UNIVERSITY OF Southampton
School of Social Sciences

Economics Department
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

The impact of labour market discrimination on welfare dependency of second generation immigrants in the UK

Armine Ghazaryan (University of Southampton)

No. 1901

This paper is available on our website
http://www.southampton.ac.uk/socsci/economics/research/papers

ISSN 0966-4246

The impact of labour market discrimination on welfare dependency of second generation immigrants in the UK*

Armine Ghazaryan†

January 2019

Abstract

Many studies show that immigrants tend to claim more benefits than natives, even when accounting for their individual characteristics. This paper suggests and tests the hypothesis that the tendency of immigrants to claim more benefits is linked to income discrimination in the labour market.

This study uses panel data from Understanding Society, a UK household longitudinal survey, to look at second generation immigrants in comparison to UK natives. By estimating labour market discrimination against immigrants using available methodology on income decomposition, the paper then uses the estimates of discrimination to study whether labour market discrimination affects welfare dependency of immigrants. This paper shows that immigrants' likelihood to move into state welfare dependency increases when there is discrimination in the labour market. The results differ for EU versus non-EU second generation immigrants.

Keywords: Immigrants, Discrimination, Welfare

JEL Codes: J15, J71, I38

*I am very grateful to my supervisors, Jackie Wahba and Corrado Giulietti, for their continuous support and guidance. I am also grateful to Jeffrey Wooldridge, Carmine Ornaghi, Michael Vlassopoulos and Thomas Gall for their valuable comments.

†University of Southampton, Department of Economics, Email: a.ghazaryan@soton.ac.uk

1 Introduction

Sustaining a large number of people on state welfare benefits is costly for a country, therefore, it is important to understand the reasons behind welfare dependency. Given its importance, the issue of state welfare dependency of immigrants is a constant topic of political discussion, including in the United Kingdom.

Due to the growing immigrant population of the UK, many studies concentrate on the patterns of state welfare dependency of immigrants compared with natives. And while the effect of recent immigration can be perceived as temporary and may fade over time as immigrants return to their home countries or assimilate, the effect of British-born second generation immigrants is persistent. According to the Office for National Statistics reports[1] the share of children born in England and Wales to foreign-born parents has been increasing since the 1990s, and currently one in three childbirths are to foreign-born parents.

Labour market outcomes of immigrants, as well as the patterns of claiming state welfare benefits by immigrants versus natives have been vastly explored, whereas the reasons immigrants claim benefits have not been studied much. This paper explores the link between income discrimination in the labour market of the UK and state welfare dependency of second generation immigrants.

This paper contributes to the literature in multiple ways. Firstly, it explores the link between income discrimination and state welfare dependency of second generation immigrants which, to the best of my knowledge, has not been explored.

Secondly, the estimations are strengthened by using second generation immigrants as a subject matter, which reduces biases associated with first generation immigrants, such as return-migration, incomparability in levels of education and the language factor, which can be a possible reason for differences in labour market outcomes for immigrants compared with natives.

And finally, the factors uncovered for second generation immigrants can be valid for first generation immigrants as well, as the paper explores patterns for different ethnic groups, which, if there for second generation immigrants are most likely to be even stronger for first

[1] www.ons.gov.uk

generation immigrants, as discussed by Brücker et al. (2002).

There are many studies on the topic of reliance on welfare benefits by first generation immigrants compared with natives (Borjas and Hilton, 1996; Hansen and Lofstrom, 2003; Barrett and McCarthy, 2008; Riphahn, 1998; Castronova et al., 2001; Bruckmeier and Wiemers, 2017). Most studies find higher welfare dependency of immigrants when looking at raw data. Yet first generation immigrants are subject to different initial conditions compared with the native population, thus making the comparison subject to biases, such as incomparable levels of education and work experience, or language skills of immigrants being different from natives. That is, as evidenced by Castronova et al. (2001) and Bruckmeier and Wiemers (2017), once the initial conditions are considered, there is no significant difference in the probabilities of claiming benefits by immigrants versus natives, and, in some cases, (Riphahn, 1998, for instance) the probabilities of claiming are lower for immigrants.

The differences in the initial conditions can make immigrants less competitive in the labour market thus moving them into a higher risk of relying on welfare support. Brücker et al. (2002) discuss that several factors might place immigrants into a situation, where they are more likely to be on welfare dependency than natives. They highlight that immigrants might self-select to countries with generous welfare systems, hence their income is likely to depend not only on observable characteristics, but also on some unobservables that result in welfare dependency. Immigrants are also likely to be affected by certain migration-related idiosyncrasies, such as psychological factors from moving to another country and language issues, which might increase the risk of welfare dependency, or weaken the welfare entitlement. Besides, immigrants might have limited transferability of their entitlements in their home countries, such as work experience; or immigrants might also have better or worse networks compared with natives, which will affect their labour market outcomes. Two more reasons the welfare dependency of immigrants might be different from natives outlined by Brücker et al. (2002) are discrimination and reduced wages. Discrimination in the labour market might push immigrants towards welfare dependency. Discrimination might also affect immigrants' incentives to look for a job if it results in reduced wages for immigrants.

The factors above make comparison of natives and first generation immigrants difficult. That is, while the probability of welfare dependency might be different for immigrants compared with natives, this might be due to these factors contributing to immigrants being different from natives, rather than being attributable to the propensity of immigrants to claim more or less benefits. These factors might also be different across countries.

For second generation immigrants, on the other hand, the factors of self-selection, migration-related idiosyncrasies, non-transferability of entitlements and networks mostly disappear. The factors of discrimination, and reduced wages as a result of discrimination, however, continue to be of great importance in explaining differences in take-up of benefits between immigrants and natives.

Many studies find significant income discrimination against both first and second generation immigrants or income gaps for certain groups of immigrants or ethnic minorities in the UK (Chiswick, 1980; Blackaby et al., 2002; Bell, 1997; Clark and Drinkwater, 2008; Dustmann and Theodoropoulos, 2010). This study uses estimates of income discrimination against second generation immigrants to test the hypothesis that discrimination affects the probability of them claiming state welfare benefits. It uses panel data from Understanding Society, the UK Household Longitudinal Study, first to estimate wage discrimination against immigrants, which is in line with findings from previous studies. It then uses these estimates to assess the impact of discrimination on the welfare dependency of immigrants.

The paper is organised as follows: Section 2 provides a review of relevant literature, Section 3 described the data, provides data analysis and describes the estimation method, Section 4 discusses the results and robustness tests, and Section 5 concludes.

2 Background studies

The issue of the reliance of immigrants on the welfare system of the host country has been widely studied in economic literature. Most studies look into probabilities of claiming benefits by first generation immigrants versus natives. Yet, there are only a few studies that discuss the reasons for immigrant dependency on welfare benefits. This paper discusses the link between income discrimination and welfare dependency of immigrants. Below is a review of the background literature on welfare dependency, followed by literature on income discrimination.

Welfare dependency of immigrants

When looking at overall probability of immigrants claiming benefits, most studies find higher probabilities for immigrants compared with natives (Borjas and Hilton, 1996; Hansen and Lofstrom, 2003; Barrett and McCarthy, 2008, for overview of related literature). When controlling for individual characteristics, however, different studies find different results.

For instance, the study by Hansen and Lofstrom (2003) looking into the case of Sweden, finds that immigrants receive more welfare benefits when considering raw data, and it is not explained by their individual characteristics.

Bird et al. (1999), looking into the case of Germany, find that immigrants are both more likely to be eligible, and also, have higher probability to take up benefits, conditional on eligibility. However, they find that, when controlling for socio-economic factors, immigrants do not tend to exhibit a higher likelihood of claiming benefits compared with natives.

On the other hand, other studies, looking into the take up of welfare benefits, conditional on eligibility, find that the immigrant take up of benefits is not significantly different from that of natives (Riphahn, 1998; Castronova et al., 2001; Bruckmeier and Wiemers, 2017). Castronova et al. (2001) and Bruckmeier and Wiemers (2017) look at the differences in patterns of claiming welfare benefits by immigrants and natives in Germany, conditional on eligibility, thus capturing the differences in behaviour between immigrants and natives. Castronova et al. (2001) find that immigrants are more likely to claim benefits. However, when controlling for a number of socio-economic characteristics, immigrant take-up of benefits is no different from that of natives. Bruckmeier and Wiemers (2017), using a microsimulation model study the probability of immigrants and natives to claim benefits, conditional on eligibility for welfare benefits. They also find no evidence that immigrants are more likely to take up benefits than natives, after controlling for eligibility, even though immigrants have higher risk to be eligible for welfare benefits.

A more recent study by Barrett and Maître (2013) estimates whether immigrants are more likely to receive welfare benefits compared with natives for a number of EU countries, including the UK, using data from European Union Statistics on Income and Living Conditions for 2007. Their findings indicate that there is little evidence that immigrants would receive more social benefits than natives.

Drinkwater and Robinson (2013) look into welfare participation in the UK. They use data from the UK Labour Force Survey for 2004-2009 to examine welfare dependency of first generation immigrants by types of benefits claimed and country of origin. They find different patterns of welfare dependency for different groups of immigrants and benefits claimed.

Brücker et al. (2002) study welfare dependency of non-EU immigrants across EU countries. They derive a residual dependency, as a difference between predicted dependency, based on individual characteristics, and immigrants' actual dependency. They study welfare dependency by three types of benefits: unemployment benefits, old-age pensions and family benefits. Their findings show that the average predicted unemployment welfare dependency of immigrants is slightly higher for immigrants than natives; the average predicted old-age pension dependency is much higher for natives (almost non-existent for immigrants); and the average predicted family welfare dependency is higher for immigrants, although differs across countries. Finally, they move to comparing the predicted welfare dependency based on the certain set of characteristics with actual welfare dependency, that is, residual dependency, to understand whether immigrants are more or less likely to be dependent on welfare than natives. They find positive and significant *unemployment welfare dependency* of immigrants for Finland, Denmark, Austria, Netherlands, France and Belgium, no *old-age pension residual dependency* for immigrants, while immigrants' *family welfare dependency* is positive and significant for France and Spain, and it is negative and significant for the UK.

The authors highlight the possible reasons for residual dependency:

- **self-selection:** immigrants with low earnings will self-select to countries with generous welfare systems, and hence their earnings in host country will not only depend on observed characteristics, but also on some unobserved individual characteristics, which will result in residual welfare dependency (this phenomenon and related literature are discussed, for instance, in Borjas (1999); Giulietti and Wahba (2013); Barrett and Maître (2013); Giulietti et al. (2013b); Razin and Wahba (2015));

- **migration-related idiosyncratic effects:** immigrants might be affected by specific factors, such as psychological and language issues, which might increase their risk of welfare dependency; or welfare entitlement might be conditional on literacy in the language of the host country, in the case of negative residual dependency;

- **networks:** ethnic networks can make it easier for immigrants to find a job, or make them depend on welfare, if they have less developed networks than natives (this topic is explored, for instance, by Munshi (2003); Frijters et al. (2005); Battu et al. (2011); Giulietti et al. (2013a));

- **non-transferability of entitlements:** if immigrants cannot transfer their entitlement from home countries, then they will have negative residual dependency compared with natives with the same characteristic (particularly prominent in the case of pensions); on the other hand immigrants might be less entitled to benefits due to non-portability of work experience;

- **discrimination:** discrimination in labour market might push immigrants towards welfare dependency;

- **reduced wages:** factors reducing wages of immigrants could result in welfare dependency (for instance, by disincentivising looking for a job). These factors can be discrimination or reduced access to public jobs.

The discussion above highlights that studies on welfare take up by first generation immigrants and the comparison with natives are prone to biases, such as incomparability of labour market outputs due to immigrants having language skills different to those of natives.

When studying second generation immigrants, the factors of self-selection, migration-related effects, non-transferability of entitlement and largely, networks, seem not to be relevant. Yet, discrimination, and reduced wages as a result of discrimination continue to be of great importance in explaining differences in take-up of benefits between immigrants and natives.

There are studies on the effect of discrimination on labour market outcomes of immigrants (Giulietti et al., 2017; Jilke et al., 2018; Neumark, 2016, for review of experimental research). To the best of my knowledge, however, the link between discrimination in labour market and welfare dependency of immigrants has not been studied.

Labour market discrimination

The issue of discrimination in labour market has been studied extensively. The first economic model on discrimination by Becker (1957) introduced "taste discrimination", according to which employers get disutility from employing minority workers. The firms, therefore, will hire minority workers only if their wage offsets the disutilty.

Later studies, Phelps (1972) and Arrow et al. (1973), discuss the notion of "statistical discrimination", according to which, when employers have limited information about productivity of an employee, they infer it from observable characteristics, for instance, gender or race, and their correlation with productivity (usually based on a group mean).

A more recent study by Bertrand et al. (2005) suggests a third concept, "implicit discrimination", when individuals are not aware of their discriminatory behaviour. In their study the discriminatory behaviour is discovered through a race Implicit Association Test.

It has been shown that persistent discrimination can be a self-fulfilling prophecy, affecting the performance and educational choices of certain groups (Glover et al., 2017, for instance).

