education Needs (SEN) status;
- IMD and IDACI scores.⁹

Linked information on each pupil's average point score at Key Stage 2 (KS2) and total capped point score at Key Stage 4 (KS4) is also available from within the NPD data.

Finally, information on each pupil's post-16 participation and attainment is also linked in. Participation records come either from School Census data one (two) years later to capture participation in school sixth forms in Year 12 (13), or from Individualised Learner Record (ILR) data to capture participation in other sixth form and further education institutions. Post-16 attainment data is taken from the Level 2/3 indicators data provided by DfE: for each pupil, indicators for whether they achieved the Level 2/3 threshold by age 18 – and if so, whether through the academic or vocational route – are linked in. The participation and attainment outcomes are recorded for 2004/05¹⁰ to 2011/12, which was the latest year available at the time of analysis.

The cohorts who were in Year 11 up to 2008/09 faced the previous EMA regime: these are referred to as EMA cohorts. The Year 11 cohort in 2009/10 is the EMA transition cohort: such pupils would have faced the EMA regime in Year 12 in 2010/11, but would have then faced the transitional arrangements in Year 13 in 2011/12. Finally, the cohort in Year 11 in 2010/11 is the bursary cohort: such pupils will have faced the new bursary arrangements in Year 12 in 2011/12 (and in Year 13 in 2012/13). This means that the analysis of Year 12 participation outcomes in this report focuses on the bursary cohort, while the analysis of Year 13 participation and attainment outcomes focuses on the transition cohort.

The analysis of common trends, as well as the impact analysis, both take into account pupil characteristics in order to control for them to the extent that they might influence the participation and attainment outcomes of interest. In particular, there may be changes over time in the relevant characteristics of different income groups: for example, while higher-income pupils tend to have higher prior attainment, it may be that the prior attainment of lower-income pupils has caught up over time, thereby narrowing the attainment gap. Any analysis would need to take *relative* changes such as these into account to avoid confounding the estimated impacts of the policy reform. The following characteristics are controlled for in the analysis by estimating their influence using

⁹ For more information on these indices, see <http://data.gov.uk/dataset/index-of-multiple-deprivation>.

¹⁰ This is the year that EMA was implemented nationally. The analysis does not use any data on post-16 outcomes preceding this point, in order to avoid having any policy changes during the pre-reform window (which could compromise the assessment of the trends in outcomes).

regression modelling in order to isolate the underlying trends (see Appendix 2 for model specifications):

- Ethnicity;
- English as an Additional Language (EAL) status;
- Special Education Needs (SEN) status;
- Attainment at KS2 and KS4.

The analysis is also split by gender. The next step is to construct the different income groups in order to identify pupils who would have been affected by the reforms and pupils who were not. However, the data used in this analysis do not contain any measures of family income that could be used to define whether a pupil would have been eligible on EMA grounds. This eligibility or otherwise therefore has to be imputed on the basis of the socio-economic characteristics available. These are an individual's Free School Meals (FSM) status, their neighbourhood's Index of Multiple Deprivation (IMD) and Income Domain Affecting Children Index (IDACI) scores, their postcode's 'ACORN' socio-economic classification¹¹ and the proportion of households in the pupil's neighbourhood which are owner-occupied (taken from the 2001 Census). All of this information is combined into an index which serves as a proxy for family income.¹²

To map this index to levels of actual family income, it is combined with information on the distribution of family income taken from the Family Resources Survey (FRS) for 2003/04 to 2009/10. In particular, the percentile points for the distribution of gross family income among households with at least one child aged 14–16 are used. Each pupil's percentile in the socioeconomic index is then mapped to the same percentile in the income distribution: for example, if a pupil's score on the socioeconomic index is at the 39th percentile of that index, they are given the income corresponding to the 39th percentile of the family income distribution according to the FRS. The assumption underpinning this is that a pupil's ranking in the socioeconomic index is the same as their ranking in the distribution of family income.

With a level of family income assigned to each pupil, they are then classified to a particular level of potential EMA eligibility using the known income thresholds for EMA, including for cohorts subject to the transitional or bursary arrangements who might otherwise have been eligible for EMA; the aim is to identify groups of pupils with different levels of exposure to the policy reform in 2011/12. Each income group corresponds to a particular level of EMA eligibility, as shown in Table 3.

¹¹ For more information on the ACORN index, see <http://acorn.caci.co.uk/>.

¹² This is done using principal components analysis. More information can be found in the Appendix.

Table 3 EMA eligibility

EMA eligibility	Gross family income (before 2005/06)	Gross family income (from 2005/06)
Full (£30)	£19,630 or less	£20,817 or less
Partial (£20)	£19,631–£24,030	£20,818–£25,521
Partial (£10)	£24,031–£30,000	£25,522–£30,810
None	£30,001+	£30,811+

Groups 1 to 3, detailed in Table 4, are the income groups affected by the reform in 2011/12, for whom the impacts will be measured. An inevitable consequence of our estimation of income is that some children will be inaccurately classified; individuals who were truly ineligible might be classified as eligible and vice versa. This is likely to be particularly true for individuals with true family incomes near to the £30,000 eligibility cut-off. To allow for this, we extend our Group 3 income eligibility to £35,000. This means that we will slightly underestimate the impact for this EMA-eligible group by systematically including some who won't have received it, but it will help to ensure our control group has a reasonably low chance of actually receiving EMA and give more accurate estimates of the overall affect for this income group. Group 4 consists of pupils whose imputed family income was slightly too high for them to be potentially eligible for EMA if they stayed in education or training. This group serves as the comparison group in all of the analysis that follows.¹³ This group was chosen specifically to be as 'similar' as possible to the Groups 1–3 in terms of pupil characteristics and education outcomes, while also being likely to ineligible for EMA and therefore likely to be unaffected by the 16 to 19 Bursary Fund.

¹³ Groups 1 to 4 constitute around 65% of a total Year 11 cohort. A fifth group, who's imputed gross family income is above £35,000, is also created in order to ensure that the full cohort is used in the statistical models and maximise the sample sizes. This fifth group can be used to assess the robustness of our estimates, for instance as an alternative control group, if the common trends assumption appears robust.

Table 4 EMA eligibility groups used in analysis

Group	Notional EMA eligibility	Gross family income (before 2005/06)	Gross family income (from 2005/06)
1	Full (£30)	£19,630 or less	£20,817 or less
2	Partial (£20)	£19,631–£24,030	£20,818–£25,521
3	Partial (£10)	£24,031–£35,000	£25,522–£35,000
4	Control Group	£35,001–£45,000	£35,001–£45,000
5	Rich Group	£45,001+	£45,001+

Using this tightly-defined comparison group (group 4) in order to measure the impacts of the reform maximises the likely validity of the assumptions that are made regarding the linearity of the trend in each income group’s outcomes (relative to the comparison group). By starting it £5,000 above the EMA threshold, we reduce the probability of including EMA recipients in the control group. The next section uses these group definitions to assess the validity of these assumptions. Finally, Group 5 is the wealthiest group consisting of individuals from households with combined gross income above £45,000. Although individuals within this group were therefore unaffected by the reform, they are likely to be more systematically different to those eligible for the EMA than individuals in Group 4, and this Group is therefore not used as the control group in our regressions. However, Group 5 is not excluded from the analysis altogether, as it serves as a useful check for the accuracy of our results. For example, in the analysis of Y12 full-time participation, the inclusion of Group 5 provides suggestive evidence that the effect in this case is an underestimate of the true effect.

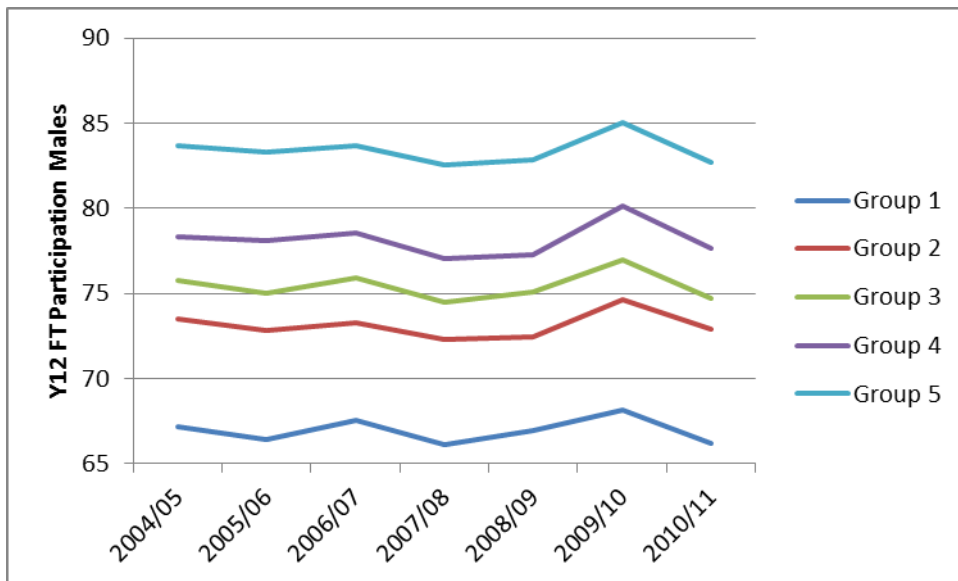
2.3. Robustness of proposed evaluation methodology

As stated above, the DiD approach is valid under the assumption that the underlying trends in outcomes for each of the income groups are parallel. In our modified approach, we assume that after stripping out *linear* trends for each group and the effects of other covariates, the underlying trends are parallel. This is assessed graphically by plotting the underlying trends in participation and academic outcomes for each group, once we strip out linear trends and the effects of other covariates using data from 2005 until 2011.

