Institute for Fiscal Studies

Ethnic Parity in Labour Market Outcomes for Benefit Claimants

IFS Working Paper W09/04

Claire Crawford
Lorraine Dearden
Alice Mesnard
Jonathan Shaw
Barbara Sianesi

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Ethnic parity in labour market outcomes for benefit claimants[1]

Claire Crawford (IFS and IoE)
Lorraine Dearden (IFS and IoE)
Alice Mesnard (IFS and CEPR)
Jonathan Shaw (IFS)
Barbara Sianesi (IFS)

November 2008

Abstract

A significant gap exists in the UK between the employment rate for Ethnic Minorities and that for Whites. From a policy perspective, it is important to know whether this gap is due to differences in the characteristics of White and Ethnic Minority groups (which reduce the employability of Ethnic Minority groups relative to Whites) or whether it results from some form of discriminatory behaviour in the labour market. In this paper, we use administrative data to estimate ethnic differences in employment and benefit receipt amongst individuals who began claiming a Jobcentre Plus benefit in 2003. In contrast to much of the previous UK literature, we use a number of different quantitative techniques to estimate this gap, and show that in a lot of cases the estimates obtained are very sensitive to the techniques used. We argue that for the questions we are interested in and the data we have, propensity score matching methods are the most robust approach to estimating ethnic parity. We compare this preferred approach with estimates derived using alternative approaches commonly used in the literature (generally regression-based techniques) to determine the extent to which more straightforward methods are able to replicate those produced by matching. In many cases, it turns out not to be possible to calculate satisfactory quantitative estimates even with matching techniques: the characteristics of Whites and Ethnic Minorities are simply too different before the Jobcentre Plus intervention to reliably estimate the parameters of interest. Moreover, for a number of the groups, results seem to be very sensitive to the methodology used. This calls into question previous results based on simple regression techniques, which are likely to hide the fact that observationally different ethnic groups are de facto being compared on the basis of parametric extrapolations. Two groups for which it was possible to calculate reasonably reliable results are incapacity benefit (IB) and income support (IS). For these groups we find that large and significant raw penalties almost always disappear once we appropriately control for pre-inflow background and labour market characteristics. There is also a good degree of consistency across methodologies.

JEL classification numbers: J08, J15
Key words: ethnic, employment, benefit, discrimination, matching

[1] The authors would like to thank the Department for Work & Pensions who funded this work and provided the administrative data that we used. Corresponding author: Jonathan Shaw (jonathan_s (at) ifs.org.uk).

1. Introduction

Significant ethnic differences or lack of ethnic parity in employment rates have existed in Great Britain for several decades (Blackaby et al, 1999; Cabinet Office, 2003; Heath & Cheung, 2006; Berthoud & Blekesaune, 2006; National Audit Office, 2008). The National Audit Office (2008) shows that over the last twenty years, the employment gap between Ethnic Minorities and Whites has fluctuated between 12.5 percentage points in 1989 and 20 percentage points in 1994. Since 1994, however, the gap has fallen steadily such that by the third quarter of 2007 it stood at 13.2 percentage points, with 74.8% of the White working-age population in Great Britain in employment compared with 61.6% of the equivalent Ethnic Minority population (Labour Force Survey).

As might be expected, this overall gap conceals considerable diversity in employment rates across ethnic groups (Blackaby et al, 1997; Blackaby et al, 1999; Heath & Cheung, 2006; National Audit Office, 2008). For example, Indian individuals have employment rates that are only slightly lower than those for White individuals, whereas Pakistani and Bangladeshi individuals have employment rates that are considerably lower. In 2007, only 47% of Pakistanis and 44% of Bangladeshis were in employment, compared to around 75% of Whites.[2] Interestingly, the employment gap for Black Caribbean individuals appears to have narrowed over time: while Blackaby et al (1997) report a gap of 13.9 percentage points in 1991, this had closed to around 6 percentage points in 2007 (National Audit Office, 2008).

These observed raw differences between Ethnic Minorities and Whites suggest that some Ethnic Minority groups may be facing significant obstacles to finding and keeping a job. In 2003, the UK government stated that '*in ten years' time, Ethnic Minority groups should no longer face disproportionate barriers to accessing and realising opportunities for achievement in the labour market*' (Cabinet Office, 2003). To achieve such aims, policymakers clearly need to be well-informed about the exact form and extent of any such barriers: in particular, they need to know whether ethnic differences in employment outcomes can be attributed to differences in the characteristics of White and Ethnic Minority groups (which reduce the employability of some ethnic groups relative to Whites), or whether they result from some form of discriminatory behaviour in the labour market. The appropriate policy response clearly differs according to whether the gap is due to differences in characteristics and/or discriminatory behaviour.

To be able to identify these effects, it is essential to compare each Ethnic Minority group with an otherwise-identical White group. Most previous studies have attempted to do this using simple multivariate regression techniques and have relied on being able to observe all characteristics likely to be relevant in determining employment outcomes. In this literature, any unexplained (residual) difference is then attributed to discrimination and the response of Minorities to discrimination. For example, Blackaby et al (1999) found that over half of the observed difference in unemployment rates between Whites and West Indians can be attributed to differences in characteristics, while Indians, Pakistanis and Bangladeshis actually have *better* characteristics (i.e. ones that make them more employable) than Whites. Hence discrimination may be a factor for some ethnic groups but not for others.

However, it is now well known that simple regression techniques may be biased if: (i) there is incomplete overlap in the *range* of values taken by the control variables for the treatment group (in this case, Ethnic Minorities) and the control group (Whites) (the so-called common support problem); (ii) the distribution of characteristics differs between the two groups (across a similar range of values), and (iii) the effect of the treatment (ethnicity) varies across individuals. Given the differences in characteristics between Ethnic Minorities and Whites observed in most studies, this is likely to be particularly problematic when trying to identify the impact of ethnicity on any outcome of interest.

[2] These figures are largely driven by very low rates of employment amongst women in these groups (of around 26% for Bangladeshis and 28% for Pakistanis) (National Audit Office, 2008).

To our knowledge no study has looked at how sensitive estimates of ethnic penalties or ethnic parity are to the quantitative method of estimation used. In this paper, we borrow heavily from the economic evaluation literature to address this issue. We argue that given our data and questions of interest, the most appropriate technique (that is the one with the most tenable assumptions) is matching. Like OLS, it depends on observing all those key determinants of the outcomes of interest that differ between ethnic groups; we argue that the detailed labour market histories we construct should substantially reduce biases due to characteristics we do not observe. Matching deals with a number of the shortcomings of OLS: it compares only comparable individuals, weighting them correctly and allowing the treatment effect to vary across individuals. We then compare these results with those derived using three alternative approaches – OLS, fully interacted OLS, as well as difference-in-differences, a method which allows for selection on unobservables under certain assumptions.

Using administrative data for UK benefit claimants[3], we find that often it is simply not possible to calculate reliable estimates of the extent of ethnic parity: Minorities are just too different from Whites to be able to compare them satisfactorily. Moreover, in lots of cases, the different estimation techniques give wildly differing answers, both in terms of sign and of significance of the penalty. This casts doubt on the reliability of previous work that has relied on a single technique without any consideration of how comparable Whites and Ethnic Minorities are. Parametric methods such as OLS, or even fully interacted OLS can easily hide from the analyst the fact that observationally different ethnic groups are de facto being compared on the basis of extrapolations purely based on the imposed functional form.

There are two groups, however, for which we are able consistently to calculate reliable results: claimants of Incapacity Benefit (IB) and claimants of Income Support (IS). For these groups, we find that controlling for relevant labour market and demographic characteristics eliminates almost all ethnic penalties. In other words, the gap in labour market outcomes can be explained entirely by differences in characteristics. There is also a good degree of consistency across methodologies.

This paper proceeds as follows: Section 2 outlines our methodology; Section 3 describes the data that we use; Section 4 describes how we selected our samples; Sections 5 and 6 present our results; Section 7 concludes.

2. Methodology

The aim of this paper is to determine the extent to which there is *ethnic parity* in the labour market outcomes of benefit claimants. A natural definition for ethnic parity is when there is no difference, on average, between the outcome for an Ethnic Minority claimant and the outcome for an 'otherwise-identical' White claimant. Where parity does not exist, there will be either an ethnic penalty – if Ethnic Minority customers experience worse outcomes than otherwise-identical White customers – or an ethnic premium – if Ethnic Minority customers experience more favourable outcomes than otherwise-identical White customers. This definition is an ideal one, and our aim is to approximate it as closely as possible empirically.

Since Whites and Ethnic Minorities may differ in terms of characteristics other than ethnicity that affect labour market outcomes, simply using the observed difference in outcomes is likely to provide a biased estimate of ethnic parity. To avoid such a bias, one needs to deal with any differences between ethnic groups in terms of characteristics that may affect their outcomes, such as the individual's family background, education and labour market history.

