In-House Research

Early Impacts of the European Social Fund 2007-13

by Paul Ainsworth and Simon Marlow

DWP Department for Work and Pensions

Department for Work and Pensions

In-House Research No 3

Early Impacts of the European Social Fund 2007-13

Paul Ainsworth and Simon Marlow

A report of research carried out by the Department for Work and Pensions

This document/publication is also available on our website at:
http://research.dwp.gov.uk/asd/asd5/ihs-index.asp

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

Department for Work and Pensions, Commercial Support and Knowledge Management Team, Work and Welfare Central Analysis Division, Upper Ground Floor, Steel City House, West Street, Sheffield S1 2GQ

First Published 2011

ISBN 978-1-84947-615-7

Contents

List of Figures

List of Tables

Acknowledgements

The authors would like to thank Andrew Thomas and Mike Daly at the Department for Work and Pensions for their ongoing support and advice throughout the duration of this research.

Special thanks also go to Alex Bryson and John Forth at the National Institute for Economic and Social Research (NIESR) for their expert guidance and thorough quality assurance of this work.

We are also grateful to everyone at the Department for Work and Pensions who took the time read through the many drafts of this report; in particular Mike Daly, Andrew Thomas, Mike Jones, Ellenor Brooks, David Oatley, Nick Gilhooly, Chris Kent and Amanda Rowlatt. The feedback and comments we received have undoubtedly led to a more informative and methodologically robust evaluation than would otherwise have been possible.

Abbreviations

ATT	Average effect of Treatment on the Treated
CFO	Co-financing Organisation
CIA	Conditional Independence Assumption
DiD	Difference-in-Differences
DLA	Disability Living Allowance
DWP	Department for Work and Pensions
ESA	Employment and Support Allowance
ESF	European Social Fund
EU	European Union
EZ	Employment Zones
HMRC	Her Majesty's Revenue and Customs
IB	Incapacity Benefit
ILM	Intermediate Labour Market
IMD	Index (or Indices) of Multiple Deprivation
IS	Income Support
JCP	Jobcentre Plus
JSA	Jobseeker's Allowance
LMS	Labour Market System
LSC	Learning and Skills Council
NBD	National Benefits Database
ND50+	New Deal 50+
NDDP	New Deal for Disabled People
NDIF	New Deal Innovation Fund
NDLP	New Deal for Lone Parents
NDLTU	New Deal for Long Term Unemployed
NDP	New Deal for Partners
NDYP	New Deal for Young People
NEET	Not in Education, Employment or Training
PSM	Propensity Score Matching
UK	United Kingdom
WBLA	Work Based Learning for Adults
WBLA (BET)	Work Based Learning for Adults (Basic Employability Training)
WBLA (LOT)	Work Based Learning for Adults (Longer Occupation Training)
WBLA (SJFT)	Work Based Learning for Adults (Short Job-Focused Training)

Summary

The European Social Fund (ESF) was set up to improve employment opportunities in the European Union and so help raise standards of living. Its aim is to help people fulfil their potential by giving them better skills and better job prospects[1]. A key feature of ESF funding is that it must be used to purchase additional provision in order to extend coverage, address gaps and complement domestic funding. The provision itself is varied and flexible, including activities such as job search guidance, basic skills training, case worker support and advice on tackling specific barriers to work.

This paper describes findings from an evaluation of the net impacts of the 2007-13 European Social Fund (ESF) Programme for England. The study is focused on the Department for Work and Pensions (DWP) ESF funded[2] employment provision part[3] of the programme, which is contracted by DWP during 2008-11, and delivered by private, public and third sector providers at an expected cost of £265 million. Our analysis focuses on participants who entered the programme between June 2008 and April 2009 and estimates the programme impacts on two broad DWP customer groups: participants in receipt of Jobseeker's Allowance and participants in receipt of Incapacity Benefit or Employment Support Allowance.

The evaluation sample and measured outcomes

Entry onto the DWP ESF employment programme is voluntary and is available to anyone who is not in employment. This broad eligibility leads to a greater heterogeneity of participant characteristics than might be expected for other DWP employment programmes where eligibility may be dependent on receipt of particular benefits and/or the duration of a benefit claim.

For this reason, when performing our primary analysis we estimate impacts separately for two broad groups of participants:

1. Jobseeker's Allowance (JSA) customers; and
2. Incapacity Benefit (IB) or Employment Support Allowance (ESA) customers.

The JSA sample comprises 25,720 ESF participants who entered the programme between June 2008 and April 2009 and were receiving Jobseeker's Allowance at the time of programme entry.

The IB/ESA sample comprises 1,970 ESF participants who entered the

[1] http://www.dwp.gov.uk/esf/about-esf/

[2] The ESF funded part of the programme uses money provided by the European Social Fund. The programme also includes an equivalent amount of 'matched' funding, which is provided by the Co-financing Organisations responsible for distributing the ESF funds. In this paper we concern ourselves only with the ESF funded part of the programme.

[3] This is defined by the ESF Programme as Priorities 1 and 4.

programme between June 2008 and April 2009 and were receiving either Incapacity Benefit or Employment Support Allowance at the time of programme entry.

Comparison groups of non-participants in receipt of JSA and IB/ESA are drawn from the population of individuals who could have entered the programme during the same time period as participants in the sample. Groups of non-participants are selected who most closely resemble ESF participants with regard to demographic characteristics, benefit and employment history and prior participation on DWP programmes.

The matched groups of participants and non-participants are compared over time with regard to the proportion of the group who are:

- In receipt of their main benefit (the benefit they were receiving on programme entry);

- In receipt of any main working age DWP benefit (Jobseeker's Allowance, Incapacity Benefit, Employment Support Allowance or Income Support); and

- In employment.

This comparison between matched groups provides estimates of the net impacts of DWP ESF employment support on the labour market prospects of those who participate (the average treatment effect on the treated, ATT). Impacts are estimated for 52 weeks following entry to the programme.

Impacts of the Programme on JSA customers

As shown in Figure 1, the **impact of the programme on JSA receipt** for the JSA customer group is positive (between +1 percentage point and +3 percentage points) in each of the 52 weeks following participation, which suggests that participation slightly ***increases an individual's chances of claiming JSA over this period***. This effect is statistically significant[4] for most of the 52 week period.

Figure 3 shows that the **impact of the programme on the receipt of any main working age benefit**[5] for this group is positive for the first 15 weeks after participation (between +1 percentage points and +3 percentage points), indicating that participation slightly ***increases an individual's chances of claiming benefit over this period*.** Beyond 15 weeks after the start of participation the impact is not significantly different from zero.

[4] Unless otherwise specified, we report results as statistically significant if they are significant at the 5% level.

[5] In this paper, the *rate of receipt of any main working age benefit* is the proportion of the stated group receiving any of the following DWP benefits: Jobseeker's Allowance, Income Support, Incapacity Benefit or Employment Support Allowance.

As shown in Figure 5, the **impact of the programme on employment** for this group is positive in each of the 52 weeks following participation, rising from 0 percentage points to +4.5 percentage points towards the end of this period. This indicates that participation ***increases an individual's chances of being in employment over this period***. This effect increases over time and is statistically significant for almost all of the 52 week period.

There is no contradiction in the finding that both the benefit and employment impact estimates are positive for the first 15 weeks following participation. This is because the impacts are not mutually exclusive and do not account for all possible outcomes. In Appendix 5 we show that the positive benefit and employment impacts are balanced by a negative impact on the labour market position 'neither receiving benefit nor in employment'.

Impacts of the Programme on IB/ESA customers

As shown in Figure 2, the **impact of the programme on IB/ESA receipt** for the IB/ESA customer group is negative in each of the 52 weeks following participation (reaching a minimum of nearly -14 percentage points and then declining to -13 percentage points at 52 weeks), suggesting that participation substantially ***decreases an individual's chances of claiming IB/ESA over this period*.** This effect is statistically significant for almost all of the 52 week period.

Figure 4 shows that the **impact of the programme on the receipt of any main working age benefit** for this group is also negative in each of the 52 weeks following participation (reaching a minimum of almost -11 percentage points and then declining to -9 percentage points at 52 weeks), suggesting that participation substantially ***decreases an individual's chances of claiming benefit over this period***. This effect is also statistically significant for most of the 52 week period.

As shown in Figure 6, the **impact of the programme on employment** for this group is positive in each of the 52 weeks following participation (peaking at +12 percentage points and then declining to +11 percentage points at 52 weeks), suggesting that participation substantially ***increases an individual's chances of being in employment over this period***. This effect is statistically significant for almost all of the 52 week period.

In summary, ESF provision has low impacts on Jobseeker's Allowance recipients, but is far more effective for Incapacity Benefit and Employment Support Allowance recipients over the 52 weeks following participation.

This paper additionally investigates whether the effectiveness of ESF support for JSA customers varies according to the demographic characteristics of participants or the type of support provided. Our findings show that the impacts of the programme are fairly homogeneous across the broad range of participant characteristics and across the range of support offered.

Figure 1: Impact on likelihood of claiming main benefit (JSA group)

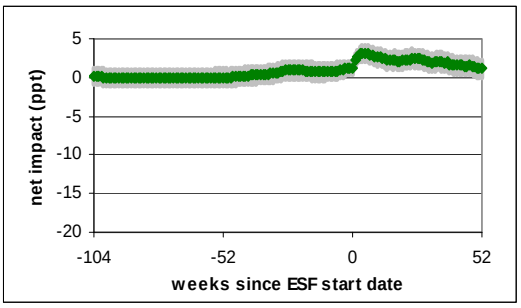


Figure 2: Impact on likelihood of claiming main benefit (IB/ESA group)

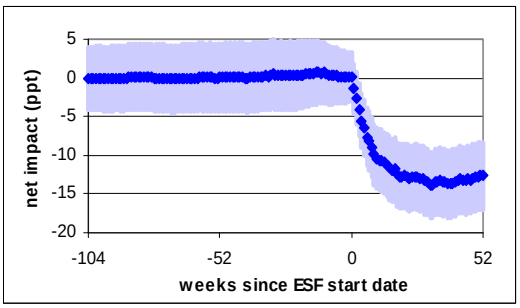


Figure 3: Impact on likelihood of claiming any working age benefit (JSA group)

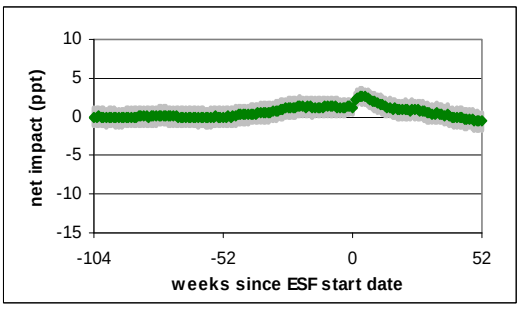


Figure 4: Impact on likelihood of claiming any working age benefit (IB/ESA group)

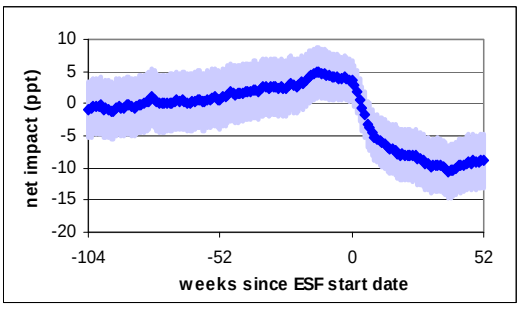


Figure 5: Impact on likelihood of being in employment (JSA group)

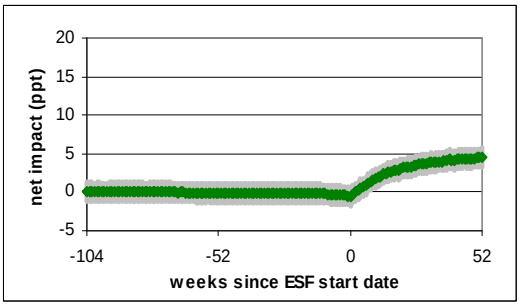
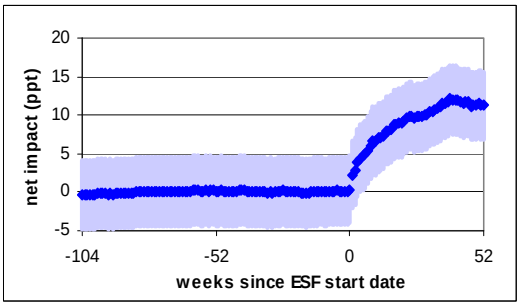


Figure 6: Impact on likelihood of being in employment (IB/ESA group)



We propose a number of possible explanations for the findings reported in this paper:

- The voluntary nature of ESF provision means that ESF participants are likely to be more 'work-ready' than non-participants. This is particularly the case among JSA customers who tend to be less disadvantaged and closer to the labour market than IB/ESA customers. It is therefore likely that many JSA customers participating on ESF could have achieved jobs without the additional support provided by ESF;

- Incapacity Benefit customers (who make up the majority of the IB/ESA participant group) have a lower base-level of employment support and

tend to be further away from the labour market than JSA customers. This could explain why impacts the of ESF provision are larger for the IB/ESA group than the JSA group;

- JSA customers tend to move away from benefit receipt quickly even without additional support. The short-term impacts for this group are negative, possibly because time spent on the programme leads to a reduction in job search activity (lock-in effect);

- There may be softer outcomes of the programme, such as increased skills, which would not necessarily be observed in our current short term impact measures, but may improve labour market prospects in the long term.

We note also three important caveats around the findings of this evaluation:

- As is the case in all non-experimental programme evaluations we can never be absolutely certain that we have fully accounted for the potential bias arising from selection onto the programme. However, the methodology we use, as described in this paper, takes all reasonable steps to minimise this bias;

- The evaluation considers only a part of the ESF programme: employment provision financed by the European Social Fund (ESF) and provided through the Department for Work and Pensions (DWP) between June 2008 and April 2009. Findings of this evaluation should not be presumed to indicate the impacts of the entire 2007-2013 ESF programme.

- The evaluation looks at net impacts in terms of benefit receipt and employment rates only. Other evaluations, such as the ESF Cohort Survey (Drever and Lloyd, 2010), have found positive effects that lie outside the scope of this analysis. For example, participants experience improved confidence and job readiness and many participants feel that the programme has helped them to gain basic skills and vocational qualifications.

1 Introduction

This paper describes findings from an evaluation of the net impacts of the 2007-13 European Social Fund (ESF) Programme for England. The study is focused on the Department for Work and Pensions (DWP) ESF funded[6] employment provision part[7] of the programme, which is contracted by DWP during 2008-11, and delivered by private, public and third sector providers at an expected cost of £265 million. Our analysis focuses on participants who entered the programme between June 2008 and April 2009 and estimates the programme impacts on two broad DWP customer groups: participants in receipt of Jobseeker's Allowance (JSA) and participants in receipt of Incapacity Benefit (IB) or Employment Support Allowance (ESA).

1.1 Rationale for analysis

Estimating the net impacts of an employment programme is vital to understanding its effectiveness in helping people to move away from benefit receipt and into employment. Impact estimates provide an objective measure of programme effectiveness, which allows policy makers to compare different employment programmes and inform policy decisions.

As is the case when estimating the impacts of any voluntary employment programme, we face a fundamental evaluation problem. Many participants who receive support subsequently enter employment or move away from receipt of benefits, but we can never be certain whether this was directly due to the support they received. We can observe the labour market outcomes (in employment, receiving benefit etc.) of each participant after they participate in the programme, but we cannot observe the 'counterfactual' outcomes that would have happened if they had not entered the programme. To understand the direct impacts of the programme, we must therefore find a way to estimate these counterfactual outcomes.

