



Department
for Education

16 to 19 Additional Hours evaluation

**Feasibility of impact evaluation:
summary report**

September 2024

**Seemanti Ghosh, Daniel Muir,
Becci Newton: Institute for Employment
Studies**

This research report was written before the new UK government took office on 5 July 2024. As a result, the content may not reflect current government policy.

Contents

Key points	2
Background	3
Data to support impact evaluation	4
Implementation of Additional Hours	5
Consideration of options for counterfactual impact analysis	7
Other evaluation approaches that may help identify the effect of Additional Hours	10
Recommendation	16

Key points

- To compensate for learners' lost time in education during the pandemic, the Department for Education provided additional funding to post-16 education providers to increase contact time from September 2022. This feasibility study explores how the impact of this on student outcomes might be identified.
- The study explores how the policy consists of two elements: the first, causing an increase of 40 hours to the minimum requirement on planned hours for students in funding bands 5-9 (and lower bandings on a pro-rata basis). For example, the minimum requirement for funding band 5 increased from 540 planned hours in the previous year to 580 planned hours. The second element of the policy expected an average increase of 40 planned hours by provider for band 5 students (and lower bandings on a pro-rata basis) compared to hours delivered two years prior.
- Applying counterfactual methods, that show the causal impact of the changed funding and additional hours in education, is problematic. The policy is universal, so a randomised controlled trial is not suitable. While building a comparison group should be possible, constraints including differences between UK nation education systems and the disruption of the pandemic affect the plausibility of comparisons between nations or of cohorts pre/post policy in England. As such, these approaches that seek to estimate the causal impact of the policy are not recommended.
- The available data can support exploratory analysis to show how the policy may have affected student outcomes, but it will not be possible to attribute the effect specifically to the policy so is a step down from assigning causality.
- There are a range of options available for exploratory analysis, which have differing benefits, but all have flaws. No single option would provide a reliable assessment but all in combination could build a picture of how the policy may have affected student outcomes. The questions that could be addressed are:
 1. What is the difference in the change in outcomes for students whose planned hours were below the new minimum requirement in 2021/22 that were then increased to (or just over) the new minimum requirement in 2022/23 compared to the change in outcomes for students whose planned hours were already at (or just over) the new minimum requirement in both 2021/22 and 2022/23? To some degree, this might reflect the effect of the first element of the Additional Hours policy.
 2. What is the difference in outcomes for providers that increased their hours from 2020/21 to 2022/23 compared to providers with similar hours in 2020/21 that did not change to 2022/23? To some degree, this might reflect the effect of the second element of the Additional Hours policy.

3. What is the difference in outcomes for students at providers that increased their hours from 2020/21 to 2022/23 in the 2022/23 cohort compared to students at the same providers in the 2021/22 cohort? To some degree, this might reflect the effect of the second element of the Additional Hours policy.
 4. How do the estimates produced to answer research questions 1, 2 and 3, which to some degree might reflect the effect of the Additional Hours policy, vary depending on how the level of qualification and non-qualification hours changed? To some degree, this might reflect the differences in effect of using the additional hours for qualification versus non-qualification time.
 5. How do the estimates produced to answer research questions 1, 2 and 3, which to some degree might reflect the effect of the Additional Hours policy, vary depending on which specific uses the respondents to the provider survey indicated they used additional hours provision for? To some degree, this might reflect the differences in effect of using the additional hours for various specific uses.
 6. How do the estimates produced to answer research questions 1, 2 and 3 (or variations of these given the intended implementation of Additional Hours for T level and high needs students), which to some degree might reflect the effect of the Additional Hours policy, vary depending on whether students are taking T levels, have high needs, are from socio-economically disadvantaged backgrounds, or failed GCSE English or maths at the end of Key Stage 4? To some degree, this might reflect differences in effect of the Additional Hours policy for different sub-groups?
- However, these analyses would produce results that would not be able to attribute causality, would be hard to interpret or draw any reliable conclusions from, and could be misconstrued. As such, our **final recommendation** is to not pursue any further exploratory analysis.

Background

Learners lost a significant amount of classroom based time and particularly time with teachers and other school staff during the pandemic, which has and will continue to have an impact on the potential of learners coming through 16-19 learning now. As part of the government's long-term education recovery plan, the Additional Hours funding was introduced from the 2022/23 academic year in 16-19 education to enable more teaching and learning hours to help compensate for this learning loss. The policy sought to affect the number of guided learning hours students receive by increasing the minimum annual required planned hours from 2022/23 by 40 for students on band 5 or higher (and lower bandings on a pro-rata basis), and by setting out an expectation for providers to increase

their average planned hours by 40 in 2022/23 compared to 2020/21 for band 5 students (and lower bandings on a pro-rata basis). The policy is universal, not targeted at disadvantaged learners or places, with the additional hours provision expected to be tailored to the needs of the student.