A number of techniques are available to control for observed differences in characteristics. Within this group of estimators, the technique which is most likely to provide reliable estimates of ethnic

[3] We consider claimants of Jobseeker's Allowance, Incapacity Benefit and Income Support, and participants in New Deal for Lone Parents, New Deal 25 Plus and New Deal for Young People. More details about our data and samples can be found in Sections 3 and 4.

parity and which therefore represents the primary approach used in this paper is matching (see e.g. Imbens, 2004, for an extensive review). Matching re-weights Whites to look as similar as possible to Minorities in terms of observable factors X that might affect labour market outcomes. The estimate of the ethnic penalty/parity/premium is then obtained by comparing the mean labour market outcome of the Ethnic Minority group with the mean outcome of the appropriately re-weighted White comparison group. To ensure comparability, Ethnic Minority customers for whom no suitable White comparator can be found (i.e. those who fall outside the 'common support') are excluded from the analysis.

Matching is more flexible than most of the alternative techniques that are available. It can also provide a series of diagnostic tests that can be used to analyse how well White and Minority samples have been matched. This is important, because when the White sample cannot be re-weighted satisfactorily, it is not clear that any method will provide unbiased estimates of ethnic parity. In this paper, we use a set of criteria to determine when a result should be considered reliable, in the sense that reasonably balanced samples have been achieved. These criteria are heuristic in nature and in line with the spirit of the different diagnostic tests used in the matching literature to ensure the groups in question are sufficiently well balanced (see, e.g., Blundell et al, 2005). The criteria are:

- Given the special importance attached in the literature to the pre-inflow history for the outcome of interest (for example, employment history for an estimate of employment parity) we require that it be successfully balanced at 5 per cent level.
- We then consider how many of roughly 120 pre-inflow characteristics still differ significantly between Whites and Minorities at the 5 per cent level after matching.
- If a substantial number of these characteristics do remain individually unbalanced, we consider their joint importance in predicting whether an individual is of Ethnic Minority or White origin (this is done in terms of the pseudo-r^2 of a probit on matched samples of Ethnic Minority group on the observables)
- As a final indicator, we consider the median bias, which gives the median percentage distance between the matched groups in terms of observed characteristics (median across all matching variables).[4]

Unlike matching, standard regression techniques (OLS) may produce biased estimates of ethnic parity because they

- may implicitly extrapolate across non-comparable individuals (the common support problem);
- may not weight comparable individuals correctly;
- typically assume that the ethnic penalty is constant across individuals.

Since previous studies estimating ethnic parity have often relied upon simple regression methods, an important contribution of this paper is in assessing the reliability of such techniques by comparing the results they produce with those produced by matching. Of course, some of the restrictions imposed by OLS can be relaxed. In particular, a fully interacted OLS model ('fully interacted linear matching' – FILM) allows the ethnic penalty to vary according to each observable characteristic rather than constraining it to be constant. Indeed, previous research has shown that FILM can often produce estimates that are very similar to those obtained by matching (Blundell *et al*, 2005). Nevertheless, the diagnostics provided by matching allow one to investigate the extent to which groups are truly comparable, which may be hidden in both simple and interacted OLS models. We investigate such issues empirically in Section 5.

It should however be noted that both matching and regression techniques are only able to control for *observed* differences between Whites and Minorities, so their reliability depends crucially on the range and quality of characteristics observed. To the extent that *unobserved* differences are

[4] In addition we restrict our analyses to ethnic minority (sub)groups with at least 400 individuals. We also imposed requirements on the maximum share of the minority sample that can be lost to common support, but this did not prove binding for the results we present.

important for explaining labour market outcomes, these will show up in the ethnic parity estimates. Our choice of sampling frame (see Section 4) and the rich set of characteristics we control for are designed to minimise the extent to which unobservables are important. Previous literature has indicated the potential for detailed labour market histories (like those we construct) to eliminate much of the bias due to unobservables (see for example, Heckman and Smith, 1999, Heckman, Ichimura, Smith and Todd, 1998 and Heckman, LaLonde and Smith, 1999).

An alternative technique, conditional difference-in-differences (DiD), is in principle able to deal both with *observed* differences in characteristics and with *unobserved* differences that remain constant over time. This technique involves controlling for observed characteristics, calculating – separately for Whites and Minorities – the average change in outcomes over time, and then taking the difference. We present some results that use conditional difference-in-differences implemented in a flexible regression framework (FILM) and compare these results with our preferred set of results. The reason for not using DiD as our preferred technique is that entitlement conditions for some of the benefits and programmes we consider require the individual to be in a given labour market status for a certain time prior to inflow, preventing a meaningful application of the technique. For example, New Deal 25 plus (a programme designed to help individuals over the age of 25 into work) is compulsory for those who have been claiming Jobseeker's Allowance (unemployment benefit) for 18 of the previous 21 months. An empirical issue with DiD is how best to deal with multiple pre-inflow time periods. Since the literature to date has not provided any established solution, we use either a 12-month moving window to capture seasonal effects or an average over the 12-month period before inflow. We present results using the former approach, but it turns out to make little difference to the results.

One important caveat of our study relates to previous (i.e. pre-inflow) discrimination, which is largely unobservable. If there has been labour-market discrimination in the past, then members of an Ethnic Minority group with the same employment history as their matched White counterparts might represent a higher 'slice' of the (unobserved) ability distribution amongst their ethnic group. As a result, any investigation risks comparing more able Ethnic Minorities with less able Whites, which would lead to an underestimate of any ethnic penalty and to an overestimate of any ethnic premium. A similar argument can be made in terms of other observables one would like to control for, such as education. If the selection process into education differs between Ethnic Minorities and Whites, then comparing Ethnic Minority and White individuals with the same level of education might still leave some unobservable differences unaccounted for.[5]

3. Data

Our strategy for estimating ethnic parity requires that we have information about labour market outcomes and that we control for all pre-inflow characteristics likely to affect these outcomes and to differ between ethnic groups. The key variables we control for across all benefits and programmes are:

- Age
- Gender
- Month of inflow to relevant benefit or programme
- Detailed employment and benefit history
- Previous participation in voluntary employment programmes
- Basic Skills attendance
- Jobcentre Plus district
- Residence in one of a group of 272 deprived wards

[5] The same issues arise in the classic case of testing for discrimination on the grounds of gender using a wage equation that controls for observed characteristics such as labour market experience and education. If selection into the labour market is non-random (for instance if it is the case that only the most able of women enter the labour market or indeed obtained higher levels of education), then men and women with the same observed experience and education will differ in terms of unobservables. As far as we are aware, this issue has not been fully addressed in this literature.

- Education (proxy)
- Housing (proxy)
- Travel-to-work-area unemployment rate

We also control for a number of variables that are only available or relevant for a subset of the benefits and programmes. These include partner, number of children, disability and illness.

Our main data source is the Work and Pensions Longitudinal Study (WPLS), an administrative dataset of benefit and employment spells for all individuals who have been on a Department for Work and Pensions (DWP) programme or benefit since June 1999. The extract of the data that we use contains spells up to early 2005. Because the WPLS data is an administrative source, it contains the entire populations we are interested in. Compared to survey data, we thus avoid problems with sample attrition, are able to use much more detailed ethnic breakdowns and provide estimates with higher precision.

We used the WPLS to construct labour market outcomes and histories. We consider two labour market outcomes, both assessed month-by-month: the probability of being in employment (E) and the probability of being on benefits (B).[6] In any one month (30-day period), individuals are defined as being employed (on benefits) if they are employed (on benefit) for at least half of the month.[7] Defining outcomes in this way explicitly takes account of the sustainability of any benefit exit or job entry and is therefore preferable to a measure that only considers first destination. Both benefit and employment outcomes are analysed to capture the extent to which individuals who are not employed remain on benefits and the extent to which individuals who are employed still collect benefits (in particular, Income Support). Individuals are followed and their outcomes measured for 12 months following inflow.

We constructed three years' worth of labour market history, including month-by-month employment and benefit status for the six months prior to inflow, as well as variables describing the proportion of time employed and the proportion spent on benefits over the two-and-a-half years prior to that. We also used the WPLS to create indicators for the fraction of time of past participation in voluntary employment programmes (as a crude indicator of willingness to improve ones circumstances) and past participation in Basic Skills (a programme designed to address basic literacy, numeracy and IT skills).

The WPLS also contains a limited number of background characteristics: ethnicity, date of birth, sex and postcode (that we used to identify Jobcentre Plus district and residence in one of a group of 272 deprived wards).