Most net impact evaluations of voluntary employment programmes rely on statistical techniques to construct a suitable comparison group of non-participants who can best represent what would have happened to participants if they had not entered the programme. This relies on having a rich data set describing the characteristics of individual participants and non-participants. In the past, such a data set has been unavailable for the ESF programme, but the recent availability of individual-level participant data has enabled us to construct a rich data set from a range of data sources. Using

[6] The ESF funded part of the programme uses money provided by the European Social Fund. The programme also includes an equivalent amount of 'matched' funding, which is provided by the Co-financing Organisations responsible for distributing the ESF funds. In this paper we concern ourselves only with the ESF funded part of the programme.
[7] This is defined by the ESF Programme as Priorities 1 and 4.

this data we are able to construct treatment and comparison groups from which we can estimate the net impacts of the programme.

Evaluation of the ESF programme is further complicated by the broad and flexible eligibility criteria and the wide variety of support offered to participants. To overcome some of the difficulties posed by this heterogeneity, we provide separate 'primary' impact estimates for two broadly defined DWP customer groups: participants in receipt of Jobseeker's Allowance and participants in receipt of Incapacity Benefit or Employment Support Allowance. We then perform 'sub-group' impact estimates to explore whether the programme impacts for Jobseeker's Allowance customers vary according to the characteristics of participants and the type of support they receive, by comparing the impacts for a number of participant and support sub-groups with our primary impact estimates.

This report makes an important step towards assessing the effectiveness of DWP employment provision funded by the ESF, which currently accounts for about a tenth of all DWP employment programme expenditure. It aims to fill some of the evidence gap around the impacts of the ESF programme, as proposed by the recent House of Lords Committee Report (2010). The report recommended that '…a robust methodolodgy for assessing effectiveness is key to the short and long term future of the European Social Fund. We conclude that there is substantial room for improvement'.

Therefore, the study is useful for a number of reasons:

- From a DWP perspective, for adding to the overall evidence base by estimating the impacts of a major employment programme, which has a distinctive quality of providing support across client groups;

- It is the first time that an impact analysis of ESF has been performed in the UK;

- It will be useful to the European Commission for providing a rare opportunity for assessing the effectiveness of the ESF relative to a comparison group[8]. It will feed into understanding the 'value for money' of the ESF, worth 76bn Euro across Europe in 2007-13, and inform the EU Budget Review for 2014-20;

- The study is among the first employment evaluations of its kind since the onset of the recent recession. It is therefore of broader interest to a range of policy makers and contributes to the research literature; and

- The study has a number of novel features in its use of data: in particular in its use of detailed benefit and work histories and the method of controlling for participation on other DWP employment programmes.

[8] For details of other impact evaluations across EU Member States on the 2000-2006 Programme, see the EC evaluation 'Study on the Return on ESF Investment in Human Capital' (2010).

We note also three important caveats around the findings of this evaluation:

- As is the case in all non-experimental programme evaluations we can never be absolutely certain that we have fully accounted for the potential bias arising from selection onto the programme. However, the methodology we use, as described in this paper, takes all reasonable steps to minimise this bias;

- The evaluation considers only a part of the ESF programme: employment provision financed by the European Social Fund (ESF) and provided through the Department for Work and Pensions (DWP) between June 2008 and April 2009. Findings of this evaluation should not be presumed to indicate the impacts of the entire 2007-2013 ESF programme; and

- The evaluation looks at net impacts in terms of benefit receipt and employment rates only. Other evaluations, such as the ESF Cohort Survey (Drever and Lloyd, 2010), have found positive effects that lie outside the scope of this analysis. For example, participants experience improved confidence and job readiness and many participants feel that the programme has helped them to gain basic skills and vocational qualifications.

1.2 Introduction to the European Social Fund

The European Social Fund (ESF) was set up to improve employment opportunities in the European Union and so help raise standards of living. Its aim is to help people fulfil their potential by giving them better skills and better job prospects. [9] A key feature of ESF funding is that it must be used to purchase additional provision in order to extend coverage, address gaps and complement domestic funding. The provision itself is varied and flexible, including activities such as job search guidance, basic skills training, case worker support and advice on tackling specific barriers to work.

1.2.1 Funding and Management of the ESF Programme

The current ESF programme for England runs from 2007-2013 and geographically covers England and Gibraltar. The total budget for the programme is £5 billion. This includes £2.5 billion of EU money provided by the European Social Fund, and £2.5 billion of national match funding provided by Co-financing Organisations (CFOs). The overall responsibility for the funds lies with the Managing Authority, which sits within the Department for Work

[9] http://www.dwp.gov.uk/esf/about-esf/

and Pensions and manages the programme at a national level. The Managing Authority delegated some functions to Government Offices[10].

At a regional level, the funds are distributed through Co-financing Organisations[11] (CFOs). The Skills Funding Agency and the DWP are the two largest CFOs (responsible for about 87% of the total fund in England) and bring together ESF and domestic funding for employment and skills provision so that ESF complements domestic programmes. The CFOs contract, through open and competitive tendering, with organisations or 'providers' that deliver ESF projects on the ground.

Of the £2.5 billion budget for the 2007-13 ESF funded part of the programme, £1.6 billion is allocated to Priorities 1 and 4 for helping people who are not in work to enter employment and for providing help to those aged 14-19 who are not in education, employment or training (NEETs), with a strong emphasis on tackling barriers faced by disadvantaged groups. Of the remaining Priorities, £0.9 billion is allocated to Priorities 2 and 5 to address the development of workforce skills, while Priorities 3 and 6 involve 'technical assistance' activities to support programme delivery. The present evaluation estimates the impacts of ESF support for participants who are not in employment when they enter the ESF programme. Impact estimates are therefore only for participants receiving support under Priorities 1 and 4.

At the beginning of the programme the England ESF Managing Authority set a range of indicators and targets for measuring programme performance. For Priorities 1 and 4 these include the proportion of participants who enter employment after leaving the programme, and the proportion of participants who are (at the time of entry onto the programme):

- unemployed;
- economically inactive;
- aged 14-19 and not in education, employment or training (NEET);
- female;
- disabled or have a health condition;
- of an ethnic minority;
- aged 50 years or above;
- lone parents.

1.2.2 DWP ESF funded employment provision

In 2007 the DWP acting in its CFO role (henceforth, reference to DWP will unless otherwise stated refer to its CFO role) contracted £265 million worth of ESF funded employment provision spread across 74 contracts and 26

[10] Regional Government Offices closed at the end of March in 2011 (a long time after the focus of this study) and their ESF responsibilities were handed back to the Managing Authority in DWP.
[11] The financing was different in the 2000-2006 programme: ESF applicant organisations had to supply their own match funding for projects, in a process known as direct bidding.

providers, lasting for three years between 2008 and 2011.[12] Our present impact analysis is focused on participants entering the programme within the first year (between June 2008 and April 2009) of this part of ESF provision. The key features of DWP ESF funded provision are:

- Entry onto the programme is voluntary;

- Support adds value to existing employment programmes, with a particular focus on disadvantaged groups;

- The key incentive for providers is job outcome payments. For most contracts 50 percent of the total funding was linked to the achievement of job outcomes. For 11 contracts this was 40 per cent and for the remaining four contracts it was 70 per cent;

- The contracts were projected to provide 240,000 places, with a job entry rate of 36% to yield 85,000 job entries. This is equivalent to an average cost of £1,100 per start and £3,100 per job outcome;

- There is substantial flexibility in the type of provision that can be offered and provision varies between contracts. The DWP co-financing body categorises the contracts into three broad types of provision:

    - Tailored (a flexible, personalised approach) - 51 contracts, cost £190m;

    - Targeted (contracts in which provision is specified to particular needs – for example helping participants with English language barriers or participants with a disability) - 19 contracts, cost £70m;

    - Intermediate Labour Market (high unit cost contracts for providing subsidised temporary employment with the aim of providing a bridge back to the labour market) – four contracts, cost £5m;

- There is significant variation in projected job entry rates and unit costs across contracts, providers and regions. For example, lower and upper quartile unit costs per job entry across contracts are £2,500 and £5,000 respectively; and

- The original intention was that the majority of referrals would be through direct recruitment by the provider. The ESF cohort survey of participants in 2009 (Drever and Lloyd, 2010) suggests this has not been the case and that about three quarters found out about ESF employment provison via Jobcentre Plus.

In the case of the DWP CFO, ‘matched’ funding is provided for by the New Deal and Pathways contracts. The analysis in this paper is concerned only

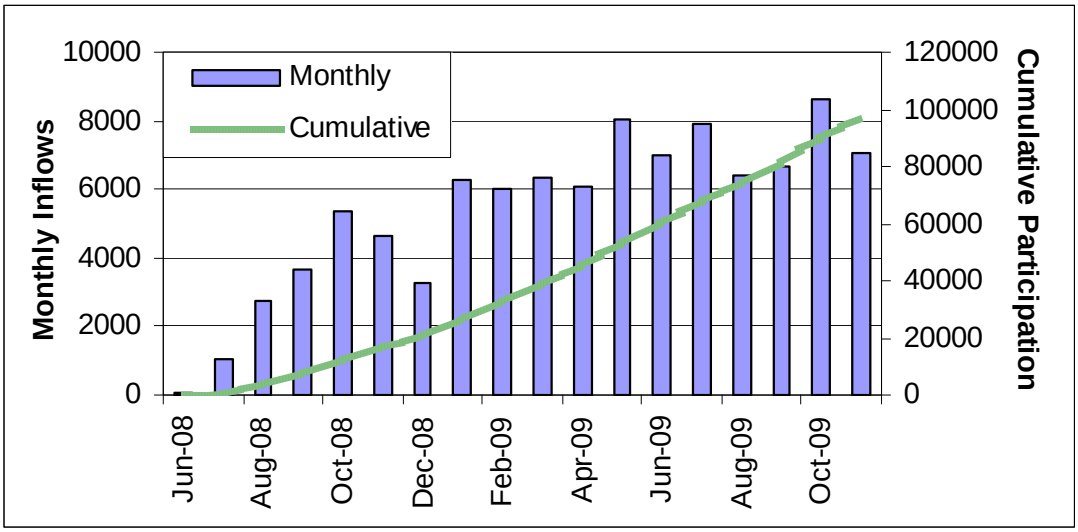
[12] 11 more contracts were contracted in 2008-09 using money released from a changing pound-euro exchange rate, but these started after the period analysed in this paper.

with estimating the impacts of DWP support offered under direct ESF funding and does not include support offered under the matched funding.

1.2.3 Participation on the Current Programme

Participation under the DWP CFO began in June 2008 in the economic context of rising unemployment caused by the recession. Figure 1.1 shows how participant inflows onto DWP ESF funded provision increased to about 7,000 per month by summer 2009, from which point the inflows remained fairly constant. The number of participant starts reached a cumulative total of about 98,000 by the end of November 2009. Of these participants, a higher than expected proportion were receiving Jobseeker's Allowance (particularly short term claimants) when they started on ESF[13]: 68,000 were claiming JSA at the start of their provision of whom 70% had been on benefit for less than six months and almost 50% had been on benefit for less than three months; 6,000 were claiming IB or ESA; 11,000 were claiming Income Support; while the remaining 12,000 were not in receipt of benefit when they started ESF provision. The average course length across all participants was about three and a half months.

Figure 1.1: Monthly inflows and cumulative participation on DWP ESF funded provision



Source: ESF evaluation database; starts between June 2008 and November 2009.

By the end of July 2010, almost all participants who started provision between June 2008 and November 2009 had left provision. Of these participants, 26% had entered employment[14]. Most of these entered employment within a few months of starting provision: of those who entered employment a fifth had

[13] The focus on JSA customers in these contracts was influenced by the recession and directives received after contracts were already in place to use ESF to support JSA customers as much as possible within the contracts' terms of delivery.
[14] Entry to employment as defined here is when the provider received a job outcome payment from DWP.

done so within one month, and over half had done within three months. This pace of entry into employment is not unusual in the context of other employment programmes; the impact study evaluation of the New Deal for Disabled People (NDDP)[15] (Orr et al., 2006) reported that of those who participated on NDDP, a third of those who secured jobs had done so within a month and 55% within three months.

ESF provision is available in all Jobcentre Plus districts in England. However, the take-up of ESF varies across these districts, ranging from a few hundred participants in districts with the lowest take-up to a few thousand in districts with the highest take-up by the end of November 2009. In terms of take-up as a proportion of the benefit caseload, JSA claimants who start on ESF provision represent between 1% and 14% of the average JSA caseload over the same time period for each district. The median and upper quartile take-up proportions are 4% and 7% respectively, which means that there is only a small group of districts that have notably higher take-up than the rest of the country.

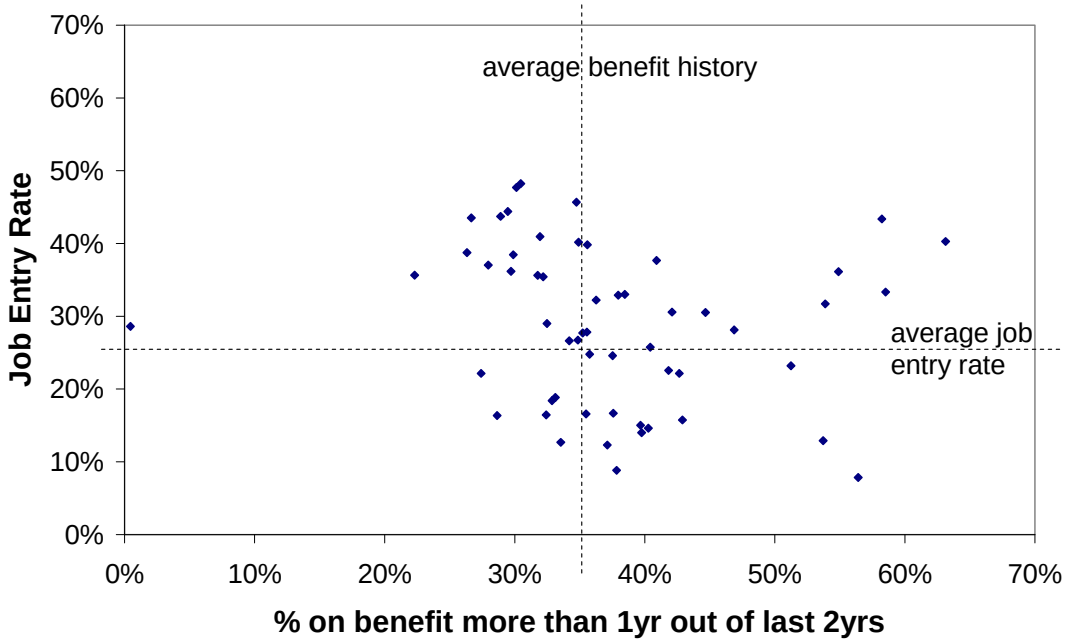
1.2.4 Performance variation across contracts

There is substantial variation across the contracts in terms of both participant characteristics and job entries. Figure 1.2 shows a basic measure of the performance of all large ESF contracts (those with more than 500 participants – 55 contracts). It plots job entry rates of all customers who started provision between June 2008 and the end of November 2009 against the proportion of participants who have spent more than one of the past two years on benefit for each contract. This measure of benefit history is a proxy for the extent to which contracts support disadvantaged participants. The average job entry rate for all participants is 26% and the average proportion of all participants who have spent more than one of the past two years on benefit is 36%. Those towards the top right are considered to be the best performing contracts as they have a high proportion of disadvantaged customers, but are nevertheless achieving high job entry rates.

