

Early impacts of work experience

April 2012

DWP Department for
Work and Pensions

Acknowledgements

Special thanks go to Helen Bewley, Richard Dorsett, Jonathan Portes and David Wilkinson at the National Institute of Economic and Social Research (NIESR) for their brief peer review of the analysis. Although this was early analysis they were able to provide some very insightful comments on the drafting, methodology and recommendations for future work to develop and extend the analysis.

We are also grateful to Mike Daly, James Whitley and Mandy Daly at the Department for Work and Pensions who took the time to read through drafts of this report.

Contents

1	INTRODUCTION	4
1.1	Rationale for the present analysis	5
1.2	Policy Background and Design.....	5
1.3	Participation on work experience	8
2	DATA AND SAMPLE DEFINITION	10
2.1	Sample definitions	10
2.1.1	Defining the work experience participant sample	10
2.1.2	Defining the non-participant sample.....	11
2.2	Data sources and variables	11
2.2.1	Description of Variables	12
2.2.2	Data quality issues	14
2.3	Comparing participants and (unmatched) non-participants.....	15
3	METHODOLOGY	19
3.1	Conditional Independence Assumption.....	19
3.2.1	Common Support for Participants	23
3.2.2	Matching Quality	24
3.3	Measuring Outcomes and Impacts	26
4.	RESULTS	28
4.1	Impacts of work experience.....	28
4.2	Sensitivity of Impact Estimates	32
5.	CONCLUSIONS	34
	APPENDICES.....	36
	Appendix 1 – Defining the Participant Sample.....	36
	Appendix 2 – Matching on other DWP Programmes.....	37
	Appendix 3 - Controlling for Labour Market History	38
	Appendix 4 – Matching Protocol.....	40
	Appendix 5 – Sensitivity Analysis	41

REFERENCES44

List of Figures

Figure 1.1: Monthly and cumulative starts on WE..... 8
Figure 1.2: WE monthly starts by age group..... 9
Figure 2.1: Benefit receipt rate among participants and non-participants..... 17
Figure 3.1: Propensity Score distributions of participants and non-participants:..... 23
Figure 4.1: Benefit receipt rate for WE participants and matched non-participants 29
Figure 4.2: Impact of WE on the likelihood of a participant claiming benefit..... 29
Figure 4.3: Employment rate for WE participants and matched non-participants..... 29
Figure 4.4: Impact of WE on the likelihood of a participant in employment..... 29

List of Tables

Table 1.1: WE starts by region 9
Table 2.1: Variables and values used in the analysis..... 13
Table 2.2: Characteristics of the participant and (unmatched) non-participant samples..... 16
Table 3.1: Specification statistics for the group matching 24
Table 3.2: Unmatched and matched means for main analysis..... 25
Table 4.1: Impacts (percentage points) on mutually exclusive outcomes at 21 weeks 31

1 Introduction

Work experience (WE) was launched in January 2011 as one of the Government's Get Britain Working Measures¹. The programme is designed to help young people on Jobseeker's Allowance (JSA) gain an insight into the world of work and provide them with the practical experience they need to secure a job before they become eligible for the Work Programme. Jobcentre Plus works with employers to offer young jobseekers the opportunity to overcome barriers to employment through offering them a WE placement usually lasting two to eight weeks, which can be extended for another four weeks. WE opportunities are brokered by Jobcentre Plus, either through a national agreement via the National Employer Service Team or through local agreements within the Jobcentre Plus District.

The decision to participate on WE is voluntary. Young people undertaking a WE placement continue to receive their benefit and, therefore, are required to continue to sign for their benefit each fortnight and be available for and actively seek employment during the period of their participation. Host employers are expected to provide time for job search and to release participants to attend interviews with employers or at the Jobcentre Plus office.

Get Britain Working official statistics² indicated that, from January 2011 up to and including November 2011, there had been 34,200 starts onto WE placements.

This report describes a quantitative analysis of WE, providing estimates of:

1. the net impact of WE on the likelihood of young participants receiving an out of work benefit³ or training allowance⁴ during the first 21 weeks following the start of their placement;
2. the net impact of WE on the likelihood of young participants being in employment during the first 21 weeks following the start of their placement; and

The main analysis focuses on the first cohort of young participants (aged 19-24) who started their WE placement between January 2011 and May 2011. The cohort was chosen to give a reasonable tracking time allowed by the data for a sizeable number of participants.

¹ Other Get Britain Working Measures are Work Clubs, Work Together, Enterprise Clubs, New Enterprise Allowance and sector-based work academies.

² http://statistics.dwp.gov.uk/asd/asd1/pwp/pwp_gbw_feb12.pdf

³ Jobseeker's Allowance, Income Support, Incapacity Benefit or Employment and Support Allowance

⁴ A training allowance is a maintenance allowance paid instead of Jobseeker's Allowance, out of public funds to people taking part in a course of training or instruction provided by the Department or under arrangements made with a partner.

The remainder of this chapter provides context for the study. Section 1.1 describes the rationale for the present analysis. Section 1.2 introduces the policy background and the overall design of the programme. Section 1.3 describes the overall participation on the programme.

1.1 Rationale for the present analysis

The challenge faced when evaluating employment programmes such as WE is that we can never be certain what would have happened to participants if they had *not* taken part in the programme, particularly when the decision to participate is voluntary. In November 2011, the Department published an ad hoc analysis⁵ that indicated for the first three months of 1,300 starts, 49% of participants were in receipt of an out of work benefit at 13 weeks following the start of their placement. However, unless we know what would have happened to these participants if they had not started their WE placement, we can not say whether the programme actually made any difference to their labour market prospects.

Therefore, the aim of the impact analysis is to compare the observed labour market outcomes of WE participants with an estimate of their 'counterfactual' outcomes (the labour market outcomes which would have occurred in an 'alternative world' if they had not started their WE placements). By comparing the benefit and employment rates of a group of WE participants with a carefully constructed group of non-participants who can best represent the counterfactual case, the net impacts of WE on the benefit and employment rates of participants can be estimated.

Our methodology draws heavily and builds on the approach used by Ainsworth and Marlow (2011). We use Propensity Score Matching (PSM) to construct a suitable comparison group of non-participants who most closely resemble the group of participants under analysis. We then use the labour market outcomes of the matched non-participant group as an estimate of the counterfactual outcomes of the WE participants and compare these with the observed participant outcomes. We use a rich data set comprising the individual characteristics of participants and non-participants to carefully construct the matched comparison group of non-participants.

1.2 Policy Background and Design

WE is part of the Government's package of Get Britain Working Measures that can be used by Jobcentre Plus to help individuals into work. The other Get Britain Working Measures are Work Clubs, Work Together, Enterprise Clubs, New Enterprise Allowance and sector-based work academies.

5

http://statistics.dwp.gov.uk/asd/asd1/adhoc_analysis/2011/work_experience_participant_outcomes.pdf

Early Impacts of Work Experience

WE placements can be flexible, lasting two to eight weeks and with attendance 25 to 30 hours per week (unless there are agreed restrictions on the Jobseeker's Agreement in which case a lower number of hours is allowed), with the aim of:

- maximising the number of young people moving into employment or training;
- providing young unemployed people with high quality voluntary work experience; and
- minimising the number of young people remaining on benefit long enough to be referred onto the Work Programme.

The WE model is targeted at young Jobseeker's Allowance (JSA) claimants, with little or no work history and low skills. The programme provides the opportunity to gain an insight into the world of work and to develop skills and experience with local employers in a 'live' environment.

Jobcentre Plus District Managers have overall ownership and accountability for the delivery of WE. The district, usually through a co-ordinator team, works with employers and advisers to promote, broker and co-ordinate WE placements. However, large employers can broker agreements at the national level through the National Employer Service Team. Host employers are required to sign an agreement that sets out clearly what is expected of all parties. This helps advisers to better describe individual opportunities and helps claimants to understand what they can expect from the placement.

A recommendation for a referral to WE placement by the adviser is discretionary but the guidance on eligibility is for JSA claimants aged 18 to 24 from week 13 of their claim up to referral to the Work Programme and that participants must be motivated and demonstrate a willingness to work. Where an adviser believes WE will benefit a claimant, there is discretion to recommend claimants earlier than week 13 and, exceptionally, those aged 25+ who have no recent work history. JSA claimants aged 16/17 are eligible for WE placements from day 1 of their claim; however, they can only be referred to host employers that have signed up to additional employment rules for this age group.

The decision as to whether to accept a place on the WE programme has always been entirely voluntary. However, prior to 29 February 2012, except for 16/17 year olds, participation became mandatory once the claimant made a decision to participate and was formally referred. Once referred, claimants were required to attend on day 1 of the placement; they could then choose to leave the placement during the first week without penalty if they decided it was not for them. However, after the end of the first week, participation became mandatory. Sanctions could be applied where the claimant could not show good cause for failing to attend or giving up a place on the programme or where a participant was dismissed by the host employer for misconduct.

Early Impacts of Work Experience

Up to and including October 2011 there had only been a very small number of sanctions⁶ (220) applied to WE participants.

From 29 February 2012, following a consultation with employers, the Department agreed to remove all sanctions in relation to WE participation with the exception of cases involving 'gross misconduct'.

Young people undertaking a WE placement remain on benefit and, therefore, are required to continue to sign for their benefit each fortnight and be available for and actively seek employment during the period of their participation. Host employers are expected to provide time for job search and to release participants to attend interviews with employers or at the Jobcentre. There is also the opportunity for some participants to have their placement extended by up to four weeks, where as a result of the participation the host employer makes an offer to take the participant onto an Apprenticeship and that offer is accepted by the participant. The extension period is to allow the host employer time to make the necessary arrangements for the apprenticeship.

During the last couple of decades there have been other employment programmes that contain an element of 'work experience', such as options within the New Deal for Young People (NDYP), more recently Work Trials (WT) and the current sector-based work academies (sbwa), where potential placements are sourced by providers contracted by DWP. However, since there are number of differences, there should be caution in making comparisons.