Information from the WPLS was supplemented with information from other administrative sources, namely the National Benefits Database (NBD) and New Deal administrative datasets. The information available varies across benefits and programmes, but includes children, partner, disability and illness. None of these sources contained reliable information on education or wealth. Since both are likely to be important determinants of labour market outcomes, we created proxies using local-area data from the 2001 Census.[8] We also used the Census to provide information about local labour market conditions (specifically, travel-to-work area unemployment rates).

---

[6] The benefits included in this definition are IB, IS, Jobseeker's Allowance (JSA), compensation whilst on a New Deal option, Basic Skills and Work-Based Learning for Adults. See Crawford et al (2008) for more details.

[7] We experimented with other definitions, but because most individuals are employed (or on benefit) either for all of the month or none of it, using an alternative threshold makes little difference to the results.

[8] To proxy education, we used SOA-level information (around 750 households) on the education level of unemployed or inactive individuals to calculate (separately for whites and minorities) the proportion whose highest qualification was: no qualifications, level 1 qualifications, level 2 qualifications, level 3 qualifications, level 4 qualifications or higher, and unknown qualifications (six categories). For wealth, we used the proportion of each ethnic group living in council or other social-rented housing, measured at Ward level (around 2,500 households).

Before moving on to describe our sample definitions, there are some limitations of the WPLS that we must note. First, there is no record of how long individuals have been living in the UK. In comparing Ethnic Minorities and Whites, immigration may be a considerable issue. An individual who has just arrived in the UK and starts claiming benefits or joins a programme will never previously have appeared in the WPLS (i.e. they will be recorded as never having previously been employed or on benefit), whereas they may, in reality, have been employed or on benefit in another country.

We have tried to minimise the problems this causes by using the time that each individual has been observed in the WPLS as the reference for calculating pre-inflow labour market histories and by controlling for whether labour market status is known or unknown. Still, the underlying assumption remains that what Ethnic Minority individuals do when not in the WPLS is, on average, equivalent (at least in terms of labour market performance) to what their observationally similar White counterparts do when not in the WPLS.

A related problem is that recent migrants to the UK may not speak English very well and are therefore likely to find it much harder to get a job than natives. At the same time, finding appropriate White comparators for them is likely to be difficult. Where possible, information on basic skills (including language) needs was used to remove such individuals from the analysis; the fraction affected was small.

A second limitation of the WPLS is that employment spells for individuals on low earnings may be missing. This is because the data are derived from employer income tax returns which are only compulsory for employees earning enough to be subject to income tax. Although some employers submit returns for all employees regardless of their earnings, this is not always the case. Therefore, individuals earning below the income tax threshold may appear to be unemployed when, in fact, they do have a job. This is problematic if it differs by ethnicity – which it might do if, say, Ethnic Minorities are more likely to work for small employers who are less likely to submit forms for employees below the income tax threshold. There is no way of telling how quantitatively important this issue is.

4. Groups of interest and sampling frame

We consider six main Jobcentre Plus benefits and programmes:

- Incapacity benefit (IB): paid to individuals who are assessed as being incapable of work and who meet certain National Insurance contributions conditions.
- Income support (IS): a benefit for individuals on low income; usually claimants are lone parents, sick or disabled, or carers.
- Jobseeker's allowance (JSA): a benefit paid to individuals of working age who are unemployed, or who work fewer than 16 hours per week and are looking for full-time work.
- New Deal for Lone Parents (NDLP): a voluntary programme whose aim is to encourage lone parents to improve their work prospects and help them into work.
- New Deal for individuals aged 25 plus (ND25plus): a programme to help unemployed individuals aged 25 and over to find and keep a job. Participation is compulsory for individuals who have been claiming JSA for at least 18 of the previous 21 months.
- New Deal for Young People (NDYP): similar to ND25plus except that it is targeted on individuals aged 18-24. Participation is compulsory for those who have been claiming JSA for at least six months.

We also present results for Jobcentre Plus overall, which combines these six benefits and programmes, together with a number of smaller programmes: New Deal for Disabled People

(NDDP), New Deal for Musicians (NDfM), Basic Skills, Work-Based Learning for Adults (WBLA), Employment Zones (Ezones) and Ethnic Minority Outreach.[9]

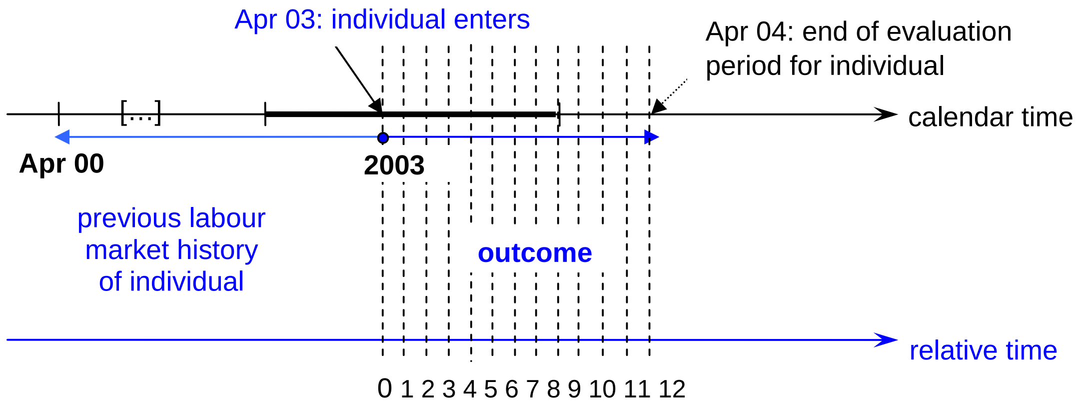
For each of these groups, the sampling frame was all individuals who

(i) started the relevant benefit or programme during 2003;[10]
(ii) on the start date, belonged to the appropriate age range for the benefit or programme in question and were aged no more than 57; and
(iii) did not have a basic skills language need.

We thus focus on the inflow of individuals starting a relevant benefit or programme during 2003. Calendar year 2003 was chosen to allow for three years' worth of labour market histories and one year of outcomes,

Figure 1 shows the sampling frame diagrammatically.

Figure 1 Individual A enters the programme in 2003



5. Overall results

Table 1 presents overall results for Jobcentre Plus and the six main benefits and programmes. Alongside the raw differences, the table sets out estimates of ethnic parity constructed using four alternative methodologies – simple OLS, FILM (which allows the penalty to vary with observed characteristics), kernel matching (our preferred technique) and conditional difference-in-differences (implemented in a flexible regression framework and taking the difference for any given outcome month relative to 12 months previously). All results are for 12 months after inflow (i.e. the end of our outcome window).

The Matching Quality column gives matching diagnostics, providing an indication of how successful re-weighting has been in making Whites look like Minorities across all observed dimensions. CS(x) indicates that x per cent of the Minority sample had to be dropped because it was not possible to find comparable Whites with whom they could be matched (i.e. it indicates the extent of the common support problem). UC(y) means that y covariates remained unbalanced after matching. UH(E,B) indicates that employment and benefit histories remain unbalanced after matching. R2(r) is the pseudo-r^2 from a probit (on matched samples) of Ethnic Minority group on

[9] For more details about these programmes, see Crawford et al (2008).
[10] For our IS and IB samples, the requirement was to have attended a Work-Focused Interview (WFI) during 2003 and that this WFI took place no more than six months after the benefit start date. WFIs are interviews carried out by Jobcentre Plus advisors to provide claimants with information and guidance about returning to work. Our analysis focuses on individuals who have attended WFIs because they are more likely to have ethnicity recorded.

the observables and MB(s) gives the median percentage distance between the matched groups in terms of observed characteristics.

