Evidence on Immigrants' Assimilation into Recipient Labour Markets using UK Longitudinal Data between 1981 and 2006

Luisa Gagliardi[†] and Sara Lemos[‡]

[†]Department of Geography and Environment, London School of Economics and Centre for Regional Economics, Transports and Tourism (CERTeT), Bocconi University
[‡] Department of Economics, University of Leicester

How well do immigrants entering the UK assimilate into recipient labour markets? Using the underexploited, sizeable and long Lifetime Labour Market Database (LLMDB) between 1981 and 2006 we investigate the evolution of the immigrant-native earnings gap – a measure of immigrants' assimilation – across the entire earnings distribution, across cohorts and across nationalities. We are able to control for observable and unobservable individual specific characteristics as well as for specific characteristics of both time periods and recipient labour markets, defined as small geographical areas, and crucially, for the interaction of the two, in a robust empirical model specification anchored in the human capital theory. We also control for cohort specific effects and nationality specific effects. Our results show little evidence of large or persistent earnings disparities across the earnings distribution, across cohorts or across nationalities. These findings are supportive evidence of successful assimilation of immigrants into the UK, suggesting that recipient labour markets primarily reward individuals' characteristics other than, and regardless of, their immigration status. Nevertheless some distinctive features emerge. When investigating the evolution of the immigrant-native earnings gap over time, our results illustrate how immigrants from different continents and cohorts have very different assimilation trajectories.

JEL Codes: F22 J31, J71
Keywords: Immigration, Geographic Labor Mobility, Wages, Assimilation, UK

Acknowledgement: We thank Mike Coombes, Alessandra Faggian, Gianni De Fraja, Steve Gibbons, Steve Hall, Jonathan Portes, Jerzy Szroeter, three anonymous referees and the editor Neil Coe for invaluable discussions, comments and suggestions. We also thank Nicola Gilpin, Andrew Needham and Eric Procter. Finally, we thank staff at the ONS and EDINA. We acknowledge and thank the financial support of the Department for Work and Pensions and the data provided. Views expressed in this paper are not necessarily those of the Department for Work and Pensions or any other Government Department.

1. Introduction

The UK has experienced a substantial increase in immigration inflows in the last two decades (Goodhart 2010), reinvigorating the immigration debate among scholars and policy makers. In this context, it has often been suggested that low skilled immigrants face disadvantages in the labour market (see for example Wadsworth 2003). More recent research, however, has suggested that this evidence is not straightforward and that immigrants may experience different degrees and trends of assimilation not only due to skill level but also due to other individual characteristics and to features of recipient labour markets and time of entry (Rodriguez-Pose and Ketterer 2012). Individuals belonging to specific nationalities might, for instance, fare differently in the labour market. So might individuals who are harder working or more able. Immigrants entering the labour market at different points in time might also fare differently, due to changing economic conditions, changing local attitudes towards immigration and changing cohort specific immigrants' characteristics (e.g. immigrants' skills, immigrants' work ethics, return migration, etc.). And finally, immigrants entering different areas might again fare differently because of specific features of recipient labour markets, including clustering, a well-known phenomenon in the literature.

In this context, the dynamic interaction between individual specific characteristics and specific features of both recipient labour markets and time of entry play a concurrent and substantial role in shaping the impact of immigration, contributing to the emergence of different degrees and trends in immigrants' assimilation (Rodriguez-Pose and Ketterer 2012; Rodriguez-Pose and Vilalta-Bufi 2005; Kanbur and Rapoport 2005; Crozet 2004). In particular, country of origin, nationality and ethnicity exert a relevant role in the assimilation process (Borjas 1995; Ottaviano and Peri 2006; Niebuhr 2010; Nathan 2014; Rodriguez-Pose and von Berlepsch 2014; Suedekum et al. 2014)), as does the actual composition of the immigrant inflow (Longhi et al. 2005 and 2010).

Yet, despite such substantial increase in UK immigration inflows, and despite suggestive evidence of different degrees and trends in immigrants' assimilation, the related literature is limited (see Dustmann et al. 2008 for a review). Furthermore, most of this literature focuses on investigating the impact of immigration inflows, often segmented by skills or ethnicity, on the wages and employment levels of UK recipient labour markets. This spatial correlation approach has three key limitations. Firstly, it says little about immigrants' assimilation as such, since it focuses on the impact of new immigration inflows on the wages and employment levels of existing recipient labour markets. Secondly, it does not account for individual heterogeneity, which is crucial in determining immigrants' assimilation into labour markets. This is potentially due to data limitations, as models accounting for individual heterogeneity aiming at identification of assimilation effects require a large and long longitudinal dataset that tracks individuals over time – and that is very scarce (Chiswick 1980; Borjas 1999; Chiswick et al. 2005).

Thirdly, it does not account for the fact that the rewards to immigrants' characteristics into specific recipient labour markets may have evolved over time, due to both changes in the composition of immigration inflows and their associated cohort-specific characteristics (e.g. nationality, education level, work ethics, etc.), and changes in attitudes towards immigration, which are intertwined with changes in national as well as local economic conditions (Card 2005; Card et al. 2005; Mayda 2006). As a result the spatial approach is limited in its scope to exploit how each of these dimensions may affect immigrants' assimilation – and in special, how the changing role of geography and its interplay with individual specific characteristics, cohorts specific characteristics and recipient labour market specific characteristics at different points in time may affect immigrants' assimilation (Dustmann et al. 2008; Rodriguez-Pose and Tselios 2010). Put differently, although the spatial correlation approach accounts, to some extent, for specific features of both time and recipient labour markets, it does not account for the interaction of the two. Neither does it account for cohort specific characteristics, such as nationality composition and skill composition, among others.[1]

Our main contribution is to address each of these three issues. We estimate a key measure of immigrants' assimilation – the immigrant-native earnings gap – controlling

[1] The UK immigrant-native earnings gap literature is very limited. 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) for males and females. They found that the wage gap for non-white immigrants was as large as -40%, although 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.

for observable and unobservable individual fixed effects, as well as for both time and recipient labour market fixed effects, and crucially, for the interaction of the two in a robust empirical model specification anchored in the human capital theory. We also control for cohort specific effects and nationality specific effects. This way we separately control for the role of each of these dimensions on immigrants' assimilation.

We use data from the Lifetime Labour Market Database (LLMDB), an underexploited, sizeable and long longitudinal dataset that has seldom been used for immigration analysis. It combines anonymised tax and social security records into a dataset that tracks a random sample of over 600,000 individuals between 1981 and 2006, providing a range of geo-referenced data on individual characteristics.

Indeed, our contribution is timely. Accounting for each of these dimensions, using such an underexploited, rich, sizeable and long longitudinal dataset, is paramount in explaining the immigrant-native earnings gap. Despite that, and to the best of our knowledge, no estimation of the gap accounting for all these dimensions, exploiting a long sample period and defining recipient labour markets as small geographical areas consistently over time is available in the literature. Such estimation would be unfeasible utilizing datasets more commonly used in the literature for UK immigration research, such as the General Household Survey and the Labour Force Survey. This is because these datasets neither follow individuals over a long period of time, nor have a large enough sample size to permit fine levels of disaggregation across small geographical areas (below the regional level) consistently over time.

We start by estimating the immigrant-native earnings gap across the entire earnings distribution. The earnings gap is a powerful, informative and direct indicator that attests to the successful integration of new labour resources into recipient labour markets. Estimating the gap not only at the average, but also across the entire distribution, enhances its informative power and provides an insightful investigation of emerging trends for different groups of immigrants that cluster at various points along the distribution. We then estimate the gap controlling for cohort of arrival and for continent of nationality. This provides further insights into how the gap is affected by immigrants' origins as well as by changing economic conditions and attitudes to immigration.

Our results show that the immigrant-native earnings gap substantially narrows down when individual observable and unobservable characteristics as well as time period and recipient labour market characteristics, and their interaction, are controlled for. The individual characteristics dimension seems to be preponderant in explaining most of the gap. These findings support the evidence of successful assimilation of immigrants into the UK, suggesting that recipient labour markets primarily reward individuals' characteristics other than, and regardless of, their immigration status. Nevertheless some distinctive features emerge. Immigrants entering the labour market at the bottom of the earnings distribution tend to have a less favorable assimilation experience. Also, immigrants entering the UK in earlier cohorts, such as in the post war period, experienced faster assimilation, suggesting, possibly, a more positive attitude towards immigration associated with the role of immigrants in the post war reconstruction effort. Earlier cohorts, such as the post war cohorts, not only fare better than more recent ones at entry, but also the earnings of immigrants in such cohorts catch up faster with natives' earnings. Similarly, North Americans, Europeans and Australians fare better at entry and their earnings catch up faster with natives' earnings. More generally, our results when investigating the evolution of the immigrant-native earnings gap over time illustrate how immigrants from different continents and cohorts have very different assimilation trajectories.

2. Data and Descriptive Statistics

We use data from the Lifetime Labour Market Database (LLMDB). The LLMDB is derived from a number of administrative datasets linked together through the National Insurance Number (NINo). Whereas natives are automatically given a NINo, immigrants typically apply for one 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 referred to as UK and overseas nationals.) Because the NINo is a unique individual identifier, the LLMDB tracks individuals over their entire working lifetime.

The LLMDB is a long, sizeable and rich longitudinal dataset. It comprises over 600,000 individuals (a 1% random sample of NINo records) followed between the tax-years 1981 and 2006. A fresh cohort of individuals enters the sample every year and is followed from then on. We restricted our sample to males aged 25 to 64 and females aged 25 to 59, as customary in the earnings gap literature, and to those with earnings between £100 and £1000000 in any one tax-year (which run from April to March). The self-employed, for whom we do not observe earnings, are excluded from the sample.[2] 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. We further restricted our sample to those observed at least twice in order to control for individual fixed effects (see Section 3). Finally, we restricted our sample to those whose address in each time period is observed in order to control for area fixed effects.[3] Our final working sample therefore consists of 354,465 individuals, 38,074 of whom are immigrants, as shown in Table 1.

The LLMDB contains date of birth, date of death, age, gender, address, nationality, country of origin (country of arrival immediately prior to NINo registration), immigrants' entry date, 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 and unemployed in the year, spells of unemployment, spells of receipt of benefits, benefit type, pension contributions, pension entitlements, etc. However as is common in the case of administrative records, no information on educational attainment is provided. We circumvent this limitation to some extent both by restricting our sample to individuals in work aged at least 25 and by controlling for individual fixed effects (see Section 3).[4]

[2] In our sample, 6.36% of immigrants (9.33% of natives) are observed as being self-employed at least part of a tax-year during their working life in the UK. This proportion is relatively stable over the assimilation process, peaking half way through: it is respectively 6.83%, 6%, 12.26%, and 8.13% after 1, 5, 10 and 20 "years since immigration". However, the proportion of observations for immigrants working at least part of a tax-year as a self-employed is much lower, 2.15% (2.42% for natives), indicating that very few immigrants resort entirely to self-employment throughout, and instead, are likely to have spells of self-employment. Although we drop observations for such spells, when earnings are not observed, we retain observations outside these spells, when individuals work as employees. This means that we lose just over 2% of our observations.

[3] Our results were robust to interpolating missing addresses if the address in both the previous and subsequent tax-year remained the same, which boosted our sample size. Given such robustness, however, we report the results without interpolation. Note that the LLMDB2 geographical distribution with our without interpolation remains remarkably similar to the one in the LFS (see Table 1 and also see below).

[4] By restricting our sample to those in work aged at least 25, who, we assume, have completed their education, we are assuming that education no longer varies over time and is just one more characteristic specific to the individual that we do not observe, such as race or ability. The standard literature accounts for such unobserved time invariant individual specific characteristics by controlling for individual fixed effects. The standard argument is that, although we cannot identify the specific effect of say, race or ability, on wages, their effect is controlled for by the fixed effects in a manner that does not bias other coefficients.

As the LLMDB records information on address, it provides a range of geo-referenced data on individuals, which we then track across the whole sample period. Immigrants display a significant degree of geographic concentration in specific areas (see Table 1). As expected, immigrants cluster in London and in the South East, emphasizing the role of labour market characteristics and multicultural environments in attracting immigrants – a well-known phenomenon in the literature. Note that the geographical distribution of natives and immigrants is remarkably similar in the LLMDB and the LFS, which is the dataset most widely used for UK immigration research (see Dustmann and Fabbri 2005). This geographical distribution pattern is fairly persistent over time; if anything, there is a slight upwards trend in the proportion of immigrants in London. This evidence reinforces the importance of accounting for recipient labour market characteristics when estimating the immigrant-native earnings gap (see Section 3).

Table 1 shows that natives are more evenly spread across the country, are older than immigrants, earn more on average, are more likely to be employed and slightly less likely to be unemployed. Figure 1 shows the immigrant-native 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.

Interestingly, substantial heterogeneity emerges when we consider the gap across the earnings distribution. Table 1 shows that immigrants at the very bottom of the earnings distribution earn less than natives whereas those at the very top earn more. This is confirmed in Figure 1. Whilst immigrants at the bottom of the distribution can earn less than a half of what their native counterparts earn, those at the top can earn up to a quarter

---

By assuming that education is a time invariant (fixed) characteristic, we extend this standard argument to education. Put differently, there is nothing particular about time-invariant completed education that makes it any less eligible than say, race or ability, to modelling via fixed effects. The main point is that we are not studying the effect of education on earnings, we are only accounting for it to prevent bias in other coefficients in our model (see Section 3). Indeed, our results (see Section 4) are qualitatively similar to other results in the literature where education was controlled for (Dustmann et al. 2013). Incidentally, even when education is observed, the decision to include it in an immigrant-native earnings gap model such as ours is not straightforward. Although earnings models commonly include education, there is an unresolved debate in the immigration literature about the interpretation of other coefficients in the model 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 the extent and quality of education varies across countries. Immigrants and natives with the same education level may have different skills and compete for different jobs. For example, there is evidence that natives and immigrants are imperfect substitutes within education groups in the UK (Manacorda et al. 2007). Also, immigrants across the education spectrum often suffer skill downgrading due to language or other labour market barriers (Card and DiNardo 2000; Friedberg 2001).

more. 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, before it slopes down towards the end of the sample period.