Data to support impact evaluation

We detail the findings from desk research and investigations of the key datasets that cover 16-19 education that would be used in an impact evaluation – the National Pupil Database and the Individualised Learner Record. These datasets record information on:

- a student's level of planned hours, which can be used to identify whether Additional Hours affected their planned hours in the ways we would expect based on the policy guidance;
- background characteristics (including outcomes from periods prior to the policy being introduced) which could be used to 'match' students that were treated by the policy (so had higher planned hours as a result of Additional Hours) to similar students that did not receive the treatment, or used to control for differences in the prevalence of these characteristics which will affect the outcomes of students (aside from policy itself) in any treatment and comparison groups (regardless of whether or not, for other reasons that will be discussed, an impact evaluation is deemed feasible);
- measures of the outcomes of interest – final versions of all the various relevant datasets would not be available for the prospective impact evaluation due to its' planned timeline, preventing specifically any analysis of students' destinations beyond 16-19 education and qualification achievement and attainment at these destinations (regardless of whether or not for other reasons that will be discussed an impact evaluation is deemed feasible).

Given the differences in the specific outcome measures that can be constructed from each dataset, any analysis would need to be run separately for those students appearing (primarily) in the NPD versus the ILR. Some of the approaches to evaluating the impact of the policy require outcomes to be measured at multiple points in time at the student level, which is not the case for some outcomes including qualification achievement and attainment.

One issue when considering our ability to measure the effect of the policy on hours (with implications for how treatment with the policy is thought of) is the distinction between *planned* and *received* hours. The datasets capture data on *planned* hours, and this is also the unit that is used to implement Additional Hours and monitor compliance with the policy. However, the theory of change underlying the policy is based on the amount of hours a student *receives*. Differences between a student's *planned* hours and the hours

they *receive* might be related to both the likelihood of them receiving treatment with Additional Hours as well as their outcomes in the absence of treatment, which introduces a potential source of bias into any approach based on this data. Other factors that might lead to differences in a student's *planned* hours and the hours they *receive* include the national (and additional regional) teacher strikes that took place in England in the 2022/23 academic year. This will affect the actual change in *delivered* hours that students receive before and after Additional Hours was introduced for reasons unrelated to the policy itself.

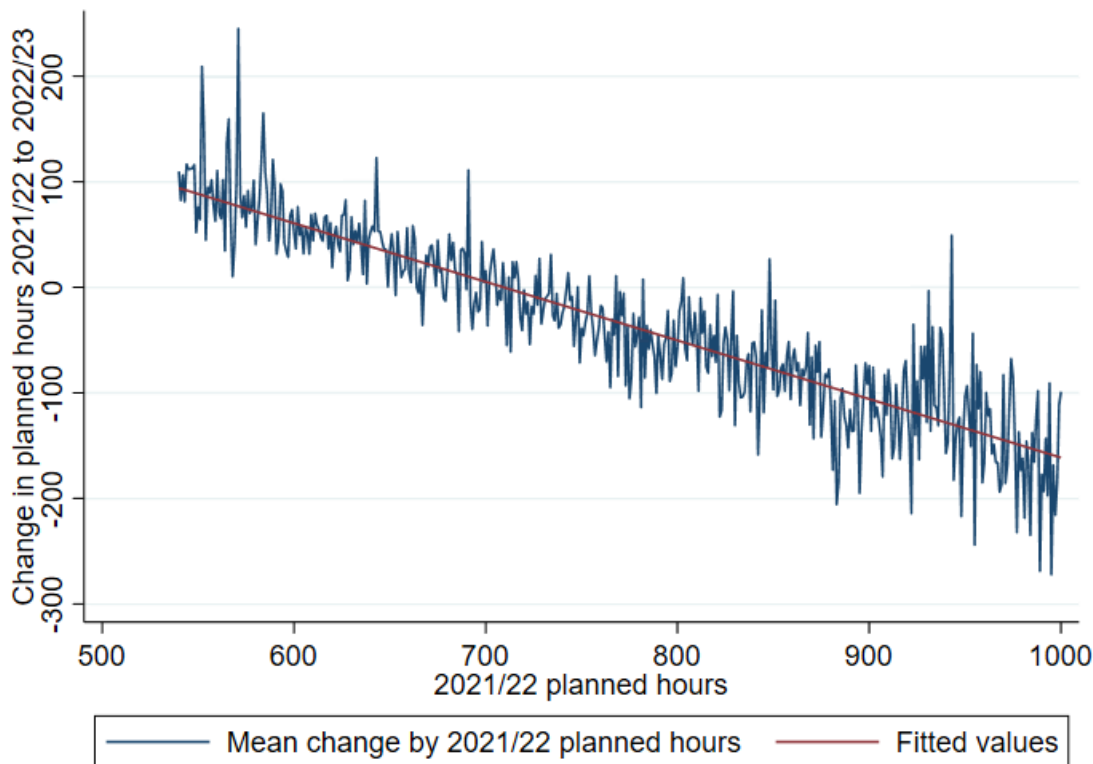
Implementation of Additional Hours

The Additional Hours policy consists of two elements. The first (raising of the minimum hours requirement by funding band) will be seen as a change in student level planned hours around the minimum requirement from the 2021/22 academic year to the 2022/23 academic year. Meanwhile, the second (the expectation to increase average planned hours by 40 from provider baseline for band 5 students and a proportionate amount for those on lower bands) will be seen as an increase in the average number of planned hours amongst band 5 students by provider in the 2022/23 academic year compared to the 2020/21 academic year. While each element should affect planned hours (either at the student or provider level) in a certain way, just because we see this expected change in planned hours does not mean that this can be attributed to the policy. At the student level, there may be natural variation in the number of planned hours across different years of a course not relating to Additional Hours, while at the provider level changes in the composition of the student body and the courses they are taking (which is outside of the direct control of the provider) will also affect average planned hours. As such, this introduces the potential for misclassification of providers/ students as being treated by the policy or not, meaning that the degree of compliance and non-compliance with the policy (considering both elements separately) could be either overstated or understated. Additionally, changes in student level hours will lead to changes in provider level hours, and vice versa. This also makes assessing compliance with each of the two elements of the policy separately challenging.