Table 1 Overall results

	Number of Minorities	Outcome	Raw	OLS	FILM	Kernel matching	DiD	Matching quality
Jobcentre Plus	210,360	Employment	-0.089**	-0.004**	-0.011**	0.038**	-0.022**	CS(2)UC(94)UH(E,B)R2(0.141)MB(3.2)
			(0.001)	(0.001)	(0.002)	(0.004)	(0.002)	
		Benefit	0.059**	0.013**	0.021**	-0.006	0.034**	
			(0.001)	(0.001)	(0.002)	(0.004)	(0.002)	
IB	5,660	Employment	–0.070**	0.004	0.006	0.006	0.008	CS(0)UC(28)R2(0.017)MB(0.9)
			(0.006)	(0.006)	(0.006)	(0.008)	(0.007)	
		Benefit	0.064**	0.011	0.010	0.007	0.006	
			(0.006)	(0.007)	(0.007)	(0.008)	(0.010)	
IS	8,560	Employment	–0.028**	0.014**	0.017**	0.021**	0.022**	CS(0)UC(35)R2(0.019)MB(1.2)
			(0.005)	(0.005)	(0.005)	(0.007)	(0.006)	
		Benefit	0.019**	–0.017**	–0.019**	–0.020**	–0.042**	
			(0.005)	(0.006)	(0.006)	(0.006)	(0.008)	
JSA	159,500	Employment	-0.090**	-0.008**	-0.015**	0.037**	-0.025**	CS(2)UC(104)UH(E,B)R2(0.143)MB(2.9)
			(0.001)	(0.002)	(0.002)	(0.005)	(0.002)	
		Benefit	0.076**	0.023**	0.030**	-0.015**	0.041**	
			(0.001)	(0.002)	(0.002)	(0.005)	(0.002)	
NDLP	11,040	Employment	-0.038**	-0.009	-0.006	0.000	-0.014*	CS(0)UC(46)UH(E,B)R2(0.029)MB(1.5)
			(0.005)	(0.005)	(0.006)	(0.007)	(0.007)	
		Benefit	0.080**	0.031**	0.027**	0.021**	0.022**	
			(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	
ND25plus	13,080	Employment	-0.004	0.016**	0.014**	0.030**	-0.007	CS(0)UC(47)R2(0.102)MB(1.7)
			(0.004)	(0.005)	(0.005)	(0.008)	(0.006)	
		Benefit	-0.040**	-0.023**	-0.013*	-0.021*	-0.005	
			(0.005)	(0.005)	(0.006)	(0.009)	(0.007)	
NDYP	26,960	Employment	-0.030**	0.006	0.002	0.066**	-0.025**	CS(1)UC(79)UH(B)R2(0.223)MB(4.1)
			(0.003)	(0.004)	(0.005)	(0.014)	(0.007)	
		Benefit	-0.033**	0.008	0.023**	0.052**	0.027**	
			(0.003)	(0.004)	(0.006)	(0.014)	(0.007)	

Notes:

1. Two stars (**) and one star (*) indicate that the difference is statistically significant at the 1 and 5 per cent level of significance respectively.
2. Figures in parentheses are standard errors. For kernel matching they are calculated using an analytical approximation that ignores the fact that the propensity score is estimated. Due to the very large sample sizes in administrative data, it was not computationally feasible to bootstrap standard errors.
3. CS(x) means that x per cent of the Ethnic Minority sample was lost to common support. UC(y) means that y covariates remained unbalanced after matching. UH(E,B) indicates that employment and benefit histories remained unbalanced after matching
4. Sample sizes have been rounded to the nearest 20.

There are two key points to take from this table. First, given our heuristic criteria set out above, only the results for IB and IS can be considered reliable. For all other groups except ND25plus, labour market histories are not satisfactorily balanced, and for all other groups except NDLP, the pseudo r^2 is above 10 per cent. In these cases, Whites and Minorities seem simply too different to be reliably compared. Crawford et al (2008) show that reliability problems persist when results are calculated for subgroups defined by sex, ethnicity and region.

Second, consistency across methods varies considerably. For four of the groups (IB, IS, NDLP and ND25plus), different methods give very similar results. The sign and significance of the estimates is usually the same and the point estimates are close (and substantially different to the raw estimates). The only exception is the difference-in-differences results for ND25plus. For the other three groups (Jobcentre Plus, JSA and NDYP), however, different techniques do not give the same answer. For example, for Jobcentre Plus, OLS, FILM and difference-in-differences all suggest employment and benefit penalties significant at the 1 per cent level, whereas matching indicates a significant 4 percentage point employment premium and insufficient evidence to reject the hypothesis of benefit parity. Likewise, for NDYP, kernel matching indicates a 6.6 percentage point employment premium and difference-in-differences a 2.5 percentage point employment penalty (both significant at the 1 per cent level). Parity in employment outcomes is not rejected by either OLS or FILM. For the benefit outcome, OLS does not reject parity, while the other three methods suggest significant employment penalties, though the size of the point estimate is much larger for matching than for FILM or DiD.

From this table, it might be tempting to conclude that, when results seem reliable, all the different techniques give the same answer. Unfortunately, as Crawford et al (2008) show, this does not hold for all subgroups: there are cases where matching diagnostics suggest the result is reliable, but alternative methods give qualitatively different answers. This is somewhat surprising given earlier work suggesting that FILM often produces results that are very similar to those obtained by matching (Blundell *et al*, 2005).

These findings raise important questions over the reliability of previous regression-based estimates of ethnic parity. When matching cannot construct a suitable comparison group, it seems unlikely that any of the other techniques we employ will provide unbiased estimates. Results can also be sensitive to the estimation technique. Since matching avoids a number of biases associated with OLS, this serves as a warning against relying on simple regression techniques to estimate the extent of ethnic parity in labour market outcomes. Earlier work that fails to address directly the issue of actual comparability of ethnic groups may therefore be misleading.

6. IB and IS results

As Table 1 shows, not all results were deemed unreliable. In particular, there are satisfactory results for IB and IS. In this section, we present more detailed results for these two groups, including breakdowns by sex and ethnic group (see Crawford et al, 2008 for detailed results for the other benefits and programmes in Table 1).

First, however, it is worth taking a step back and considering the characteristics of the two samples. Table 2 below shows the ethnic breakdown. In total, there are 72,600 claimants in the IB sample and 86,700 in the IS sample (these figures exclude individuals whose ethnicity is missing – see notes to Table 2). This is considerably lower than the number of individuals claiming IS or IB during 2003 because we consider only those flowing onto benefit, not the existing stock. Whites make up 92.2 per cent of the IB sample, Blacks 2.4 per cent, and Asians 3.6 per cent. For IS, the ethnic breakdown is similar, though with a slightly higher fraction of Minorities (9.9 per cent). Most of the difference is due to a greater number of Black individuals, who make up 3.9 per cent of the IS sample. Although not shown in the table, there is very little difference in ethnic composition between males and females (females make up just under 40 per cent of the IB sample and just under 60 per cent of the IS sample).

Table 2 Ethnic breakdown of IB and IS samples

Ethnic subgroup	IB		IS	
	%	*Number*	*%*	*Number*
White	92.2	66,920	90.1	78,140
Ethnic Minority	7.8	5,660	9.9	8,560
Black	2.4	1,760	3.9	3,360
Black Caribbean	1.1	800	1.8	1,520
Black African	1.0	720	1.6	1,420
Other Black	0.3	240	0.5	420
Asian	3.6	2,600	3.9	3,380
Indian	1.2	880	1.0	900
Pakistani	1.8	1,320	2.2	1,920
Bangladeshi	0.2	160	0.3	240
Other Asian	0.3	240	0.4	320
Other	1.8	1,300	2.1	1,820
Mixed	0.5	340	0.7	580
Chinese	0.1	80	0.1	100
Other Ethnic Group	1.2	900	1.3	1,160
All groups	100.0	72,600	100.0	86,700

Notes:
1. Individuals of unknown ethnic origin have been excluded. 8.4 per cent of the IB sample and 7.7 per cent of the IS sample are of unknown ethnic origin.
2. Sample sizes have been rounded to the nearest 20.

Whites and Minorities in the two samples differ considerably in terms of a number of characteristics that are likely to be important for labour market outcomes. For our sample of IB claimants, Table 3 shows mean values for some key characteristics. Ethnic Minorities are on average younger than Whites and have more children; they are more likely to have exhibited a basic skills need and participated in a voluntary programme in the three years prior to inflow, to be claiming Income Support at WFI date (used as a proxy for low income) and to live in higher unemployment areas. In terms of labour market histories, Ethnic Minorities have, on average, spent a smaller proportion of the three years prior to inflow in employment and a larger proportion on benefits than Whites. There is also considerable variation within the Ethnic Minority sample. For example, Blacks on average spent a larger proportion of the three years prior to WFI date on benefits than Asians.

Table 3 IB: mean characteristics of the sample by ethnicity

	All	White	Difference relative to White			
			Ethnic Minority	Black	Asian	Mixed, Chinese and Other
Female	0.391	0.389	0.002	-0.016	0.004	0.022
Age at inflow	36.7	36.7	0.927**	1.642**	0.078	1.653**
Number of kids at inflow	0.010	0.010	-0.006*	0.008**	-0.020**	0.002
Proportion of time employed, years 1–3 before inflow	0.459	0.464	0.113**	0.096**	0.120**	0.120**
Proportion of time on benefits, years 1–3 before inflow	0.460	0.458	-0.106**	-0.175**	-0.049**	-0.125**
Participated in a voluntary programme before inflow	0.057	0.057	-0.015**	-0.035**	0.001	-0.021**
On Income Support at inflow	0.560	0.558	-0.097**	-0.154**	-0.044**	-0.128**
Unemployment rate in travel-to-work area	0.054	0.054	-0.001**	-0.002**	0.001**	-0.001*

Notes:
1. Two stars (**) and one star (*) indicate that the difference is statistically significant at the 1 and 5 per cent level of significance respectively.