A key feature of the WE programme is that Jobcentre Plus works directly with employers to secure placements, both locally and at national level. Also unlike NDYP and sbwa, which contain various options and 'modules' that need to be completed, and which are mandatory, work experience is a 'stand alone' voluntary programme providing the opportunity for young unemployed people to develop much needed work related skills and disciplines in a 'real life' working environment.

The evaluation of NDYP (Beale et al. 2008) showed that NDYP participants spent on average 64 fewer days claiming JSA and training allowance than a group of JSA claimants slightly too old to be eligible for NDYP over a four year period. The estimated impact of NDYP gradually fell over the four year period from 24 days in the first year to 10 days in the fourth year. However, given the differences to WE there should be a degree of caution in making comparisons of impacts, least of all due to the short run of data available for our present analysis.

Work Trials (WT) are also different in that there must be a vacancy for which the claimant is being considered. A WT is available where there are doubts about suitability for both or either the employer and claimant and is not about gaining work related experience. There can only be one WT per vacancy (i.e.

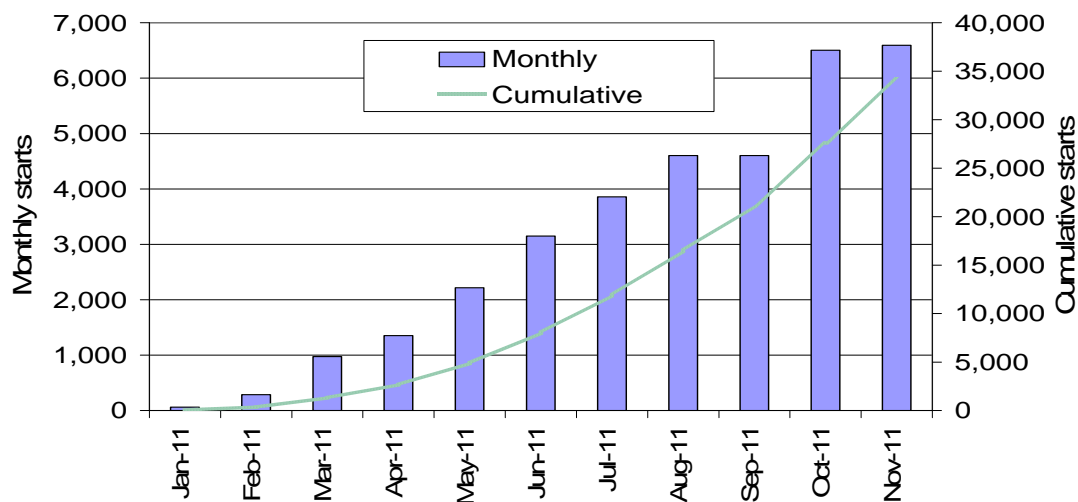
⁶ <http://83.244.183.180/sanction/sanction/LIVE/tabtool.html>

they are non-competitive), whereas with WE there are no vacancies and the employer can host as many placements as they want, safeguards exist to reduce the risk of displacement and to ensure employers do not use the scheme as a means of staffing their business. To date, there hasn't been a quantitative evaluation of WT.

1.3 Participation on work experience

In the period from January 2011 up to and including November 2011, there were 34,200 starts⁷ on WE placements. Figure 1.1 shows that participation began quite slowly with an inflow of 60 participants in January 2011 increasing to 6,500 by October 2011.

Figure 1.1: Monthly and cumulative starts on WE

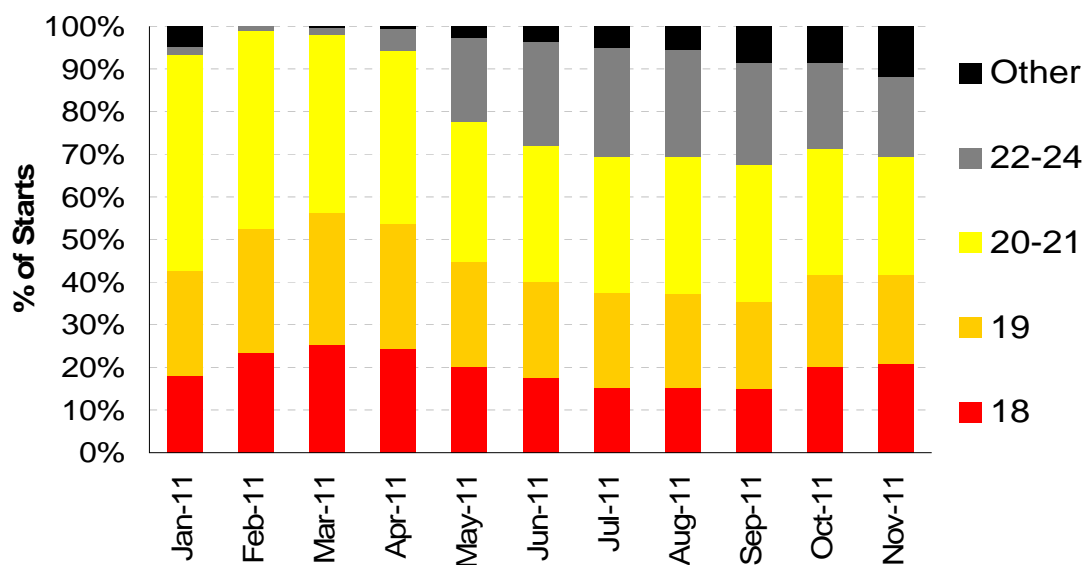


Source: Pre-Work Programme and Get Britain Working Official Statistics, February 2012

For the whole period, 92% of WE starts, at the time of starting on the programme, were aged 18-24; a sizeable majority (72%) were aged 18-21 and 40% were aged 18-19. Figure 1.2 shows that over time there was significant variation in the age distribution. In particular there was a very high proportion (88%) of 18- to 21-year-olds in the first five months of the programme compared to the proportion (69%) from June onwards. This needs to be considered for the present analysis which is focused on the cohort of starts from January to May. Therefore, on this basis alone, care needs to be taken on extrapolating findings for later cohorts.

⁷ As defined in the official statistics, a work experience start is defined as the date the claimant accepted their work experience placement with the employer. This is inputted by an adviser to Jobcentre Plus' Labour Market System.

Figure 1.2: WE monthly starts by age group



Starts to WE placements were available in all regions across Great Britain. Table 1.1 shows the spread of starts across the regions in comparison to a proxy to the eligible population: the proportion of 18- to 24-year-old JSA claimants who passed through 3 months duration (3 month threshold flows) during January-November 2011. The data shows that there have been more starts on WE placements in the South than the North with respect to the numbers of eligible people passing through 3 months.

Table 1.1: WE starts by region

Region	WE Starts (Jan-Nov 2011)	3 month JSA 18- to 24-year-old threshold flows (Jan – Nov 2011)
Central England	27%	21%
London and the Home Counties	22%	19%
North East	14%	17%
North West	8%	13%
Scotland	7%	10%
Southern England	17%	12%
Wales	4%	6%
Unknown	1%	3%
Great Britain	34,200	428,000

Source: Pre-Work Programme and Get Britain Working Official Statistics Official Statistics, February 2012 and National Benefits Database

2 Data and Sample Definition

This section outlines the data and sample definition used in the evaluation. Section 2.1 describes the method of drawing samples from which we construct groups of participants and non-participants for comparison. Section 2.2 describes the variables used in the evaluation. Finally, Section 2.3 compares the participant and non-participant samples.

2.1 *Sample definitions*

This section describes the participant and non-participant samples selected for the main impact analysis (Section 4.2 describes a number of sensitivity tests which use alternative participant and non-participant samples).

2.1.1 **Defining the work experience participant sample**

The aim of this paper is to estimate the impacts of WE on the 18- to 24-year-olds who start on the programme. A WE start is defined as the date the claimant accepted their placement with the employer; this is inputted by an adviser to Jobcentre Plus' Labour Market System.

The main analysis was performed on a participant sample of 19- to 24-year-olds who started a WE placement between January 2011 and May 2011 and were receiving JSA in the week their placement started. The cohort was chosen to give a reasonable tracking time allowed by the data for a sizeable number of participants.

18 year olds were not included in the main analysis because the analysis described later controls for work and benefit history over the previous year, which by definition will be only partial for this age group, and we do not have other data which could serve the same purpose, such as educational achievements. However, one of several sensitivity tests described in Appendix 5 includes all 18- to 24-year-old participants in the analysis.

All participants meeting the following criteria were included:

- the participant must be claiming Jobseeker's Allowance (JSA) one week before starting their WE placement;
- the participant must be aged between 19 and 24 at the start of their WE placement; and
- the WE start date must be between January 2011 and May 2011;⁸

⁸ This provided a cohort of participants for whom we had a minimum of 21 weeks of outcome data.

The resulting participant sample size was 3,490. Appendix 1 shows in more detail how this sample of 3,490 participants was selected from the total of 4,890 who started WE during the cohort period.

2.1.2 Defining the non-participant sample

In defining the non-participant sample, the aim is to select non-participants who can best represent what would have happened to WE participants if they had not participated on WE.

To compare the outcomes of participants and non-participants over a time period such that non-participants can represent what would have happened to WE participants if they had not participated, pseudo start dates⁹ were assigned to each non-participant. The pseudo start dates for non-participants were subsequently treated as equivalent to the actual start date for participants.

The non-participant sample was drawn from DWP administrative data sets. All non-participants meeting the following conditions were included:

- The non-participant must be claiming JSA in the week of their assigned pseudo start date;
- The non-participant must be aged between 19 and 24;
- The pseudo start date must be between January 2011 and May 2011;

Before including the last condition for pseudo starts the number of non-participants was 637,690. After applying the condition on pseudo starts the resulting sample size was 378,210 non-participants.

Section 3 describes how suitable 'matched' groups of participants and non-participants were selected from these samples and compared to estimate the impacts of WE on employment and benefit receipt. This selection was carried out using Propensity Score Matching.

