**Mind the Gap:
A Detailed Picture of the Immigrant-Native
Earnings Gap in the UK Using Longitudinal Data
Between 1978 and 2006**

Sara Lemos

October 2011

Mind the Gap:
A Detailed Picture of the Immigrant-Native Earnings Gap in the UK Using Longitudinal Data Between 1978 and 2006

Sara Lemos
University of Leicester
and IZA

ABSTRACT

Mind the Gap:
A Detailed Picture of the Immigrant-Native Earnings Gap in the UK Using Longitudinal Data Between 1978 and 2006[*]

Using the underexplored, sizeable and long Lifetime Labour Market Database (LLMDB) we estimated the immigrant-native earnings gap across the entire earnings distribution, across continents of nationality and across cohorts of arrival in the UK between 1978 and 2006. We exploited the longitudinal nature of our data to separate the effect of observed and unobserved individual characteristics on earnings. This helped us to prevent selectivity biases such as cohort bias and survivor bias, which have been long standing unresolved identification issues in the literature. In keeping with the limited existing UK literature, we found a clear and wide dividing line between whites and non-whites in simple comparable models. However, in our more complete models we found a much narrower and subtler dividing line. This confirms the importance of accounting for unobservable individual characteristics, which is an important contribution of this paper. It also suggests that the labour market primarily rewards individual characteristics other than immigration status. We also found that the lowest paid immigrants, whom are disproportionately non-white, suffer an earnings penalty in the labour market, whereas higher paid immigrants, whom are disproportionately white, do not. Finally, we found less favourable earning gaps for cohorts that witnessed proportionately larger non-white and lower paid white immigration.

Corresponding author:

Sara Lemos
University of Leicester
Economics Department
University Road
Leicester LE1 7RH
United Kingdom
E-mail: sl129@leicester.ac.uk

1. Introduction

The immigrant-native earnings gap matters. It matters because is a measure of the absorption of economic immigrants into the labour market and into society more generally. For example, the immigration debate is more heated in countries where economic immigrants are perceived as a threat to natives' job opportunities. This is the case, for instance, if immigrants are unskilled and work for comparatively lower wages. In contrast, in countries where economic immigrants are perceived as filling up vacancies where there is labour shortage, they are seen as contributing to the economy. In this case, their skills might be favourably rewarded and they might work for comparatively higher wages.

The direction of such earnings gap, therefore, is a measure of how immigrants fare in the labour market – and this informs policymaking. Therefore, a careful and detailed analysis of the immigrant-native earnings gap right across the board is the first step to understand the effects of immigration on the labour market and on the economy more generally. Yet, there is very limited evidence on such a crucial labour market issue in the UK. In his pioneering paper, Chiswick (1980) alerted for such scarcity in the UK literature. Nevertheless, 30 years on and less than a handful of papers have followed (Bell 1997; Dustmann and Fabbri 2005; Dickens and McKnight 2008). Furthermore, this rather small literature is sometimes hindered by the use of unsophisticated models of average gaps and by limitations in the data.[2]

The main contribution of this paper is to help to fill this blank in the literature and to inform policymaking in the face of continuing public debate on immigration policy in the UK. We estimate the immigrant-native earnings gap across the entire earnings distribution. This allows us to uncover potential earnings gaps at particular points along the distribution that might have been masked by the average gap. We also estimate the immigrant-native earnings gap by gender, by continent of nationality and by cohort of arrival. This allows us to investigate whether the earnings gap is affected by immigrants' origins or by economic conditions in the UK at the time of their arrival.

---

[2] Although the corresponding literature is quite large for the US (Chiswick 1978; Borjas 1985, 1995 and 1999; Butcher and DiNardo 2002; Cortes 2004; Chiswick et al. 2005; etc.), it is limited for the UK. Using data from the 1972 General Household Survey (GHS) to estimate a standard human capital earnings model, Chiswick (1980) found no earnings gap for white but a -25% gap for non-white male immigrants. In an attempt to model cohort and assimilation effects separately, Bell (1997) used 1973-1992 GHS data and broadly confirmed these earlier findings. Dustmann and Fabbri (2005) estimated a simple model using data from the 1979-2004 Labour Force Survey (LFS) and expanded the analysis to include females. They found that the wage gap for non-white immigrants was as large as -40%, though this varied with immigrants' region of origin. Dickens and McKnight (2008) estimated an unrefined model using data from the 1978-2003 Lifetime Labour Market Database (LLMDB) and found surprisingly large and negative wage gaps for all immigrants. In particular, they found a large wage penalty for white (European) immigrants, which is not in line with the UK or international literature.

Interestingly, there is something of a correspondence between immigrants' continent of nationality, their cohort of arrival and the section of the earnings distribution where they end up at. For example, whereas many of the high skilled North Americans that arrived during the 1990s and 2000s ended up at the top of the earnings distribution, many of the low skilled Eastern Europeans that arrived in the 2000s ended up at the bottom of the distribution. Therefore, our detailed and comprehensive analysis provides an invaluable insight into the same phenomenon from three different angles.

Another contribution of this paper is to exploit a sizeable and long longitudinal dataset that has been yet largely unused in the UK immigration literature. The Lifetime Labour Market Database (LLMDB) combines anonymised administrative tax records and social security records into a dataset that tracks a random sample of 647,000 individuals between 1978 and 2006. Given that data limitations is one of the main reasons for scarce evidence on immigration for the UK, exploiting such a unique dataset is a timely contribution. For example, while the LLMDB is large enough to allow disaggregation by continent of nationality and by cohort of arrival, as well as by small geographical areas, the more commonly used GHS (General Household Survey) and LFS (Labour Force Survey) do not permit such fine levels of disaggregation.

Another important contribution of this paper is that the longitudinal nature of our data helps circumvent identification issues that have long posed difficulties in the literature. One such identification issue is controlling for individual specific time invariant (fixed) effects. If immigrants are more able, more motivated or work harder than natives, then these unobserved individual characteristics could bias the earnings gap estimates. Other such selection arguments include changes over time in unmeasured dimensions of immigrants' skills and return migration of immigrants that are less able (along with other types of data attrition). Concern with biases implied by these various types of selectivity in the data, such as cohort bias and survivor bias, have occupied much of the literature for the last 30 years (Chiswick 1978 and 1980; Carliner 1980; Borjas 1985, 1994 and 1999; Duleep and Regets 1997; Chiswick et al. 2005). This literature has long recognized that the ideal way to address such selectivity biases is to use sufficiently large and long longitudinal data – which has, however, been very scarce. Therefore, using the LLMDB to control for individual fixed effects when estimating the immigrant-natives earnings gap is a timely contribution, as it enables us to separate the effect of individual unobserved characteristics from the effect of other observed variables (such as gender, age, continent of nationality, cohort of arrival, etc.) on earnings.

Indeed, the results are striking. In keeping with the existing UK literature (Chiswick 1980; Bell 1997; Dustmann and Fabbri 2005), we found a clear and wide dividing line between whites and non-whites when we estimated the earnings gap using simple models. In our base model, non-white immigrants earned between 16.5% and 21.8% less than natives, whereas white immigrants earned up to 26.5% more than natives. However, as we estimated more complete models, our earnings gap estimates narrowed and became smaller than those available in the existing UK literature. In our preferred model, where we control for observable and unobservable individual characteristics, the gap for non-whites is 0%, whereas it is between 2.9% and 10.7% for whites. This suggests a much narrower and subtler dividing line, confirming the importance of accounting for observable as well as unobservable individual characteristics, which is an important contribution of this paper. It also suggests that the labour market primarily rewards individual characteristics other than immigration status.

When we estimated the gap across the earnings distribution, we found that it was between -7.4% and 0.7% below the median, where non-white immigrants are overrepresented, and between 2.4% and 8.7% above the median, where white immigrants are overrepresented. This suggests that most immigrants do not seem to suffer an earnings penalty in the labour market. It confirms that, except for the lowest and highest paid immigrants, the labour market primarily rewards individual characteristics other than the immigration status.

Finally, we found less favourable gaps for cohorts that witnessed proportionately larger non-white and lower paid white immigration. For example, the earnings gap is between -1.7% and 0% for the 1985-1994 cohorts, when there was proportionately large non-white immigration (although this period also coincides with greater, lower paid, white immigration following the enlargement of the EU in the mid 1980s when Greece, Portugal and Spain joined in). This contrasts with an earnings gap between 3.7% to 8.6% following proportionately large white immigration in the 1995-2004 cohorts, when EU immigration increased – accelerating dramatically after a further enlargement in the mid 2000s when Eastern European countries joined in – along with increased immigration from North America, Australasia and Oceania (though this period also witnessed greater non-white immigration from Africa, Asia and the Middle East).

We thoroughly discuss the above issues in the remainder of this paper. In Section 2 we depict our data. In Section 3 we specify our empirical model and carefully discuss several identification issues. In Section 4 we summarise the results. In Section 5 we discuss our

results in light of the existing literature before we conclude in Section 6.

2. Data

We use data from the Lifetime Labour Market Database (LLMDB). The LLMDB is derived from several administrative datasets linked together by a unique individual identifier, the national insurance number (NINo). Individuals must apply for a NINo in order to pay tax (income tax, national insurance contributions, self-employment contributions, etc.), receive retirement pension, or claim social security benefits (e.g. unemployment benefit, incapacity benefit, sickness benefit, maternity benefit, child benefit, housing benefit, etc.). Whereas natives are automatically given a NINo just before they turn 16 years of age, which is derived from their child benefit number, immigrants typically apply for a NINo when they start interacting with the system, either by paying taxes or by claiming benefits. ("Natives" and "Immigrants" here and throughout the paper are respectively UK born and overseas born nationals.) Because individuals need to produce their NINo in every interaction with the system, the LLMDB effectively tracks individuals throughout their lifetime – and more crucially, throughout their working lives.

The main advantage of the LLMDB, therefore, is that it is a rich, long and large longitudinal dataset. It has high levels of accuracy and relatively low levels of attrition (individuals only drop out of the sample if they neither pay taxes nor claim any benefit nor receive retirement pension for more than 12 months; they re-enter the sample when they again interact with the system). Our LLMDB sample comprises 647,068 individuals (a 1% random sample of NINo records) followed between the tax-years 1978 and 2006 (which run from April to March) resulting in 11,061,433 observations (a fresh cohort of individuals enters the data every year and is followed from then on). We restricted our sample to those aged 25 to 64, as is common in the earnings gap literature (though the results were robust when including those aged 16 to 64). We also restricted our sample to those earning between £100 and £1000000 in any one tax-year (this excludes the self-employed, for whom we do not observe earnings). We also restricted our sample to immigrants arriving from 1945 onwards, because the number of immigrants arriving previously was relatively very low and because restricting the sample facilitates cohort modelling, which we discuss in Section 3. Finally, we restricted our sample to those observed at least twice, due to the way we control for individual fixed effects, which we discuss in detail in Section 3 (although our results were robust to including individuals observed only once). Our final working sample therefore consists of 433,069

individuals, 45,309 of whom are immigrants, as shown in Table 1. The total number of observations is 5,373,086 and the average number of observations for natives (immigrants) is 15.76 (18.71). Figure 1 shows the distribution of natives and immigrants across tax-years.

The LLMDB contains well over 400 variables, including date of birth, date of death, age, gender, address, nationality, country which immigrants arrived from, immigrants' entry date,[3] immigrants' age at entry, number of jobs in the year, annual earnings per job, type of employment (employee or self-employed), number of weeks employed (unemployed) in the year, dates of spells of unemployment, dates of spells of receipt of benefits, benefit type, pension contributions, pension entitlements, etc. As is common with administrative data, the LLMDB does not contain information on education. In this paper we circumvent this issue to some extent by restricting our sample to those in work aged 25 to 64, who, we assume, have completed their education, and by modelling individual fixed effects, as we discuss in Section 3. In addition, the LLMDB does not contain information on the immigrants' entry route (work permit, student visa, family reunification, etc.) or on their departure date. Since in this paper we focus on those already in work, their entry route is not crucial, although some limited information on such a route can be gauged from their first few interactions with the system. Similarly, since we are focusing on those in work, differentiating whether an immigrant left the labour force or left the country is not crucial here, although it might be very relevant elsewhere.

Table 1 shows that natives are older than immigrants, are more evenly spread across the country, are more likely to be employed and slightly less likely to be unemployed, and earn more on average. Figure 2 shows the immigrant-native average real earnings gap across tax-years, confirming that on average immigrants earned less than natives during most of the sample period, although the variation is large. Figure 2 also shows that the earnings gap is greater and more negative for male immigrants.

Table 1 also shows that immigrants arrive young and many remain for several years. Those at the very bottom of the earnings distribution earn less whereas those at the very top earn more than natives. This is confirmed in Figure 3, which shows the immigrant-native real earnings gap across tax-years for several percentiles of the earnings distribution. Whilst immigrants at the bottom of the distribution can earn less than a half of what natives at the bottom of the distribution earn, those at the top can earn up to a quarter more than natives at

---

[3] The entry date is only recorded when the immigrant applies for a NINo, which depends on her individual circumstances and might not happen immediately upon arrival. Existing internal checks in the data and our own analysis suggest that the associated measurement error is fairly limited, especially after 1997.

the top earn. The earnings gap for the lower paid becomes more negative over time, especially after 2003, which coincides with the inflow of low paid Eastern Europeans. In contrast, the earnings gap for the higher paid becomes more positive over time, especially around 2000, following the inflow of high paid North Americans during the 1990s and 2000s. This is also confirmed in Figure 4, which shows that the immigrants' earnings distribution becomes relatively more dispersed over time.