There have been many empirical studies on ethnic discrimination and the earnings gap in the UK labour market. One of the first studies on ethnic discrimination in the UK labour market by Chiswick (1980) and McNabb and Psacharopoulos (1981) discusses earnings of the white and non-white UK population and finds that the earnings of the non-white population are lower, not attributable to education and potential experience. McNabb and Psacharopoulos (1981) argues that the disadvantage in the earnings gap is attributable to lower return to education and return to experience for the non-white population.

Blackaby et al. (1994) study wage and employment gaps between the white and black population in the UK for the periods of 1970s and 1980s using General Household Surveys (GHS). They decompose probit equations for the probabilities of employment, and log-linear equations for income of both groups. They find not only a significant income gap and a gap in employment prospects for the black population versus the white population, but also that the gaps tend to deteriorate in 1980s compared with 1970s. Blackaby et al. (1998) and Blackaby et al. (2002) update the results based on the data from 1990s and further explore the question using the Labour Force Survey (LFS), which makes it possible for them to also look at different UK-born ethnic groups. These studies confirm disadvantaged positions in employment and income of ethnic minorities in the UK, which cannot be explained by observable characteristics, including qualifications or region.

Similar findings are discussed by Bell (1997), who uses GHS data of 1973-1992 to study the performance of first generation immigrants to the UK by country of origin, while accounting for their education, cohort, years since migration and foreign experience. He finds that the most disadvantaged group is black immigrants with work experience abroad. The gap remains, but gets smaller as they assimilate over time. He also finds that white immigrants, in contrast, are better positioned compared with natives, but the difference disappears after a short time.

Clark and Drinkwater (2008) study labour market performance of first generation immigrants to the UK in comparison with the UK natives, using data from LFS. They find that all immigrants perform worse compared with natives in terms income and employment, particularly after accounting for individual characteristics, although the scale differs across groups. However, English language proficiency varies across groups and is likely to cause the disadvantage compared with UK native born.

Dustmann and Theodoropoulos (2010) discuss economic performance of both first and second generation immigrants in the UK using LFS data and compare them with UK white natives. Their findings indicate that even though ethnic minorities are better educated than UK natives, they have lower employment rates. They also find that both male and female second generation immigrants, when accounting for their observable characteristics, receive lower earnings compared with UK natives. They did not find any relationship between employment rates and self-reported perceptions of discrimination.

Algan et al. (2010) compare economic performance of first and second generation immigrants in Germany, France and the UK. They find that the UK has higher income and employment gaps of first generation immigrants, but also considerable improvements for second generation immigrants, even though the gaps persist for some groups of immigrants.

The studies discussed highlight a general pattern of income gap between natives and immigrants. The gap is usually bigger for first generation immigrants, which is to be expected considering the different initial conditions for immigrants versus natives, such as language skills or education. The gap, however, persists for some groups of second generation immigrants as well, which is likely to affect immigrants' behaviour and their labour market outcomes. Therefore, this paper contributes to the literature in understanding the consequences of a persistent income gap, particularly how it is linked with welfare take-up by second generation immigrants.

3 Methodology and data

3.1 Methodology

This study firstly uses existing methods on estimating income gap or discrimination between natives and immigrants, and then uses the estimates to study the effect of the gap on the benefit take-up by immigrants. Particularly, it uses Blinder-Oaxaca decomposition (B-O) method to estimate discrimination in labour market (Blinder, 1973; Oaxaca, 1973; Jann et al., 2008), as the B-O method allows for direct comparison and estimation of a value of income discrimination. The latter is important as we need to estimate a numerical value of discrimination to use it for further analysis of welfare dependency. The comparison of decomposition methods and the details of B-O method are described in Appendix A.

In order to estimate the productivity of natives and immigrants we include the following individual characteristics: potential experience = age - years of education − 6, squared potential experience, years of education, squared years of education (highest educational qualification achieved converted to years), occupations, job type: part-time/full-time, industry, UK government office region, gender, urban versus rural area, health issues.

Based on Blinder-Oaxaca decomposition, the difference in labour market outcomes for the groups of natives (N) and immigrants (M) is:

$$R = E(Y_N) - E(Y_M), \tag{1}$$

where $E(Y_N)$ and $E(Y_M)$ are expected value of log earnings of natives and immigrants, accordingly, the estimates of which are derived by estimating the following equation for natives and immigrants:

$$Y_k = \boldsymbol{X}'_k \boldsymbol{\beta}_k + \epsilon_k, \;\; where \;\; E(\epsilon_k) = 0, \;\; \boldsymbol{X}_k - a\ set\ of\ explanatory\ variables \;\; and \;\; k \in \{N, M\} \tag{2}$$

Substituting (2) in (1) and rearranging as described in Section 2, we get:

$$\hat{R} = (\bar{X}_N - \bar{X}_M)'\hat{\beta}_N + \bar{X}'_N(\hat{\beta}_N - \hat{\beta}_M) \tag{3}$$

On the other hand, since the data under consideration is a panel data, the equation (2) for panel data has the following form:

$$y^k_{it} = b^k_t + \boldsymbol{X}^{k'}_{it}\boldsymbol{\beta}^k + c^k_i + e^k_{it}, \tag{4}$$

where b^k_t is the time intercept, and c^k_i is the time-invariant unobserved effect.

The choice of the estimation method of (4) largely depends on the relationship between $\{\boldsymbol{X}_{it} : t = 1, 2, ..., T\}$ and c_i (Wooldridge, 2010, 2015; Hsiao, 2014). Considering that in our case y_{it} is the log income, and $\boldsymbol{X}_{it}$'s are trying to capture productivity, making an assumption that $Cov(\boldsymbol{X}_{it}, c_i) = 0$ will be too strong. That is, we need to allow correlation between $\boldsymbol{X}_{it}$ and c_i. Therefore, the estimation of (4) will be consistent when using Fixed Effects method (FE). However, since the Fixed Effects method removes c_i, all time-invariant variables are also removed. The latter, as pointed out by Heitmueller (2005), can potentially be an issue for Blinder-Oaxaca decomposition as it can result in an omitted variable issue when interpreting the unexplained component.

In order to tackle the above mention issue, Correlated Random Effects method (CRE) is applied, in which case, rather than removing c_i, the relationship between $\boldsymbol{X}_{it}$ and c_i is modelled (Mundlak, 1978; Wooldridge, 2010, 2015):

$$c_i = \varrho + \bar{\boldsymbol{X}}_i \boldsymbol{\xi} + a_i, \tag{5}$$

where $\bar{\boldsymbol{X}}_i = T^{-1}\sum_{t=1}^{T}\boldsymbol{X}_{it}$. CRE produces exactly the same results for $\boldsymbol{\beta}$, but also allows for time-invariant variables. Thus, substituting (5) and allowing for time-invariant variables Q_i, (4) is modified into the following:

$$y_{it}^k = b_t^k + \boldsymbol{Q}_i^{k\prime}\boldsymbol{\delta}^k + \boldsymbol{X}_{it}^{k\prime}\boldsymbol{\beta}^k + \varrho^k + \bar{\boldsymbol{X}}_i^k\boldsymbol{\xi}^k + a_i^k + e_{it}^k, \tag{6}$$

I then estimate (6) using Random Effects, as $Cov(\boldsymbol{X}_{it}, a_i) = 0$ and $Cov(\boldsymbol{X}_i, a_i) = 0$.

Another major issue to consider is that the panel under consideration is unbalanced. Hence, it is important to understand whether the attrition / sample selection is uncorrelated with the idiosyncratic error, e_{it}, as well as the time-invariant unobserved effect, c_i. If we define an indicator variable, s_{it}, as follows:

$$s_{it} = \begin{cases} 1 & if\ all\ of\ (\boldsymbol{X}_{it}, y_{it})\ are\ observed \\ 0 & otherwise \end{cases}$$

,

then FE allows $Cov(s_{it}, c_i) \neq 0$, while for consistency it requires that $Cov(s_{it}, e_{it}) = 0$, in addition to $Cov(\boldsymbol{X}_{it}, e_{it}) = 0$. The same assumptions apply to CRE, provided one accounts for the panel being unbalanced. The paper follows Wooldridge (2010, 2015), Mundlak (1978) in applying CRE method. It uses the observation for which complete set of data are observed, that is, when $s_{it} = 1$. It then include time averages of the variables for complete set of data only:

$$\bar{\boldsymbol{X}}_i = T^{-1}\sum_{t=1}^{T} s_{it}\boldsymbol{X}_{it}.$$

Furthermore, time averages of time effects are also included:

$$\bar{b}_i = T^{-1}\sum_{t=1}^{T} s_{it}b_t.$$

Thus, after the adjustments for unbalanced panel for CRE, (6) looks like follows:

$$y_{it}^k = b_t^k \mu + \boldsymbol{Q}_i^{k'} \boldsymbol{\delta}^k + \boldsymbol{X}_{it}^{k'} \boldsymbol{\beta}^k + \varrho^k + \bar{\boldsymbol{X}}_i^k \boldsymbol{\xi}^k + \bar{b}_i^k \eta^k + a_i^k + e_{it}^k, \tag{7}$$

where $\bar{\boldsymbol{X}}_i = T^{-}1 \sum_{t=1}^{T} s_{it} \boldsymbol{X}_{it}$ and $\bar{b}_i = T^{-}1 \sum_{t=1}^{T} s_{it} b_t$.

However, as mentioned above, the consistency of CRE requires that $Cov(s_{it}, e_{it}) = 0$, that is the panel is unbalanced due to randomly missing data. In the data under consideration the main reason for the panel to be unbalanced is because the dependent variable, log income from labour, is observed only if an individual is employed and receives a positive income. That is, if we denote $\boldsymbol{Z}$ the full set of independent variables regardless of whether income from labour is observed or not, we have: $s_{it} = 1[\boldsymbol{Z}_{it}\boldsymbol{\gamma} + \nu_{it} \geq 0]$, assuming that $E(\nu_{it}|\boldsymbol{Z}_{it}) = 0$ and $\nu_{it} \sim N(0,1)$. The latter indicates that the observations are not randomly missing from the panel and creates a potential sample selection bias. As shown in Tables 8 and 9, the share of labour force participation varies across natives and different groups of immigrants, as well as for males and females. Therefore, in order to correct for the potential sample selection bias, I follow Heckman's two-step approach for sample selection correction (Heckman, 1979; Wooldridge, 2010, 2015).

Considering that $\boldsymbol{X}_{it}$ and $\boldsymbol{Q}_i$ are sub-samples of $\boldsymbol{Z}_{it}$, the model (7) can be written as follows:

$$E(y_{it}|\boldsymbol{Z}_{it}, a_i, s_{it} = 1) = E(y_{it}|\boldsymbol{Z}_{it}, a_i, y_{it} \geq 0) = b_t\mu + \boldsymbol{Q}_i^{'}\boldsymbol{\delta} + \boldsymbol{X}_{it}^{'}\boldsymbol{\beta} + \varrho + \bar{\boldsymbol{X}}_i\boldsymbol{\xi} + \bar{b}_i\eta + a_i + E(e_{it}|\nu_{it} \geq -\boldsymbol{Z}_{it}\boldsymbol{\gamma}) \tag{8}$$

If we represent $E(e_{it}|\nu_{it} \geq -\boldsymbol{Z}_{it}\boldsymbol{\gamma})$ as $\rho E(e_{it}|\boldsymbol{Z}_{it}, s_{it})$, given that $s_{it} = 1[\boldsymbol{Z}_{it}\boldsymbol{\gamma} + \nu_{it} \geq 0]$ and $\nu_{it} \sim N(0,1)$, then $E(e_{it}|\boldsymbol{Z}_{it}, s_{it}) = \lambda(\boldsymbol{Z}_{it}\boldsymbol{\gamma}) = \phi(\boldsymbol{Z}_{it}\boldsymbol{\gamma})/\Phi(\boldsymbol{Z}_{it}\boldsymbol{\gamma})$, the inverse Mills ratio, when $s_{it} = 1$. Thus, the estimable version of (8) is the following:

$$E(y_{it}|\boldsymbol{Z}_{it}, a_i, s_{it} = 1) = b_t\mu + \boldsymbol{Q}_i^{'}\boldsymbol{\delta} + \boldsymbol{X}_{it}^{'}\boldsymbol{\beta} + \varrho + \bar{\boldsymbol{X}}_i\boldsymbol{\xi} + \bar{b}_i\eta + a_i + \rho\lambda(\boldsymbol{Z}_{it}\boldsymbol{\gamma}) \tag{9}$$

In the two step approach, as a first step γ is estimated and $\hat{\lambda}_{it} = \lambda(\boldsymbol{Z}_{it}\hat{\boldsymbol{\gamma}})$ is computed for each i and t. Since $P(s_{it} = 1|\boldsymbol{Z}_{it})$ follows a probit model, γ is estimated from the following probit model:

$$P(s_{it} = 1|\boldsymbol{Z}_{it}) = \Phi(\boldsymbol{Z}_{it}\boldsymbol{\gamma}) \tag{10}$$

My exclusion restriction is achieved by including four variables in the first stage: number of children under 16; a binary variable if a person is married or lives with a partner, and mother's and father's educational qualifications.

