# Central London Works

## Impact scoping study

Helen Gray

# Institute for Employment Studies

IES is an independent, apolitical, international centre of research and consultancy in public employment policy and HR management. It works closely with employers in all sectors, government departments, agencies, professional bodies and associations. IES is a focus of knowledge and practical experience in employment and training policy, the operation of labour markets, and HR planning and development. IES is a not-for-profit organisation.

# Contents

# 1    Introduction

## 1.1    Purpose of the report

The main aim of this report is to assess the feasibility of estimating the impact Central London Works (CLW) on the outcomes it seeks to effect, given the likely availability of data, both on programme participants and a potential comparison group.

## 1.2    The intervention

CLW aims to assist those who face barriers to work to find and sustain good quality employment. It operates as an alternative to the national Work and Health programme (WHP), with some differences in design between the two schemes. The national WHP is being evaluated through a randomised control trial (RCT). Jobcentre Plus Work Coaches are responsible for assessing eligibility and making referrals to CLW or assigning individuals to a control group who receive Jobcentre Plus business as usual support.

CLW provides bespoke support for up to 21 months, including 15 months of support in finding work and six months of support after starting a job. The programme started in March 2018 and is expected to support around 21,000 residents in 12 Central London boroughs referred to the programme in the period to March 2023. The intervention is targeted at three main groups:

- people with health conditions or disabilities;
- the long-term unemployed, defined as those out of work for over 24 months;
- and those with other barriers to work, such as carers, ex-offenders, homeless people and those with English language needs, referred to as the early access disadvantaged group.

Participation in CLW is voluntary.[1] At the outset, the expectation was that three-quarters of participants would have a health condition or disability, with two-thirds of these on Jobseekers' Allowance (JSA) and one-third in the Employment and Support Allowance (ESA) work-related activity group.

The main aims of the intervention are to increase the likelihood of participants entering work and to raise their earnings. It is delivered by service providers (operating under a main provider) who are rewarded for non-employment outcomes, in-work earnings in

---

[1] The specification for the WHP stated that CLW was mandatory for the long-term unemployed, but our understanding is that this is not in fact the case.

excess of a national earnings threshold common to the WHP, and an additional threshold to recognise earnings above the London Living Wage.

### 1.2.1  Expected outcomes

The primary outcomes that CLW seeks to effect are:

- the probability of being in paid employment;

- earnings from paid employment.

In addition to these, it can be expected to have an impact on other measures such as working hours and the probability of being in receipt of out-of-work benefits. The intention is that it would also have an impact on other activities which would improve the likelihood of entering work at some point in the future, such as skills acquisition and work-readiness, as well as the health and well-being of participants.

## 1.3  Overview of potential data sources

### 1.3.1  Management Information

Providers are responsible for populating a detailed management information (MI) database on programme participants. The MI data supplied to IES in early December 2019 included 5,319 individuals who were referred to CLW between the start of the programme in March 2018 and 25 October 2019 and who attended an initial appointment. These individuals also had to consent to data sharing to be included in the extract supplied. The MI covers programme participants only and not trial participants assigned to the control group by Jobcentre Plus.

Although all individuals included in the MI extract had attended an initial appointment, just under a quarter (23.0 per cent) had not completed the detailed assessment which is carried out in the early interviews. For those who did complete the assessment, the MI database includes their responses, as well as information on their demographic characteristics and personal circumstances prior to starting on the programme. The assessment is designed to identify barriers to work to assist in drawing up an action plan which can be used to increase job readiness. The MI record is updated with details of any interventions completed with the main provider, or another specialist provider. It also records details of any jobs attained, including start and end dates, the type of contract, average weekly working hours and the hourly wage. Details of up to three jobs were recorded on the data extract supplied in December 2019.

As the MI covers individuals who participated in an initial interview and not those in the control group, it is not suited to estimating the impact of CLW. The reasons for this are explained in the following chapter.

### 1.3.2   Administrative data

The main administrative datasets which are likely to be relevant to this study, given the primary aims of getting participants into work and progressing into better-paid jobs, are Department for Work and Pensions (DWP) data on benefit claims and Her Majesty's Revenue and Customs (HMRC) data on employment and earnings. Records on individuals from each of these datasets can be linked together using National Insurance (NI) numbers to build up a detailed picture of an individual's history of employment and claiming benefits. They also include basic demographic information, such as gender, ethnicity and age. These key fields are generally well-recorded for all individuals who have claimed benefits at some point in the past 20 years.