Table 1 shows that immigrants predominantly come from the European Union (EU), Asia and the Middle East and Africa. The composition of the inflows has changed over time with a large share of EU immigrants (mainly Irish) and immigrants from former colonies (India, Pakistan, Bangladeshi, South Africa, Nigeria etc.) being disproportionally represented during the 1950s and 1960s. In the 1970s there was an increase in EU immigration after the UK joined the Union. During the 1980s and 1990s there was a steady increase in inflows of immigrants from the EU, mainly due to the accession of Greece, Spain and Portugal and an increase in the number of individuals coming from North America, Australasia and Oceania. In this period, immigration, mainly from Africa, Asia and the Middle East, also increased. Finally there was a large inflow of EU immigrants in concomitance with 10 Eastern European countries (A10) joining the EU in the 2000s.[5] Figure 2 shows the immigrant-native earnings gap across tax-years by continent of nationality. This 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. This evidence reinforces the importance of accounting for nationality and cohort characteristics when estimating the earnings gap (see Section 3).

Note that 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

[5] The EU was successively enlarged at various points during our sample period, so for consistency we use the 2006 membership. We separately define the A10 countries, as is common in the literature, which are: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovenia and Slovakia, Malta and Cyprus.

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 (50th) percentile of the distribution. However, although earnings are consistently lower in the LLMDB, the average earnings trend over time is similar.[6] Comparisons between the LLMDB and ASHE also show that annual earnings are lower in the LLMDB but that the trend of average earnings across both datasets is similar over time (Dickens and McKnight 2008).

3. Estimation Strategy

Our descriptive statistics in Section 2 provide evidence of an unconditional immigrant-native earnings gap in the UK between 1981 and 2006. This gap is quite sizeable for some groups of immigrants, though it varies greatly across nationalities and across the earnings distribution. Our descriptive statistics also provide evidence of a distinctive pattern in the geographical distribution of immigrants. More importantly, our descriptive statistics provide suggestive evidence that some of the variability in the gap might be explained by the dynamic interaction of immigrants' geographical distribution, cohort of arrival, continent of nationality and position in the earnings distribution. As we argue in the Introduction, the interaction of each of these dimensions plays a role in explaining the gap. For example, whereas many of the highly skilled North Americans that arrived during the 1990s and 2000s ended up at the top of the earnings distribution and experience a more favourable gap, many of the Eastern Europeans that arrived in the 2000s ended up at the bottom of the distribution and experience a less favourable gap (see Figure 2).

We now account for each of these dimensions by estimating the conditional immigrant-native earnings gap using a standard human capital model (see for example

[6] On the one hand, although a small number of workers who earn too little to incur a national insurance contribution liability (those working part-time for very small employers), are not included in the LLMDB, non-liable employees of medium and large employers are included. On the other hand, the LFS does not cover communal establishments or individuals living in the UK for less than 6 months, which include many low paid immigrants. As a result, the LLMDB better captures low paid immigrant workers, who tend to be younger. Therefore, although both datasets exhibit broadly similar patterns overall (see period 1997-2007 in Table 1 for this comparison), and although the age distribution is remarkably similar for natives, a larger proportion of immigrants is younger in the LLMDB (also note that we tabulate observations, not individuals). More broadly, note that even though the LLMDB is a 1% random sample of a large, long, and accurate administrative dataset, it is nevertheless, a sample, and thus prone to measurement error.

Chiswick 1980; Dustmann and Fabbri 2005). In the human capital model, individuals' earnings are a function of characteristics that influence individuals' productivity:

$$E_{iat} = \alpha + \beta I_i + \lambda X_{iat} + f_i + f_a + f_t + f_{at} + \varepsilon_{iat} \qquad (1)$$

where E_{iat} is log real earnings of individual $i = 1,...,354465$ in area $a = 1,...,49$ and time $t = 1981,...,2006$; I_i is an indicator variable that is 1 if the individual is an immigrant; X_{iat} is a vector of observable individual characteristics including sex, age, age squared, number of employed weeks in the year and number of jobs in the year; f_i is individual specific effects; f_a is area specific effects; f_t is time specific effects; f_{at} is the interaction of area and time fixed effects; and ε_{iat} is the error term.

The ideal estimation approach here is to use the fixed effects model. Such approach is, however, unfeasible in the presence of a large number of parameters to be estimated.[7] So it is first-difference transformation – a common variation of the fixed effects model when there are so many parameters – in the presence of dummy variables such as I_i, which is, incidentally, our variable of interest. An alternative variation of the fixed effects model is Nakamura and Nakamura's (1985) inertia model, later re-worked by Chiswick et al. (2005), where individual specific effects are modelled as a function of lagged log real earnings, lagged number of employed weeks in the year X^W_{iat}, and an error term ξ_{iat}. Thus, instead of using individual dummies to model f_i, we parameterize it as follows:

$$f_i = aE_{iat-1} + bX^W_{iat-1} + \xi_{iat} \qquad (2)$$

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, ability, etc. This is because these lagged variables capture individual specific time invariant characteristics that have the same impact on earnings year after year. That is, these lagged variables account for characteristics and

[7] The random effects model, ideal in the presence of such a large number of parameters, is also unfeasible here because it is unrealistic to assume that individual, area and time fixed effects are independent of one another. For example, as discussed in Section 2, individuals of particular nationalities cluster in specific geographical areas. In the fixed effects model instead, the components f_i, f_a, f_t and f_{at} are fixed parameters to be estimated.

circumstances, specific to individuals, that affect earnings year after year, and these individual fixed effects are captured by a and b.

In sum, we control for unobservable individual fixed effects via lagged log real earnings and lagged number of employed weeks. Controlling for individual fixed effects enables us to separately account for the effect of individual specific time invariant characteristics and circumstances on earnings. This way, we account for earnings differentials due, for example, to workers who are more motivated or who suffer more discrimination. Furthermore, by controlling for the lagged number of employed weeks in the year, we account for lower earnings for individuals with historically long spells of unemployment. Finally, by including these two lagged variables we also account for the effect of dynamics in the model and alleviate problems arising from serial correlation in the residuals. Controlling for individual fixed effects using a sufficiently large and long longitudinal dataset, such as the LLMDB, is an important improvement on the existing UK immigrant-native earnings gap literature.

We model area fixed effects using county dummies. This way, we remove any permanent differences across counties and make them equally attractive to immigrants and natives. In other words, we control for specific factors in a county (such as more schools, more housing, lower prices, multiculturalism, 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. Note that most available models in the immigrant-native earnings gap literature do not control for area fixed effects, except Dustmann and Fabbri (2005), where region fixed effects are included. Here, we model area fixed effects using 49 counties instead of 12 regions, which is a more flexible approach (see Section 2). (We further relax this assumption in Section 4d and model area fixed effects using four other geographies, including Local Authorities (LAs) and Travel To Work Areas (TTWAs), and find remarkably robust results).

We model time fixed effects using tax-year dummies. 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 on the existing UK earnings gap literature.

Finally, we 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 control for age, which, albeit imperfectly, captures overall experience to a certain extent.[8]

This is a robust empirical specification that significantly improves over available specifications in the existing immigrant-native earnings gap literature. By controlling for individual observable and unobservable specific characteristics, area specific characteristics, time specific characteristics, and the interaction of the last two, we largely prevent certain common selectivity biases. Firstly, individual heterogeneity may introduce various types of biases in the model, such as ability bias, sorting bias, survivor bias, etc. (the long sample period and low levels of attrition in our data are particularly important to prevent survivor bias). Secondly, different cohorts of arrival can introduce bias such as cohort bias and return migration bias.[9] Thirdly, different levels of attractiveness across recipient labour markets can further introduce bias such as simultaneity bias (i.e. immigrants are more attracted to high wage and low unemployment

[8] Our results were robust to controlling for "age at entry", which captures the human capital endowment at arrival and identifies immigrants who arrived as children and, therefore, have labour market characteristics of natives (Borjas 1999). In our data, most immigrants arrived as adults. Around 10% entered as children, most of who arrived before 1969 from ex-colonies and from Europe.

[9] Return migration bias is a concern if large-scale, non-circular, non-random outflows are extensive. In this case, immigrant-native earnings gap estimates are biased in favor of immigrants and might indicate more of a weeding out process than an assimilation process (Chiswick and Hatton 2003; Dustmann 2003; LaLonde and Topel 1997). Firstly, while return migration of the less able or less motivated is clearly non-random, other reasons for return migration include maximizing consumption or returns to human capital acquired in host countries (Dustmann 1994). Some studies suggest that non-random outflows in recent waves of immigration to the UK are not too severe, mitigating any potential return migration bias (Gilpin et al. 2006; Lemos and Portes 2013). Secondly, although the scale of return migration is roughly between 30% and 50% after 5 to 50 years (LaLonde and Topel 1997; Dustman and Weiss 2007) (tracking a sample of immigrants in our own data shows that around 30% return after 10 years), circular migration can be around 60% (Constant and Zimmermann 2007). Unlike most data in the immigration literature, our data picks up circular-migration (see Section 2), mitigating further any potential return migration bias. Finally, Table 1 shows that our own sample has a relatively even distribution of immigrants across "years since immigration" and is not severely biased towards those who remain after any weeding out process.

areas). Fourthly, specific macroeconomic conditions as well as attitudes to immigration vary over time, and, if uncontrolled, can also introduce simultaneity bias in the model.

Finally, we 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). The interpretation of our coefficient of interest is that immigrants on average earn $\beta\%$ more than natives.[10]

4. Results

Column 1 of Table 2 shows a negative and significant estimate for the immigrant-native earnings gap, controlling for area fixed effects, time fixed effects and the interaction of the two. This estimate suggests that immigrants earn 10.7% less than natives on average. However the poor explanatory power of this base specification indicates that the impact of important dimensions might have been unaccounted for. Indeed, once we control for observable individual characteristics through sex, age, age squared, number of employed weeks in the year and number of jobs per year as well as for unobservable individual fixed effects through lagged log real earnings and lagged number of employed weeks, the earnings gap estimate narrows down substantially. Column 2 shows that the earnings gap is now positive and significant. This estimate suggests that immigrants earn 2.3% more than natives on average. The explanatory power of this specification is now high, confirming the importance of controlling for individual characteristics. Indeed, individual heterogeneity is the major determinant of the gap, since we had already controlled for immigrants clustering in specific areas in our base specification. This is a key finding, as the majority of previous studies have not controlled for individual characteristics to the extent that we have, mainly due to data limitations.

The estimates of all other controls are significant and robust here as well as in the remaining models in the paper. Women earn 13.5% less than men on average. An extra year of experience (proxied by age) in the labour market increases earnings by 1.8% on average (we interpret the estimate of the squared term as zero, though it is a significant

[10] 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 the ones deriving from our preferred 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).

-0.00023, which confirms the usual inverted U shape relation between age and earnings found in the literature; this small estimate needs to be interpreted in light of the fact that our sample does not include retirement age individuals, for whom earnings decline faster). An extra week employed in the year increases earnings by 3.2% on average. Holding a second job has a very marginal effect, increasing earnings by 0.1% on average. Increasing past earnings has a positive effect on current earnings: a 1% increase in earnings the year before increases current earnings by 0.726%.

The estimate of the lagged number of employed weeks is negative. The corresponding past employment (lagged hours worked per week) estimate in Chiswick et al. (2005) is also negative. This variable is capturing the effect on earnings of unobserved individual specific characteristics and circumstances, such as motivation, race, ability, etc., via associated dimensions of the labour contract, such as part-time or full-time work, high or low turnover jobs, overtime work, etc., that in turn reflect labour market conditions, including labour competition, discrimination, market imperfect information, labour force composition, productivity shocks, demand and supply shocks, etc. The wide range of omitted (unobserved) time invariant characteristics and circumstances at the individual level captured by this variable exert different and opposite effects on earnings and that is why this variable cannot be interpreted directly[11] – in the same way that the estimate of other fixed effect dummies also cannot be interpreted directly (see Section 3). (Note that this variable is negative and significant across all percentiles in Table 3, rising monotonically, confirming that it is systematically capturing unobserved individual fixed effects across the earnings distribution, which is reassuring).

Our main result here is that immigrants do not suffer an earnings penalty and are successfully assimilated into the UK labour market between 1981 and 2006. Our preferred estimate suggests that immigrants earn 2.3% more than natives on average. This suggests that the labour market primarily rewards (observable and unobservable)

[11] More technically, the associated -0.22 estimate cannot be directly interpreted as the effect of an extra lagged employed week on earnings, since an extra lagged employed week also increases lagged real earnings, whose estimate is 0.726. Thus the effect of an extra lagged employed week on earnings will be in fact -2.2% plus a positive fraction of 7.26% (which depends on a model of lagged log real earnings as a function of the lagged number of employed weeks). (Note that the raw correlation between log real earnings and both current and lagged number of employed weeks is positive.) However, as discussed above, we can interpret the effect of lagged log real earnings directly on current earnings.

individual characteristics other than immigration status. This, in turn, facilitates the assimilation of immigrants into the UK labour market.

The average immigrant-native earnings gap estimate might, however, conceal distinctive patterns of immigrants' assimilation across the earnings distribution. There is, for example, wide consensus that unskilled immigrants do not compete with skilled natives and that any detrimental effect on wages is likely to be at the lowest tier of the distribution (Dustmann et al. 2008 and 2013; Gagliardi 2014). To account for this, we re-estimate our model using quantile regression estimation. This way we uncover potentially larger or smaller earnings gaps along the distribution that might have been concealed by the average gap. This is a particularly appealing approach where immigrants concentrate at the bottom and top of the earnings distribution, as is the case for the UK over our sample period (see Section 2). Estimating the earnings gap in such a flexible yet robust manner across the earnings distribution is an improvement on the existing UK earnings gap literature, where only estimates of the average gap are available.