Figure 1.2 indicates broad variation in contract performance when assessed using this basic measure. This could suggest a high degree of heterogeneity in the effectiveness of ESF provision across the range of support offered by different contracts. We explore this possible explanation in Section 4.3 where we compare impacts achieved by groups of contracts with similar characteristics.

[15] NDDP is a major employment programme for people claiming IB and other disability benefits in the UK. It is similar in size to ESF (in 2005, 5000 participants/month, with a cost of £75m).

Figure 1.2: Contract Performance – Job entry rate by Benefit History



Source: ESF evaluation database; contract performance of starts between June 2008 and November 2009.

2 Data and Sample Definition

This section outlines the data and sample definition used in the impact evaluation. In Section 2.1, we describe the method of drawing samples from which we derive groups of participants and non-participants for comparison. Section 2.2 describes the administrative data sources and variables used in the evaluation. Section 2.3 provides some descriptive statistics for our participant and non-participant samples. Finally, Section 2.4 describes the benefit rates of participants and non-participants within these samples.

2.1 Sample definition

A key feature of ESF employment provision is that it is available to anyone who is not in employment, regardless of their current benefit status or prior interaction with the benefits system. This broad eligibility leads to a greater heterogeneity of participant characteristics than might be expected for other DWP employment programmes where eligibility may be dependent on receipt of particular benefits and/or the duration of a benefit claim.

It is difficult to estimate impacts for a highly heterogeneous group of participants for two main reasons. Firstly, the impacts of the programme are likely to vary across the range of participant characteristics. Secondly, if a group of participants is highly heterogeneous with regard to observed characteristics then it is also likely to be highly heterogeneous with regard to unobserved characteristics. Estimating impacts for a programme with a highly heterogeneous participant group is therefore more likely to result in a biased estimate than for a programme with specific eligibility requirements.

For this reason, when performing our primary analysis we estimate impacts separately for two broad groups of participants:

1. Jobseeker's Allowance (JSA) customers;
2. Incapacity Benefit (IB) or Employment Support Allowance (ESA) customers.

We describe in Sections 2.1.1 and 2.1.2 the method used for drawing samples of JSA and IB/ESA customers for use in this primary analysis.

In Section 4.3, we additionally explore the heterogeneity of impacts of the ESF programme on the primary JSA customer group, by estimating impacts for specific participant sub-groups and different types of employment support offered under the ESF programme.

2.1.1 Defining the ESF participant samples

Both the JSA and the IB/ESA participant samples are drawn from the ESF administrative data set. For inclusion in our primary analysis, participants must meet the following conditions:

- Participants must be claiming the appropriate benefit (JSA or IB/ESA) at the start date of their ESF provision;
- The benefit spell must start after June 2005;[16]
- The ESF start date must be between June 2008 and April 2009;[17]

The resulting participant sample sizes are 25,720 for JSA recipients and 1,970 for IB and ESA recipients.

2.1.2 Defining the non-participant samples

To provide a suitable comparison pool of benefit customers who did not participate on the ESF programme, we also draw two samples (one for JSA customers and one for IB/ESA customers) of non-participants using equivalent selection criteria to those used to draw the participant samples.

To compare the outcomes of participants and non-participants over a time period such that non-participants can represent what would have happened to ESF participants if they had not participated, we assign a pseudo start date to each non-participant. The pseudo start date for non-participants is treated as equivalent to the actual start date for participants. More details of how these dates are generated are provided in Appendix 2.

For inclusion in our primary analysis, non-participants must meet the following conditions:

- The non-participants must be claiming the appropriate benefit (JSA or IB/ESA) at their pseudo start date;
- The benefit spell must start after June 2005;
- The pseudo start date must be between June 2008 and April 2009;

The resulting non-participant sample sizes are 732,600 for JSA customers and 406,430 for IB/ESA customers.

In Section 3 we describe how suitable 'matched' groups of participants and non-participants are selected from these samples and compared to estimate

[16] This excluded a small number of records (2%), in which the benefit start date was prior to June 2005. This decision was taken for pragmatic reasons – extending the analysis to older claims would have resulted in a much larger non-participant sample and a corresponding increase in computational requirements.

[17] This provided a cohort of participants for whom we had a minimum of 52 weeks of outcome data.

the impacts of ESF support. This selection is carried out using a Propensity Score Matching methodology.

2.2 Data sources and variables

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

- *DWP administrative databases*, which provide details of spells on DWP benefits, characteristics of DWP customers and spells on employment programmes including ESF; and

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

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 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 heterogeneity of the programme impacts with regard to participant characteristics and types of support (see Section 4.3); and

- The cost and time associated with the analysis are substantially less when using only administrative data.

2.2.1 Description of Variables

We outline below in Table 2.1 the variables used in the analysis. We discuss the importance of these variables in controlling for selection onto the programme in Section 3.

Table 2.1: Variables and values used in the analysis

Variable	Type	Values
Gender	Categorical	*Male; Female*
Age, (and Aged squared)[18]	Numerical	Integer values
Disability	Categorical	*Not disabled; Disabled; Unknown*
Ethnicity	Categorical	*White; Black; Asian; Mixed; Chinese; Other; Unknown*
Marital status	Categorical	*Single; Married; Widowed; Divorced; Separated; Cohabiting; Unknown*
Sought occupation	Categorical	26 broad categories: e.g. *"Administrative"; "Health Professionals"; "Sales Occupations"*
Government Office Region	Categorical	9 regions in England
Lone Parent	Categorical	*Lone Parent; Couple; Not a Parent; Unknown*
Low Qualified	Categorical	*No; Yes; Unknown*
Jobcentre Plus District	Categorical	38 districts in England; and *Unknown*
Benefit start month[19]	Categorical	47 months from June 2005 to April 2009
ESF start month	Categorical	11 months from June 2008 to April 2009
Indices of multiple deprivation[20]: - Income - Employment - Health - Education - Housing - Crime - Living environment	Numerical	Continuous values, available from Government data website[21]
Benefit labour market history	Categorical	104 binary variables – one representing each of the 104 weeks prior to ESF start date. Values are: *receiving benefit; not receiving benefit*
Work labour market history	Categorical	104 binary variables – one representing each of the 104 weeks prior to ESF start date. Values are: *in work; not in work*

[18] Age squared is included as the literature extensively shows a non linear relationship between employment and age.
[19] Benefit start and end dates refer to the benefit spell during which the ESF support is received.
[20] The Indices of Multiple Deprivation are variables which describe local deprivation in each of the 32,482 super output areas across the country. They cover a range of economic, social and housing issues.
[21] http://data.gov.uk/dataset/index_of_multiple_deprivation_imd_2007

JSA Programme[22] History	Numerical	Number of weeks on JSA programmes in two years prior to ESF start date: *Integers from 0 to 104*
Other Programme[23] History	Numerical	Number of weeks on other programmes in two years prior to ESF start date: *Integers from 0 to 104*

To capture the labour market history of individuals, we use two series of binary variables which indicate whether each person was in/out of work and receiving/not receiving benefit in each of 104 weeks prior to their ESF start date. In our analysis of JSA customers, we use individuals' JSA receipt over this period, while in our analysis of IB and ESA customers, we use individuals' IB and ESA receipt. In Appendix 1, we compare the advantages and disadvantages of this approach with those of a more commonly used method in the literature and adopted by, for example, Card and Sullivan (1988).

In our analysis of IB/ESA customers, we additionally include the variables outlined in Table 2.2, which provide more information about an individual's benefit claim and the nature of their health condition and/or disability.

Table 2.2: Variables and values additionally used in analysis of IB/ESA recipients

Variable	**Type**	**Values**
Benefit Type	Categorical	*Incapacity Benefit; Employment Support Allowance*
Health Condition	Categorical	*Circulatory/Respiratory; Injury/Poisoning; Mental Health Condition; Musculo-skeletal; Nervous System; Other; None recorded*
Disability Living Allowance (DLA) recipient	Categorical	*Receiving DLA; Not receiving DLA* (at any time since June 2005)
Mental Health Group	Categorical	*Alcohol/Drug abuse; Mood disorder; Stress; Mental Development disorder; Other; None recorded*
DLA care component	Categorical	*Higher; Lower; Medium; None; Unknown*
DLA mobility component	Categorical	*Higher; Lower; None; Unknown*

The *health condition* and *Mental Health Group* variables are derived from GP certificates submitted as part of the benefit claim.

[22] JSA programmes include: New Deal for Young People (NDYP), New Deal for the Long Term Unemployed (NDLTU), Basic Skills and Work Based Learning for Adults (WBLA).
[23] Other programmes include: Pathways, New Deal for Lone Parents (NDLP), New Deal 50 Plus, New Deal Innovation Fund (NDIF), New Deal for Disabled Peple (NDDP), Employment Zones (EZ), Action Teams, Outreach and New Deal for Partners (NDP).

2.2.2 Data quality issues

HMRC employment data

There are a number of well documented issues with the quality of the HMRC employment data. 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 income tax Personal Allowance for the relevant year, are therefore not recorded;

2. Any employment spells which are known to have started in a particular tax year, but on an unknown date during that year, are automatically given a start date of 6th April. This is the earliest date on which they could actually have started. Similarly, any employment spells which are known to have ended in a particular tax year, but at an unknown point during that year, are automatically given an end date of 5th April. This is the latest date on which they could actually have ended. The net effect of this is that the length of many employment spells will be overestimated; and

3. A small number of records contain other known errors, such as missing start dates or missing end dates.

Within our sample, approximately 35% of employment spells have a suspected error, as defined above. This proportion is the same among participants and non-participants. Moreover, the proportion of each type of error identified is also the same among participants and non-participants. Therefore we would not expect any systematic bias to result from these errors. 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 appropriate tax year for records in which they are unknown.

Appendix 6 discusses an alternative method of cleaning the HMRC employment data which we explored in our model development stage.

DWP administrative data

The DWP administrative data sets also contain missing values because advisers do not always fill in some characteristic fields during client interviews. This is particularly the case for variables identifying lone parents, ethnicity and disability. In the case of variables with missing values, 'unknown' is treated as a valid category for controlling for participant characteristics.

The proportion of missing values among JSA customers is around 9% for *ethnicity*, 7% for *occupational choice*, 2% for *marital status*, 2% for *disability,* 2% for *district,* 52% for *lone parent* and 75% for *low qualified*. The proportion

of missing values for each variable is broadly similar among ESF participants and non-participants.

The proportion of missing values among IB/ESA customers is higher and there are greater differences between participants and non-participants. Among IB/ESA customers for example, 11% of participants have unknown *ethnicity* compared with 24% of non-participants; 31% of participants have unknown *occupational choice* compared with 47% of non-participants; 28% of participants have unknown *marital status* compared with 38% of non-participants; and 12% of participants have unknown *(health) condition* compared with 5% of non-participants.

2.3 Descriptive Statistics

In this section we describe the basic characteristics of ESF participants and non-participants in the JSA and IB/ESA samples. Table 2.3 lists summary statistics detailing personal and demographic characteristics, benefit receipt, and participation on DWP employment programmes other than ESF. This table includes only a few of the most pertinent characteristics – for a full list of variables and their value ranges, refer back to Tables 2.1 and 2.2. We note below some of the key differences between participants and non-participants within each of the JSA and IB/ESA samples. We also describe the key differences between the JSA and IB/ESA non-participant samples to gain an understanding of the overall differences between customers receiving each of these benefits.

Table 2.3: Characteristics of the JSA and IB/ESA primary samples

	JSA Participants	JSA Non-Participants	IB/ESA Participants	IB/ESA Non-Participants
Observations	25,720	732,610	1,970	406,430
	Personal / Demographic Characteristics			
Age (mean years)	37	34	40	43
Male (%)	80	73	67	56
Disabled (%)	28	21	73	67
Ethnic minority (%)	17	16	8	10
Low Qualified (%)	18	20	20	10
Lone Parent (%)	2	3	5	8
Deprivation score above 50 (%)	21	17	16	16
Single (%)	59	63	42	35
Married (%)	19	17	15	13
Seeking Professional / Management job (%)	7	12	8	5
	Benefit Receipt			
Receiving JSA at ESF start (%)	100	100	0	0
Receiving IB at ESF start (%)	0	0	83	91
Receiving ESA at ESF start (%)	0	0	17	9
Proportion of past year spent receiving JSA (%)	45	41	8	6
Proportion of past year spent receiving IB /ESA (%)	4	3	62	73
Proportion of past year spent in employment (%)	44	42	43	31
Benefit Duration at start of programme (mean weeks)	20	22	47	66
	DWP Programme Participation			
Proportion of past year spent on JSA programmes (%)	3	2	1	1
Proportion of past year spent on other programmes (%)	5	4	45	28
	Additional characteristics known only for IB/ESA customers			
Mental Health Condition (%)	-	-	45	39
Receives IB Credit only (%)	-	-	16	46
Received Disability Living Allowance (DLA) within 2 years (%)	-	-	29	36
High DLA care component (%)	-	-	3	7
High DLA mobility component (%)	-	-	7	17

Notes:

Deprivation scores range from 0 to 100. Higher deprivation scores are associated with more deprived areas, so a higher proportion with scores over 50 is indicative of more deprivation among the participants within the sample.

2.3.1 JSA customers: comparing ESF participants with non-participants

The mean age of an ESF participant in receipt of JSA is 37 years, compared with 34 years for a non-participant. Fewer participants (20%) than non-participants (27%) are female and there is a higher rate of disability among participants (28%) than non-participants (21%).

There are similar proportions of ethnic minorities, low qualifications and lone parents among participants and non-participants. More participants (21%) than non-participants (17%) live in highly deprived areas. A lower proportion

of participants (59%) than non-participants (63%) are single. A lower proportion of participants (7%) are looking for professional or management jobs than non-participants (12%).

Participants tend to have spent more of the past year receiving JSA (45%) than non-participants (41%) and more of the past year on programmes other than ESF than non-participants.

Overall, these statistics suggest that JSA customers who participate on ESF are *more disadvantaged* than JSA customers who do not participate.

2.3.2 IB/ESA customers: comparing ESF participants with non-participants

The mean age of an ESF participant in receipt of IB or ESA is 40 years, compared with 43 years for IB and ESA non-participants. Fewer participants (33%) than non-participants (44%) are female and there is a higher rate of disability[24] among participants (73%) than non-participants (67%).

There is a slightly lower proportion of ethnic minorities among participants (8%) than non-participants (10%). More participants have low qualifications (20%) and fewer are lone parents (5%) than non-participants (10% and 8% respectively). A similar proportion of participants (16%) and non-participants (16%) live in highly deprived areas. A higher proportion of participants (42%) than non-participants (35%) are single and a higher proportion of participants (8%) are looking for professional or management jobs than non-participants (5%).

Participants tend to have spent less of the past year receiving IB or ESA (62%) than non-participants (73%) and more (45%) of the past year on programmes other than ESF than non-participants (28%). Participants are also more likely to be in receipt of ESA (17%) than non-participants (9%) and correspondingly less likely to be in receipt of IB (83%) than non-participants (91%).

A higher proportion of participants (45%) than non-participants (39%) have a mental health condition. A much lower proportion of participants (16%) than non-participants (46%) are in receipt of Incapacity Benefit credits only, indicating less employment history among non-participants.

Overall, these statistics suggest that IB/ESA customers who participate on ESF are *less disadvantaged* than IB/ESA customers who do not participate.

[24] This is the proportion with a disability marker on the DWP Labour Market System. Other studies suggest that actual proportions with disability may be higher.