2.2 Data sources and variables

The evaluation was carried out using administrative data derived from two main sources:

- 1) *DWP administrative databases*, which provide details of spells on DWP benefits, characteristics of DWP customers (drawn from the Jobcentre

⁹ The pseudo starts were generated using the same methodology employed by Ainsworth and Marlow (2011) in their assessment of the European Social Fund. The method aligns the non-participants and participants to two time dimensions: calendar time and length of time on benefit so that the distribution of monthly WE starts mirrors the distribution of monthly pseudo starts.

Plus Labour Market System which relies on inputs from advisers) and spells on employment programmes, including WE; and

- 2) *Her Majesty's Revenue and Customs (HMRC) Tax System*, which provides details of spells in employment.

Given the age group concerned it would have also been highly desirable to draw on data on educational attainment. However data on qualifications is not routinely collected by Jobcentre Plus advisors.

It is widely recognised that there are both advantages and disadvantages to using administrative data compared with, for example, survey data. We outline below some of the broad differences between these two methods:

- administrative data allows for a much larger sample size (close to the population) than survey data;
- survey data tends to suffer from non-response;
- administrative data can also suffer from omissions and errors – notably, there are substantial flaws in the HMRC employment data, as set out in section 2.2.2 below;
- administrative data allows variables and outcomes to be tracked over a longer period than survey data, which generally offers only a snapshot in time; however
- administrative data is limited to a pre-defined set of variables, while survey data can provide a richer data set tailored to a specific research question.

While survey data could provide additional variables with which to control for participant characteristics (as found by, for example, Dolton and Smith, 2011), the present study uses purely administrative data for the following reasons:

- the larger sample size allows us to explore the sensitivity and heterogeneity of the estimated impacts with regard to using different participant and non-participant groups (see Section 4.2); and
- the costs and time involved in undertaking fieldwork to collect survey data are high. Administrative data is readily available on DWP systems.

2.2.1 Description of Variables

Table 2.1 outlines the variables used in the analysis. The importance of these variables in controlling for selection onto the WE programme is described in Section 3; Appendix 3 outlines the method used for constructing the benefit and employment history variables and explains the advantage of the approach over alternative methods.

Table 2.1: Variables and values used in the analysis

Variable	Type	Values
Gender	Categorical	<i>Male; Female</i>
Age	Numerical	<i>19-24 years old</i>
Disability ¹⁰	Categorical	<i>Not disabled; Disabled; Unknown</i>
Ethnicity	Categorical	<i>White; Black; Asian; Mixed; Chinese; Other; Unknown</i>
Sought occupation	Categorical	<i>26 broad categories: e.g. "Administrative"; "Health Professionals"; "Sales Occupations"</i>
Lone Parent ¹¹	Categorical	<i>Lone Parent; Not a Lone Parent (at any time within 2 years prior to start date/pseudo start date)</i>
Jobcentre Plus District	Categorical	<i>48 Jobcentre Plus districts in Great Britain and Unknown</i>
Low Qualified	Categorical	<i>No; Yes; Unknown</i>
Local Authority labour market characteristics: <ul style="list-style-type: none"> - Employment rate; - Unemployment rate; - Economic inactivity rate; - Average pay; - Job density; - Vacancy density; 	Numerical	Employment, unemployment and economic inactivity rate can range between 0 and 1. Average pay, job density and vacancy density can take any positive value.
Benefit ¹² history	Categorical	<i>52 binary variables – representing each of the 52 weeks prior to WE start date/pseudo start date. Values are: in receipt of benefit; not in receipt of benefit</i>
Employment history	Categorical	<i>52 binary variables – representing each of the 52 weeks prior to WE start date/pseudo start date. Values are: in work; not in work</i>
WE start/pseudo start month	Categorical	<i>Months from January 2011 to May 2011 are given distinct values</i>

¹⁰ Since disability is set by a Jobcentre Plus adviser based on claimant self -disclosure, this variable is not a systematic identification of disability as defined by the Disability Discrimination Act (DDA).

¹¹ 'Lone parent' defined by marital status information to infer partner status and HMRC Child Benefit data to infer parental status

¹² 'Benefit' is defined as any of four out of work benefits (Jobseeker's Allowance, Incapacity Benefit, Employment and Support Allowance, Income Support) or training allowance. Other benefits are not included in the benefit history variables or outcomes.

Early Impacts of Work Experience

Benefit start month ¹³	Categorical	All months up to May 2011 are given distinct values for the benefit spell prior to WE start /pseudo start
Other programme participation. (See Appendix 2 for a full list of programmes included.)	Numerical	Number of days spent on each DWP programme in two years prior to WE start date: <i>Integers from 0 to 730</i>

2.2.2 Data quality issues

Employment data

The employment data used in this analysis comes from Her Majesty's Revenue and Customs (HMRC) data systems. In most circumstances, employers are obliged to notify HMRC when an employee starts or ends a spell of employment. Employment history and outcomes of individuals were derived using the recorded start and end dates of these notified employment spells. However, there are a number of documented issues with the quality of this data.¹⁴ These are briefly described below.

1. Employment spells are only recorded when a tax form is submitted. Some employment spells, such as those corresponding to self employment and individuals not earning higher than the National Insurance contributions threshold, are therefore not recorded;
2. If HMRC do not know the date on which an employment spell started, they assign a start date of the 6th April in the year that they become aware of the employment spell. This may not be the actual year in which the spell began. A similar process occurs when HMRC do not know the date on which an employment spell ended. In this case they assign an end date of the 5th April; and
3. A small number of records contain other known errors, such as missing start dates or missing end dates.

As in other evaluation studies, such as Beale et al. (2008) we have followed advice to mitigate the problem of all dates with errors, by randomly assigning start and end dates within the assigned tax year for records in which they are unknown.

We acknowledge that our estimates of the impact of WE on employment rely on imperfect data. However, they still provide an important estimate of how

¹³ Benefit start and end dates refer to the benefit spell leading up to the start of the WE placement.

¹⁴ <http://research.dwp.gov.uk/asd/asd5/rports2007-2008/rrep432.pdf>

effectively WE impacts on the future employment prospects of participants, if we assume that there is no systematic bias between recording of employment for participants and non-participants. Of course this assumption may not hold if, for example, employers who take on individuals who were on WE placements are more likely than employers on average to record employment starts. This could lead to the employment effect being overestimated due to instances of employment being more frequently recorded for those individuals in the treatment group than for those in the comparison group.

DWP administrative data

On the whole, we believe that the recording of out of work benefit and training allowance spells to be accurate and therefore the estimates for the impact of WE on benefit receipt to be more reliable than employment impacts. However, characteristics data from the Jobcentre Plus Labour Market System contains a number of missing values because advisers do not routinely fill in all of the fields during client interviews or client does not disclose the information. This is particularly the case for variables identifying ethnicity, disability and low qualified. In the case of variables with missing values, 'unknown' is treated as a valid category for controlling for participant characteristics.

There are no statistically significant differences in the proportions of missing or unknown values for the participant and non-participant samples, with *ethnicity* reporting around 5% missing in both groups, *low qualified* reporting 74% and 75% respectively, *occupational choice* reporting levels of 1% (participants) and 2% (non-participants) and less than 1% of missing or unknown values reported for all other variables.

2.3 Comparing participants and (unmatched) non-participants

This section compares the basic characteristics of individuals in our main WE participant and (unmatched) non-participant samples.

Table 2.2 lists summary statistics detailing personal and demographic characteristics, benefit receipt, and participation on DWP employment programmes other than WE. The table includes only a number of summary characteristics – for a full list of variables included in the analysis, refer back to Table 2.1.

Table 2.2: Characteristics of the participant and (unmatched) non-participant samples

	WE Participants	Non-Participants
Observations	3,489	378,207
Personal/ Demographic Characteristics		
Age (mean years)	20	21
Male(%)	63	69
Disabled(%)	13	14
Ethnic Minority (%)	15	16
Low Qualified (%)	25	22
Lone Parent (%)	0	1
Average Local Authority Unemployment Level (%)	8	9
Benefit Receipt		
Weeks of past year spent receiving JSA (mean weeks)	29	27
Weeks of past year spent in employment (mean weeks)	11	17
Benefit duration at start of programme (mean weeks)	21	20
Benefit duration more than six months (%)	28	25
DWP Programme Participation		
Weeks of the past two years spent on other DWP programmes (mean weeks)*	9	16

* Based on the cumulative number of days spent on programmes. Therefore, an individual being on more than one programme at the same time could cause this figure to exceed 104 weeks

In terms of demographics, participants are quite similar to non-participants apart from fewer participants (63%) than non-participants (69%) being male. Also more participants (25%) than non-participants (22%) are recorded as having low qualifications (although the majority are recorded as unknown).