Table 4 presents the same figures for our sample of IS claimants. Ethnic Minorities as a whole are more likely than Whites to be female, older, married/cohabiting and disabled; they tend to have more children and tend to live in higher unemployment areas. On average, they have also spent a smaller proportion of the three years prior to inflow in employment and a larger proportion on benefits. Like the IB sample, there is significant variation within the Ethnic Minority sample. For example, Asians were, on average, employed for a considerably smaller proportion of the three years prior to inflow than Blacks.

Table 4 IS: mean characteristics of the sample by ethnicity

	All	White	Difference relative to White			
			Ethnic Minority	Black	Asian	Mixed, Chinese and Other
Female	0.582	0.581	-0.017**	-0.073**	0.014	0.030*
Age at inflow	33.0	32.9	-0.559**	0.006	-1.422**	0.007
Married/cohabiting	0.106	0.098	-0.053**	0.063**	-0.184**	-0.021**
Disabled	0.252	0.246	-0.008	0.052**	-0.044**	-0.053**
Number of kids at inflow	0.843	0.821	-0.218**	-0.113**	-0.455**	0.029
Age of youngest child at inflow	5.209	5.221	0.050	0.032	0.011	0.156*
Proportion of time employed, years 1–3 before inflow	0.379	0.382	0.057**	-0.017*	0.127**	0.065**
Proportion of time on benefits, years 1–3 before inflow	0.517	0.515	-0.056**	-0.062**	-0.039**	-0.077**
Participated in a voluntary programme prior to inflow	0.114	0.118	0.003	-0.041**	0.046**	0.006
Unemployment rate in travel-to-work area	0.054	0.054	-0.002**	-0.003**	-0.000	-0.001**

Notes: see notes to Table 3

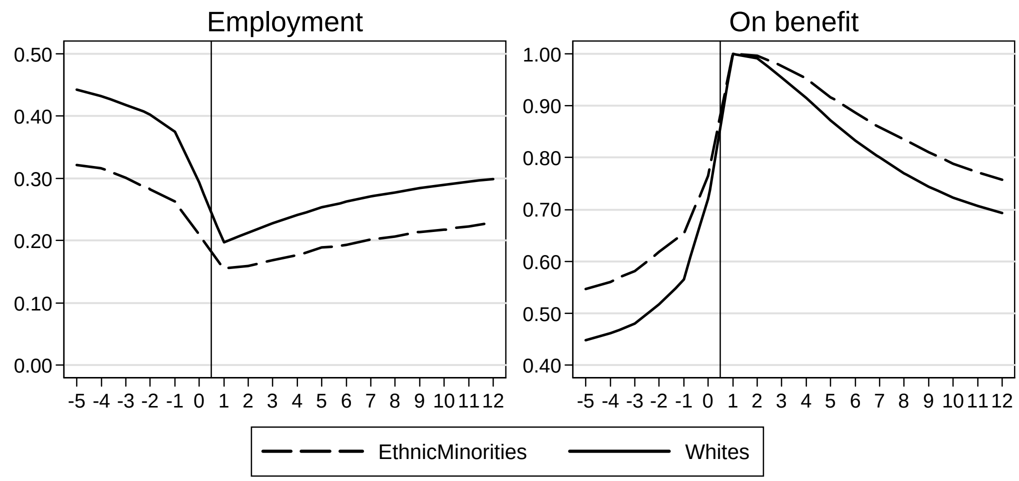
These tables illustrate clearly that in order to be confident of any estimates of the extent of ethnic parity between Minorities and Whites, we must adequately take into account differences in composition. They also suggest it may be important to consider different ethnic groups separately.

Overall IS and IB results

Figures 1 and 2 show raw employment and benefit status over time for our IB and IS samples respectively. Months are measured along the horizontal axis, from six months prior to inflow to 12 months after (months -5 to 12). Inflow is indicated by the vertical line between months 0 and 1. The vertical axis is the proportion of individuals employed or on benefit. For both IB and IS, Whites are more likely to be employed and less likely to be on benefits at all points in time, but the gap narrows around the time of inflow and tends to be smaller for IS than IB. As expected, the proportion of individuals employed rises over time after inflow and the proportion on benefit falls, but the fall in receipt of benefit is faster than the rise in employment. Around 20 per cent of individuals appear to be employed at inflow.[11]

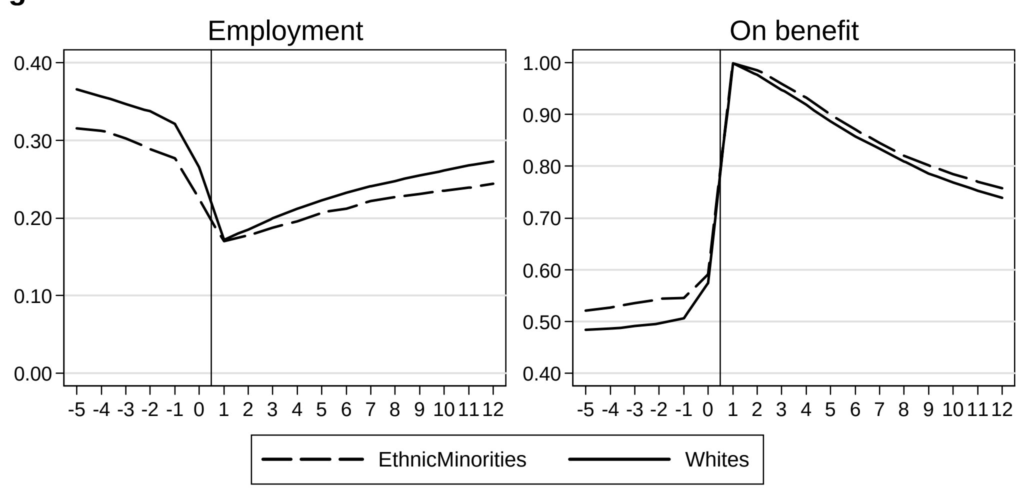
[11] There are a number of possible explanations for this. First, although individuals claiming IB should be incapable of work, this does not mean that they are not employed: for example, it may be the case that they are in a period of temporary absence from their job, but were not entitled to Statutory Sick Pay (and hence have started claiming IB). Similarly, individuals claiming IS may be working (so long as it is less than 16 hours per week), but this is not a valid explanation in our case as these individuals will not have had a WFI and hence will not have been included in our sample. A second possible explanation relates to the fact that the figure refers not to the day of interview itself, but to employment over the next 30 days. It is therefore possible that nobody was working on the actual day of interview, but did work for at least 15 of the next 30 days. Other work using the WPLS, however, suggests that this does not fully explain the difference (see, for example, Brewer et al, 2008). Third, it may also reflect uncertainties over start and end dates in the employment data, or fraudulent benefit claims. It is not possible to quantify the relative importance of each of these explanations.

Figure 1 IB: raw labour market status over time



Notes:
1. The x-axis shows the six months before entry into the programme (x=–5 to 0) and the 12 months after (x=1 to 12).
2. The y-axis shows the proportion of the sample employed or on benefit.

Figure 2 IS: raw labour market status over time



Notes: see notes to Figure 1

Figures 3 to 6 show the raw and matched results for the employment and benefit outcomes of IS and IB claimants overall, all of which do reasonably well on our heuristic reliability criteria described in Section 2. As before, months are along the horizontal axis. The vertical axis measures the percentage point *difference* between Minorities and Whites in the proportion employed/on benefit.[12] A positive difference for employment (benefit) is described as an ethnic premium (penalty) and a negative difference as an ethnic penalty (premium). Large and small circles indicate that the difference is statistically significantly different from zero at the 1 and 5 per cent level.

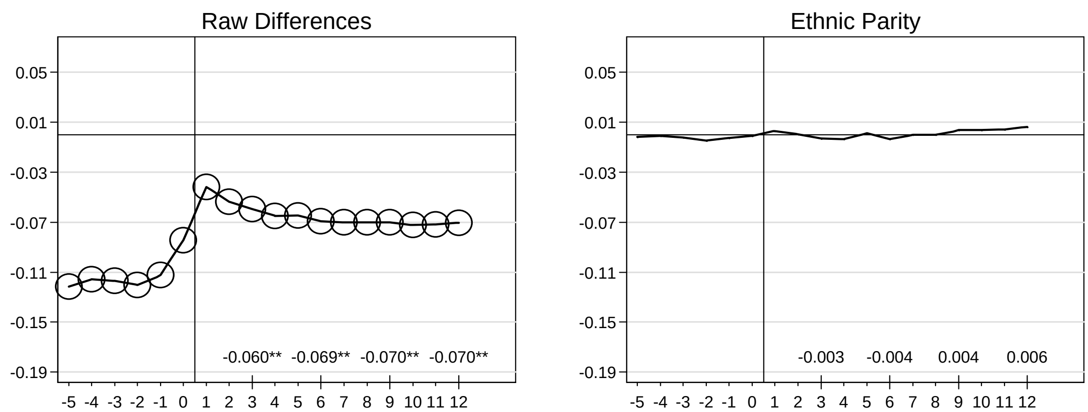
The six months prior to inflow are included in the graphs to give an indication of matching quality: if Whites have been successfully rebalanced to look like Minorities, then the estimates will lie on the horizontal line at zero and there will be no circles indicating statistical significance. Obviously, this is not sufficient to meet all of the above reliability criteria, which also depend on how well all

[12] So, for example, the left-hand panel in Figure 3 shows the difference between the two lines in the left-hand panel of Figure 1.

characteristics (not just previous employment and benefit history) have been balanced (Tables 5 and 6 below provide all the necessary details).