2.3.3 Comparing the JSA and IB/ESA non-participant samples

Non-participants in receipt of IB or ESA tend to be much older (with a mean of 43 years) than non-participants in receipt of JSA (with a mean of 34 years). A higher proportion of the JSA sample is male (73%) than the IB/ESA sample (56%). Disabled customers make up a much higher proportion of the IB/ESA sample (67%) than the JSA sample (21%). There are fewer ethnic minorities among the IB/ESA sample (10%) than the JSA sample (16%).

There are more low qualified (20%) and fewer lone parents in the JSA sample (3%) than the IB/ESA sample (10% and 8% respectively). A similar proportion of the JSA and IB/ESA samples live in areas of high deprivation (17% and 16% respectively). A higher proportion of the JSA sample are single (63%) than the IB/ESA sample (35%) and more are seeking professional or management jobs (12%) than the IB/ESA sample (5%).

IB and ESA recipients tend to have spent more of the past year receiving benefits, less in employment and more on employment programmes other than ESF than JSA recipients.

2.4 Comparing benefit rates of participants and non-participants

Figures 2.1 and 2.2 show the benefit rates of participants and non-participants in our primary analysis samples of JSA customers and IB/ESA customers. 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 the ESF programme, but rather a combination of programme impacts and differences in the characteristics of those who participate and those who do not.

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

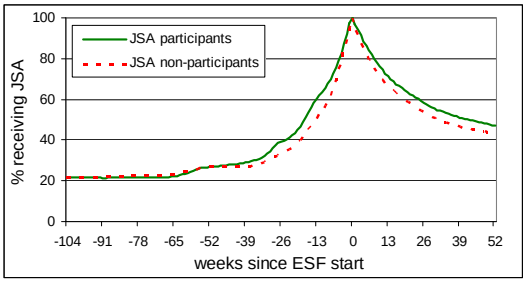
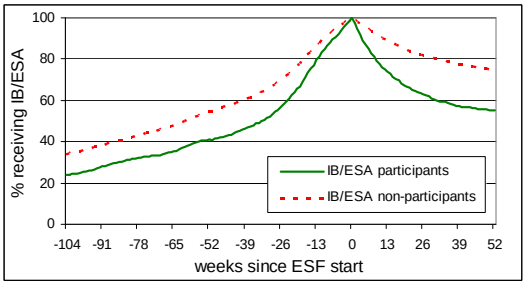


Figure 2.2: IB/ESA receipt rate among participants and non-participants



For JSA customers, we note that ESF participants and non-participants have similar rates of benefit receipt until around 40 weeks before the ESF programme start (or pseudo start for non-participants) date. Over the 40 week period leading up to the ESF start date, there is a higher rate of benefit receipt among those who go on to participate compared with those who do not.

Similarly, in the 52 weeks after the start date, there is a higher rate of benefit receipt (typically four percentage points higher) among participants than non-participants.

For IB/ESA customers a very different situation is observed. There is a much lower rate of benefit receipt among participants than non-participants over the full 104 weeks prior to the ESF programme start (or pseudo start) date. Similarly, in the 52 weeks after the start date, there is a lower rate of benefit receipt among participants than non-participants. The difference between the participant and non-participant groups increases more rapidly after the ESF start date, such that 52 weeks after the programme start, the difference (-20 percentage points) is greater than at any time in the pre-programme period (a maximum difference of -15 percentage points).

Overall, these observations suggest that among JSA customers, participants tend to be *further from the labour market* than those who do not participate. Conversely, among IB/ESA customers, participants tend to be *closer to the labour market* than non-participants.

3 Methodology

In this section, we outline the methodology used in this study to estimate the average effect of the ESF employment programme on its participants (*average effect of treatment on the treated, ATT*). In Section 3.1 we describe the Conditional Independence Assumption, which forms the foundation of impact evaluations of this type. We describe in Section 3.2 the Propensity Score Matching (PSM) methodology used to control for selection bias and construct suitable counterfactuals for our ESF participant groups. 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 ESF employment support. Finally in Section 3.4, we briefly outline an extensive sensitivity analysis conducted during our model development stage to establish the robustness of our methodology.

3.1 Conditional Independence Assumption

Our aim is to estimate the average effect of the ESF employment programme on its participants (*average effect of treatment on the treated, ATT*). A simple comparison between the employment and benefit outcomes of participants with those of non-participants could be biased if there are systematic differences between these groups which may be related to the labour market outcomes of interest (see Table 2.3). However, if we can control for observable differences in characteristics between the participant and non-participant groups, the outcome that would result in the absence of treatment is the same in both cases (Bryson, et al. 2002). This identifying assumption for matching is known as the Conditional Independence Assumption (CIA). It enables us to infer the counterfactual outcome for ESF participants, and therefore to attribute any differences between carefully matched participant and non-participant groups to the effect of the programme. We control for characteristics using a Propensity Score Matching methodology, as described in Section 3.2.

The following section discusses how our individual level data (as described in Section 2) allows us to try to control for the difference in typical characteristics between those who participate on the ESF programme and those who do not. Differences in characteristics between these two groups arise as a result of the way participants are selected onto the programme, and are therefore known as *selection bias*.

3.1.1 Controlling for selection bias

The majority of ESF participants are referred onto the programme by Jobcentre Plus with a minority recruited by providers. Therefore selection onto the programme is determined by two main interconnecting factors:

1. Selection decisions made by Jobcentre Plus advisors; and
2. Self-selection by the individual.

In deciding whether to recommend ESF provision to a potential participant, a Jobcentre Plus advisor is likely to consider whether the person needs additional help to move away from benefit and into employment and the extent to which the ESF programme is likely to meet these needs.[25] It will ultimately be a judgment by the adviser but is likely to depend on the demographic characteristics of the participant, the type of job sought, the perceived motivation of the participant, the availability of ESF in the local area, the familiarity of the particular Jobcentre Plus advisor with ESF, and the availability of other employment programmes in the area. The factors which influence a potential participant's self-selection onto the programme are similar to those which influence the advisor - it will largely depend on whether they feel that ESF will improve their labour market prospects and the extent to which they actually want to improve their labour market prospects. These factors will be driven by demographic characteristics, motivation and circumstance. Since ESF is a voluntary programme, the final participation decision will be a mutual decision reached after a discussion between the participant and the advisor. [26]

To allow us to control for the selection bias which results from the complex interaction between advisor and participant, we have collected demographic characteristics such as age, gender, ethnic group, disability, qualification, marital status and lone parent status (see Section 2.2) for a large number of participants and non-participants. We have also obtained each individual's stated primary occupation choice. Since labour market prospects may be highly dependent on dynamic local labour market characteristics and the local availability of employment provision, we have also collected each individual's geographical district and their Index of Multiple Deprivation (IMD) scores[27] at a local super output area level.

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

[25] In the smaller number of cases where participants are directly recruited onto the programme by providers, we expect similar criteria to be considered in the selection process.
[26] The criteria for selection suggested in this section are broadly supported by ongoing qualitative research into Priority 1 and 4 provision, which is expected to be published in the summer of 2011.
[27] Separate IMD scores are included for: employment, Income, Health, Education, Housing, Crime and Living environment.

variables we have not observed. For example, while we have not been able to observe personality type, life experience, experience of discrimination, confidence, health, language skills or happiness, we believe that by controlling for gender, age, ethnicity, disability, local deprivation, labour market history, prior programme participation etc. the model will capture some of their influence by proxy.

3.2 Propensity Score Matching

The aim of the matching process is to construct a comparison group of individuals who did not participate on the ESF programme, but who have identical characteristics to those who did participate. If this is successfully achieved, we can then use the labour market outcomes of non-participants in the comparison group as an approximation for 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 becomes large compared to the number of observations (Rosenbaum and Rubin, 1983). Therefore, we follow the literature in using a single balancing score on which to match, which is a function of all the observed covariates. The balancing score we use is a propensity score, which is the probability of an individual participating on the programme given their observed characteristics. We describe below the Propensity Score Matching (PSM) protocol used in this evaluation to construct suitable treatment and comparison groups from our participant and non-participant samples. Appendix 3 shows a step-by-step guide to the protocol.

For each participant and non-participant in the sample, a 'propensity score' is calculated. This defines the probability of the individual participating on the ESF programme, given their observed characteristics. The score is calculated using a probit regression model, where the binary participation variable is the dependent variable, and the observed individual characteristics are the independent variables. In the probit model used for our analysis of JSA recipients, all variables except *low qualified* were significant at the 1% level. In the probit model used for our analysis of IB/ESA recipients, all variables except *DLA Care Component, ESF start date* and *marital status* were significant at the 5% level.

We construct matched treatment and comparison groups using a Kernel 'one-to-many' PSM approach. For each participant in the sample, all non-participants with propensity scores within the Kernel bandwidth are selected and weighted using an Epanechnikov distribution. The result is a matched comparison group of non-participants with a total weighting equal to the number of participants in the final treatment group. 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).

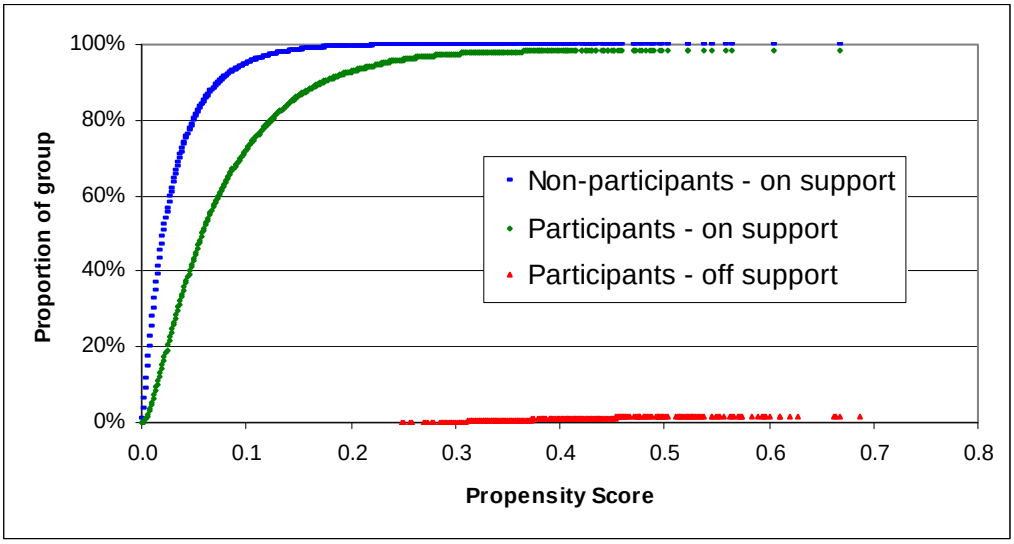
3.2.1 Common Support for Participants

For Propensity Score Matching to be a viable methodology, 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 bandwidth used in the Kernel matching 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. After testing the sensitivity of our impact estimates and common support to a range of bandwidths in the model development stage (see Appendix 6), we chose a bandwidth of 0.0001 for our primary analysis. This was felt to provide a model which retained a high level of common support, whilst also ensuring a tight match between non-participants and participants.

The propensity score distribution for our primary JSA estimate is given below; each point is the cumulative proportion of the group with propensity scores in increments of 0.001.

Figure 3.1: Primary JSA Analysis Propensity Score Distribution:
(25,720 treatment records; 732,610 comparison records)



The graph shows that propensity scores tend to be higher among participants, although the mean score for both groups is low. The mean propensity score is 0.03 (the proportion of all records which receive ESF support); 34% of non-participants have a score of more than 0.03 compared with 68% of participants; also 4% of non-participants have scores more than 0.1 compared with 25% of participants. The small proportion of participants for whom no common support is available are those with the highest propensity scores.

The propensity score distribution for JSA recipients provides sufficient overlap between participants and non-participants, with over 98% of treatment records finding common support.

The propensity score distribution for our primary IB and ESA estimate is provided in Appendix 4. We found a similar distribution of scores among participants and non-participants to that found among JSA recipients. As with the JSA scenario, our propensity score distribution for IB and ESA recipients provides sufficient overlap between participants and non-participants, with almost 99% of participants finding common support.

3.2.2 Matching Quality

We found our propensity score model to be highly effective in constructing treatment and comparison groups that are well balanced on the observed characteristics. Tables 3.1 and 3.2 show specification statistics for the matching of the JSA and IB/ESA groups. The chi squared test shows that prior to the match, there was approximately zero probability that the treatment and comparison samples had the same set of characteristics. After matching there is statistically no difference between the matched groups in terms of observed variables.

Table 3.1: Specification statistics for the JSA group matching

Sample	Pseudo R-sq	LR chi-sq	p>chi-sq
Unmatched	0.12	25,780	0.000
Matched	0.00	100	1.000

Table 3.2: Specification statistics for the IB/ESA group matching

Sample	Pseudo R-sq	LR chi-sq	p>chi-sq
Unmatched	0.14	3,460	0.000
Matched	0.00	20	1.000

Table 3.3 below shows the unmatched and matched means of a few select variables for the JSA treatment and comparison groups, while Table 3.4 shows unmatched and matched means for the IB/ESA treatment and comparison groups. It can be seen that the PSM methodology has been extremely effective in balancing the groups on the observed covariates.

Table 3.3 shows, for example, that the mean age of ESF participants in receipt of JSA is 37, while the mean age of the JSA non-participants is 34. Our Propensity Score Matching method produces matched treatment and comparison groups from the JSA samples, which both have a mean age of 34. A similarly high quality of matching is observed across almost all variables within both our JSA and IB/ESA groups.