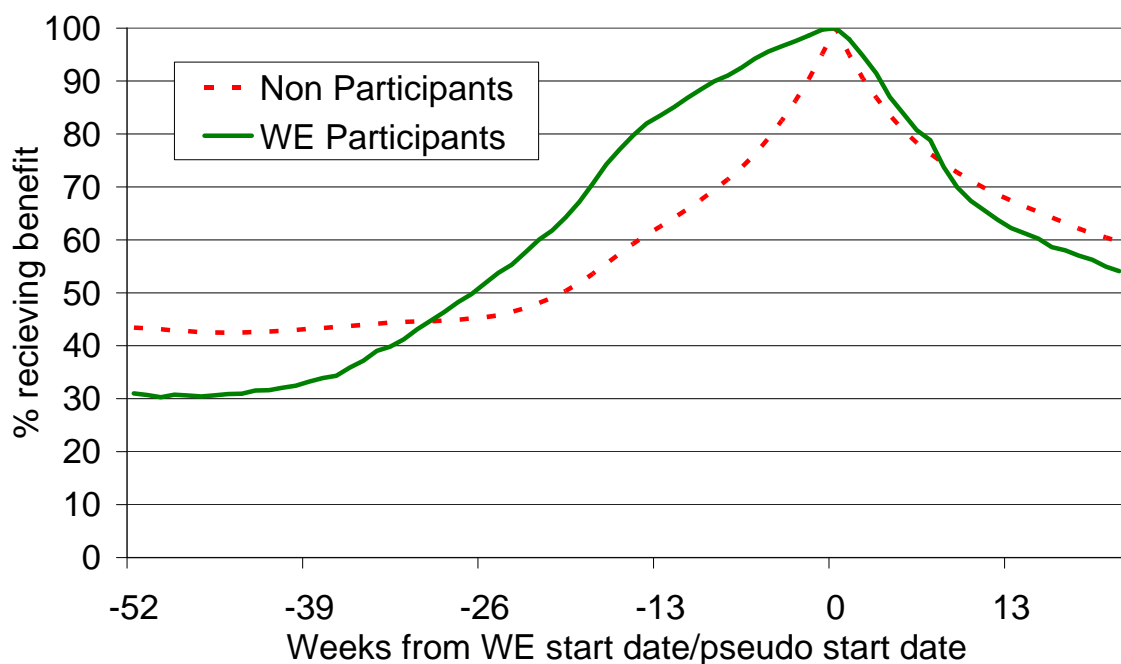
Participants tend to have spent more of the previous year before starting WE on JSA (29 weeks) than non-participants (27 weeks). However, participants spent less of the past two years on other DWP programmes (9 weeks) than non-participants (16 weeks).

Overall, these statistics suggest that participants are more likely to be low qualified, to have spent less time in employment and more likely to have spent more than six months on benefit before participating than non-participants. However, they have spent less time in the last year on other DWP

programmes before participating than non-participants, possibly because they are younger and have had less time to participate on DWP programmes.

Figure 2.1 compares the benefit¹⁵ receipt rates of participants and non-participants at each week before and after the WE start/pseudo start date. The benefit rates shown are for the complete sample in each case, i.e. before any attempt is made to select groups of participants and non-participants with similar characteristics. Differences between participants and non-participants are therefore not attributable to impacts of WE, but rather a combination of WE impacts and differences in the characteristics of those who participate and those who do not.

Figure 2.1: Benefit receipt rate among participants and non-participants



52 weeks before the WE start/pseudo start date, the proportion on benefit was lower among those who participated (31%) compared with those who did not (43%). This might be, in part, down to the age of WE participants, as the vast majority are from the lower end of the age range of eligibility; in some cases individuals might have been unable to claim working age benefits in the period prior to a WE start as they fell below the age limit or were still in education.

Towards 30 weeks before starting, the proportion of participants on benefit increased dramatically so that it exceeded that of those who did not participate. By 13 weeks the proportion of participants on benefit was much higher (84%) than that of non-participants (63%).

¹⁵ 'Benefit' is defined as any of four out of work benefits (JSA, IB, ESA, IS) or training allowance.

Early Impacts of Work Experience

Overall, prior to participation, these observations suggest that WE participants tended to be slightly further away from the labour market than those who did not participate.

Over the first 8 weeks following the start/pseudo start date, the proportion of participants receiving benefits was higher than the proportion of non-participants. Between 8 and 16 weeks after the start of the WE placement, the proportion of participants in receipt of benefit decreased from 73% to 59%. This is different to the non-participant group where the proportion on benefit receipt did not fall as much; it decreased from 75% to 65% over the same period. By 21 weeks after the start of the WE placement, the proportion on benefit was lower among participants (54%) than non-participants (60%).

The next section describes the methodology used to select the comparison group of non-participants who are similar to WE participants with regard to demographic characteristics, labour market history and previous programme participation. This enables us to attempt to isolate the effect of the WE placement on the labour market prospects of participants.

3 Methodology

This section outlines the methodology used to estimate the impacts of WE on the labour market prospects of participants. The impact estimates are specifically for the *average effect of treatment on the treated, ATT*.

Section 3.1 describes the Conditional Independence Assumption, which forms the foundation of impact evaluations of this type. Section 3.2 describes the Propensity Score Matching (PSM) methodology used to control for selection bias and construct a suitable counterfactual for our WE participant group. Finally, Section 3.3 describes the method used to measure the labour market outcomes of participants and non-participants in our samples, and of using these outcomes to estimate the net impacts of the WE programme.

3.1 Conditional Independence Assumption

The aim is to estimate the average effect of WE on those who participate. We can not use a simple comparison between the benefit and employment outcomes of participants with those of non-participants as this could be biased if there are systematic differences between these groups which may be related to the labour market outcomes of interest (see Table 2.2). However, if we can control for all characteristics that influence selection onto the programme, then the outcome that would result in the absence of treatment is the same in both cases (Bryson, et al. 2002). Our identifying assumption is that conditional on the variables we have observed (as defined in Table 2.1), the counterfactual outcome is independent of participation. This is known as the 'Conditional Independence Assumption' (CIA). It enables us to infer the counterfactual outcomes for WE participants, and therefore to attribute any differences between carefully matched participant and non-participant groups to the effect of WE. We control for characteristics using a Propensity Score Matching methodology, as described in Section 3.2.

The remainder of this section discusses how the individual level data (as described in Chapter 2) allows us to try to control for the difference in typical characteristics between those who start a WE placement and those who do not. Differences in characteristics between these two groups arise as a result of the way participants are selected (voluntarily referred to and accepted) onto placements, and are therefore known as *selection bias*.

Controlling for selection bias

To understand the differences between WE participants and non-participants we must consider the mechanism by which jobseekers are referred to and accepted onto WE placements.

The policy design describing the process from referral to acceptance was described in Section 1.2. For the present discussion, selection onto the programme appears to be determined by three interconnected factors:

1. suggestion by Jobcentre Plus advisors, guided by eligibility criteria, to participate in WE;
2. voluntary decision (i.e. no obligation to accept) by the individual to be referred onto the programme and decision to start on the programme; and
3. employers may also be involved in the participant selection process if they wish. This could include, for example, an informal chat to ensure they select people who they feel are suitable for the placement.

At this early stage of the programme there is little evidence as to how these factors operate and which factors are more significant in driving the selection. Therefore the following is a conjecture of potential factors, based on the policy design, behind how people start on the programme.

In deciding whether to refer a jobseeker to WE, a Jobcentre Plus advisor might consider the extent to which a young person needs additional help to gain employment and the extent to which WE is likely to meet these needs amongst other options for support. They will make a judgement on whether the individual appears motivated and willing to work. This judgement will be influenced by the guidance on eligibility criteria, the individual characteristics of the participant, the type of job sought, the number and type of WE placements available in the local area, the suitability of the participant for specific posts, the perceived motivation of the participant, the familiarity of the particular Jobcentre Plus advisor with WE, and the availability of other employment support options.

It is important to note that whilst the adviser might have decided that a place on the programme would be suited to a claimant, it was always the claimant's decision whether or not to accept the offer of a place. The programme only became mandatory once the offer was accepted by the claimant and the referral made.

We know, from an early cohort of referrals, that about a third of referrals started a placement; this gives a sense of the scale of the selection by the individual and employer after being referred by the adviser. There are probably a number of reasons for this drop out including a reluctance of claimants to tell advisers who recommend a WE placement that they are not

interested but also because people leave benefits, possibly because in some cases WE acted as a deterrent.

There will be some similar factors which influence a potential participant's decision to be referred and start on a placement to those which influence the advisor – i.e. whether they feel that WE will improve their labour market prospects and the extent to which they actually want to improve their labour market prospects. However, there will be some other factors driving the participant's decision, which include the attractiveness of the placement on offer and a claimant's drive to undertake a short period of unpaid WE. All these factors will be driven by individual characteristics, motivation and circumstance. Since WE is one of a range of support options available to potential participants, the final decision to apply for a WE placement will be a mutual decision reached after a discussion between the participant and the advisor.

In cases where employers choose to be involved in the selection process, they will probably choose participants who the employer believes will benefit most from the experience and will be of most benefit to the employer organisation. In either case, they are likely to choose claimants who appear most motivated and keen to take up a WE placement.

To allow us to control for the selection bias which results from this process of referral and recruitment, we have brought together demographic characteristics such as age, gender, ethnic group, disability, qualification and lone parent status (Section 2.3) for the participant and non-participant samples. We have also obtained each individual's stated preferred choice of occupation. Since labour market prospects may be highly dependent on dynamic local labour market characteristics and the local availability of employment support, we have also collected each individual's geographical district and the labour market characteristics of the Local Authority where each individual lives.

While we have not been able to directly observe each individual's level of motivation, we have collected a number of proxy variables which attempt to indirectly capture this characteristic. For example, we have constructed weekly labour market history variables for each individual and also collected data detailing time spent on other DWP employment programmes.

There are likely to be many other unobserved variables, which to varying extents play a role in the participation decision. The value of having a rich data set is that, as described above for the *motivation* characteristic, some of the variables which we have observed will indirectly capture the influence of variables we have not observed. For example, while we have not been able to observe an individual's qualifications, personality type, life experience, experience of discrimination, confidence, health, language skills or happiness, we hope that by controlling for gender, age, ethnicity, disability, local labour market characteristics, labour market history, prior programme participation etc. the model will capture virtually all of their influence by proxy.

However, there may well be factors that influence whether or not an individual chooses to go on WE for which we cannot control. We try to construct the counterfactual as accurately as we can, using the observable characteristics for which we do have data.

Participation on other DWP programmes before WE start/pseudo start date might influence an individual's motivation or how easily they find work. In principle we could control for this by removing those people who were on another DWP programme before their WE start date. However, this would bias the sample by deliberately removing a particular type of claimant because the remaining sub sample may no longer represent those on WE. Therefore, a measure of previous participation on other DWP programmes has also been used as a matching variable. In the same light, it is not evident that we should remove claimants who have been on programmes after WE, because it is possible that we might omit claimants who are more likely not to have found work, thus affecting the size of the impact.