Table 1 also shows that immigrants predominantly come from the European Union (EU), Asia and the Middle East, and Africa, and that there are marginally proportionately more white than non-white immigrants. Figure 5 confirms that these overall patterns persist across tax-years, although the proportion of non-whites increases over time. Interestingly, Table 1 shows that immigration intensified after the mid 1970s.[4] Figure 6 plots the inflow of immigrants by year of arrival and continent of nationality.[5] Most cohorts display a mix of white and non-white immigrants. For example, during the 1950s and 1960s, white EU immigrants (mainly Irish) and non-white immigrants from former colonies (India, Pakistan, Bangladeshi, South Africa, Nigeria, etc.) came to the UK to help with the post-war reconstruction effort. There were also other minority groups, such as Jews fleeing to the UK and refugees from Kenya, Uganda, etc. During the 1970s this overall trend continued, with increased EU immigration, especially after 1973, when the UK joined the EU. In this period there was also an increased inflow of immigrants from India (mainly because Gujarati Indians were expelled from Uganda) and from Vietnam and South East Asia. During the 1980s and 1990s white immigration increased proportionately more, following the enlargement of the EU when Greece (in 1981) and Portugal and Spain (in 1986) joined in, and following increases in the inflow of immigrants from North America, Australasia and Oceania. In that period non-white immigration, mainly from Africa, Asia and the Middle East, also increased. In the late 1990s and 2000s white immigration from the EU, North America, Australasia and Oceania continued to increase strongly, more dramatically after 2004, when ten Eastern

---

[4] Although our sample period is between 1978 and 2006, the last cohort is four years short, by construction, since all those arriving in 2006 were dropped from the sample because they were only observed once.

[5] Although the LLMDB is more reliable after 1975, and more reliable still after 1997, Figure 6 plots the inflow from 1945 onwards. The overall number of immigrants is fairly reliable, but their continent of nationality was not always imputed. The rather large spikes in the series of "unknown" continent of nationality in Figure 6 reflects the introduction of new computing systems when inputting fields such as country of nationality of immigrants was not a priority. For example, the spike in 1975 reflects immigrants that arrived in the immediately preceding years (that is also true for the spikes in 1948 and 1997). Luckily, this affects only three of our 13 cohorts of interest. Thus, despite this caveat in the data, it is possible to identify some overall patterns of immigration.

European countries (A10) joined in a further enlarged EU.[6] In this period non-white immigration from Africa, Asia and the Middle East increased sharply, and immigration from Central and South America also became more pronounced.

Finally, Figure 7 shows the immigrant-native real earnings gap across tax-years by continent of nationality and is another way to see the earnings gap becoming more positive for North Americans during the late 1990s and early 2000s and more negative for Eastern Europeans after 2000 (although it bounces back after 2004 when the minimum wage was increased, as shown in Figure 3 – also see the spike at the minimum wage level on the 2005 earnings distribution in Figure 4).

The above figures illustrate that disaggregation by continent of nationality and by cohort of arrival is another main advantage of the LLMDB. The LLMDB also permits disaggregation by small geographical levels.[7] This is in contrast with the more widely used Labour Force Survey (LFS) – which is a rotating panel survey that interviews around 60,000 households with about 140,000 respondents every quarter available since 1992 (more comprehensive wages and hours worked data is available after 1997) – where immigration analysis across years and continents or below the regional level is not feasible due to sample size limitations. Table 1 shows LFS variables. Dustmann and Fabbri (2005), using LFS data for a roughly comparable sample period, report descriptive statistics that are in line with our own descriptives here.

Table 1 shows that both the LLMDB and the LFS exhibit broadly similar patterns (we use the 1997-2007 sample period for this comparison). Women are slightly underrepresented in the LLMDB, perhaps reflecting their labour market participation decisions. The age distribution in both datasets is remarkably similar for natives, though a larger proportion of immigrants is younger in the LLMDB (note that we tabulate observations, not individuals). This is because a larger share of observations in the LLMDB is for immigrants (8.2% versus 7.7%) – and because the LLMDB better captures low paid immigrant workers, who tend to be younger. For example, the LLMDB might capture working foreign students and working illegal immigrants who might not have been captured in the LFS.[8] As a result, average

---

[6] Although the EU was successively enlarged at various points during our sample period, for consistency our definition of countries belonging to the EU throughout the whole period is that of 2006. We separately define the A10 countries, which are: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovenia and Slovakia, Malta and Cyprus.

[7] The ONS-defined geographical areas we use are: 409 Local Authority Districts, 49 counties and 12 Government Regions (ONS 2003) (see Table 1).

[8] A small number of workers who earn too little or work too few hours to incur a national insurance contribution liability, for example those working part-time for very small employers, are not included in the LLMDB.

earnings are lower in the LLMDB.

Importantly, the LLMDB records annual earnings (within the tax-year) – i.e. total annual earnings including any part-time and/or unemployment spells – whereas the LFS records weekly earnings in a given week, which are extrapolated for the year ignoring any part-time and/or unemployment spells (which are unknown). As a result, the LFS figures in Table 1 overestimate earnings, which are higher for every percentile of the distribution. The difference is larger at the bottom and smaller at the top of the distribution, confirming that the LLMDB captures more low paid workers (who either earn lower wages or work fewer hours). In particular, the LFS figures overestimate earnings for immigrants, who are more likely to be low paid, and thus the gap between natives and immigrants is less (more) persistent in the LFS (LLMDB), with immigrants earning more than natives up to the 20th (60th) percentile of the distribution. However, although earnings are consistently lower in the LLMDB, the average earnings trend over time is similar. (Detailed comparisons across years, available on request, depict a very similar pattern of average earnings, percentile earnings, employment and unemployment rates across both datasets.) Dickens and McKnight (2008) carried out a similar analysis comparing the LLMDB and the ASHE and also concluded that annual earnings are lower in the LLMDB but that the trend of average earnings, and of selected percentiles of the earnings distribution, across both datasets is similar over time.

Finally, Table 1 shows that the regional distribution both in the LLMDB and in the LFS is remarkably similar for natives – and for immigrants, if those with unknown or abroad locations are excluded from the analysis. The distribution of immigrants' country of origin is also very similar in both datasets – again if those from unknown origin are excluded from the analysis. The distribution of immigrants' cohort of arrival is also very similar in both datasets (for a discussion of the spike at the 1975-1979 cohort in the LLMDB, see Footnote 5).

3. Model Specification

Figures 2, 3 and 7 suggest that there is indeed an immigrant-native earnings gap in the UK between 1978 and 2006, which is quite sizeable for some groups of immigrants. However, such raw unconditional earnings gap estimates need to be proved robust when accounting for the effect of other individual characteristics (such as gender, age, continent of nationality, cohort of arrival, ability, motivation, etc.) on earnings. We now account for this by estimating

---

Medium and larger employers are captured and their non-liable employees are included in the LLMDB. The LFS includes earnings for the self-employed, which are not recorded in the LLMDB (see above).

the conditional immigrant-native earnings gap (conditional on such characteristics) using a standard human capital earnings model (see for example, Chiswick 1980, Dustmann and Fabbri 2005). In the human capital model, individuals' earnings are a function of characteristics that influence individuals' productivity:

$$E_{iat} = \alpha + \beta I_i + \lambda X_{iat} + f_i + f_a + f_t + \varepsilon_{iat} \tag{1}$$

where E_{iat} is log real earnings of individual $i = 1,...,433069$ in area $a = 1,...,49$ and time $t = 1978,...,2006$; I_i is an indicator variable that equals 1 if the individual is an immigrant; X_{iat} is a vector of observable individual characteristics; f_i is individual fixed effects; f_a is area fixed effects; f_t is time fixed effects; and ε_{iat} is the error term. The interpretation of our coefficient of interest is that immigrants on average earn $\beta\%$ more than natives.[9]

We model area fixed effects using county dummies (see Section 2). This way we remove any permanent differences across counties and make them equally attractive. In other words, we control for specific factors in a county (such as more schools, more housing, lower prices, etc.) that may make it more attractive to immigrants or natives or both. This enables us to separately account for the effect of county specific time invariant factors on earnings. We model time fixed effects using tax-year dummies (see Section 2). This way we control for the effect of tax-year specific macroeconomic effects (such as seasonal shocks, national and international macroeconomic shocks, etc.) on earnings. This enables us to separately account for the effect of time specific factors on earnings. Controlling for area and time fixed effects in this flexible manner (across counties and tax-years) is an improvement to the existing UK earnings gap literature.[10]

We also control for observable individual characteristics such as sex, age, age squared number of employed weeks in the year and number of jobs in the year (see Table 1). This enables us to separately account for the effect of such characteristics on earnings. For example, this way we account for earnings differentials due to workers being younger or less experienced in addition to being immigrants. Although we do not observe experience, we

---

[9] More precisely, immigrants on average earn $b = 100[\exp(\beta) - 1]$ more than natives. As most of our β estimates are close to zero – in particular our preferred ones deriving from our most complete specifications (see Section 4) – β is a good approximation of b, so for simplicity we report β throughout the paper. Strictly speaking, immigrants on average earn β more than natives in logarithmic units (Halvorsen and Palmquist 1980).

[10] The available models in this literature do not control for area fixed effects, except Dustman and Fabbri (2005), where region (not county) fixed effects are included. Here we model area fixed effects using 49 counties instead of 12 regions, which is a more flexible approach.

control for age, which, albeit imperfectly, captures overall experience to a certain extent.[11,12]

We control for unobservable individual characteristics using the lagged number of employed weeks in the year and lagged real earnings (instead of using individual dummies). These two lagged variables together embed all the relevant information on unobservable individual characteristics that affects earnings, such as motivation, race, immigrant's age at arrival, education (recall, as discussed in Section 2, that our sample is restricted to those aged 25 to 64 who have already completed their education), etc.[13] This is because these two lagged variables capture individual specific time invariant characteristics that have the same impact on earnings year after year (Nakamura and Nakamura 1985; Chiswick et al. 2005). This way we also circumvent the problem of perfect collinearity between individual dummies necessary to model individual fixed effects f_i and our variable of interest I_i. Finally, controlling for the lagged number of employed weeks in the year accounts for lower earnings for individuals with historically long spells of unemployment. Including these two lagged variables also allows for the effect of dynamics in the model and alleviates problems arising from serial correlation in the residuals.

Controlling for individual fixed effects enables us to separately account for the effect of

---

[11] Our results were robust to including extra controls such as receipt of maternity benefit, child benefit, incapacity benefit, sickness benefit, etc. Most of these, however, are not consistently available for the entire sample period, since they were introduced or changed at various points in time.

[12] We also experimented with controlling for "age at entry" – which captures the human capital endowment at arrival and is particularly important for identifying immigrants who arrived as children and thus have the labour market characteristics of natives – but that did not alter our main results (also see Borjas 1994 and 1999). In our data sample, most immigrants arrived as adults; around 10% entered as children, and the vast majority of these arrived before 1969 from ex-colonies and from Europe. Lemos (2011) estimated an "immigrant economic assimilation" model using the same data sample and controlling for "years since immigration", in an attempt to disentangle cohort and assimilation effects, and found results in line with ours. That model is conceptually different from ours because here we simply want to describe the existing immigrant-native earnings gap when comparing natives and immigrants as like for like. That is, we compare immigrants and natives with the same gender, age and employment level to see how differently they perform in the labour market, whatever their assimilation histories. In contrast, Lemos (2011) attempts to separately estimate assimilation effects. In addition, she allowed the effect of observable individual characteristics such as sex, age, number of employed weeks in the year and number of jobs in the year to differ between natives and immigrants. However, she restricted the effect of area and time fixed effects to be the same for immigrants and natives, as we also do here. This is a restrictive assumption, since macroeconomic conditions most likely affected both groups differently. This is particularly so if immigrants are concentrated in the low or high end of the earnings distribution and if the earnings distribution changed over time, as suggested by Table 1 and Figure 4 (see Section 2) (LaLonde and Topel 1992; Butcher and DiNardo 2002). However, this is a common restriction in the literature (Borjas 1999).

[13] Although earnings models commonly control for education, there is an unresolved debate in the immigration literature about what the interpretation of other coefficients in the model should be when controlling for education (Borjas 1999). Excluding education implies that we are comparing the earnings of immigrants and natives, and not the earnings of immigrants and natives with the same education level. This is important because, as we discuss in Section 4.1, the extent and quality of education varies across countries. Therefore, immigrants and natives with the same education may have different skills and compete for different jobs. Furthermore, immigrants across the education spectrum often suffer skill downgrading due to language or other labour market barriers (see Section 4.1).

individual specific time invariant characteristics on earnings. This enables us to account for any earnings differential due, for example, to workers who are more motivated. Controlling for individual fixed effects using a sufficiently large and long longitudinal dataset such as the LLMDB is an important improvement to the existing UK immigrant-native earnings gap literature. This is because controlling for individual fixed effects ensures that sources of selectivity bias, such as cohort bias or survivor bias, are accounted for. For example, by controlling for individual fixed effects we account for immigrants being more able, more motivated or hardworking than natives as well as for changes over time in unmeasured dimensions of immigrants' skills and return migration (and for other types of data attrition).

We use generalized least square estimation and correct for intragroup serial correlation, as standard errors are assumed to be independent across groups of individuals but not within groups (i.e. for a particular individual over time).