Analysis of NPD Autumn Census pupil level data from 2021/22 and 2022/23 shows that, despite an overall fall in planned hours year-to-year for students in 16-19 education in both years that are in band 5 in 2021/22 and with recorded hours data in both years, students with lower planned hours in 2021/22 and particularly those with planned hours below 580 do see a substantial increase in their planned hours year-to-year. Figure 1 shows the average (mean) change in planned hours for these students (aged 17 at the start of the 2022/23 academic year and in band 5 in 2021/22) from 2021/22 to 2022/23 by their level of 2021/22 planned hours, excluding the approximately 1% of students with high outlier planned hours above 1000 in 2021/22. The blue line displays the average change in hours for all students with each level of 2021/22 planned hours (i.e., the mean

change in hours for students with 540 planned hours in 2021/22, with 541 planned hours in 2021/22, etc.). Overlaid in red is the line of best fit.

Figure 1: Change in planned hours from 2021/22 to 2022/23 by student, by the students' level of planned hours in 2021/22 (for 2021/22 band 5 students) ¹



Source: ONS

The graph demonstrates a clear strong negative relationship between the change in planned hours by students from 2021/22 to 2022/23 and the student's level of planned hours in 2021/22. It also shows that students with fewer planned hours in 2021/22 typically experience an increase in their planned hours, of around a 100 hour increase for those with planned hours below 580 (i.e., the new minimum requirement) in 2021/22. This suggests that broadly speaking the increase in the minimum requirement had the desired effect – band 5 students with planned hours below 580 in 2021/22 largely had their hours increased to above 580 in 2022/23.

Next, looking at the provider level, it is possible to observe the change in average planned hours that would be expected as a result of Additional Hours. According to data from the NPD and ILR on provider level average planned hours provided by the Department for Education, there was a mean increase of 43 hours across band 5 students from 2020/21 to 2022/23, with 55% of providers seeing average planned hours

¹ Note: a low underlying count prevents reporting for students with 952 planned hours in 2021/22.

increase by at least 40, and 79% seeing some increase in hours. However, 1% saw no change in average planned hours and 20% saw planned hours fall. This significant degree of possible 'non-compliance' with the second element of the policy creates the potential for selection bias, whereby characteristics associated with student outcomes are also associated with the likelihood of receiving treatment with Additional Hours, which would bias any estimates of the impact of the second element of the policy.

The implementation and process evaluation produced several findings with implications for an impact evaluation. It highlighted how the flexibility in the guidance led to wide variation in how providers implemented the additional hours, with some providers implementing them in ways which would likely lead to a misclassification of the extent to which they complied with the two elements of the policy based on the data that would be used for the impact evaluation (for instance, by using the funding to staff previously unsupervised but timetabled hours). Internal resource availability and financial pressures were the two most common factors cited in the provider survey that made delivery of Additional Hours more challenging (both 45%). These factors will affect both the likelihood of a student/ provider 'complying' with the policy (that is, receiving treatment in the form of additional hours) as well as the student's/ provider's untreated outcomes (so what would have happened to them in the absence of any treatment), which will result in estimates being biased.

The issues highlighted in this section, as well as those relating to the data available, affect the ability of a quantitative evaluation to produce a robust estimate of the impact of Additional Hours on student outcomes (broadly) regardless of the specific approach used. As such, these issues are in addition to any further issues highlighted with regards to the specific approaches discussed below.

Consideration of options for counterfactual impact analysis

Of the options for a counterfactual impact evaluation of Additional Hours, a randomised controlled trial would provide the most robust evidence. Randomising the allocation of students to either the treatment (that is, to receive Additional Hours) or control (so they do not receive Additional Hours) group would mean that the two groups are in theory balanced in terms of characteristics that might, alongside the intervention, affect their outcomes. Under these conditions, the outcomes of the control group can be seen to reflect the counterfactual outcomes of the treatment group. Therefore, any differences in the outcomes between the two groups can be considered the causal impact of the Additional Hours policy.

However, due to the universal nature of the policy, use of a randomised controlled trial to evaluate Additional Hours is not feasible.

As such, we explored various quasi-experimental design approaches to evaluate the impact of the policy. These seek to replicate the experimental conditions of a randomised controlled trial by identifying a comparison group of similar students/ providers that did not receive the treatment (in this case because of ineligibility based on location or age) and then comparing their outcomes to those of the treatment group. There are a number of analytical approaches that can be used to construct the comparison group and to compare their outcomes to the treatment group. Following these approaches, the differences in outcomes between the two groups can, under various assumptions, be considered to estimate the causal impact of the Additional Hours policy.