Being able to obtain access to benefits and employment records for individuals from administrative data sources depends on having sufficient information on programme participants to be able to identify them in DWP and HMRC datasets, as well as permission to use personal data for the purposes of data linking. The process of data linking would be most straightforward if NI numbers were available for all trial participants. If alternatively NI numbers are not available for all, it would be necessary to use other personal data (first name, surname, date of birth, gender and postcode) to find trial participants in the administrative datasets. The feasibility of this approach depends on whether all of these items of personal data are well-recorded for all trial participants, but the match rate to administrative data sources using this approach is likely to be lower than if NI numbers were available for all.

In addition to the question of whether data permissions and the information collected on trial participants are adequate to allow data linking, it is important to note that there are a number of limitations to the DWP and HMRC administrative data, set out below.

1. The data on benefits receipt are constructed from regular snapshots of the live benefits system. As a result, short periods of time on benefits which fall between two scans may be omitted from the database.

2. Whilst the benefits data in particular contain many fields, there are differences in the quality and comprehensiveness of the information recorded for different types of benefit. In practice, only a fairly limited amount of information is likely to be available for all trial participants. The fact that some important drivers both of programme participation and the outcomes experienced as a result may be unobserved means some methods of estimating impact are unlikely to be robust.

3. The roll out of Universal Credit (UC) to replace existing working age benefits started in 2013. By the end of 2018 all Jobcentres were due to replace legacy benefits with UC for all new claimants, but the process of transferring existing claimants to UC is not due to be completed until 2023. Within the 12 London boroughs participating in CLW, by August 2019 79,738 individuals who were not in employment were claiming UC, whilst a further 94,814 individuals remained on ESA and 11,568 individuals were on JSA. Those eligible for CLW would be split across UC claimants and those on legacy benefits. Whilst our current understanding is that it is possible to observe UC claims within DWP administrative datasets, less is known about the details of what is recorded for those on different conditionality regimes. It is also uncertain how compatible

information on claims for legacy benefits and claims for UC are, so this may potentially make it difficult to explore the impact of CLW as an increasing number of individuals move from legacy benefits to UC.

4. The HMRC data do not cover self-employed workers who provide details of their employment and earnings through the self-assessment system. Estimates from the Labour Force Survey (LFS) suggest around 15 per cent of the UK workforce were self-employed in the final calendar quarter of 2019,[2] so this omission is a significant limitation of the HMRC data. Even if it is possible to obtain access to the self-assessment data, there are long lags in this information becoming available, as the tax return is not due until the January after the end of the tax year. Even for those in employment, the start and end dates of employment spells are unknown in a large proportion of cases, or only the tax year when the spell started or ended is known.

5. Whilst the HMRC data are now based on real-time information (RTI), there are nevertheless time delays of around two months between an individual starting or leaving employment and this information being made available to analysts. DWP benefits data are subject to a greater lag of around three or four months. This is likely to mean that the analysis will only provide information on benefit or employment spells up to around four months before drawing the data extract.

### 1.3.3  Publically available data sources

A range of publically-available data sources could potentially be used to identify a suitable set of comparison areas, including information from the LFS on unemployment rates, the MHCLG index of multiple deprivation, Census data on ethnicity and DWP data on benefit receipt. Again, the potential value of this approach is considered in the following chapter.

## 1.4  Report structure

The next chapter outlines the aims of the impact evaluation of CLW. It set out the main challenges which are likely to be encountered in the course of the impact evaluation and reviews quasi-experimental methods which may provide a feasible way of estimating impact as an alternative to using the RCT data. The chapter also summarises the main outcome measures which are likely to be affected by CLW, given the objectives of the programme. It sets out the intentions with regards to exploring the relative effectiveness of CLW and any variation in its impact on each of the three main target groups.

The final chapter summarises the overall feasibility of estimating the impact of CLW and sets out the proposed approach to impact evaluation, given the findings from the review of possible methods. It concludes by setting out the steps required to conduct an impact evaluation, highlighting the immediate priorities.

---

[2] Figures calculated from the Office for National Statistics (2020) EMP14:Employees and self-employed by industry dataset, available at:
https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/employeesandselfemployedbyindustryemp14

# 2   Appropriate methods

## 2.1   Introduction

This chapter starts by providing a brief description of the primary aims of any impact evaluation. It explains the likely advantages and disadvantages of seeking access to data from the national evaluation of the WHP. It then moves on to consider how it would be possible to explore the relative effectiveness of CLW compared with alternative options, such as Jobcentre Plus business as usual or the national WHP. It also highlights some of the potential difficulties in estimating the impact of CLW.

The chapter provides a brief review of alternative methods to estimate the impact of CLW, given the potential limitations inherent in the RCT data from the national evaluation. This section focuses on the methods which are most likely to provide a robust estimate of impact, whilst highlighting any aspects of each approach which are likely to undermine the credibility of the findings.

The chapter provides a list of outcome measures appropriate to capturing the impact of CLW and concludes with an assessment of the likely feasibility of carrying out a subgroup analysis to explore how the impact of CLW varies between participants, as well as to explore the relative effectiveness of CLW compared to the wider WHP.