3.2 Propensity Score Matching

The aim of the Propensity Score Matching process is to construct a comparison group of individuals who did not participate in the WE programme, but who in aggregate have identical characteristics to those who did participate, in those characteristics which influence selection and outcomes. If this is successfully achieved, we can then use the labour market outcomes of non-participants in the comparison group as an approximation for the counterfactual, i.e. what the labour market outcomes of participants in our treatment group would have been if they had not participated.

When there are a large number of observed characteristics, as is the case in the present evaluation, direct matching on all characteristics becomes a limited device as the number of dimensions relative to the number of observations increases (Rosenbaum and Rubin, 1983)¹⁶. Therefore, we follow the literature in using a single balancing score on which to match, which is a function of all the observed variables. The balancing score used is a propensity score, which is the probability of an individual participating in the programme given all of their observed characteristics.

Below is a summary of the Propensity Score Matching (PSM) protocol used in this evaluation to construct suitable treatment and comparison groups from the participant and non-participant samples (Appendix 4 shows a step-by-step guide to the protocol).

Firstly, the probability of participation (dependent variable) was modelled¹⁷ using the observed individual characteristics of participants and non-participants, as independent variables. From this model, the predicted

¹⁶ An accessible explanation of how it can be applied to evaluation of labour market policy can be found in Bryson, et al. 2002.

¹⁷ Using a probit regression

probability of participation - the 'propensity score' - was calculated for each participant and non-participant in the sample. Secondly, a matched comparison group was constructed by matching¹⁸ each participant with the average of all non-participants with similar propensity scores, giving more weight to those whose score was nearest.

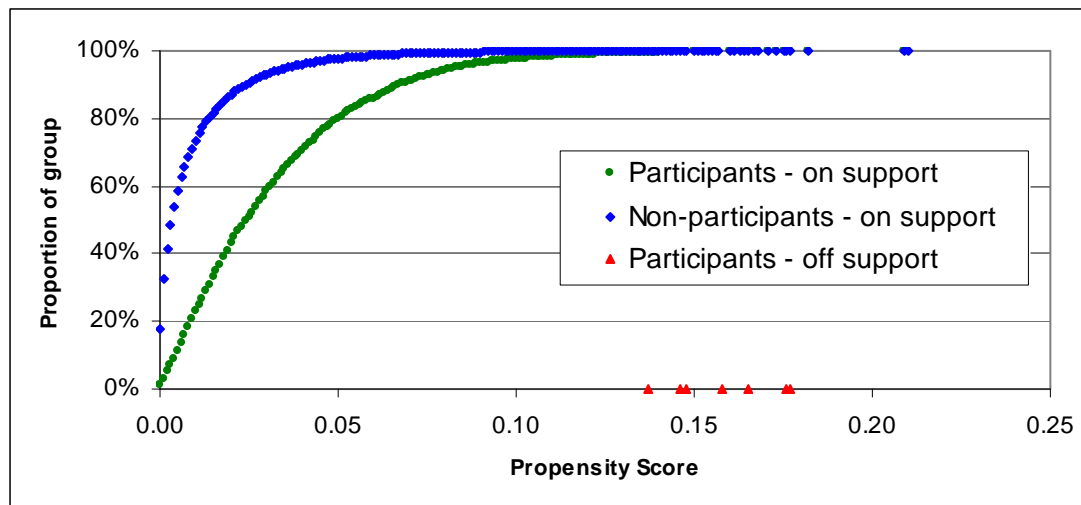
3.2.1 Common Support for Participants

For Propensity Score Matching to be a successful methodology for estimating the counterfactual, there must be sufficient common support for participants among the non-participant sample. This means that we must be able to find matching non-participants for the vast majority of our participants. This is important as any impact estimates are only valid for those participants for whom common support is available.

The propensity score distribution for the sample in our main analysis is given below; each point is the cumulative proportion of the group with propensity scores in increments of 0.001.

Figure 3.1: Propensity Score distributions of participants and non-participants:

(3,490 participants; 374,360 non-participants)



As expected, Figure 3.1 shows that propensity scores tend to be higher among participants. The overall mean propensity score is 0.009 (this is the

¹⁸ Matching used 'Kernel' matching. For each participant in the sample, all non-participants with propensity scores within the Kernel bandwidth were selected and weighted using an Epanechnikov distribution. The bandwidth determines how closely the propensity score of a non-participant must be to that of a participant for selection into the final matched comparison group. A bandwidth of 0.0001 was used for our analysis. This bandwidth was shown by Ainsworth and Marlow (2011) to provide a model that retained a high level of common support whilst also ensuring a tight match between non-participants and participants. The matching was carried out using an adaptation of the *Stata* code applied by Thomas (2006), which relies on the *Stata* module *psmatch2* written by Leuven and Sianesi (2003).

proportion of the sample receiving WE support). 24% of non-participants have a score of more than the mean propensity score (0.009) compared with 79% of participants. The small proportion of participants (<1%) for whom no common support is available all have propensity scores between 0.1 and 0.2.

The propensity score distribution provides a very high degree of overlap between participants and non-participants, with over 99% of participants finding common support. We can be confident therefore in generalising from those participants for whom there was a good match to the entire cohort.

3.2.2 Matching Quality

We found the propensity score model to be highly effective in constructing treatment and comparison groups that are well balanced on the observed characteristics. Table 3.1 shows specification statistics for the matching of the participant and non-participant groups. The chi-squared test shows that prior to the match, there was approximately zero probability that the participant and non-participant samples had the same set of characteristics. After matching there is statistically no difference between the matched groups in terms of observed variables at the 5% level.

Table 3.1: Specification statistics for the group matching

Sample	Pseudo R-sq	LR chi-sq	p > chi-sq
Unmatched	0.14	5,600	0.00
Matched	0.00	14	1.00

Table 3.2 below shows the unmatched and matched means of a range of variables for the participant and non-participant groups. The variables in this table are non-exhaustive and are provided to illustrate the extent to which the PSM methodology selects non-participants with similar characteristics across a range of variables. A full list is not provided here as there are over 250 individual variables included in the propensity score model. The table illustrates that the PSM methodology has been extremely effective in balancing the groups on the listed covariates. This high quality of matching is also observed across the range of variables not listed.

Early Impacts of Work Experience

Table 3.2: Unmatched and matched means for main analysis

Variable	Sample	Treatment	Comparison	% bias	% reduction in bias	t	p> t
Age (mean years)	Unmatched	20	21	-67.8		-35.8	0.0
	Matched	20	20	0.0	100	0.0	1.0
Male (%)	Unmatched	63	69	-12.8		-7.7	0.0
	Matched	63	62	1.1	92	0.4	0.7
Disabled (%)	Unmatched	13	14	-2.5		-1.4	0.2
	Matched	13	13	-0.7	72	-0.3	0.8
White ethnicity (%)	Unmatched	80	80	1.4		0.8	0.4
	Matched	80	80	0.2	85	0.1	0.9
Asian ethnicity (%)	Unmatched	7	6	2.3		1.4	0.2
	Matched	7	7	-0.5	80	-0.2	0.8
On JSA, ESA, IB, IS or TA 4 weeks before WE start date (%)	Unmatched	97	83	46.3		21.4	0.0
	Matched	97	96	1.7	96	1.2	0.3
On JSA, ESA, IB, IS or TA 30 weeks before WE start date (%)	Unmatched	47	47	-0.9		-0.5	0.6
	Matched	47	47	-0.2	80	-0.1	0.9
In work 4 weeks before WE start date (%)	Unmatched	12	23	-28.2		-14.9	0.0
	Matched	13	13	-0.9	97	-0.5	0.7
Chosen occupation is Sales (%)	Unmatched	35	24	23.6		14.7	0.0
	Matched	35	35	-0.4	98	-0.2	0.9
Local Authority Unemployment Rate (%)	Unmatched	8	9	-12.2		-7.3	0.0
	Matched	8	8	0.9	93	0.4	0.7
Days out of past two years spent on Flexible New Deal (mean)	Unmatched	7	23	-23.8		-11.7	0.0
	Matched	7	8	-0.9	96	-0.5	0.6
Lone Parent (%)	Unmatched	0.1	0.7	-8.5		-3.9	0.0
	Matched	0.1	0.2	-0.2	98	-0.1	0.9
Low Qualified (%)	Unmatched	25	22	7.4		4.5	0.0
	Matched	25	25	-0.6	92	-0.2	0.8

Notes:

The % bias is the difference between the sample means in the treatment and comparison groups as a percentage of the square root of the average of the sample variances in the treated and non-treated groups (Rosenbaum and Rubin, 1983).

Table 3.2 shows, for example, that before matching the proportion of participants who are male is 63%, while the proportion of non-participants who are male is 69%. After matching, the treatment and comparison groups both comprise very similar proportions of males (63% and 62% respectively).

The matching results in a reduction in bias for 245 out of the 251 variables used in the propensity score model¹⁹. Prior to matching, there were significant differences (at the 5% level) between participants and non-participants for 192 variables. After matching there were *no longer* significant differences (at the

¹⁹ These include the following dummy variables: 2 start/pseudo start month variables, 1 benefit start month variable, 2 district variables and the job density labour market variable.

5% level) between participants and non-participants for any of the 251 variables used. While the matching appears to be of good quality for the observed variables we cannot know if the same is true for the unobserved variables.

3.3 Measuring Outcomes and Impacts

To estimate the average effect of the WE employment programme on its participants (the *average effect of treatment on the treated*, *ATT*), we have measured two main sets of outcome data for each participant and non-participant. In each of the 21 weeks following a participant's WE start date (pseudo start date in the case of non-participants), we measure whether the individual was:

- claiming any out of work benefits (JSA, IS, IB, ESA) or training allowance; and
- in employment;

In all cases, we use DWP administrative data to determine individual benefit spells, and data originally sourced from the HM Revenue and Customs (HMRC) tax system to determine employment spells. Outcomes are identified at weekly intervals following a WE start or a pseudo start date (comparison group). A definitive outcome is assigned where a weekly point in time (7, 14, 21..... days after the programme start) is identified as being between a benefit spell start and end or an employment spell start and end (or there is a start but no accompanying end recorded). The outcomes measured are not mutually exclusive, so in any given week an individual may appear as both 'in employment' and 'in receipt of benefit'. The outcome period covers an independently calculated period of time for each individual, spanning from the individual's WE/pseudo start date to the end of the maximum period of the data available for the cohort (21 weeks later).