Figure 2 shows Year 12 full-time participation of males for Groups 1 to 5, after controlling for the characteristics described above and stripping out linear trends. Under the common trends assumption these lines would be parallel, maintaining the same difference with the comparison group over time. We see that all groups, even Group 5,

appear to move in parallel over the entire period¹⁴. Group 4, those from families earning between £35,000 and £45,000 per year appears to follow common trends with all of our groups of interest. An equivalent chart for females can be found in Appendix 1; again the common trends assumption appears to hold.

Figure 2 Pre-reform common trends in male full-time Year 12 participation



In Figure 3 the analysis is repeated for male part-time participation in Year 12 using the same methodology. This time the common trends assumption does not seem to hold, which suggests the proposed empirical methodology is unlikely to be valid for this outcome. Female part-time participation in Year 12 is not presented (they are available on request), though that pattern has similar implications.

¹⁴ We also repeated the analysis using quadratic trends, but these were rarely significant and did not perform as well as simple linear trends.

Figure 3 Pre-reform common trends in male part-time Year 12 participation

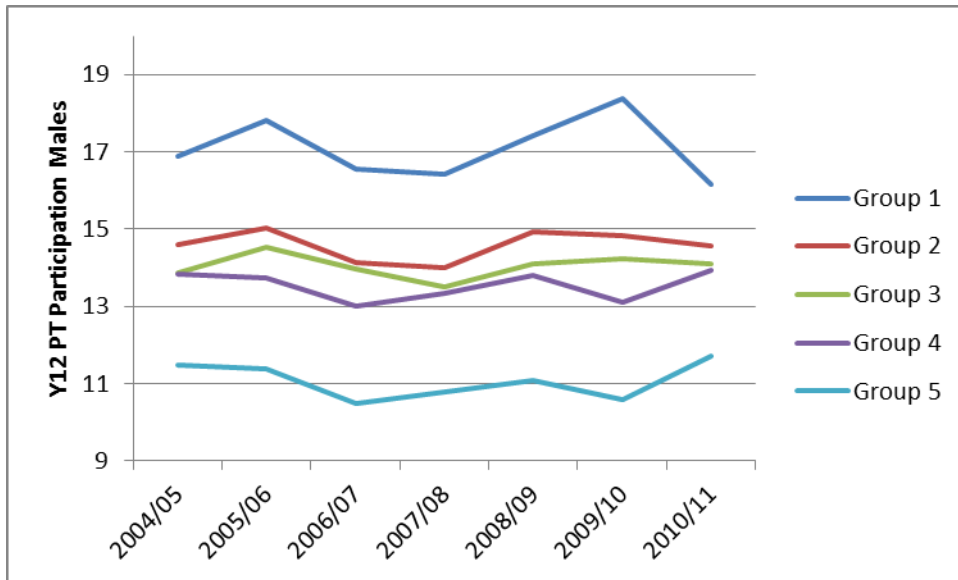
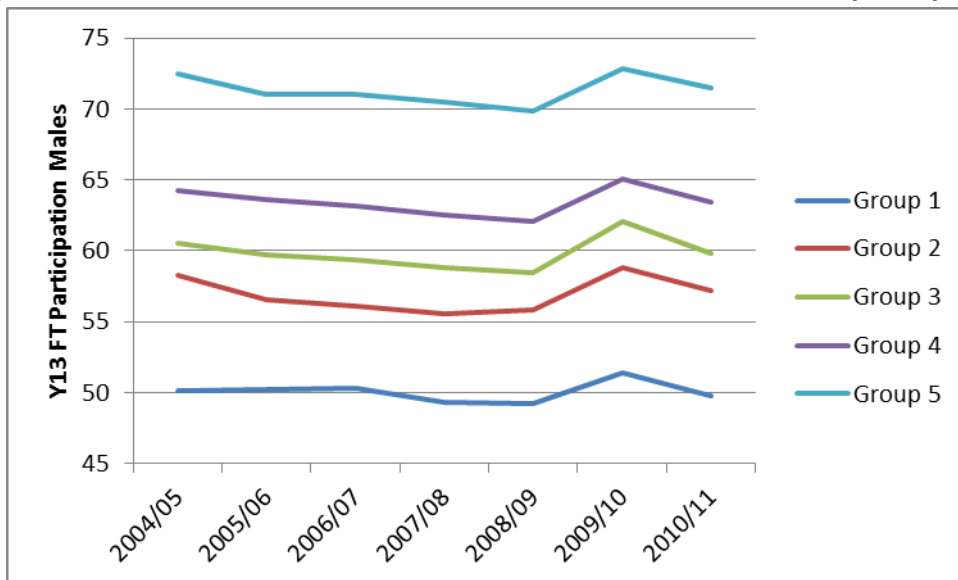


Figure 4 presents the pre-reform trends in full-time Year 13 participation for males. As with Year 12, the common trends assumptions appear to hold, suggesting that linear underlying trends are sufficient in explaining any apparent convergences or divergences between groups. The same is true for females (see Appendix 1 for chart).

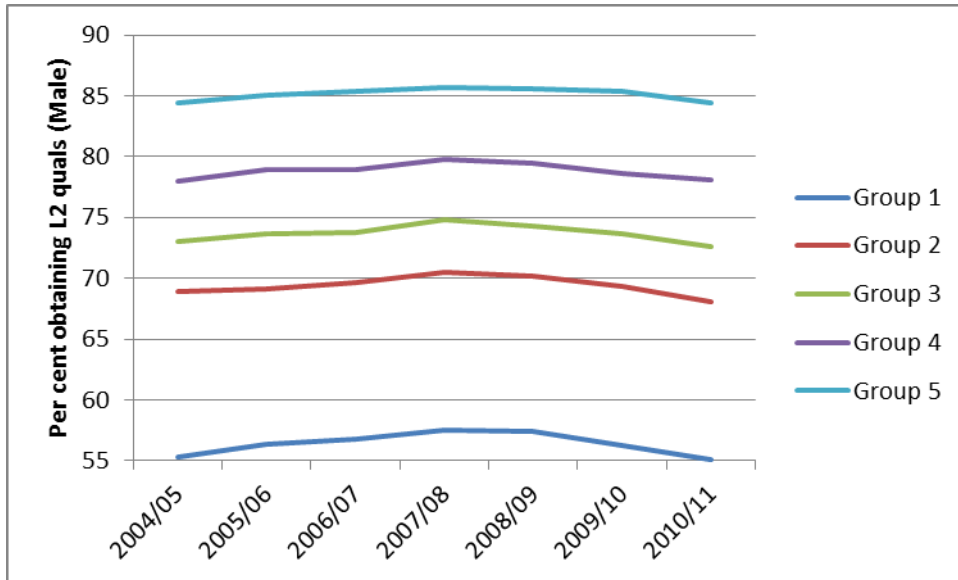
Figure 4 Pre-reform common trends in male full-time Year 13 full-time participation



As for Year 12, for part-time participation the common trends assumption clearly does not hold for either males or females, which make it difficult to draw conclusions for this group. The relevant graphs are not presented but are available on request.

Moving on to investigate attainment, Figure 5 plots the trends in the proportion of male pupils achieving at least the Level 2 threshold¹⁵ by 18 (through any route) during the pre-reform window.. While the trends for Groups 1 to 3 are not parallel to the horizontal axes, the assumption of linearity does not appear to be unsuitable. The equivalent for females shows similar patterns (see Appendix 1). Hence this outcome can be analysed with the proposed empirical approach.