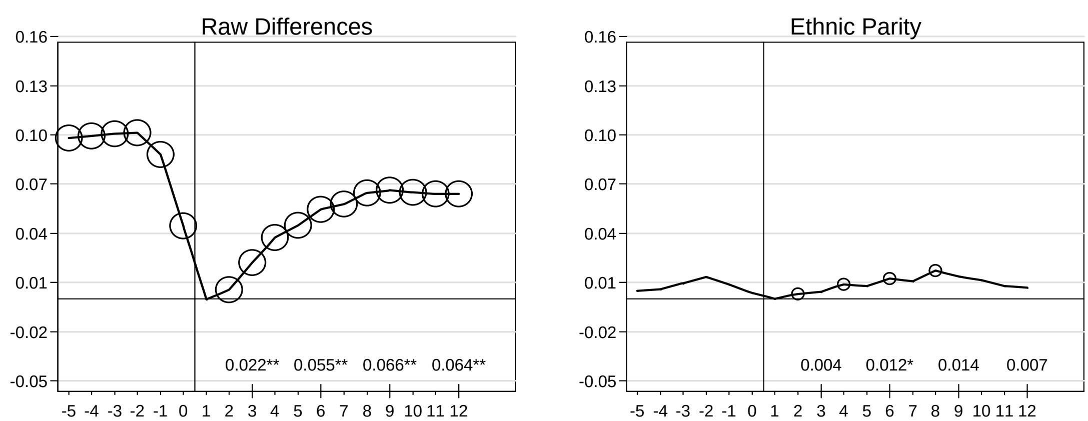
In all four cases, there is a large difference between raw and re-weighted outcomes. For IB, what was roughly a 7 percentage point employment penalty 12 months after inflow is completely eliminated by matching, so that ethnic parity cannot be rejected once characteristics have been controlled for. The IB benefit penalty is eliminated for all but a few months in the middle of the year, and these penalties at around 1 percentage point are considerably smaller than the raw differences. For IS, raw employment and benefit penalties become premia of 2 percentage points by the end of the 12-month outcome window, meaning that Minorities were significantly more likely to be in work and significantly less likely to be claiming benefits than comparable Whites.

Figure 3 IB: ethnic parity in employment outcomes



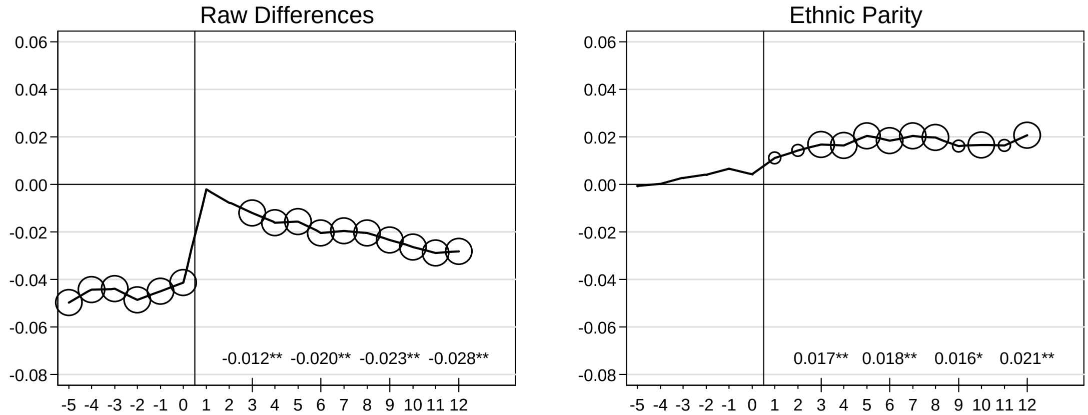
Notes:
1. The x-axis shows the six months before entry into the programme (x=–5 to 0) and the 12 months after (x=1 to 12).
2. The y-axis shows the difference between Ethnic Minorities and Whites in the proportion employed or on benefit.
3. The vertical line shows the time that individuals start claiming the benefit.
4. Two stars (**) and one star (*), and large and small circles, indicate that the difference is statistically significant at the 1 and 5 per cent levels respectively.

Figure 4 IB: ethnic parity in benefit outcomes



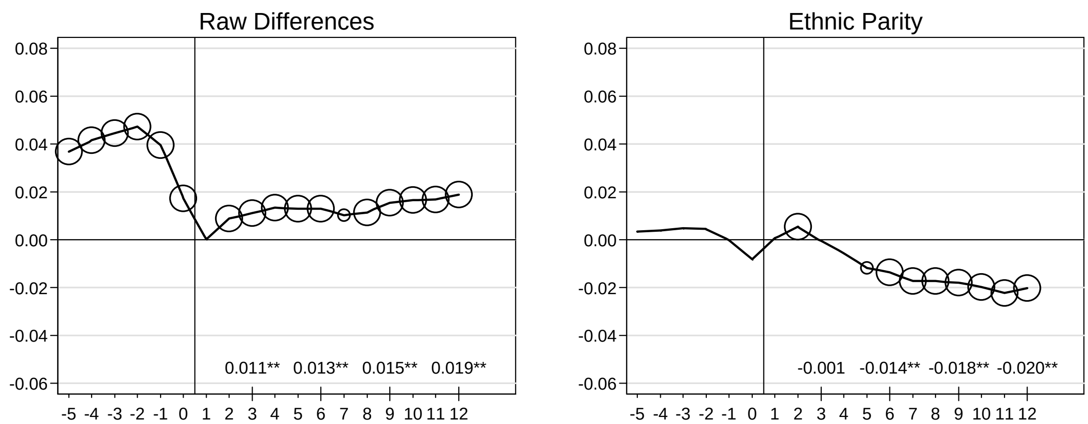
Notes: see notes to Figure 3

Figure 5 IS: ethnic parity in employment outcomes



Notes: see notes to Figure 3

Figure 6 IS: ethnic parity in benefit outcomes



Notes: see notes to Figure 3

Results by ethnic group

For comparative purposes, the top lines of Tables 5 and 6 set out the overall results, with the remainder of the Table providing estimates for different ethnic groups, further broken down by gender (note that the ethnic subgroups are not mutually exclusive). Due to insufficient sample sizes, analyses have not been performed for all subgroups (half of the ethnicity-sex combinations for IB and just under a third for IS contained fewer than the required 400 Minorities). As in Table 1, the figures in the Raw and Matched columns relate to month 12 (the end of the outflow window) and the Matching Quality column provides an indication of how well matching has performed in balancing the observed characteristics of Ethnic Minority and White claimants.

For both IB and IS, all results seem reliable perhaps with the exception of Black and Black African groups. Almost all groups exhibit significant raw employment penalties: for example, the raw results suggest that Blacks overall are 5.4 percentage points less likely to be in employment 12 months after inflow than Whites, while individuals of Mixed, Chinese and Other (MCO) ethnic origin are 7.3 percentage points less likely to be in work. The only two groups without significant employment penalties are Black African males and Indian males. Among Black ethnic groups, raw penalties are generally larger for men than women, while the reverse is true for other ethnic groups. For benefits, there are significant raw penalties for Black and MCO groups (12.8 percentage points for Blacks overall and 8.1 percentage points for MCO overall), and for Black groups, this penalty is generally much larger than the raw employment penalty. For Asian groups, in contrast, there is no significant raw benefit penalty, despite their having a significant employment penalty.

Matching away observable differences between Minorities and Whites again eliminates almost all the penalties, both for employment and benefit outcomes. Black Caribbeans are the only exception, with around a 4.5 percentage point benefit penalty both overall and among the male subgroup. It is however interesting to note that these were the groups who displayed by far the largest differences in raw benefit collection rates (15.2 per cent and 16.6 per cent respectively); contrasting them to comparable White claimants has thus decreased the observed penalty by well over two thirds. On the other hand, two groups experience significant employment premia: males overall and Indian males. This means that Ethnic Minority males (Indian males) are 1.9 percentage points (6.7 percentage points) *more* likely to be in work than comparable White males 12 months after inflow.