## 2.2   Estimating impact

To estimate the impact of any intervention it is necessary to form a credible estimate of outcomes if the programme had not been introduced. This is known as the **counterfactual**. Estimated counterfactual outcomes can be compared to observed outcomes for participants - the **treatment group** - to estimate the impact of the intervention. For the estimate of the counterfactual to be accurate it is necessary to adjust for any changes in outcomes over time that might have occurred even without the intervention. If the counterfactual does not take account of changes in outcomes that might have occurred anyway, the estimate of impact will not reflect the true impact of the intervention.

In the case of CLW, individuals who meet the eligibility criteria are randomised to treatment and control groups by Jobcentre Plus Work Coaches. With eligible individuals assigned at random to either group, outcomes for the control group should in theory provide an unbiased estimate of counterfactual outcomes for the treatment group. However, in practice the experience of randomisation for the WHP is that at least some

individuals in the control group have ended up receiving the treatment at a later point in time.[3] This is known as **contamination**.

If it is not possible to distinguish between members of the control group who do participate in the WHP and those who are never treated, it is likely that the counterfactual would be biased upwards, reducing the chances of detecting any impact from CLW. However, as the MI database allows us to observe those participating in CLW, it may be possible to use this source in conjunction with individual-level data from the national WHP RCT to identify and exclude the subset of those initially assigned to the control group who subsequently participate in CLW. The feasibility of this approach depends on the personal data recorded in both the CLW MI and national WHP RCT datasets being sufficient to identify members of the control group who end up participating in CLW. Unless all individuals found in the CLW MI are matched to the national RCT data, it will be uncertain whether all members of the control group not found in the CLW MI are untreated or whether the failure to find a record of participation is due to inadequacies in the recording of personal data in either source.

Even assuming it is possible to link the CLW MI with individual-level data from the national RCT and exclude members of the control group who are actually treated, it is possible that there are systematic differences between the treatment group and members of the control group who do not receive support which have a bearing on the outcomes they experience. For example, it seems probable that members of the control group who are unable to find work are more likely to end up being referred to CLW at a later date, whereas those who do find employment are less likely to receive support. If this is the case, restricting the control group to those who never participate in CLW could reduce the estimated impact of CLW. If it was possible to link the CLW MI with the data from the national RCT, it would at least be possible to calculate the proportion of the control group who did receive the treatment. If a very low proportion of the control group were referred to CLW at a later date, there would be less reason to believe that outcomes for the control group would provide a poor estimate of the counterfactual.

A further risk in comparing outcomes for the treatment group with that of the untreated control group is that the size of the untreated control group may be so small that it is not possible to detect any impact from the programme. It would be difficult to judge the likelihood that this will be the case before linking the datasets.

If it is feasible to link individual level data from the national RCT with the CLW MI database and if this suggests the extent of any contamination is very limited, it would be possible to compare the characteristics of the treatment and control groups prior to randomisation to explore whether the two groups appeared similar on a range of observable characteristics before participation. Provided this analysis suggests that the characteristics of the treatment and control groups are similar, this would increase

---

[3] Our understanding is that this has occurred because it is difficult for Jobcentre Plus Work Coaches to know whether individuals coming forward to participate in the WHP have previously been assigned to the control group. This means that members of the control group may end up receiving the treatment.

confidence that comparing outcomes for the control group with those for the treatment group would provide a defensible way of estimating the impact of the intervention. Reanalysis of the existing data from the RCT for the 12 boroughs participating in CLW is likely to provide the most robust estimate of impact at the lowest cost provided it is possible to establish that the extent of any contamination is low and the treatment and control groups were similar prior to randomisation.

## 2.3    Comparison with alternative interventions

Those coming forward to participate in CLW are already eligible to receive support designed to increase their chances of finding work. Whilst the intention is that the control group should be excluded from participating in CLW, they do receive business as usual support from Jobcentre Plus. This means that the impact of CLW will be measured against existing Jobcentre Plus provision, rather than nothing at all.

As well as estimating the impact of CLW compared against existing provision, with access to data from the national WHP for areas outside of the 12 boroughs offering CLW, in theory it would be possible to estimate the impact of CLW compared to the version of the WHP running in other areas. However, while it may be possible to use CLW MI to explore and potentially overcome any contamination of the control group in the 12 boroughs participating in CLW, it may not be feasible to investigate the extent of this problem in other areas.

The only potential way of estimating the relative effectiveness of CLW would be if other areas similar to the 12 boroughs participating in CLW were able to supply MI on participants in their area which could be linked to data from the national RCT. If it were possible to construct a linked MI and RCT database for other areas, the scale of contamination could be investigated  and if this was low, members of the control group who participated in the WHP in these areas could be excluded from the analysis. However, it may not be possible to obtain access to MI from other areas under existing data sharing arrangements.