Figure 3 and Table 3 show that the immigrant-native earnings gap narrows down substantially in our preferred specification, as before. It is non-negative, except for those below the 30th percentile, and it increases monotonically across the distribution. The gap is respectively -0.016, 0.002, 0.012, 0.033 and 0.089 for the 10th, 30th, 50th, 70th and 90th percentiles. That is, among the 10% worst paid workers, immigrants earn 1.6% less than natives; whereas among the 10% best paid workers, immigrants earn 8.9% more. Given that immigration to the UK has been of predominantly unskilled or highly skilled labour, it is unsurprising that the gap is larger at the bottom and top of the distribution.

Thus, on the one hand, the lowest paid immigrants suffer an earnings penalty in relation to the lowest paid natives with comparable individual characteristics. This suggests that, for this group, assimilation has been slower. On the other hand, other immigrants do not seem to suffer an earnings penalty and seem to have been well assimilated into the labour market – the gap is fairly small in the middle of the distribution and is in favour of higher paid immigrants at the top.

According to standard human capital theory, non-negligible gap estimates could be due to unaccounted for productivity differentials. Our model is quiet comprehensive. For example, it seems to have captured most such productivity differentials in the middle of the distribution, where the gap is fairly small. Dustmann et al. (2005) argue that the immigrants' skill distribution resembles that of natives, which suggests that such productivity differentials might not be very large in the UK. Nonetheless, our models might not have fully captured productivity differentials for some groups of workers, for example, those at the top and bottom of the distribution. In addition to supply side productivity differentials, possible demand side explanations are that non-negligible gap estimates are due to imperfect information, friction, discrimination or market power of individual firms, again, perhaps not fully captured in our models.

In sum, our main result here is that the immigrant-native earnings gap in the UK between 1981 and 2006 varies substantially across the earnings distribution, increasing monotonically, and this variability is concealed when solely the average gap is considered. Nevertheless, the gap is still relatively small, at most ±3.5% for almost the entire distribution (except at the very top). Although no estimates of the gap across the distribution for the UK are available, our results are in line with the international literature, which shows that the immigrant-native earnings gap is also more favourable higher up the distribution for the US (Butcher and DiNardo 2002; Chiswick et al. 2008).

a. Excluding London

It is customary in the UK literature, especially the strands concerned with introducing more geography into the economic analysis, to perform robustness checks excluding London from the model (Dustman et al. 2003; Gagliardi 2014). This is because, as discussed in Section 2, immigrants display a significant degree of geographic concentration, heavily clustering in London (see Table 1). London is atypical, as it attracts large shares of both highly skilled and unskilled immigrants, and whether the immigrant-native earnings gap is larger or smaller in London than in the rest of the country is an empirical matter that hinges (a) on the relative magnitudes of such shares in and outside London and (b) on their associated level of earnings in and outside London.

Table 4 shows that the pattern of estimates' significance and magnitude across percentiles is the same whether London is excluded or not (compare with Table 3). As expected, for the lowest paid, the gap is now smaller (in absolute terms), less adverse. Put differently, the lowest paid immigrants outside London still earn less than their native counterparts, but not by quite so much. This means that the lowest paid immigrants outside London are better off, possibly because ethnic minorities, for whom the gap is often less favourable (see Sections 4b and 4e), are overrepresented in London (Nathan 2014); and possibly because competition from newly arrived immigrants, whom often enter the labour market as low paid workers (see Section 2), is more fierce in London.

Conversely, for the highest paid, the immigrant-native earnings gap is larger, more favourable, when excluding London. Put differently, the highest paid immigrants outside London now earn even more than their native counterparts (compare Tables 3 and 4). This means that the highest paid immigrants outside London are again better off, possibly because the highest paid natives in London are more educated than in the rest of the UK and command a higher skill premium, which shrinks the gap.

Our main result from before is thus maintained: the immigrant-native earnings gap in the UK between 1981 and 2006 again varies substantially across the earnings distribution, increasing monotonically, when excluding London. Although, again, no comparable estimates are available in existing studies, our results relate to a growing literature documenting lower wage inequality in large urban areas (Black et al. 2009, Lee 2010, Moretti 2013).

b. The role of "diversity"

Diversity, defined in terms of nationality or ethnic group, is an important source of heterogeneity when studying immigrants' assimilation, as discussed in the Introduction. Starting with the work of Chiswick (1978 and 1980) and Borjas (1985), through to the work of Ottaviano and Peri (2006), and beyond, ethnicity has been shown to be a significant dimension of immigrants' assimilation. More recently, Rodriguez Pose and von Berlepsch (2014) showed that nationality also drives immigrants' degree of assimilation, and in turn, their contribution to economic development.

Our sample of immigrants is characterised by substantial heterogeneity in terms of continent of nationality (i.e. the continent where the country of nationality is located), allowing us to exploit the role of this dimension when estimating the immigrant-native earnings gap. Although we implicitly account for continent of nationality to some extent when we control for unobserved individual characteristics in Section 4 above, we now re-estimate our model 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). It is also a way of gaining further insight into the immigrant-native earnings gap. As discussed in Section 2, the various immigration waves to the UK between 1981 and 2006 happened in such a manner that it is possible that immigrants from particular nationalities broadly cluster in specific segments of the earnings distribution. Estimating the earnings gap in such a flexible yet robust 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 4 and Table 5 show the immigrant-native earnings gap for our base (left panel) and preferred (right panel) specifications. The gap is positive and significant for most nationalities, although it is insignificant for immigrants from Africa, Central and South America and negative (-1.5%) for immigrants from Asia and the Middle East. In contrast, the gap is 2.4% for immigrants from the A10. The gap then ranges from 3.4% for immigrants from the EU to 11.8% for immigrants from North America.

Non-negative earnings gap estimates for most continents of nationality suggest that, on the whole, immigrants do not suffer an earnings penalty and are well assimilated into the UK labour market. Furthermore, the gap estimates vary across continents of nationality, and this suggests that some nationalities, such as North Americans, fare better in the labour market. Our analysis suggests that continent of nationality is a significant source of heterogeneity in explaining the earnings gap.

Our results are in line with previous research, which reports a gap estimate for non-whites between -40% and -10% (Chiswick 1980, Bell 1997, Dustmann et al. 2005). In fact, our base specification, which is closer in nature to those in this previous research, yields estimates between -26.3% to -19.1% for Africa, Asia and the Middle East, Central and South America. This range narrows down substantially in our preferred specification: -1.5% to 0%.

We can also compare continent of nationality estimates with estimates across the earnings distribution. For example, immigrants from Africa, Asia and the Middle East, Central and South America (gap between -1.5% and 0%) are overrepresented at the bottom of the distribution (gap between -3.4% and -0.3%). In contrast, immigrants from North America, Europe and EU, Australasia and Oceania (gap between 2.4% and 11.8%) are overrepresented at the top of the distribution (gap between 3.3% and 8.9%). This confirms that nationalities tend to cluster in segments of the distribution (see Section 2).

Our main result here is that the immigrant-native earnings gap in the UK between 1981 and 2006 varies across continents of nationality. Immigrants from Africa, Asia and the Middle East, Central and South America, in the main, do not seem to suffer much of an earnings penalty in the labour market as a result of their immigrant status. This suggests that this group is assimilated into the labour market. In contrast, immigrants from North America, Europe and EU, Australasia and Oceania experience a favourable gap, again suggesting that this group is assimilated – with perhaps a more auspicious assimilation experience. It is, however, worth noting the literature that suggests very different fortunes for immigrants in the UK coming from different parts within a continent. This suggests that assimilation varies within as well as across continents of nationality, and so the results here should be viewed with caution. For example, Dustman and Fabbri (2003) find that Indians, Afro-Asians and Chinese have higher employment probabilities once in the UK than Pakistanis and Bangladeshis. Bell (1997) shows that entry wages in the UK are higher for Indians than for West Indians; while Shields and Wheatley Price (2002) suggest that African-Asians perform better in the UK labour market than Indians, Pakistanis and Bangladeshis.

c. Heterogeneity across cohorts of arrival

Cohort of arrival is also an important source of heterogeneity when studying immigrants' assimilation, as discussed in the Introduction. Starting with the work of pioneers such as Chiswick (1978 and 1980) and Borjas (1985), immigrants' time of entry into the host country has been shown to be a significant dimension of immigrants' assimilation. Immigrants entering the labour market at different points in time fare differently because of changing economic conditions, changing local attitudes towards immigration and changing cohort specific immigrants' characteristics. We largely account for local and national macroeconomic conditions in the UK affecting earnings, and to some extent, for changes in attitudes to immigration, when we control for area and time fixed effects and their interaction. However, it is still possible that the earnings gap estimates are affected by immigrants' characteristics specific to their cohort of arrival.

We thus exploit the long sample period in our data and re-estimate our model including an explicit indicator for 13 five-year cohorts of arrival as a proxy for such "cohort" individual characteristics. This way, we account for distinctive features that vary across cohorts such as unmeasured dimensions of immigrants' skills or return migration of immigrants that are more or less able. This also allows us to account to some extent for changes in attitudes towards immigration over time, which might affect immigrants' geographical distribution. Finally, this allows us to account for the nationality composition of each cohort, which might be a driving factor affecting immigrants' geographical distribution across local labour markets. Estimating the earnings gap in such a flexible yet robust manner across 13 cohorts of arrival is a contribution to the existing UK earnings gap literature, where such estimates are as yet unavailable. 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 further insight into the immigrant-native earnings gap. As discussed in Section 2, the various immigration waves to the UK between 1981 and 2006 happened in such a manner that it is possible to associate immigrants from particular nationalities with particular cohorts.

Figure 5 and Table 6 show the immigrant-native earnings gap for our base (left panel) and preferred (right panel) specifications. The gap is positive and significant for most cohorts, although it is insignificant for immigrants arriving in 1985-1989 and slightly negative for immigrants arriving in 1990-1994. Interestingly, these are cohorts that witnessed greater immigration of lower paid EU workers, following the accession of Greece, Portugal and Spain. In contrast, the gap is between 2.5% and 3.6% for immigrants arriving in 1995-2004. These cohorts received a mix of A10 workers, higher paid workers from the EU and North America, and lower paid workers from Africa, Asia and the Middle East. The gap is between 3.8% and 6.1% in the 1950s and 1960s, when workers mainly came from Ireland and former colonies. A number of distinctive features of these immigrant inflows, such as knowledge of the English language and British work ethics, may contribute to explaining their faster assimilation. In addition, immigration in the post war, aimed at the reconstruction effort, might have attracted a more positive attitude towards immigrants in recipient labour markets.[12]

An attempt to compare cohort of arrival estimates with our earlier continent of nationality estimates indicate that, broadly speaking, cohorts with a greater share of immigrants from North America, Europe and EU, Australasia and Oceania performed better. However, although this comparison is a worthwhile exercise for very broad patterns, it is only suggestive evidence, since in the majority of cases each cohort is characterized by a mix of immigrants from several continents of nationality. The analysis is further confounded when we differentiate between lower paid and higher paid immigrants, which might also affect the direction and magnitude of the gap.

Non-negative earnings gap estimates for most cohorts of arrival suggest that, on the whole, immigrants do not suffer an earnings penalty and are well assimilated into the UK labour market. Furthermore, the gap estimates vary across cohorts of arrival, and this suggests that immigrants arriving in particular cohorts, such as during the post war, fare better in the labour market. Our analysis suggests that cohort of arrival is a significant source of heterogeneity in explaining the earnings gap.

Although no directly comparable estimates of the immigrant-native 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 (Borjas 1999).

d. Area fixed effects

We modelled area fixed effects above using counties as geographical units to proxy local labour markets. We now address two concerns about the choice of such geographical units, which, in turn, enable us to purposely exploit the geography of the UK to further check the robustness of our results. Firstly, the UK geography over the period we study is not straighforward. Even when we use administrative geographies, such as counties, instead of more dynamic geographies (see Annex), we run into difficulties. This is largely because local government in the UK has been the subject of a

[12] The gap estimate is a large 0.142 for the 2005 cohort, which should be viewed with caution. This cohort is four years short and has only one observation per individual (after calculating lagged log real earnings).

constant re-structuring process. As we discuss in detail in the Annex, different definitions of counties are used in the UK literature. We summarize such main definitions in Table A1. As our sample period spans the 1980s through to the 2000s, we are restricted to using the definition prior to the Local Government Reorganization (pre-LGR) (see Table A1 and Annex), which amounts to 66 counties for Great Britain (as opposed to the 49 counties for the UK we used above). Table 7 shows that our results are remarkably robust to this alternative definition, despite a smaller sample that now excludes Northern Ireland (see Annex). Our immigrant-native earnings gap estimate remains unchanged at 2.3%.

Secondly, ideally, geographical units should conform to the actual radius of local labour markets, where both natives and immigrants compete for work. The boundaries of the actual radius of such local labour markets are unlikely to coincide with counties. As our data allows us to choose alternative geographical units, we experiment with three of them: counties, LAs and TTWAs (see Table A1). Counties and LAs are first and second tier administrative geographies and TTWAs are local labour market functional geographies (see Annex). Table A1 shows that there are 408 LAs, as well as 232 TTWAs in the 2001 definition and 297 TTWAs in the 1991 definition in Great Britain (as opposed to the 49 counties for the UK we used in the main analysis) (see Annex). We already showed above that our results are robust to two alternative definitions of counties. Table 7 now also shows that our results are again remarkably robust to using LAs and TTWAs, despite a smaller sample that again excludes Northern Ireland. Our immigrant-native earnings gap estimates change only very marginally, now ranging between 2.1% and 2.2%.

e. The earnings gap over time

We have so far discussed assimilation as a snapshot for the whole sample period, pulling together immigrants that have been in the UK for different lengths of time (see Table 1). We now look at different immigrant profiles to build a picture of different assimilation trajectories over time across continents of nationality (i.e. the continent where the country of nationality is located) and cohorts of arrival. To do this, we re-estimate our model with two alterations. Firstly, we include an extra term, $X_{iat}I_i$ in the model. By interacting each control in X_{iat} with our immigrant indicator I_i (see Section

3), we allow the effect of each control to differ across natives and immigrants. Secondly, we include the variable "years since immigration", and its squared value, as an additional control in the model to account for the effect of time spent in the UK on earnings. As we already control for experience, via age, and since "years since immigration" is often identical to experience in the UK, this enables us to gauge the weight that employers attach to experience gained in the UK. The resulting is a standard model in the literature used to estimate the immigrant-earnings gap at entry and over time (Chiswick 1980; Bell 1997; Borjas 1999; Lubotsky 2007).