By comparing the outcome data of the matched treatment and comparison groups, we are able to estimate the impact of the WE programme on each outcome over time. When we present our results in Section 4, we therefore refer to the programme impacts on the likelihood of participants claiming benefit, and the likelihood of participants being in employment.

To calculate the net impacts of the programme on a particular outcome for a given week, we first take the mean outcome value of the treatment group (i.e. the proportion of the group who are receiving benefit or are in employment) and subtract the weighted mean outcome value of the comparison group. Thus a raw net impact measure is the absolute percentage point difference between the treatment and weighted comparison groups for the corresponding outcome.

However, some small differences in labour market history between the groups may exist after matching. Therefore, we use a difference-in-differences

Early Impacts of Work Experience

approach to adjust our impact measure to further reduce this bias. The estimated impacts are adjusted by the average pre-programme differences in labour market outcomes. However, for this particular analysis, this works out at an average of only 0.1 percentage points across the weekly benefit history variables. For example, if the estimated impact on benefit receipt were found to be -7.1 percentage points, but the average pre-programme difference in benefit receipt were -0.1 percentage points, then the impact estimates would be adjusted to gain a final difference-in-differences adjusted estimate of -7 percentage points (-0.1 subtracted from -7.1).

The impacts presented in Section 4 are therefore the difference-in-differences adjusted impacts of the programme on each of the outcomes of interest.

4. Results

In this section, estimates of the average effect of the WE programme on participants (the *average effect of treatment on the treated, ATT*) are presented.

Section 4.1 presents our main impact estimates for the WE programme. Section 4.2 explores the sensitivity and heterogeneity of the estimated impacts by performing the analysis with several different groups of participants and non-participants.

4.1 Impacts of work experience

The main estimates describe the impact of WE, for each of the 21 weeks following starting a placement, on a participant's probability of being:

- in receipt of benefit (JSA, IS, ESA, IS or training allowance); and
- in employment.

Impact graphs are extended over a 52-week period prior to starting on WE to illustrate the extent to which the Propensity Score Matching controls for labour market history over the pre-programme period. The impact graphs also show a 95% confidence interval around the central impact estimates²⁰.

Figure 4.1 shows the comparison of the proportion on benefit between the participants and matched non-participants and Figure 4.2 shows the impacts after the “difference in differences” adjustment.

Before describing the impacts we confirm the finding in Section 3.2 that the benefit history of non-participants in the 52 weeks prior to participation matches very closely to the benefit history of WE participants. This gives us confidence that the matching has been successful. However, it is important to keep in mind that the closeness of the match is probably more of a reflection of the fact that we have matched on that specifically, so there is no guarantee that we have eliminated selection bias.

In the first 8 weeks after starting, participants were more likely to be on benefit than non-participants. Since the period of WE placements is usually 2-8 weeks this is likely to reflect a ‘lock-in’ period when participants were engaged in WE, which reduced the time spent on job search activity. This effect is often seen in employment programmes.

²⁰ Standard errors are calculated using a linear probability model. The standard errors suffer from heteroscedasticity and non-normality. There is some debate in the literature as to the best way to calculate errors without being too computationally intensive. Our method probably gives an overly cautious approach to errors; errors calculated by `psmatch2` are up to half the size that we have reported and errors in other evaluations appear to be smaller for similar sample sizes, for example Lechner and Wunsch (2009).

Early Impacts of Work Experience

After this 8 week period, estimates show that WE had a strong and beneficial impact on the likelihood of a participant receiving benefit compared to non-participants - participants became much less likely to be on benefit than non-participants; the central estimate is about -6 percentage points by week 13. The impact then continued at that level until the end of the measurement period at week 21. Since about 46% of participants were off benefits compared to 40% for non-participants at week 21 this means that participants were about 16% more likely to be off benefits than non-participants at week 21.

Overall, the benefit impact over the first 21 weeks equates to participants being off benefit for an average of about 5 days more than if they had not participated.

Figure 4.1: Benefit receipt rate for WE participants and matched non-participants

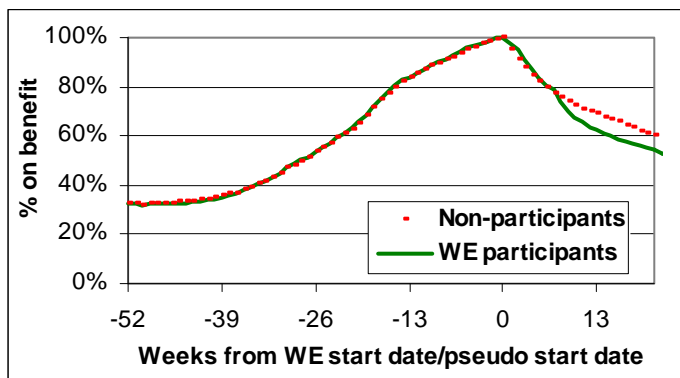


Figure 4.2: Impact of WE on the likelihood of a participant claiming benefit

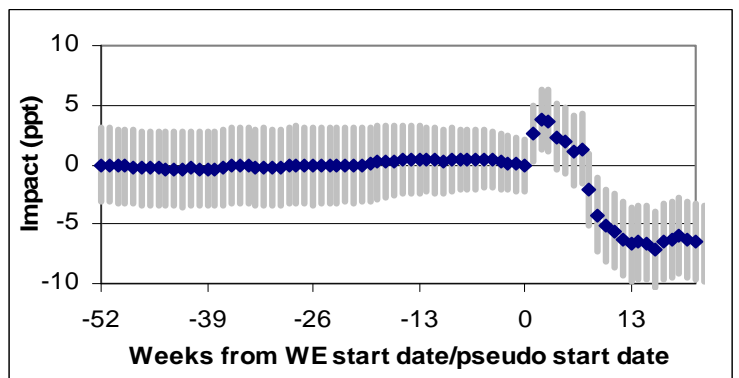


Figure 4.3 shows the comparison for the proportion in employment between the participants and matched non-participants and Figure 4.4 shows the impacts after the difference in differences adjustment.

Figure 4.3: Employment rate for WE participants and matched non-participants

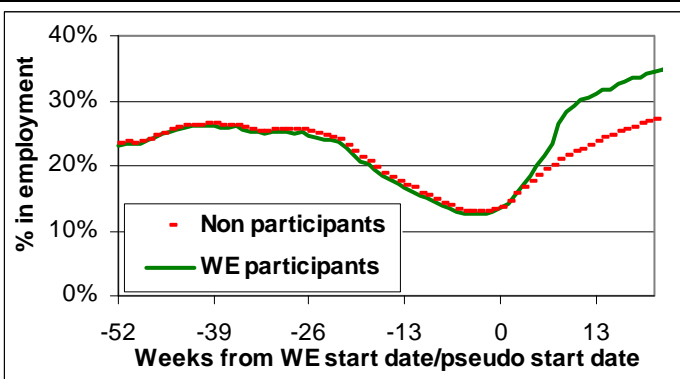
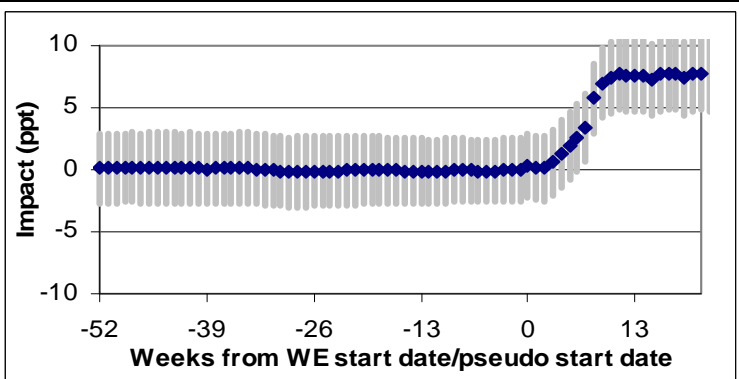


Figure 4.4: Impact of WE on the likelihood of a participant in employment



Important note on employment impacts

Figure 4.3 highlights imperfections in the P45 HMRC employment data and exposes gaps in our understanding of an individual's status. In Figures 4.1 and 4.3, at 52 weeks before starting, 23% were in employment and 33% were on benefit (there will be some overlap between these two figures). This means that at least 44% (how many more depends on the degree of overlap) who are recorded as neither on benefit nor in employment. Their status is unknown; some might be in education but others will be in employment but are not recorded to be in work at that time. Moreover, in the week of starting WE, 13% were employed; we would expect the real employment rate to be close to zero because all participants were actually claiming benefit in this week. As discussed in Section 2.2.2, we do not believe that the imperfections in the employment data will systematically bias our impact estimates. However, we acknowledge that the employment impact estimates rely on imperfect data and should be treated with a degree of caution.

Similar to benefit history, the employment history of non-participants in the 52 weeks prior to participation matches closely that of WE participants, giving us confidence that the estimated impacts from starting on the programme are due to the programme.

Over the first 8 weeks after starting, participants became more likely to be in employment than non-participants. This suggests that although more people were on benefit in this period, which we proposed to be due to programme 'lock in' effects, the data suggests that slightly more people were finding employment. It is difficult to tell whether this is a true empirical finding or a result of the imprecision of the employment data, so some caution should be attributed to this particular observation.

After the 8 week period, the estimates follow the pattern of the benefit impacts. WE had a strong and beneficial impact on the likelihood of a participant being in employment compared to non-participants - participants were more likely to be in employment than non-participants: the central estimate reached +8 percentage points by week 13. The impact then continued at that level until the end of the measurement period at week 21.