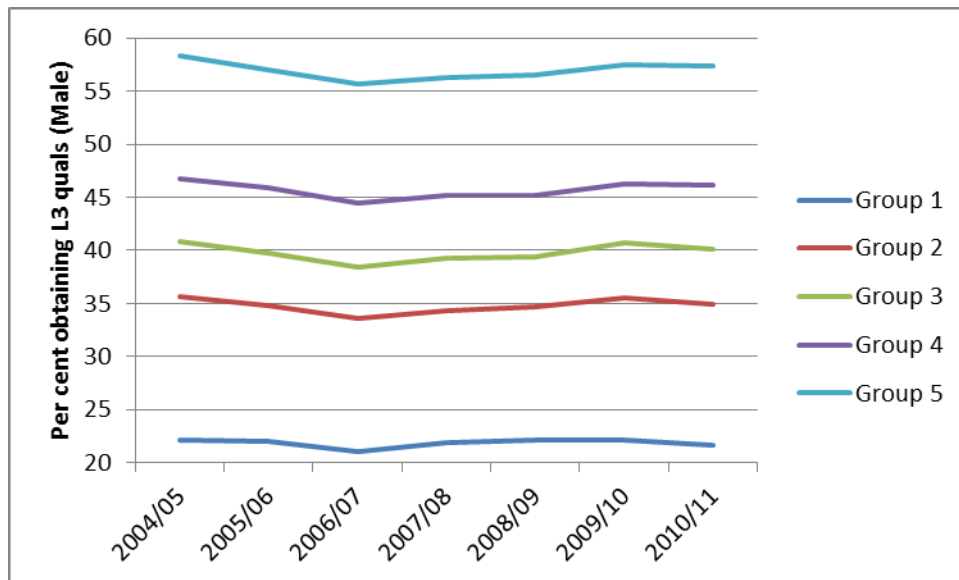
Figure 5 Pre-reform common trends in Male Level 2 threshold achievement



Finally, Figure 6 shows the underlying trends in the proportion of reaching the Level 3 threshold by 18 (through any route) for male (see Appendix 1 for females). Again, the assumption of a differential linear trend appears reasonable, hence this outcome can be analysed with the proposed empirical approach.

¹⁵ This includes higher qualifications.

Figure 6 Pre-reform common trends in Male Level 3 threshold achievement



Based on this analysis, the impact analysis in the next section focuses on the outcomes for which the assumption of differential linear trends appears most valid, in which case a DiD model stripping out those trends and the effect of background characteristics is most likely to reliably uncover the policy impact. These outcomes are:

- Full-time (FT) participation in Year 12;
- Full-time (FT) participation in Year 13;
- Achievement of the Level 2 (L2) threshold by 18;
- Achievement of the Level 3 (L2) threshold by 18.

For all four of these categories the results are split by gender, meaning separate regressions are run for males and females in each case. Based on these illustrations and independent testing of the common trends assumption, our preferred specification involves not using 2005 data¹⁶. We also report results when we also exclude 2006 data.

¹⁶ If we use 2005 data, the common trend assumption does not hold for Group 1 for a number of outcomes.

3. Research Analysis Findings

This section presents the estimates of the impacts of the 16 to 19 Bursary Fund in 2011/12 on Year 12 and 13 FT participation, and attainment at L2 and L3 by 18, based on the empirical approach previously described. The impacts are presented in percentage point terms and should be interpreted as the effect on participation or attainment in 2011/12 of replacing EMA with the 16 to 19 Bursary Fund. For example, an impact of -0.5 percentage points (ppts) on FT participation means that FT participation was 0.5 ppts lower in 2011/12 as a result of the reform; in absence of the reform to the EMA, participation would have been 0.5 ppts higher in 2011/12.

3.1. Headline impacts on participation and attainment

The headline estimates are presented below. These figures show, for each outcome, the impact on each EMA eligibility group (1, 2 or 3), the average impact across Groups 1–3 and the average impact across the cohort as a whole. We are slightly concerned that our comparison group, Group 4 may be slightly contaminated, and include people who were receiving EMA. As a way of checking this, we included the impact of the change in policy on Group 5. Our hypothesis would be that this group should not have experienced similar changes to Group 4. It is clear that this is not true for our year 12 sample in particular, and the positive significant effects we find for this group suggests that we may be underestimating the effects of changes in the policy for Groups 1, 2 and 3. There is also a significant difference between Groups 4 and 5 when considering Level 2 attainment, though the magnitude of the effect is smaller.

Table 5 Overall impacts of 16 to 19 Bursary Fund

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Group 1 impact (ppts)	-1.194**	-1.751**	-1.825**	-0.086
Group 2 impact (ppts)	-0.647*	-1.647**	-1.178**	-0.136
Group 3 impact (ppts)	-0.834**	-0.506	-0.588**	0.140
Impact across Groups 1–3 (ppts)	-1.073**	-1.498**	-1.520**	-0.047
Impact across whole cohort (ppts)	-0.649**	-0.884**	-0.898**	-0.028
Group 5 impacts (ppts)	0.898**	-0.044	0.426**	0.361
Actual outcome level (%)	83.3	69.7	82.8	48.0
Predicted level without reform (%)	83.9	70.6	83.7	48.0

Notes: * = statistically significant at 5% level; ** = statistically significant at 1% level. Standard errors clustered at school level – see Section 4.3 for further discussion of this.

These results are derived from separate regressions for males and females, the results of which are given in Tables 6 and 7. The tables show that the effect is generally stronger for males, both in terms of participation and attainment. This suggests that the gap in participation and attainment between males and females grew as a result of the reform. However it should be noted that in almost all cases the gender gap is not statistically significantly different to zero – that is to say there is no statistically significant difference between males and females.

Table 6 Impacts of 16 to 19 Bursary Fund: Males

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Group 1 impact (ppts)	-1.296**	-2.152**	-1.954**	-0.190
Group 2 impact (ppts)	-0.770	-1.996**	-1.392**	-0.276
Group 3 impact (ppts)	-1.152**	-1.026*	-0.678*	-0.186
Impact across Groups 1–3 (ppts)	-1.217**	-1.916**	-1.648**	-0.198
Impact across whole cohort (ppts)	-0.736**	-1.129**	-0.971**	-0.117
Group 5 impacts (ppts)	0.949**	-0.567	0.410	-0.407
Actual outcome level (%)	81.3	66.8	79.6	42.9
Predicted level without reform (%)	82.1	67.9	80.6	43.0

Notes: * = statistically significant at 5% level; ** = statistically significant at 1% level. Standard errors clustered at school level. See section 4.3 for further discussion of this.

Table 7 Impacts of 16 to 19 Bursary Fund: Females

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Group 1 impact (ppts)	-1.089**	-1.336**	-1.692**	0.022
Group 2 impact (ppts)	-0.519	-1.287*	-0.957**	0.008
Group 3 impact (ppts)	-0.505	0.032	-0.494	0.477
Impact across Groups 1–3 (ppts)	-0.924	-1.065**	-1.388**	0.109
Impact across whole cohort (ppts)	-0.558**	-0.631**	-0.823**	0.065
Group 5 impacts (ppts)	0.846**	0.497	0.442*	1.155**
Actual outcome level (%)	85.3	72.7	86.0	53.2
Predicted level without reform (%)	85.9	73.4	86.8	53.2

Notes: * = statistically significant at 5% level; ** = statistically significant at 1% level. Standard errors clustered at school level. See section 4.3 for further discussion of this.

According to our estimates, the reform reduced the Year 12 FT participation rate of pupils in Group 1 – those who would have been eligible for the full EMA award had it been available – by 1.2 ppts, and this estimate is statistically significant at the 1% level.¹⁷ The effect for Groups 2 or 3, which consists of individuals who would have been eligible for only a partial EMA, meanwhile is smaller at approximately three-quarters of a percentage point. These estimates are statistically significantly different to zero, but not to each other (i.e. we would reject the hypothesis that there was a different effect on group 2 to Group 3). Consequently the average impact across the three groups is estimated at -1.1 ppts, which is statistically significant.¹⁸

¹⁷ This means that if the analysis was repeated many times with different populations, there would be less than a 1% chance of finding an impact of at least this magnitude if the ‘true’ effect was actually zero.

¹⁸ Since the data in question cover the full population of state school pupils, the interpretation of a standard error and associated confidence interval is more nuanced than it would be in a study involving a sample of students (e.g. from a survey). In this analysis there is no immediate sampling error since the data cover the full population of interest. However, the population that is observed could be thought of as drawn from a broader set of potential populations over a longer time period; in future years, the relevant population may look different. Moreover, even if the full population is observed, the standard errors around the point estimates also reflect the presence of an error term in the model. To the extent that it is impossible these statistical models to perfectly fit the data, the impacts they provide are estimated with a certain amount of error.

An important point to consider when looking at the Y12 participation results is the positive, statistically significant coefficient on Group 5.¹⁹ Were this to be used instead of Group 4 as our control group, estimates for all groups would have increased by 0.9 percentage points. This would involve an estimate of closer to -2 percentage points for all three groups.