4. Results

The first column in bold of Table 2 shows a significant -0.132 β estimate in our base model, whose predicted residuals are plotted in the top left panel of Figure 8. This estimate suggests that immigrants on average earn 13.2% less than natives. This base model accounts for area specific time invariant factors that may simultaneously affect both the earnings and the area choice of individuals, such as the fact that more multicultural or higher earnings areas (e.g. London) attract both immigrants and natives. In addition, this model also accounts for macroeconomic tax-year specific effects that may simultaneously affect both earnings and immigration decisions, such as interest rate changes or international shocks. As expected, this simple base model has a low R^2 (0.02).

Further controlling for other individual characteristics yields a significant -0.023 estimate in the second column in bold of Table 2 (also see top right panel of Figure 8) and a substantially higher R^2 (0.49). This estimate shows that the initial immigrant-natives earnings gap has considerably narrowed, suggesting that the earlier more negative gap was driven by omitted variables. This indicates that characteristics other than the immigration status (such as sex, age, number of employed weeks and number of jobs in the year, which, in the main are significant and of the expected sign here as well as in the remaining models in the paper) are important factors in explaining the immigrant-natives earnings gap in the UK, as expected.

Finally, when controlling for unobservable individual characteristics – through lagged number of employed weeks and lagged log real wages (see Sections 2 and 3) – the immigrant-

natives earnings gap estimate is a significant 0.023, as shown in the last column in bold of Table 2. This estimate suggests that immigrants earn 2.3% more than natives on average. This is our preferred model, as it separates out the effect of unobservable individual characteristics, which were being unsatisfactorily proxied by observable individual characteristics in the models above (particularly, perhaps, by the number of employed weeks and the number of jobs in the year). Its residuals are well behaved, as shown in the bottom panel of Figure 8, and the associated R^2 is substantially higher (0.72).

Our results suggest a small average immigrant-native earnings gap in the UK between 1978 and 2006. Our preferred estimate suggests that immigrants earn 2.3% more than natives on average. Put differently, immigrants do not seem to suffer an earnings penalty in the labour market as a result of their immigrant status. This suggests that the labour market primarily rewards observable and unobservable individual characteristics other than the immigration status. We explore these findings further using three different alternative and complementary estimation approaches in Sections 4.1 to 4.3.

4.1   Across the Earnings Distribution

Our models in Section 4 only estimate the average immigrant-natives earnings gap. The implicit approach was to compare the earnings of all immigrants with the earnings of all natives, which may not be realistic. This is because unskilled immigrants do not compete with skilled natives, for instance. So, an inflow of low paid unskilled immigrants will not directly affect earnings in the high paid highly skilled end of the labour market.

We thus re-estimate Equation (1) for several percentiles across the earnings distribution using quantile regression estimation. This is to uncover potentially larger or smaller earnings gaps at particular points along the distribution that might have been masked by the average gap.[14] This is a particularly appealing approach where immigrants concentrate at the bottom and top of the earnings distribution, as it is the case for the UK over the sample period, where immigration has been of predominantly unskilled or highly skilled labour (see Section 2). Estimating the immigrant-native earnings gap in such a flexible manner across the earnings

---

[14] Several skill definitions have been used in the literature, e.g. occupation, education, education-experience, percentiles of the earnings or wage distribution, etc. (see for example Card 2001; Borjas 2003). Using percentiles of the earnings distribution is a more accurate measure than perhaps education or experience. Firstly, the extent and quality of education varies across countries. Therefore, immigrants and natives in the same education cell may have different skills and compete for different jobs. Secondly, earnings measure the effective reward that immigrants obtain, after usual skill downgrading due to language or other labour market barriers. Thirdly, there is evidence that natives and immigrants are imperfect substitutes within education groups in the UK (Manacorda et al. 2006).

distribution is an improvement to the existing UK earnings gap literature, where only estimates of the average gap are available.

Figure 9 shows significant immigrant-native earnings gap estimates for our base model and for our preferred model (also see Table A1 in the appendix). As before, the gap substantially narrows in our most complete and preferred specification. It also becomes non-negative, except below the 30th percentile. Interestingly, the gap increases monotonically across the distribution: it is respectively -0.030, 0.000, 0.014, 0.037 and 0.087 for the 10th, 30th, 50th, 70th and 90th percentiles. This suggests that the more skilled immigrants are, the more favourable the earnings gap they experience; and the less skilled immigrants are, the larger the earnings penalty they suffer. For example, among the 10% best paid workers, immigrants earn 8.7% more than natives; whereas among the 10% worst paid workers, immigrants earn 3% less than natives. This compares with our earlier average gap estimate of 2.3% (compare the last column in bold of Tables 2 and A1).

Thus, our conclusion is that the immigrant-native earnings gap in the UK between 1978 and 2006 varies substantially across the distribution, increasing monotonically, and this variability is masked when solely the average gap is considered. While the gap was zero at the 30th percentile, it was negative at the bottom of the distribution (-7.4% at the 5th percentile), and positive above the median (between 2.4% and 8.7%). Given that immigration to the UK has been of predominantly unskilled or highly skilled labour (see Section 2), we indeed expected larger gaps at the bottom and top of the distribution.

In sum, on the one hand, the lowest paid immigrants suffer an earnings penalty in relation to the lowest paid natives with comparable individual characteristics. On the other hand, other immigrants do not seem to suffer an earnings penalty in the labour market as a result of their immigrant status – the gap is fairly small in the lower middle of the distribution and it is in favour of higher paid immigrants. In line with our conclusion from before, this suggests that, except for the lowest and highest paid immigrants, the labour market primarily rewards observable and unobservable individual characteristics other than the immigration status.

According to standard human capital theory, non-negligible gap estimates could be due to unaccounted productivity differentials. However, our model is quiet comprehensive. For example, our model seems to have captured most such productivity differentials in the lower middle of the distribution, where the gap was fairly small. In the absence of such supply side productivity differentials, possible demand side explanations are that the relatively large gaps at the bottom and top of the distribution are due to imperfect information or discrimination.

Although no estimates of the immigrant-natives earnings gap across the distribution for the UK are available to compare our estimates with, our results are in line with the limited international literature, which shows that the immigrant-native gap is also more favourable higher up the distribution for the US (Butcher and DiNardo 2002; Chiswick et al. 2008).

4.2   By Continent of Nationality

Although estimates across the earnings distribution are much more informative than our earlier average estimates, they still pool together very diverse groups of immigrants who differ widely in several individual characteristics (such as English proficiency, work ethics, skills (formal education) transferability, etc.) and might not always be perfect labour substitutes. Although we implicitly account for these individual characteristics that are not observed in our dataset when we control for unobservable individual characteristics, we now re-estimate Equation (1) including an explicit indicator for continent of nationality as a proxy for such "group" individual characteristics. Producing estimates by continent of nationality is, of course, informative in itself, as recognized in the existing UK and international literature (Chiswick 1980; Borjas 1994; Butcher and DiNardo 2002; Dustmann and Fabbri 2005; Dickens and McKnight 2008). It is also a way to gain a further insight into the immigrant-native earnings gap from a different angle. This is because, as discussed in Section 2, the various immigration waves to the UK between 1978 and 2006 happened in such a manner that it is possible that immigrants from particular nationalities broadly cluster in different segments of the earnings distribution. We check this by contrasting our new estimates across continents of nationalities with our earlier estimates across the earnings distribution. Estimating the earnings gap in such a flexible manner across continents of nationality is a contribution to the existing UK earnings gap literature, where only estimates by race and ethnicity are available.

Figure 10 shows immigrant-native earnings gap estimates for our base model (left panel) and for our preferred model (right panel) (also see Table A2 in the appendix). As before, the gap substantially narrows in our most complete and preferred specification. It also becomes positive and remains significant for most nationalities, though it is insignificant for immigrants from Africa, Asia and the Middle East and Central and South America. The gap is 2.9% for EU immigrants, 4.6% for immigrants from Australia and Oceania, 5.6% for immigrants from Europe, 7.9% for immigrants from the A10 countries and 10.7% for immigrants from North America. This compares with our earlier average gap estimate of 2.3%

(compare the last column in bold of Tables 2 and A2).

In keeping with the existing UK literature (Chiswick 1980; Bell 1997; Dustmann and Fabbri 2005), our base model estimates suggest a clear and wide dividing line between whites and non-whites (see left panel of Figure 10). However, our preferred model estimates, where we control for observable as well as unobservable individual characteristics, suggest that the gap for non-whites is 0%, whereas it is between 2.9% and 10.7% for whites (see right panel of Figure 10). This suggests a much narrower and subtler dividing line, confirming the importance of accounting for observable as well as unobservable individual characteristics. It also suggests yet again that the labour market primarily rewards individual characteristics other than immigrant status.

We can compare continent of nationality estimates with estimates across the earnings distribution. For example, non-white immigrants (gap of 0%) are overrepresented below the median of the distribution (gap between -7.4% and 0.7%). In contrast, white immigrants (gap between 2.9% and 10.7%) are overrepresented in the top quantile of the distribution (gap between 3.7% and 8.7%). This suggests that nationalities tend to cluster in segments of the distribution, in line with our descriptive analysis in Section 2 (see Figures 3 and 7).

Thus, our conclusion is that the immigrant-native earnings gap in the UK between 1978 and 2006 varies across continents of nationality. Unlike previous research, we found a 0% gap for non-whites when estimating our most complete and preferred specifications. The existing UK literature (Chiswick 1980; Bell 1997; Dustmann and Fabbri 2005) reports earnings gap estimates between -40% and -10% for non-whites when estimating models that are closer in nature to our less complete specifications, which produced comparable estimates magnitudes (between -22% and -16.5%). We also found a positive gap for whites that was smaller than in previous research (between 2.9% and 10.7%).

In sum, non-whites do not seem to suffer an earnings penalty in the labour market as a result of their immigrant status. In contrast, whites experience a favourable gap. We thus maintain our main conclusion from before, that the labour market primarily rewards observable and unobservable individual characteristics other than the immigration status. Again, explanations for this gap could be either supply side productivity differentials or demand side imperfect information and/or discrimination by employers (see Section 4.1).

4.3 By Cohort of Arrival

Allowing our immigrant-native earnings gap estimate to vary across the earnings

15

distribution and across continents of nationality in Sections 4.1 and 4.2 consists in two different identification strategies which allowed an interesting comparison. Another way to re-estimate our Equation (1) is to allow for estimates to vary across cohorts of arrival. Producing estimates by cohort of arrival is, of course, informative in itself, as widely recognized in the literature (Borjas 1985 and 1999; Bell 1997). It is also a way to gain a further insight into the immigrant-native earnings gap here from a third different angle. This is because the nature of the immigration phenomenon to the UK between 1978 and 2006 is such that a comparison between earnings gap estimates across continents of nationality and across cohorts of arrival is particularly appealing. As discussed in Section 2, the various immigration waves to the UK between 1978 and 2006 happened in such a manner that it is possible to associate immigrants from particular nationalities with particular cohorts.

If all cohorts had the same mix of immigrants in terms of continent of nationality, skills, motivation, work ethics, etc., and if the economic conditions in the UK were the same throughout the whole sample period (along with natives' attitudes to migration), then there would be no reason to expect earnings gap estimates to vary across cohorts of arrival. We largely account for local and national macroeconomic conditions in the UK affecting earnings when we control for county and tax-year fixed effects. However, it is still possible that the earnings gap estimates are affected by immigrants' characteristics specific to their cohort of arrival. Thus, we now re-estimate Equation (1) including an explicit indicator for 13 five-year cohorts of arrival as a proxy for such "cohort" individual characteristics. This way we account for characteristics that vary across cohorts such as unmeasured dimensions of immigrants' skills or return migration of immigrants that are more or less able. We also account for the nationality composition of the cohort. This is because most cohorts display a mix of nationalities, and thus a very diverse group of immigrants is pooled together in each cohort. However, in most cohorts there are one or two relatively dominant nationalities. Because of this, we can loosely compare the estimates by cohort of arrival (by pinpointing the dominant continent of nationality in that cohort) with the estimates by continent of nationality. Estimating the earnings gap in such a flexible manner across 13 cohorts of arrival is a contribution to the existing UK earnings gap literature, where such estimates are yet unavailable.

Figure 11 shows immigrant-native earnings gap estimates for our base model (left panel) and for our preferred model (right panel) (also see Table A3 in the appendix). As before, the gap substantially narrows in our most complete and preferred specification, when it is

between -1.7% and 8.6%. It is positive and remains significant for all cohorts, except for 1990-1994, when it is a significant -0.017, and for 1985-1989 and 1945-1949, when it is insignificantly different from zero. This compares with our earlier average gap estimate of 2.3% (compare the last column in bold Tables 2 and A3).

Between 1945 and 1965 the gap is between 0% and 2.6%. These cohorts were dominated by the arrival of white (mainly from Ireland) and non-white (mainly from former colonies) immigrants who came to help with the post-war reconstruction effort. The gap surprisingly jumps to 5.5% and 3.9% for 1965-1969 and 1970-1974 respectively, before coming down again. This should be viewed with caution. As discussed in Section 2, with the new computer system introduced in 1975, data for immigrants who arrived in the immediately preceding years was inputted with a significant delay. Thus the estimates for 1970-1974 (the same is the case for 1945-1948), and to a lesser extent for 1965-1969, might be based on a biased sample, where perhaps whites are incorrectly overrepresented. For example, non-white immigration of Indians peaked between 1965 and 1972, in part because of Gujarati Indians expelled from Uganda, though this does not seem to be accurately recorded in our data (see Figure 6).