We should be cautious about translating these results in terms of how much more likely participants are to be in employment than non-participants because the calculation requires absolute proportions in employment (35% for participants compared to 27% for non-participants at week 21) which in turn relies on accurate recording of HMRC employment outcomes. If we assume these figures, based on limited employment data, are correct then we can say that participants were roughly 28% more likely to be in employment than non-participants at week 21, but we should not place too much emphasis on this figure.

Early Impacts of Work Experience

Overall the employment impact over the first 21 weeks equates to participants being in employment for an average of about 8 days more than if they had not participated.

At 21 weeks there is no sign of a decline in benefit or employment impacts so we would expect these impacts to continue beyond the period that can be tracked. Of course it is a matter of speculation how long these impacts will persist but to date 5 days off benefit/8 days in employment is a minimum estimate (assuming of course that the impact does not become negative in the future).

In the following we postulate possible factors behind the difference between the benefit and the employment impacts. We note first that the actual impacts observed were -6.4 percentage points for benefit and 7.7 percentage points for employment, so the difference is not actually as large as 2 percentage points as implied by the rounded figures. We also stress that we believe we should not read too much into why there is a difference due to the unreliability of the employment data.

First, we note that the outcomes and therefore the impacts are not mutually exclusive; table 4.1 illustrates the impacts at week 21 with regard to the four mutually exclusive labour market groups. It shows that participants are slightly less likely (-2.7 percentage points) to be neither in employment nor on benefit, which indicates that participants were less likely to leave benefit for a destination other than employment.

Table 4.1: Impacts (percentage points) on mutually exclusive outcomes at 21 weeks

Benefit only	Work Only	Neither	Both
-5.0	+9.1	-2.7	-1.4

One explanation, mentioned in Section 2.2.2, for the difference between the employment and the off-benefit impact may be in the nature of the employers participating in WE. For example, if employers participating were disproportionately large this would probably mean that they would be more likely to submit employment data to HMRC for all their employees, whereas smaller employers may only submit data on employees earning above the PAYE threshold; they may also submit their data more quickly and with fewer errors. This would mean that employment for employees working for WE employers would be more likely to be recorded in the data, and/or recorded more quickly, and hence may bias upwards the estimates of the employment impact of WE.

Currently there is no evidence to suggest WE is attracting disproportionately larger employers, so it would therefore be useful to explore in future whether employers participating in WE are different from other employers.

The effects of other support on the impact estimates

The impact estimates provided in Section 4.1 do not describe the impact of WE against a baseline of 'no support', but rather the impact of WE against a baseline of other Jobcentre Plus support and time on other DWP employment programmes.

The Department does not hold centrally comprehensive data for the amount of Jobcentre Plus support that each claimant receives. Therefore, as with other evaluations of this type, there is no other option but to assume that participants and non-participants receive similar levels of standard Jobseeker support of fortnightly signings and contact with Jobcentre Plus advisors.

As outlined in Sections 2 and 3, we have used a number of variables which capture time spent on other DWP programmes *before* the WE start or pseudo start date in our propensity score model. The aim of including these variables was both to balance the groups with regard to any positive effects that previous programmes may have had on labour market outcomes, and also to capture any associated characteristics, such as motivation, by proxy.

However, participants and non-participants may also have spent time on alternative DWP programmes *after* their WE start/pseudo start date. We found that, on average, non-participants spent an additional five days on other DWP programmes in the 21 weeks following their pseudo start date than participants spent over the equivalent period.

If we estimate the impacts of WE against 'no support' we may therefore expect the impacts to be slightly different. Over this short period of time, whether the impacts are weaker or stronger will largely depend on the strength of the 'lock-in' on alternative programmes relative to WE. However, as it is difficult to disentangle other support from our estimates of WE impacts and the difference appears quite small (given that time on WE ranges from 14 to 56 days), we do not attempt to adjust the impacts for this effect.

4.2 Sensitivity of Impact Estimates

So far the analysis has been focused on the cohort of starts aged 19-24 between January and May 2011 compared to a group of non-participants taken from 19- to 24-year-old JSA claimants. Four additional analyses are presented in Appendix 5 to explore the sensitivity and the heterogeneity of the estimates:

- Impacts of WE on an earlier cohort of participants from January to March;

Early Impacts of Work Experience

- Impacts of WE on sub-groups aged 19-20 and 21-24;
- Impacts of WE on a larger cohort that includes 18-year-olds; and
- Impacts of WE using a comparison group of non-participants aged 25.

Overall we found that the impact estimates were largely insensitive to each of these alternative implementations. This provides increased confidence that the methodology is robust and that our findings are not biased by the definition of our participant and non-participant samples. Moreover, the small early cohort offers evidence for continued persistence of the impacts.

5. Conclusions

We have performed an impact analysis of work experience (WE) by comparing the benefit and employment outcomes of a cohort of participants with those of a matched comparison group, as an estimate of the counterfactual.

The central impact estimates for starts between January and May 2011 show that WE decreased the likelihood of claiming benefit by 6 percentage points after 21 weeks following starting on placements. This means that WE participants were about 16% more likely to be off benefits than non-participants after 21 weeks.

WE also increased the likelihood of being in employment by nearly 8 percentage points after 21 weeks following starting on placements. This means that WE participants were about 28% more likely to be in employment after 21 weeks. However, we note (as we describe amongst the caveats below) that we believe that the employment impacts are less robust than the benefit impacts.

This also means that for this short period, after accounting for initial programme 'lock-in', participants were off benefits for an average of 5 days more and in employment 8 days more, per participant, compared to non-participants.

There is no evidence of a decline in impacts so we would expect these impacts to persist for a much longer time. In particular, impacts from a smaller earlier cohort showed that the impacts persisted at similar levels until 30 weeks. However, it is really too early to speculate on how long the impact might continue in order to estimate the cost-effectiveness of the programme.

Some sensitivity tests were performed using different cohorts and sub-groups. We found that the impact estimates were largely insensitive to each of these alternative implementations. This provides increased confidence that the methodology was robust and that the findings are not biased by the definition of the chosen participant and non-participant samples.

Nevertheless, the analysis is complex and caution should be applied to the results, least of all because this is a first impact analysis, based on a small cohort of starts from the early months of WE. The caveats to note are:

- Participants have to demonstrate a willingness to work with the adviser. This is extremely difficult to capture in the matching approach. We have tried very hard to design the matched comparison group so as to mimic the counterfactual as accurately as possible and capture the motivational differences between participants and non-participants. However, we recognise that although the matching on observed variables appears to be of good quality there is a strong possibility that we have not captured all self selection bias from unobserved variables. This is a common risk for any evaluation of labour market policies where individuals actively choose to participate.

Early Impacts of Work Experience

- Impact is not against 'nothing' but against a background of support, which includes Jobcentre Plus support and other programmes. Consideration of the time spent on other programmes suggests that the impacts against a background of standard Jobcentre Plus support might increase or depress the impacts marginally. As the availability of other support increases over the latter half of 2011, any future analysis must deal with this issue more comprehensively.
- The cohort of participants up to the end of May 2011 includes a high proportion of 18- to 21-year-olds (88%) compared to the level of representation on the programme since then until November (69%). Therefore, just on this basis, care should be taken in extrapolating any results.
- In common with other evaluations, the employment impacts are less robust than benefit impacts because the HMRC tax data does not capture all employment outcomes, and cannot always be matched to benefit records. This means the benefit and employment impacts cannot be directly compared, and the benefit impact is likely to be a more robust estimate. The analysis presented here provides assurance that there is an impact on employment as well as benefit exits.
- Another reason for being cautious about extrapolating the estimates is that we cannot be sure if the first employers to participate in WE were more likely than employers who participated later to have available jobs or apprenticeship places. They may have been keen to participate because they had vacancies and could use WE to screen potential employees/apprentices. It is possible that employers that participated later may have been reluctant to participate because they had no such vacancies to fill. The Department does not have centralised data on employers but analysis of whether the percentage of WE participants who ended up working for their WE employers after the programme ended had changed over time could give some indication of whether earlier employer participants were more likely to have permanent opportunities than later employer participants.
- These are estimates of the additional effect on participants but **not** on the overall labour market, which might be less (particularly in the short-term) if there are knock-on effects on other groups. For example it is possible that the programme may have incurred some substitution effects (i.e. an employee taken on as part of WE by an employer is substituted for someone else who would have been hired). It is possible that this has occurred but our current belief is that for this type of programme the net effect is likely to be quite small.

Appendices

Appendix 1 – Defining the Participant Sample

Table A.1 below shows how the participant sample of 3,490 is derived from the total number who started a WE placement between January 2011 and November 2011. It shows that of the 34,200 WE participants, 4,890 started during our cohort period (January 2011 to May 2011). Of these, 4,580 were recorded on our systems as receiving Jobseeker's Allowance in the week of their WE start date. Finally, restricting the sample to participants aged between 19 and 24 years old left 3,490 participants. This was the final participant sample used in our main analysis.

Table A.1 Sampling valid participants using selection conditions

Condition for selection	Valid participants remaining
WE starts (January 2011 to November 2011)	34,200
Participant starts during the cohort period (January 2011 to May 2011)	4,890
Participants recorded as receiving Jobseeker's Allowance in the week that their WE starts	4,580
Participants aged 19 to 24	3,490

Appendix 2 – Matching on other DWP Programmes

Participants and non-participants may have been on employment programmes other than WE prior to their start or pseudo start date. To ensure that the impacts we measure are balanced with regard to the amount of past additional support received, we match on the time spent on each of 13 DWP employment programmes in the year prior to the start or pseudo start date. Programmes included are as follows²¹:

- Flexible New Deal (FND);
- New Deal for Young People (NDYP);
- New Deal for Long Term Unemployed (NDLTU)
- New Deal for Disabled People (NDDP);
- Pathways to Work (PtW);
- European Social Fund (ESF).
- New Deal for Lone Parents (NDLP);
- Basic Skills (BS);
- Work Based Learning for Adults (WBLA);
- Employment Zones (EZ);
- New Deal for Partners (NDP);
- Young Person's Guarantee (YPG); and
- Six Month Offer (6MO).