For Year 13 FT participation the estimated impact of the policy change – using Group 4 as the control group – is stronger. Amongst Group 1, participation is estimated to have dropped by 1.8 ppts as a result of the policy change. The average impact across Groups 1-3 is -1.5 ppts, while the overall effect is on the cohort is -0.9 ppts. All of these estimates are statistically significant. While this analysis cannot explore the mechanisms behind the impacts, one potential explanation for the stronger impact on Year 13 participation is that Groups 1 to 3 in this cohort, which faced the transitional arrangements, would have experienced an actual loss of financial support since they would have received EMA the previous year (2010/11), though the extent of this will have varied across individuals. The bursary cohort, for whom the Year 12 impacts are measured, would have not had any experience of receiving EMA and would therefore not have experienced a loss of financial support. However, the effectively zero coefficient on Group 5 in these regressions suggests that were Group 5 used as the control group in both the Y12 and Y13 regressions the observed difference in the average fall in participation between Y12 and Y13 would be eradicated, since using Group 5 as the control in the Y13 regressions would not affect the results, but using it as the control in the Y12 regressions would increase the estimated effect by almost 1 percentage point.

For attainment, the analysis estimates that among Group 1, the proportion who had achieved at least the L2 threshold by 18 fell by 1.8 ppts, with a corresponding fall across all groups potentially eligible for EMA of 1.5 ppts. As a result, the overall proportion of the cohort achieving the L2 threshold by 18 is estimated to be around 0.9ppts below the

¹⁹ Given the suggestive evidence of common trends between Groups 4 and 5 presented in Figures 5-9, this finding is unexpected. There are two plausible explanations. First, we could be badly misallocating income, and Group 4 could include a number of individuals affected by the reform. This would suggest our estimated impact is conservative, as participation within our control group might have been negatively affected by the reform. Second, common trends between Groups 4 and 5 might have broken down between 2010/11 and 2011/12 due to a factor that differentially affected Groups 4 and 5. If this factor also differentially impacted Groups 1,2 and 3, this represents a cause for concern. One candidate factor is the increase in University tuition fees from £3,000 to £9,000 per year that affected students starting University in 2012 onwards. For this to explain the observed effect, students would have to have opted out of education between ages 16 and 18 due to the increased cost of University. This seems less likely considering the high labour market returns to post-16, pre-University qualifications (e.g see Conlon and Patrignani (2010)). Further, for the effect of this to bias the negatively bias the coefficient, it would require participation in Group 4 to be more negatively affected than in Groups 1,2 and 3 by the tuition change. This seems unlikely; suggesting any bias on the estimates cause by the change in tuition fees would be towards zero. Thus in both cases, the positive coefficient on Group 5 suggests the effect on participation is being underestimated in this setting.

counterfactual level, at 82.8%. If Group 5 rather than Group 4 was used as the control group, all these estimates would have increased by more than 0.4 ppts, suggesting the effects we are estimating may be conservative.

The overall impact on the L3 attainment rate is effectively zero, and statistically insignificant.²⁰ Even in the extreme case of using group 5 as the control, the effect would be approximately -0.4 percentage points, and statistically insignificantly different to zero.

²⁰ Further analysis, not shown here, reveals that the small, statistically insignificant effect on overall L3 attainment (through any route) consists of a statistically significant reduction in L3 attainment through the academic route which is offset by a small (and not statistically significant) increase through the vocational route.

3.2. Subgroup impacts on participation and attainment

The previous section provided an estimation of the overall impact of replacing the EMA with the 16 to19 Bursary for the entire cohort. In this section the analysis is extended to investigate the possibility of variation in the effect for different subgroups. This analysis rests on the assumption of within group common trends, which may not be valid. Subsequently all of the results presented in the section should be treated with caution.²¹

We allow the effect of the policy on participation and attainment to differ for ‘Non-white’, SEN and EAL individuals²² by including interaction terms between the subgroup variables and the treatment dummy variable (see the appendix for more details on this). To improve the interpretability of the results, Groups 1, 2 and 3 are clustered together into one ‘eligible’ group, called ‘EMA’. Groups 4 and 5 are as defined in the previous section.

The results for the subgroup analysis are given in Table 8. As in the main analysis, the effects on FT participation in Y12 and Y13 and on Level 2 and Level 3 attainment are all investigated. The subgroup effects are estimated overall and separately by gender. The ‘baseline effect’ in this table is the estimated effect of the policy on white, non-SEN, non-EAL individuals (in the male (female) regressions the baseline case is white, non-SEN, non-EAL males (females)). All other coefficients are then relative to this group within each column; i.e. the remaining coefficients should be added to the estimated Baseline Effect to arrive at the individual subgroup effects (for example in column 1, the overall average effect on FT Y12 participation for non-white males is equal to approximately $-0.877 + -2.763 = -3.64$ ppts).²³

Table 8 Subgroup analysis: males, females and overall

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Overall				
Baseline Effect	-0.665**	-0.843**	-0.777**	0.730**
Non-White Effect	-2.002**	-3.062**	-0.163	2.247**
SEN Effect	0.662**	0.438*	-1.389**	-4.483**
EAL Effect	-0.968**	-0.736*	-2.099**	0.124

²¹ An investigation into the validity of this assumption will be presented in the final version of the report.

²² We do not present results for more subgroups due to our concerns with the assumption of common trends within subgroups highlighted above. In particular, whilst we acknowledge the possibility of variation within the general ‘non-white’ subgroup, we do not provide analysis investigating further ethnic subgroups due to this concern.

²³ A fully interacted model is not estimated, meaning the effects of each subgroup are assumed to be additive and independent of the status of the remaining subgroups (for example, the FT Y12 participation effect of being a non-white male rather than a white male is -3.6 ppts, regardless of SEN or EAL status in this model). A fully interacted model would allow more flexibility in this context, but is less favourable in terms of interpretability of the coefficients.

Male				
Baseline Effect	-0.877**	-1.450**	-0.635*	0.809*
Non-White Effect	-2.763**	-3.813**	-0.166	1.568**
SEN Effect	1.655**	1.369**	-1.838**	-3.857**
EAL Effect	-1.676**	-0.813	-2.133**	0.223
Female				
Baseline Effect	-0.522	-0.365	-0.812**	0.725*
Non-White Effect	-1.250**	-2.224**	-0.205	2.947**
SEN Effect	-0.336	-0.454	-0.958**	-5.144**
EAL Effect	-0.208	-0.728	-2.037**	0.008

The results show interesting differences in the effect for different subgroups. For Y12 FT participation, the effect is largest for non-white individuals, with the overall effect around -3.6ppts for males and -1.8 ppts for females in this group. The participation effect is also negative relative to the base case for EAL students, though this result is much stronger for males. Perhaps surprisingly, the effect for male SEN students is positive overall for Y12 students, estimated at approximately +0.8ppts. It is plausible that this effect could be attributable to the method of allocation of the Bursary, with initial analysis suggesting that some SEN students may actually be better off under the new scheme. However, the case for this argument is weakened by the lack of an equivalent result for females. The results for the participation of the transition cohort (Y13) are similar to the participation results for the Y12 cohort.

The effect on attainment is rather different, and in some cases seemingly contradictory to the participation effects. For example, for non-whites, whilst Level 2 attainment is not significantly different to the Level 2 attainment of whites (and is negative overall), Level 3 attainment is estimated to have been positively affected by the policy change, for both males and females. This result is rather surprising, but could plausibly be explained by spill over effects, which would require non-white individuals who remained in education after the Bursary to perform better as a result of their cohort consisting of fewer individuals from the lower end of the income distribution. However, it would be surprising if this effect could explain the entire effect, and this is an example where the common trends between ethnicity groups could be questioned.

The SEN individuals the attainment effect is also surprising, with large negative effects on the proportion obtaining Level 2 and Level 3 qualifications – particularly for males – despite the increase in participation amongst this group. A similar line of argument through spill over effects to that given above could theoretically explain the effect, though again, it appears unlikely that this could explain the full effect.

Overall we observe that the participation effect is more negative for non-whites and for non EAL students, and Level 2 attainment is generally worse for all three subgroups than in the base cases for both males and females. The signs and magnitudes of the Level 3 attainment results are such that it might be necessary to acquire a further year of data before any reliable conclusions are drawn in this area.

4. Robustness Checks

As discussed previously, for a difference-in-difference analysis to be valid, the common trends assumption must hold. In this case, the key identifying assumption is that trends in education participation would be the same in Groups 1, 2 and 3 as in Group 4, were it not for the policy change, once underlying trends and group composition are taken into account. The figures presented in Section 2.3 provide suggestive evidence that this assumption does indeed hold. However in this section we investigate the common trends assumption more formally. In section 4.1, we run a placebo difference-in-difference test with 2011 as the treatment year, and in Section 4.2, we test the sensitivity of our results to the exclusion of observations from 2005/06. In both cases we find evidence that supports our assumption of common trends.

4.1. DiD analysis on a placebo treatment year

In this section we present results from a placebo difference-in-difference test in which 2011 is used as the treatment year. If there are truly common trends between groups in the pre-treatment period, the coefficient estimates for this effect should be insignificantly different to zero.

The results from this test are presented in Table 9. The table shows that the coefficients are considerably smaller in magnitude than the corresponding effects reported in Figures 12 and 13, and that of the 32 estimated coefficients, 5 are significantly different to zero, which we feel is sufficient evidence in favour of our assumption of common trends.²⁴

²⁴ Uncertainty in the estimation process means that in expectation, 1 in 20 zero coefficients would be estimated as being significantly different to zero. Although five coefficients is perhaps more than one would expect through natural variation, the overall results are still favourable to our assumptions. This will be investigated in more detail in the final version of the report.