In the second stage (9) is estimated for $\{N, M\}$, using the estimates $\hat{\lambda}_{it}$. In the case when the missing data in the unbalanced panel is random, that is, if there is no sample selection bias, then $\rho = 0$.

Let $\boldsymbol{\chi}_{it}^k = \{b_t^k, \boldsymbol{Q}_i^k, \boldsymbol{X}_{it}^k, \bar{\boldsymbol{X}}_i^k, \bar{b}_i^k, \hat{\lambda}_{it}^k\}$ and $\boldsymbol{B}^k = \{\mu^k, \boldsymbol{\delta}^k, \boldsymbol{\beta}^k, \varrho^k, \boldsymbol{\xi}^k, \eta^k, \rho^k\}$ with $k = \{N, M\}$, then (3) can be written as:

$$\hat{R} = (\bar{\boldsymbol{\chi}}_{it}^N - \bar{\boldsymbol{\chi}}_{it}^M)'\hat{\boldsymbol{B}}^N + \bar{\boldsymbol{\chi}}_{it}^{N\prime}(\hat{\boldsymbol{B}}^N - \hat{\boldsymbol{B}}^M) \tag{11}$$

(11) is the final version of B-O decomposition I estimate, where $(\bar{\boldsymbol{\chi}}_{it}^N - \bar{\boldsymbol{\chi}}_{it}^M)'\hat{\boldsymbol{B}}^N$ is the explained component, and $\bar{\boldsymbol{\chi}}_{it}^{N\prime}(\hat{\boldsymbol{B}}^N - \hat{\boldsymbol{B}}^M)$ is the unexplained difference in labour market outcomes.

After Blinder-Oaxaca decomposition, I use the results of decomposition to estimate the effect of labour market discrimination on the welfare dependency of immigrants compared with natives. The estimate of discrimination is the unexplained income differential from B-O decomposition in region τ in period $(t-1)$: D_{t-1}^{τ}. The welfare dependency is the probability of claiming benefits in period t. We expect D_{t-1}^{τ} to affect immigrants' propensity to claim benefits in period t, yet immigrants' behaviour and circumstances in period t should have no effect on income discrimination in period $(t-1)$. D_{t-1}^{τ} is the demeaned value of discrimination in region τ.

I estimate the effect of labour market discrimination on the probability of claiming benefits using panel data and a linear probability model. As in the case of B-O decomposition, the relevant method is the fixed effects method due to similar assumptions. In order to compare immigrants with natives by using a dummy variable for immigrants, I proceed with estimating the linear probabilities model by CRE to allow for time invariant variables. Therefore, the effect of discrimination on claiming welfare benefits is estimated based on the following equation (Wooldridge, 2010, 2015):

$$P(y_{it}=1|\boldsymbol{X}_{it},D^{\tau}_{t-1},M_i,a_i,s_{it}=1)=\boldsymbol{X}'_{it}\boldsymbol{\alpha}+D^{\tau}_{t-1}\beta+D^{\tau}_{t-1}M_i\theta+M_i\lambda+b_t\mu+\bar{\boldsymbol{X}}_i\boldsymbol{\xi}+\bar{b}_i\eta+a_i+\epsilon_{it}, \tag{12}$$

assuming $E(\boldsymbol{X}'_{it}\epsilon_{it})=0$ and $Cov(s_{it},\epsilon_{it})=0$; and where M_i is a binary variable for an individual being an immigrant; $\bar{\boldsymbol{X}}_i=T^{-1}\sum_{t=1}^{T}s_{it}\boldsymbol{X}_{it}$ and $\bar{b}_i=T^{-1}\sum_{t=1}^{T}s_{it}b_t$.

In (12), the impact of discrimination on welfare dependency of immigrants is given by θ.

3.2 Data

In order to test the research question, this paper uses the data from Waves 1 to 6 of the Main survey of the UK Household Longitudinal Study, Understanding Society (USoc), which covers years 2009-2014. I narrow down the sample to natives and second generation immigrants. Natives are defined as white individuals born in the UK, whose parents and grandparents were born in the UK. Since we look at discrimination, I include only white individuals in the definition of natives to limit any bias from heterogeneity of native population. Immigrants are defined as individuals born in the UK with parents being born outside the UK.

Summary statistics

The age range of individuals in the sample is limited to native and immigrant males aged 18 to 67, and females aged 18 to 60-65, depending on the year of birth. The age 18 is chosen since individuals are eligible to claim benefits from that age. I also limit the sample to under pension age to have only working age individuals in the sample as the research topic concerns individuals in the labour market. State pension age in the UK for the period under consideration was 65[2]. Pension age for women was 60 before and up to 2009 (women born in December 1953) and 65 starting year 2010 (women born between 6 April 1950 and 5

[2] https://www.nidirect.gov.uk/articles/check-your-state-pension-age

December 1953), that is, all women born before 1950 are in the retirement age for the period under consideration.

In addition to individuals of state pension age, I exclude self-reported retirees, according to w_jbstat. I also exclude individuals in full-time education.

Since the survey data is prone to attrition, only those individuals who stay in the survey over the six waves are included in the study, so that the sample included is strongly balanced. I also exclude any observation with missing data points for any of the variables considered, thus including only complete cases, which is further discussed in the next subsection.

I use positive log net monthly income from labour to measure labour market outcomes of natives and immigrants. Income from labour includes net monthly earnings from main job, net monthly income from self-employment and net monthly earnings from a second job.

Table 1: Summary statistics on monthly income from labour and benefits

Natives										
	Income from labour					**Benefits**				
year	**mean**	**max**	**min**	**sd**	**N**	**mean**	**max**	**min**	**sd**	**N**
2009	1506	15000	0.1	1263	6539	449	3201	1.1	435	3416
2010	1511	15000	0.1	1215	6585	479	4617	2.5	475	3595
2011	1522	15000	0.1	1142	6453	516	15000	0.1	556	3584
2012	1558	15000	0.1	1210	6353	533	4677	3.2	527	3441
2013	1568	15000	0	1181	6286	571	15000	1.1	649	3255
2014	1639	15000	0.8	1293	6144	529	4343	1.7	530	3411
Immigrants										
	Income from labour					**Benefits**				
year	**mean**	**max**	**min**	**sd**	**N**	**mean**	**max**	**min**	**sd**	**N**
2009	1543	8333	1	1023	567	567	2671	4.3	490	424
2010	1585	15000	14.1	1349	567	630	5004	13.0	617	453
2011	1564	15000	2.5	1111	572	617	3627	11.0	587	456
2012	1542	15000	5.8	1237	579	635	3458	20.0	600	435
2013	1582	15000	4.3	1209	591	659	3802	10.0	615	414
2014	1640	9944	12	1082	576	631	5246	8.3	622	423

Notes: Natives are white individuals born in the UK, whose parents and grandparents were born in the UK. Immigrants are individuals born in the UK with parents being born outside the UK. All individuals included are of working age - from 18 years old to the retirement age.
Income from labour includes monthly net positive earnings from first and second jobs and positive net self-employment income in GBP.
Benefits include total monthly state benefits in GBP, that comprise of the sum of the following: income support, job seeker's allowance (unemployment benefit); child benefits; maternity allowance; tax credits; housing benefit, council tax benefit; sickness, disability and incapacity benefits; state retirement pension; a widow's or war widow's pension; a widowed mother's allowance / widowed parent's allowance; income from any other state benefit.
Source: Understanding Society.

For estimating probabilities of claiming benefits, I use the data on the positive value of social benefits. Social benefits include total monthly benefits, that comprise of the sum of the following: income support, job seeker's allowance (unemployment benefit), child benefits (including lone-parent child benefit payments), maternity allowance, tax credits, housing benefit, council tax benefit (offset against council tax); sickness, disability and incapacity benefits; state retirement pension; a widow's or war widow's pension; a widowed mother's allowance / widowed parent's allowance; income from any other state benefit. Since the sample excludes individuals of pension age, individuals receiving state retirement (old-age) pension are excluded form the sample. All the tables in the paper are based on the sample as defined above.

Table 1 shows the statistics on net personal income from labour and social benefits. Both variables are top coded up to 15000.

Average income of natives and immigrants are similar on average, although varies over years. Average benefits, on the other hand, is higher for natives.

In addition to total amount of income from benefits reported, USoc also reports data on types of benefits claimed, without specifying the amount. Table 2 shows the breakdown of types of benefits claimed by natives and immigrants. Child and family benefits constitute equally the largest part of benefits for both natives and immigrants, followed by tax credits. Slightly higher share of immigrants claims unemployment benefits compared with natives, as well as slightly higher share claims housing or council tax benefits. Lower share of immigrants, compared with natives, claims sickness, disability or incapacity benefits.

Transition matrices

In order to utilise panel data and analyse the impact of discrimination on welfare dependency, we need to check whether there is any transition in and out of welfare from year to year. Table 3 shows the transition in and out of welfare by natives and immigrants. If an individual of working age has positive social benefits in period t, then they are considered to be on welfare (Yes), and not - otherwise (No). For instance, from 2009 to 2010, 8.3% of natives transitioned from not being on welfare to being on welfare, and 6.7% transitioned from being on welfare to not being on welfare. For the same year, more immigrants, 9.7%, transitioned into welfare, and 6.5% - out of welfare. Generally, there is a trend of decreasing welfare dependency for both natives and immigrants following the post-2008 crisis, except for 2014,

Table 2: Breakdown of shares of social benefits by source

Natives						
	2009	**2010**	**2011**	**2012**	**2013**	**2014**
Unemployment benefits	4.1	4.4	4.2	4.2	4.2	3.1
Income support	5.3	5.5	5.2	4.8	4.6	4.5
Child or family benefits	37.0	36.0	36.0	37.5	37.9	38.8
Tax credits	30.8	30.6	29.0	25.9	23.7	22.8
Sickness, disability or incapacity benefits	9.0	9.1	9.9	10.9	11.9	13.5
Housing or council tax benefits	11.3	11.9	12.8	13.4	14.4	13.0
Other benefits	2.5	2.5	3.0	3.1	3.4	4.3
Total	100.0	100.0	100.0	100.0	100.0	100.0
Immigrants						
	2009	**2010**	**2011**	**2012**	**2013**	**2014**
Unemployment benefits	5.6	6.3	6.3	6.0	5.2	4.0
Income support	5.6	6.6	5.4	5.1	5.0	4.3
Child or family benefits	37.5	35.8	35.9	37.0	36.9	38.2
Tax credits	31.2	30.4	29.2	27.4	26.9	26.8
Sickness, disability or incapacity benefits	4.8	5.3	6.5	7.8	7.8	9.5
Housing or council tax benefits	13.8	14.2	14.7	15.0	15.8	14.0
Other benefits	1.4	1.2	1.9	1.8	2.3	3.2
Total	100.0	100.0	100.0	100.0	100.0	100.0

Note: Each row shows the percentage share of respective types of benefits in total benefits for each year. Natives are white individuals born in the UK, whose parents and grandparents were born in the UK. Immigrants are individuals born in the UK with parents being born outside the UK. All individuals included are of working age - from 18 years old to the retirement age.

when there is a slight increase in welfare dependency. Table 3 shows that a higher share of immigrants are on welfare benefits, compared with natives.

One factor to consider is whether the transition is different for males and females, as the higher share of immigrants on welfare dependency might be attributable to lower share of females in the labour force and higher levels of child benefits for women. Tables 4-5 show the transition into and out of welfare by native and immigrant males and females. Indeed, a higher share of females of both natives and immigrants are on welfare benefits compared with males, with the share of immigrant women being around 12 percentage points higher than for natives. The transition in and out of welfare is higher for men - both natives and immigrants.

Another factor to consider is whether the proportions of young people are different for immigrants versus natives and whether the differences in welfare dependency of immigrants and natives are attributable to that. To look at that question, I split the sample into two age groups: 40 years and under, and 41 years and over. Tables 6-7 are on transition matrices of natives and immigrants in the two age groups. Immigrants have a higher share of younger individuals aged 40 and under - 58%, versus 41% for natives. Younger people tend to claim more benefits in the case of both immigrants and natives. However, the shares are higher for the younger group of immigrants compared with natives and the group of immigrants aged 41 and over.

Thus, when looking at raw statistics of welfare dependency, a larger share of immigrants tend to claim benefits compared with natives, which is consistent with previous studies. The next question to discuss is whether these patterns are the same when considering the observable characteristics of natives and immigrants, and most importantly, whether the patterns are dependent on income discrimination in the labour market.