Table 8 (see also its counterpart Table A2) shows significant immigrant-native earnings gap estimates at entry and over time. The first column in Table 8 shows that immigrants on average earn 51.5% less than comparable natives at the point of entry in the UK. The gap estimates narrow to -2% after one year and -0.7% after two years in the UK. It is then 9.8% after 10 years and 22% after 20 years. This suggests that immigrants' earnings catch up with natives' earnings in a little over two years.

There is, however, considerable variability among continents of nationality. The second column in Table 8 shows that the earnings gap estimate for Asians and Middle Easterners goes from -61.2% at entry, to -24.2% after one year, -11.1% after 10 years, and 2% after 20 years. In contrast, the earnings gap for North Americans goes from -37.3% at entry, to 23.8% after one year, 36.9% after 10 years, and 50% after 20 years. This suggests that North Americans not only have a smaller negative gap at entry, but also a faster catch up rate over time. While Asians and Middle Easterners' earnings took almost 20 years to catch up with natives' earnings, North Americans' earnings substantially surpassed natives' earnings in less than a year. More generally, column 2 in Table 8 shows that the earnings gap estimates are less negative for North Americans, Europeans and Australians at entry and that their earnings catch up faster with natives'. Bell (1997) estimated a comparable model using GHS data for the 1970s and 1980s, and also found that the gap at entry, and the subsequent catch-up, is more negative for Asians (Indians) than for Europeans. Dustmann and Fabbri (2005) estimated a simpler model using LFS data for the 1980s through to the 2000s and also found that the gap was more negative for Asians and Africans than for Europeans.

There is also considerable variability among cohorts of arrival. The third column in Table 8 shows that the earnings gap estimate for those arriving in the 1945-1949 post war, goes from -18.1% at entry, to 54.5% after one year, 68.6% after 10 years, and 85.6% after 20 years. In contrast, the earnings gap for those arriving most recently, in 2000-2004, goes from -53.9% at entry, to -2.8% after one year, 11.2% after 10 years, and 28.3% after 20 years. This suggests that those arriving in the post war not only have a smaller negative gap at entry, but also a faster catch up rate over time. While the earnings of those arriving most recently took a little over two years to catch up with natives' earnings, the earnings of those arriving in the post war greatly surpassed natives' earnings in less than a year. More generally, column 3 in Table 8 shows that the earnings gap estimates are less negative for earlier cohorts of arrival. For example, immigrants that arrived in the 1940s and 1950s earned around 20%-30% less than comparable natives at entry on average; whereas immigrants that arrived in the 1990s and early 2000s earned around 50% less than comparable natives at entry on average. Bell (1997) also found successively larger cohort estimates for some groups of immigrants but not for others.

Our estimates contrast with entry earnings gap estimates between -15% and -35% and with slower catch up rates for the US in a roughly comparable model, which however does not control for country of nationality (Borjas 1999). There are many reasons why our estimates are larger, including differences in the labour market and in the immigrant population composition, as well as differences in model specification, data type and sample period. The US model uses the wage rate, instead of annual earnings (see Section 2), different controls and data from the 1970, 1980 and 1990 Census. Other obvious differences include the fact that the US has had substantial low skilled immigration whereas the UK has had comparatively larger highly skilled immigration (Borjas 1994). In addition, perhaps because the US labour market is more flexible, low paid immigrants in the UK have a more negative earnings gap at entry and thus have faster earnings growth (Chiswick et al. 2008). Furthermore, such low paid workers in the UK might be more skilled and hence might overcome the usual earnings and occupation downgrading they suffer at entry more quickly (Friedberg 2001; Manacorda et al. 2007).

Our main conclusions are that North Americans, Europeans and Australians fare better at entry and their earnings catch up faster with natives'. Similarly, earlier cohorts fare better than more recent ones at entry, and the earnings of immigrants from such cohorts catch up faster with natives' earnings. These examples of trajectories of immigrant-native earnings gap over time across continents of nationality and cohorts of arrival illustrate how the pooled estimates mix together very diverse groups of immigrants who differ widely in a range of individual characteristics (such as English proficiency, work ethics, skills transferability, motivation, etc.). This is in line with our earlier results and completes the picture of our earlier analysis.

5. Conclusions

This paper investigates the degree to which immigrants have been able to assimilate into UK recipient labour markets by estimating the immigrant-native earnings gap across the earnings distribution, across continents of nationality and across cohorts of arrival between 1981 and 2006. By using a rich, underexploited, sizeable and long longitudinal dataset that has rarely been used for immigration analysis, we track a large sample of immigrants over 25 years. We are able to control for observable and unobservable individual specific characteristics as well as for specific characteristics of both time periods and recipient labour markets, defined as small geographical areas, and crucially, for the interaction of the two, in a robust empirical model specification. We also control for cohort specific effects and nationality specific effects. This way, we separately control for the role of each of these dimensions on immigrants’ assimilation. The individual characteristics dimension seems to be preponderant in explaining most of the gap.

This paper is an important contribution, as previous studies have not, possibly due to data limitations, estimated the immigrant-native earnings gap using such a robust empirical specification, tracking such a large sample of individuals and defining recipient labour markets as small geographical areas consistently over such a long time period.

Our results show little evidence of large or persistent earnings disparities across the earnings distribution, across cohorts or across nationalities. These findings are supportive evidence of successful assimilation of immigrants into the UK. Recipient labour markets primarily reward individuals’ characteristics other than, and regardless of, their immigration status. This in turn, facilitates assimilation. Nevertheless some distinctive features emerge. Immigrants entering the labour market at the bottom of the earnings distribution tend to have a less favorable assimilation experience. Also, immigrants entering the UK in the post war period experienced faster assimilation, suggesting, possibly, a more positive attitude towards immigration associated with the role of immigrants in the post war reconstruction effort. Earlier cohorts, such as the post war cohorts, not only fare better than more recent ones at entry, but also the earnings of immigrants in such cohorts catch up faster with natives' earnings. Similarly, North Americans, Europeans and Australians fare better at entry and their earnings catch up faster with natives' earnings. More generally, investigating the evolution of the immigrant-native earnings gap over time reveals how immigrants from different continents of nationality and cohorts of arrival have different assimilation trajectories.

The emergence of these distinctive features highlights that assimilation effects – and immigration effects more generally – feed through complex channels in the economy that include factor equalization as well as industry structure and output mix adjustments. In other words, assimilation effects depend on how native workers respond to competition from immigrants, the degree of substitution or complementarity between immigrant and native labour, and how firms alter their production function and production mix in response to immigration-led labour supply shifts. These, in turn, affect productivity, wages, employment and growth in recipient labour markets. Needless to say, these have been, and continue to be, fruitful avenues for future research.

6. Annex – Geographical Units Definitions

We model area fixed effects exploiting the geography of the UK (see Sections 3 and 4d). However, the UK geography over the period we study is not straighforward:

Administrative Geographies. Regions. The UK is divided into Statistical Standard Regions (SSRs), which roughly coincide with Government Office Regions (GORs). These are: London, South West, South East, East of England, West Midlands, East Midlands, North West, North East, Yorkshire and the Humber, Wales, Scotland and Northern Ireland. The last three are not technically GORs, but they are often reported alongside GORs (ONS 1999). Because sample size limitations in commonly used UK datasets prevent levels of disaggregation below GORs (see Introduction), most of the UK immigration literature uses this geography. As our data allows us to choose alternative geographical units, we experiment with three of them: counties, LAs and TTWAs. Counties and LAs are first and second tier administrative geographies and TTWAs are local labour market functional geographies.

Counties. Local government in the UK has been the subject of a constant re-structuring process. In the mid 1970s a radical overhaul introduced a two tier system of local government across the country. In England and Wales, counties were the top tier and Local Authority Districts (LADs), the bottom tier. In Scotland, the upper tier was Regions and the lower tier, Districts. In contrast, in Northern Ireland, the then existing two tier system was replaced by a single tier of District Council Areas (DCAs) (ONS 1999).

This system was in force until the mid 1990s, when the Local Government Reorganization (LGR) committee recommended that some non-metroplitan areas in England be re-set as single tier Unitary Authorities (UAs) (ONS 1999). In the same period, Wales also underwent substantial changes with UAs being introduced, while Scotland was divided into geographical units called Council Areas (CAs). Northern Ireland was unaffected by the reorganization and retained its structure based on DCAs.

Table A1 summarizes the pre-LGR and post-LGR systems described above. Since our sample period spans the 1980s through to the 2000s, we are restricted to using the pre-LGR definition, where England is divided into 39 counties, 6 metropolitan counties and Greater London, totalling 46 geographical units; Wales is divided into 8 counties;

Scotland, into 12 regions (not to be counfused with GORs); and Northern Ireland, into 26 DCAs. This definition is really a mix of 46-8-12-26 "counties, regions and DCAs". This definition is close in nature to that used in the 1981 Census, where London is split into Inner and Outter London, resulting in a mix of 47-8-10 "counties and regions" (Northern Ireland is not included in this definition) (see http://casweb.mimas.ac.uk/step1_81.cfm). In contrast, the 2001 Census definition is a mix of 42-22-32-26 "counties, UAs, CAs and DCAs" (see http://casweb.mimas.ac.uk/2001/start.cfm).

Other competing definitions exist, however (see Table A1). One argument is that the geographical units in the pre-LGR are not evenly defined across the different UK countries, either in terms of population size or in terms of territory size. For example, Nothern Ireland (26 geographical units) is much more finely split than, say, London (one geographical unit). So, one competing definition maintains the 46 geographical units for England, and treats Wales, Scotland and Northern Ireland each as a GOR. This definition is then a mix of 46-1-1-1 "counties and GORs" (see Table A1). This matters when estimating empirical models such as ours above because area fixed effects modelling differs across geographical units. For instance, at the county-region-DCA level, different parts of Northern Ireland, say, are modelled as a number of small local labour markets; in contrast, at the county-GOR level, Northern Ireland is treated as one single labour market. We checked the robustness of our results by modelling area fixed effects using the 46-1-1-1 county-GOR and 46-8-12-26 county-region-DCA definitions above,[13] and found remarkably robust results (see Table 7 and Section 4d).

[13] The counties we use are as follows. **England**: 1 Avon 2 Bedfordshire 3 Berkshire 4 Buckingamshire 5 Cambridgeshire 6 Cheshire 7 Cleveland 8 Cornwell and Isle of Scilly 9 Cumbria 10 Derbyshire 11 Devon 12 Dorset 13 Durham 14 East Sussex 15 Essex 16 Gloucestershire 17 Hampshire 18 Herefordshire and Worcester 19 Hertfordshire 20 Humberside 21 Isle of Wight 22 Kent 23 Lancashire 24 Leicestershire 25 Lincolnshire 26 London 27 Manchester 28 Merseyside 29 Norfolk 30 North Yorkshire 31 Northamptonshire 32 Northumberland 33 Nottinghamshire 34 Oxfordshire 35 Shropshire 36 Somerset 37 South Yorkshire 38 Staffordshire 39 Suffolk 40 Surrey 41 Tyne and Wear 42 Warwickshire 43 West Midlands 44 West Sussex 45 West Yorkshire 46 Wiltshire. **Wales**: 1 Clwyd 2 Dyfed 3 Gwent 4 Gwynedd 5 Mid Glamorgan 6 Powys 7 South Glamorgan 8 West Glamorgan. **Scotland**: 1 Borders 2 Central 3 Dumfries and Galloway 4 Eilean Siar 5 Fife 6 Grampian 7 Highland 8 Lothian 9 Orkney Islands 10 Shetland Islands 11 Strathclyde 12 Tayside. **Northern Ireland**: 1 Antrim 2 Ards 3 Armagh 4 Ballymena 5 Ballymoney 6 Banbridge 7 Belfast 8 Carrickfergus 9 Castlereagh 10 Coleraine 11 Cookstown 12 Craigavon 13 Down 14 Dungannon 15 Fermanagh 16 Larne 17 Limivady 18 Lisburn 19 Londonderry 20 Magherafelt 21 Moyle 22 Newry+Mourne 23 Newtownabbey 24 North Down 25 Omagh 26 Strabane (see ONS 1999, p. 47-55; and http://casweb.mimas.ac.uk/step1_81.cfm; https://geoportal.statistics.gov.uk/geoportal/catalog/main/home.page; http://www.ons.gov.uk/ons/guide-

Local Authorities. The definitions above mix first and second tier geographical units. An alternative is to use single tier geographical units, such as LADs, created in England between the mid 1960s and mid 1970s (DCLG 2010). LAs are defined to include LADs and UAs (DEFRA 2005). We combine LAs with the single tier CAs and DCAs, discussed above, to obtain the 354-22-32-26 LAD-UA-CA-DCA definition (see Table A1). This definition counters some of the criticisms above, as now all areas of the UK are more evenly split, both in terms of population size and in terms of territory size. This allows us to treat each unit as a small local labour market, eliminating the disparity of modelling the whole of London as one single labour market, for example. We checked the robustness of our results by modelling area fixed effects using this definition and found remarkably robust results (see Table 7 and Section 4d).