In order to be considered a counterfactual based impact evaluation, these approaches need to achieve the same balance in characteristics between the treatment and comparison groups that an experimental approach achieves through randomisation using an alternative method. These characteristics can be observable, that is to say captured within the data (including for example, measures of socio-economic disadvantage such as Free School Meal eligibility) or unobservable so not captured within the data (for example, the priorities of the provider's senior management team). We can examine the balance in the observable characteristics between the treatment and comparison groups through descriptive analysis and testing for statistically significant differences. However, this is not possible for unobservable characteristics. Therefore, for each quasi-experimental design approach we must consider to what extent there might be differences in unobservable characteristics between the groups, and to what extent balance in observable characteristics might reflect balance in unobservable characteristics.

One potential method to construct a comparison group that is similar in characteristics to the treatment group is propensity score matching (PSM). This approach would 'match' each treated student/ provider to one or more students/ providers from the comparison group that have sufficiently similar characteristics so that their outcomes can be compared to provide an estimate of the causal impact of the policy.

PSM using devolved administrations as a comparison group

One potential choice of a comparison group would be the same cohort as those receiving treatment with Additional Hours in 2022/23 from one of the devolved administrations (Wales, Scotland and Northern Ireland). As education policy is devolved, only students in England are eligible for Additional Hours, and in combination with the relative similarities between the UK nation's education sectors this means that in theory this is the best prospect for a suitable comparison group. However, the main obstacle to this approach is that the relevant datasets covering England and the devolved administrations currently cannot be brought into the same research environment to perform analysis. This is necessary in order to match students between the two groups. The Welsh government is in discussions with ONS about bringing data into the same environment as data covering

England for research purposes in the future, but this is only at a nascent stage and will not happen on a timeline suited to the prospective impact evaluation.

On top of this practical issue, there are also important differences between England and the devolved administrations that would present challenges for an impact evaluation, including: in their education systems (the compulsory age of participation in education is 16 in the devolved administrations compared to 18 in England, meaning there are not the same requirements on the minimum number of hours); the data collected (one of the key Welsh datasets is focussed on sessions attended as opposed to planned hours); and the impact of other factors affecting hours (teacher strikes also took place at various points across the devolved administrations, although at different times and for differing durations, with a lack of a combined, detailed dataset on strike action by provider limiting our ability to account for the differential impact of strikes). Additionally, the devolved administrations implemented their own policy responses to the COVID-19 pandemic in the 16-19 education sector, including a 6.6% increase in funding in Wales to support additional teaching time tailored to the needs of full-time post-16 learners. As such, this affects the extent to which the outcomes of similar students in these settings can be considered the counterfactual outcomes for students in England receiving the Additional Hours treatment.

In summary, this means a counterfactual based impact evaluation approach using PSM and based on a comparison group from a devolved administration is not feasible.

PSM using a past cohort comparison group

Another potential option for a comparison group to be used in a PSM approach would be to use a past cohort of students that was in the English 16-19 education sector before the implementation of Additional Hours in 2022/23. The cohort could be drawn from within the same providers as those receiving treatment in the post-policy cohort, which would eliminate any impact of factors relating to the provider that affect outcomes that do not change over time, including those affecting the likelihood of a provider implementing Additional Hours and the way in which they implement Additional Hours. However, any factors affecting outcomes in a way that varies over time will bias the estimates of the impact of the policy. There are various examples that will have significantly affected outcomes to differing degrees in 2021/22 and 2022/23. These include: how the pandemic has directly affected outcomes, through the dissipation of its direct impact on education over time and given that the different cohorts would have been at different stages in their education during the period of peak disruption by the pandemic; the rollout of the 16-19 Tuition Fund, which includes changes to the eligibility criteria in 2022/23 and increases in funding over time; and the rollout of T Levels, whereby the range of courses offered has shifted significantly over time, and any provider learning regarding implementation issues will result in improved quality of provision over time.

As such, a counterfactual based impact evaluation using a past cohort as the comparison group is not feasible.

Given the various problems with the approaches to construct a robust estimate of impact, as well as potential sources of bias introduced by issues relating to how the policy is implemented and the data available, a counterfactual-based impact evaluation of Additional Hours is not feasible.

Other evaluation approaches that may help identify the effect of Additional Hours

Next, we explore some alternative methodologies using the data sources outlined above, which could be used in a quantitative evaluation of what the potential effect of Additional Hours on the outcomes of interest might have been. This is a step down from inferring a causal impact of the policy.