Table 3: Year on year transition matrices on welfare dependency: immigrants vs. natives

Natives

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N	2009	4,164	676	4,840	2010	4,201	509	4,710	2011	4,296	463	4,759	2012	4,427	462	4,889	2013	4,214	787	5,001
	%		51.0	8.3	59.3		51.4	6.2	57.7		52.6	5.7	58.3		54.2	5.7	59.9		51.6	9.6	61.2
Yes	N		546	2,780	3,326		558	2,898	3,456		593	2,814	3,407		574	2,703	3,277		561	2,604	3,165
	%		6.7	34.0	40.7		6.8	35.5	42.3		7.3	34.5	41.7		7.0	33.1	40.1		6.9	31.9	38.8
Total	N		4,710	3,456	8,166		4,759	3,407	8,166		4,889	3,277	8,166		5,001	3,165	8,166		4,775	3,391	8,166
	%		57.7	42.3	100.0		58.3	41.7	100.0		59.9	40.1	100.0		61.2	38.8	100.0		58.5	41.5	100.0

Immigrants

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N	2009	331	81	412	2010	321	64	385	2011	340	42	382	2012	359	44	403	2013	350	73	423
	%		39.6	9.7	49.3		38.4	7.7	46.1		40.7	5.0	45.7		43.0	5.3	48.3		41.9	8.7	50.7
Yes	N		54	369	423		61	389	450		63	390	453		64	368	432		63	349	412
	%		6.5	44.2	50.7		7.3	46.6	53.9		7.5	46.7	54.3		7.7	44.1	51.7		7.5	41.8	49.3
Total	N		385	450	835		382	453	835		403	432	835		423	412	835		413	422	835
	%		46.1	53.9	100.0		45.7	54.3	100.0		48.3	51.7	100.0		50.7	49.3	100.0		49.5	50.5	100.0

Note: Natives are white individuals born in the UK, whose parents and grandparents were born in the UK. Immigrants are individuals born in the UK with parents being born outside the UK.
All individuals included are of working age - from 18 years old to the retirement age, and the sample is strictly balanced.
(No) Was not on welfare/ had zero monthly state benefits. (Yes) Was on welfare/ had positive amount of monthly state benefits.
(N) Number of individuals. (%) Share in total for the period.

Table 4: Year on year transition matrices on welfare dependency of natives: males vs. females

Males

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N	2009	2,364	313	2,677	2010	2,389	248	2,637	2011	2,462	195	2,657	2012	2,490	221	2,711	2013	2,322	393	2,715
	%		66.4	8.8	75.2		67.1	7.0	74.1		69.2	5.5	74.7		70.0	6.2	76.2		65.2	11.0	76.3
Yes	N		273	609	882		268	654	922		249	653	902		225	623	848		219	625	844
	%		7.7	17.1	24.8		7.5	18.4	25.9		7.0	18.3	25.3		6.3	17.5	23.8		6.2	17.6	23.7
Total	N		2,637	922	3,559		2,657	902	3,559		2,711	848	3,559		2,715	844	3,559		2,541	1,018	3,559
	%		74.1	25.9	100.0		74.7	25.3	100.0		76.2	23.8	100.0		76.3	23.7	100.0		71.4	28.6	100.0

Females

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N	2009	1,800	363	2,163	2010	1,812	261	2,073	2011	1,834	268	2,102	2012	1,937	241	2,178	2013	1,892	394	2,286
	%		39.1	7.9	47.0		39.3	5.7	45.0		39.8	5.8	45.6		42.0	5.2	47.3		41.1	8.6	49.6
Yes	N		273	2,171	2,444		290	2,244	2,534		344	2,161	2,505		349	2,080	2,429		342	1,979	2,321
	%		5.9	47.1	53.0		6.3	48.7	55.0		7.5	46.9	54.4		7.6	45.1	52.7		7.4	43.0	50.4
Total	N		2,073	2,534	4,607		2,102	2,505	4,607		2,178	2,429	4,607		2,286	2,321	4,607		2,234	2,373	4,607
	%		45.0	55.0	100.0		45.6	54.4	100.0		47.3	52.7	100.0		49.6	50.4	100.0		48.5	51.5	100.0

Note: Natives are white individuals born in the UK, whose parents and grandparents were born in the UK.
All individuals included are of working age - from 18 years old to the retirement age, and the sample is strictly balanced.
(No) Was not on welfare/ had zero monthly state benefits. (Yes) Was on welfare/ had positive amount of monthly state benefits.
(N) Number of individuals. (%) Share in total for the period.

Table 5: Year on year transition matrices on welfare dependency of immigrants: males vs. females

Males

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N		199	49	248		194	27	221		205	18	223		211	27	238		198	40	238
	%		58.5	14.4	72.9		57.1	7.9	65.0		60.3	5.3	65.6		62.1	7.9	70.0		58.2	11.8	70.0
Yes	N	2009	22	70	92	2010	29	90	119	2011	33	84	117	2012	27	75	102	2013	32	70	102
	%		6.5	20.6	27.1		8.5	26.5	35.0		9.7	24.7	34.4		7.9	22.1	30.0		9.4	20.6	30.0
Total	N		221	119	340		223	117	340		238	102	340		238	102	340		230	110	340
	%		65.0	35.0	100.0		65.6	34.4	100.0		70.0	30.0	100.0		70.0	30.0	100.0		67.6	32.4	100.0

Females

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N		132	32	164		127	37	164		135	24	159		148	17	165		152	33	185
	%		26.7	6.5	33.1		25.7	7.5	33.1		27.3	4.8	32.1		29.9	3.4	33.3		30.7	6.7	37.4
Yes	N	2009	32	299	331	2010	32	299	331	2011	30	306	336	2012	37	293	330	2013	31	279	310
	%		6.5	60.4	66.9		6.5	60.4	66.9		6.1	61.8	67.9		7.5	59.2	66.7		6.3	56.4	62.6
Total	N		164	331	495		159	336	495		165	330	495		185	310	495		183	312	495
	%		33.1	66.9	100.0		32.1	67.9	100.0		33.3	66.7	100.0		37.4	62.6	100.0		37.0	63.0	100.0

Note: Immigrants are individuals born in the UK with parents being born outside the UK.
All individuals included are of working age - from 18 years old to the retirement age, and the sample is strictly balanced.
(No) Was not on welfare/ had zero monthly state benefits. (Yes) Was on welfare/ had positive amount of monthly state benefits.
(N) Number of individuals. (%) Share in total for the period.

Table 6: Year on year transition matrices on welfare dependency of natives by age groups

Aged 40 years and under

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N		1,463	283	1,746		1,459	241	1,700		1,458	233	1,691		1,505	214	1,719		1,418	358	1,776
	%		43.2	8.4	51.6		43.1	7.1	50.3		43.1	6.9	50.0		44.5	6.3	50.8		41.9	10.6	52.5
Yes	N	2009	237	1,400	1,637	2010	232	1,451	1,683	2011	261	1,431	1,692	2012	271	1,393	1,664	2013	270	1,337	1,607
	%		7.0	41.4	48.4		6.9	42.9	49.7		7.7	42.3	50.0		8.0	41.2	49.2		8.0	39.5	47.5
Total	N		1,700	1,683	3,383		1,691	1,692	3,383		1,719	1,664	3,383		1,776	1,607	3,383		1,688	1,695	3,383
	%		50.3	49.7	100.0		50.0	50.0	100.0		50.8	49.2	100.0		52.5	47.5	100.0		49.9	50.1	100.0

Aged 41 years and over

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N		2,666	283	2,949		2,733	269	3,002		2,834	222	3,056		2,918	244	3,162		2,789	431	3,220
	%		55.7	5.9	61.7		57.1	5.6	62.8		59.3	4.6	63.9		61.0	5.1	66.1		58.3	9.0	67.3
Yes	N	2009	336	1,498	1,834	2010	323	1,458	1,781	2011	328	1,399	1,727	2012	302	1,319	1,621	2013	271	1,292	1,563
	%		7.0	31.3	38.3		6.8	30.5	37.2		6.9	29.2	36.1		6.3	27.6	33.9		5.7	27.0	32.7
Total	N		3,002	1,781	4,783		3,056	1,727	4,783		3,162	1,621	4,783		3,220	1,563	4,783		3,060	1,723	4,783
	%		62.8	37.2	100.0		63.9	36.1	100.0		66.1	33.9	100.0		67.3	32.7	100.0		64.0	36.0	100.0

Note: Natives are white individuals born in the UK, whose parents and grandparents were born in the UK.
All individuals included are of working age - from 18 years old to the retirement age, and the sample is strictly balanced.
Individuals are considered "aged 40 and under" and "aged 41 and over" based on their age in year 2009.
(No) Was not on welfare/ had zero monthly state benefits. (Yes) Was on welfare/ had positive amount of monthly state benefits.
(N) Number of individuals. (%) Share in total for the period.

Table 7: Year on year transition matrices on welfare dependency of immigrants by age groups

Aged 40 years and under

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N	2009	187	56	243	2010	173	40	213	2011	181	34	215	2012	187	34	221	2013	177	45	222
	%		38.4	11.5	49.9		35.5	8.2	43.7		37.2	7.0	44.1		38.4	7.0	45.4		36.3	9.2	45.6
Yes	N		26	218	244		42	232	274		40	232	272		35	231	266		39	226	265
	%		5.3	44.8	50.1		8.6	47.6	56.3		8.2	47.6	55.9		7.2	47.4	54.6		8.0	46.4	54.4
Total	N		213	274	487		215	272	487		221	266	487		222	265	487		216	271	487
	%		43.7	56.3	100.0		44.1	55.9	100.0		45.4	54.6	100.0		45.6	54.4	100.0		44.4	55.6	100.0

Aged 41 years and over

			2010				2011				2012				2013				2014		
			No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total		No	Yes	Total
No	N	2009	144	25	169	2010	148	24	172	2011	159	8	167	2012	172	10	182	2013	173	28	201
	%		41.4	7.2	48.6		42.5	6.9	49.4		45.7	2.3	48.0		49.4	2.9	52.3		49.7	8.0	57.8
Yes	N		28	151	179		19	157	176		23	158	181		29	137	166		24	123	147
	%		8.0	43.4	51.4		5.5	45.1	50.6		6.6	45.4	52.0		8.3	39.4	47.7		6.9	35.3	42.2
Total	N		172	176	348		167	181	348		182	166	348		201	147	348		197	151	348
	%		49.4	50.6	100.0		48.0	52.0	100.0		52.3	47.7	100.0		57.8	42.2	100.0		56.6	43.4	100.0

Note: Immigrants are individuals born in the UK with parents being born outside the UK.
All individuals included are of working age - from 18 years old to the retirement age, and the sample is strictly balanced.
Individuals are considered "aged 40 and under" and "aged 41 and over" based on their age in year 2009.
(No) Was not on welfare/ had zero monthly state benefits. (Yes) Was on welfare/ had positive amount of monthly state benefits.
(N) Number of individuals. (%) Share in total for the period.

Heterogeneity

Table 8 includes summary statistics of natives as defined above and groups of immigrants by country of origin of the father. The following breakdown of the countries is due to the sample sizes of immigrants. Where the sample size is enough to have the country as a separate group, I include it separately, otherwise, I group them according to country groupings used by Office for National Statistics for International Passenger Survey[3]. "EU(EEA)" includes all European Union member-countries (excluding the UK), and Iceland, Liechtenstein, Norway and Switzerland. "Other Africa" includes Sub-Saharan African countries. "Latin America" includes Central and South American countries. "Other" includes all other countries not included in the previous categories.

Here, the variables of average monthly income from labour and benefits are averages of the entire sample, that is, total of individuals that are receiving income from labour and/or benefits, as opposed to Table **??**, where the statistics are from sub-samples of individuals who has income from labour, and individuals who receive income from benefits.

Table 8: Summary statistics of immigrant versus native characteristics

	Natives	**Immigrants by country of origin of father**						
		EU(EEA)	**India**	**Pakistan**	**Bangladesh**	**Other Africa**	**Latin America**	**Other**
Avg. monthly labour income	1194.3	1242.7	1133.2	640.1	733.3	1366.1	1112.2	1525.3
Avg. monthly benefits	213.5	224.8	243.0	488.2	397.4	321.9	389.9	219.3
Share of labour force participation, %	80.7	81.0	79.2	60.3	69.2	80.9	81.7	84.4
Avg. age	45	49	40	35	31	36	45	42
Avg. years of school	11.1	11.9	12.7	12.1	10.9	14.1	12.1	13.6
Avg. number of children under 16	0.6	0.5	0.9	1.6	1.3	0.9	0.8	0.7
Share of females, %	55.3	56.9	57.9	55.8	60.4	66.3	63.0	52.6
N	49679	756	1062	807	364	517	1004	449

Notes: Average monthly income and benefits are computed based on the entire sample, including zero values.
"EU(EEA)" includes all European Union member-countries (excluding the UK), and Iceland, Liechtenstein, Norway and Switzerland. "Other Africa" includes countries in Sub-Saharan Africa. "Latin America" includes countries in Central and South America. "Other" includes all other countries.
Source: Understanding Society.

Natives and EU immigrants have, in general, similar characteristics, whereas there is a lot of heterogeneity across non-EU immigrants. EU immigrants have, on average, slightly higher income from labour than natives. Income of non-EU immigrants varies significantly depending on the country of origin of immigrants. Non-EU immigrants from *other* countries have the highest average income followed by immigrants from Other Africa. Average monthly benefits exceeds that of natives for all immigrant groups.