Between 1975 and 1984 the gap is between 0.9% and 1.2%. These cohorts saw proportionately large white immigration, following the decision of the UK to join the EU – although the earlier part of this period coincides with greater refugee immigration from India, Vietnam and South East Asia.

Between 1985 and 1994 the gap is smallest. It is insignificant for 1985-1989 and it is -1.7% for 1990-1994. These are cohorts where immigration of both whites and non-whites accelerated. Following the enlargement of the EU in the mid 1980s when Greece, Portugal and Spain joined in, there was greater immigration of lower paid EU whites (also see Figure 7). There was also a modest increase of relatively low paid white immigrants from Australia and Oceania. In addition, there was greater immigration of lower paid non-whites from Africa and Asia and the Middle East.

Between 1995 and 2004 the gap is largest, between 3.7% and 8.6%. These are cohorts when the immigration of both whites and non-whites accelerated steeply. White EU immigration increased sharply, now including higher paid workers (for example from Germany and France) alongside workers from Portugal and Spain as well as from Eastern European countries following further enlargement of the EU. White immigration of higher paid workers from North America and from Australasia and Oceania also contributed to a larger gap in favour of immigrants. In contrast, non-white immigration of lower paid workers

also increased, for example from Africa, Asia and the Middle East, and more modestly from Central and South America.[15]

An attempt to compare cohort of arrival estimates with continent of nationality estimates suggests that, broadly speaking, cohorts with proportionately larger white immigration have a larger positive gap in favour of immigrants. For example, the earnings gap is between -1.7% and 0% for the 1985-1994 cohorts, when there was proportionately large non-white immigration (although this period also coincides with greater lower paid white immigration following the enlargement of the EU in the mid 1980s when Greece, Portugal and Spain joined in). This contrasts with an earnings gap between 3.7% to 8.6% following proportionately large white immigration in the 1995-2004 cohorts, when EU immigration increased along with increased immigration from North America, Australasia and Oceania (although this period also witnessed greater non-white immigration from Africa, Asia and the Middle East).

These are, however, at best tentative conclusions, since most cohorts display a mix of white and non-white immigration. The analysis is further confounded when we differentiate between lower paid and higher paid white immigration, which might also affect the direction and magnitude of the gap. Nonetheless, this is a worthwhile exercise for very broad patterns. For example, it is relatively safe to contrast the -1.7% to 0% gap during the 1985-1994 cohorts, which was dominated by non-whites together with lower paid whites, with the 3.7% to 8.6% gap during the 1995-2004 cohorts, when the proportion of higher paid whites increased. This is in line with our earnings gap estimates for non-whites (0%) and for whites (2.9% to 10.7%) in Section 4.2 and with our descriptive statistics in Section 2 (see Figures 6 and 7).

In sum, our conclusion is that the immigrant-native earnings gap in the UK between 1978 and 2006 varies across cohorts of arrival. We found small (positive and negative) gaps clustered around zero for cohorts that witnessed proportionately larger non-white and lower paid white immigration. Thus, non-whites might sometimes suffer a small earnings penalty in the labour market as a result of their immigrants' status. In contrast, we found a more favourable gap for cohorts that experienced proportionately larger white immigration. We thus maintain our main conclusion from before, that this suggests that the labour market

---

[15] For the single-year-cohort 2005, the gap estimate is an unreasonable 0.440. This is probably due to the fact that this cohort is four years shorter, so it constitutes a smaller sample than the other five-years cohort, for which furthermore, there is only one usable observation for each individual (since the observation from the year before is used up when calculating lagged earnings). Therefore, estimates for this cohort should be treated with caution. For example, the estimate swings wildly across columns in Table A3 (going from -0.765 to 0.440).

primarily rewards observable and unobservable individual characteristics other than simply the immigration status. Again, explanations for this gap could be either supply side productivity differentials or demand side imperfect information and/or discrimination by employers (see Section 4.1).

The earnings gap did not increase monotonically over time. Some of the post-war cohorts do as well, and sometimes better, than more recent cohorts (1945-1959 versus 1985-1994). In contrast, some more recent cohorts, only a few years apart, register gaps which are relatively far apart (1990-1994 and 1995-1999 respectively). This suggests that return migration of more or less able immigrants might not be very severe, because if the least able were more likely to return to their country of origin, earlier cohorts would have more favourable earnings gaps than more recent ones. Nevertheless, it is possible to identify an overall trend of more positive gaps over time if we ignore the 1975-1995 cohorts.

Even though no directly comparable estimates of the immigrant-natives earnings gap across cohorts of arrival are available for the UK (Bell 1997 offers cohort estimates by ethnicity using a different model specification), our results are broadly in line with the international literature – although Borjas (1999), using a different model specification and US data, shows that cohort effects increase monotonically over time.

4.4   By Gender

We now re-estimate Equation (1) separately for males and females, in turn. This is a final way to gain further insight into the immigrant-native earnings gap here. Moreover, although most of the UK and international literature concentrates on estimates for male workers, producing estimates for female workers is informative in itself and is an important area for research (Dustmann and Fabbri 2005; Chiswick and Adsera 2007). One important distinction is the different mobility behaviour of male and female immigrants, both into and across the host country, as well as of male and female natives across the country in response to immigration flows. Another important distinction is the different labour market participation decisions made by men and women, both immigrants and natives, as well as the different employers' decisions in relation to these groups. For example, employers may substitute away from native mothers with small children and towards male immigrants. Also, entry into some occupations in which women are overrepresented might be easier for immigrants, as for instance those that accommodate low paid or part-time work. For example, some of the available evidence in the literature suggests that the minimum wage plays a role in shaping the

earnings distribution for women (Butcher and DiNardo 2002). Our own descriptive evidence in Section 2 also suggests that the unconditional earnings gap is different for men and women (see Figure 2). Thus, estimating the earnings gap by gender is a contribution to the existing UK earnings gap literature, where estimates for female workers are scarcely available.

Figure 12 shows immigrant-native earnings gap estimates for our preferred model across the earnings distribution for male and female workers (also see Table A4 in the appendix). The earnings gap is between -7.9% and 10.3%, in line with the results for the pooled sample (see Section 4.1). As before, the gap is non-negative above the 20th (30th) percentile of the distribution for male (female) and it again increases monotonically across the distribution. It is respectively -0.031, 0.018 and 0.103 for the 10th, 50th and 90th percentiles for males, and -0.022, 0.010, and 0.066 for females. This compares with an average estimate of 0.024 and 0.026 for males and females (see Table A5 in the appendix). Thus, although the average gap is fairly close for males and females, it masks different patterns across the distribution: the estimates are fairly close for males and females in the lower middle of the distribution (20th to 50th percentiles) although they are more negative at the bottom and more positive at the top of the distribution for males. This suggests that male immigrants are at disadvantage at the bottom and at advantage at the top of the distribution in relation to female immigrants, for whom the gap is less variable throughout the distribution.

Figure 13 shows immigrant-native earnings gap estimates for our preferred model across continents of nationality (left panel) and cohort of arrival (right panel) for male and female workers (compare Table A5 with the last column in bold of Tables 2, A2 and A3 in the appendix). Across continents of nationality, the earnings gap is between 0% and 11.8%, in line with the results for the pooled sample (see Section 4.2). They follow the same pattern, although the estimates for females are now more often significant. As before, the gap is 0% for non-white males and females (except for 1.2% for Asia and the Middle East female immigrants) and between 2.8% and 11.8% for whites. The gap is again more positive for some male immigrants (from North America and A10) but fairly close, in the main, between males and females.

Across cohorts of arrival, the earnings gap is between -2.8% and 11%, in line with the results for the pooled sample (see Section 4.3). They follow the same pattern, although the estimates for males (females) are now more (less) often significant. As before, broadly speaking, cohorts with proportionately larger white immigration have a larger positive gap in favour or immigrants. For example, the earnings gap is between -2.8% and 0% for the 1985-

1994 cohorts, when there was proportionately large non-white immigration, and between 1.7% and 11% for the 1995-2004 when there was proportionately large white immigration (see Section 4.3).

In sum, our main conclusions from before are maintained. The immigrant-native earnings gap in the UK between 1978 and 2006 for male and female workers varies across the earnings distribution, across continents of nationality and across cohorts of arrival. This is in line with the limited UK evidence, which also shows that the immigrant-native gap varies across gender (Dustmann and Fabbri 2005).

5. Discussion

Our results suggest that there is indeed an immigrant-native earnings gap in the UK between 1978 and 2006. We initially found that immigrants earn 2.3% more than natives on average. However, this gap varied substantially across the distribution, increasing monotonically, and this variability was masked when solely the average gap was considered. While the gap was zero at the 30th percentile, it was between -7.4% and 0.7% below the median, and between 2.4% and 8.7% above the median. We also found a gap for non-white immigrants, who are overrepresented below the median, of 0%; and a gap for white immigrants, who are overrepresented in the top quantile of the distribution, between 2.9% and 10.7%. This suggests that immigrants from certain nationalities tend to cluster in segments of the distribution. Finally, we found less favourable gaps for cohorts that witnessed proportionately larger non-white and lower paid white immigration (e.g. between -1.7% and 0% in 1985-1994) and more favourable gaps for cohorts experiencing larger white immigration (e.g. between 3.7% and 8.6% in 1995-2004).

Thus, on the one hand, the lowest paid immigrants, whom are disproportionately non-white, suffer an earnings penalty in relation to the lowest paid natives with comparable individual characteristics. On the other hand, other immigrants do not seem to suffer a penalty – the gap was fairly small in the lower middle of the distribution and it was in favour of higher paid immigrants, whom are disproportionately white.

Our main conclusion deriving from these findings is that most immigrants do not seem to suffer an earnings penalty in the labour market. This suggests that, except for the lowest and highest paid immigrants, the labour market primarily rewards observable and unobservable individual characteristics other than the immigration status.

Our main conclusion stems from the estimates in our more complete models. When we

estimated the earnings gap using simple models, comparable to those in the limited existing UK literature, we found comparable results that confirmed the previously documented clear and wide dividing line between whites and non-whites. However, when we estimated the earnings gap using more complete models, where we control not only for observable, as it is common in the literature, but also for unobservable individual characteristics, we found a much narrower and subtler dividing line than that in the existing UK literature. This confirms the importance of accounting for observable as well as unobservable individual characteristics, which is an important contribution of this paper. And this is what drives our main conclusion that the labour market primarily rewards individual characteristics other than immigration status.

Our estimates are in line with our earlier descriptive analysis and are robust to a number of specification checks as well as to different stratifications of the labour market and to different sub-samples of workers. Although our main conclusion is that most immigrants do not seem to suffer a largely favourable or unfavourable earnings gap, our results confirm that, especially at the bottom and top of the earnings distribution, there is indeed an immigrant-native earnings gap in the UK between 1978 and 2006. According to standard human capital theory, non-negligible immigrant-native earnings gap estimates could be due to supply side productivity differentials between natives and immigrants, perhaps unaccounted for in our model. This could be due to imperfect transferability of immigrants' country-specific human capital or immigrants' skill downgrading due to language or other labour market barriers. For example, employers attach a different weight to experience gained prior to immigration. Even in the absence of such supply side productivity differentials, there could be demand side explanations for non-negligible earnings gap, such as imperfect information or discrimination in the labour market. These considerations point towards fruitful areas for future research which inherently raise a wealth of policy implications.

6. Conclusion

Exploiting a sizeable and long longitudinal dataset we estimated the immigrant-native earnings gap across the entire earnings distribution, across continents of nationality and across cohorts of arrival in the UK between 1978 and 2006.

In keeping with the limited existing UK literature, we found a clear and wide dividing line between whites and non-whites when we estimated the earnings gap using simple comparable models. However, when we estimated the earnings gap using more complete models, where

we control not only for observable, as it is common in the literature, but also for unobservable individual characteristics, which is an important contribution of this paper, we found a much narrower and subtler dividing line than that in the existing UK literature: a gap of 0% for non-whites and a gap between 2.9% and 10.7% for whites. This confirms the importance of accounting for observable as well as unobservable individual characteristics. It also suggests that the labour market primarily rewards individual characteristics other than immigration status.

When we estimated the gap across the earnings distribution, we found that it was between -7.4% and 0.7% below the median, where non-white immigrants are overrepresented, and between 2.4% and 8.7% above the median, where white immigrants are overrepresented. This suggests that most immigrants do not seem to suffer an earnings penalty in the labour market as a result of their immigrants' status. Finally, we found less favourable gaps for cohorts that witnessed proportionately larger non-white and lower paid white immigration.

This new evidence is an important contribution to the very limited UK immigration literature – in particular, it helps to fill a blank in the literature on the immigrant-native earnings gap in the UK. This new evidence is also an important contribution to the international immigration literature because it applies a thorough empirical estimation approach to an underexplored, long and sizeable longitudinal dataset to carefully study a crucial labour market issue: the immigrant-native earnings gap, which is an objective measure of how immigrants fare in the labour market. We exploit the longitudinal nature of our dataset to separate the effect of individual unobserved characteristics from the effect of other individual observed characteristics on earnings – which in turn helps prevent biases implied by various types of selectivity in the data, such as cohort bias and survivor bias. Given that such biases constitute an identification issue that has occupied much of the literature for the last 30 years, and given that this literature has long recognized that the ideal solution for this is to use sufficiently large and long longitudinal data – always very scarce – this paper is a timely contribution.

This new evidence is, most crucially, an important and timely contribution to informing policymaking in the face of continuing and heated public debate on immigration policy in the UK and internationally.

7. References

BELL, B. D. (1997): "The Performance of Immigrants in the United Kingdom: Evidence from the Ghs," *The Economic Journal*, 107, 333-44.