Aside from the issue of potential control group contamination, as WHP trial participants in other areas would receive an intervention with similar aims to CLW, any differences in impact between the two interventions are likely to be small in magnitude. This means it may be difficult to say with certainty that there are clear differences in impact between CLW and the version of the WHP running in other similar areas.

## 2.4    Challenges

There are a number of challenges that any evaluation of CLW will potentially face. Firstly, a lack of access to individual-level data from the national RCT will reduce the ability to form a robust estimate of impact. As participation in CLW is voluntary, there is a risk that those who choose to participate are more motivated to find work than those who do not. Non-participants include both those who are assigned to the control group at random and those who do not come forward to participate in the trial. It is likely that those who are motivated to join the trial are also more motivated to work than those who choose not to

participate. Comparing outcomes between those who receive support from CLW and those who do not may inflate the estimated impact of CLW if participants and non-participants differ in their levels of motivation. For this reason, if it is necessary to rely on quasi-experimental methods to estimate impact, rather than using the RCT data, there is a risk that the impact estimates will be overstated, unless there is good reason to believe that any differences in motivation levels are observed.

If it were not possible to gain access to data from the national RCT, an alternative option would be to request access to DWP and HMRC administrative data on the wider population of individuals eligible to participate in CLW. Outcomes for individuals eligible for CLW within the 12 participating boroughs would be compared with outcomes for individuals who could be observed to meet the eligibility criteria in other similar areas not participating in CLW. However, it is likely to be difficult to detect any impact from CLW if it is not possible to distinguish between those in the treatment and control groups as those eligible for CLW in the 12 participating boroughs will include those receiving Jobcentre Plus business as usual support as well as those participating in CLW. In the comparison areas, those meeting the eligibility criteria for CLW would include those participating in the national WHP as well as those subject to Jobcentre Plus business as usual. As the comparison areas would include individuals receiving an intervention similar to CLW, it may be difficult to say with certainty whether CLW has had an impact.

It would be necessary to obtain agreement from DWP and HMRC to use their administrative data in the evaluation of CLW and so access is not guaranteed. The current process to gain approval is to draw up a proposal for the DWP External Data Sharing Advisory Committee (EDSAC). The proposal would need to explain why the data are required and set out the methodology being used in the evaluation. The Committee assesses the feasibility of the proposed research and the legal basis for granting access. A data protection impact assessment would also need to be carried out. Access to the data would be likely to be through the Office for National Statistics (ONS) Secure Research Service which limits access to ONS accredited researchers working in a secure environment.

A further challenge faced in any impact evaluation is whether the number of individuals coming forward to participate in the intervention is sufficient to be able to say with certainty whether the intervention has had a discernible impact. This can be more difficult in the case of an RCT as some of those agreeing to participate in the trial are assigned to a control group. In the case of CLW it is understood that Jobcentre Plus determines which individuals are assigned to the treatment or control groups and this decision is made before individuals are referred to the CLW provider.

The initial expectation was that around 20,800 referrals to CLW would be made over the lifetime of the programme, with a further 5,290 individuals assigned to the control group. In practice, by 25 Oct 2019, 5,319 individuals had been referred to CLW and attended an initial appointment, compared with an expected volume of around 8,000 referrals by this point. It is uncertain how many individuals were assigned to the control group over this period, but the expectation was that it would amount to approximately 2,000 individuals by around October 2019.

As the volume of referrals to CLW is currently below initial expectations, it is uncertain whether it will be feasible to achieve the targets for the size of the treatment and control groups over the remainder of the trial. However, this seems unlikely as the time available to increase referral volumes is also limited by the current aim of producing the final evaluation report by January 2023. Given that it will be necessary to complete the impact evaluation, cost-benefit analysis and the synthesis of research findings across the different strands of the evaluation prior to submission of the final report, the data extracts required for the impact evaluation will be needed by early 2022. Allowing for time lags in administrative data sources being updated, it is likely to be possible to observe outcomes for a period of 12 months following the initial assessment for those starting on CLW by Autumn 2020.[4] If the numbers referred to CLW were in line with expectations by this point, this would mean that the analysis could focus on a cohort of around 13,000 individuals, with a further 3,245 assigned to the control group.  However, actual volumes may well be lower than this unless the shortfall in referrals in the period to October 2019 is made up in the year to October 2020. Failing to meet the expected volumes of referrals will reduce the likelihood of detecting any impact from CLW.

Early experiences were that some clients experienced a delay in starting to receive support from CLW. In some cases the programme may not have been delivered according to the intended design, or clients may have disengaged if they were not offered support soon after referral. This may mean that for some clients CLW has had a smaller impact than would have been the case if it had been delivered as intended. This could potentially mean that it is more difficult to detect any impact from the programme.