Table 5 IB: estimates of ethnic parity for different ethnic groups

Ethnic Group	*Sex*	*N*	*Employment (Month 12)*		*Benefit (Month 12)*		*Matching quality*
			Raw	*Matched*	*Raw*	*Matched*	
All	M+F	5660	-0.070**	0.006	0.064**	0.007	CS(0)UC(28)R2(0.017) MB(0.9)
			(0.006)	(0.008)	(0.006)	(0.008)	
	M	3480	-0.058**	0.019*	0.063**	0.003	CS(0)UC(19)R2(0.018) MB(1.1)
			(0.007)	(0.010)	(0.008)	(0.010)	
	F	2200	-0.090**	-0.013	0.065**	0.012	CS(0)UC(0)R2(0.020) MB(1.5)
			(0.009)	(0.012)	(0.009)	(0.012)	
Black	M+F	1760	-0.054**	0.027	0.128**	0.023	CS(2)UC(23)R2(0.105) MB(3.0)
			(0.010)	(0.019)	(0.009)	(0.019)	
	M	1040	-0.066**	0.027	0.139**	0.025	CS(1)UC(12)R2(0.088) MB(2.6)
			(0.013)	(0.019)	(0.012)	(0.019)	
	F	720	-0.037*	0.028	0.111**	0.014	CS(16)UC(9)R2(0.093) MB(3.1)
			(0.017)	(0.028)	(0.015)	(0.026)	
Black Caribbean	M+F	800	-0.062**	0.008	0.152**	0.046*	CS(5)UC(12)R2(0.085) MB(2.3)
			(0.015)	(0.021)	(0.013)	(0.020)	
	M	500	-0.085**	0.010	0.166**	0.045*	CS(4)UC(0)R2(0.075) MB(2.7)
			(0.018)	(0.023)	(0.016)	(0.022)	
Black African	M+F	720	-0.044**	0.029	0.116**	0.008	CS(5)UC(13)R2(0.143) MB(5.6)
			(0.016)	(0.035)	(0.015)	(0.035)	
	M	400	-0.034	0.006	0.116**	0.005	CS(2)UC(6)R2(0.111) MB(5.2)
			(0.022)	(0.039)	(0.020)	(0.039)	
Asian	M+F	2600	-0.080**	0.013	0.012	-0.022	CS(1)UC(17)UH(B) MB(2.8)
			(0.008)	(0.014)	(0.009)	(0.015)	
	M	1600	-0.052**	0.028	0.006	-0.032	CS(1)UC(6)R2(0.066) MB(2.4)
			(0.011)	(0.019)	(0.012)	(0.020)	
	F	1000	-0.124**	-0.034	0.022	0.001	CS(3)UC(7)R2(0.071) MB(3.3)
			(0.013)	(0.021)	(0.014)	(0.022)	
Indian	M+F	880	-0.044**	0.024	0.002	-0.013	CS(1)UC(5)R2(0.070) MB(3.6)
			(0.015)	(0.023)	(0.016)	(0.024)	
	M	480	-0.011	0.067*	0.005	-0.002	CS(2)UC(0)R2(0.073) MB(4.2)
			(0.021)	(0.032)	(0.022)	(0.033)	

	F	420	-0.089** (0.021)	-0.022 (0.033)	-0.006 (0.023)	-0.012 (0.033)	CS(5)UC(0)R2(0.072) MB(3.6)
Pakistani	M+F	1320	-0.107** (0.011)	-0.019 (0.020)	0.020 (0.013)	-0.004 (0.021)	CS(4)UC(9)R2(0.075) MB(3.4)
	M	860	-0.087** (0.014)	-0.019 (0.025)	0.014 (0.016)	-0.015 (0.027)	CS(4)UC(4)R2(0.066) MB(4.0)
	F	460	-0.139** (0.018)	-0.023 (0.030)	0.035 (0.021)	0.006 (0.031)	CS(8)UC(0)R2(0.074) MB(3.5)
Mixed, Chinese & Other	M+F	1300	-0.073** (0.012)	0.007 (0.015)	0.081** (0.012)	0.020 (0.015)	CS(1)UC(9)R2(0.055) MB(1.8)
	M	820	-0.058** (0.015)	0.021 (0.018)	0.077** (0.015)	0.006 (0.019)	CS(3)UC(0)R2(0.058) MB(2.1)
	F	480	-0.097** (0.019)	0.002 (0.026)	0.089** (0.019)	0.039 (0.025)	CS(0)UC(0)R2(0.063) MB(2.5)
Other	M+F	900	-0.089** (0.014)	-0.001 (0.017)	0.083** (0.014)	0.021 (0.018)	CS(2)UC(3)R2(0.065) MB(2.6)
	M	600	-0.067** (0.017)	0.021 (0.021)	0.076** (0.018)	0.004 (0.022)	CS(3)UC(0)R2(0.061) MB(2.9)

Notes:
1. Three stars two stars (**) and one star (*) indicate that the difference statistically significant at the 1 and 5 per cent level of significance respectively.
2. Figures in parentheses are standard errors. These are calculated using an analytical approximation that ignores the fact that the propensity score is estimated. It was not computationally feasible to bootstrap standard errors.
3. CS(x) means that x per cent of the Ethnic Minority sample was lost to common support. UC(y) means that y covariates remained unbalanced after matching. UH(B,E) indicates that benefit and employment histories remained unbalanced after matching
4. Sample sizes have been rounded to the nearest 20.
5. Ethnic groups are not mutually exclusive. For example, Black Caribbean is a subset of Black, and Other is a subset of Mixed, Chinese and Other.

Results are similar for IS, though less uniform. The raw employment penalties observed for the overall group of claimants are driven by the Asian groups and, to a lesser extent, the Mixed, Chinese and Other groups. Indeed, Black groups display either no raw employment penalties (males) or a premium (females). The overall premia of 4.1 percentage points for Blacks, 5.9 percentage points for Black Caribbeans and 2.6 percentage points for Black Africans are in fact driven by the corresponding female subgroups. As was the case with IB, raw employment outcomes tend thus to be worse for men than for women amongst Black groups, while the reverse is true for other ethnic groups.

In terms of benefit receipt, there are raw penalties for Black and Mixed, Chinese and Other groups. For Asians, there are no raw penalties and a few raw premia (i.e. the raw results suggest that IS claimants of Asian ethnic origin are less likely to be employed but also less likely to be on benefits than White claimants).

Comparing Ethnic Minority claimants with comparable Whites again eliminates all the raw penalties – both in terms of benefit and employment outcomes – and uncovers premia for some groups. In terms of employment, for example, there is a premium of 5.9 percentage points for Blacks overall and of 5.4 percentage points for Black Caribbeans overall (almost three times larger than the 2.1 percentage point premium across all ethnic groups). In terms of benefit receipt, the premia are concentrated among Asian groups: Asians overall are 5.5 percentage points less likely to be claiming benefits than comparable Whites 12 months after inflow, while Indians overall are 6.0 percentage points less likely.

In short, the results for IB and IS indicate that differences in characteristics explain almost all of the observed raw employment and benefit penalties, such that most ethnic groups are found to be at least as likely to be in work and off benefits 12 months after inflow as comparable White benefit claimants. This finding is in marked contrast to previous studies which have sought to estimate the ethnic (un)employment gap, in which significant penalties have been found, even after controlling for a range of characteristics.