BORJAS, G. J. (1985): "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants," *Journal of Labor Economics*, 3, 463-489.

— (1994): "The Economics of Immigration," *Journal of Economic Literature*, 32, 1667-1717.

— (1995): "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?," *Journal of Labor Economics*, 13, 201-245.

— (1999): "The Economic Analysis of Immigration," in *Handbook of Labor Economics, Vol. 3a*, ed. by O. Ashenfelter, and D. Card. Amsterdam: North-Holland, 1697-1760.

— (2003): "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *The Quarterly Journal of Economics*, 118, 1335-1374.

BUTCHER, K. F., and J. DINARDO (2002): "The Immigration and Native-Born Wage Distributions: Evidence from United States Censuses," *Industrial and Labor Relations Review*, 56, 97-121.

CARD, D. (2001): "Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19, 22-64.

CARLINER, G. (1980): "Wages, Earnings and Hours of First, Second and Third Generation American Males," *Economic Inquiry*, 18, 87-102.

CHISWICK, B. R. (1978): "The Effects of Americanization on the Earnings of Foreign-Born Men," *Journal of Political Economy*, 86, 897-921.

— (1980): "The Earnings of White and Coloured Male Immigrants in Britain," *Economica*, 47, 81-87

CHISWICK, B. R., and A. ADSERA (2007): "Are There Gender and Country of Origin Differences in Immigrant Labor Market Outcomes across European Destinations," *Journal of Population Economics* 20, 495-526.

CHISWICK, B. R., A. T. LE, and P. W. MILLER (2008): "How Immigrants Fare across the Earnings Distribution in Australia and the United States," *Industrial and Labor Relations Review*, 61, 353-373.

CHISWICK, B. R., Y. L. LEE, and P. W. MILLER (2005): "Immigrant Earnings: A Longitudinal Analysis," *Review of Income and Wealth*, 51, 485-503.

CORTES, K. E. (2004): "Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States," *The Review of Economics and Statistics*, 86, 465-480.

DICKENS, R., and A. MCKNIGHT (2008): "Assimilation of Migrants into the British Labour Market," *CASE Discussion Paper*, 133.

DULEEP, H., and M. REGETS (1997): "Measuring Immigrant Wage Growth Using Matched Cps Files," *Demography*, 34, 239-249.

DUSTMANN, C., and F. FABBRI (2005): "Immigration in the British Labour Market," *Fiscal Studies*, 26, 423-470.

HALVORSEN, R., and R. PALMQUIST (1980): "The Interpretation of Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 70, 474-475.

LALONDE, J. R., and H. R. TOPEL (1992): "The Assimilation of Immigrants in the U.S. Labor Market," in *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, ed. by G. Borjas, and R. Freeman. Chicago: University of Chicago Press, 67-92.

LEMOS, S. (2011): "Immigrant Economic Assimilation: Evidence from Uk Longitudinal Data between 1978 and 2006," *Manuscript*.

MANACORDA, M., A. MANNING, and J. WADSWORTH (2006): "The Impact of Immigration on the Structure of Male Wages: Theory and Evidence from Britain," *IZA Discussion Paper*, 2352.

NAKAMURA, A., and M. NAKAMURA (1985): *The Second Paycheck: A Socioeconomi Analysis of Earnings*. New York: Academic Press.

ONS (2003): "United Kingdom: Local Authority Districts, Counties and Unitary Authorities, 1998," *ONS Geography GIS & Mapping Unit*.

Table 1 - DESCRIPTIVE STATISTICS						
VARIABLES	**LLMDB**		**LLMDB**		**LFS**	
	April 1978 - March 2007		**April 1997 - March 2007**		**January 1997 - March 20**	
	Natives	Immigrants	Natives	Immigrants	Natives	Immigrants
I - POPULATION VARIABLES						
%aged:						
25 to 34 years old	31.34%	43.83%	29.47%	43.24%	29.06%	36.59%
35 to 64 years old	68.66%	56.17%	70.53%	56.76%	70.94%	63.41%
%of women	43.43%	44.85%	46.67%	44.37%	48.47%	47.51%
%from:						
EU (except A10)	-	32.20%	-	30.99%	-	25.66%
A10	-	4.03%	-	5.41%	-	5.43%
Europe (except EU)	-	3.15%	-	3.64%	-	2.65%
Asia and Middle East	-	20.89%	-	22.38%	-	27.94%
North America	-	6.09%	-	5.04%	-	4.63%
Latin America	-	3.24%	-	3.49%	-	6.59%
Africa	-	13.80%	-	16.16%	-	21.39%
Australasia and Oceania	-	7.30%	-	6.39%	-	4.74%
Unknown	-	9.31%	-	6.51%	-	na
Average age at arrival	-	23.03	-	24.43	-	na
Average nb of years since immigration	-	14.73	-	13.30	-	na
%with lenght of immigration						
0 to 1 years	-	8.64%	-	11.15%	-	na
2 to 3 years	-	10.17%	-	12.54%	-	na
4 to 5 years	-	8.48%	-	9.73%	-	na
6 to 10 years	-	16.96%	-	17.06%	-	na
11 to 15 years	-	15.24%	-	14.02%	-	na
16 to 20 years	-	11.70%	-	10.46%	-	na
over 20 years	-	28.82%	-	25.04%	-	na
%arrived during:						
1945-1949	-	3.87%	-	0.04%	-	1.03%
1950-1954	-	2.63%	-	0.46%	-	1.77%
1955-1959	-	4.54%	-	1.69%	-	3.45%
1960-1964	-	5.79%	-	2.84%	-	7.23%
1965-1969	-	6.05%	-	3.38%	-	8.57%
1970-1974	-	4.38%	-	2.79%	-	9.16%
1975-1979	-	17.45%	-	11.77%	-	7.61%
1980-1984	-	9.62%	-	8.48%	-	5.57%
1985-1989	-	12.72%	-	13.47%	-	7.69%
1990-1994	-	10.54%	-	14.22%	-	9.10%
1995-1999	-	10.59%	-	18.95%	-	13.49%
2000-2004	-	10.76%	-	19.92%	-	13.26%
2005	-	1.08%	-	1.99%	-	1.39%
%located in:						
East Midlands	7.64%	4.26%	7.68%	4.28%	7.66%	4.79%
East of England	9.50%	7.28%	9.55%	7.62%	9.69%	9.12%
London	8.73%	29.87%	9.03%	35.14%	9.61%	41.76%
North East	4.46%	2.08%	4.35%	1.88%	4.32%	1.31%
North West	11.72%	5.54%	11.56%	5.57%	11.97%	5.57%
Northern Ireland	2.39%	1.26%	2.51%	1.29%	2.51%	1.22%
Scotland	9.43%	5.16%	9.39%	4.78%	9.03%	3.67%
South East	13.73%	12.37%	13.93%	12.87%	13.99%	13.87%
South West	8.79%	5.25%	8.76%	5.09%	8.51%	5.24%
Wales	4.87%	2.17%	4.79%	1.95%	4.87%	1.80%
West Midlands	9.16%	5.10%	9.10%	5.26%	9.20%	6.90%
Yorkshire and the Humber	8.68%	4.12%	8.66%	4.13%	8.66%	4.74%
Unknown or Abroad	0.90%	15.55%	0.69%	10.15%	-	-
						(continues)

Table 1 - DESCRIPTIVE STATISTICS (continued)

VARIABLES	LLMDB		LLMDB		LFS	
	April 1978 - March 2007		**April 1997 - March 2007**		**January 1997 - March 20**	
	Natives	Immigrants	Natives	Immigrants	Natives	Immigrants
II - LABOUR MARKET VARIABLES						
% in work:						
1 to 25 weeks in the year	16.74%	26.05%	18.84%	30.33%	na	na
26 to 50 weeks in the year	13.93%	18.96%	15.45%	20.95%	na	na
51 to 52 weeks in the year	69.09%	54.76%	65.55%	48.54%	na	na
Average number of employed weeks in the year	43.18	38.52	42.10	36.40	na	na
Average number of unemployed weeks	1.37	1.39	1.03	1.17	na	na
Average number of jobs in the year	1.33	1.53	1.44	1.66	na	na
5th percentile of the log real earnings distribution	7.60	7.27	7.65	7.25	9.22	9.15
10th percentile of the log real earnings distribution	8.26	7.97	8.35	7.99	9.39	9.34
20th percentile of the log real earnings distribution	8.87	8.65	8.93	8.68	9.59	9.57
30th percentile of the log real earnings distribution	9.26	9.09	9.30	9.12	9.73	9.74
40th percentile of the log real earnings distribution	9.50	9.40	9.55	9.43	9.86	9.88
50th percentile of the log real earnings distribution	9.68	9.64	9.74	9.68	9.99	10.03
60th percentile of the log real earnings distribution	9.84	9.86	9.91	9.90	10.12	10.17
70th percentile of the log real earnings distribution	9.99	10.06	10.08	10.13	10.26	10.32
80th percentile of the log real earnings distribution	10.16	10.29	10.27	10.37	10.42	10.51
90th percentile of the log real earnings distribution	10.41	10.62	10.52	10.74	10.65	10.80
Average of the log real earnings distribution	9.48	9.45	9.55	9.50	10.00	10.04
Standard deviation of the log real earnings distribution	0.95	1.13	0.99	1.18	0.53	0.61
Number of observations	5053659	319427	1935699	172466	507606	42230
Number of individuals	387760	45309	277532	35415	na	na
Average number of times an individual is observed	15.76	18.71	24.53	25.16	na	na
% observations per year:						
1978	2.89%	1.69%	-	-	-	-
1979	3.16%	2.02%	-	-	-	-
1980	3.07%	1.97%	-	-	-	-
1981	3.09%	1.96%	-	-	-	-
1982	3.08%	1.95%	-	-	-	-
1983	3.07%	1.95%	-	-	-	-
1984	3.14%	2.05%	-	-	-	-
1985	3.08%	2.03%	-	-	-	-
1986	3.11%	2.08%	-	-	-	-
1987	3.11%	2.14%	-	-	-	-
1988	3.26%	2.35%	-	-	-	-
1989	3.34%	2.52%	-	-	-	-
1990	3.42%	2.67%	-	-	-	-
1991	3.38%	2.74%	-	-	-	-
1992	3.42%	2.90%	-	-	-	-
1993	3.44%	3.02%	-	-	-	-
1994	3.50%	3.19%	-	-	-	-
1995	3.56%	3.32%	-	-	-	-
1996	3.56%	3.45%	-	-	-	-
1997	3.68%	3.67%	9.62%	6.80%	10.51%	9.23%
1998	3.73%	3.91%	9.75%	7.24%	11.55%	10.50%
1999	3.81%	4.30%	9.95%	7.96%	11.11%	10.15%
2000	3.86%	4.69%	10.08%	8.69%	10.56%	9.27%
2001	3.88%	5.00%	10.13%	9.26%	7.72%	7.30%
2002	3.88%	5.41%	10.13%	10.01%	10.14%	10.27%
2003	3.90%	5.91%	10.18%	10.94%	9.73%	9.86%
2004	3.83%	6.52%	9.99%	12.08%	9.23%	9.54%
2005	3.90%	7.43%	10.17%	13.75%	8.87%	10.23%
2006	3.83%	7.16%	10.00%	13.26%	8.44%	10.79%
Average number of observations per year	174264	11015	193570	17247	50761	4223

Source: Lifetime Labour Market Database and Labour Force Survey.
(1) Sample includes those aged 25 to 64 employed and earning between £100 and £1000000 in the year.
(2) LFS figures are courtesy of the DWP.

Table 2 - Immigrant-Native Earnings Gap

Variable	coefficient	s. errors	coefficient	s. errors	coefficient	s. errors
Intercept	**9.647**	0.009	**6.052**	0.013	**1.728**	0.008
Immigrant (=1)	**-0.132**	0.006	**-0.023**	0.004	**0.023**	0.002
Sex (male=1)			**0.595**	0.002	**0.175**	0.001
Age			**0.048**	0.000	**0.020**	0.000
Age squared			**-0.001**	0.000	**0.000**	0.000
Number of employed weeks			**0.040**	0.000	**0.023**	0.000
Number of jobs			**0.347**	0.003	**0.202**	0.002
Lagged number of employed weeks					**-0.005**	0.000
Lagged log real earnings					**0.668**	0.001
Adjusted R-squared	0.02		0.49		0.72	
Sample size	5373086		5366162		4893023	
Number of individuals	433069		433069		433069	

(1) The sample data used includes those aged 25 to 64 earning between £100 and £1000000 in any one tax year who are observed at least twice (it excludes the self-employed). It includes immigrants arriving from 1945 onwards. See text for details.

(2) All models include area fixed effects (49 county dummies) and time fixed effects (29 tax-year dummies). Only the the model in the right-most column controls for individual fixed effects via lagged log real earnings and lagged number of employed weeks. See text for details.

(3) All models are corrected for intragroup correlation, as standard errors are assumed independent across groups of individuals but not within groups (i.e. errors are not assumed independent for a particular individual over time). See text for details.