A further issue which may affect the estimated impact of CLW is the fact that the numbers of individuals in the health and disability group referred to the programme have been more heavily skewed towards those on JSA than was expected at the outset. Initial expectations were that around two-thirds of the health and disability group would be on JSA, with one-third from the ESA work-related activity group. In practice it is thought that around 95 per cent of individuals in this group referred to CLW have been on JSA, with very low numbers in the ESA work-related activity group. Given that those on JSA are likely to be closer to the labour market and have fewer barriers to work than those on ESA, it seems probable that having a higher than expected volume of referrals to CLW from those on JSA would result in an upward bias in outcomes, compared to what was anticipated at the outset.

Overall it is uncertain how the difference between the initial intentions behind CLW and how it has operated in practice are likely to affect the likelihood of being able to detect any impact from the programme. This is because some of these divergences from the intended design and caseload are likely to inflate the magnitude of any observed impacts whilst others are likely to have the opposite effect.

---

[4] Note, however, that some of these individuals may continue to receive support for a further 9 months after this point.

## 2.5    Review of suitable methods

Due to the likely differences in motivation between those who come forward to participate in CLW and those who are eligible but do not wish to participate in the trial (noted in section 2.4) the most robust approach to estimating the impact of CLW is likely to be to use data from the RCT, provided it is possible to establish that the scale of any control group contamination is low. The following sections review alternative options in the event that analysis of the RCT data is not viable. There are two potential scenarios in this case:

1. It is only possible to identify those who may be eligible for CLW based on what can be observe from administrative data sources.
2. It is possible to use MI to identify those who participate in CLW in administrative data sources.

In the first scenario we would be restricted to estimating the impact of eligibility for CLW, rather than the impact of actually participating in the programme. It is likely that a sizeable proportion of individuals who are eligible for the programme do not participate and so the impact of CLW would be diluted if the estimate of impact includes outcomes for eligible non-participants. For example, the specification for CLW showed the numbers of individuals in receipt of JSA and ESA or IB across the 12 participating boroughs. This indicated that around 150,000 individuals were claiming these benefits in 2016. However, only a proportion of these would meet the health and disability or long-term unemployment criteria. More recent figures (from August 2019) for the numbers of individuals in the 12 boroughs in the ESA work-related activity group, the UC work search group and on JSA suggested that the numbers likely to be eligible for CLW would be closer to half this figure (around 76,000), although even this would overstate the eligible population, since the number of JSA or UC claimants with health problems is not shown in published figures. If around 50,000 individuals were eligible for CLW and the aim of supporting around 20,000 individuals is achieved, approximately 40 per cent of the eligible population could be expected to participate in CLW over the programme's lifetime.

While the fact that a large proportion of the eligible population would not participate in CLW dilutes the estimated effect of the intervention, it would at least provide robust and defensible estimate of the impact of CLW on those who are eligible. If the focus was instead on estimating the impact of CLW on those who choose to participate (under scenario 2), it is likely to be difficult to identify a comparison group who would be similarly motivated to take part and would be likely to achieve similar outcomes. In this case, the impact of participation in CLW is likely to be overstated. The following sub-sections review the methods best suited to evaluating the impact of eligibility for, or participation in, CLW and highlight the main strengths and weaknesses of each approach.

### 2.5.1    Difference-in-differences (DiD)

This method would estimate the impact of CLW by comparing the change in outcomes for those eligible for the programme between the periods before and after its introduction with the change in outcomes for a similar group of individuals who meet the eligibility criteria in

comparison areas outside of the 12 participating boroughs. As outcomes are likely to vary over time with fluctuations in the labour market and other macroeconomic changes, this approach is able to adjust for any changes in outcomes that would have occurred even without the introduction of CLW.

For the approach to be robust, it is important to demonstrate that the comparison areas experienced similar trends in outcomes to the CLW areas over a period of time prior to the introduction of CLW. There must also be reason to believe that eligible individuals in both the CLW areas and the comparison areas would have continued to experience a similar trend in outcomes in the period following the introduction of CLW if the programme had not been introduced.[5]

With DiD methods it is only likely to be possible to estimate the impact of being eligible for CLW, not the impact of participating in CLW. This is because it is likely to be difficult to identify the subset of the comparison group who would have been likely to volunteer to participate in CLW if it had been available in their area.

As noted in section 2.4 an approach which involves estimating the impact of CLW by comparing outcomes for eligible individuals in the 12 participating boroughs with those for individuals who meet the eligibility criteria in non-participating areas is complicated by the fact that some eligible individuals who wish to participate in CLW will be assigned to a control group which receives Jobcentre Plus business as usual support. Also, within potential comparison areas, some eligible individuals will receive support from the WHP, whilst others are subject to Jobcentre Plus business as usual. In practice this means that a DiD approach would involve estimating the impact CLW or Jobcentre Plus business as usual compared with the WHP or Jobcentre Plus business as usual.