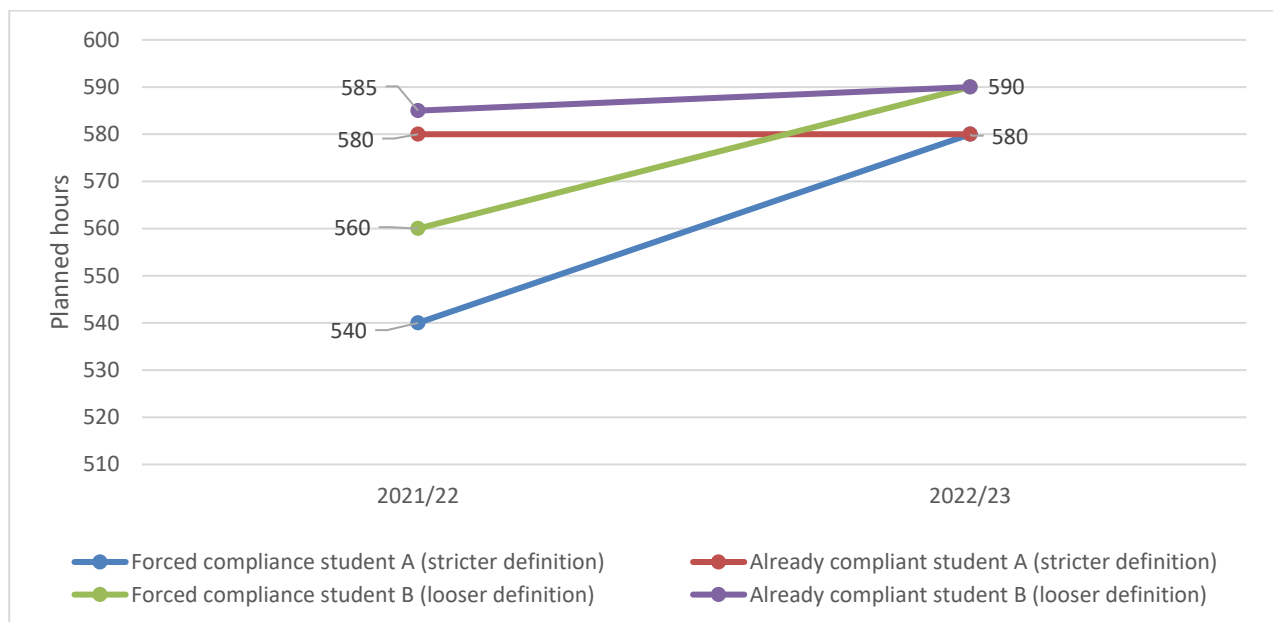
Comparing ‘incentivised compliance’ students to those students ‘already compliant’ with the new minimum requirement

One approach to examining the potential effect of Additional Hours would be to compare the outcomes of students for whom providers were ‘incentivised’ to comply with the first element of the policy (raising of the minimum hours requirement by funding band) against those that were already compliant with it. Any discussion of ‘compliance’ in this section is thus in relation to this element of the policy and not the second element (the expectation to increase average planned hours by 40 from provider baseline for band 5 students and a proportionate amount for those on lower bands) or the policy overall (the first and second elements in combination). By incentivised, we mean it was necessary for providers to increase a student’s planned hours to remain above the minimum requirement, which increased by 40 hours for band 5-9 students (and by a proportionate amount for lower funding bands), in order to maintain the same funding rate for the student. Were the student to fall into a lower funding band if their hours were not increased to above the new minimum requirement, the funding loss to the provider would be significant. There are 12,226 unique students in the NPD that were in 16-19 education in both 2021/22 and 2022/23, not studying T Levels and not high needs, and with planned hours between 540 and 579 in 2021/22 and above 580 in 2022/23 i.e., the pool of eligible students for whom their providers complied with the first element of the policy.

The comparison could be drawn from students who in 2021/22 had planned hours already above the new minimum requirement of 580 for band 5 students in 2022/23 i.e., those students that were ‘already compliant’. The way these groups are defined could be varied to examine how this affects the findings and based on how compliance with the

first element is conceptualised. As an example demonstrated in Figure 2 (focussing on band 5 students for simplicity), a strict approach could compare the change in outcomes of ‘incentivised compliance’ band 5 student A with 540 planned hours in 2021/22 and 580 planned hours in 2022/23 to the change in outcomes of ‘already compliant’ band 5 student A with 580 planned hours in both years, while a looser approach could compare the change in outcomes of ‘incentivised compliance’ band 5 student B whose planned hours went from 560 in 2021/22 to 590 in 2022/23 to the change in outcomes of ‘already compliant’ band 5 student B whose planned hours were 590 in both 2021/22 and 2022/23. Restricting the analysis in this way, eliminates to an extent differences in unobservable characteristics that may relate to the decision to comply with the policy and untreated outcomes.

Figure 2: Examples of potential comparisons between ‘incentivised compliance’ and ‘already compliant’ students



Source: IES, 2023

This approach would compare the change in outcomes between 2021/22 and 2022/23 for the ‘incentivised compliance’ and ‘already compliant’ groups, with the difference in these changes being attributed to the potential effect of the policy. Whilst this panel data approach (by which we mean tracking the same units over time) considers the change in outcomes over time, it does not constitute a Difference-in-Differences approach. This is a form of counterfactual-based impact evaluation approach, which also compares the change in outcomes of a treatment and comparison group over time from before to after the policy was implemented. For this to be considered a counterfactual-based impact evaluation, the parallel trends assumption, the key assumption underpinning the approach, is required to hold. This requires that the outcomes of the treatment and comparison groups are following a similar path prior to the intervention, so that we can

assume that they would have continued on a similar path in the absence of the intervention, and therefore the outcomes of the comparison group can be considered as the counterfactual outcomes of the treatment group. However, we are unable to test this assumption in this context, given that any outcomes from 2020/21 and 2019/20 will be significantly affected by the pandemic, both in terms of the outcomes themselves but also the collection of outcome data (note, for instance, the use of teacher assessed grades in these two academic years).