Functional Geographies. Local Labour Markets. The official definition of local labour markets in the UK is Travel To Work Areas (TTWAs). TTWAs are defined to encompass both the homes and workplaces of most workers. That is, TTWAs are defined so that most commuting flow is contained within their boundaries. TTWAs were introduced in the 1960s but are reviewed with each new Census to account for changes in the dynamic patterns of commuting, due, for example, to improvements in transport infrastructure and individual's preferences reflecting trends and shifts in the geographical space (Coombes and Bond 2008). As a result, TTWAs are ever changing geographical units, which are regularly updated with dramatic boundary discontinuities (see Table A1). Because of this, it is conceptually misleading to use TTWAs over a long period of time, such as the one in our analysis, as we would be implicitly assuming that dynamic labour markets are unchanging. Furthermore, ongoing changes in their definition criteria[14] make TTWAs non-comparable over time (Coombes and Bond 2008).

---

method/geography/products/postcode-directories/-nspp-/index.html). Although our data records postcode and LAs for those in Great Britain, it does not record any information on address other than country for those in Northern Ireland. As a result, we are unable to use the 46-8-12-26 county-region-DCA definition including Northern Ireland, and instead use this definition for Great Britain only (46-8-12 county-region). The same is the case below for LAs (we use 408 geographical units) and TTWAs (we use respectively 232 and 297 geographical units with the 2001 and 1991 definitions) (see Tables A1 and 7 and Section 4d).

[14] For example, the 2001 definition relaxes the criteria that TTWAs be contained within national borders and also lowers the self-containment proportion criteria (see Table A1). Other improvements in the 2001 definition include commuting data covering 100% of those in work, up from 10% in earlier definitions. Also, the data in 2001 is aggregated to smaller geographies than the earlier wards (see Table A1). Finally,

Bearing in mind these caveats, we checked the robustness of our results by modelling area fixed effects using both the 2001 and 1991 TTWA definitions, formed, respectively, of 243 and 308 geographical units (see Footnote 13) – the postcode-TTWAs mapping for the 1981 definition proved impossible to obtain.[15] The idea, as some argue, is that the underlying labour market economic activity in, say, the 2001 TTWAs is an enduring geographical reality that existed before the 2001 Census Day and outlives changes in the definition criteria. Put differently, the 2001 TTWA were an aspect of the labour economic activity back in the 1980s and 1990s, though they might have been defined then as, say, cities or regions, and the intensity and dynamics of labour market activities then might have been different (Coombes 2010). By using both the 1991 and 2001 TTWAs definitions, each in turn, we allow for two different "weighing systems" in the intensity and dynamics of such labour market activities, and test whether our results are robust to this. Indeed, we found remarkably robust results (see Table 7 and Section 4d).

unlike in the rest of the UK, the number of TTWAs in Northern Ireland remained unchanged between the 1991 and 2001 definitions (ONS 2003 and 2007, OECD 2002, Coombes and Bond 2008).

[15] Although this mapping is available for 2001 and 1991 from https://geoportal.statistics.gov.uk/geoportal/catalog/content/filelist.page, http://geoconvert.mimas.ac.uk and http://census.edina.ac.uk/easy_download.html (see also http://www.nomisweb.co.uk), it is not available for 1981. We directly contacted ONS and EDINA and confirmed that such early mapping is indeed unavailable. Also note that when mapping our data postcodes into TTWAs using the 1991 definition, observations that had either new post-1991 or obsolete pre-1991 postcodes were dropped. Similarly, when using the 2001 definition, obsolete pre-2001 postcodes observations were again dropped. However the loss of observations was relatively very small (see Table 7) and the results remarkably robust (see Section 4d).

7. References

Bell, B. D. 1997. The performance of immigrants in the United Kingdom: Evidence from the GHS. *Economic Journal* 107 (441):333-344.

Black, D., Kolesnikova, N. Lowell J. 2009. Earnings Functions When Wages and Prices Vary by Location. *Journal of Labor Economics*, 27: 21-47.

Borjas G. 1985. Assimilation, changes in cohort quality and the earnings of immigrants. *Journal of Labor Economics* 3: 463-489.

Borjas G. 1994. Economics of immigration. *Journal of Economic Literature* 32: 1667-1717.

Borjas, G. J. 1995. Ethnicity, Neighborhoods, and Human-Capital Externalities. *American Economic Review* 85 (3):365-390.

Borjas, G. J. 1999. Immigration and welfare magnets. *Journal of Labor Economics* 17 (4):607-637.

Butcher K. and J. DiNardo. 2002. The immigrant and native-born wage distributions: Evidence from the United States censuses, *Industrial and Labor Relations Review* 56 (1): 97-121

Card, D. 2005. Is the new immigration really so bad? *Economic Journal* 115 (507):F300-F323.

Card, D., and J. E. Dinardo. 2000. Do Immigrant Inflows Lead to Native Outflows? *American Economic Review*, 90, 360-367.

Chiswick B. 1978. The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy* 86: 897-921.

Chiswick, B. R. 1980. The Earnings of White and Coloured Male Immigrants in Britain, *Economica*, 47, 81-87

Chiswick, B. R. and T. J. Hatton. 2003. International Migration and the Integration of Labor Markets. *Globalization in Historical Perspective*. M. Bordo, A. Taylor and J. G. Williamson. Chicago/London, University of Chicago Press: 65-117.

Chiswick, B. R., Y. L. Lee, and P. W. Miller. 2005. Immigrant earnings: A longitudinal analysis. *Review of Income and Wealth* (4):485-503.

Chiswick, B. R., A. T. Lee, 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.

Constant, A. and K. F. Zimmermann. 2011. Circular and Repeat Migration: Counts of Exits and Years Away from the Host Country. *Population Research and Policy Review*, 30(4), 495-515.

Coombes, M. and S. Bond. 2008. *Travel-to-Work Areas: The 2007 Review*. London: Office for National Statistics.

Coombes, M. 2010. Defining Labour Market Areas by Analysing Commuting Data: Innovative Methods in the 2007 Review of Travel-to-Work Areas, J. Stillwell, O. Duke-Williams and A. Dennett, *Technologies for Migration and Commuting Analysis: Spatial Interaction Data Applications.* New York: Information Science Reference.

Crozet, M. 2004. Do migrants follow market potentials? An estimation of a new economic geography model. *Journal of Economic Geography* 4 (4):439-458.

DCLG. 2010. *Local Government Financial Statistics England*. London: Department for Communities and Local Government.

DEFRA. 2005. *Defra Classification of Local Authority Districts and Unitary Authorities in England*. London: Birkbeck College.

Dickens, R. and A. McKnight. 2008. Assimilation of migrants into the British labour market. *CASE Discussion Papers* 133.

Dustmann, C. 1994. Speaking Fluency, Writing Fluency and Earnings of Migrants. *Journal of Population Economics*, 7(2), 133-56.

Dustmann, C. 2003. Return Migration, Wage Differentials and the Optimal Migration Duration. *European Economic Review*, 47, 353-69.

Dustmann, C., F. Fabbri, I. Preston, and J. Wadsworth. 2003. Labour Market Performance of Immigrants in the UK Labour Market, *Home Office Report*, 05/03.

Dustmann, C., and F. Fabbri. 2003. Language Proficiency and Labour Market Performance of Immigrants in the UK, *Economic Journal* 113, 695-717.

Dustmann, C., and F. Fabbri. 2005. Immigrants in the British labour market. *Fiscal Studies* 26 (4):423-470.

Dustmann, C., F. Fabbri, I. Preston. 2005. The Impact of Immigration on the British Labour Market, *The*

Economic Journal, 115, F324-F341.

Dustmann, C. and Y. Weiss. 2007. Return Migration: Theory and Empirical Evidence from the UK. *British Journal of Industrial Relations*, 45(2), 236-56.

Dustmann, C., A. Glitz, and T. Frattini. 2008. The labour market impact of immigration. *Oxford Review of Economic Policy* 24 (3):478-495.

Dustmann, C., T. Frattini, and I. P. Preston. 2013. The Effect of Immigration along the Distribution of Wages. *Review of Economic Studies* 80 (1):145-173.

Faggian, A.; P. McCann and S. Sheppard. 2006. An Analysis of Ethnic Differences in UK Graduate Migration Behaviour. *The Annals of Regional Science*, 40, 461-71.

Faggian, A., P. McCann and S. Sheppard. 2007. Some Evidence That Women Are More Mobile Than Men: Gender Differences in UK Graduate Migration Behaviour. *Journal of Regional Science*, 47(3), 517-39.

Friedberg, R. M. 2001. The Impact of Mass Migration on the Israeli Labor Market. *The Quarterly Journal of Economics*, November, 1373-1408.

Gagliardi, L. 2014. Does skilled migration foster innovative performance? Evidence from British local areas. *Papers in Regional Science*, forthcoming.

Gallagher, A. M. 1991. *Majority Minority Review 2: Employment, Unemployment and Religion in Northern Ireland*. Coleraine: University of Ulster.

Gilpin, N., M. Henty, S. Lemos, J. Portes and C. Bullen. 2006. The Impact of Free Movement of Workers from Central and Eastern Europe on the UK Labour Market. *Department for Work and Pensions Working Paper*, 29.

Goodhart, D. 2010. A Crowded Island. *Prospect* 95 (February).

Halvorsen, R., and R. Palmquist. 1980. The Interpretation of Dummy Variables in Semilogarithmic Equations. *American Economic Review* 70: 474-475.

Kanbur, R. and H. Rapoport. 2005. Migration selectivity and the evolution of spatial inequality. *Journal of Economic Geography* 5 (1):43-57.

LaLonde, R. and R. Topel. 1997. Economic Impact of International Migration and the Economic Performance of Migrants, M. Rosenzweig and O. Stark, *Handbook of Population and Family Economics.* New York: Elsevier Science, 799-850.

Lee, S.C. 2010. Ability sorting and consumer city. *Journal of Urban Economics* 68(1): 20-33.

Lemos, S. and J. Portes. 2013. New labour? The Impact of Migration From the New European Union Member States on the UK Labour Market. *The B.E. Journal of Economic Analysis and Policy* 14(1): 299-338.

Longhi, S., P. Nijkamp, and J. Poot. 2005. A meta-analytic assessment of the effect of immigration on wages. *Journal of Economic Surveys* 19 (3):451-477.

Longhi, S., P. Nijkamp, and J. Poot. 2010. Meta-analyses of labour-market impacts of immigration: key conclusions and policy implications. *Environment and Planning C-Government and Policy* 28 (5):819-833.

Lubotsky, D. 2007. Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings. *Journal of Political Economy*, 115, 820-867.

Manacorda, M., A. Manning and J. Wadsworth. 2007. The Impact of Immigration on the Structure of Male Wages: Theory and Evidence from Britain. *Research in Labour Economics*, 26(125-155).

Mayda, A. M. 2006. Who Is Against Immigration? A Cross-Country Investigation of Individual Attitudes toward Immigrants. *Review of Economics and Statistics* 88(3): 510-530.

Moretti, E. 2013. Real Wage Inequality. *American Economic Journal: Applied Economics*, 5(1): 65-103.

Nakamura, A., and M. Nakamura. 1985. *The Second Paycheck: A Socioeconomic Analysis of Earnings*. Edited by N. Y. A. Press.

Nathan, M. 2014. Same Difference? Minority Ethnic Inventors, Diversity and Innovation in the UK. *Journal of Economic Geography*, forthcoming.

Niebuhr, A. 2010. Migration and innovation: Does cultural diversity matter for regional R&D activity? *Papers in Regional Science* 89 (3):563-585.

OECD. 2002. *Redefining Territories the Functional Regions: The Functional Regions*. Paris: OECD Publishing.

ONS. 1999. *Gazetteer of the Old and New Geographies of the United Kingdom*. London: Office for National Statistics.

ONS. 2003. *Guide to Regional and Local Labour Market Statistics*. London: National Statistics Publication.

ONS. 2007. *Introduction to the 2001-Based Travel-to-Work Areas*. London: National Statistics Publication.

Ottaviano, G. I. P., and G. Peri. 2006. The economic value of cultural diversity: evidence from US cities. *Journal of Economic Geography* 6 (1):9-44.

Rodriguez-Pose, A., and T. D. Ketterer. 2012. Do Local Amenities Affect the Appeal of Regions in Europe for Migrants? *Journal of Regional Science* 52 (4):535-561.

Rodriguez-Pose, A., and V. Tselios. 2010. Returns to migration, education and externalities in the European Union. *Papers in Regional Science* 89 (2):411-434.

Rodriguez-Pose, A., and M. Vilalta-Bufi. 2005. Education, migration, and job satisfaction: the regional returns of human capital in the EU. *Journal of Economic Geography* 5 (5):545-566.

Rodriguez-Pose, A., and V. von Berlepsch. 2014. When migrants rule: the legacy of mass migration on economic development in the US. Annals of the Association of American Geographers 104, 3, 628-651.

Shields, M., and S. Wheatley Price. 2002. The English Language Fluency and Occupational Success of Ethnic Minority Immigrant Men Living in English Metropolitan Areas. *Journal of Population Economics*, 137-160.

Suedekum, J., K. Wolf, and U. Blien. 2014. Cultural Diversity and Local Labour Markets. *Regional Studies* 48 (1):173-191.

Wadsworth, J. 2003. The Labour Market Performance of Ethnic Minorities in the Recovery. In *The labour Market Performance under New Labour: The State of Working Britain*, edited by R. Dickens, P. Gregg and J. Wadsworth: Palgrave Macmillan.