Provided a similar proportion of individuals meeting the criteria for CLW and in each of the client groups in the treatment and comparison areas are receiving Jobcentre Plus business as usual support, DiD methods could be used to estimate the impact of CLW. However, the comparison will be against the WHP, rather than against Jobcentre Plus business as usual. As the WHP and CLW have similar aims, the magnitude of any impact from CLW is likely to be smaller than if the comparison were made against Jobcentre Plus business as usual, and therefore it is less likely to be possible to say with certainty that CLW has had a clear impact.

## 2.5.2  Propensity Score Matching

Propensity score matching (PSM) involves first estimating the likelihood that an individual with a given set of characteristics chooses to participate in CLW. Having calculated the probability of a given individual taking part, it is possible to match treated and untreated individuals with a similar propensity and compare the outcomes of the treated with those of matched comparators. This provides an estimate of the impact of participating in CLW.

---

[5] This would ideally include ensuring that the comparison areas had a similar timetable for the roll-out of UC.

For PSM to provide a robust and credible estimate of impact it is necessary to ensure that treated and untreated individuals are well-matched on the full range of characteristics likely to determine both the likelihood that they come forward to participate in CLW and the outcomes experienced as a result. If there are unobserved characteristics which affect the probability of eligible individuals choosing to participate in CLW, or their likelihood of finding paid employment and attaining a higher rate of pay, this could mean that estimated outcomes for the matched comparison group provide a poor estimate of counterfactual outcomes for the treatment group. It is debatable whether DWP and HMRC administrative data sources alone would provide a sufficient level of detail on individuals eligible for CLW to ensure that treatment and comparison groups were well-matched on all characteristics related to both participation and outcomes.

With PSM, potentially any comparison group could be drawn from individuals living within the 12 boroughs participating in CLW who met the eligibility criteria, or from those meeting the criteria but living in other areas. If it is possible to use MI to identify those who receive support from CLW, it may be possible to draw the matched comparison group from non-participants within the 12 participating boroughs. This is likely to be preferable to constructing a matched comparison group from those living in other areas, as non-participants within the 12 areas offering CLW would be subject to Jobcentre Plus business as usual, whereas a comparison group drawn from other areas would potentially be participating in either the national WHP or Jobcentre Plus business as usual. However, irrespective of whether the comparison group are drawn from areas participating in CLW or not, it would nevertheless be important to observe all factors likely to determine participation and outcomes for the treatment and comparison groups.

### 2.5.3  Other quasi-experimental methods

It is understood that in the early stages of CLW there were delays in carrying out the initial appointments with individuals referred to the programme. Survival analysis can be used to estimate the impact of interventions where there are delays in offering support due to the volume of referrals exceeding provider capacity. With survival analysis, outcomes for individuals who are eligible to use a service but who are not offered support for an extended period can be compared with those of individuals who do receive the intervention. However, this approach rests on a sizeable number of individuals not receiving the intervention for a lengthy period. It seems unlikely that the delays experienced in the operation of CLW would be sufficient to make survival analysis viable.

A further difficulty which reduces the likely viability of survival analysis is the fact that detailed information on those referred to the service is only collected once the individual has completed a full assessment. The analysis of MI showed that even basic information, such as gender, is missing for a high proportion of those referred to the service until they have participated in the initial assessment. This means that it would be difficult to verify that those referred to CLW who had to wait many months for the initial assessment were similar to those who were assessed and started to receive support shortly after referral.

While individuals referred to CLW are taken through a detailed assessment to identify barriers to work, this is used to inform the type of support that is offered, rather than to

determine whether they should be offered any support at all. If potential participants were ranked in terms of their responses to the assessment and support were only available to those above a certain threshold, a regression discontinuity design (RDD) may have offered a further way of estimating the impact of CLW. With this approach, outcomes for those just above the threshold, and therefore receiving support from CLW, would be compared to outcomes for those just below the threshold, to estimate the impact of CLW around the margin of treatment. However, as responses to the initial assessment are not used to determine access to the CLW, a RDD approach is not viable in this case.