Table 3.3: Unmatched and matched means for primary JSA analysis

Variable	Sample	Treatment	Comparison	% bias	% reduction in |bias|	t	p>|t|
Age (mean years)	Unmatched	37	34	22		34.29	0.00
	Matched	37	37	-1	96	-1.07	0.28
Female (%)	Unmatched	20	27	-17		-25.52	0.00
	Matched	20	21	-1	97	-0.66	0.51
Disabled (%)	Unmatched	28	21	16		27.02	0.00
	Matched	28	29	-3	85	-2.70	0.01
White ethnicity (%)	Unmatched	74	77	-6		-9.35	0.00
	Matched	74	74	0	99	-0.06	0.96
Black ethnicity (%)	Unmatched	7	6	4		6.65	0.00
	Matched	7	7	0	94	0.28	0.78
Married (%)	Unmatched	19	17	5		7.72	0.00
	Matched	19	18	1	87	0.70	0.48
On JSA 26 weeks before ESF start date (%)	Unmatched	39	33	12		19.35	0.00
	Matched	38	38	-1	96	-0.59	0.56
In work 26 weeks before ESF start date (%)	Unmatched	47	44	4		7.01	0.00
	Matched	47	46	1	81	0.94	0.35
Occupation choice - Corporate Managers (%)	Unmatched	1	3	-11		-14.52	0.00
	Matched	1	1	0	100	-0.05	0.96
ESF start date in August 2008 (%)	Unmatched	5	8	-11		-15.92	0.00
	Matched	6	6	-2	84	-2.17	0.03
Benefit Start Date in November 2008 (%)	Unmatched	10	7	8		13.56	0.00
	Matched	10	9	0	95	0.45	0.65
Region - North East (%)	Unmatched	9	7	6		10.67	0.00
	Matched	9	8	1	87	0.90	0.37
IMD income score (mean)	Unmatched	24	23	14		21.32	0.00
	Matched	24	24	0	99	-0.21	0.83
IMD employment score (mean)	Unmatched	15	14	17		26.46	0.00
	Matched	15	15	0	100	-0.09	0.93
JSA programme history (mean weeks)	Unmatched	3	2	4		7.51	0.00
	Matched	3	3	-5	-8	-4.82	0.00
Other programme history (mean weeks)	Unmatched	5	4	8		14.09	0.00
	Matched	5	5	-1	90	-0.81	0.42
District - Bedfordshire & Hertfordshire (%)	Unmatched	1	2	-10		-13.05	0.00
	Matched	1	1	0	99	-0.14	0.89
Lone Parent (%)	Unmatched	2	3	-6		-9.39	0.00
	Matched	2	2	0	96	-0.35	0.73
Low Qualified (%)	Unmatched	18	20	-4		-6.83	0.00
	Matched	18	18	0	95	-0.25	0.80

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 3.4: Unmatched and matched means for primary IB/ESA analysis

Variable	Sample	Treatment	Comparison	% bias	% reduction in |bias|	t	p>|t|
Age (mean years)	Unmatched	40	43	-22		-8.96	0.00
	Matched	40	40	1	96	0.27	0.79
Female (%)	Unmatched	33	44	-21		-8.98	0.00
	Matched	34	34	0	100	0.02	0.98
Disabled (%)	Unmatched	73	67	13		5.79	0.00
	Matched	73	73	0	98	0.08	0.94
White ethnicity (%)	Unmatched	81	66	35		14.23	0.00
	Matched	81	81	0	100	0.06	0.96
Married (%)	Unmatched	15	13	4		1.89	0.06
	Matched	15	15	0	94	0.08	0.93
On IB/ESA 26 weeks before ESF start date (%)	Unmatched	57	70	-27		-12.59	0.00
	Matched	57	57	0	99	0.09	0.93
In work 26 weeks before ESF start date (%)	Unmatched	45	32	27		12.33	0.00
	Matched	44	44	1	97	0.24	0.81
Occupation choice - Corporate Managers (%)	Unmatched	1	1	1		0.24	0.81
	Matched	1	1	0	28	-0.12	0.91
ESF start date in August 2008 (%)	Unmatched	6	8	-9		-3.81	0.00
	Matched	6	6	0	98	-0.06	0.95
Benefit Start Date in November 2008 (%)	Unmatched	7	3	16		8.72	0.00
	Matched	7	7	0	100	0.00	1.00
Region - North East (%)	Unmatched	9	6	8		3.66	0.00
	Matched	9	8	2	79	0.48	0.63
IMD employment score (mean)	Unmatched	14	14	3		1.29	0.20
	Matched	14	14	0	85	0.14	0.89
JSA programme history (mean weeks)	Unmatched	1	1	5		2.81	0.01
	Matched	1	1	-1	91	-0.13	0.89
Other programme history (mean weeks)	Unmatched	34	22	35		15.91	0.00
	Matched	33	33	1	97	0.28	0.78
District - Bedfordshire & Hertfordshire (%)	Unmatched	2	2	-3		-1.43	0.15
	Matched	2	2	0	88	0.13	0.89
Lone Parent (%)	Unmatched	5	8	-10		-3.91	0.00
	Matched	6	6	-1	90	-0.33	0.74
Low Qualified (%)	Unmatched	20	10	30		15.59	0.00
	Matched	20	20	0	99	-0.10	0.92
Employment Support Allowance Customer (%)	Unmatched	17	9	26		13.29	0.00
	Matched	17	17	-1	97	-0.22	0.82
Incapacity Benefit Customer (%)	Unmatched	83	91	-26		-13.29	0.00
	Matched	83	83	1	97	0.22	0.82
Circulatory or Respiratory Health Condition (%)	Unmatched	3	7	-15		-5.81	0.00
	Matched	3	3	0	97	0.15	0.88
Alcohol or Drug Abuse Health Condition (%)	Unmatched	55	61	-12		-5.27	0.00
	Matched	55	54	1	93	0.26	0.79
Receiving Disability Living Allowance (%)	Unmatched	29	36	-15		-6.31	0.00
	Matched	29	29	-1	96	-0.20	0.84

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).

For the JSA group, the matching results in a reduction in bias for 463 out of the 585 variables used in the propensity score model; while for the IB/ESA group, the matching results in a reduction in bias for all 606 variables used in the propensity score model. Within the JSA group, it is mostly the variables representing JSA receipt between 45 and 104 weeks before participation which do not incur any reduction in bias as a result of the matching. This simply reflects the fact that the treatment group and comparison population are actually very similar with regard to these variables even before matching.

3.3 Measuring Outcomes and Impacts

To estimate the average effect of the ESF employment programme on its participants (the *average effect of treatment on the treated, ATT*), we have measured three sets of outcome data for each participant. In each of the 52 weeks following the start of the ESF programme, we measure whether the individual was:

1. in receipt of their primary benefit[28];
2. in receipt of any main working age DWP benefit[29]; and
3. in employment.

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. 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 ESF/pseudo start date to 52 weeks later.

By comparing the outcome data of the matched treatment and comparison groups, we are able to determine an estimate of the impact of the ESF 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 their primary benefit, the likelihood of participants claiming any main working age benefit, and the likelihood of 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.

[28] For our JSA recipient group this is Jobseeker's Allowance; for our IB/ESA recipient group it is either of Incapacity Benefit or Employment Support Allowance.
[29] This refers to any of JSA, IB, ESA or IS.

However, it is possible that there is some residual bias which has not been fully captured by the variables used in our propensity score model. Therefore, we adjust our impact measure with a difference-in-difference adjustment between the pre- and post-programme periods to reduce any remaining bias.

We calculate the difference-in-difference adjusted impacts using a linear probability model.[30] The difference-in-differences adjustment to the impact uses the mean percentage point difference over the period from 104 weeks to 52 weeks before the programme start date. This period was chosen as it is likely to be unaffected by any pre-programme anticipation effects, which may be responsible for an increase in benefit receipt among participants in the run up to the programme.

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

3.4 Sensitivity to methods

During the development stages of our analysis, we explored the sensitivity of our impact estimates to the precise implementation of the PSM methodology. We performed the following sensitivity tests on subsets of our final primary JSA group:

- sensitivity to time-based variables;
- sensitivity to alternative pseudo start generation methods;
- sensitivity to the Kernel bandwidth; and
- sensitivity to alternative methods of cleaning the HMRC employment data;

We found that our impact estimates were largely insensitive to each of these variations. This provided us with increased confidence that our methodology is robust and that our findings are not biased by the precise specification of our model. Further details of these sensitivity tests are provided in Appendix 6.

[30] This produces consistent estimates for the difference in difference impacts but 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).

4 Results

In this section, we present our estimates of the average effect of the ESF employment programme on participants (the average effect of treatment on the treated, ATT). As discussed in Section 2.1, we have performed separate primary analyses for two main groups of participants:

1. Jobseeker's Allowance (JSA) customers; and
2. Incapacity Benefit (IB) or Employment Support Allowance (ESA) customers.

In Section 4.1 we present our primary impact estimates for these participant groups. Section 4.2 then discusses a number of possible explanations for our findings. In Section 4.3 we explore the heterogeneity of the impacts of the ESF programme, with regard to its impact on specific sub-groups and also the heterogeneity of support provided.

4.1 Primary Estimates for JSA and IB/ESA groups

We present below our primary estimates for the average treatment effect of the ESF employment programme on our two main analysis groups: JSA customers and IB/ESA customers. Figures 4.1 to 4.6 show the net impacts of provision on each group with regard to receipt of primary benefit, receipt of any main working age DWP benefit and employment status over a 52 week period following ESF participation. We extend the impact graphs over a 104 week period prior to ESF participation to illustrate the extent to which labour market history is controlled for over the pre-programme period. The graphs also show a 95% confidence interval around the central impact estimates.

Figure 4.1: Impact on likelihood of claiming main benefit (JSA group)

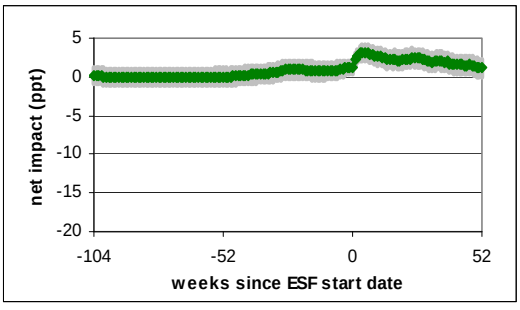


Figure 4.2: Impact on likelihood of claiming main benefit (IB/ESA group)

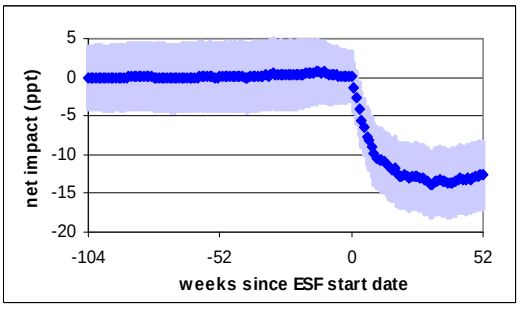


Figure 4.3: Impact on likelihood of claiming any working age benefit (JSA group)

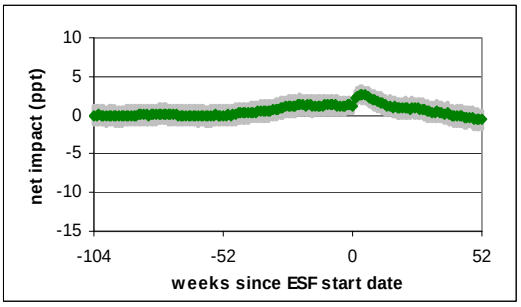


Figure 4.4: Impact on likelihood of claiming any working age benefit (IB/ESA group)

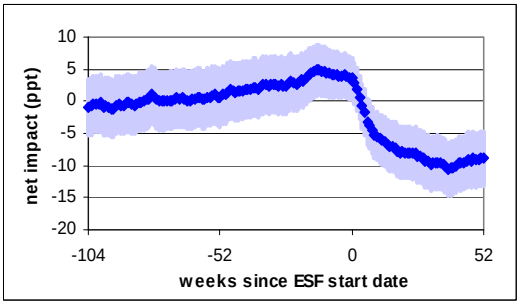


Figure 4.5: Impact on likelihood of being in employment (JSA group)

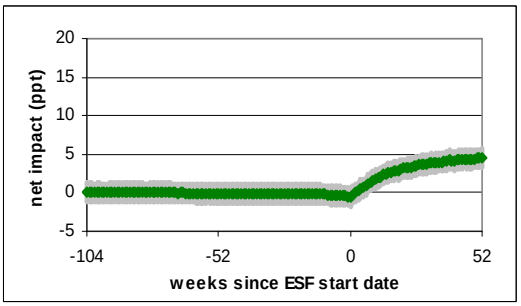
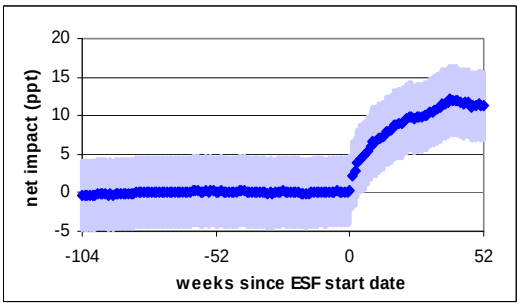


Figure 4.6: Impact on likelihood of being in employment (IB/ESA group)



4.1.1 Net impacts on the JSA customer group

As shown in Figure 4.1, the **impact of the programme on JSA receipt** for the JSA customer group is positive (between +1 percentage point and +3 percentage points) in each of the 52 weeks following participation, which suggests that participation slightly ***increases an individual's chances of claiming JSA over this period***. This effect is statistically significant[31] for most of the 52 week period.

Figure 4.3 shows that the **impact of the programme on the receipt of any main working age benefit** for this group is positive for the first 15 weeks after participation (between +1 percentage points and +3 percentage points), indicating that participation slightly ***increases an individual's chances of claiming benefit over this period*.** Beyond 15 weeks after the start of participation the impact is not significantly different from zero.

As shown in Figure 4.5, the **impact of the programme on employment** for this group is positive in each of the 52 weeks following participation, rising from 0 percentage points to +4.5 percentage points towards the end of this period. This indicates that participation ***increases an individual's chances***

[31] Unless otherwise specified, we report results as statistically significant if they are significant at the 5% level.

of being in employment over this period. This effect increases over time and is statistically significant for almost all of the 52 week period.

There is no contradiction in the finding that both the benefit and employment impact estimates are positive for the first 15 weeks following participation. This is because the impacts are not mutually exclusive and do not account for all possible outcomes. In Appendix 5 we show that the positive benefit and employment impacts are balanced by a negative impact on the labour market position 'neither receiving benefit nor in employment'.

4.1.2 Net impacts on the IB/ESA customer group

As shown in Figure 4.2, the **impact of the programme on IB/ESA receipt** for the IB/ESA customer group is negative in each of the 52 weeks following participation (reaching a minimum of nearly -14 percentage points and then declining to -13 percentage points at 52 weeks), suggesting that participation substantially ***decreases an individual's chances of claiming IB/ESA over this period***. This effect is statistically significant for almost all of the 52 week period.

Figure 4.4 shows that the **impact of the programme on the receipt of any main working age benefit** for this group is also negative in each of the 52 weeks following participation (reaching a minimum of almost -11 percentage points and then declining to -9 percentage points at 52 weeks), suggesting that participation substantially ***decreases an individual's chances of claiming benefit over this period***. This effect is also statistically significant for most of the 52 week period.

As shown in Figure 4.6, the **impact of the programme on employment** for this group is positive in each of the 52 weeks following participation (peaking at +12 percentage points and then declining to +11 percentage points at 52 weeks), suggesting that participation substantially ***increases an individual's chances of being in employment over this period***. This effect is statistically significant for almost all of the 52 week period.

4.2 Discussion of Primary Estimates

It is clear from the primary impact estimates that the effects of the ESF programme are not homogeneous across the two participant groups analysed (JSA customers and IB/ESA customers). The effect of ESF provision on the JSA group is fairly small: the impact on benefit receipt is close to zero, while the impact on employment reaches just +4.5 percentage points one year after the start of the programme. For the IB/ESA group, however, the programme's effect is substantial: the impact on benefit receipt is -9 percentage points, while the impact on employment is +11 percentage points after one year.

We discuss in this section a number of possible explanations for the observed impacts of ESF provision on participants in receipt of JSA and IB/ESA.

Voluntary JSA participants may not need ESF support

The voluntary nature of ESF provision means that ESF participants are likely to be more 'work-ready' than non-participants. In addition to this, evidence has shown that over half of new JSA customers end their benefit claim within three months and three quarters end their benefit claim within six months (for example Thomas, 2007). It is therefore likely that many JSA customers who volunteer to participate on ESF could have achieved jobs without the additional support provided by ESF.

While IB/ESA participants are also voluntary, and therefore likely to be more work-ready than non-participants, IB customers (who make up the majority of the IB/ESA participant group) have a lower base-level of employment support than JSA customers and also tend to be further away from the labour market. This could explain why impacts the of ESF provision are larger for the IB/ESA group than the JSA group.

Programme 'Lock In' Effects

Participants typically spend three and a half months receiving support from the ESF programme. This time on the programme reduces the time that participants may otherwise have spent on job-search activities. For participants who would have moved away from benefit quickly without support, the programme may therefore act to delay their move away from benefit. This could explain the low impacts observed for JSA customers, and in particular the initial sharp rise in benefit receipt among this group following the start of participation (shown in Figures 4.1 and 4.3). Similar evidence of so-called 'lock-in' effects was found in the recent evaluation by the European Commission (2010) into the impacts of ESF provision across four EU Member States. The report suggested that, "Impact in the short term is negative because people stop seeking jobs when they participate in a training programme (lock-in effect)."