Table 1 - Descriptive Statistics						
VARIABLES	**LLMDB**		**LLMDB**		**LFS**	
	April 1981 - March 2006		**April 1997 - March 2007**		**January 1997 - March 2007**	
	Natives	Immigrants	Natives	Immigrants	Natives	Immigrants
I - POPULATION VARIABLES						
% aged:						
25 to 34 years old	30.17%	43.81%	29.87%	43.71%	29.06%	36.59%
35 to 64 years old	69.83%	56.19%	70.13%	56.29%	70.94%	63.41%
% of women	45.09%	44.53%	47.38%	45.15%	48.47%	47.51%
% from:						
EU (except A10)	-	31.62%	-	30.19%	-	25.66%
A10	-	4.47%	-	5.26%	-	5.43%
Europe (except EU)	-	3.43%	-	3.65%	-	2.65%
Asia and Middle East	-	22.17%	-	22.87%	-	27.94%
North America	-	5.76%	-	4.96%	-	4.63%
Latin America	-	3.32%	-	3.56%	-	6.59%
Africa	-	15.67%	-	16.70%	-	21.39%
Australasia and Oceania	-	6.97%	-	6.32%	-	4.74%
Unknown	-	6.59%	-	6.49%	-	na
Average age at arrival		23.73		24.48	-	na
Average nb of years since immigration		13.73		13.08	-	na
% with length of immigration						
0 to 1 years	-	9.24%	-	10.49%	-	na
2 to 3 years	-	11.56%	-	13.07%	-	na
4 to 5 years	-	9.37%	-	10.09%	-	na
6 to 10 years	-	17.82%	-	17.37%	-	na
11 to 15 years	-	15.15%	-	14.27%	-	na
16 to 20 years	-	11.64%	-	10.47%	-	na
over 20 years	-	25.21%	-	24.25%	-	na
% arrived during:						
1945-1949	-	0.67%	-	0.03%	-	1.03%
1950-1954	-	1.22%	-	0.42%	-	1.77%
1955-1959	-	2.93%	-	1.51%	-	3.45%
1960-1964	-	4.02%	-	2.66%	-	7.23%
1965-1969	-	4.55%	-	3.24%	-	8.57%
1970-1974	-	3.40%	-	2.70%	-	9.16%
1975-1979	-	15.52%	-	11.72%	-	7.61%
1980-1984	-	10.05%	-	8.44%	-	5.57%
1985-1989	-	14.64%	-	13.69%	-	7.69%
1990-1994	-	13.51%	-	14.52%	-	9.10%
1995-1999	-	13.97%	-	19.12%	-	13.49%
2000-2004	-	14.08%	-	19.90%	-	13.26%
2005	-	1.45%	-	2.05%	-	1.39%
% located in:						
East Midlands	7.42%	4.96%	7.53%	4.78%	7.66%	4.79%
East of England	9.16%	8.09%	9.27%	7.97%	9.69%	9.12%
London	10.82%	39.43%	10.31%	40.89%	9.61%	41.76%
North East	4.54%	2.27%	4.46%	2.11%	4.32%	1.31%
North West	12.07%	6.62%	11.83%	6.29%	11.97%	5.57%
Northern Ireland	2.50%	1.42%	2.53%	1.40%	2.51%	1.22%
Scotland	9.38%	5.58%	9.46%	5.37%	9.03%	3.67%
South East	13.16%	13.46%	13.54%	13.54%	13.99%	13.87%
South West	7.93%	5.75%	8.23%	5.58%	8.51%	5.24%
Wales	4.67%	2.14%	4.70%	2.08%	4.87%	1.80%
West Midlands	9.45%	5.70%	9.29%	5.61%	9.20%	6.90%
Yorkshire and the Humber	0.08904	0.04589	0.08849	0.04389	0.08659	0.04736
						continues…

						...continued
II - LABOUR MARKET VARIABLES						
% in work:						
1 to 25 weeks in the year	8.50%	14.08%	8.17%	14.13%	na	na
26 to 50 weeks in the year	13.67%	20.37%	13.70%	20.87%	na	na
51 to 52 weeks in the year	72.94%	62.92%	78.05%	64.93%	na	na
Average number of employed weeks in the year	44.43	42.50	47.17	43.82	na	na
Average number of jobs in the year	1.36	1.60	1.43	1.67	na	na
5th percentile of the log real earnings distribution	7.77	7.42	7.82	7.42	9.22	9.15
10th percentile of the log real earnings distribution	8.40	8.13	8.45	8.14	9.39	9.34
20th percentile of the log real earnings distribution	8.97	8.78	8.99	8.78	9.59	9.57
30th percentile of the log real earnings distribution	9.34	9.21	9.35	9.19	9.73	9.74
40th percentile of the log real earnings distribution	9.57	9.50	9.58	9.48	9.86	9.88
50th percentile of the log real earnings distribution	9.74	9.73	9.76	9.72	9.99	10.03
60th percentile of the log real earnings distribution	9.90	9.94	9.93	9.93	10.12	10.17
70th percentile of the log real earnings distribution	10.06	10.14	10.10	10.15	10.26	10.32
80th percentile of the log real earnings distribution	10.23	10.36	10.28	10.38	10.42	10.51
90th percentile of the log real earnings distribution	10.47	10.71	10.53	10.75	10.65	10.80
Average of the log real earnings distribution	9.56	9.55	9.60	9.56	10.00	10.04
Standard deviation of the log real earnings distribution	0.93	1.12	0.95	1.14	0.53	0.61
Number of observations	3313138	221877	1835337	156896	507606	42230
Number of individuals	316391	38074	265849	33500	na	na
Average number of times an individual is observed	19.63	22.08	24.50	25.15	na	na
Average number of observations per year	114246	7651	63287	5410	50761	4223

Source: Lifetime Labour Market Database and Labour Force Survey.
(1) Sample includes males aged 25 to 64 and females aged 25 to 60 employed and earning between £100 and £1000000 in the year, observed at least twice, with non-missing address.
(2) LFS figures are courtesy of the DWP.

Table 2 - Immigrant-Native Earnings Gap

Variable	coefficient		s. errors	coefficient	s. errors	
Intercept	**9.526**	***	0.054	**2.078**	0.053	***
Immigrant (=1)	**-0.107**	***	0.006	**0.023**	0.002	***
Sex (female=1)				**-0.135**	0.001	***
Age				**0.018**	0.000	***
Age squared				**0.000**	0.000	***
Number of employed weeks				**0.032**	0.000	***
Number of jobs				**0.001**	0.000	***
Lagged number of employed weeks				**-0.022**	0.000	***
Lagged log real earnings				**0.726**	0.001	***
Individual fixed effects	*no*			*yes*		
Area fixed effects	*yes*			*yes*		
Time fixed effects	*yes*			*yes*		
Interaction of area and time fixed effects	*yes*			*yes*		
Adjusted R-squared	0.017			0.765		
Sample size	3535015			3381717		
Number of individuals	354465			354465		

(1) The sample data used includes males aged 25 to 64 and females aged 25 to 59 earning between £100 and £1000000 in any one tax year who are observed at least twice with non-missing address (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) and their interaction. Only the model in the right-most columns in bold control 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.

(4) Level of significance: *p<0.10, **p<0.05, ***p<0.01

Table 3 - 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	**8.382** ***	0.126	**9.046** ***	0.071	**9.484** ***	0.055	**9.634** ***	0.040	**9.728** ***	0.029	**9.811** ***	0.025	**9.869** ***	0.023	**9.941** ***	0.022	**10.034** ***	0.023	**10.208** ***	0.029
Immigrant (=1)	**-0.445** ***	0.009	**-0.361** ***	0.005	**-0.288** ***	0.004	**-0.233** ***	0.003	**-0.165** ***	0.002	**-0.111** ***	0.002	**-0.057** ***	0.002	**-0.007** ***	0.002	**0.045** ***	0.002	**0.152** ***	0.002
Individual fixed effects	*no*		*no*		*no*		*no*		*no*		*no*		*no*		*no*		*no*		*no*	
Area fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Time fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Interaction of area and time fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
R-squared	0.008		0.007		0.007		0.008		0.009		0.011		0.013		0.016		0.021		0.032	
Sample size	3535015		3535015		3535015		3535015		3535015		3535015		3535015		3535015		3535015		3535015	
Number of Individuals	354465		354465		354465		354465		354465		354465		354465		354465		354465		354465	
Intercept	**-1.072** ***	0.038	**-0.740** ***	0.025	**-0.280** ***	0.014	**0.046** ***	0.012	**0.272** ***	0.008	**0.417** ***	0.006	**0.660** ***	0.007	**1.057** ***	0.009	**1.671** ***	0.014	**2.681** ***	0.023
Immigrant (=1)	**-0.034** ***	0.003	**-0.016** ***	0.002	**-0.003** **	0.001	**0.002** ***	0.001	**0.006** ***	0.001	**0.012** ***	0.000	**0.020** ***	0.001	**0.033** ***	0.001	**0.052** ***	0.001	**0.089** ***	0.002
Sex (female=1)	**-0.070** ***	0.001	**-0.044** ***	0.001	**-0.022** ***	0.001	**-0.015** ***	0.000	**-0.012** ***	0.000	**-0.014** ***	0.000	**-0.023** ***	0.000	**-0.038** ***	0.000	**-0.062** ***	0.001	**-0.103** ***	0.001
Age	**0.016** ***	0.001	**0.008** ***	0.000	**0.004** ***	0.000	**0.002** ***	0.000	**0.001** ***	0.000	**0.000** ***	0.000	**0.000**	0.000	**0.001** ***	0.000	**0.004** ***	0.000	**0.009** ***	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	**0.051** ***	0.000	**0.043** ***	0.000	**0.035** ***	0.000	**0.030** ***	0.000	**0.028** ***	0.000	**0.028** ***	0.000	**0.027** ***	0.000	**0.027** ***	0.000	**0.026** ***	0.000	**0.022** ***	0.000
Number of jobs	**-0.103** ***	0.001	**-0.061** ***	0.001	**-0.032** ***	0.000	**-0.018** ***	0.000	**-0.007** ***	0.000	**0.004** ***	0.000	**0.015** ***	0.000	**0.026** ***	0.000	**0.039** ***	0.000	**0.057** ***	0.001
Lagged number of employed weeks	**-0.008** ***	0.000	**-0.012** ***	0.000	**-0.018** ***	0.000	**-0.022** ***	0.000	**-0.025** ***	0.000	**-0.026** ***	0.000	**-0.027** ***	0.000	**-0.027** ***	0.000	**-0.028** ***	0.000	**-0.029** ***	0.000
Lagged log real earnings	**0.858** ***	0.001	**0.905** ***	0.001	**0.943** ***	0.000	**0.954** ***	0.000	**0.962** ***	0.000	**0.959** ***	0.000	**0.939** ***	0.000	**0.904** ***	0.000	**0.851** ***	0.000	**0.771** ***	0.001
Individual fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Area fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Time fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Interaction of area and time fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
R-squared	0.653		0.663		0.672		0.669		0.664		0.653		0.636		0.608		0.568		0.508	
Sample size	3381717		3381717		3381717		3381717		3381717		3381717		3381717		3381717		3381717		3381717	
Number of Individuals	354465		354465		354465		354465		354465		354465		354465		354465		354465		354465	

(1) Notes as in Table 2.

Table 4 - Immigrant-Native Earnings Gap Across the Earnings Distribution - Excludes London

Variable	coefficient	s.errors	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	**average**		**5th**		**10th**		**20th**		**30th**		**40th**		**50th**		**60th**		**70th**		**80th**		**90th**	
Intercept	**2.166** ***	0.091	**-1.060** ***	0.040	**-0.730** ***	0.026	**-0.281** ***	0.015	**0.042** ***	0.012	**0.258** ***	0.008	**0.399** ***	0.006	**0.635** ***	0.007	**1.016** ***	0.010	**1.607** ***	0.014	**2.599** ***	0.024
Immigrant (=1)	**0.039** ***	0.002	**-0.019** ***	0.003	**-0.005** **	0.002	**0.004** **	0.001	**0.007** ***	0.001	**0.009** ***	0.001	**0.014** ***	0.001	**0.023** ***	0.001	**0.036** ***	0.001	**0.056** ***	0.001	**0.093** ***	0.002
Sex (female=1)	**-0.142** ***	0.001	**-0.073** ***	0.001	**-0.047** ***	0.001	**-0.023** ***	0.001	**-0.016** ***	0.000	**-0.013** ***	0.000	**-0.015** ***	0.000	**-0.023** ***	0.000	**-0.038** ***	0.000	**-0.062** ***	0.001	**-0.104** ***	0.001
Age	**0.019** ***	0.000	**0.017** ***	0.001	**0.009** ***	0.000	**0.004** ***	0.000	**0.002** ***	0.000	**0.001** ***	0.000	**0.000** *	0.000	**0.000**	0.000	**0.002** ***	0.000	**0.004** ***	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	**0.000** ***	0.000
Number of employed weeks	**0.032** ***	0.000	**0.050** ***	0.000	**0.042** ***	0.000	**0.034** ***	0.000	**0.030** ***	0.000	**0.027** ***	0.000	**0.027** ***	0.000	**0.027** ***	0.000	**0.027** ***	0.000	**0.025** ***	0.000	**0.022** ***	0.000
Number of jobs	**0.000**	0.000	**-0.105** ***	0.001	**-0.062** ***	0.001	**-0.033** ***	0.000	**-0.019** ***	0.000	**-0.008** ***	0.000	**0.003** ***	0.000	**0.014** ***	0.000	**0.026** ***	0.000	**0.039** ***	0.000	**0.056** ***	0.001
Lagged number of employed weeks	**-0.022** ***	0.000	**-0.008** ***	0.000	**-0.012** ***	0.000	**-0.018** ***	0.000	**-0.021** ***	0.000	**-0.024** ***	0.000	**-0.026** ***	0.000	**-0.026** ***	0.000	**-0.027** ***	0.000	**-0.027** ***	0.000	**-0.028** ***	0.000
Lagged log real earnings	**0.723** ***	0.001	**0.857** ***	0.001	**0.904** ***	0.001	**0.942** ***	0.000	**0.953** ***	0.000	**0.962** ***	0.000	**0.959** ***	0.000	**0.939** ***	0.000	**0.903** ***	0.000	**0.849** ***	0.000	**0.767** ***	0.001
Individual fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Area fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Time fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
Interaction of area and time fixed effects	*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*		*yes*	
R-squared	0.767		0.656		0.667		0.674		0.672		0.666		0.656		0.638		0.610		0.569		0.506	
Sample size	2962451		2962451		2962451		2962451		2962451		2962451		2962451		2962451		2962451		2962451		2962451	
Number of Individuals	310493		310493		310493		310493		310493		310493		310493		310493		310493		310493		310493	

(1) Notes as in Table 2.