Additionally, there are 1,297 unique students in the NPD that were in 16-19 education in both 2021/22 and 2022/23, not studying T Levels and not high needs, and with planned hours between 540 and 579 in 2021/22 whose hours were not increased to above 580 in 2022/23 i.e., students whose providers did not respond to the first element of the policy by not having their hours increased to above the new minimum requirement. This means that (in combination with the 12,226 students identified previously) of the total pool of students whose providers should have been incentivised to comply with the first element of the policy, 9.6% 'defied' this. This therefore detracts somewhat from the assertion that providers were 'incentivised' to comply with the first element of the policy and therefore the extent to which this approach avoids the selection issues associated with the policy.

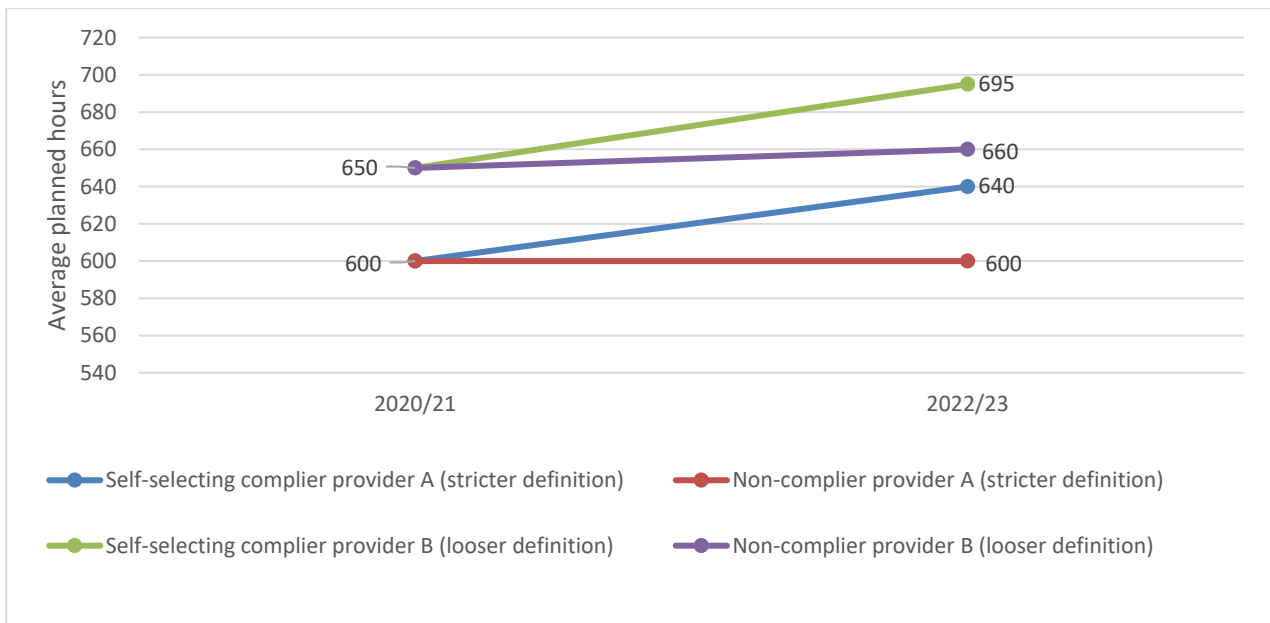
The panel data approach comparing the outcomes of 'incentivised compliance' and 'already compliant' students would also only be able to eliminate the selection issues to the extent to which the 'incentivised compliance' students would have not seen their hours changed in the absence of the increased minimum requirement of hours like the 'already compliant' students and vice versa, while potential differences in unobservable characteristics relating to the difference in the level of planned hours prior to the policy introduce an alternative source of bias. Using the 'already compliant' group to compare the change in outcomes of the 'incentivised compliance' group to would also not allow outcomes at baseline to be used to match students in the two groups, as based on the theory of change (putting aside differences between *planned* and *delivered* hours) we would expect to see differences in outcomes at baseline – the extent to which this gap is closed post-policy is what would be taken as the potential effect of Additional Hours. Matching on baseline outcomes is one way of reducing differences in unobservable characteristics where these are related to outcomes. There may also be a residual effect on outcomes post-intervention of the 'incentivised compliance' group having previously received less hours pre-intervention than the 'already compliant' group, potentially over an extended period. The estimates of the potential effect of Additional Hours produced by this approach would also only be applicable to students with a similar level of planned hours as those used in the analysis. Additionally, as this approach tracks the outcomes of the same students over time, this approach could not be used to analyse outcomes only measured at a single point in time, such as qualification achievement or attainment.

Comparing ‘self-selecting complier’ and ‘non-complier’ providers

Next, we consider two potential approaches to non-causal analysis that would seek to evaluate the second element of the policy i.e., the ‘expected’ 40 hour increase in provider level average band 5 hours compared to baseline. Any discussion of ‘compliance’ in this and the next section is thus in relation to this element of the policy and not the first element or the policy overall. First, we could compare the outcomes of ‘self-selecting compliers’ (those that selected to comply with the second element of the policy) with ‘non-compliers’ (those that appear from the data not to do this). We describe compliance with the second element of the policy as ‘selection’ because it was an *expectation* for providers to increase their average planned hours by 40 from 2020/21 to 2022/23. Regarding the first element of the policy, the increase in the minimum number of planned hours is a *requirement* on providers in order for them to continue receiving the same level of funding per student, hence why we describe compliance with that element of the policy as being ‘incentivised’. To reiterate the points made in the sections ‘Data to support impact evaluation’ and ‘Implementation of additional hours’, defining compliance and non-compliance with the policy based solely on the available data is problematic.