## 2.6   Outcome measures

Under the current evaluation timetable, the final report is due in January 2023. However, the period from October to December 2022 will be used to synthesise findings across the different strands of research and the cost-benefit analysis will be carried out between July and September 2022. This means that the impact analysis will need to start in early 2022. Participants may receive support for a period of up to 21 months following the initial assessment. Estimating the impact of CLW for the subset of individuals for whom it is possible to observe impacts at least 21 months following their start on CLW would reduce the sample for analysis, as it would only be possible to observe 21-month outcomes for those starting on the programme up to the end of 2019, allowing for lags in updating administrative data sources. To maximise sample sizes, it would be advisable to focus on outcomes over a period of 12 months following the initial assessment. It seems likely that if CLW is effective, at least some impacts would be apparent by this point. Therefore, for all individuals whose outcomes can be observed for a period of at least 12 months after the initial assessment, the following outcome measures are proposed:

■ Proportion in employment in each month following the initial assessment;

■ Number of weeks employed in the 12 months following the initial assessment;

■ Earnings in the 12 months following the initial assessment.

In addition to these outcomes, it may be feasible to look at longer-term outcomes for the subset of individuals whose outcomes can be observed over a period of 21 months following their assessment. For this subset, the following outcome measures would be likely to capture the intended impact of the programme:

■ Proportion in employment in each month following the initial assessment;

■ Number of weeks employed in the 21 months following the initial assessment;

■ Earnings in the 21 months following the initial assessment.

In addition to measures which seek to capture the primary intended effects of CLW, it may be feasible to look at the impact of the programme on benefit receipt, including the proportion on out-of-work benefits in each month following the initial assessment and the number of weeks on out-of-work benefits over the first 12 or 21 months. These outcome measures could be observed with access to DWP and HMRC administrative data.

While the intention is that CLW should affect a wider range of outcomes for participants, such as health, wellbeing and skills acquisition, it would not be possible to estimate the

impact of CLW on these measures with DWP and HMRC administrative data sources alone. It would however be important to consider likely impacts on these outcomes as part of the cost-benefit analysis, and so again, it would be extremely beneficial to gain access to information from the RCT on the impact of CLW on these outcomes.

## 2.7   Subgroup analysis

As noted previously, as well as estimating the impact of CLW on the outcomes set out in the previous section across the full range of eligible participants, the analysis will seek to explore the relative effectiveness of CLW compared to the version of the WHP operating in other areas. With access to individual-level data from the national RCT of the WHP and access to MI from other similar areas, as well as from CLW, a subgroup analysis would be used to explore whether the impact of CLW differed from that of the WHP. This would include testing the statistical significance of any apparent differences in impact between CLW and the WHP. However, as noted previously, it is less likely to be possible to detect any differences in impact between CLW and the national WHP due to the similarities between the schemes. As a result, this analysis may be unable to provide a clear insight into the relative effectiveness of CLW compared to the national WHP.

To ensure that any differences in impact are due to differences in the effectiveness of the schemes, rather than the nature of local demographic and labour market characteristics, we would seek to make comparisons with a selection of areas which are similar to the 12 boroughs participating in CLW. As mentioned in section 1.3.3, a range of publically-available data sources will be used to identify suitable comparison areas. Given the differences between the London labour market and the rest of the country, in practice this means making comparisons with other London boroughs not participating in CLW. Also, the analysis would only be meaningful if the level of control group contamination was low both within the 12 boroughs participating in CLW and in the comparison areas.

In addition to examining the overall impact of CLW and its relative effectiveness compared to the wider WHP, the aim is to estimate the impact of CLW for each of the three target groups. Analysis of MI to date suggests that the numbers in the long-term unemployed and the early access disadvantaged groups may be insufficient to detect any impact from the programme on these subsets of participants, even if one exists.[6] Nevertheless, with access to individual-level data from the RCT, the evidence that CLW is more effective for particular target groups will be assessed.

---

[6] In the period to 25 October 2019, 3,270 individuals from the health and disability group had attended an initial interview compared to 1,445 from the long-term unemployed group and 604 individuals from the early access disadvantaged group.

# 3    Summary and conclusions

## 3.1    Impact evaluation feasibility

A robust and defensible impact evaluation of CLW is most likely to be possible with access to individual-level data from the national RCT, linked to the CLW MI. The option of an impact evaluation using DWP and HMRC administrative data sources and using the CLW MI to identify programme participants would be another option if it were not feasible to access the RCT data or to link it to the CLW MI, but this is less likely to be robust. It would also be less likely to be possible to detect any impact from CLW with this approach.

## 3.2    Proposed approach to impact evaluation

### 3.2.1    Recommended approach

The most credible option for estimating the impact of CLW would be to obtain access to the RCT data on both treatment and control groups from the national evaluation of the WHP for the 12 London boroughs participating in the programme and to link it to the CLW M. In addition to this, it would be helpful to seek access to the RCT data on other London boroughs along with their MI to explore the relative effectiveness of CLW compared to the version of the WHP running in other similar areas. However, as it is less likely to be possible to detect any differences in impact between CLW and the WHP more generally, it is suggested that this analysis should be a lower priority. If the data on other areas are readily available, there may be value in carrying out the analysis, but given the likelihood that the findings will be inconclusive, it would not be worth pursuing in circumstances where it is very time-consuming or costly to negotiate data access.