[3] www.ons.gov.uk

EU immigrants are, on average, 3 years older than natives, whereas non-EU immigrants are about 6 years younger. All immigrants have a higher average years of schooling than natives do, except for immigrants for Bangladesh, for whom schooling is similar to natives. Immigrants have, on average, more children under 16 compared with natives, except for EU immigrants, who have slightly fewer. Since a higher proportion of females claim welfare benefits, the next important indicator is the share of females across groups. Natives have a lower share of females compared with all groups of immigrants except for immigrants from the "other" group.

Labour force participation is similar for natives and EU immigrants, whereas it varies a lot across non-EU immigrants, with the highest being for second generation immigrants from the "other" group. This variation might be an important issue when discussing income from labour across groups as it yields potential sample selection bias. Labour force participation here and in the rest of the paper is defined as the share of individuals with positive income or full or part-time employment, and individuals who are self-reported as unemployed.

Table 9: Labour force participation by groups

	Male	**Female**
Natives	84.1	77.9
EU	79.4	82.1
India	88.1	72.7
Pakistan	88.8	37.8
Bangladesh	84.0	59.5
Other Africa	90.2	76.1
Latin America	84.4	80.1
Other	89.2	80.1

Notes: The country groups of immigrants are based on father's country of birth.
Labour force participation is computed as share of individuals who are employed/have positive earnings or are unemployed to total individuals in the group.

Table 9 includes breakdown of labour force participation rates (LFP) by natives and immigrant groups and male versus female. Labour force participation varies substantially across groups, particularly for females. This creates an issue of sample selection bias, as we do not observe income of individuals who are not in labour force. Therefore, we need to correct for the selection before comparing the income of natives and immigrants across groups and for males and females.

USoc includes Ethnic Minority Boost sample, where they oversample individuals from certain ethnic minority groups, including Indian, Pakistani, Bangladeshi, Caribbean, and African. Table 17 in Appendix B includes LFP with the adjusted weights, which accounts for oversampling. LFP adjusted for sample weights are, on average, similar to the unadjusted LFP. Therefore, we proceed with the unadjusted sample.

From the discussion above we can see that the average labour income of immigrants and natives is similar. However, immigrants and natives have different characteristics. Therefore, in order to assess the differences in labour income and to understand whether there is an income gap between natives and immigrants, we should consider the differences in characteristics of immigrants and natives.

Furthermore, we can see that a higher share of immigrants tend to claim benefits compared with natives. As the next step we need to understand whether immigrants are more likely to claim benefits given eligibility, that is, when controlling for observables, and whether immigrants' probability of claiming benefits is affected by income gap in the labour market, if there is one.

4 Results

Income discrimination

I start by decomposing income from labour for natives and immigrants by Blinder-Oaxaca decomposition following the methodology described in Section 3.1. As discussed, Fixed Effects method is the appropriate estimation method to decompose log wages of natives and immigrants, which is also confirmed by the Hausman test. I apply Correlated Random Effect method, as CRE results in the same coefficients for time-variant variables as Fixed Effects, but also allows for time-invariant variables, such as education, including parental education, gender, industry. Therefore, as part of CRE, time averages of all time-variant variables are included, as well as time averages of year effects, while using only complete cases of data. I also use robust standard errors to account for heteroskedasticity.

Table 10 shows the results of the decomposition, which includes the results of the decomposition without the adjustment for sample selection bias, and with the adjustment as in equations (9) and (11). For Heckman correction I use four exclusion restrictions, number of children aged under 16, a binary variable for being married or living with a partner, and mother's and father's educational qualifications. Since the patterns of labour force participation might differ for natives and different groups of immigrants, as well as for men and women, I conduct the 1st stage separately for different groups. The results of the first stage of Heckman correction are included in Table 18 of Appendix B. The coefficients across groups indeed vary significantly. I construct the final Mills ratio for the second stage of Heckman correction from these subgroups.

The difference in income of natives and immigrants is not significantly different from zero when the decomposition results are not adjusted for sample selection. When adjusted, however, natives' income exceeds that of immigrants' by 12%.

In individual regressions of B-O decomposition, the coefficients of the inverse Mills ratio, ρ from (9), are different from zero (negative) at 1% significance level for immigrants. Hence, since the sample selection is not random, there is a negative selection, and I will proceed with the model corrected for sample selection bias.

The Heckman corrected income gap of around 12% is due to unexplained income differen-

Table 10: Blinder-Oaxaca decomposition for natives and immigrants

	Correlated random effects			CRE, corrected for selection bias		
	overall	explained	unexplained	overall	explained	unexplained
group_1: natives	7.093***			7.291***		
	(0.006)			(0.026)		
group_2: migrants	7.111***			7.173***		
	(0.020)			(0.046)		
difference	–0.018			0.118**		
	(0.021)			(0.053)		
explained	–0.119***			–0.124***		
	(0.014)			(0.014)		
unexplained	0.100***			0.242***		
	(0.022)			(0.053)		
potential experience (years)		0.106**	0.548		0.092*	0.599
		(0.049)	(0.678)		(0.050)	(0.677)
squared potential experience (years)		–0.112***	0.035		–0.100***	0.019
		(0.015)	(0.122)		(0.015)	(0.126)
years of education		0.037***	–0.373*		0.037***	–0.346*
		(0.006)	(0.194)		(0.006)	(0.195)
years of education squared		–0.063***	0.283		–0.055***	0.214
		(0.007)	(0.173)		(0.007)	(0.173)
male		0.000	0.021**		0.000	0.018*
		(0.001)	(0.010)		(0.001)	(0.010)
female		0.000	–0.025**		0.000	–0.021*
		(0.001)	(0.012)		(0.001)	(0.012)
urban area		0.002	–0.064		0.002	–0.055
		(0.002)	(0.091)		(0.002)	(0.096)
rural area		0.002	0.003		0.002	0.002
		(0.002)	(0.004)		(0.002)	(0.004)
Occupational controls		X	X		X	X
Industry controls		X	X		X	X
Regional controls		X	X		X	X
Time effects		X	X		X	X
Time averages		X	X		X	X
Other controls		X	X		X	X
N	40899			40873		

Note: CRE, corrected for selection bias - correlated random effects with Heckman correction.
The dependent variable is log income from labour.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

tial. Attributable to explained characteristics, immigrants' income would have been around 12.4% higher than natives'. However, that advantage for immigrants is cancelled due to unexplained difference of around 24.2% of income of natives.

Immigrants' income, attributable to the potential experience, is lower by 9.2% (significant at 10%) compared with natives up until potential experience is higher, when the situation reverses, attributable to older native population.

In terms of education, natives get lower return to education until their education is about 7.5 years, from which point onwards return to education for them increases. This relation does not hold for immigrants. That is likely to be due to native population being from an older generation, who were more likely to leave education early for work. This results in an unexplained difference in education of 34.6% in favour of immigrants.

There seems to be a small unexplained difference (significant at 10%) in favour of female immigrants and and a disadvantage against male immigrants.

EU and non-EU. To understand whether the trend holds for different groups of immigrants, I look at B-O decomposition for EU and non-EU immigrants. Table 11 shows the results of B-O decomposition of log earnings of natives versus EU immigrants and natives versus non-EU immigrants separately for men and women.

The results are very different for EU and non-EU immigrants. The difference in log income of natives and immigrants is not statistically significant for either men or women. However, attributable to the difference explained by observable characteristics, female EU immigrants' income on average exceeds that of native females by around 17%. There seems to be no unexplained income differential for EU immigrants versus natives.

The picture is different for non-EU immigrants. Non-EU men have around 18% lower income compared with native men. Based on the individual characteristics, non-EU men would have had around 6.7% higher income than native men. However, the income of non-EU men is lower due to the unexplained difference of 24.5%.

Female non-EU immigrants' income difference, on the other hand, is not statistically significantly different from zero since the explained and unexplained difference neutralise each other. Non-EU women would have had income, explained by observable characteristics, by 18.6% higher than native women. However, this is offset by unexplained difference of around 24% in favour of natives.

Even though the scale of income discrimination is slightly higher against non-EU men compared with women, the pattern is the same for non-EU second generation immigrants.

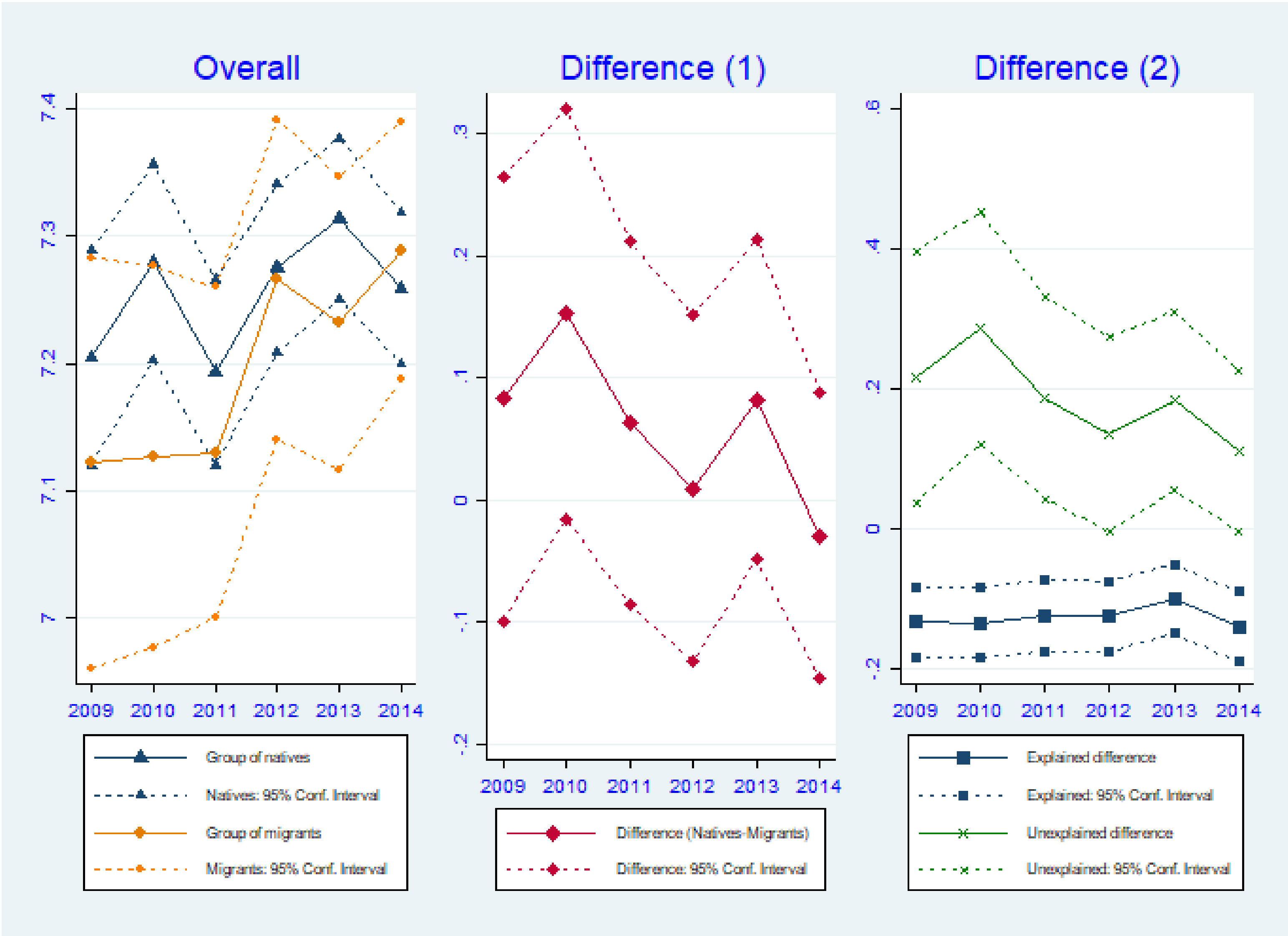
The results of B-O decomposition by individual countries are included in Table 19 in Appendix C.