For IB and ESA recipients, benefit claims tend to last much longer, with many individuals claiming these benefits for several years. ESF participants in receipt of these benefits would therefore be much less likely to move away from benefit quickly in the absence of ESF support. We would therefore expect any 'lock-in' effect from time spent on the programme to be much less pronounced for this group of participants.

Unobserved Benefits of the Programme

It is possible that the programme is successful in moving some ESF participants a step closer to the labour market, but not to the extent that they enter employment or move away from benefit receipt. 'Softer outcomes', such as increased skills, confidence, motivation etc. would not be observed in our current short term impact measures, but may enhance the lives of participants in other ways. Such effects are also likely to manifest themselves in the longer term labour market outcomes of participants, which are beyond our current analysis period.

It is also possible that as participants learn new skills and develop confidence, they develop higher employment expectations, and therefore turn down less desirable jobs which they may otherwise have accepted. The effect of this is that participants may take longer to find work, but ultimately gain a more desirable job. This may be reflected by sustained outcomes over a longer term period of several years, but again may not be observable in our current short term impact measures.

As more data becomes available it will be possible to explore this further by investigating the longer term impacts of the programme.

The 2008-2009 Recession

The cohort of participants analysed in the present study entered the ESF employment programme between June 2008 and April 2009. We have measured their labour market outcomes to April 2010. Our impact estimates therefore span a time period in which the UK was in a recession (2008-2009).

As participants and non-participants were tracked over the same period, we do not expect the recession to have caused any systematic bias in our impact estimates. However, it is important to note that our estimates can only tell us what the impacts of the ESF programme were during the analysis period, which included the period of recession. The effectiveness of the programme may have been very different had it operated under different economic conditions. As more data becomes available, it will be possible to compare the impacts of the programme during and after the period of recession to gain more understanding of how the recession affected the impacts of the ESF programme.

4.3 Sub-Group and Sub-Treatment analysis

Our primary estimates described in Section 4.1 have shown that the impacts of the ESF employment programme on its participants (the average effect of treatment on the treated, ATT) vary substantially according to whether the participants are in receipt of JSA or IB/ESA.

There are many reasons to expect that the impacts of the ESF programme are highly heterogeneous even within these broadly defined JSA and IB/ESA groups. The universal eligibility of the programme means that there is huge diversity in the characteristics of participants. It seems likely, for example, that a low-qualified 18 year old male from a poor neighbourhood will have a different participation experience from a highly qualified 50 year old female with a disability. There is also likely to be diversity in the type and effectiveness of support offered across the 74 DWP ESF contracts.

In this section we investigate a number of sub-groups within our original primary JSA sample, which define specific characteristics of participants that we expect to influence both propensity to participate and labour market

outcomes. We do this by 'hard matching' on the most pertinent participant characteristics. This means that we ensure that non-participants in the comparison group are identical to participants in the treatment group with regard to the characteristic of interest. For example, when we consider the impacts of the programme on different age groups, we separate the participants and non-participants by age group before using PSM to select non-participants from the resulting pools based on other characteristics.

In addition to allowing us to explore the heterogeneity of impacts, this 'hard matching' on the most pertinent participant characteristics has the additional advantage of providing a cleaner counterfactual. That is to say, because our groups are less heterogeneous with regard to observed characteristics, they are also likely to be less heterogeneous with regard to unobserved characteristics, thereby minimising any selection bias further prior to our difference-in-differences adjustment.

We also investigate the effects of different treatments provided to the JSA cohort. Firstly, we estimate the impacts of groups of ESF contracts which share characteristics that could influence the type of provision offered. For example, we consider the impacts of provision offered by private, public and third sector providers respectively. Secondly, we investigate whether the effectiveness of ESF support has varied over time by estimating the impacts of provision over two different cohort periods.

Table 4.1 below briefly introduces the sub-groups and sub-treatments considered in our analysis, while Sections 4.3.1 and 4.3.2 provide a more in depth discussion of the social and economic rationale for selecting these specific groups and treatments.

Table 4.1: Sub-groups and Sub-treatments for analysis

Sub-Groups	Sub-Treatments
Grouped by *Age*	Treated by *Sector*
Grouped by *District*	Treated by *Funding model*
Grouped by *Benefit duration*	Treated by *Provision type*
Grouped by *Gender*	Treated by *Unit cost*
Grouped by *Ethnicity*	Treated by *Contract size*
Grouped by *Disability*	Treated by *Cohort period*

We note that our sub-group and sub-treatment analyses are restricted to JSA customers only. We were unable to explore the IB and ESA group in this way, due to the relatively small number of participants within this group.

It is important to note that even if there are differences between sub-groups and treatments, this analysis cannot establish causation. Any difference in impacts between, say, participants of large versus small contracts is not necessarily attributable to contract size per se – it could be that the participants served by different sized contracts are different on other observable factors, such as age. Or it could be that the support itself varies in other ways - for example larger contracts may also be more likely to be run by private organisations. All we could say in this instance would be that there

might be a relationship between contract size and impacts. Establishing the reasons for differences would require further quantitative and qualitative analysis.

4.3.1 Rationale for selected sub-group categories

In this section, we describe the rationale for evaluating each of our selected participant sub-group categories. While there are very specific reasons for each category selection, they were broadly motivated by two overarching objectives:

1. we are interested in the heterogeneity of the impacts with regard to the most pertinent participant characteristics; and

2. by hard-matching on the most pertinent characteristics, we gain a counterfactual that is more precise with regard to each characteristic of interest, thereby lending more credibility to our Conditional Independence Assumption.

The selected categories also allow us to estimate the impact for each of the four ESF target groups: over 50s, female, ethnic minorities and disabled participants. This provides an unprecedented insight into the effectiveness of the programme on some of the groups it is most committed to helping.

Jobcentre Plus District

The importance of controlling for local labour market characteristics is strongly emphasised in much of the PSM literature (for example Bryson, et al., 2002).[32] Our primary analysis aimed to achieve this control by including three sets of geographical variables in the propensity score model: the Government Office Region of the individual; the district of the individual; and the individual Index of Multiple Deprivation scores of the individual (indicative of employment, income, health, education, housing, crime and living environment at a local super output area level).

To ensure that our primary model fully controls for local labour market characteristics we have also performed our analysis for a number of district sub-groups, by hard-matching on the *district* variable. In this analysis, participants can only be matched against non-participants from the same district. We have selected six districts from among those with high participation numbers, deliberately choosing districts which would provide us with a reasonable geographical spread across England. [33] These districts constitute about one third of our total participant sample for JSA recipients. In Section 4.3.3, we present our findings and compare the weighted mean of impacts across these districts with our primary estimate.

[32] However, we note that Dolton and Smith (2011) found that conditioning on variables relating to local labour markets had little effect on their resulting estimates.
[33] We do not identify the individual districts in this paper for reasons of data protection.

Benefit Duration

As discussed in Section 4.2, DWP benefits data shows that the majority of new JSA recipients move away from benefit receipt within a few months even without the help of employment programmes such as ESF. It is plausible that the extent to which participants need support is related to the duration of their current benefit claim. As the ESF fund has been used predominantly for short term JSA recipients during the recent recession, it is important to know whether this could offer a partial explanation of the low impacts of the programme on the JSA group. We have therefore estimated the impacts of the programme for participants, by hard-matching on the duration of their most recent JSA claim at the time of entering the programme. We consider the following groups:

- 0 to 3 months duration;
- 3 to 6 months duration;
- 6 to 9 months duration;
- over 9 months duration;

Age, disability, ethnicity and gender

Age, disability, ethnicity and gender are important characteristics of identity in the UK. The variables which relate to each of these characteristics in our model are therefore likely to be particularly important in determining a participant's probability of participation and their expected employment outcomes in the absence of support. Each variable is also indicative of a number of unobserved factors which are not directly captured in our model: for example, age is likely to be related to life experience and maturity, gender may relate to personality and attitudes, ethnicity could be indicative of cultural background, while a disability may relate to specific types of qualities and challenges which are not experienced in the wider population.

Hard-matching on each of these characteristics allows us to achieve our two primary objectives. Firstly, we are able to achieve a more precise counterfactual with regard to each pertinent characteristic, for comparison with our treatment sub-groups. By comparing the weighted impact across each category with our original primary estimate, we can gain confidence in the success of our original matching approach in controlling for these pertinent characteristics. Secondly, it is important to understand the impacts of ESF provision on each of the sub-groups within these categories.

The ESF programme aims to help five specific target groups, which include individuals aged over 50, females, ethnic minorities and individuals with a disability. Each of these groups is regarded as being at a disadvantage in the current UK labour market, indicated for example by low employment rates. The fifth ESF target group is lone parents. We were unable to analyse this group explicitly owing to the small number of participants who fall into this category.

We have therefore analysed the programme impact on participants falling into the following groups:

Table 4.2: Group categories for age, disability, ethnicity and gender

Age	Disability	Ethnicity	Gender
Under 25	Disabled	White	Female
25 to 49	Not Disabled	Black	Male
Over 50		Asian	

We note that only three ethnic groups were analysed in this way, as all other ethnic groups had too few participants to gain robust results.

4.3.2 Rationale for selected sub-treatment categories

In this section, we describe the rationale for evaluating specific groups of contracts with regard to their shared characteristics. Broadly speaking we aim to uncover the heterogeneity of impacts with regard to two dimensions:

1. the type of support which participants receive;
2. the incentives which drive providers;

Sector

It is likely that the incentives which drive organisations in each of the private, public and third sectors are very different. The three sectors are also likely to utilise different methods of implementing provision and to have different structures in place. In thinking about future policy for employment programmes such as ESF, it is therefore important to know the impacts achieved by providers from each sector.

Funding Model and Unit Cost

As discussed in Section 1.2.2, the ratio between service fee payments and outcome payments varies between contracts. The unit cost to DWP for each participant also varies between contracts. It is plausible that the incentives provided by different funding arrangements leads to different ways of operating among providers.

We have therefore analysed the programme impact of contracts subdivided into the treatment types shown in Table 4.3.

Table 4.3: Treatment by Unit Cost and Funding Model

By unit cost	By service fee: job outcome ratio
Lower Quartile	50:50 contract
Mid Quartiles	60:40 contract
Upper Quartile	30:70 contract

Provision Type

As discussed in Section 1.2.2, the DWP co-financing body categorises contracts into three broad types of provision:

- Tailored (a flexible, personalised approach similar to Flexible New Deal);
- Targeted (provision specified and aimed at particular support requirements);
- Intermediate Labour Market (ILM);

There are too few participants receiving ILM type provision to gain any meaningful impact results, so our sub-treatment analysis focuses only on participants treated under either *Tailored* provision or *Targeted* provision. By estimating the impacts of contracts which fall under each of these categories, we can begin to explore the relative effectiveness of the two approaches. This is particularly important in a policy environment where future provision looks likely to move towards the *Tailored* type of provision.

Contract Size

There is huge variability in the projected number of participants supported under each of the 74 DWP ESF contracts. These range from fewer than 100 participants under the smallest contracts to almost 20,000 participants under the largest. It seems highly likely that the methods adopted and the structures in place will vary according to the expected number of participants. The size of the contract may also be indicative of how well established a provider is. The current plans for future employment programmes are to move towards larger contracts, so it is particularly important to explore the relative effectiveness of large and small contracts on the current programme.

We have therefore analysed the programme impacts of contracts subdivided into the following contract sizes:

- Lower quartile (less than 2700 projected participants[34]);
- Mid quartiles (2700 to 7500 projected participants);
- Upper Quartile (more than 7500 projected participants).

Cohort Period

Feedback from the DWP CFO suggests that many ESF contracts have taken several months to 'get off the ground', and may therefore have been less effective in the early stages of the programme. This idea is loosely supported by management information data, which shows that the job entry rate of participants has improved over time since the start of the programme. If the programme effectiveness is indeed improving over time, then it is possible that our primary impact estimates may underestimate the true impact of the programme as a whole. To begin to explore whether support in the early and

[34] We note that the actual number of participants starting within our cohort period correlates strongly with the projected number of participants (R-squared=0.68).

later stages of the programme have had different impacts on participants, we have analysed the programme impact for participants in each of two narrower cohort periods:

- June 2008 to November 2008; and
- December 2008 to April 2009.

4.3.3 Results of the Sub-group and Sub-treatment analysis

Tables 4.4 and 4.5 show the benefit and employment impacts at 26 and 52 weeks[35] for sub groups and sub-treatments. Since the impacts fluctuate to varying degrees over time, we have also presented the mean impacts for the whole period of 52 weeks and for the period 26 to 52 weeks.[36]

It is essential to realise that the subgroup estimates are based on a smaller number of observations than used in our primary analysis. Consequently the impact estimates of the subgroups are less certain and have larger confidence intervals than the equivalent estimates described in our primary analysis of the entire JSA group, and it becomes more difficult to detect an effect which is statistically different from zero or from other groups, should one exist. Therefore, the lack of statistical significance should not necessarily be taken to mean that there was no impact for a particular subgroup; rather it should be viewed indicating that ESF has not had a sufficiently large effect, should one exist, for it to be captured statistically.

While all sub-group and sub-treatment impact estimates described in Tables 4.4 and 4.5 are for JSA customers only, we have included the impacts for the primary IB/ESA group at the bottom of each table for reference.

[35] Indicating statistical significance at the 90% and 95% level compared to zero impact.
[36] We do not calculate statistical significance for these mean impacts.

Table 4.4: Sub-group Impacts

Sub-Group Category	Treatment obs (Nt)	All Benefit Impact: Impact at 26 weeks	Impact at 52 weeks	Mean (0-52wk)	Mean (26-52wk)	Employment Impact: Impact at 26 weeks	Impact at 52 weeks	Mean (0-52wk)	Mean (26-52wk)
JSA baseline	25,720	0.9	-0.6	0.8	0.1	3.5**	4.5**	3.1	4.1
By Gender									
Female	5,200	-0.5	-1.3	-0.3	-1.2	4.1**	5.0**	3.7	4.7
Male	20,530	0.9	-0.6	1.0	0.3	3.5**	4.4**	3.0	4.0
TOTAL	25,720	0.6	-0.7	0.7	0.0	3.6	4.5	3.1	4.1
By Age									
Aged <25	5,000	1.9	0.6	1.7	1.2	2.9**	3.2**	2.6	3.2
Aged 25-49	16,050	0.3	-1.0	0.4	-0.4	3.8**	4.7**	3.3	4.3
Aged 50+	4,680	1.1	-0.4	0.7	0.0	3.9**	5.5**	3.5	4.8
TOTAL	25,720	0.8	-0.6	0.7	0.0	3.6	4.6	3.2	4.2
By Disability									
Disabled	7,230	1.6	-0.3	1.2	0.7	3.1**	4.2**	2.8	3.6
Not Disabled	18,060	-0.1	-1.4*	0.1	-0.8	3.8**	4.7**	3.3	4.4
TOTAL	25,300	0.4	-1.1	0.4	-0.4	3.6	4.6	3.2	4.2
By Ethnicity									
White	19,050	0.0	-1.5**	-0.2	-0.9	3.7**	4.6**	3.3	4.3
Black	1,820	1.5	-1.3	1.5	0.4	3.3	6.5**	3.8	5.1
Asian	1,480	4.1	4.2	4.2	4.0	2.7	3.5	2.7	3.4
TOTAL	22,340	0.4	-1.1	0.2	-0.5	3.6	4.7	3.3	4.3
By benefit duration									
0-3 months	11,920	-0.4	-1.9*	0.0	-1.0	3.9**	5.1**	3.3	4.5
3-6 months	6,980	0.4	-0.7	0.2	-0.6	4.0**	4.5**	3.6	4.5
6-9 months	3,540	0.6	-1.0	0.3	-0.4	2.5	4.1**	2.3	3.3
Over 9 months	3,290	0.1	-0.4	0.3	-0.1	3.1*	3.7**	2.7	3.5
TOTAL	25,720	0.0	-1.3	0.1	-0.7	3.6	4.6	3.2	4.2
By District									
District 1	2,000	2.0	0.7	2.7	2.7	0.1	2.9	1.4	1.7
District 2	1,000	2.7	4.1	3.3	3.4	4.0	2.8	2.7	2.9
District 3	1,000	-2.8	-1.0	-1.5	-1.8	5.1*	6.1**	3.7	5.5
District 4	2,000	0.9	-2.4	-0.3	-1.0	2.6	3.4	1.9	3.3
District 5	1,000	2.3	0.3	2.6	1.9	3.1	3.9	2.5	3.6
District 6	1,000	-0.3	-0.5	0.1	-0.5	3.8	4.0	2.9	3.8
TOTAL	9,000	0.9	-0.1	1.0	0.7	2.4	3.3	2.0	2.9
IB Baseline	1,970	-8.5**	-8.9**	-7.4	-9.5	9.8**	11.2**	9.2	11.1

Notes:

All participant observations are rounded to the nearest ten, except for those associated with specific districts which are rounded to the nearest thousand for data protection reasons.