Table A1 - Immigrant-Native Earnings Gap Across the Earnings Distribution

Variable	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.
Percentile	**5th**		**10th**		**20th**		**30th**		**40th**		**50th**		**60th**		**70th**		**80th**		**90th**	
Intercept	8.133	0.018	8.637	0.011	9.217	0.008	9.511	0.006	9.671	0.004	9.799	0.004	9.910	0.003	10.011	0.003	10.136	0.003	10.325	0.004
Immigrant (=1)	-0.393	0.008	-0.372	0.005	-0.318	0.004	-0.274	0.003	-0.203	0.002	-0.140	0.002	-0.084	0.001	-0.033	0.001	0.017	0.001	0.107	0.002
R-squared	0.01		0.01		0.01		0.01		0.01		0.01		0.01		0.02		0.03		0.04	
Sample size	5373086		5373086		5373086		5373086		5373086		5373086		5373086		5373086		5373086		5373086	
Number of Individuals	433069		433069		433069		433069		433069		433069		433069		433069		433069		433069	
Intercept	-1.613	0.007	-1.222	0.004	-0.673	0.002	-0.208	0.002	0.192	0.002	0.421	0.001	0.740	0.002	1.218	0.002	1.845	0.003	2.746	0.005
Immigrant (=1)	-0.074	0.001	-0.030	0.001	-0.008	0.000	0.000	0.000	0.007	0.000	0.014	0.000	0.024	0.000	0.037	0.000	0.054	0.001	0.087	0.001
Sex (male=1)	0.097	0.001	0.062	0.000	0.041	0.000	0.036	0.000	0.033	0.000	0.031	0.000	0.041	0.000	0.055	0.000	0.074	0.000	0.109	0.000
Age	0.016	0.000	0.008	0.000	0.005	0.000	0.005	0.000	0.004	0.000	0.003	0.000	0.003	0.000	0.004	0.000	0.006	0.000	0.010	0.000
Age squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Number of employed weeks	0.050	0.000	0.040	0.000	0.028	0.000	0.020	0.000	0.013	0.000	0.010	0.000	0.008	0.000	0.007	0.000	0.006	0.000	0.006	0.000
Number of jobs	0.075	0.001	0.078	0.000	0.083	0.000	0.081	0.000	0.071	0.000	0.071	0.000	0.083	0.000	0.091	0.000	0.102	0.000	0.129	0.000
Lagged number of employed weeks	0.000	0.000	-0.001	0.000	-0.002	0.000	-0.003	0.000	-0.004	0.000	-0.005	0.000	-0.006	0.000	-0.008	0.000	-0.012	0.000	-0.019	0.000
Lagged log real wages	0.835	0.001	0.884	0.000	0.909	0.000	0.915	0.000	0.920	0.000	0.924	0.000	0.906	0.000	0.876	0.000	0.837	0.000	0.775	0.000
R-squared	0.63		0.64		0.64		0.62		0.61		0.60		0.58		0.55		0.52		0.48	
Sample size	4893023		4893023		4893023		4893023		4893023		4893023		4893023		4893023		4893023		4893023	
Number of Individuals	433069		433069		433069		433069		433069		433069		433069		433069		433069		433069	

(I) Notes as in Table 2.

Table A2 - Immigrant-Native Earnings Gap by Continent of Nationality						
Variable	coefficient	s. errors	coefficient	s. errors	coefficient	s. errors
Intercept	**9.617**	0.010	**6.020**	0.013	**1.722**	0.008
EU (except A10)	**-0.067**	0.010	**0.031**	0.007	**0.029**	0.003
A10	**-0.424**	0.022	**-0.140**	0.015	**0.079**	0.008
Europe (except EU)	**-0.019**	0.030	**0.028**	0.022	**0.056**	0.008
Asia and Middle East	**-0.166**	0.012	**-0.116**	0.009	**0.001**	0.003
North America	**0.265**	0.023	**0.287**	0.017	**0.107**	0.007
Central and South America	**-0.218**	0.026	**-0.097**	0.018	**0.005**	0.008
Africa	**-0.165**	0.013	**-0.108**	0.010	**-0.006**	0.004
Australasia and Oceania	**-0.036**	0.017	**0.109**	0.013	**0.046**	0.006
Unknown	**-0.396**	0.015	**-0.085**	0.011	**0.003**	0.004
Sex (male=1)			**0.595**	0.002	**0.175**	0.001
Age			**0.048**	0.000	**0.021**	0.000
Age squared			**-0.001**	0.000	**0.000**	0.000
Number of employed weeks			**0.040**	0.000	**0.023**	0.000
Number of jobs			**0.347**	0.003	**0.202**	0.002
Lagged number of employed weeks					**-0.005**	0.000
Lagged log real earnings					**0.667**	0.001
Adjusted R-squared	0.02		0.49		0.72	
Sample size	5373086		5366162		4893023	
Number of individuals	433069		433069		433069	

(I) Notes as in Table 2.

Table A3 - Immigrant-Native Earnings Gap by Cohort of Arrival						
Variable	coefficient	s. errors	coefficient	s. errors	coefficient	s. errors
Intercept	**9.604**	0.010	**6.051**	0.013	**1.722**	0.008
1945-1949 arrivals	**-0.247**	0.022	**-0.002**	0.015	**0.008**	0.006
1950-1954 arrivals	**-0.043**	0.030	**0.009**	0.020	**0.016**	0.008
1955-1959 arrivals	**0.022**	0.026	**0.009**	0.018	**0.017**	0.007
1960-1964 arrivals	**0.074**	0.024	**0.044**	0.016	**0.026**	0.006
1965-1969 arrivals	**0.174**	0.023	**0.137**	0.016	**0.055**	0.006
1970-1974 arrivals	**0.071**	0.027	**0.091**	0.019	**0.039**	0.007
1975-1979 arrivals	**-0.042**	0.014	**0.010**	0.010	**0.009**	0.004
1980-1984 arrivals	**-0.080**	0.017	**-0.008**	0.013	**0.012**	0.005
1985-1989 arrivals	**-0.138**	0.014	**-0.058**	0.011	**-0.005**	0.004
1990-1994 arrivals	**-0.257**	0.014	**-0.136**	0.012	**-0.017**	0.005
1995-1999 arrivals	**-0.250**	0.013	**-0.049**	0.011	**0.037**	0.005
2000-2004 arrivals	**-0.408**	0.010	**-0.094**	0.008	**0.086**	0.004
2005 arrivals	**-0.765**	0.022	**-0.211**	0.017	**0.440**	0.018
Sex (male=1)			**0.595**	0.002	**0.175**	0.001
Age			**0.048**	0.000	**0.021**	0.000
Age squared			**-0.001**	0.000	**0.000**	0.000
Number of employed weeks			**0.040**	0.000	**0.023**	0.000
Number of jobs			**0.347**	0.003	**0.202**	0.002
Lagged number of employed weeks					**-0.005**	0.000
Lagged log real earnings					**0.668**	0.001
Adjusted R-squared	0.02		0.49		0.72	
Sample size	5373086		5366162		4893023	
Number of individuals	433069		433069		433069	

(1) Notes as in Table 2.

Table A4 - Immigrant-Native Earnings Gap Across the Earnings Distribution by Gender

Variable	5th coef	s.e.	10th coef	s.e.	20th coef	s.e.	30th coef	s.e.	40th coef	s.e.	50th coef	s.e.	60th coef	s.e.	70th coef	s.e.	80th coef	s.e.	90th coef	s.e.
Male																				
Intercept	**-1.915**	0.008	**-1.405**	0.004	**-0.818**	0.002	**-0.357**	0.002	**0.085**	0.002	**0.398**	0.002	**0.786**	0.002	**1.302**	0.002	**1.958**	0.003	**2.880**	0.006
Immigrant (=1)	**-0.079**	0.002	**-0.031**	0.001	**-0.007**	0.001	**0.003**	0.000	**0.010**	0.000	**0.018**	0.000	**0.029**	0.000	**0.044**	0.000	**0.064**	0.001	**0.103**	0.001
Age	**0.013**	0.000	**0.007**	0.000	**0.005**	0.000	**0.004**	0.000	**0.003**	0.000	**0.003**	0.000	**0.003**	0.000	**0.005**	0.000	**0.008**	0.000	**0.014**	0.000
Age squared	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000
Number of employed weeks	**0.052**	0.000	**0.041**	0.000	**0.029**	0.000	**0.021**	0.000	**0.014**	0.000	**0.009**	0.000	**0.007**	0.000	**0.005**	0.000	**0.004**	0.000	**0.004**	0.000
Number of jobs	**0.073**	0.001	**0.078**	0.000	**0.083**	0.000	**0.086**	0.000	**0.075**	0.000	**0.069**	0.000	**0.067**	0.000	**0.067**	0.000	**0.073**	0.000	**0.099**	0.000
Lagged number of employed weeks	**0.002**	0.000	**0.000**	0.000	**-0.001**	0.000	**-0.002**	0.000	**-0.003**	0.000	**-0.004**	0.000	**-0.005**	0.000	**-0.008**	0.000	**-0.013**	0.000	**-0.020**	0.000
Lagged log real wages	**0.861**	0.001	**0.898**	0.000	**0.919**	0.000	**0.923**	0.000	**0.925**	0.000	**0.927**	0.000	**0.910**	0.000	**0.880**	0.000	**0.843**	0.000	**0.785**	0.001
R-squared	0.62		0.61		0.58		0.55		0.54		0.53		0.51		0.50		0.47		0.44	
Sample size	2777556		2777556		2777556		2777556		2777556		2777556		2777556		2777556		2777556		2777556	
Number of Individuals	234982		234982		234982		234982		234982		234982		234982		234982		234982		234982	
Female																				
Intercept	**-0.937**	0.011	**-0.921**	0.007	**-0.492**	0.004	**-0.079**	0.004	**0.250**	0.003	**0.424**	0.003	**0.726**	0.003	**1.188**	0.004	**1.803**	0.005	**2.748**	0.008
Immigrant (=1)	**-0.041**	0.002	**-0.022**	0.001	**-0.008**	0.001	**-0.002**	0.001	**0.004**	0.001	**0.010**	0.001	**0.019**	0.001	**0.030**	0.001	**0.044**	0.001	**0.066**	0.001
Age	**0.014**	0.000	**0.007**	0.000	**0.005**	0.000	**0.005**	0.000	**0.004**	0.000	**0.004**	0.000	**0.003**	0.000	**0.005**	0.000	**0.006**	0.000	**0.007**	0.000
Age squared	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000	**0.000**	0.000
Number of employed weeks	**0.048**	0.000	**0.039**	0.000	**0.027**	0.000	**0.019**	0.000	**0.013**	0.000	**0.011**	0.000	**0.010**	0.000	**0.009**	0.000	**0.009**	0.000	**0.009**	0.000
Number of jobs	**0.082**	0.001	**0.083**	0.001	**0.084**	0.000	**0.080**	0.000	**0.070**	0.000	**0.078**	0.000	**0.099**	0.000	**0.117**	0.000	**0.134**	0.000	**0.166**	0.000
Lagged number of employed weeks	**-0.002**	0.000	**-0.002**	0.000	**-0.003**	0.000	**-0.004**	0.000	**-0.006**	0.000	**-0.006**	0.000	**-0.007**	0.000	**-0.009**	0.000	**-0.012**	0.000	**-0.017**	0.000
Lagged log real wages	**0.769**	0.001	**0.855**	0.000	**0.895**	0.000	**0.907**	0.000	**0.916**	0.000	**0.921**	0.000	**0.900**	0.000	**0.865**	0.000	**0.822**	0.000	**0.753**	0.001
R-squared	0.61		0.61		0.61		0.60		0.60		0.59		0.57		0.55		0.52		0.47	
Sample size	2115467		2115467		2115467		2115467		2115467		2115467		2115467		2115467		2115467		2115467	
Number of Individuals	198087		198087		198087		198087		198087		198087		198087		198087		198087		198087	

(l) Notes as in Table 2.

Variable	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.
Table A5 - Immigrant-Native Earnings Gap by Gender												
	Male		Female		Male		Female		Male		Female	
Intercept	1.736	0.011	1.848	0.012	1.731	0.011	1.843	0.012	1.729	0.011	1.842	0.012
Immigrant (=1)	0.024	0.002	0.026	0.002								
EU (except A10)					0.029	0.004	0.028	0.004				
A10					0.096	0.011	0.059	0.010				
Europe (except EU)					0.059	0.010	0.062	0.014				
Asia and Middle East					-0.001	0.004	0.012	0.005				
North America					0.118	0.009	0.089	0.010				
Central and South America					0.004	0.011	-0.003	0.010				
Africa					-0.006	0.006	-0.002	0.006				
Australasia and Oceania					0.032	0.008	0.056	0.008				
Unknown					0.016	0.011	0.018	0.005				
1945-1949 arrivals									0.046	0.013	0.017	0.006
1950-1954 arrivals									0.036	0.011	-0.003	0.011
1955-1959 arrivals									0.032	0.009	0.000	0.009
1960-1964 arrivals									0.038	0.008	0.007	0.009
1965-1969 arrivals									0.067	0.008	0.038	0.010
1970-1974 arrivals									0.051	0.010	0.030	0.010
1975-1979 arrivals									0.013	0.005	0.011	0.005
1980-1984 arrivals									0.016	0.007	0.008	0.007
1985-1989 arrivals									-0.011	0.006	0.003	0.006
1990-1994 arrivals									-0.028	0.007	-0.004	0.007
1995-1999 arrivals									0.017	0.007	0.060	0.007
2000-2004 arrivals									0.067	0.006	0.110	0.006
2005 arrivals									0.466	0.023	0.415	0.030
Age	0.027	0.000	0.012	0.000	0.028	0.000	0.012	0.000	0.028	0.000	0.012	0.000
Age squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Number of employed weeks	0.023	0.000	0.024	0.000	0.023	0.000	0.024	0.000	0.023	0.000	0.024	0.000
Number of jobs	0.202	0.002	0.205	0.003	0.202	0.002	0.205	0.003	0.202	0.002	0.205	0.003
Lagged number of employed	-0.004	0.000	-0.006	0.000	-0.004	0.000	-0.006	0.000	-0.004	0.000	-0.006	0.000
Lagged log real earnings	0.665	0.001	0.664	0.001	0.664	0.001	0.664	0.001	0.665	0.001	0.664	0.001
R-squared	0.65		0.72		0.65		0.72		0.65		0.72	
Sample size	2777556		2115467		2777556		2115467		2777556		2115467	
Number of individuals	234982		198087		234982		198087		234982		198087	

(1) Notes as in Table 2.