Table 11: B-O decomposition for natives versus EU / non-EU immigrants: men and women

	EU						Non-EU					
	male			female			male			female		
	overall	explained	unexplained	overall	explained	unexplained	overall	explained	unexplained	overall	explained	unexplained
group_1: natives	7.476***			7.125***			7.476***			7.125***		
	(0.042)			(0.034)			(0.042)			(0.034)		
group_2: migrants	7.600***			7.179***			7.299***			7.072***		
	(0.496)			(0.270)			(0.078)			(0.070)		
difference	–0.123			–0.054			0.178**			0.053		
	(0.497)			(0.273)			(0.088)			(0.078)		
explained	–0.022			–0.173***			–0.067***			–0.186***		
	(0.029)			(0.032)			(0.021)			(0.022)		
unexplained	–0.101			0.119			0.245***			0.240***		
	(0.497)			(0.271)			(0.090)			(0.079)		
potential experience (years)		–0.064*	2.413		–0.019	0.962		0.249**	–1.131		0.024	1.721**
		(0.036)	(5.330)		(0.061)	(2.053)		(0.110)	(1.156)		(0.077)	(0.795)
squared potential experience (years)		0.043*	–0.932		0.025**	–0.567		–0.237***	0.122		–0.046**	–0.043
		(0.025)	(0.822)		(0.011)	(0.696)		(0.038)	(0.221)		(0.018)	(0.170)
years of education		0.022**	–0.564		–0.003	–0.617		0.047***	–0.803**		0.041***	–0.040
		(0.011)	(2.736)		(0.008)	(0.517)		(0.012)	(0.360)		(0.009)	(0.331)
years of education squared		0.000	0.000		0.000	0.000		0.000	0.000		0.000	0.000
		(.)	(.)		(.)	(.)		(.)	(.)		(.)	(.)
urban area		0.000	0.000		0.000	0.000		0.000	0.000		0.000	0.000
		(.)	(.)		(.)	(.)		(.)	(.)		(.)	(.)
rural area		0.000	0.000		0.000	0.000		0.000	0.000		0.000	0.000
		(.)	(.)		(.)	(.)		(.)	(.)		(.)	(.)
Occupational controls		X	X		X	X		X	X		X	X
Industry controls		X	X		X	X		X	X		X	X
Regional controls		X	X		X	X		X	X		X	X
Time effects		X	X		X	X		X	X		X	X
Time averages		X	X		X	X		X	X		X	X
Other controls		X	X		X	X		X	X		X	X
N	17534			20542			18598			21708		

Note: The dependent variable is log income from labour.
The estimation method is Correlated random effects, corrected for selection bias.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

Figure 1: Dynamics of the results of Blinder-Oaxaca decomposition of log income from labour



Note: Decomposition of income of natives and immigrants by Blinder-Oaxaca method for each year.
Overall - log income of natives / immigrants.
Difference(1) - difference in long income of natives and immigrants.
Difference(2) - breakdown of Difference(1) into "explained difference" and "unexplained difference".

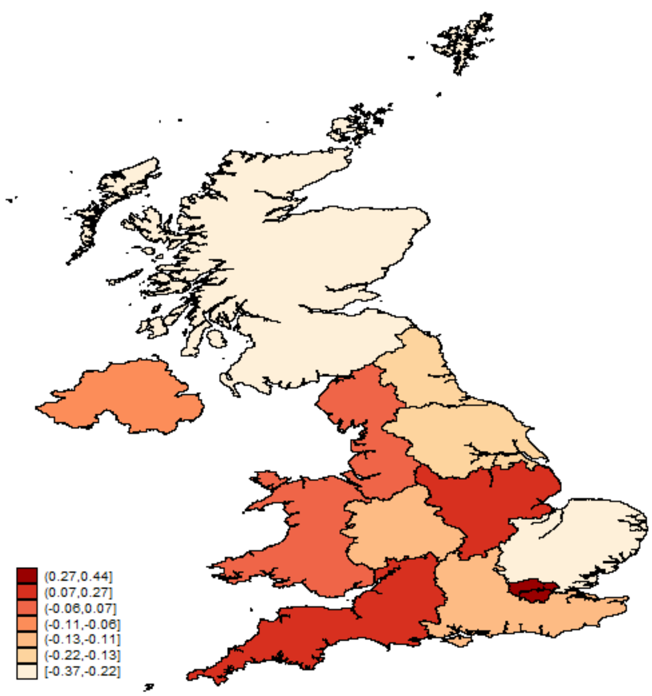
Figure 1 show the results of B-O decomposition by years, adjusted for selection bias. The average income of natives and immigrants is volatile over the years even though the explained difference is quite stable. The major part of the volatility is due to unexplained income differential between natives and immigrants. In the next stage, I use this yearly volatility to explore the effect of income discrimination on immigrants' welfare dependency.

Welfare dependency

Figure 2 includes average annual results of B-O decomposition by region. Discrimination is estimated as average unexplained difference for each region in year t and is expressed as a percent of income from labour of natives. That is, for instance, immigrants in London receive 44% less income not explained by their observable characteristics than natives, whereas immigrants in Scotland get around 37% higher income that is not explained by observables.

I now turn to analysing the welfare dependency of immigrants versus natives, and the

Figure 2: The map of average income discrimination by region, %



impact of discrimination on it. I use unexplained income difference in the UK regions in time $(t-1)$ to explore the impact of labour market discrimination on the probability of immigrants claiming benefits compared with natives. As discussed in Section 3.1, I use linear CRE for estimation. The Hausman test also confirms that the appropriate model is FE. I use linear regression for my estimations, although the results are robust to using logit regressions as well.

In this stage I introduce a new control variable, the share of individuals claiming benefits in the corresponding region, expecting the probability of claiming benefits to be positively affected by this variable. I also include the following variables as controls: number of children aged under 16, a binary variable for an individual being married or living with a partner, and parents' educational qualifications.

Table 12 shows the results of linear regressions on the probability of natives and immigrants to claim benefits, including results with all immigrants, EU immigrants only and non-EU immigrants only versus natives. In all three results the dummy variables for immigrants, including EU and non-EU immigrants, are not statistically significant, signifying that the likelihood of immigrants claiming benefits is overall not different from natives.

The effect labour market discrimination has on natives is given by the coefficient of the variable "discrimination in $(t-1)$", which is not statistically different from zero. That is, the propensity of natives to claim benefits is unaffected by discrimination against immigrants.

The share of individuals claiming benefits, as expected, positively affects the probability of claiming benefits. Potential experience, on the other hand, reduces the probability of claiming benefits, so does education squared. Squared potential experience positively affects welfare dependency since it is associated with older individuals who are more likely to claim some types of benefits.

The coefficient of interest, the coefficient of the interaction variable of the binary variables for immigrant groups and discrimination, is positive and significant (at 5% level) for total immigrants. Other things being equal, a 10% increase in income discrimination against immigrants results in an increase in welfare dependency by immigrants of 0.40%.

When looking at EU and non-EU immigrants separately, the welfare dependency of EU immigrants is unaffected by discrimination. This is in line with the results of Blinder-Oaxaca

Table 12: The impact of dicrimination on the probability of claiming benefits

	Natives / all immigrants	Natives / EU immigrants	Natives / non-EU immigrants
Discrimination in t-1	–0.005	–0.005	–0.004
	(0.005)	(0.005)	(0.005)
Immigrants × Discrimination in t-1	0.040**		
	(0.018)		
Immigrants	0.010		
	(0.013)		
EU(EEA) × Discrimination in t-1		–0.044	
		(0.044)	
EU(EEA)		0.034	
		(0.032)	
Non-EU × Discrimination in t-1			0.052***
			(0.020)
Non-EU			0.005
			(0.014)
share of individuals claiming benefits by region	0.608**	0.618**	0.580**
	(0.275)	(0.284)	(0.277)
potential experience (years)	–0.030***	–0.033***	–0.030***
	(0.008)	(0.009)	(0.008)
squared potential experience (years)	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
years of education	–0.005**	–0.005**	–0.005**
	(0.002)	(0.002)	(0.002)
female	0.148***	0.144***	0.148***
	(0.008)	(0.008)	(0.008)
urban area	0.011	0.010	0.011
	(0.008)	(0.008)	(0.008)
no of children aged under 16	0.066***	0.068***	0.066***
	(0.003)	(0.003)	(0.003)
married or lives with partner	–0.074***	–0.075***	–0.075***
	(0.007)	(0.007)	(0.007)
Occupational controls	X	X	X
Industry controls	X	X	X
Regional controls	X	X	X
Time effects	X	X	X
Time averages	X	X	X
Other controls	X	X	X
N	45508	41998	44876

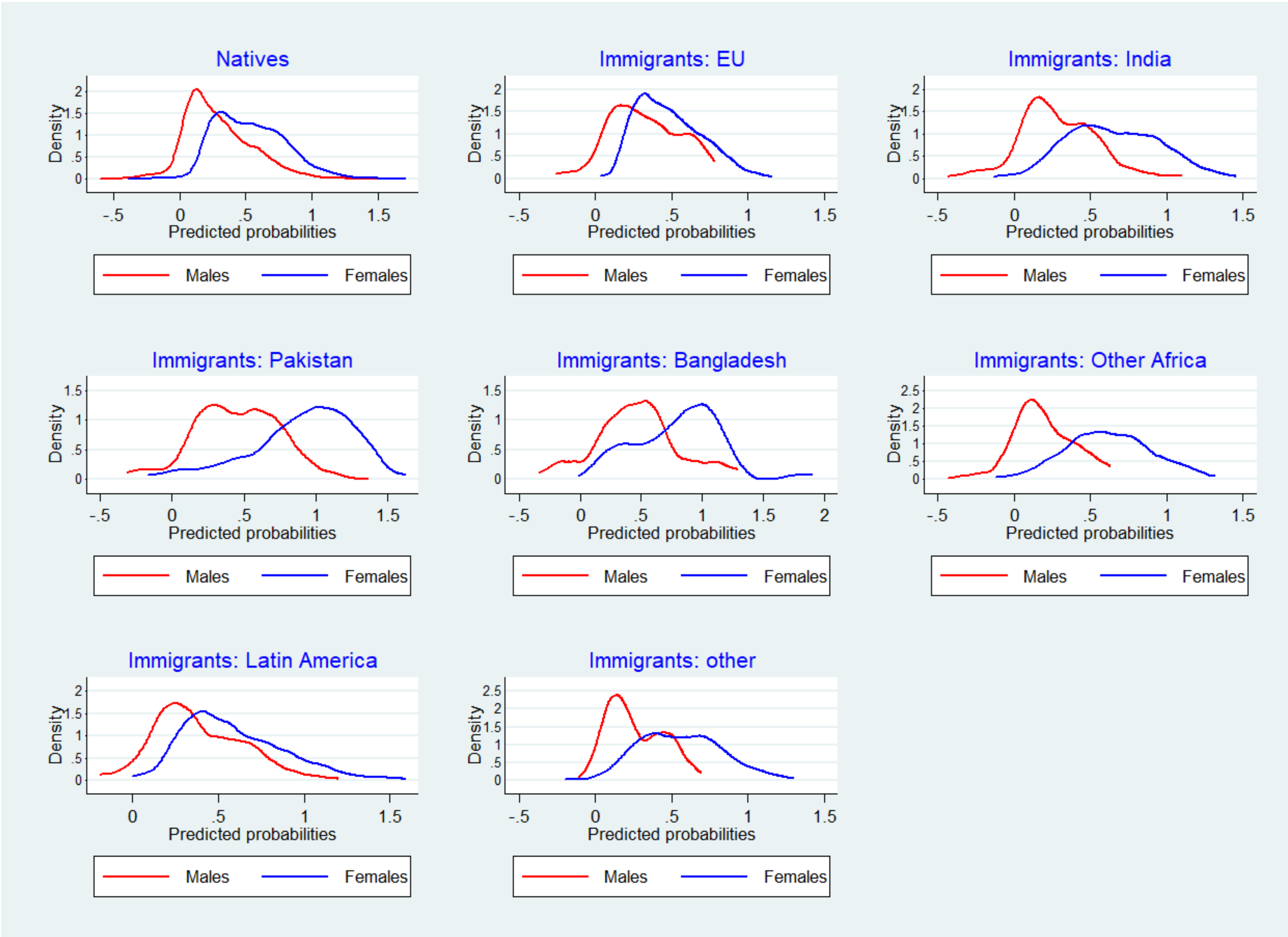
Note: The dependent variable is a binary variable, equal to 1 if an idividual claims benefits.
The estimation method is correlated random effects.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

decomposition, since we did not observe discrimination against EU immigrants.

The effect of discrimination on welfare dependency of non-EU immigrants, on the other hand, is positive and statistically significant. That is, a 10% increase in income discrimination against immigrants results in an increase in the probability of non-EU immigrants claiming benefits by 0.52%. This is also in line with the B-O decomposition results, which exhibit income discrimination against non-EU immigrants.

If we consider the situation of no discrimination, in which case the overall probability of welfare dependency of non-EU immigrants is 54%, then the highest observed income discrimination in a single year, for instance in East Midlands, increases the probability of claiming benefits to 65% in the region in that year.

Figure 3: Distributions of predicted probabilities of claiming benefits



Note: Predictions are based on model (12).

Figure 3 shows distributions of predicted probabilities of claiming benefits by male and female natives and groups of immigrants based on model (12), that is, estimates in Table 12. The likelihood of welfare dependency is very heterogeneous across and within groups. Women, in general, are more likely to claim benefits compared with women. In terms of highest and lowest probability distributions, men from Sub-Saharan Africa have the lowest probability of welfare dependency amongst men, which, on average, is 19.7%, whereas men from Bangladesh have the highest probability of claiming benefits, with the average of 46.8%. It, however, varies a lot within the groups. EU women are the least likely to be on welfare dependency amongst women, with the average of 48.8%, whereas Pakistani women are the most likely - 90.8%.