For impacts at 26 weeks and 52 weeks:
* significantly different from zero at the 10 percent level;
** significantly different from zero at the 5 percent level.

Table 4.5: Sub-treatment Impacts

Sub-Treatment Category	Treatment obs (Nt)	All Benefit Impact: Impact at 26 weeks	Impact at 52 weeks	Mean (0-52wk)	Mean (26-52wk)	Employment Impact: Impact at 26 weeks	Impact at 52 weeks	Mean (0-52wk)	Mean (26-52wk)
JSA baseline	25,720	0.9	-0.6	0.8	0.1	3.5**	4.5**	3.1	4.1
By Sector									
Private sector	17,640	0.2	-0.9	0.2	-0.4	3.5**	4.2**	3	4
Public sector	3,130	1	-2.2	0.2	-0.8	3.8**	4.9**	3.3	4.6
Third Sector	4,880	1.6	-0.2	1.9	0.8	3.4**	4.9**	3	4
TOTAL	25,650	0.6	-0.9	0.5	-0.2	3.5	4.4	3.0	4.1
By Funding model									
50:50 contract	22,320	0.8	-0.5	0.8	0	3.7**	4.6**	3.3	4.3
60:40 contract	2,610	1.5	-0.7	2	1.6	0	1.2	0.6	1
30:70 contract	790	1	-5.9*	-2.8	-5	6.5*	8.0**	5.1	6.8
TOTAL	25,720	0.9	-0.7	0.8	0.0	3.4	4.4	3.1	4.0
By Cohort period									
Jun 08 to Nov 08 cohort	10,120	1	-0.4	0.6	0.3	2.4**	3.6**	2.4	3
Dec 08 to Apr 09 cohort	15,610	0.3	-0.7	0.7	-0.4	4.5**	5.1**	3.7	5
TOTAL	25,720	0.6	-0.6	0.7	-0.1	3.7	4.5	3.2	4.2
By provision type									
Tailored	21,080	1.2*	-0.3	1	0.3	3.1**	4.2**	2.8	3.8
Targeted	4,560	-2.4*	-3.4**	-1.6	-2.8	5.6**	5.9**	4.6	5.8
TOTAL	25,640	0.6	-0.9	0.5	-0.3	3.5	4.5	3.1	4.2
By unit cost									
Lower Quartile	6,550	1.7	0.2	1.4	0.8	2.2*	3.8**	2.3	3.3
Mid Quartiles	12,560	1.7	0	1.6	1.1	3.2**	4.0**	2.8	3.7
Upper Quartile	6,600	-1.7	-3.1**	-1.7	-3	5.3**	5.7**	4.5	5.5
TOTAL	25,710	0.8	-0.7	0.7	0.0	3.5	4.4	3.1	4.1
By Contract Size									
Large Contracts	6,110	1.8	-0.1	1.9	1.1	2.2*	3.6*	1.9	2.9
Medium Contracts	13,010	0.2	-1.2	0.3	-0.5	3.5**	4.0**	3	3.8
Small Contracts	6,600	0.5	-1	-0.1	-0.9	4.4**	5.8**	4.2	5.4
TOTAL	25,720	0.7	-0.9	0.6	-0.2	3.4	4.4	3.0	4.0
IB Baseline	1,970	-8.5**	-8.9**	-7.4	-9.5	9.8**	11.2**	9.2	11.1

Notes:

All participant observations are rounded to the nearest ten.

For impacts at 26 weeks and 52 weeks:
* significantly different from zero at the 10 percent level;
** significantly different from zero at the 5 percent level.

Overall, by far the largest difference in impacts is between the IB/ESA and JSA customer groups already described in our primary estimates. Within the JSA customer group there appears to be relatively little variation in impacts according to the characteristics of participants or the type of support; however, as we have discussed, the subgroup size is often too small to detect differences from zero, particularly for benefit impacts. Nonetheless, our findings suggest that on the whole, the impacts on the JSA treatment group are reasonably homogeneous across the sub-groups and sub-treatments that we have described.

In addition to understanding the heterogeneity of impacts across our sub-groups and sub-treatments, this analysis also enables us to gain more confidence in our original primary estimates. As already noted, by hard-

matching on the most pertinent participant characteristics, we gain a more precise counterfactual with regard to those characteristics, which thereby lends more credibility to our Conditional Independence Assumption. We have estimated the weighted impacts across each of the sub-groups and sub-treatments, as shown in Tables 4.4 and 4.5.

The overall impact estimates gained when hard-matching within each sub-group or by investigating the impacts of each sub-treatment category are broadly similar to our primary JSA estimates (all mean benefit impacts over the full 52 weeks lie between 0 and +1.5 percentage points; all mean employment impacts over the same period lie between +2 percentage points and +3.5 percentage points). This provides additional assurance that our initial approach to including these pertinent variables in our Propensity Score Matching resulted in a suitable counterfactual, and therefore that our primary estimates are robust.

Notwithstanding these two prime conclusions, some small differences appear to be emerging within the sub-group and sub-treatment categories, which we describe below. We note, however, that we cannot ascribe statistical significance[37] or causal explanations to these differences; this would require further investigation as more data becomes available.

By sub-group:

- Impacts are slightly better for women than men;

- Impacts are slightly better for participants aged 25 or over than participants aged under 25;

- Impacts are slightly better for participants who do not have a disability or health condition than participants who do have a disability or health condition;

- Impacts are slightly better for participants from white or black ethnic minority backgrounds than participants from an Asian ethnic minority background; and

- There is little difference in impacts between long and short term benefit recipients. This suggests that the reason for the poor impacts of the JSA group **cannot** be explained by the high number of short term JSA claimants flowing onto the programme from the recession.

By sub-treatment:

- There is little difference in impacts between public, private and third sector provision;

[37] As described in Footnote 30, this may be in part due to an overly cautious approach to calculating errors.

- Contracts with higher job outcome funding models are delivering better impacts than those with lower job outcome funding; although the subgroup is small, the 30:70 funding model shows particularly good benefit (-6 percentage points) and employment impacts (+8 percentage points) at 52 weeks. This suggests that the incentive offered for higher job outcome payments may be working for these contracts;

- Impacts are slightly better for participants from the later cohort period than the earlier cohort period;

- Targeted contracts are delivering slightly better impacts than tailored contracts. This suggests that contracts categorised to be more 'black box' in design may be having less impact than contracts focused on specific types of participant support;

- Contracts with higher unit costs are delivering slightly better impacts than those with lower unit costs; and

- Smaller contracts are delivering slightly better impacts than larger contracts.

5 Conclusions

This paper investigated the impacts of employment provision financed by the European Social Fund (ESF) and provided through the Department for Work and Pensions (DWP) on the labour market outcomes of unemployed and economically inactive participants in England. The analysis relied on a comprehensive individual database derived from administrative data sources. A Propensity Score Matching methodology was used to determine the labour market outcomes of participants and matched non-participants over a 52 week period following participation.

The main findings are:

- For **Jobseeker's Allowance** customers (67% of DWP ESF participants) the impacts of the programme are low. There is an increase in benefit receipt of up to 3 percentage points among participants over the first four months following programme entry, most likely due to reduced job-search activity while participating on the programme (commonly known as lock-in effect). In the following eight months there is almost no difference in the rate of benefit receipt between participants and non-participants. There is a small increase in employment rate among participants, which reaches 4.5 percentage points one year after programme entry.

- For **Incapacity Benefit and Employment Support Allowance** customers (6% of DWP ESF participants) the impacts of the programme are higher. There is a reduction in the rate of benefit receipt of 9 percentage points and an increase in employment rate of 11 percentage points among participants one year after programme entry.

In summary, ESF provision has low impacts on Jobseeker's Allowance customers, but is far more effective for Incapacity Benefit and Employment Support Allowance customers over the 52 weeks following participation.

In Section 4.2 we discussed a number of possible explanations for our findings, which we briefly outline below:

- The voluntary nature of ESF provision means that ESF participants are likely to be more 'work-ready' than non-participants. This is particularly the case among JSA customers who tend to be less disadvantaged and closer to the labour market than IB/ESA customers. It is therefore likely that many JSA customers participating on ESF could have achieved jobs without the additional support provided by ESF;

- Incapacity Benefit customers (who make up the majority of the IB/ESA participant group) have a lower base-level of employment support and tend to be further away from the labour market than JSA customers. This

could explain why impacts the of ESF provision are larger for the IB/ESA group than the JSA group;

- JSA customers tend to move away from benefit receipt quickly even without additional support. The short-term impacts for this group are negative, possibly because time spent on the programme leads to a reduction in job search activity (lock-in effect);

- There may be softer outcomes of the programme, such as increased skills, which would not necessarily be observed in our current short term impact measures, but may improve labour market prospects in the long term.

The sub-group and sub-treatment analysis of JSA recipients, discussed in section 4.3, suggests that the impacts of ESF are fairly homogeneous across the broad range of participant characteristics, and across the range of support offered by ESF providers. We describe some emerging observations in section 4.3.3. However, further investigation would be required before statistical significance can be ascribed to these findings.

It is important to emphasise that the scope of our analysis lies in estimating net impacts of DWP ESF employment provision on the benefit receipt and employment rate of participants. The programme may have a number of other benefits to its participants and to society as a whole which are not captured in our outcome measures. For example, participants may gain skills and confidence which move them a step closer to the labour market, without resulting in a change in labour market status. Or indeed, participants may attain higher paying or otherwise more desirable jobs than they would have achieved without ESF help. While the present analysis makes no attempt to directly capture such outcomes, some of these effects may influence our outcome measures in the longer term through their indirect influence on the benefit receipt and employment rate of participants. These may therefore be captured in any future studies investigating the longer term impacts of ESF support.

Finally, we outline below a number of possibilities for future research which would further our understanding of the effectiveness of ESF provision:

- further sub-group and sub-treatment analysis to see if any statistically significant differences emerge as more data becomes available;

- longer term impact estimates for the JSA and IB/ESA customer groups;

- impact analysis of participants entering the programme across the entire programme period;

- impact analysis of ESF for the other main client group (Income Support customers);

- impact analysis of Skills Funding Agency ESF funded adult employment provision.

Appendix 1 - Controlling for Labour Market History

As discussed in section 3.1, 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. In the early stages of our analysis, we explored the effectiveness of this approach in balancing our JSA participant and non-participant samples with regard to benefit and employment history. We constructed two labour market history strings for each individual to represent a timeline of their benefit and employment histories in the two years before entering the programme.

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. However, we found that our implementation of this method did not enable us to satisfactorily control for labour market history in the few months leading up to programme entry. We describe the reasons for this below, before outlining an alternative method which we developed and used throughout our analysis.

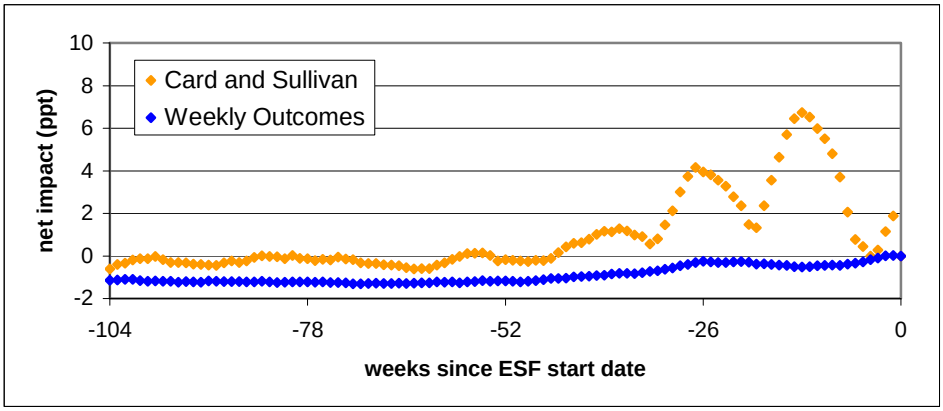
In the approach utilised by Card and Sullivan, the number of permutations of the constructed string variable is 2^N, where N is the number of time periods independently represented in the history string. Therefore, each additional time period included in the string doubles the number of possible permutations. In our implementation of the Card and Sullivan approach, we divided the labour market history into eight periods of three months. This allowed us to include two years of labour market history, with only 256 possible permutations. For example, the benefit string ‘11110000’ would indicate that an individual was receiving benefit for the first year (four periods of three months) and was not receiving benefit for the second year of the two year period before entering the programme. We considered an individual to be *receiving benefit* in a period if they were receiving benefit for more than half of the three month period.

We found that our implementation of the approach adopted by Card and Sullivan was limited by the large time periods represented by each binary indicator in the string (in our case three months). The method therefore failed to control for labour market history over sufficiently small periods of time, which was particularly apparent in the six month period immediately prior to the start of ESF provision, where differences between participants and non-participants are greatest. This is shown below in Figure 5.1.

We have therefore devised and implemented an alternative approach to control for labour market history, which is adapted to control for differences over smaller time periods. We generate 104 independent binary variables which represent an individual's benefit receipt or non-receipt in each of the 104 weeks prior to ESF start. We then generate a further 104 independent binary variables which represent whether an individual is in or out of employment in each of the 104 weeks prior to ESF start. Our probit model controls for labour market history using the resulting 208 binary variables.

We found that using this 'weekly outcomes' approach to control for labour market history resulted in a much tighter match with regard to benefit receipt between the two groups over the pre-programme period. Figure 5.1 below shows the percentage point difference between the proportion of the final matched treatment and comparison groups receiving JSA in each pre-programme week using both our implementation of the methodology adopted by Card and Sullivan and our alternative approach.

Figure 5.1: Comparing two alternative controls for labour market history



As shown, our alternative model constructs a comparison group which is more consistently matched with regard to the entire two years of benefit history. In our 'weekly outcomes' model, the percentage point difference between the treatment and comparison groups is approximately -1 percentage point in the period between 104 and 26 weeks prior to starting the ESF programme. In the 26 weeks immediately before the programme start, this difference reduces to less than half a percentage point. We conclude that, at least for the purposes of this ESF evaluation, our weekly history matching approach produces overall a more tightly matched comparison group with regard to benefit history than our implementation of the Card and Sullivan approach. We have therefore used this method throughout our analysis.