Table 9 Common Trends Assumption: 2011 as the placebo treatment year

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Male				
Group 1	-0.362	-0.361	-0.572**	-0.398
Group 2	0.257	0.340	-0.587	-0.230
Group 3	-0.309	0.010	-0.323	-0.099
Group5	-0.363	0.129	0.094	0.076
Female				
Group 1	-0.458*	-0.397	-0.467*	-0.217
Group 2	-0.743*	-0.140	0.018	-0.524
Group 3	-0.199	-0.246	0.149	-0.004
Group5	-0.062	0.136	0.258	0.147

An equivalent check using 2010 as the placebo treatment year is found to yield very similar results.

4.2. Exclusion of 2006

In this section we test the robustness of our results to the exclusion of 2006 from the estimation. The results for the male and female regressions are presented in Table 10. The results are very similar to the overall results presented in Tables 6 and 7, which supports our assumption of common trends. A repeat of the residuals test given in Table 9 with this shortened sample is similarly favourable (and is available upon request).

Table 10 Common Trends Check: Excluding 2006

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Male				
Group 1 impact (ppts)	-0.972**	-1.928**	-1.748**	0.179
Group 2 impact (ppts)	-0.713	-2.152**	-1.156**	-0.226
Group 3 impact (ppts)	-0.851*	-1.004*	-0.561	-0.176
Impact across Groups 1-3 (ppts)	-0.925**	-1.773**	-1.465	0.073
Impact across whole cohort (ppts)	-0.559**	-1.045**	-0.863**	0.043
Group 5 impact (ppts)	1.052**	-0.341	0.371	-0.278
Female				
Group 1 impact (ppts)	-0.802*	-1.131**	-1.359**	0.091
Group 2 impact (ppts)	-0.193	-1.411*	-0.804*	-0.079
Group 3 impact (ppts)	-0.137	0.329	-0.243	0.455
Impact across Groups 1-3 (ppts)	-0.624*	-0.881	-1.095**	0.144
Impact across whole cohort (ppts)	-0.377*	-0.522*	-0.649**	0.085
Group 5 impact (ppts)	0.885**	0.512	0.296	1.196**

5. Assessment of Value for Money

Section 3 presented impact estimates of the 16 to 19 Bursary Fund on participation and attainment rates. While that analysis is sufficient to answer questions about whether the policy reform led to changes in these outcomes, it is not sufficient to make judgements about the overall effectiveness of the reform. This requires weighing up the impacts caused by the reform against their likely economic implications, both in terms of savings to the Government and in terms of net social benefits. In this section, the headline impacts presented above are used to estimate whether the policy reform delivered value for money (VfM). Also discussed is the method that will be used to estimate the overall net social benefits in the next version of this report due in October 2014.

As shown in Section 3, the introduction of the 16 to 19 Bursary Fund is estimated to have led to reductions in the participation and attainment rates of pupils who would have been eligible for some EMA support; 17 provides a recap of these estimated impacts.

Table 11 Recap of impacts on participation and attainment

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Impact across Groups 1–3 (ppts)	-1.073**	-1.498**	-1.520**	-0.047

Notes: Estimates are taken from. * = statistically significant at 5% level; ** = statistically significant at 1% level.

Overall assessment of VfM of policies such as this is a speculative exercise, particularly when attempting to quantify changes in net social benefits, which will be included in the later version of this report. However even in the current form, the calculations presented here rely on departmental estimates of costs and benefits which are themselves based on many simplifying assumptions, and therefore much uncertainty; these caveats should be borne in mind throughout.

An additional source of uncertainty is the statistical uncertainty in the impacts themselves, which are all estimated with a margin of error (known as a ‘confidence interval’) and should be thought of as a range of plausible values rather than a single number. Throughout this analysis, VfM will be assessed using not only the central estimates of the policy impacts, i.e. those in, but also based on the upper and lower bounds of those impacts using their 95% confidence intervals.²⁵ Table 12 presents these figures: the central estimates are the same numbers as in as 5, while the upper and

²⁵ The confidence interval represents the range of values within which the ‘true’ level of the impact would lie 95% of the time, if the analysis was repeated many times with different populations.

lower bounds represent each end of the 95% confidence interval around those estimates. The analysis that follows will be based on this range of impact estimates.

Table 12 Range of impacts across all EMA-eligible pupils

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Upper bound of impact (ppts)	-1.484	-1.979	-1.838	-0.494
Central estimate of impact (ppts)	-1.073	-1.498	-1.520	-0.047
Lower bound of impact (ppts)	-0.662	-1.016	-1.202	0.400

5.1. Assessment of cost effectiveness

A less speculative assessment of VfM can be carried out by comparing the impacts of the policy reform to its budgetary implications. Normally, when assessing a new policy involving additional expenditure, the cost effectiveness is measured as the improvement in outcomes per unit of additional spending incurred; a higher ratio is more favourable. Here, the reform in question involves a financial *saving* to the Government and overall reductions in participation and attainment outcomes, so a lower ratio – a smaller reduction in outcomes per unit of spending saved – will be more favourable.

The first step in this analysis is to calculate the total headcount changes in participation and attainment outcomes as a result of the policy reform, under not only the central estimates of the policy impacts but also their upper and lower bounds. To do this, the impacts above in 7 are scaled up by the total size of the EMA-eligible group (across Groups 1–3) in the relevant cohorts. For Year 12 outcomes, this is the 2010/11 Year 11 cohort, while for Year 13 outcomes it is the 2009/10 Year 11 cohort. These totals are presented in 19 below.

Table 13 Size of total EMA-eligible group

Year 12 outcomes	314,094
Year 13 outcomes	318,037

Combining these figures leads to the estimated headcount impacts in Table 14. These are the numbers of young people not participating or achieving the Level 2/3 threshold in 2011/12, who would have done so had there been no policy reform.²⁶

Table 14 Headcount impacts on participation and attainment

	Participation		Attainment	
	Y12 FT	Y13 FT	L2 overall	L3 overall
Upper bound of impact	-4662	-6293	-5773	-1571
Central estimate of impact	-3370	-4763	-4774	-150
Lower bound of impact	-2079	-3233	-3777	1272

To calculate a cost-effectiveness ratio, these headcount impacts must be compared to the departmental savings from the reform. The total spending allocated for the 16 to 19 Bursary Fund was £180m in 2011/12; in 2010/11, the last year of EMA, £554m was spent on EMA (as per figures provided by DfE). Assuming that the same amount would have been spent on EMA in 2011/12 had it not been replaced, there is therefore a saving of £374m in 16–19 financial support. However, this is not the only Government saving, since any reductions in post-16 participation also involve a saving in the form of reduced expenditure on the provision of post-16 education places; this saving is greater the larger the reduction in participation. A figure of £4,644 was provided by DfE as its estimate of the 16 to 19 unit of funding in 2011/12.²⁷ Applying this figure to the reductions in Year 12 and 13 FT participation, and adding it to the £374m saving in financial support, leads to the total departmental savings in Table 15.

Table 15 Departmental savings due to 16 to 19 Bursary Fund (£m)

	Financial support	Cost of education provision	Total
Upper bound of impact	374	50.9	424.9
Central estimate of impact	374	37.8	411.8
Lower bound of impact	374	24.7	398.7

²⁶ The numbers presented in this table do not take account of those induced to leave part time education as a consequence of the policy change, since we cannot provide a reliable estimate of this. If the proportion in part time education dropped as a result of the policy, these – and all subsequent figures – represent an underestimate of the true numbers.

²⁷ This estimate was derived from 2011/12 provider funding and volumes that will include some part-time provision. The funding also excludes the Teachers' Pay Grant allocated to local authorities, who then pass it on to schools using their own allocation methodology.

Finally, the estimated cost-effectiveness ratio is calculated by dividing the headcount impacts in 20 by the departmental savings in 21. The ratios are presented below in Table 16, and are expressed as the change in the total number of participants (summed across Year 12 and Year 13) and the number of Level 2/3 achievers, per £1m in departmental savings.

Table 16 Change in number of participants and achievers per £1m saved

	Participation	Attainment	
	Y12 + Y13 FT	L2 overall	L3 overall
Upper bound of impact	-25.8	-13.6	-3.7
Central estimate of impact	-19.8	-11.6	-0.4
Lower bound of impact	-13.3	-9.5	3.2

Based on the central impact estimate, there were 20 fewer participants in FT education, 12 fewer L2 achievers per £1m saved. The estimated change in the number of L3 achievers is zero. The ‘worst-case scenario’, based on the upper bound of the impacts, is 26 fewer FT participants, 14 fewer L2 achievers and 4 fewer L3 achievers²⁸ per £1m saved. The ‘best-case’ scenario, based on the lower bound of the impacts, is 13 fewer FT participants, 10 fewer L2 achievers and 3 more L3 achievers²⁹ per £1m saved.