Specifically, ‘self-selecting compliers’ would constitute those providers with average planned hours above 580 hours in the 2020/21 academic year that increased their average hours by 40 to the 2022/23 academic year, whilst ‘non-compliers’ would be institutions with the same level of hours in 2020/21 as the ‘self-selecting’ compliers that did not increase their average hours from baseline to post-policy. Again, the way these groups are defined could be varied to examine how this affects the findings. As an example demonstrated in Figure 3, a strict approach could compare the change in outcomes of ‘self-selecting complier’ provider A with 600 average planned hours in 2020/21 and 640 average planned hours in 2022/23 to the change in outcomes of ‘non-complier’ provider A with 600 planned hours in both years, while a looser approach could compare the change in outcomes of ‘self-selecting complier’ provider B whose average planned hours went from 650 in 2020/21 to 660 in 2022/23 to the change in outcomes of ‘non-complier’ provider B whose planned hours went from 650 in 2020/21 to 660 in 2022/23.

Figure 3: Examples of potential comparisons between ‘self-selecting complier’ and ‘non-complier’ providers



Source: IES, 2023

The key problem with this approach is the issue of selection bias – that differences in characteristics between those providers that ‘select’ to comply with the second element of the policy and those that do not might also relate to differences in the untreated outcomes of these two groups. This would affect the validity of using the outcomes of the ‘non-complier’ providers as the counterfactual outcomes of the ‘self-selecting complier’ providers. Matching on pre-intervention outcomes and other observable characteristics would control for differences in unobservable factors that lead to this selection bias to the extent that the unobservable factors are correlated with observable characteristics. As these factors are unobservable to an impact evaluation, we cannot be certain of the extent to which selection bias has been accounted for. Therefore, this approach cannot be considered a counterfactual-based impact evaluation of Additional Hours.

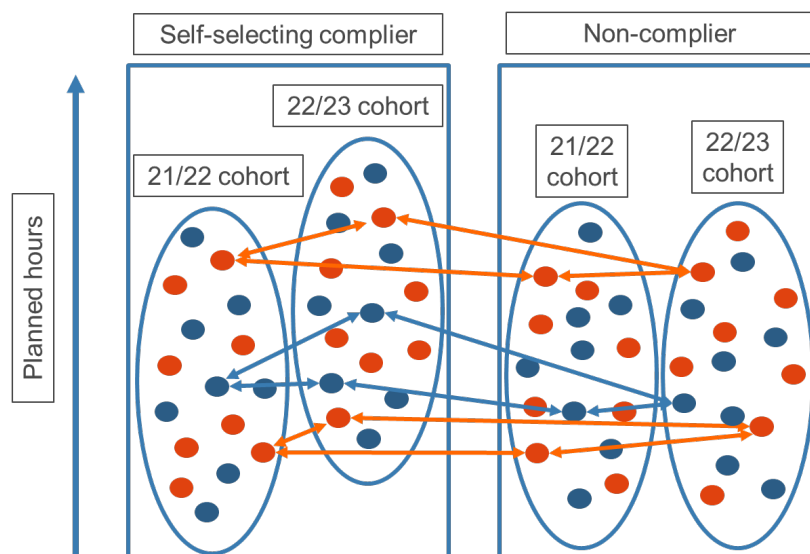
Comparing across cohorts within ‘self-selecting complier’ providers

An alternative approach to assessing the potential effect of Additional Hours on outcomes that could eliminate this selection bias would be to compare the outcomes of students at ‘self-selecting compliers’ (i.e., those providers that chose to comply with the second element of the policy) to a past cohort of students from within the same providers. This would remove any self-selection issues given that the students whose outcomes would be compared would attend the same provider, assuming that the factors affecting these institution’s decisions to select to comply with the second element of the policy remained fixed across cohorts. Once a pool of ‘self-selecting complier’ providers was defined, students within a provider in the 2022/23 cohort would be matched to students within the same provider in the past cohort on student-level observable characteristics. One point to note here is that due to the direct impact of the pandemic on outcomes and

their measurement in 2020/21, whilst it would be the 2020/21 that would be used to define whether a provider was a 'self-selecting complier' or not, it would be the 2021/22 cohort from that provider that would constitute the past cohort used in the analysis. Changes in hours between 2020/21 and 2021/22 would need to be considered to check that the provider's assigned treatment status still held. The difference in outcomes between matched students across cohorts across all providers would then constitute the estimate of the potential effect of Additional Hours.

The main issue with this approach is that it would suffer from bias relating to factors other than the policy affecting outcomes differently over time. This is because by assigning treatment status based on cohort, the estimates of the effect of the policy will also capture the impact of other time effects (including the residual impact of the pandemic, and the rollouts of the 16-19 Tuition Fund and T levels), which are likely to have had a large impact on outcomes. This could to some extent be accounted for by examining the change in outcomes across cohorts within 'non-complier' providers, that is, those providers where average planned hours did not increase. As there was no change in average planned hours for these providers and therefore they did not comply with the second element of Additional Hours, any change in outcomes across cohorts within these 'non-complier' providers will not be due to the second element of the policy. Instead, any changes will reflect the effect of the factors other than the policy that we want to account for in our estimates of the potential effect of Additional Hours. Figure 4 visualises the idea of comparing students (represented by the dots) with similar characteristics (denoted by the colour of the dot) across cohorts and across provider types in order to difference out the potential effect of the policy.