Table 5 - Immigrant-Native Earnings Gap by Continent of Nationality

Variable	coefficient		s.errors	coefficient		s.errors
Intercept	**9.515**	***	0.054	**2.079**	***	0.053
EU (except A10)	**-0.008**		0.010	**0.034**	***	0.003
A10	**-0.431**	***	0.023	**0.024**	***	0.007
Europe (except EU)	**-0.034**		0.032	**0.031**	***	0.008
Asia and Middle East	**-0.220**	***	0.013	**-0.015**	***	0.003
North America	**0.369**	***	0.026	**0.118**	***	0.007
Central and South America	**-0.263**	***	0.028	**0.000**		0.007
Africa	**-0.191**	***	0.014	**0.006**	*	0.004
Australasia and Oceania	**-0.005**		0.019	**0.051**	***	0.006
Unknown	**-0.303**	***	0.022	**0.021**	***	0.006
Sex (female=1)				**-0.135**	***	0.001
Age				**0.018**	***	0.000
Age squared				**0.000**	***	0.000
Number of employed weeks				**0.032**	***	0.000
Number of jobs				**0.001**	***	0.000
Lagged number of employed weeks				**-0.022**	***	0.000
Lagged log real earnings				**0.726**	***	0.001
Individual fixed effects	*no*			*yes*		
Area fixed effects	*yes*			*yes*		
Time fixed effects	*yes*			*yes*		
Interaction of area and time fixed effects	*yes*			*yes*		
Adjusted R-squared	0.019			0.765		
Sample size	3535015			3381717		
Number of individuals	354465			354465		

(1) Notes as in Table 2.

Table 6 - Immigrant-Native Earnings Gap by Cohort of Arrival

Variable	coefficient		s. errors	coefficient		s. errors
Intercept	**9.489**	***	0.054	**2.078**	***	0.053
1945-1949 arrivals	**-0.208**	***	0.069	**0.058**	***	0.022
1950-1954 arrivals	**0.042**		0.046	**0.038**	***	0.012
1955-1959 arrivals	**0.146**	***	0.035	**0.038**	***	0.009
1960-1964 arrivals	**0.217**	***	0.031	**0.053**	***	0.007
1965-1969 arrivals	**0.339**	***	0.029	**0.061**	***	0.007
1970-1974 arrivals	**0.257**	***	0.033	**0.059**	***	0.008
1975-1979 arrivals	**0.053**	***	0.016	**0.019**	***	0.004
1980-1984 arrivals	**-0.005**		0.019	**0.022**	***	0.005
1985-1989 arrivals	**-0.115**	***	0.015	**0.003**		0.004
1990-1994 arrivals	**-0.262**	***	0.014	**-0.013**	***	0.004
1995-1999 arrivals	**-0.267**	***	0.013	**0.025**	***	0.004
2000-2004 arrivals	**-0.416**	***	0.010	**0.036**	***	0.004
2005 arrivals	**-0.816**	***	0.023	**0.142**	***	0.015
Sex (female=1)				**-0.135**		0.001
Age				**0.018**		0.000
Age squared				**0.000**		0.000
Number of employed weeks				**0.032**		0.000
Number of jobs				**0.001**		0.000
Lagged number of employed weeks				**-0.022**		0.000
Lagged log real earnings				**0.726**		0.001
Individual fixed effects	*no*			*yes*		
Area fixed effects	*yes*			*yes*		
Time fixed effects	*yes*			*yes*		
Interaction of area and time fixed effects	*yes*			*yes*		
Adjusted R-squared	0.021			0.765		
Sample size	3535015			3381717		
Number of individuals	354465			354465		

(1) Notes as in Table 2.

Table 7 - Immigrant-Native Earnings Gap - Robustness Checks

Variable	coefficient		s.errors	coefficient		s.errors	coefficient		s.errors	coefficient		s.errors
	Pre-LGR			**LAs**			**TTWA 2001**			**TTWA 1991**		
Intercept	**1.830**	***	0.229	**2.257**	***	0.255	**2.293**	***	0.218	**2.129**	***	0.215
Immigrant (=1)	**0.023**	***	0.002	**0.021**	***	0.002	**0.022**	***	0.002	**0.022**	***	0.002
Sex (female=1)	**-0.135**	***	0.001	**-0.136**	***	0.001	**-0.136**	***	0.001	**-0.136**	***	0.001
Age	**0.018**	***	0.000	**0.018**	***	0.000	**0.018**	***	0.000	**0.018**	***	0.000
Age squared	**0.000**	***	0.000	**0.000**	***	0.000	**0.000**	***	0.000	**0.000**	***	0.000
Number of employed weeks	**0.032**	***	0.000	**0.032**	***	0.000	**0.032**	***	0.000	**0.032**	***	0.000
Number of jobs	**0.001**	**	0.000	**0.001**	***	0.000	**0.001**	**	0.000	**0.001**	**	0.000
Lagged number of employed weeks	**-0.022**	***	0.000	**-0.022**	***	0.000	**-0.022**	***	0.000	**-0.022**	***	0.000
Lagged log real earnings	**0.726**	***	0.001	**0.724**	***	0.001	**0.723**	***	0.001	**0.723**	***	0.001
Individual fixed effects	*yes*			*yes*			*yes*			*yes*		
Area fixed effects	*yes*			*yes*			*yes*			*yes*		
Time fixed effects	*yes*			*yes*			*yes*			*yes*		
Interaction of area and time fixed effects	*yes*			*yes*			*yes*			*yes*		
Adjusted R-squared	0.765			0.766			0.765			0.765		
Sample size	3299294			3299294			3225633			3209953		
Number of individuals	346171			346171			342164			340724		

(1) Notes as in Table 2, except that area fixed effects are now, respectively, 66 Pre-LGR counties, 408 LADs, 232 TTWAs in 2001and 297 TTWAs in 1991(see Annex and Table A

Table 8 - Entry Earnings and Earnings Growth

Immigrant-Native Earnings Gap at Entry			
Pooled	-0.515		
By Continent of Nationality			
EU (except A10)		-0.537	
A10		-0.610	
Europe (except EU)		-0.548	
Asia and Middle East		-0.612	
North America		-0.373	
Central and South America		-0.608	
Africa		-0.595	
Australasia and Oceania		-0.479	
Unknown		-0.580	
By Cohort of Arrival			
1945-1949 arrivals			-0.181
1950-1954 arrivals			-0.296
1955-1959 arrivals			-0.333
1960-1964 arrivals			-0.315
1965-1969 arrivals			-0.302
1970-1974 arrivals			-0.326
1975-1979 arrivals			-0.395
1980-1984 arrivals			-0.415
1985-1989 arrivals			-0.466
1990-1994 arrivals			-0.514
1995-1999 arrivals			-0.499
2000-2004 arrivals			-0.539
2005 arrivals			-0.582
Immigrant-Native Earnings Gap			
Chosen continent or cohort		*North America*	*1945-1949*
After 1 year	-0.020	0.238	0.545
After 2 years	-0.007	0.253	0.560
After 3 years	0.007	0.268	0.575
After 4 years	0.020	0.283	0.591
After 5 years	0.033	0.298	0.606
After 10 years	0.098	0.369	0.686
After 20 years	0.220	0.500	0.856
Chosen continent or cohort		*Asia and Middle East*	*2000-2004*
After 1 year		-0.242	-0.028
After 2 years		-0.227	-0.013
After 3 years		-0.212	0.002
After 4 years		-0.197	0.017
After 5 years		-0.182	0.033
After 10 years		-0.111	0.112
After 20 years		0.020	0.283

(1) Estimates derived from coefficients in Table A2.

Table A1 - Geo Units

Definition	Number of Units	Total UK	Total GB	Reference
county-region-DCA	England: 46 (39 counties + 6 metropolitan counties + Greater London)	92	66	Faggian et al. (2007)
46-8-12-26	Wales: 8 Counties			ONS (1999)
Pre-LGR	Scotland: 12 Regions			
	N Ireland: 26 DCAs			
county-UA-CA-DCA	England: 87 (34 counties + 46UAs + 6 metropolitan counties + Greater London)	167	141	Faggian et al. (2006)
87-22-32-26	Wales: 22 UAs			ONS (1999)
Post-LGR	Scotland: 32 CAs			
	N Ireland: 26 DCAs			
county-GOR	England: 46 Counties	49	48	Lemos and Portes (2013)
46-1-1-1	Wales: 1 GOR			ONS (1999)
	Scotland: 1 GOR			
	N Ireland: 1 GOR			
county-region	England: 47 Counties (Inner and Outter London)	N/A	65	http://casweb.mimas.ac.uk/step1_81.cfm
47-8-10	Wales: 8 Counties			
Census 1981	Scotland: 10 Regions			
	N Ireland: N/A			
county-UA-CA-DCA	England: 42 (32 counties + 6 metropolitan counties + Inner and Outter London)	122	96	http://casweb.mimas.ac.uk/2001/start.cfm
42-22-32-26	Wales: 22 UAs			
Census 2001	Scotland: 32 CAs			
	N Ireland: 26 DCAs			
LAD-UA-CA-DCA	England: 354 LADs	434	408	Lemos and Portes (2013)
354-22-32-26	Wales: 22 UAs			DEFRA (2005)
	Scotland: 32 CAs			
	N Ireland: 26 DCAs			
TTWA 1981	Geo Units: Wards (Census Sectors in Scotland)	334	322	Gallagher (1991)
publication: 1984	TTWAs must be at least 70% self-cointained			ONS (2003)
	TTWAs cannot span the borders between England and either Wales of Scotland			
	TTWAs based on commuting data that covers 10% sample of those in work			
TTWA 1991	Geo Units: Wards (Census Sectors in Scotland)	308	297	http://census.edina.ac.uk/easy_download.html
publication: 1998	TTWAs must be at least 69.5% self-cointained			Coombes and Bond (2008)
	TTWAs cannot span the borders between England and either Wales of Scotland			ONS (2003 and 2007) OECD (2002)
	TTWAs based on commuting data that covers 10% sample of those in work			
TTWA 2001	Geo Units: Lower Layer Super Output Areas (Zones in Scotland and Super Output Areas in Northern Ireland)	243	232	http://www.nomisweb.co.uk
publication: 2007	TTWAs must be at least 66.67% self-cointained			Nathan (2014)
	TTWAs have no borders constraints			ONS (2007)
	TTWAs based on commuting data that covers 100% sample of those in work			

(1) Although our data records postcode and LA for those in England, Wales and Scotland, it does not record any information on address other than country for those in Northern Ireland. As a result, except for the definition county-GOR above, we are unable to use data for the whole of the United UK (united Kingdom) and instead use data for GB (Great Britain) only (excluding Northern Ireland).

Table A2 - Entry Earnings and Earnings Growth

Variable	coefficient		s. errors	coefficient		s. errors	coefficient		s. errors
Intercept	**7.771**	***	0.074	**7.765**	***	0.075	**7.739**	***	0.073
Immigrant (=1)	**-0.724**	***	0.067						
EU (except A10)				**-0.770**	***	0.067			
A10				**-0.943**	***	0.069			
Europe (except EU)				**-0.795**	***	0.071			
Asia and Middle East				**-0.947**	***	0.068			
North America				**-0.467**	***	0.069			
Central and South America				**-0.937**	***	0.071			
Africa				**-0.903**	***	0.068			
Australasia and Oceania				**-0.652**	***	0.068			
Unknown				**-0.866**	***	0.070			
1945-1949 arrivals							**-0.200**	*	0.107
1950-1954 arrivals							**-0.352**	***	0.085
1955-1959 arrivals							**-0.405**	***	0.081
1960-1964 arrivals							**-0.378**	***	0.079
1965-1969 arrivals							**-0.360**	***	0.077
1970-1974 arrivals							**-0.394**	***	0.077
1975-1979 arrivals							**-0.503**	***	0.071
1980-1984 arrivals							**-0.536**	***	0.070
1985-1989 arrivals							**-0.628**	***	0.069
1990-1994 arrivals							**-0.721**	***	0.068
1995-1999 arrivals							**-0.692**	***	0.069
2000-2004 arrivals							**-0.774**	***	0.070
2005 arrivals							**-0.872**	***	0.072
Years since migration	**0.007**	***	0.001	**0.004**	***	0.001	**0.008**	***	0.002
Years since migration squared	**0.000**	***	0.000	**0.000**	***	0.000	**0.000**	***	0.000
Sex (female=1) x Immigrant (=1)	**0.162**	***	0.009	**0.154**	***	0.008	**0.167**	***	0.009
Age x Immigrant (=1)	**0.007**	*	0.004	**0.011**	***	0.004	**0.006**	*	0.004
Age squared x Immigrant (=1)	**0.000**	***	0.000	**0.000**	***	0.000	**0.000**	***	0.000
Number of employed weeks x Immigrant	**0.012**	***	0.000	**0.012**	***	0.000	**0.012**	***	0.000
Number of jobs x Immigrant (=1)	**0.010**	***	0.003	**0.015**	***	0.003	**0.014**	***	0.003
Sex (female=1)	**-0.542**	***	0.002	**-0.542**	***	0.002	**-0.543**	***	0.002
Age	**0.059**	***	0.001	**0.059**	***	0.001	**0.059**	***	0.001
Age squared	**-0.001**	***	0.000	**-0.001**	***	0.000	**-0.001**	***	0.000
Number of employed weeks	**0.033**	***	0.000	**0.033**	***	0.000	**0.033**	***	0.000
Number of jobs	**-0.027**	***	0.001	**-0.027**	***	0.001	**-0.027**	***	0.001
Area fixed effects	*yes*			*yes*			*yes*		
Time fixed effects	*yes*			*yes*			*yes*		
Interaction of area and time fixed effects	*yes*			*yes*			*yes*		
R-squared	0.455			0.456			0.455		
Sample size	3381717			3381717			3381717		
Number of individuals	354465			354465			354465		

(1) Notes as in Table 2.