Does this represent good value for money? At this stage it is hard to say as in order to fully understand this, we need to have some assessment of the long-term costs of increasing school drop-out rates and reducing overall attainment (in terms of the likely reduction in taxation receipts and increase in benefit payments) and then comparing the present value of these costs to the departmental savings. This will also enable calculate an effective Internal Rate of Return (IRR)³⁰ of the cut in the program and make direct comparisons with other programs in terms of cost effectiveness under different assumptions. The way we plan to do this is highlighted in the next section.

5.2. Assessment of net social benefits

In the final report, we plan to look in more detail at the likely long-term implications for individuals and the exchequer so that we can assess the overall costs and benefits of the policy. This will involve a number of stages.

²⁸ The increase in L3 achievement under the ‘best-case’ scenario is due to the fact that the impact on L3 achievement is not statistically significant: its lower bound involves a slight increase in attainment.

²⁹ The increase in L3 achievement under the ‘best-case’ scenario is due to the fact that the impact on L3 achievement is not statistically significant: its lower bound involves a slight increase in attainment.

³⁰ The IRR is a standardised measure used to compare the value for money of policies by considering the long run benefits and the short term costs of a policy. In this case the IRR would consider the long run costs against the short run savings of the policy.

The first will involve estimating the longer term *education* effects of the change in policy – most importantly the likely impact on higher education participation (other education levels will be directly estimated as part of our impact analysis). To do so, we will use the Longitudinal Study of Young People in English (LSYPE) because the LSYPE is the closest contemporary cohort for whom we observed HE outcomes. Another possibility is to use linked NPD/HESA data for slightly older cohorts (which the IFS already has access to). For each individual we will estimate the probability of achieving an educational outcome under the old EMA scheme and with the new 16to 19 Bursary. The four different educational levels we plan to use are: (1) less than 5 GCSEs A*-C (< L2) (2) 5 or more GCSEs A*-C (L2) (3) A-levels (L3), and (4) a university degree.

Having done this, we will simulate lifetime earnings and employment profiles for individuals who we predict will eventually have attained one of four different educational levels: (1) less than 5 GCSEs A*-C, (2) 5 or more GCSEs A*-C, (3) A-levels, and (4) a university degree under the old EMA system and under the new 16 to 19 Bursary scheme.

In order to do so, we will estimate, for each educational group and gender, a rich statistical model of earnings and employment dynamics that allow us to take into account the likely persistence of earnings and employment shocks (e.g. a recession). The model will be estimated using two large data sets that contain survey information on British individuals' labour market outcomes – the Labour Force Survey (LFS) and the British Household Panel Survey (BHPS). The statistical model will generate cross-sectional earnings distributions that are consistent with the high-quality cross-sectional data from the LFS. Transitions between employment and non-employment, and year-on-year earnings fluctuations, are consistent with the dynamics observed in the BHPS. The IFS has used this approach in assessing the cost-effectiveness of other programmes (e.g. for the Effective Pre-School, Primary & Secondary Education (EPPSE) project study for DfE³¹ and also for assessing the implications of HE funding reforms³².

For each educational category and gender, we will use the estimates of the corresponding earnings and employment model to simulate artificial earnings and employment paths of a cohort of 10,000 individuals.

The next step will consist of computing, for each simulated lifetime profile of gross earnings, the corresponding profile of net earnings, along with the amount of tax paid and benefit received. To do so, we will use The Institute for Fiscal Studies' tax and benefit model to calculate the amount of tax paid and benefits received by each individual in our data assuming firstly they were under the old EMA scheme and then secondly that they

³¹ See Cattan, Crawford and Dearden (2013), 'The long term consequences of pre-school education and quality', forthcoming DfE report.

³² Chowdry, Dearden, Goodman and Jin (2012), 'The Distributional Impact of the 2012 Higher Education Funding Reforms in England', *Fiscal Studies*, Vol. 33, No. 2.

were under the new 16 to19 Bursary Scheme using the current tax regime (2014/15). This will create an additional difficulty because the tax and benefit system in the UK is intrinsically dependent on a number of characteristics, which we do not observe for individuals in our cohorts. In particular, it depends on the individuals' future family structure (marital status, partner's age, number and age of his/her dependent children), his/her number of hours worked, his/her partner's number of hours worked and earnings, his/her region and housing situation (whether he/she is a renter, the value of the rent, and council tax band). We will therefore need to predict all these characteristics for each period the individuals in our dataset will spend in the labour market.

To do so, we will first estimate the empirical distribution of these characteristics in the BHPS for each gender-education-age groups (we use weights so as to approximate a nationally representative distribution). We will then use these estimates to predict these characteristics for each period the individuals in the 16 to19 bursary dataset spend in the labour market (under both scenarios). Doing so is the only way we will be able to compute the net earnings profiles for these individuals, but it adds another level of uncertainty to our calculations of the effect of the introduction of the 16 to 19 bursary on lifetime net earnings and savings to the Exchequer.

We will perform two calculations. Firstly we will assume that the individual remains single and without children all his/her life in order to compute the change in net earnings from the change in policy. This is a less realistic scenario, but it requires us to rely only on a limited set of predicted characteristics and therefore to minimize the error we are adding through this step. This will also give us an estimate of the savings to the Exchequer at the individual level.

Second, we will allow individuals to marry and to have children. We will use the predicted earnings and hours of work of the partner to compute the gross earnings of the household and use all the predicted characteristics to compute the net earnings of the household. This is a more realistic scenario, especially because it accounts for the fact that, if the introduction of the program reduces the final educational outcome level an individual reaches, he/she will have a family structure and partner's labour outcomes that are likely to be different from the one he/she would have otherwise.

The final step will involve predicting the average effect on gross and net lifetime earnings and employment associated with the introduction of the 16 to19 Bursary. First we will predict the probability that each individual in our sample would have to attain each of the four educational levels if she had been entitled to the EMA and then if only entitled to the 16 to19 Bursary Scheme. To each individual in our sample, we will then match 400 (100 per education group) possible profiles of earnings and employment from each of the artificial cohorts simulated from our model. We will compute the discounted present value of lifetime gross and net earnings conditional on being in each educational category. Finally, we will compute a weighted average of discounted present value of lifetime earnings using the probabilities associated with each individual and each

counterfactual. For each counterfactual, we then average these results across all individuals. The difference between the averages in gross (net) earnings is the predicted average effect of the 16 to19 bursary on lifetime gross (net) earnings. We will compute the losses to the Exchequer by comparing the difference between lifetime gross and net earnings in the case all individuals were under the EMA scheme and the case all individuals are instead under the 16 to19 Bursary scheme. The net present value of this loss to the Exchequer can be directly compared to the savings from the introduction of the scheme. Of course predicting earnings 40 years into the future, involves a lot of uncertainty and speculation. We will perform a number of sensitivity checks to these estimates in order to gauge the magnitude of some of this uncertainty. For instance, we can look at the sensitivity of these estimates using different discount rates and high and low earning growth scenarios. Although the results will be necessarily speculative, the exercise remains interesting to conduct, and will help us ascertain the cost effectiveness of the programme and allow direct comparisons with other program interventions.

6. Conclusions

The 16 to 19 Bursary Fund was implemented nationally in September 2011 as the replacement for the previous Education Maintenance Allowance (EMA). The policy intention was to provide more efficient targeted support for post-16 learning, by delegating responsibility to sixth forms and colleges and allowing them to allocate it the learners deemed most in need of additional support.

The analysis in this report uses statistical techniques to provide estimates of the impact of replacing EMA with the 16 to 19 Bursary Fund. The outcomes analysed are whether learners stayed in full-time (FT) participation in post-16 education and whether they had achieved the Level 2 or 3 threshold by age 18. The estimated impacts in this research briefing should be interpreted as the changes in participation and attainment rates compared to a hypothetical no-reform scenario where the EMA had been retained in 2011/12. They are not the impacts compared to a scenario of no 16 to 19 financial support.

Measuring the impacts of the policy reform is complicated by the fact that the reform was implemented across England at the same time (i.e. there was no control group against which to measure the impact). It is therefore challenging to identify what participation and attainment rates would have been in 2011/12 *had there been no reform to EMA*. Reliably estimating this quantity is crucial, as it is the baseline to which the actual levels of participation and attainment in 2011/12 should be compared, in order to isolate impacts that can be attributed to the policy reform itself.

Briefly, the analysis in this report compares the change in post-16 education outcomes for lower-income pupils – who would have been eligible for EMA in 2011/12 had it been retained – against the change in the same outcomes for pupils whose family income was slightly too high for them to be eligible for EMA in 2011/12 (had it been retained). The latter group is assumed to have been unaffected by the reform, since they would not have been eligible for EMA. Furthermore, it is assumed that the change in their observed post-16 education outcomes is a reliable guide to the change in education outcomes that would have been seen among lower-income pupils in 2011/12 had there been no policy reform. Pupils with family incomes just above the EMA eligibility income threshold are used as the basis for comparison, in order to maximise the validity of this assumption.