Figure 4: Examples of potential comparisons between students within 'self-selecting complier' and 'non-complier' providers



Source: IES, 2023

This extension to the approach would not fully deal with issues relating to other factors (including the residual impact of the pandemic, and the rollouts of the 16-19 Tuition Fund and T levels) affecting outcomes differently over time, given that these may well also affect outcomes differently by provider type over time. Therefore, this approach cannot generate a counterfactual impact assessment of Additional Hours.

Examining differences in how Additional Hours was implemented and different effects amongst sub-groups

Each of the non-causal, correlational approaches above can be altered to allow for analysis of how the potential effect of Additional Hours on outcomes varies based on how providers used the Additional Hours – such as whether the hours were used for qualification or non-qualification time. Differences in the potential effect based on more specific uses of the Additional Hours could also be performed by linking the provider survey undertaken in the process evaluation to the NPD and ILR datasets, although this would significantly reduce the sample size. This analysis would also be open to another source of bias, given that there might be differences in characteristics between providers that use Additional Hours for different purposes which also relate to the providers untreated outcomes.

Nonetheless, various combinations of these approaches could also be adjusted to allow for some analysis of differences in the potential effect Additional Hours for different sub-groups of interest, including students taking T levels, those with high needs, those from socio-economically disadvantaged background or those that failed GCSE English or maths at the end of Key Stage 4.

Recommendation

This report details the findings from a study of the feasibility of an impact evaluation of Additional Hours. Before examining various specific approaches to assess the policy, we highlight various overarching issues with the data that is available to use for any impact evaluation as well as how the policy was implemented. These issues could introduce sources of bias into any impact evaluation, as well as making classification of whether a student/ provider complied with the two elements of the policy challenging.

Applying any of the counterfactual methods, that show the causal impact of the changed funding and additional hours in education, is problematic. The policy is universal, so a randomised controlled trial is not suitable. While building a comparison group should be possible, constraints including differences between UK nation education systems and the disruption of the pandemic, affect the plausibility of comparisons between nations or of cohorts pre/post policy in England. As such, these approaches that seek to assign causality are not recommended.

Therefore, a number of non-causal approaches were also investigated. Each approach attempts to provide different insights, and each is flawed in varying and differing ways. We would recommend either not pursuing any given that they will not be able to provide robust estimates of the impact of Additional Hours or different approaches to implementing Additional Hours, or pursuing all of them to build a picture of what effect the policy might have had on the outcomes of interest. Specifically, as the results these analyses would produce would not be able to attribute causality, would be challenging to interpret or draw any reliable conclusions from, and therefore have the potential to be misconstrued, **our final recommendation is to not pursue any of the non-causal, correlational approaches outlined.**

If these approaches were to be pursued, the following research questions could be answered:

1. What is the difference in the change in outcomes for band 5 students whose planned hours were below 580 in 2021/22 that were then increased to 580 or marginally above 580 in 2022/23 compared to the change in outcomes for students whose planned hours were 580 or marginally above 580 in both 2021/22 and 2022/23? To some degree, this might reflect the effect of the first element of the Additional Hours policy.
2. What is the difference in outcomes for providers that increased their hours from 2020/21 to 2022/23 compared to providers with similar hours in 2020/21 that did not change to 2022/23? To some degree, this might reflect the effect of the second element of the Additional Hours policy.
3. What is the difference in outcomes for students at providers that increased their hours from 2020/21 to 2022/23 in the 2022/23 cohort compared to students at the same providers in the 2021/22 cohort? To some degree, this might reflect the effect of the second element of the Additional Hours policy.
4. How do the estimates produced to answer research questions 1, 2 and 3, which to some degree might reflect the effect of the Additional Hours policy, vary depending on how the level of qualification and non-qualification hours changed? To some degree this might reflect the differences in effect of using the additional hours for qualification versus non-qualification time.
5. How do the estimates produced to answer research questions 1, 2 and 3, which to some degree might reflect the effect of the Additional Hours policy vary depending on which specific uses the respondents to the provider survey indicated they used additional hours provision for? To some degree, this might reflect the differences in effect of using the additional hours for various specific uses.
6. How do the estimates produced to answer research questions 1, 2 and 3 (or variations of these given the intended implementation of Additional Hours for T level

and high needs students), which to some degree might reflect the effect of the Additional Hours policy vary depending on whether students are taking T levels, have high needs, are from socio-economically disadvantaged backgrounds, or failed GCSE English or maths at the end of Key Stage 4? To some degree, this might reflect differences in effect of the Additional Hours policy for different sub-groups.

References

This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.



Department
for Education

© Department for Education copyright 2024

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3.

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned.

Reference: RR1457

ISBN: 978-1-83870-573-2

For any enquiries regarding this publication, contact us at:
www.education.gov.uk/contactus

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