All the analysis would be carried out within a regression framework to control for the characteristics individuals in the treatment and control groups and to estimate the average impact of CLW on those who choose to participate. It would also be possible to carry out a subgroup analysis to estimate the impact of the WHP on each of the three main target groups and to explore whether it is more or less effective for any particular group. However, it is likely to be more difficult to detect any impact, or differences in impact, in the case of the long-term unemployed or the early access disadvantaged groups, due to the more limited sample sizes, particularly for the early access group, where only around 600 individuals attended an initial interview in the period to the end of October 2019.

### 3.2.2    Alternatives

If it is not possible to access the individual-level RCT data from the national evaluation of the WHP, or to link it to the CLW MI, the most viable alternative is likely to be PSM, drawing the comparison group from within the 12 London boroughs participating in CLW.

In this case, the CLW MI could be used to distinguish between CLW participants and those who are either assigned to the control group or who do not choose to participate. With access to linked DWP and HMRC data on all individuals meeting the eligibility criteria for CLW within the 12 London boroughs it would be possible to use PSM to identify a matched comparison group. However, it is probable that it will not be possible to observe all the factors which are likely to determine both whether an individual chooses to participate in CLW and the outcomes they experience as a result within DWP and HMRC data alone. For this reason, this approach may overstate the impact of CLW as at least some matched comparators may not have chosen to participate in CLW and may have been less motivated to find work than participants.[7]

The other limitation of PSM relying on within-area comparators is that it would not be possible to explore the relative effectiveness of CLW compared to the WHP running in other areas. For this to be feasible it would be necessary to have access to DWP and HMRC data on the eligible population in other London boroughs, as well as a way of identifying those participating in the WHP programme in these areas. It would be possible to carry out a subgroup analysis for each of the three main client groups using PSM. However, again this would depend on sample sizes being adequate to detect any impacts, or differences in impact, between client groups.

PSM with matching to comparators within CLW boroughs is likely to be preferable to DiD analysis, as with DiD methods it would only be possible to estimate the impact of eligibility for CLW, rather than participation. As many of those eligible will not actually participate (either because they are assigned to the control group, or because they do not volunteer for CLW), it would be more difficult to detect any impact from CLW on the eligible population. Also, a DiD approach would effectively estimate the impact of CLW and Jobcentre Plus business as usual within the 12 London boroughs compared with the impact of the WHP and Jobcentre Plus business as usual in other areas. As CLW and the national WHP have similar aims, CLW would need to have a substantial impact compared to the national WHP for any impact to be manifest.

---

[7] If it were possible to identify those assigned to the control group in the 12 CLW boroughs, as well as those receiving support, that would substantially reduce the risk of unobserved differences between the treatment and matched comparison group driving any differences in outcomes, as the matched comparison group could be drawn from those who had volunteered to participate in CLW, but had been denied the treatment. However, our understanding is that this is unlikely to be possible under current data sharing agreements with Jobcentre Plus.

## 3.3    Next steps

This section summarises the immediate priorities for action in relation to the impact evaluation of CLW:

1. The first task is to establish whether it is likely to be possible to obtain access to data from the national RCT of the WHP for the 12 boroughs participating in CLW and for the wider Greater London area.

2. If access to the RCT data is not possible, the next task will be to explore whether it is likely to be feasible to obtain DWP and HMRC administrative data extracts on those eligible for CLW in the 12 participating boroughs and the Greater London area.

    a. This should include the feasibility of adding a flag to the DWP and HMRC data extracts to identify records for CLW participants and WHP participants.

    b. If current data sharing agreements permit the flagging of records for members of the control group as well, this would be highly desirable.

3. Assuming it is necessary to use the DWP and HMRC administrative data for the evaluation, the next step would be to draft a proposal for EDSAC outlining the reasons for the data request and the evaluation approach. It would also be necessary to carry out a data protection impact assessment. Any data request would include a list of all the variables required for the analysis.

4. Provided EDSAC grant access to the DWP and HMRC administrative data, the evaluation would then be carried out in the way outlined in section 3.2.2.

5. The aim would be to obtain access to the RCT or administrative data linked to the CLW MI and MI from other areas in early 2022. This will make it possible to complete all evaluation activities and supply the final report in January 2023.

Whilst it is not necessary or desirable to start work on the impact analysis for around 18 months or more, it is strongly recommended that Central London Forward initiates discussions with DWP and other London boroughs regarding data access at the earliest possible opportunity. The process of agreeing access in principle and making an application to EDSAC is likely to be a lengthy one and timeframes for preparing and supplying data extracts will depend on staff availability at DWP. It is therefore desirable to ensure that access arrangements are agreed well in advance of the data being required for the evaluation.