Figure 1 - Natives and Immigrants
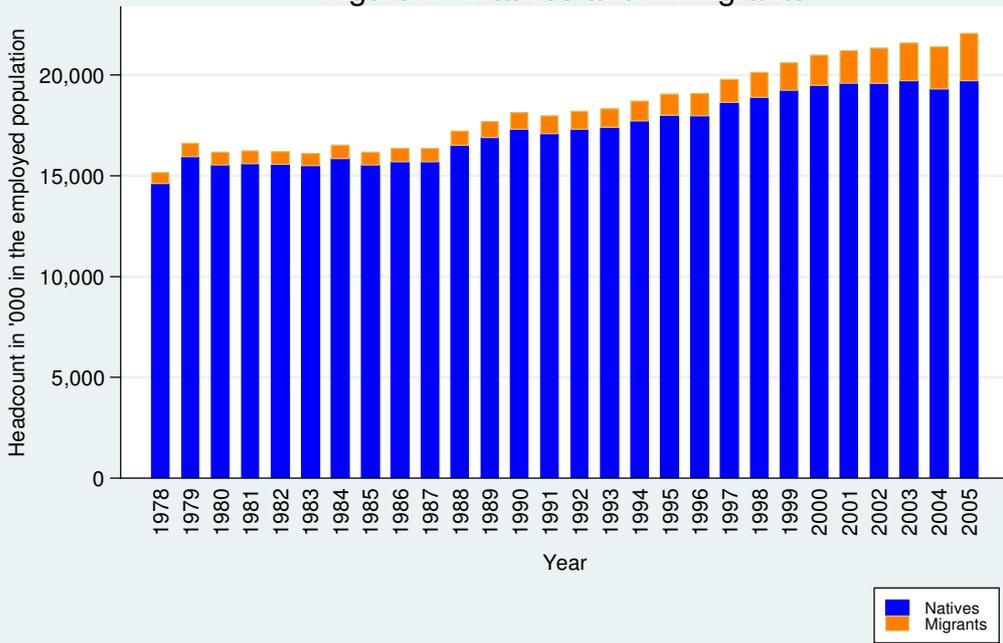
Source: Lifetime Labour Market Database

Figure 2 - Immigrant-Native Earnings Gap
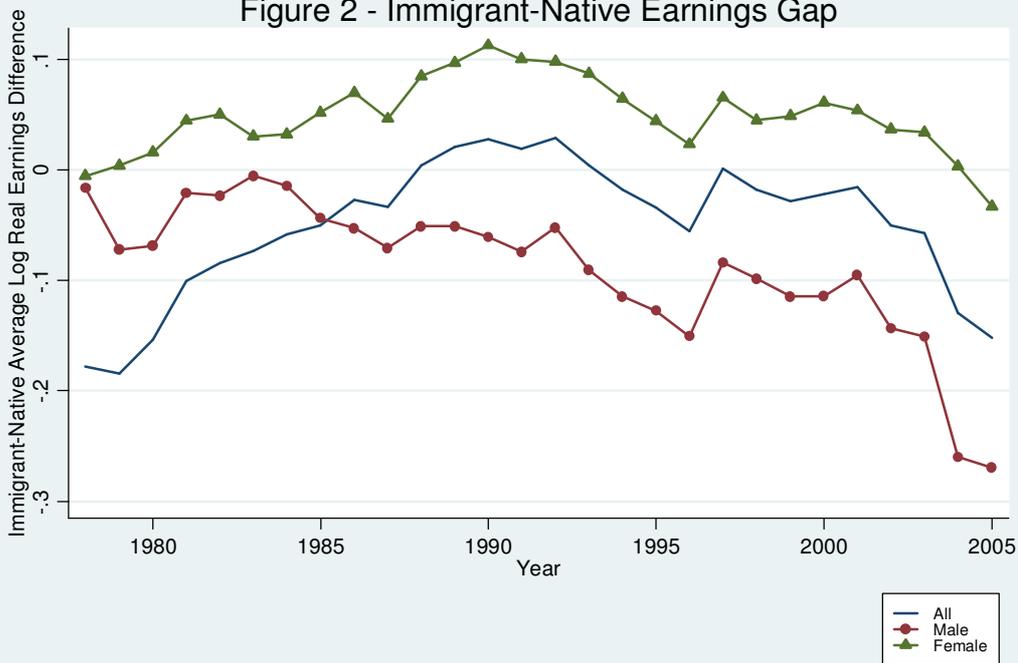
Source: Lifetime Labour Market Database

Figure 3 - Immigrant-Native Earnings Gap Across the Distribution

Source: Lifetime Labour Market Database
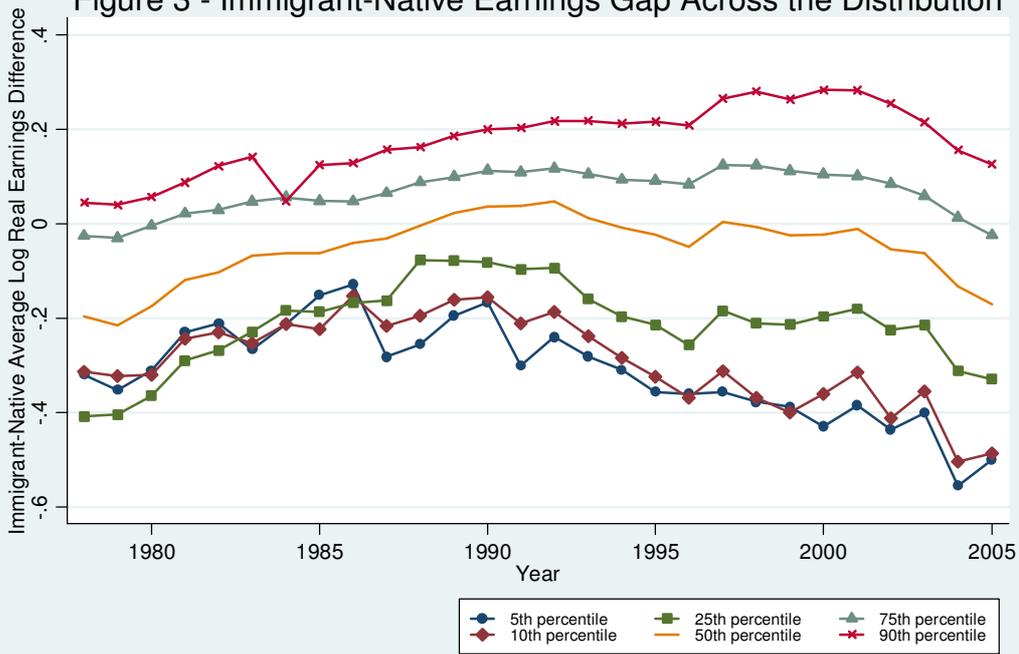
Figure 4 - Real Earnings Distribution

Source: Lifetime Labour Market Database
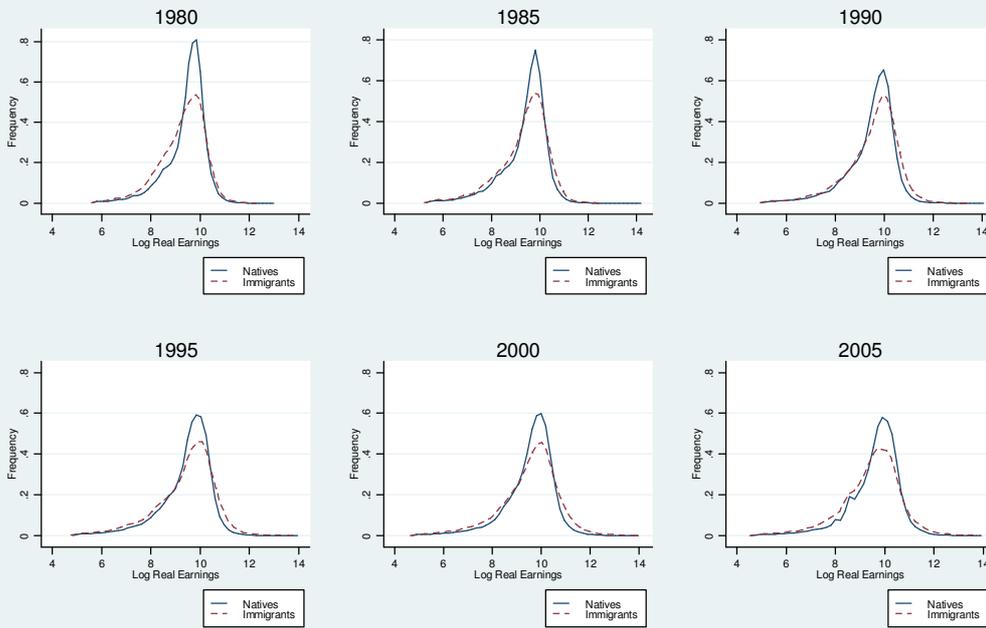
Figure 5 - Immigrants by Continent of Nationality
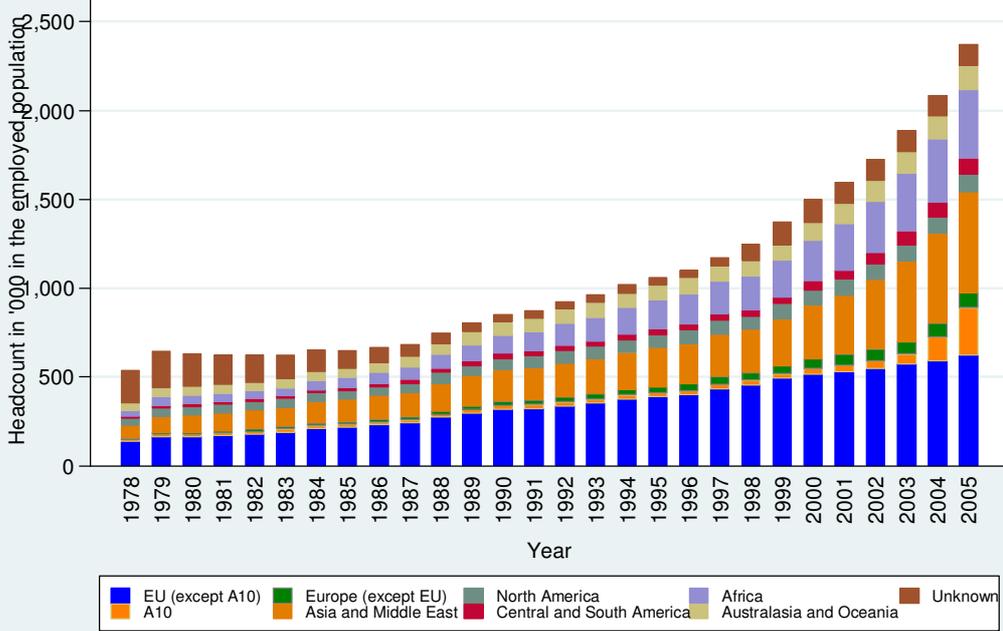
Source: Lifetime Labour Market Database

Figure 6 - Immigrants' Inflow in the Year by Continent of Nationality
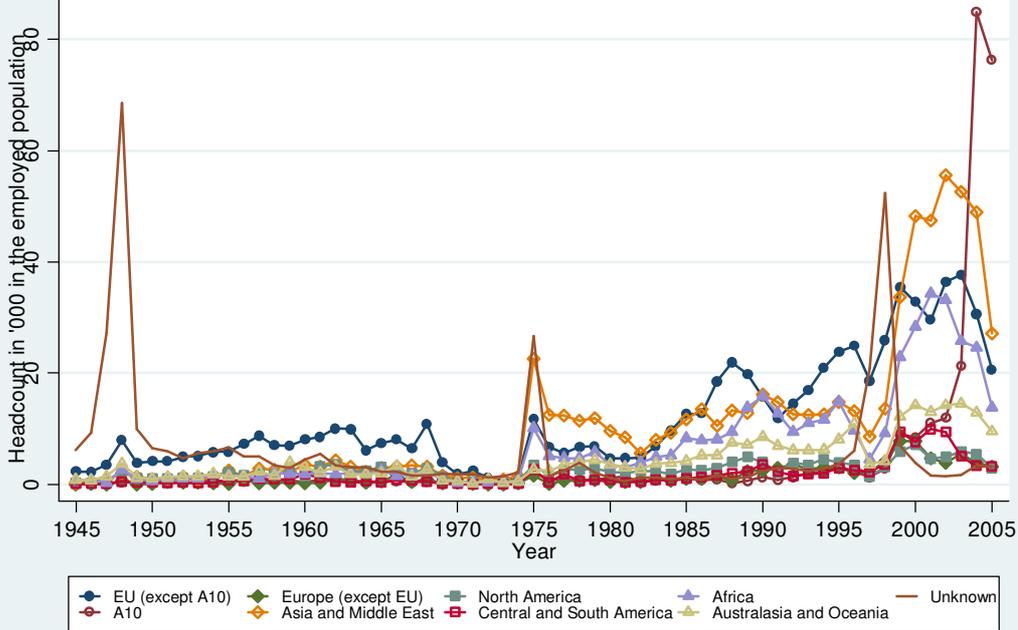
Source: Lifetime Labour Market Database