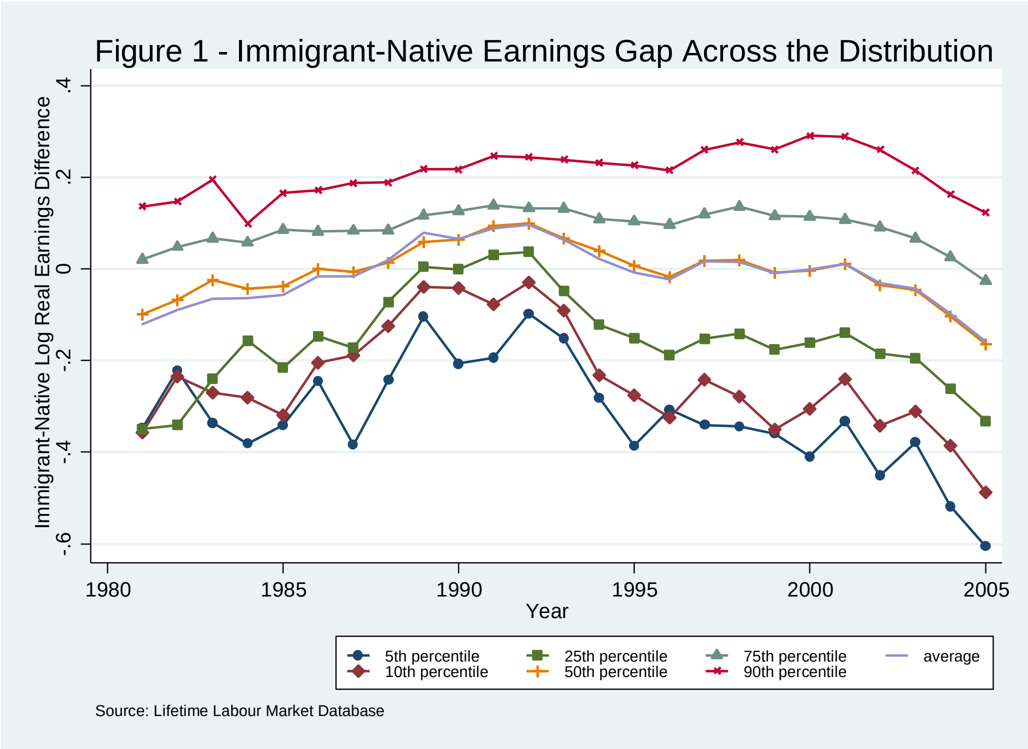


Figure 1 - Immigrant-Native Earnings Gap Across the Distribution
Immigrant-Native Log Real Earnings Difference
-.6
-.4
-.2
0
.2
.4
1980
1985
1990
1995
2000
2005
Year
5th percentile
10th percentile
25th percentile
50th percentile
75th percentile
90th percentile
average
Source: Lifetime Labour Market Database

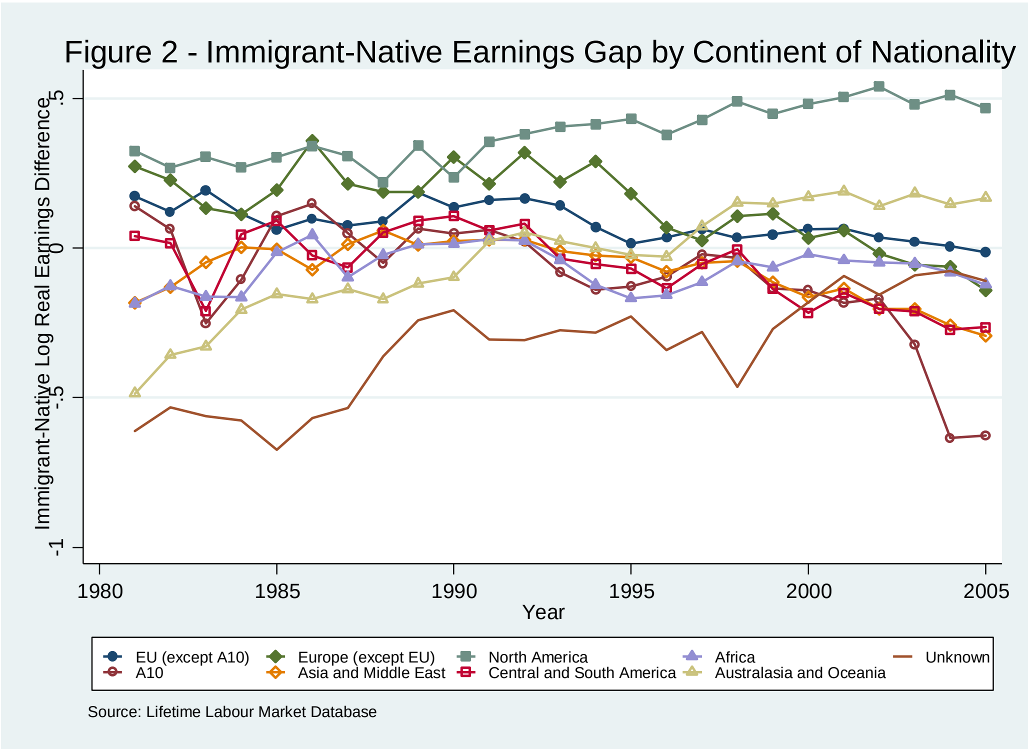


Figure 2 - Immigrant-Native Earnings Gap by Continent of Nationality
Immigrant-Native Log Real Earnings Difference
-1
-.5
0
.5
1980
1985
1990
1995
2000
2005
Year
EU (except A10)
A10
Europe (except EU)
Asia and Middle East
North America
Central and South America
Africa
Australasia and Oceania
Unknown
Source: Lifetime Labour Market Database

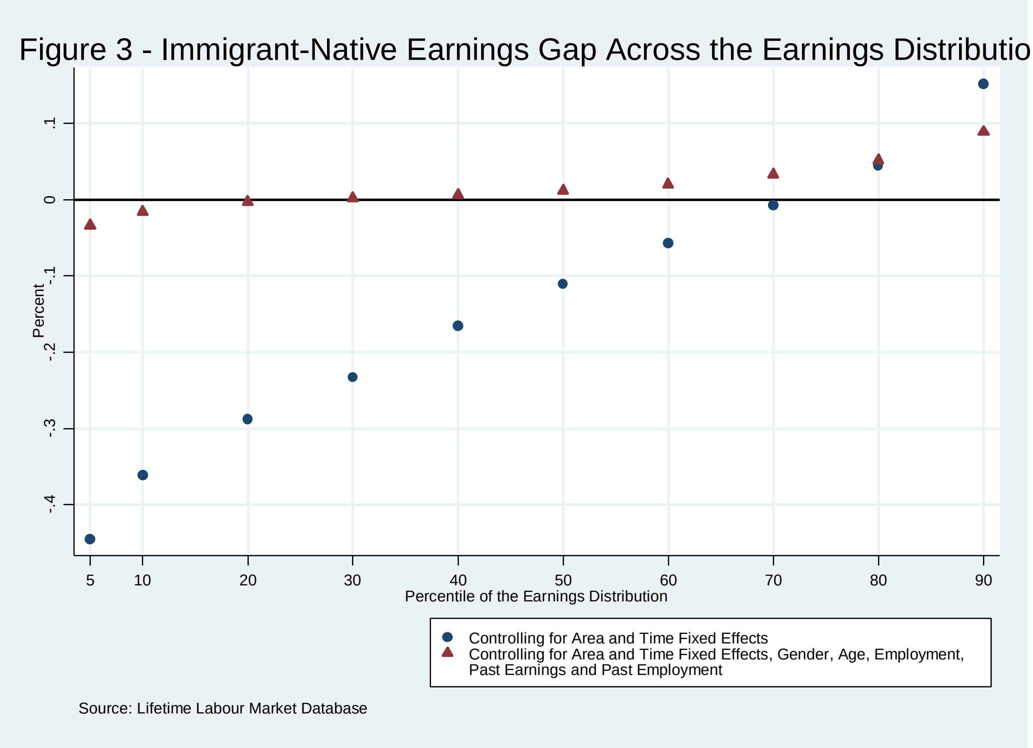


Figure 3 - Immigrant-Native Earnings Gap Across the Earnings Distribution
Percent
Percentile of the Earnings Distribution
Controlling for Area and Time Fixed Effects
Controlling for Area and Time Fixed Effects, Gender, Age, Employment, Past Earnings and Past Employment
Source: Lifetime Labour Market Database

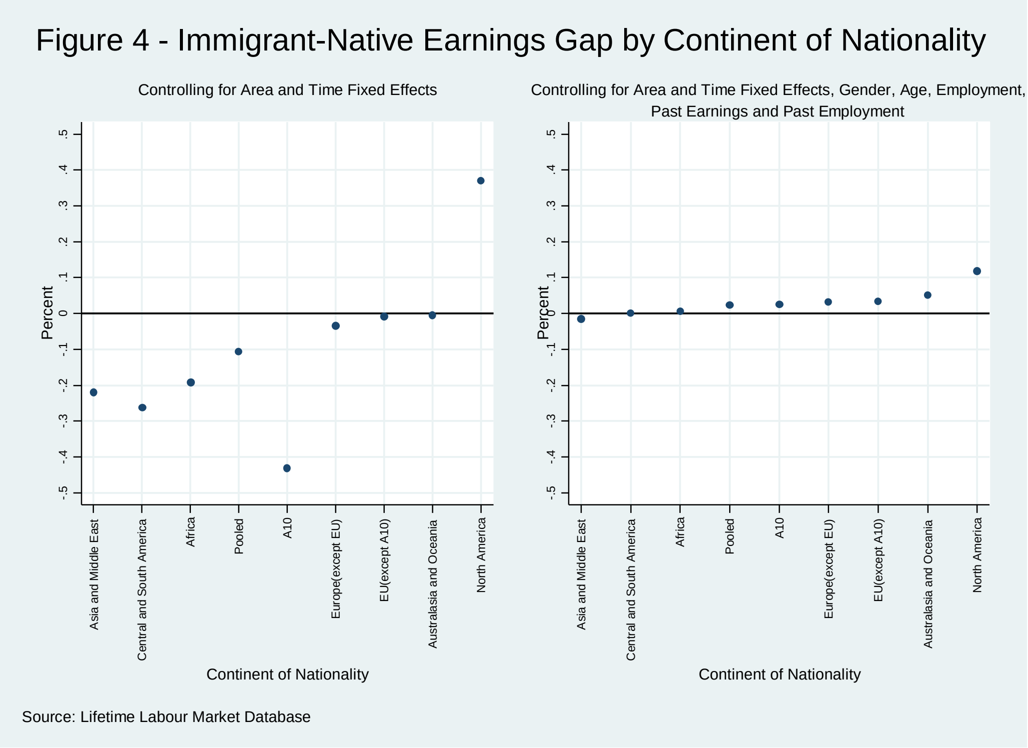
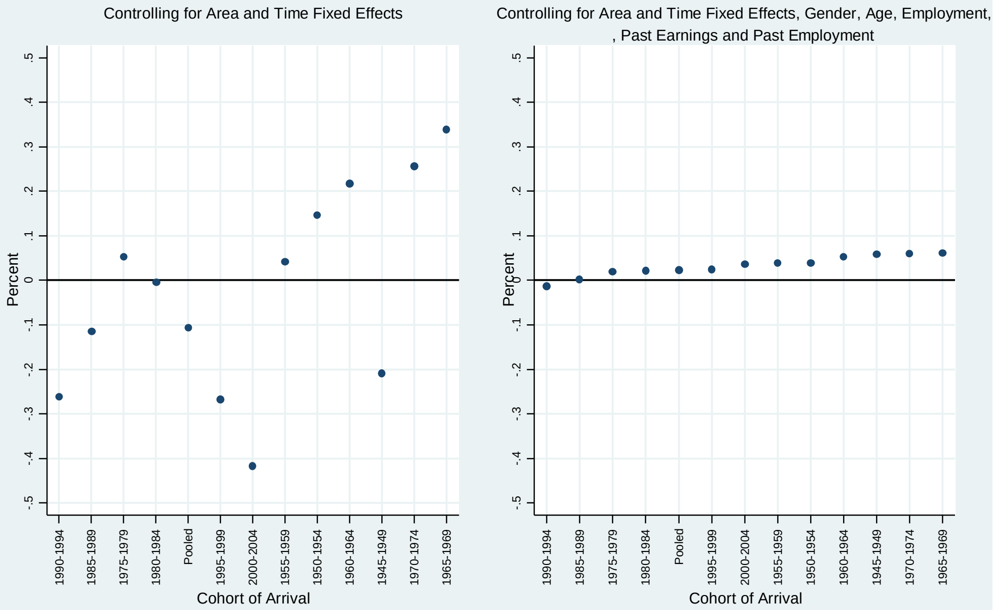


Figure 4 - Immigrant-Native Earnings Gap by Continent of Nationality
Controlling for Area and Time Fixed Effects
Controlling for Area and Time Fixed Effects, Gender, Age, Employment, Past Earnings and Past Employment
Percent
Asia and Middle East
Central and South America
Africa
Pooled
A10
Europe(except EU)
EU(except A10)
Australasia and Oceania
North America
Continent of Nationality
Source: Lifetime Labour Market Database

Figure 5 - Immigrant-Native Earnings Gap by Cohort of Arrival



Source: Lifetime Labour Market Database