Table 13: The impact of dicrimination on the probability of claiming benefits by types of benefits

	I	II	III	IV	V	VI
Discrimination in t-1	–0.004**	0.001	0.001	0.002	–0.006**	–0.003
	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)
Immigrants × Discrimination in t-1	0.027**	0.012	0.008	–0.011	0.017	–0.004
	(0.013)	(0.011)	(0.016)	(0.017)	(0.013)	(0.011)
Immigrants	0.007	–0.026***	0.029**	0.038***	–0.024**	–0.040***
	(0.006)	(0.007)	(0.012)	(0.013)	(0.011)	(0.010)
share of individuals claiming benefits by region	0.108	0.246**	0.950***	0.590**	0.338***	0.174*
	(0.093)	(0.099)	(0.264)	(0.265)	(0.126)	(0.103)
female	–0.031***	0.007*	0.201***	0.119***	–0.001	–0.004
	(0.004)	(0.004)	(0.007)	(0.006)	(0.006)	(0.005)
no of children aged under 16	0.005***	0.007***	0.046***	0.035***	0.015***	0.014***
	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Occupational controls	X	X	X	X	X	X
Industry controls	X	X	X	X	X	X
Regional controls	X	X	X	X	X	X
Time effects	X	X	X	X	X	X
Time averages	X	X	X	X	X	X
Other controls	X	X	X	X	X	X
N	45508	45508	45508	45508	45508	45508

Note: The dependent variable is a binary variable, equal to 1 if an idividual claims benefits.
(I) unemployment benefits, (II) income support, (III) child benefits, (IV) tax credit, (V) housing or council tax, (VI) sickness, disability or incapacity benefits.
The estimation method is correlated random effects.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

When looking at the probability of claiming different types of benefits in Table 13 (detailed Table 20 in Appendix E), the effect of labour market discrimination on the probability of immigrants to claim benefits is significant in the case of unemployment benefits only. The latter is expected since, of all the benefits, unemployment benefits are most closely related to the labour market.

Interestingly, the probabilities of natives claiming unemployment and housing or council tax benefits are slightly lower with higher income discrimination against immigrants.

When looking at the probabilities by types of benefits, while controlling for individual characteristics, immigrants exhibit different behaviour when compared with natives. Immigrants are 2.9% more likely to claim child benefits compared with natives and 3.8% more likely to claim tax credit. However, immigrants are 2.6% less likely to claim income support benefits, 4% less likely to claim housing or council tax benefits, and 2.4% less likely to claim sickness, disability or incapacity benefits.

The probability of claiming benefits also tends to increase with higher share of individuals claiming the corresponding type of benefits in the region.

4.1 Robustness tests

I check the robustness of Blinder-Oaxaca decomposition by conducting tests to check the effect of top-coding of the data on the results of B-O decomposition. By trimming the data on income from labour and assigning different values to top-coding, I conclude that the results of B-O decomposition are not sensitive to top-coding.

Table 14: The impact of dicrimination on the probability of claiming benefits: natives versus immigrants

	Natives	EU migrants	Non-EU migrants
Discrimination in t-1	–0.004	–0.064	0.053***
	(0.005)	(0.047)	(0.019)
share of individuals claiming benefits by region	0.587**	2.343	1.271
	(0.286)	(2.386)	(1.211)
N	41366	632	3510

Note: The dependent variable is a binary variable, equal to 1 if an idividual claims benefits.
The estimation method is fixed effects.
Time effects and occupational, industry, regional and other controls are included.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

As a robustness exercise for probabilistic models, I estimate probabilities of claiming benefits for separate samples of natives, EU and non-EU immigrants. The results, included in Table 14 (a detailed Table 21 in Appendix F), confirm the estimations in Table 12. The likelihood of natives and EU immigrants to claim benefits is unaffected by income discrimination in corresponding region in time $(t-1)$, whereas it increases for non-EU immigrants. Interestingly, immigrants are not responding to the share of individuals who claim benefits in corresponding region, whereas natives are more likely to claim benefits in regions with higher share of claimants.

Table 15: The impact of dicrimination on the probability of claiming benefits by men: robustness check

	All	EU	Non-EU
Discrimination in t-1	–0.084	–0.102	–0.087
	(0.229)	(0.230)	(0.230)
Immigrants × Discrimination in t-1	0.736		
	(0.518)		
Immigrants	0.021		
	(0.019)		
EU(EEA) × Discrimination in t-1		–1.477	
		(0.937)	
EU(EEA)		–0.006	
		(0.044)	
Non-EU × Discrimination in t-1			0.998*
			(0.557)
Non-EU			0.030
			(0.021)
share of individuals claiming benefits by region	0.246	0.184	0.256
	(0.397)	(0.406)	(0.401)
Occupational controls	X	X	X
Industry controls	X	X	X
Regional controls	X	X	X
Time effects	X	X	X
Time averages	X	X	X
Other controls	X	X	X
N	20059	18634	19787

Note: The dependent variable is a binary variable, equal to 1 if an individual claims benefits.
The variable for discrimination is unexplained difference of regional dummy variables.
The estimation method is correlated random effects.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

I also check the robustness of the results in Table 12 by using unexplained difference of regional dummy variables as a measure of discrimination in period $(t-1)$ instead of total unexplained difference of regressions for each region. The problem with regional dummies is that the the total difference does not get adjusted for sample selection bias. Since LFP is particularly heterogeneous in the case of women, which makes comparison between groups difficult, I limit the sample to men only for this exercise. Table 15 shows the results of this robustness exercise. Here, the coefficients of the interaction term is significant (at 10% level) for non-EU men, which is in line with the results in Table 12. The higher values of these coefficients are due to different scale of the variable of discrimination in this exercise compared with the original variable.

Table 16: The impact of contemporaneous dicrimination on welfare dependency

	All	EU	Non-EU
Discrimination by regions	–0.009	–0.010	–0.009
	(0.008)	(0.008)	(0.008)
Immigrants × Discrimination by regions	0.064**		
	(0.030)		
Immigrants	0.008		
	(0.013)		
EU(EEA) × Discrimination by regions		0.026	
		(0.067)	
EU(EEA)		0.034	
		(0.034)	
Non-EU × Discrimination by regions			0.070**
			(0.033)
Non-EU			0.004
			(0.014)
share of individuals claiming benefits by region	0.603**	0.609**	0.578**
	(0.275)	(0.284)	(0.277)
Occupational controls	X	X	X
Industry controls	X	X	X
Regional controls	X	X	X
Time effects	X	X	X
Time averages	X	X	X
Other controls	X	X	X
N	45508	41998	44876

Note: The dependent variable is a binary variable, equal to 1 if an idividual claims benefits.
The estimation method is correlated random effects.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

As another robustness exercise I use contemporaneous measure of discrimination instead of the lagged (Table 16). The results are robust to this exercise as well, with contemporaneous discrimination increasing likelihood of claiming benefits by immigrants, and particularly, non-EU immigrants.

I also check how sensitive the results are to removing a major region from the regression, as for instance, London. The results are robust to dropping a major region.

5 Conclusions

By using Blinder-Oaxaca decomposition to estimate labour market discrimination against second generation immigrants in the UK, after correcting for sample selection bias, this paper shows that there is significant discrimination against non-EU second generation immigrants, while there seems to be none against EU immigrants. These results are in line with previous studies.

I estimate discrimination in the UK regions, by decomposing income from labour by regions of the UK. I then use the estimates to analyse the impact of discrimination on the probability of welfare dependency of immigrants versus natives. The results show that discrimination does not affect welfare dependency of EU immigrants. However, it increases the probability of non-EU immigrants to claim benefits. Compared with non-EU immigrants' overall probability of claiming benefits of 54% in the situation of no discrimination, the highest observed income discrimination in a single year increases the probability of non-EU immigrants to be on welfare dependency to 65% in the region in that year.

By looking at the probability of claiming different types of benefits, while controlling for individual characteristics, I find that discrimination increases the likelihood of claiming unemployment benefits by immigrants, whereas it decreases the likelihood of claiming unemployment and housing/council tax benefits by natives. The findings also show that immigrants are more likely to claim child benefits and tax credits compared with natives, while they are less likely to claim income support, housing/council tax benefits and sickness/disability/incapacity benefits.

These results yield important potential policy implications, particularly in the areas of welfare dependency and unemployment. The link between income discrimination and dependency on welfare benefits, and above all, unemployment benefits, can be used as a tool for policy makers, also given the opposite effect of discrimination on welfare dependency between natives and immigrants.

Appendix A. Methods of measuring income discrimination. Blinder-Oaxaca decomposition

The main notion behind measuring labour income discrimination is that individuals with similar levels of productivity should be paid similarly. The task of measuring income discrimination, therefore, is reduced to measuring productivity. And here, one would expect that the observable characteristics of individuals will capture productivity. Thus, individuals with the same observable characteristics are expected to be paid similarly.

There are two major approaches to measuring income discrimination. One approach, suggesdted by Neal and Johnson (1996), is to estimate income gap between majority and minority groups by estimating wage equations that include individual characteristics and adding a dummy variable for minority groups:

$$lny_{it} = \boldsymbol{x}_{it}\boldsymbol{\beta} + g_i\alpha + e_{it},$$

where y_{it} is income, $\boldsymbol{x}_{it}$ is a vector of individual characteristics, and g_i is a dummy variable for a minority group i.

The second approach, suggested by Blinder (1973) and Oaxaca (1973) is based on estimating wage equations separately and then comparing the results of the estimates. For two groups, minority group (A) and majority group (B), the following two equation is estimated:

$$Y_k = \boldsymbol{X}'_k\boldsymbol{\beta}_k + \epsilon_k, \ \ where \ \ E(\epsilon_k) = 0, \ \ \boldsymbol{X}_k - a \ set \ of \ explanatory \ variables \ \ and \ \ k \in \{A, B\} \tag{13}$$

The differences in labour market outcomes for the two groups are derived as follows:

$$R = E(Y_A) - E(Y_B), \tag{14}$$

where R is the difference between labour market outcomes of the minority and majority groups, $E(Y_A)$ and $E(Y_B)$ are expected value of an outcome variable of natives and immigrants, accordingly.

Substituting (13) in (14), we get:

$$R = E(X_B)^{'}\beta_B - E(X_A)^{'}\beta_A, \tag{15}$$

After estimating the equations as in (13), substituting the estimates into (15) and rearranging, the authors derive the following expression:

$$\hat{R} = (\bar{X}_B - \bar{X}_A)^{'}\hat{\beta}_B + \bar{X}_B^{'}(\hat{\beta}_B - \hat{\beta}_A) \tag{16}$$

In (16), $(\bar{X}_B - \bar{X}_A)^{'}\hat{\beta}_B$ is the explained difference in labour market outcomes, and $\bar{X}_B^{'}(\hat{\beta}_B - \hat{\beta}_A)$ is the unexplained component.

There are two of issues associated with estimating income inequality through wage equations. Firstly, income equations are prone to sample selection bias as income from labour is observed only for those individuals who are employed. Secondly, productivity of individuals might depend on characteristics that might not necessarily be observed.

Appendix B. Labour force participation: adjusted

Table 17: Labour force participation by groups: adjusted for sample weights

	Male	**Female**
Natives	84.5	77.9
EU	78.2	79.9
India	90.0	75.5
Pakistan	85.4	46.8
Bangladesh	83.3	74.1
Other Africa	86.1	77.6
Latin America	81.0	80.5
Other	90.6	81.0

Notes: The country groups of immigrants are based on father's country of birth.
Labour force participation is computed as share of individuals who are employed/have positive earnings or are unemployed to total individuals in the group, after adjusting the sample for longitudinal weights.