Appendix 2 – Generating Pseudo Start Dates

The benefit and employment outcomes of individual participants in our treatment group are measured weekly from the date on which they start provision. However, because participation is voluntary and can occur at any time during an individual's claim, those in our comparison group have no natural start date from which outcomes can be measured. It is therefore necessary to assign a 'pseudo start date' to each non-participant so that a time-based comparison between groups can be made. We must ensure that these pseudo start dates identify a period of time over which non-participants can best represent what would have happened to ESF participants if they had not participated in the programme.

The method used for generating pseudo starts is based on the technique described in the evaluation of New Deal for Disabled People (NDDP) (Orr et al, 2007). It aims to align the non-participants and participants with respect to two time dimensions: calendar time and length of time on benefit.

The method used for generating pseudo start dates is as follows:

1. All participant records were separated according to the benefit start month of the participant – i.e. separate data sets were created for participants starting benefit in June 2007 [38], July 2007, Aug 2007, and so on up to the latest ESF start month included in the analysis;

2. For each of these participant data sets, we calculated the proportion of ESF starts occurring in each possible ESF start month. For example, for participants with a benefit start month of September 2007 (Table 5.1 is representative only):

   Table 5.1: Illustrative example of the proportion and cumulative proportion of participants starting ESF provision by month

ESF start month	% of ESF starts	Cumulative %
June 2008	1.2	1.2
July 2008	7.3	8.4
Aug 2008	6.1	14.5
.	.	.
.	.	.
April 2009[39]	9.4	100

[38] Note that records with benefit start dates prior to June 2007 were kept together as a single file. This was because it was assumed that the decision of when to begin ESF participation would not depend on benefit start month for those who had been on benefit for such a long period.

[39] April 2009 was the latest month from which 52 weeks worth of outcome data were available. Participants who started after this date were not included in our analysis.

3. We then separated all non-participant records according to the benefit start month of the non-participant, in the same manner as carried out in Step 1 for participants;

4. For each non-participant benefit start month file, we randomly assigned a pseudo start month from the distribution of ESF start months gained from the participant file with the equivalent benefit start month.

   E.g. for participants with a benefit start month of June 2007 (again, figures are representative only):

   Table 5.2: Illustrative example of the assignment of pseudo start dates to non-participants by month

ESF start month	% of ESF starts	% of pseudo starts
June 2008	1.2	1.2
July 2008	7.3	7.3
Aug 2008	6.1	6.2
.	.	.
.	.	.
April 2009	9.3	9.4

5. We then assigned a random date in the assigned month from a flat distribution (i.e. all dates within the month were equally likely).

6. If an assigned pseudo start date occurred at a time when the non-participant was not claiming their primary benefit (JSA or IB/ESA), then the pseudo start was considered 'invalid' and the record was removed from the sample.

In Appendix 6 we additionally describe a sensitivity test performed in our model development stage to test the sensitivity of our impact estimates to the method of generating pseudo start dates.

Appendix 3 – Matching Protocol

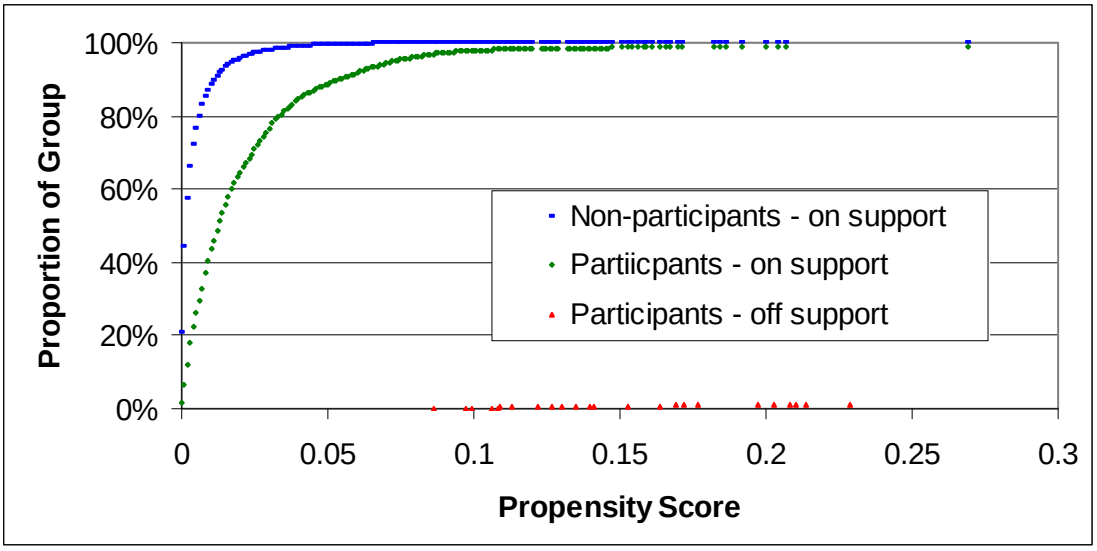
We describe below the matching protocol used in this evaluation to construct suitable treatment and comparison groups from our participant and non-participant samples.

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 higher 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 4 - Propensity Score Distribution for primary IB and ESA group:

The propensity score distribution for our IB and ESA primary estimate is shown in Figure 5.2 below; each point is the cumulative proportion of the group with propensity scores in increments of 0.001.

Figure 5.2: Primary IB/ESA Analysis Propensity Score Distribution: (1,970 treatment records; 406,430 comparison records)



The graph shows that propensity scores tend to be higher for participants, although the mean score for both groups is low. The mean propensity score is 0.005 (proportion of records receiving ESF support); 28% of non-participants have a score of more than the mean compared with 77% of participants; also 3% of the non-participants have scores more than 0.025 compared to 31% of participants. The small proportion of participants for whom no common support is available are those with the highest propensity scores.

We find that our propensity score distribution for IB/ESA recipients provides sufficient overlap between participants and non-participants, with almost 99% of participants finding common support.

Appendix 5 - Mutually exclusive outcomes and impacts

As discussed in Section 4.1.1, our primary estimates for the JSA group show that both the benefit and employment impact estimates are positive for the first 15 weeks following participation. This suggests that ESF participation increases an individual's chances of both claiming benefit and being in employment. As noted, this result does not present a contradiction because the impacts we consider are not mutually exclusive and do not account for all possible outcomes.

To illustrate this point, Table 5.3 below shows the impacts of ESF support on JSA customers with regard to four mutually exclusive labour market outcomes:

- *Benefit only* - receiving benefit and not in employment;
- *Work only* - in employment and not receiving benefit;
- *Neither* - neither in employment nor receiving benefits.
- *Both* - both in employment and receiving benefit.

Table 5.3: Impacts (percentage points) on mutually exclusive outcomes at 26 weeks for JSA customers

Benefit only	Work only	Neither	Both
-0.3	2.3	-3.2	1.2

It can be seen from Table 5.3 that the arithmetic sum of all mutually exclusive impacts is zero. The negative impacts of the programme on *receiving benefit only* and *neither in employment nor receiving benefit* precisely balance the positive impacts on *employment* and *both receiving benefit and in employment*. The impacts we describe throughout the rest of this paper are not mutually exclusive, so the impact on benefit receipt is the arithmetic sum of the impact on 'benefit only' and on 'both', while the impact on work is the arithmetic sum of the impact on 'work only' and on 'both'.

Appendix 6 - Sensitivity to methods

As discussed in Section 3.4, during the development stages of our model we explored the sensitivity of our impact estimates to the precise implementation of the PSM methodology. These sensitivity tests helped to inform our decisions in producing a final preferred methodology, which we then used in our primary estimates, sub-group and sub-treatment analyses. We performed the following sensitivity tests on subsets of the JSA group.

- sensitivity to time-based variables;
- sensitivity to alternative pseudo start generation methods;
- sensitivity to the Kernel bandwidth;
- sensitivity to alternative methods of cleaning the HMRC employment data;

Sensitivity to time-based variables

As discussed in Appendix 2, for our comparison group to represent a suitable counterfactual we need to ensure that ESF participants and non-participants are aligned in time (both calendar time and length of time on benefit). Our pseudo start generation method provides us with a comparison sample whose starts distribution is similar to that of the treatment group. However, we still need to ensure that our inclusion of time-based variables in the propensity score model produces a final comparison group which is aligned in time with the treatment group.

To this end our primary analysis uses two time-based variables: benefit start month and ESF/pseudo start month.

To investigate whether our impact estimates were sensitive to the choice of time-based variables, we performed exploratory analysis using the following alternative variable sets in our probit model:

1. ESF/pseudo start month only;
2. ESF/pseudo start month, benefit start month and benefit duration at start date;
3. ESF/pseudo start month and benefit duration at start date.

We found that altering the time-based variables used in the PSM made very little difference to the final impact measure. We therefore conclude that our impact estimates are not strongly dependent on the exact method of incorporating time-based variables into the probit model.

Sensitivity to the method of generating pseudo starts

We carried out two sensitivity tests on the method with which pseudo starts were generated to address two key analytical questions:

1. Do our impact estimates depend on the random allocation of pseudo starts from the actual starts distribution – i.e. if we had used a different series of random numbers in the assignment, would we have achieved a different result?

2. Does it matter if the distribution of pseudo starts for the comparison group differs from the distribution of actual starts in the treatment group?

To address the first question, we simply used a different random seed to generate pseudo start dates. This resulted in each record being assigned an alternative pseudo start date, and consequently different non-participants being retained in the sample once we had removed those with 'invalid' pseudo start dates (dates on which the non-participant was not receiving their primary benefit).

The use of an alternative random seed produced no change to our impact estimates.

To address the second question, we used a flat distribution to randomly assign pseudo start dates to non-participants. Each non-participant in the comparison sample was therefore equally likely to receive a pseudo start date in any of the start months, regardless of their duration on benefit. We then used time based variables in the propensity score model, as in our primary analysis: *ESF (pseudo) start month* and *benefit start month*.

Using a flat distribution to generate pseudo starts produced no change to our impact estimates. We conclude that the precise method of generating pseudo starts makes no difference to the impact estimate, provided that:

- there are enough non-participants who align with participants on the dimensions of *ESF/pseudo start date* and *benefit duration*;

- we include *ESF/pseudo start date* and *benefit duration* on time based variables when performing the PSM.

Sensitivity to the Kernel bandwidth

The bandwidth used in the Kernel matching determines how closely the propensity score of a non-participant must be to that of a participant in order for the non-participant to be selected as a match for the participant. The bandwidth used in our primary analysis was 0.0001. Thus a participant will be matched with any non-participants who have a propensity score within +/- 0.0001. The closer the non-participant score is to the participant score, the higher it will be weighted in the final matched comparison group.

If the bandwidth is increased, it means that each participant will match with more non-participants. This has the advantage that it can increase common support for participants with extreme propensity scores, but has the disadvantage that we are giving greater weight to non-participants who are

less precisely matched. Conversely, if the bandwidth is decreased, it means that participants will more tightly match to non-participants. This has the advantage of ensuring a more closely matched comparison group, but the disadvantage that we may not find matches for some participants – a lack of common support. Effectively the choice of bandwidth represents a trade-off between tightness of the matching and level of common support.

In our model development stage, we investigated the effect of varying the bandwidth on both the impact estimates and the level of common support for a sub-sample of our JSA primary sample. The bandwidths investigated were 0.0001, 0.00001 and 0.000001. We found that using a smaller bandwidth had the effect of reducing the level of common support (down to 83% for a bandwidth of 0.000001 from 98% for our primary analysis bandwidth of 0.0001) but no substantial effect on the impact estimates (all three estimates were within 1.5 percentage points of one another for JSA, benefit and employment impacts at a time 34 weeks after the start of participation).

We conclude that our impact measure is sufficiently insensitive to changes in the bandwidth used.

Sensitivity to the method of cleaning employment data

As discussed in Section 2.2.2, a large proportion of employment spells on the data set provided by HMRC have missing start or end dates. Our primary estimates were achieved after a process of randomising the missing start and end dates across the tax year in which they were known to occur. While we do not expect any systematic bias to result from this process, we note that this leads to a large proportion of individuals (up to 30%) being identified as both ‘receiving benefit’ and ‘in work’ in the same week. To provide additional assurance that our employment impact estimates were not biased by our randomisation process, during our model development stage we tested the sensitivity of our impact estimates to an alternative method of randomisation for the JSA group.

The alternative approach makes two main assumptions:

1. that the DWP benefits data is correct;
2. that an individual cannot be simultaneously in employment and receiving JSA.

The first step of this approach is, as previously, to randomise missing start and end dates across the tax year in which they were known to occur. We then adjust the randomised dates such that they do not fall within a known JSA spell of the individual, as described below:

- If a JSA spell overlaps with the employment end date, we assign a new employment end date equal to the JSA start date;

- If a JSA spell overlaps with the employment start date, we assign a new employment start date equal to the JSA end date;

- If a JSA spell fully encloses the employment spell, the employment spell is assumed to be incorrect and is removed;

We found that, while using this alternative method of cleaning the employment data led to lower proportions of individuals in both the treatment and comparison groups being identified as in employment for any given week, the impact on employment was similar for both methods (to within 0.5 percentage points).

We therefore conclude that the method with which the employment data was cleaned does not appear to systematically bias our impact estimates.

References

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

Blundell R. and Costa Dias M. (2000), Evaluation Methods for Non-Experimental Data, Fiscal Studies 21, 4, 427-468.

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.

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.

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

Drever, E. and Lloyd, C. (2010), European Social Fund Cohort Survey: Wave 1, DWP Research Report 647.

European Commission Evaluation (2010), Study on the Return on ESF Investment in Human Capital.

Lechner, M. and Wunsch, C. (2009), Active labour market policy in East Germany: waiting for the economy to take off, Economics of Transition Volume 17 (4) pp 661-702.

Leuven, E. and Sianesi, B. (2003), psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Lords EU Committee Report (2010), Making it work: the European Social Fund (9th Report, Session 2009-10, HL Paper 92).

Orr L., Bell S. and Lam K. (2007), Long term impacts of the New Deal for Disabled People, DWP Research Report 432.

Reinowski E., Schultz B. and Wiemers J. (2004), Evaluation of Further Training Programmes with an Optimal Matching Algorithm, Halle Institut for Economic Research (IWH).

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.

Speckesser, S. and Bewley, H. (2006), The Longer term impacts of Work Based Learning for Adults, DWP Research Report 390.

Thomas A. (2006), Evaluating the econometric evaluations of active labour market programmes using administrative data, DWP Working Paper 50.

Thomas A. (2007), Flexible New Deal Evidence Paper, DWP publication.

The European Social Fund (ESF) was set up to improve employment opportunities in the European Union and so help raise standards of living. Its aim is to help people fulfil their potential by giving them better skills and better job prospects. Employment support provided through the European Social Fund is varied and flexible, including activities such as job search guidance, basic skills training, case worker support and advice on tackling specific barriers to work.

This paper describes findings from an evaluation of the net impacts of the 2007-13 European Social Fund (ESF) Programme for England. The study is focused on the Department for Work and Pensions (DWP) ESF funded employment provision part of the programme. In this paper, we estimate the impacts of ESF support on participants who entered the programme between June 2008 and April 2009. We provide separate impact estimates for two broad DWP customer groups: participants in receipt of Jobseeker's Allowance and participants in receipt of Incapacity Benefit or Employment Support Allowance.

If you would like to know more about DWP research, please contact:
Kate Callow, Commercial Support and Knowledge Management Team,
Work and Welfare Central Analysis Division, Upper Ground Floor, Steel City House, West Street, Sheffield S1 2GQ.
http://research.dwp.gov.uk/asd/asd5/rrs-index.asp

DWP Department for Work and Pensions

Published by the
Department for Work and Pensions
May 2011
www.dwp.gov.uk
In-House Research no. 3
ISBN 978-1-84947-615-7