Figure 7 - Immigrant-Native Earnings Gap by Continent of Nationality
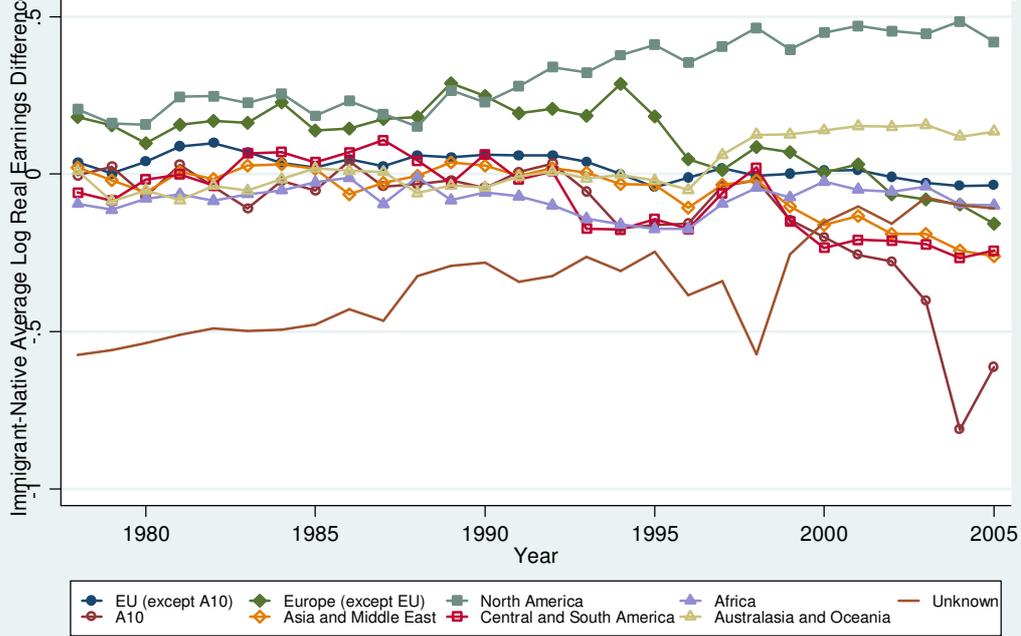
Source: Lifetime Labour Market Database

Figure 8 - Predicted Residuals
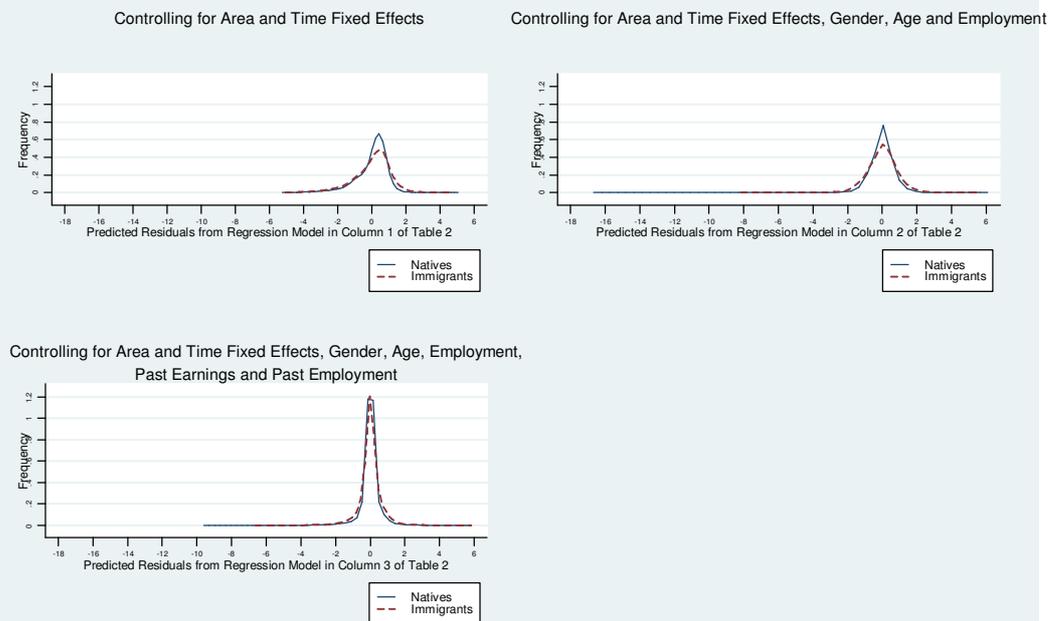
Source: Lifetime Labour Market Database

Figure 9 - Immigrant-Native Earnings Gap Across the Earnings Distribution

Source: Lifetime Labour Market Database
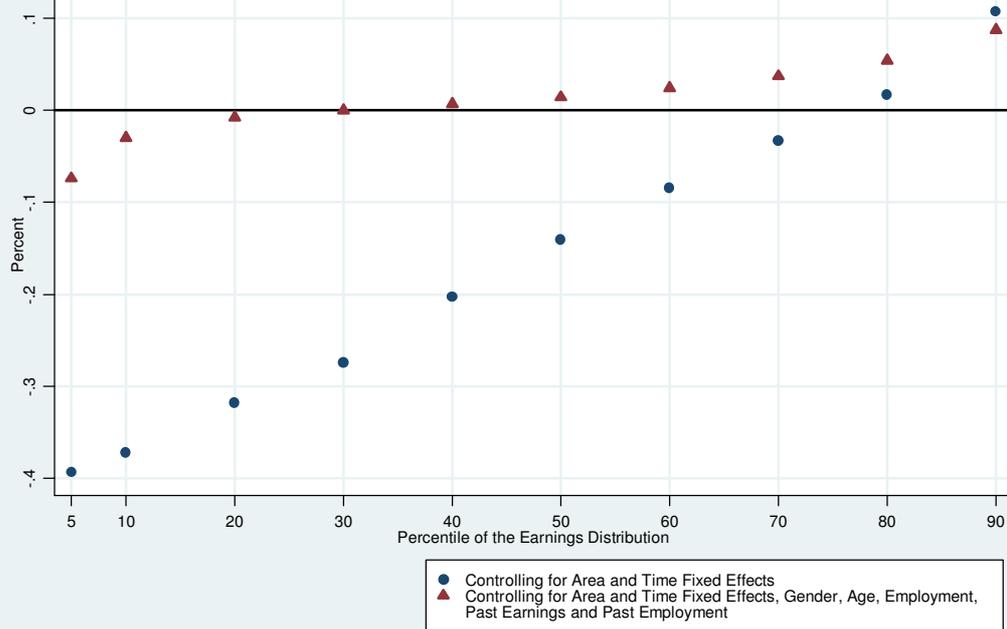
Figure 10 - Immigrant-Native Earnings Gap by Continent of Nationality

Source: Lifetime Labour Market Database
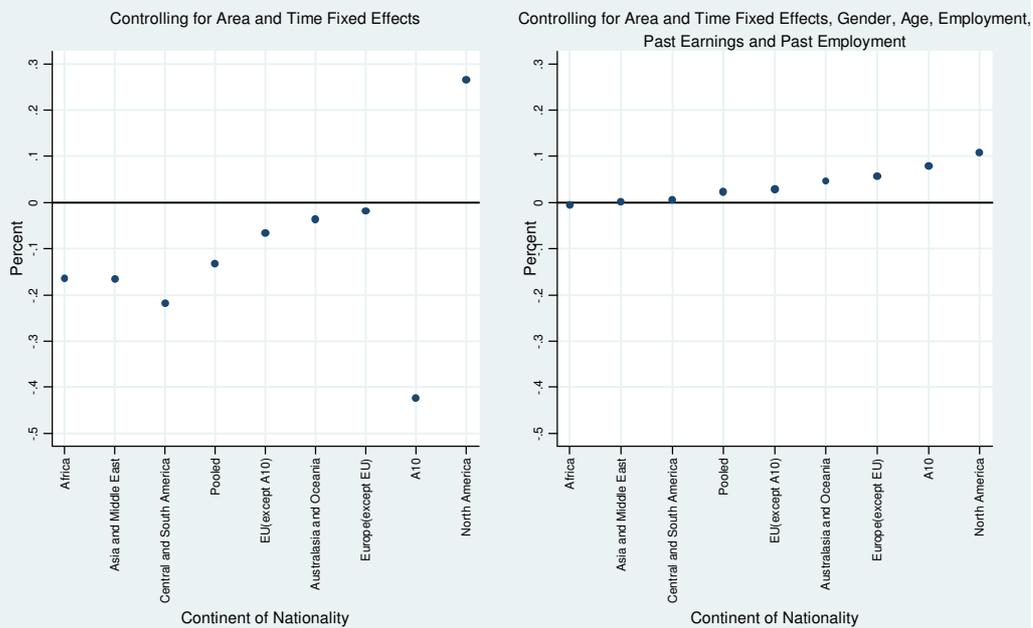
Figure 11 - Immigrant-Native Earnings Gap by Cohort of Arrival
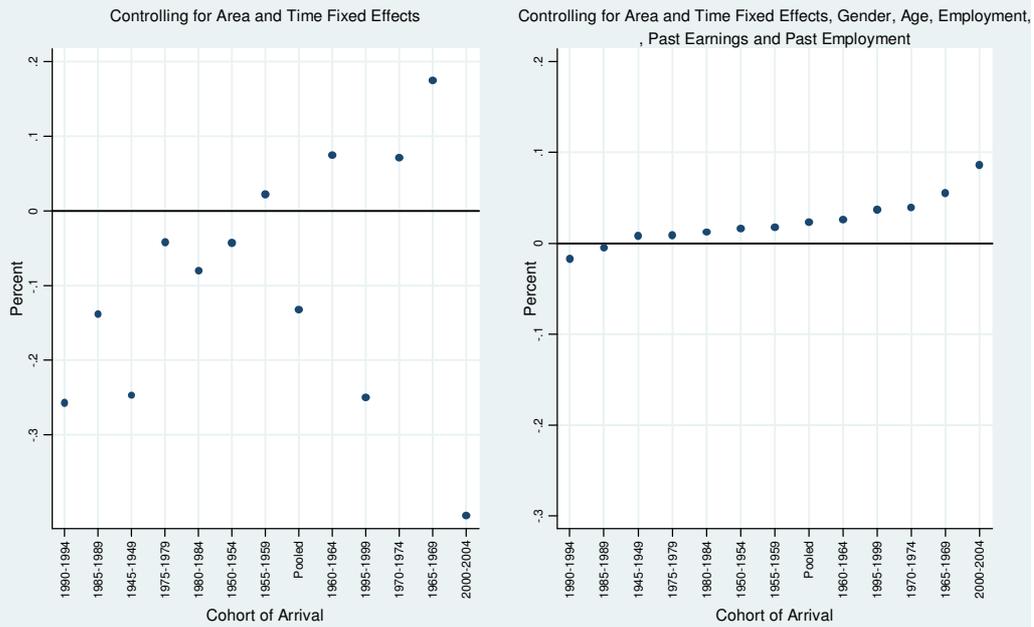
Figure 12 - Immigrant-Native Earnings Gap Across the Earnings Distribution by Gender
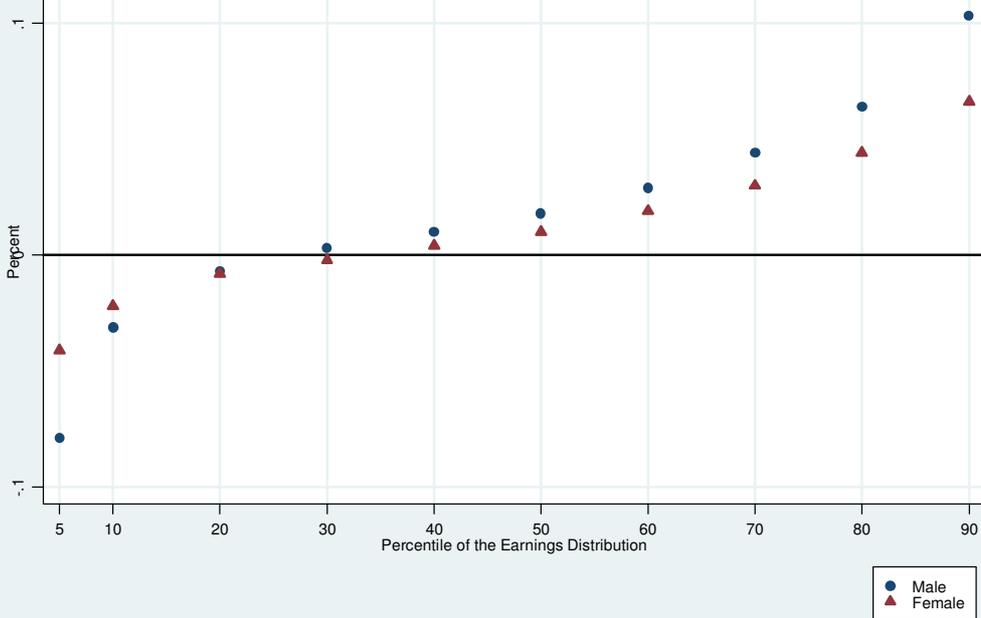
Figure 13 - Immigrant-Native Earnings Gap by Gender

Source: Lifetime Labour Market Database