Appendix C. Heckman correction - 1st stage

Table 18: The 1st stage of Heckman correction

	Natives		EU immigrants		Non-EU: India		Non-EU: Pakistan		Non-EU: Bangladesh		Non-EU: Africa		Non-EU: S.America		Non-EU: ther	
	male	female	male	female	male	female	male	female	male	female	male	female	male	female	male	female
no of children aged under 16	–0.001	–0.153***	0.021	–0.067	0.010	–0.148***	0.088*	–0.311***	0.374***	–0.104	–0.175	–0.140**	0.063	–0.015	–0.118	–0.048
	(0.009)	(0.007)	(0.078)	(0.074)	(0.050)	(0.042)	(0.048)	(0.047)	(0.095)	(0.070)	(0.116)	(0.058)	(0.062)	(0.045)	(0.125)	(0.093)
married or lives with partner	0.192***	0.134***	0.199	–0.046	–0.220*	0.050	0.122	–0.017	0.285	0.580***	0.515**	0.244*	0.701***	0.024	0.038	0.010
	(0.017)	(0.014)	(0.147)	(0.122)	(0.124)	(0.092)	(0.131)	(0.130)	(0.213)	(0.174)	(0.216)	(0.135)	(0.135)	(0.091)	(0.195)	(0.191)
father's educational qualifications	0.122***	0.124***	0.235***	0.165***	0.152***	0.107***	–0.011	–0.076	–0.026	0.242***	0.032	0.095**	0.099**	0.016	0.131*	0.297***
	(0.006)	(0.005)	(0.066)	(0.059)	(0.046)	(0.036)	(0.046)	(0.057)	(0.095)	(0.075)	(0.059)	(0.040)	(0.044)	(0.035)	(0.071)	(0.063)
mother's educational qualifications	0.028***	0.020***	–0.027	0.070	0.082*	–0.023	0.124**	0.039	0.113	–0.041	0.107**	0.061	0.044	0.186***	0.057	0.100*
	(0.006)	(0.005)	(0.063)	(0.060)	(0.045)	(0.038)	(0.058)	(0.061)	(0.125)	(0.082)	(0.053)	(0.038)	(0.041)	(0.034)	(0.073)	(0.059)
Regional controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Other controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
N	32301	40710	502	632	726	1034	621	923	352	443	356	578	654	1083	335	352

Note: The estimation method is Probit regression.
The dependent variable is labour force participation.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

Appendix D. Blinder-Oaxaca decomposition by country of origin of immigrants

Table 19: B-O decomposition for natives and immigrants: by country of origin

	EU	India	Pakistan	Bangladesh	Other Africa	Latin America	Other
group_1: natives	7.116***	7.116***	7.116***	7.116***	7.116***	7.116***	7.116***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
group_2: migrants	7.241***	7.173***	6.941***	7.046***	7.138***	7.175***	7.263***
	(0.076)	(0.053)	(0.061)	(0.049)	(0.082)	(0.051)	(0.062)
difference	–0.125	–0.057	0.175***	0.070	–0.022	–0.060	–0.147**
	(0.076)	(0.054)	(0.061)	(0.051)	(0.082)	(0.052)	(0.063)
explained	–0.122***	–0.119***	0.102***	–0.006	–0.259***	–0.227***	–0.232***
	(0.019)	(0.017)	(0.020)	(0.028)	(0.026)	(0.019)	(0.022)
unexplained	–0.003	0.062	0.073	0.077	0.237***	0.167***	0.085
	(0.074)	(0.053)	(0.059)	(0.048)	(0.081)	(0.052)	(0.059)

Note: Dependent variable is log income from labour.
The estimation method is correlated random effects, corrected for sample selection bias.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

Table 19 shows the result of income decomposition by country of origin of immigrants. The breakdown is limited by the sample sizes of immigrants, based on which, the following groups of immigrants are identified: EU, India, Pakistan Bangladesh, Other Africa (excluding North Africa), Latin America (Central and South America), Other (any other country not included in the previous groups). Based on individual country group decomposition, there are two countries with statistically significant unexplained difference - Other Africa and Latin America, with unexplained difference against immigrants of 23.7% and 16.7% of income of natives, respectively.

Appendix E. Probabilities by types of benefits: detailed

Table 20: The impact of dicrimination on the probability of claiming benefits by types of benefits

	I	II	III	IV	V	VI
Discrimination in t-1	–0.004**	0.001	0.001	0.002	–0.006**	–0.003
	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)
Immigrants × Discrimination in t-1	0.027**	0.012	0.008	–0.011	0.017	–0.004
	(0.013)	(0.011)	(0.016)	(0.017)	(0.013)	(0.011)
Immigrants	0.007	–0.026***	0.029**	0.038***	–0.024**	–0.040***
	(0.006)	(0.007)	(0.012)	(0.013)	(0.011)	(0.010)
share of individuals claiming benefits by region	0.108	0.246**	0.950***	0.590**	0.338***	0.174*
	(0.093)	(0.099)	(0.264)	(0.265)	(0.126)	(0.103)
potential experience (years)	0.001	0.004	0.003	–0.006	–0.002	0.003
	(0.003)	(0.003)	(0.007)	(0.007)	(0.006)	(0.004)
squared potential experience (years)	–0.000**	–0.000***	–0.000***	0.000	–0.000	–0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
years of education	–0.006***	–0.008***	–0.008***	–0.001	–0.011***	–0.010***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
years of education squared	0.000*	0.000	–0.000*	–0.001***	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
female	–0.031***	0.007*	0.201***	0.119***	–0.001	–0.004
	(0.004)	(0.004)	(0.007)	(0.006)	(0.006)	(0.005)
urban area	0.001	0.007*	0.004	0.002	0.018***	0.010*
	(0.003)	(0.004)	(0.007)	(0.007)	(0.006)	(0.005)
no of children aged under 16	0.005***	0.007***	0.046***	0.035***	0.015***	0.014***
	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
married or lives with partner	–0.030***	–0.064***	–0.000	–0.055***	–0.128***	–0.032***
	(0.003)	(0.004)	(0.006)	(0.007)	(0.006)	(0.005)
Occupational controls	X	X	X	X	X	X
Industry controls	X	X	X	X	X	X
Regional controls	X	X	X	X	X	X
Time effects	X	X	X	X	X	X
Time averages	X	X	X	X	X	X
Other controls	X	X	X	X	X	X
N	45508	45508	45508	45508	45508	45508

Note: The dependent variable is a binary variable, equal to 1 if an idividual claims benefits.
(I) unemployment benefits, (II) income support, (III) child benefits, (IV) tax credit, (V) housing or council tax, (VI) sickness, disability or incapacity benefits.
The estimation method is correlated random effects.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

Appendix F. Robustness: probabilities regressions by groups

Table 21: The impact of dicrimination on the probability of claiming benefits: natives versus immigrants

	Natives	EU migrants	Non-EU migrants
Discrimination in t-1	–0.004	–0.064	0.053***
	(0.005)	(0.047)	(0.019)
share of individuals claiming benefits by region	0.587**	2.343	1.271
	(0.286)	(2.386)	(1.211)
potential experience (years)	–0.032***	–0.120	–0.015
	(0.009)	(0.074)	(0.024)
squared potential experience (years)	0.000***	0.001	–0.000
	(0.000)	(0.001)	(0.000)
no of children aged under 16 that resp is parent of	0.068***	0.087***	0.049***
	(0.003)	(0.029)	(0.008)
N	41366	632	3510

Note: The dependent variable is a binary variable, equal to 1 if an idividual claims benefits.
The estimation method is fixed effects.
Time effects and occupational, industry, regional and other controls are included.
Significance levels: *:10% **:5% ***:1%
Robust standard errors in parentheses.

References

Algan, Yann, Christian Dustmann, Albrecht Glitz, and Alan Manning, "The economic situation of first and second-generation immigrants in France, Germany and the United Kingdom," *The Economic Journal*, 2010, *120* (542).

Arrow, Kenneth et al., "The theory of discrimination," *Discrimination in labor markets*, 1973, *3* (10), 3–33.

Barrett, Alan and Bertrand Maître, "Immigrant welfare receipt across Europe," *International Journal of Manpower*, 2013, *34* (1), 8–23.

_ **and Yvonne McCarthy**, "Immigrants and welfare programmes: exploring the interactions between immigrant characteristics, immigrant welfare dependence, and welfare policy," *Oxford Review of Economic Policy*, 2008, *24* (3), 542–559.

Battu, Harminder, Paul Seaman, and Yves Zenou, "Job contact networks and the ethnic minorities," *Labour Economics*, 2011, *18* (1), 48–56.

Becker, Gary S, "The theory of discrimination," 1957.

Bell, Brian D, "The performance of immigrants in the United Kingdom: evidence from the GHS," *The Economic Journal*, 1997, *107* (441), 333–344.

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan, "Implicit discrimination," *American Economic Review*, 2005, *95* (2), 94–98.

Bird, Edward, Hilke Kayser, Joachim Frick, and Gert Wagner, "The immigrant welfare effect: take-up or eligibility?," *IZA Discussion Paper 66*, 1999.

Blackaby, David H, Derek G Leslie, Philip D Murphy, and NC O'Leary, "White/ethnic minority earnings and employment differentials in Britain: evidence from the LFS," *Oxford Economic Papers*, 2002, *54* (2), 270–297.

_ **,** _ **,** _ **, and Nigel C O'Leary**, "The ethnic wage gap and employment differentials in the 1990s: evidence for Britain," *Economics Letters*, 1998, *58* (1), 97–103.

_ , **Ken Clark, Derek G Leslie, and Philip D Murphy**, "Black-white male earnings and employment prospects in the 1970s and 1980s evidence for Britain," *Economics Letters*, 1994, *46* (3), 273–279.

Blinder, Alan S, "Wage discrimination: reduced form and structural estimates," *Journal of Human resources*, 1973, pp. 436–455.

Borjas, George J, "Immigration and welfare magnets," *Journal of labor economics*, 1999, *17* (4), 607–637.

_ **and Lynette Hilton**, "Immigration and the welfare state: Immigrant participation in means-tested entitlement programs," *The quarterly journal of economics*, 1996, *111* (2), 575–604.

Brücker, Herbert, Gil S Epstein, Barry McCormick, Gilles Saint-Paul, Alessandra Venturini, and Klaus F Zimmermann, "Managing migration in the European welfare state," *Immigration policy and the welfare system*, 2002, *74.*

Bruckmeier, Kerstin and Jürgen Wiemers, "Differences in welfare take-up between immigrants and natives–a microsimulation study," *International Journal of Manpower*, 2017, *38* (2), 226–241.

Castronova, Edward J, Hilke Kayser, Joachim R Frick, and Gert G Wagner, "Immigrants, Natives and Social Assistance: Comparable Take-Up Under Comparable Circumstances," *International Migration Review*, 2001, *35* (3), 726–748.

Chiswick, Barry R, "The earnings of white and coloured male immigrants in Britain," *Economica*, 1980, *47* (185), 81–87.

Clark, Ken and Stephen Drinkwater, "The labour-market performance of recent migrants," *Oxford Review of Economic Policy*, 2008, *24* (3), 495–516.

Drinkwater, Stephen and Catherine Robinson, "Welfare participation by immigrants in the UK," *International Journal of Manpower*, 2013, *34* (2), 100–112.

Dustmann, Christian and Nikolaos Theodoropoulos, "Ethnic minority immigrants and their children in Britain," *Oxford Economic Papers*, April 2010, *62* (2), 209–233.

Frijters, Paul, Michael A Shields, and Stephen Wheatley Price, "Job search methods and their success: a comparison of immigrants and natives in the UK," *The Economic Journal*, 2005, *115* (507), F359–F376.

Giulietti, Corrado and Jackline Wahba, "26 Welfare migration," *International handbook on the economics of migration*, 2013, p. 489.

_ , Christian Schluter, and Jackline Wahba, "With a Lot of Help from my Friends: Social Networks and Immigrants in the UK," *Population, Space and Place*, 2013, *19* (6), 657–670.

_ , Martin Guzi, Martin Kahanec, and Klaus F Zimmermann, "Unemployment benefits and immigration: evidence from the EU," *International Journal of Manpower*, 2013, *34* (1), 24–38.

_ , Mirco Tonin, and Michael Vlassopoulos, "Racial Discrimination in Local Public Services: A Field Experiment in the United States," *Journal of the European Economic Association*, 2017.

Glover, Dylan, Amanda Pallais, and William Pariente, "Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores," *The Quarterly Journal of Economics*, 2017, p. qjx006.

Hansen, Jorgen and Magnus Lofstrom, "Immigrant Assimilation and Welfare Participation Do Immigrants Assimilate Into or Out of Welfare?," *Journal of Human Resources*, 2003, *38* (1), 74–98.

Heckman, James J, "Sample Selection Bias as a Specification Error," *Econometrica*, 1979, *47* (1), 153–161.

Heitmueller, Axel, "A note on decompositions in fixed effects models in the presence of time-invariant characteristics," 2005.

Hsiao, Cheng, *Analysis of panel data* number 54, Cambridge university press, 2014.

Jann, Ben et al., "A Stata implementation of the Blinder-Oaxaca decomposition," *Stata journal*, 2008, *8* (4), 453–479.

Jilke, Sebastian, Wouter Van Dooren, and Sabine Rys, "Discrimination and Administrative Burden in Public Service Markets: Does a Public-Private Difference Exist?," *Journal of Public Administration Research and Theory, Forthcoming*, 2018.

McNabb, Robert and George Psacharopoulos, "Racial earnings differentials in the UK," *Oxford Economic Papers*, 1981, *33* (3), 413–425.

Mundlak, Yair, "On the pooling of time series and cross section data," *Econometrica: journal of the Econometric Society*, 1978, pp. 69–85.

Munshi, Kaivan, "Networks in the modern economy: Mexican migrants in the US labor market," *The Quarterly Journal of Economics*, 2003, *118* (2), 549–599.

Neal, Derek A and William R Johnson, "The role of premarket factors in black-white wage differences," *Journal of political Economy*, 1996, *104* (5), 869–895.

Neumark, David, "Experimental research on labor market discrimination," Technical Report, National Bureau of Economic Research 2016.

Oaxaca, Ronald, "Male-female wage differentials in urban labor markets," *International economic review*, 1973, pp. 693–709.

Phelps, Edmund S, "The statistical theory of racism and sexism," *The american economic review*, 1972, *62* (4), 659–661.

Razin, Assaf and Jackline Wahba, "Welfare magnet hypothesis, fiscal burden, and immigration skill selectivity," *The Scandinavian Journal of Economics*, 2015, *117* (2), 369–402.

Riphahn, Regina T, "Immigrant participation in the German welfare program," *FinanzArchiv/Public Finance Analysis*, 1998, pp. 163–185.

Wooldridge, Jeffrey M, *Econometric analysis of cross section and panel data*, MIT press, 2010.

_ , *Introductory econometrics: A modern approach*, Nelson Education, 2015.