We use the number of days spent on each of these programmes in the year prior to start or pseudo start date as matching variables in the propensity score model. One exception to this is that for YPG and 6MO we do not have accurate end dates for all spells. Therefore, we use a single binary variable, which indicates whether an individual has had one or more spells on either of these programmes.

²¹ Future analysis would need to include the effect of other Get Britain Working Measures and the Work Programme.

Appendix 3 - Controlling for Labour Market History

As discussed in section 3, we believe that the labour market history of an individual provides an important proxy for unobserved characteristics, such as motivation to work, which will jointly influence both the participation decision and the outcomes in the absence of participation. It is therefore important to control for benefit and employment history in our propensity score model.

A common method described in the literature for controlling for labour market history is the approach adopted by Card and Sullivan (1988), in which a single variable is constructed to describe the labour market position of each individual over time. However, Ainsworth and Marlow (2011) proposed an alternative method of controlling for labour market history, which we have adopted in the present analysis. This appendix outlines the advantages of using the Ainsworth and Marlow approach.

In the approach utilised by Card and Sullivan, a single variable is constructed to represent an individual's labour market history. For example, a string variable of eight binary characters could represent whether an individual was in or out of employment in each of eight time periods. This approach has the advantage that a single variable can indicate not just the length of time an individual has spent receiving benefit or in employment, but also represents a timeline of moving in and out of each labour market state.

A disadvantage of this approach is that the number of permutations of the constructed string variable is 2^N , where N is the number of time periods independently represented in the history string. Therefore, each additional time period included in the string doubles the number of possible permutations. Constructing a variable which describes eight periods of three months (i.e. two years of benefit history) therefore results in 255 ($2^8 - 1$) dummy variables. Using such labour market history variables therefore requires a trade-off between ensuring the quality of the labour market variable (in terms of describing labour market history with sufficient resolution over a sufficiently long duration) and ensuring that the variable is not over-specified by producing too many dummy variables in the propensity score model.

We have therefore adopted the alternative approach proposed by Ainsworth and Marlow to control for labour market history, which is adapted to control for labour market history with higher resolution over longer durations. To implement this method we generate 52 independent binary variables which represent an individual's benefit receipt or non-receipt in each of the 52 weeks prior to the WE start or pseudo start date. We then generate a further 52 independent binary variables which represent whether an individual is in or out of employment in each of the 52 weeks prior to the WE start or pseudo start date. In this way, we are able to control for 52 weeks of labour market history using the resulting 104 variables in the propensity score model. To gain equivalent resolution and duration using the approach adopted by Card and Sullivan would have required approximately 2^{52} variables in the model.

Early Impacts of Work Experience

A disadvantage of our approach is that it is not clear that it allows for identification of differences in labour market dynamics between the groups. So further work could be performed in augmenting these variables with some additional summary measures to capture churn (such as number of employment spells, average length of spells).

Appendix 4 – Matching Protocol

We describe below the matching protocol used in this evaluation to construct suitable treatment and comparison groups from our participant and non-participant samples.

1. Define a participant (treatment) sample within the analysis cohort period, as specified in Section 2.1.1;
2. Define a non-participant (comparison) sample within the same cohort period, as specified in Section 2.1.2;
3. Combine the records from steps 1 and 2 to produce a single sample comprising treatment and comparison records;
4. Code an indicator variable Z , which is 1 for treatment records and 0 for comparison records;
5. Specify and estimate a binary probit for $p(x) := P(Z=1 | X=x)$;
6. Restrict the sample to common support: remove all treatment records for which no comparison record falls within the Kernel bandwidth (a bandwidth of 0.0001 was used in our primary analysis);
7. Implement a Kernel ‘one-to-many’ matching approach:
 - a. Select a treatment record and identify all comparison records with propensity scores lying within the Kernel bandwidth of the treatment record score;
 - b. Apply a weighting to the comparison records using an Epanechnikov distribution such that those with closer propensity scores to the treatment record are weighted more than those with more distant propensity scores;
 - c. Repeat steps *a* and *b* until all treatment records have been selected. The weighting applied to comparison records for each repeated step is added to the cumulative weighting from all previous steps (the total weighting of all comparison records is therefore equal to the number of treatment records).
8. Use the final weights for each comparison record to calculate a weighted mean for each outcome variable in $Z=0$;

Appendix 5 – Sensitivity Analysis

Impacts of WE on an early cohort of participants

The rationale for estimating the impacts from an early cohort from January to March is threefold:

- the characteristics of the early cohort were slightly different (e.g. they were younger) and so it is important to observe whether the impacts were different;
- sometimes it can take a while before the programme can start working effectively - so it is important to observe whether there was a difference in impacts; and
- it allows for a longer measurement period: outcomes of up to 30 weeks.

With this in mind, the impacts for a small subset of 940 WE starts on benefit at the start of the programme between January and March were compared to the subset of non-participants with pseudo starts from the same period.

After an initial lock-in period the benefit impact reached -8 percentage points after 10 weeks and then remained between -7 and -9 percentage points from week 10 to the end of the measure period at week 30. Overall the benefit impact over the first 21 weeks equated to participants being off benefit for an average of 6 days more than non-participants; over the first 30 weeks this equated to 11 days more.

Similarly the employment impact reached +10 percentage points after 10 weeks and then remained between +9 and +11 percentage points from week 10 to the end of the measurement period at week 30. Overall the employment impact over the first 21 weeks equated to participants being in employment for an average of 11 days more than if they had not participated; over the first 30 weeks this equated to an additional 17 days.

At week 30 there was still no sign of a decline in benefit and employment impacts so again we would expect these to continue beyond the measurement period. Of course it is a matter of speculation as to how long these impacts will persist but to date an additional 11 days off benefit/17 days in employment is a minimum, assuming that the impacts do not become negative.

In conclusion, the impacts followed a similar pattern to that of the main cohort, and is a little more evidence of the persistence of the impacts.

Impact of different age groups

The rationale for this analysis is to explore the sensitivity of impacts by age. This is particularly important because a higher proportion of starts in January

to May were from a younger group than starts from the whole period of January to November.

The analysis explored the impacts of two age groups: a 19- to 20-year-old participant group (66% of the main cohort) and a 21- to 24-year-old participant group (34% of the main cohort). Both these groups were compared with non-participants of the corresponding age groups.

The impacts for jobseekers aged 19 to 20 and aged 21 to 24 were similar to those for the main cohort. At 21 weeks:

- for 19- to 20-year-olds the benefit and employment impacts reached -6 percentage points and +8 percentage points respectively.
- for 21- to 24-year-olds the benefit and employment impacts reached -9 percentage points and +8 percentage points respectively.

This analysis offers some indication that the main estimates are likely to be valid for 18- to 24-year-old participants as well as for 19- to 24-year-olds.

Impacts using non-participants aged 25

The main analysis compared a cohort of 19- to 24-year-old participants with a matched group of 19- to 24-year-old non-participants. The rationale of using non-participants of the same age range is that this group is likely to be more similar to the WE participant sample with regard to both observed and unobserved characteristics.

The disadvantage of using non-participants of the same age range as WE participants is that, in addition to people who were never offered WE as an option, the comparison group also includes:

- non-participants who were offered WE and decided not to volunteer; and
- non-participants who agreed to take a WE placement, but then didn't actually start the placement.

We think that any selection bias introduced by such non-participants will be relatively small, as these individuals will only make up a small proportion of the total non-participant sample.

However, in order to compare WE participants with non-participants who are very unlikely to have had the opportunity to go on WE, we have compared the 19- to 24-year-old participant group with a non-participant group aged 25.

The main disadvantage (and the principal reason why we have not used it in our main analysis) is that the mean ages of the groups are very different (20 years old for the participant group and 25 for the comparison group) and therefore the two groups are likely to differ in many other observed and unobserved characteristics.

The impacts for participants aged 19 to 24 against non-participants aged 25 were very similar estimates to those for the main analysis. At 21 weeks, the benefit and employment impacts reached -8 percentage points and +9 percentage points. However, we note that as expected common support is lower (84%) than for the main analysis because of the difference in age groups.

Impacts of 18- to 24-year-olds

The main analysis does not include participants aged 18. As discussed in Section 2.1, this is because they do not have as much work and benefit history as older participants to control for through the Propensity Score Matching methodology, and there is no other data which could serve the same purpose, such as educational achievements. However, while we may expect the matched groups to be less well balanced on characteristics than in our main analysis, it is still useful to test the sensitivity of including 18-year-olds on the estimated impacts.

The impacts for jobseekers aged 18 to 24 were very similar estimates to those for the main analysis. At 21 weeks, the benefit and employment impacts reached -7 percentage points and +8 percentage points. Although there will always be an increased concern as to whether characteristics have been sufficiently controlled for, this analysis offers some indication that the main estimates are likely to be valid for 18- to 24-year-old participants as well as for 19- to 24-year-olds.

References

Ainsworth, P. and Marlow, S. (2011), Early Impacts of the European Social Fund 2007-13, DWP In-House Research Report No. 3.

Beale I., Bloss C. and Thomas A. (2008), The longer term impact of the New Deal for Young People, DWP Working Paper 23.

Dolton, P. and Smith, J. (2011), The Impact of the UK New Deal for Lone Parents on Benefit Receipt, IZA DP No. 5491.

Bryson A., Dorsett R. and Purdon, S. (2002), The Use of Propensity score matching in the Evaluation of Active Labour Market Policies, DWP Working Paper No. 4.

Rosenbaum, P. R. and Rubin, D. B. (1983), The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* 70, 41–55.

Card, D. and Sullivan, D. (1988), Measuring the effect of subsidized training programmes on movements in and out of employment, *Econometrics* 56(3), 497-530.

Authors

Paul Ainsworth
Adam Hillmore
Simon Marlow
Stuart Prince

Department for Work and Pensions