Table 6 IS: estimates of ethnic parity for different ethnic groups

Ethnic Group	*Sex*	*N*	*Employment (Month 12)*		*Benefit (Month 12)*		*Matching quality*
			Raw	*Matched*	*Raw*	*Matched*	
All	M+F	8560	-0.028**	0.021**	0.019**	-0.020**	CS(0)UC(35)R2(0.019) MB(1.2)
			(0.005)	(0.007)	(0.005)	(0.006)	
	M	3440	-0.022**	0.024*	0.024**	-0.008	CS(0)UC(22)R2(0.019) MB(1.1)
			(0.007)	(0.009)	(0.007)	(0.010)	
	F	5120	-0.035**	0.020*	0.016*	-0.028**	CS(0)UC(29)R2(0.023) MB(1.6)
			(0.007)	(0.009)	(0.007)	(0.009)	
Black	M+F	3360	0.041**	0.059**	0.053**	-0.018	CS(1)UC(34)UH(B) R2(0.117)MB(3.0)
			(0.008)	(0.020)	(0.007)	(0.019)	
	M	1160	-0.018	0.040*	0.085**	0.008	CS(5)UC(9)R2(0.076) MB(2.1)
			(0.012)	(0.017)	(0.011)	(0.016)	
	F	2200	0.063**	0.051	0.041**	-0.025	CS(1)UC(27)R2(0.141) MB(3.9)
			(0.011)	(0.029)	(0.009)	(0.028)	
Black Caribbean	M+F	1520	0.059**	0.054**	0.042**	-0.022	CS(9)UC(15)UH(B) R2(0.090)MB(3.5)
			(0.012)	(0.018)	(0.011)	(0.017)	
	M	540	-0.030	0.040	0.099**	0.020	CS(5)UC(0)R2(0.078) MB(2.3)
			(0.017)	(0.024)	(0.015)	(0.023)	
	F	980	0.100**	0.047	0.015	-0.042	CS(7)UC(12)R2(0.103) MB(5.2)
			(0.016)	(0.034)	(0.014)	(0.033)	
Black African	M+F	1420	0.026*	0.056	0.068**	-0.014	CS(2)UC(19)R2(0.142) MB(4.8)
			(0.012)	(0.034)	(0.011)	(0.033)	
	M	460	0.002	0.038	0.076**	0.011	CS(4)UC(6)R2(0.116) MB(4.5)
			(0.019)	(0.032)	(0.017)	(0.031)	
	F	960	0.026	0.070	0.069**	-0.011	CS(1)UC(14)R2(0.153) MB(5.5)
			(0.015)	(0.042)	(0.013)	(0.040)	
Other Black	M+F	420	0.026	0.041	0.046*	0.010	CS(1)UC(0)R2(0.102) MB(3.9)
			(0.023)	(0.024)	(0.020)	(0.021)	
Asian	M+F	3380	-0.089**	0.010	-0.028**	-0.055**	CS(1)UC(22)R2(0.073) MB(2.3)
			(0.007)	(0.015)	(0.008)	(0.015)	
	M	1460	-0.024*	0.031	-0.022	-0.010	CS(2)UC(10)R2(0.072) MB(3.3)
			(0.011)	(0.023)	(0.012)	(0.024)	
	F	1920	-0.137**	-0.020	-0.034**	-0.073**	CS(3)UC(16)R2(0.074) MB(2.1)
			(0.009)	(0.017)	(0.011)	(0.018)	
Indian	M+F	900	-0.036*	0.034	-0.028	-0.060*	CS(2)UC(13)R2(0.081) MB(2.7)
			(0.014)	(0.025)	(0.015)	(0.025)	
	F	520	-0.059**	0.004	-0.024	-0.061*	CS(7)UC(0)R2(0.068) MB(3.3)
			(0.019)	(0.026)	(0.020)	(0.027)	
Pakistani	M+F	1920	-0.116**	-0.012	-0.031**	-0.026	CS(4)UC(13)R2(0.076) MB(3.1)
			(0.008)	(0.020)	(0.010)	(0.021)	
	M	840	-0.053**	-0.004	-0.014	-0.014	CS(6)UC(4)R2(0.070) MB(3.6)
			(0.013)	(0.029)	(0.015)	(0.031)	
	F	1100	-0.162**	-0.049	-0.045**	-0.024	CS(3)UC(4)R2(0.086) MB(3.6)
			(0.011)	(0.029)	(0.014)	(0.030)	
Mixed, Chinese & Other	M+F	1820	-0.043**	0.013	0.043**	-0.003	CS(1)UC(11)R2(0.049)

							MB(1.5)
			(0.010)	(0.012)	(0.010)	(0.012)	
	M	820	-0.024	0.033	0.021	-0.024	CS(1)UC(6)R2(0.061) MB(2.4)
			(0.014)	(0.018)	(0.015)	(0.018)	
	F	1000	-0.053**	-0.005	0.059**	0.011	CS(1)UC(0)R2(0.046) MB(1.9)
			(0.014)	(0.017)	(0.013)	(0.016)	
Mixed	M+F	580	0.003	0.012	0.039*	0.021	CS(0)UC(0)R2(0.077) MB(2.9)
			(0.019)	(0.019)	(0.018)	(0.018)	
Other	M+F	1160	-0.070**	0.006	0.050**	-0.012	CS(1)UC(5)R2(0.059) MB(2.0)
			(0.012)	(0.015)	(0.012)	(0.015)	
	M	580	-0.026	0.049*	0.028	-0.026	CS(1)UC(0)R2(0.061) MB(2.8)
			(0.017)	(0.022)	(0.017)	(0.022)	
	F	580	-0.099**	-0.023	0.066**	-0.003	CS(2)UC(0)R2(0.055) MB(1.9)
			(0.017)	(0.022)	(0.017)	(0.021)	

Notes: see notes to Table 5

7. Conclusion

From a policy perspective, it is important to know whether the gap that exists in labour market outcomes between Ethnic Minorities and Whites is due to differences in the characteristics of Ethnic Minority and White groups or whether it results from some form of discriminatory behaviour in the labour market.

Satisfactory results could be calculated for claimants of Incapacity Benefit and Income Support. For these two groups we found that large and significant raw penalties disappeared once we controlled for background and labour market characteristics. In other words, the gap in labour market outcomes can be explained entirely by differences in pre-programme characteristics and local conditions. Black Caribbeans on IB are the only group for which penalties, though largely reduced in size, persist after matching. For males and females together, the penalty is 4.6 percentage points. For IS, premia are more widespread, being found for Ethnic Minorities overall, as well as for Black employment and Asian benefit outcomes.

The analysis in this paper does not allow one to disentangle the various potential channels that might lead Ethnic Minorities to experience different outcomes from observationally similar Whites. At least three factors could be at work if an ethnic penalty is found: labour market discrimination by employers; differential effects of assistance received whilst on IS or IB for Minorities compared to Whites; and self-discriminatory behaviour by Ethnic Minorities (such as not applying for a job) because they anticipate (rightly or wrongly) that they will be discriminated against. These effects are conflated in our estimates of ethnic parity.

The most important conclusion of this paper, however, is that for many groups it turns out not to be possible to calculate reliable quantitative estimates of the extent of ethnic parity using existing administrative data combined with local neighbourhood information. Whites and Minorities accessing Jobcentre Plus programmes and services are often simply too different in terms of pre-programme characteristics to answer this question. This will pose a problem regardless of the estimation method used: indeed, our analysis has shown that, for some programmes, the estimates of ethnic parity differ in sign and significance depending on the quantitative approach taken. These issues raise serious questions about the reliability of previous work that has used simple regression methods without carefully considering whether the assumptions required to produce unbiased estimates have been satisfied.

8. References

Berthoud, R. & M. Blekesaune (2006), *Persistent employment disadvantage, 1974 to 2003*, ISER Working Paper, 2006-9

Blackaby, D., S. Drinkwater, D. Leslie & P. Murphy (1997), *A picture of male and female unemployment amongst ethnic minorities in Britain*, Scottish Journal of Political Economy, Vol. 44, pp. 182-197

Blackaby, D., D. Leslie, P. Murphy & N. O'Leary (1999), *Unemployment among Britain's ethnic minorities*, Manchester School, Vol. 67, pp. 1-20

Blundell, R., L. Dearden & B. Sianesi (2005), *Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey*, Journal of the Royal Statistical Society: Series A, Vol. 168, No. 3, pp. 473-512

Brewer, M., J. Browne, H. Chowdry & C. Crawford (2008), The lone parent pilots after 24-36 months: an updated impact assessment of In Work Credit, Work Search Premium, Extended Schools Childcare, Quarterly Work Focused Interviews and New Deal Plus for Lone Parents, Draft Report for Department for Work and Pensions, October 2008

Cabinet Office (2003), *Ethnic minorities and the labour market: final report*, HMSO, London

Crawford, C., L. Dearden, A. Mesnard, J. Shaw, B. Sianesi, and P. Urwin (2008), *Estimating ethnic parity in Jobcentre Plus programmes: A quantitative analysis using the Work and Pensions Longitudinal Study (WPLS)*, DWP Research Report No. 491

Heath, A. & S. Cheung (2006), *Ethnic penalties in the labour market: employers and discrimination*, DWP Research Report No. 341

Heckman, J. and J. Smith (1999). "The Pre-Programme Dip and the Determinants of Participation in a Social Programme: Implications for Simple Programme Evaluation Strategies." Economic Journal 109(457): 313-348.

Heckman, J., H. Ichimura, J. Smith and P. Todd (1998) "Characterising Selection Bias Using Experimental Data." Econometrica 66(5): 1017-1098.

Heckman, J., R. LaLonde, and J. Smith (1999). "The Economics and Econometrics of Active Labor Market Programs." In Orley Ashenfelter and David Card (eds.), Handbook of Labor Economics, Volume 3A. 1865-2097.

Imbens, G. (2004), 'Semiparametric estimation of average treatment effects under exogeneity: a review', *Review of Economics and Statistics*, 86, 4-29

National Audit Office (2008), *Department for Work & Pensions: increasing employment rates for ethnic minorities*, Report by the Comptroller and Auditor General, HC 206, Session 2007-08