Based on this empirical approach, the headline impacts indicate that the implementation of the 16 to 19 Bursary Fund in 2011/12 led to a 1.2 ppt fall in FT participation amongst Year 12 students who would otherwise have been eligible for the full EMA award. In other words, their participation rate in FT education in Year 12 would have been 1.2 ppts higher in 2011/12 had there been no reform to EMA. The effect on those eligible for a partial EMA are smaller, resulting in an estimated overall effect on those eligible for any form of EMA of -1.1ppts. Looking at the Year 13 transition cohort, there was a 1.8 ppt fall in FT participation among the poorest students who would have previously been eligible for the full EMA, and a smaller but still statistically significant effect on those who would

have been eligible for a partial EMA. Overall, there was a statistically significant fall in FT participation of 1.5 ppts across all EMA-eligible Year 13 pupils, and 0.9 ppts across the whole cohort.

The findings for attainment again suggest that the impacts were the most negative among the poorest students: among those who would have been eligible for the full EMA award, there was a 1.8 ppt fall in the L2 achievement rate, leading to a 0.9 ppt fall across the whole Year 13 cohort. The effects on L3 attainment were considerably smaller and insignificantly different to zero.

Combining the estimated overall participation effects suggests that full time Year 12 and Year 13 participation dropped by 8,100 individuals as a consequence of the policy change.

Exploration into the variation in the impacts of the reform by pupil characteristics finds that there was no strong difference between males and females in the overall effect on either participation or attainment. Further subgroup analysis suggests the reduction in participation was significantly larger amongst non-whites than amongst whites, and amongst EAL students against non-EAL students.

For males, the analysis finds that there is a positive overall effect on participation of SEN students in both Y12 and Y13, which is potentially attributable to increased eligibility for grants for SEN students under the Bursary. However, the result does not hold for females.

Investigating Level 2 attainment by subgroup suggests negative overall effects and stronger negative effects for both SEN and EAL students, for both males and females. There is no statistically significant difference in Level 2 attainment for whites and non-whites.

For Level 3 attainment, meanwhile, the results of the subgroup analysis are quite inconsistent, and are therefore difficult to extract conclusions from. For non-whites Level 3 attainment is estimated to have improved under the Bursary for both males and females. For SEN students, Level 3 attainment appears to have reduced, while there is no significant difference between Level 3 attainment for EAL and non-EAL students.

Importantly, it is likely that the overall impact estimates presented in this report are conservative, and may *underestimate* the true impacts of the policy reform in question. This is a consequence of the empirical strategy, in particular the use of proxy information on socioeconomic characteristics in the absence of actual income data to determine potential EMA eligibility. Doing so carries a risk of misclassification: some pupils may have been incorrectly deemed to be eligible for EMA while others may have been incorrectly deemed to be ineligible for EMA. This could bias upwards the outcomes of the EMA-eligible group and bias downwards the outcomes of the comparison group, thereby understating the negative impacts of the reform.

Combining the headline impact estimates with the departmental savings arising from the reform, this report finds that the reform led to 20 fewer FT participants and 12 fewer Level 2/3 achievers per £1m saved in 2011/12. In the 'worst-case' scenario, based on the largest plausible impacts, the reform led to 26 fewer FT participants and 17 fewer Level 2/3 achievers per £1m saved. In the 'best-case' scenario, based on the smallest plausible impacts, the reform led to 13 fewer FT participants and 6 fewer Level 2/3 achievers per £1m saved.

Whilst these figures are encouraging, it is only possible to assess the comparative effectiveness of the programme by looking at long-term costs and benefits under different assumptions. This will be an important component of our final report and the method we propose to use should give a much clearer idea of the overall effectiveness of the introduction of the new 16 to 19 Bursary. This will also enable direct comparisons of its effectiveness compared with other initiatives.

References

Conlon, G. and Patrignani, P. (2010), *Returns to BTEC vocational qualifications*, London Economics.

Dearden, L., Emmerson., C., Frayne, C. and Meghir, C. (2009), “Conditional Cash Transfers and School Dropout Rates”, *Journal of Human Resources*, Vol. 44, No. 4, pp. 827–857.

Kendall, L., O'Donnell, L., Golden, S., Ridley, K., Machin, S., Rutt, S., McNally, S., Schagen, I., Meghir, C., Stoney, S., Morris, M., West, A. and Noden, P. (2005), *Excellence in Cities: The National Evaluation of a Policy to Raise Standards in Urban Schools 2000-2003*, Department for Education and Skills Research Report RR675A.

Tanner, E., Brown, A., Day, N., Kotecha, M., Low, N., Morrell, G., Turczuk, O., Brown, V., Collingwood, A., Chowdry, H., Greaves, E., Harrison, C., Johnson, G., and Purdon, S. (2011), *Evaluation of Every Child a Reader (ECaR)*, Department for Education Research Report RR114.

Appendix 1

Additional figures on pre-reform participation trends

Figure 7 Pre-reform common trends in female full-time Year 12 participation

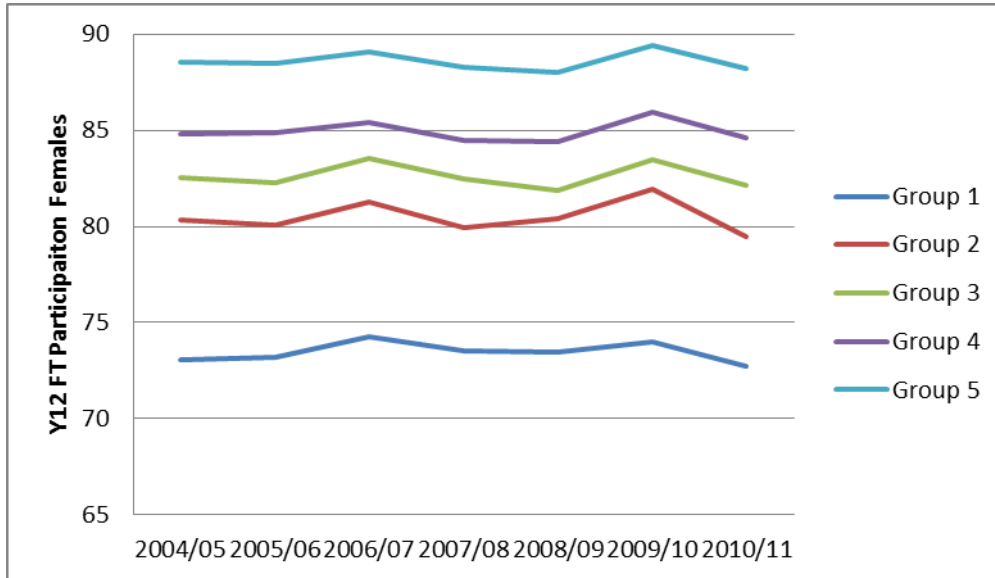
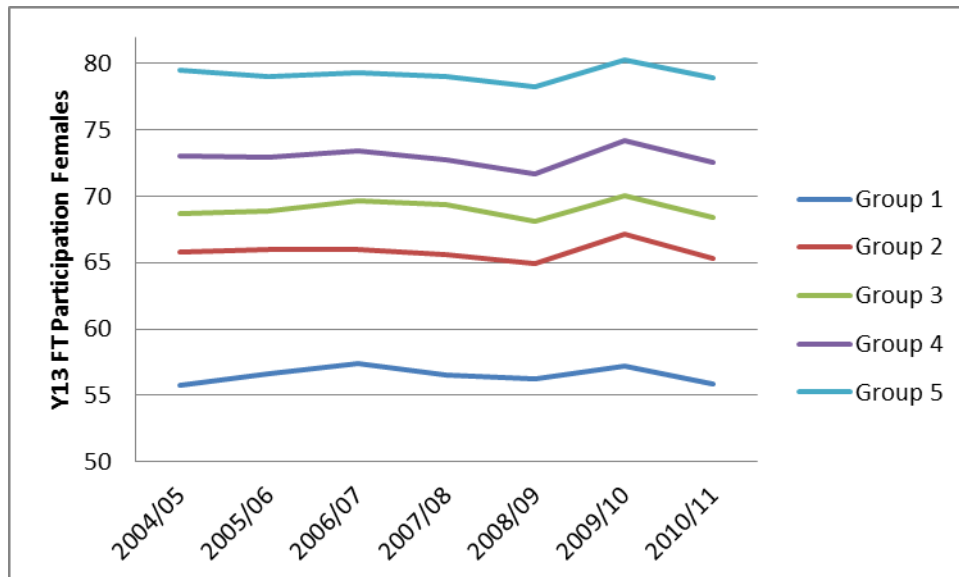


Figure 8 Pre-reform common trends in female full-time Year 13 full-time participation



Additional figures on pre-reform attainment trends

Figure 9 Pre-reform common trends in Female Level 2 threshold achievement

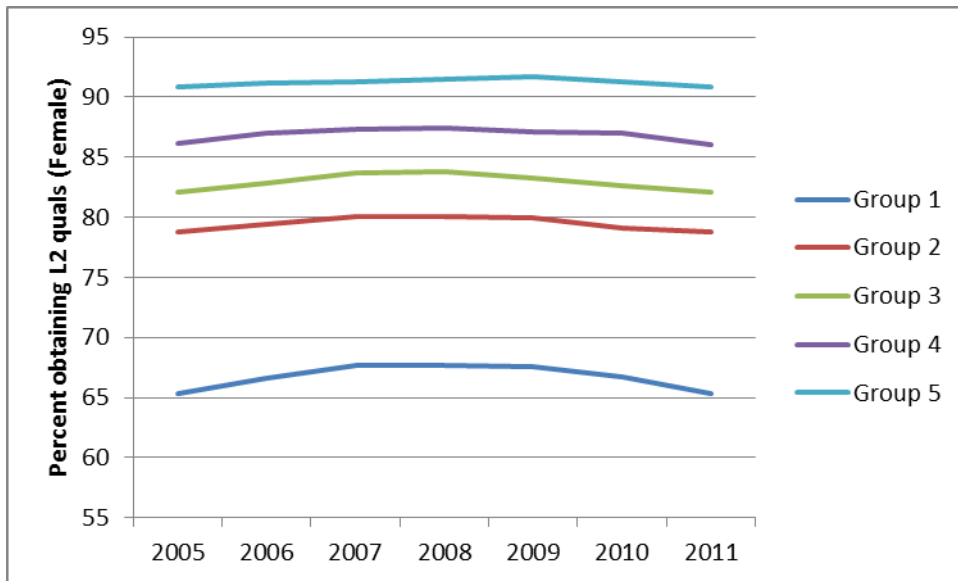
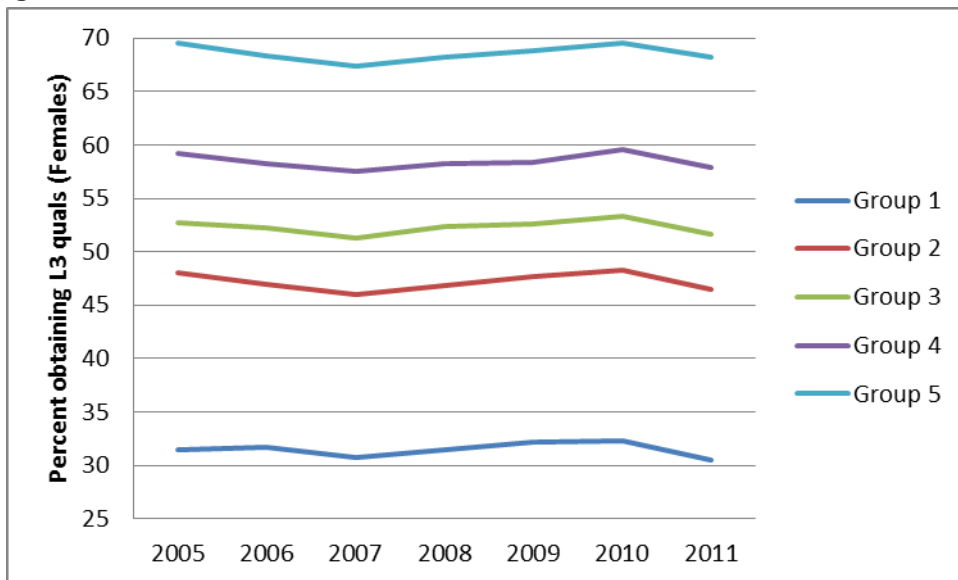


Figure 10 Pre-reform common trends in Female Level 3 threshold achievement



Appendix 2

Creation of socioeconomic index

The index was created by principal components analysis, combining the following factors:

- FSM status;
- IMD score;
- IDACI score;
- ACORN group;
- Neighbourhood proportion of owner-occupied households.

The first principal component was extracted from this, using the factor loadings given in Table 17. The principle component explained 66% of the total variance in the above factors.

Table 17 Scores used to create socioeconomic index

Factor	Factor loading
FSM	0.27
IMD score	0.50
IDACI score	0.51
ACORN group	0.44
% owned households	-0.47

Econometric specification

The statistical models used in this analysis are all least squares regression models (since the outcome indicators are all binary, they are also linear probability models). The estimating equation for the headline impact analysis is as follows:

$$y_{it} = \alpha + v_s + \beta X_{it} + \gamma_g D_g + \lambda P_t + \delta_g (D_g \cdot P_t) + \eta t + \theta_g (D_g \cdot t) + \varepsilon_{it} \quad (1)$$

Where y_{it} is the participation or attainment outcome for pupil i at time t ;

α is an intercept;

v_s is the school fixed effect;

X_{it} is the set of characteristics for pupil i at time t controlled for in the model. These are gender, ethnicity, SEN status, EAL status, KS2 average point score (entered linearly), KS4 capped point score (entered linearly), and the neighbourhood proportion of residents with a Level 4/5 qualification, obtained from the 2001 Census (entered linearly);

D_g is a set of income group indicators (the comparison group, Group 4, being the omitted category);

P_t is an indicator for the post-reform period, i.e. 2011/12;

$D_g \cdot P_t$ is an interaction term which identifies the outcomes observed among group g in the post-reform period 2011/12. The coefficients on this interaction term, δ_g , are the impact estimates;

t is an aggregate linear trend in the outcome (effectively the trend for the comparison group);

$D_g \cdot t$ is a separate linear trend for group g ;

ε_{it} is an error term representing all other unmeasured influences on y_{it} for pupil i at time t .

For the subgroup analysis, the statistical model is the same except that the subgroup characteristic in question (e.g. gender) is omitted from X_{it} since it is collinear.

In all the models, the standard errors are clustered at the school (in Year 11) level. This assumes that ε_{it} is independent across schools, but can be correlated across different pupils in the same school, and can also be correlated over time within the same school.

Subgroup Analysis

For the subgroup analysis, the estimated equation is as above, with the additional inclusion of interaction terms between the Non-white, EAL and SEN dummies with the post treatment dummy, as shown below. To simplify the analysis Groups 1, 2 and 3 are clustered together to create an 'EMA' dummy, set equal to one if the individual is eligible for any form of the EMA.

$$\begin{aligned}
 y_{it} = & \alpha + v_s + \beta X_{it} + \gamma EMA + \lambda P_t + \delta_1(EMA \cdot P_t) + \delta_4(D_4 \cdot P_t) + \delta_5(D_5 \cdot P_t) \\
 & + \eta t + \theta_1(EMA \cdot t) + \theta_4(D_4 \cdot t) + \theta_5(D_5 \cdot t) \\
 & + \rho_1(SEN \cdot EMA \cdot P_t) + \rho_1(EAL \cdot EMA \cdot P_t) + \rho_1(NonW \cdot EMA \cdot P_t) + \varepsilon_{it}
 \end{aligned}
 \tag{2}$$

Where $SEN \cdot EMA \cdot P_t$ is equal to the interaction between a SEN dummy, a post dummy and the EMA dummy, where as previously, post is set equal to one for 2011/12. Similarly $EAL \cdot EMA \cdot P_t$ is an EAL-EMA-Post interaction term and $NonW \cdot EMA \cdot P_t$ is a Non White-EMA-Post interaction term. The ρ_i 's are the coefficients of interest. All remaining variables are as described in equation (1).

Bootstrapped Standard Errors

A potential area of concern is the fact that 'Group' is only estimated from our process. The standard errors from the main specification regressions assume Group is accurately allocated and therefore do not incorporate this additional degree of uncertainty. We

investigate the validity of the standard errors (which will affect the statistical significance of our findings) by recalculating them using a bootstrap technique.

The technique involves taking a subset of all individuals included in the regressions and re-estimating the entire model. This includes re-running the principle components analysis which is used to allocate individuals an income (and hence Group). Because income is allocated based on relative rank in socio economic indicator estimated in the PCA, income might differ for the same individual in two different sub-samples depending on the composition of the remainder of the subsample. The regression is re-estimated with each repetition, with variation across subsamples creating variation in the regression coefficients. The standard errors are then equal to the standard deviations of these regression coefficients.

This process is extremely computationally burdensome due to the large sample sizes, and consequently not all standard errors presented in this report have been bootstrapped: all standard errors will be checked in the final version of the report. In any case, our initial results suggest that the standard errors are not particularly affected.



Department
for Education

© Institute of Education & Institute for Fiscal Studies [2014]

Ref: DFE- RR346

ISBN: 978-1-78105-332-4

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education

Any enquiries regarding this publication should be sent to us at:

Jonathan.Johnson@education.gsi.gov.uk or www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications