University of Kent

School of Economics Discussion Papers

English Deficiency and the Native-Immigrant Wage Gap

Alfonso Miranda and Yu Zhu

September 2012

KDPE 1213

English Deficiency and the Native-Immigrant Wage Gap

Alfonso Miranda[a]

Yu Zhu[b,*]

Abstract

We focus on the effect of English deficiency on the native-immigrant wage gap for employees in the UK using the first wave of the UK Household Longitudinal Survey (Understanding Society). We show that the wage gap is robust to controls for age, region of residence, educational attainment and ethnicity, particularly for men. However, English as Additional Language (EAL) is capable of explaining virtually all the remaining wage gap between natives and immigrants. Using the interaction of language of country of birth and age-at-arrival as instrument, we find strong evidence of a causal effect of EAL on the native-immigrant wage gap.

JEL classification: J15, J61

Keywords: native-immigrant wage gap, English as Additional Language (EAL), age-at-arrival

---

[a] Economics Division, Center for Research and Teaching in Economics (CIDE), Mexico. *Email: alfonso.miranda@cide.edu.*
[b] School of Economics, University of Kent, Canterbury, CT2 7NP, UK
[*] Corresponding author: Tel.: +44-1227-827438; fax: +44-1227-827850; *Email:* yz5@kent.ac.uk (Y. Zhu).
Note: A short version of the paper focusing exclusively on male results is forthcoming in the ***Economics Letters***.

1 Introduction

This paper investigates to what extent English deficiency explains the native-immigrant wage gap for employees in the UK, after controlling for age, region of residence, educational attainment and ethnicity. Previous work in the field have found that, upon immigration, good command of the language of the destination county carries a positive return of around 15-25% for migrants in the labour market (see e.g. Chiswick 1991, Chiswick and Miller 1999, Dustmann 1994). In the UK Dustmann and Fabbri (2003) report a return of about 18-20% using the Family and Working Lives Survey (FWLS) as well as the Fourth National Survey on Ethnic Minorities (FNSEM). However, after controlling for the potential endogeneity of English proficiency they find no statistically significant effect. This may be due to the small sample size the wage equations were fitted on, with 250 observations for the FWLS and 920 observations for the FNSEM.

Our contribution to the literature is threefold. The Dustmann and Fabbri (2003) study analysed data collected in the first half of the 1990s. Since that time, the UK has received a significant inflow of new migrants from Eastern Europe, following the expansion of the European Union. Hence, it is interesting to revisit the topic and investigate whether the returns to English proficiency remain at the same level as those reported nearly 20 years ago. Second, our analyses are based on a larger sample of immigrants, enabling us to estimate tighter confidence intervals. Finally, we use an IV strategy to address the issue of potential endogeneity of our English language proficiency indicator. Endogeneity may arise from two different sources: (i) self-selection into treatment, and (ii) measurement error. Innate skill, an unobservable, is expected to increase wage and at the same time to raise the probability that an individual will self-select into the English proficient group.

This causes a positive self-treatment. In addition, measurement error of the English proficiency variable causes yet another problem of endogeneity which induces a negative (attenuation) bias. Hence, the total endogeneity bias is a combination of the two aforementioned problems (see Dustmann and Fabbri 2003). We address the 'aggregate' endogeneity problem using an IV strategy, without making an attempt to disentangle the contribution of the two sources of bias. Following Bleakley and Chin (2004, 2010) and van Ours and Veenman (2006), we use language of the origin country as well as its interaction with age-at-arrival as instrumental variables for the subpopulation of immigrants. Our identification strategy effectively compares older and younger arrivers from non-English-speaking countries, after controlling for an age-at-arrival effect which is the same for all immigrants regardless of their native language. This identification strategy delivers a Local Average Treatment Effect (LATE) that is straightforward to interpret for the subpopulation of (first-generation) immigrants affected by the instrument and offers a suitable and meaningful control group. Hence, this is an improvement to alternative identification strategies that have been used in the past.

The remainder of the paper is organised as follows. Section 2 introduces the data and sets up the anaysis. The empirical results are presented and discussed in Section 3. Finally, Section 4 concludes.

2 Data and set-up of the analysis

Our empirical analysis is based on the first wave of the UK Household Longitudinal Survey, also known as Understanding Society. This is a longitudinal survey of just over 30,000 households in the UK over the period 2009-2011, including around 4,000 from the ethnic minority boost sample. The survey contains not only information on ethnicity and country of birth of the immigrant and both parents, but also crucially direct measures of English proficiency. Moreover, the large sample also allows analysis of immigrants at a rather disaggregate level, by for example gender and whether born in the UK. So this is an ideal data for the study of the effect of language on labour market outcomes of immigrants.

We focus on the native-immigrant wage gap of employees aged 19-60. Natives are defined as ethnic whites who were born in the UK to two UK-born parents, and who speak English as first language. Conversely, immigrants are defined as people who were born abroad to two non-native parents. We only include first-generation immigrants (i.e. non UK-born) in the treatment group, in order to exploit the variation in English deficiency induced by the variation in the age-at-arrival of immigrants from non-English-speaking versus English-speaking countries.[1] After excluding missing values and outliers of log real hourly wage,[2] we end up with a sample of 6,959 males and 8,423 females, of which 1,203 and 954 respectively are immigrants.

Table 1 reports summary statistics by gender and immigrant status. It is evident that there are systematic differences across gender in our outcome variable, the native-immigrant wage gap, and its potential determinants. So we follow the standard practice in labour economics of analysing males and females separately throughout the paper.

In the upper panel of Table 1, the raw native-immigrant wage gap for men is a highly statistically significant 0.159 log points (or 17.2 percentage point)[3] in favour of natives.

1 Blackaby et al. (2005) have studied the employment and earnings differentials between whites and UK-born ethnic minorities using a decomposition approach.

2 We exclude the top and bottom 1% of wages by the level of highest qualification within each ethnicity and gender cell.

3 A β log points gap can be transformed into differences in percentages by using the formula: 100*(exp(β)-1). For small values of β (say less than 0.20), 100*β gives a reasonable approximation of the actual percentage change.

However, this does not take into account any composition effect. Indeed, these two samples have very different characteristics. For instance, male immigrants in the UK are on average younger, hold higher qualifications, and live disproportionately in London compared to white natives. 71% of the male immigrants declare speaking English as Additional Language (EAL), while 87% of them were born in developing countries. Whereas all natives are white by construction, there is significant heterogeneity in the ethnicity composition of male immigrants, with 55% classified as Asians and 14% as blacks. A significant minority (18.5%) of male immigrants are white, of which 71% were born in European countries.[4]

In contrast, the raw native-immigrant wage gap for women in the bottom panel of Table 1 is only a statistically insignificant 0.017 log points. On average, female immigrants are also younger, better educated, and more likely to live in London, compared to their native counterparts. Almost 62% of female immigrants are classified as (speaking) EAL, whereas nearly three-quarters of them were from developing countries.

In Figure A1, we explore alternative measures of English deficiency. In our survey, if a person declares EAL, questions are asked about whether she has difficulty in speaking day-to-day English, difficulty in speaking on the phone, difficulty reading English, and difficulty completing forms in English. For each of those four aspects of English difficulties, the degree of difficulty is also asked, with possible answers of a little difficult, fairly difficult, very difficult and cannot speak (read) at all.

It turns out that of all first-generation immigrants who declare EAL, only 23% of men and 20% of women report having any difficulty in English, with the highest incidence in

4 Whites account for a larger share of more recent immigrants. For example, 27% of immigrants arriving in the UK since 2004 (the year the expansion of the European Union took place) are white, of which 61% are from Poland.

reading and the lowest incidence in speaking on the phone for both genders. When we convert the degree of difficulty into scores with 1 for a little difficult and 4 for cannot speak (read) at all, the total mean score is only 4.9 for men and 4.0 for women for immigrants who report having any difficulty. This implies that even for those who report having difficulties in English, the mean level of English deficiency is not much more than finding it a little difficult in each aspect of the language. The low incidence of self-reported difficulties and the low degree of difficulty suggest that there might be considerable measurement errors in this highly subjective measure of language deficiency.

In the following section, we will explore the extent to which the native-immigrant wage gap depends on the inclusion of various controls, and in particular, on how EAL helps to explain the composition-adjusted gap.

3 Results and discussions

3.1 Least Squares

In a wage equation, we measure the immigrant wage effect by a dummy variable, with a negative coefficient indicating a regression-adjusted native-immigrant wage gap in favour of natives. In Table 2 we successively introduce control variables. The male native-immigrant wage gap increases by 0.10 log points when differences in age profiles and region of residence are accounted for in column 1. Moreover, once highest qualifications as well as a dummy for highest qualification is obtained abroad are controlled for in column 2 the gap widens by another 0.02 log points.[5] When ethnicity dummies are added in column 3, the gap is reduced to 0.117 log points, indicating that the adjusted native-

5 The interaction terms of the highest qualifications dummies with the foreign dummy are jointly insignificant at any conventional level.

immigrant wage gap is only one-quarter smaller than the raw differential.[6] This implies that even white male immigrants (the reference ethnicity category) suffer from large wage penalties compared to male natives, holding age, region and education constant.

We then explore to what extent English deficiency explains this remaining wage gap in the next two columns. When EAL is added in column 4, the gap becomes a statistically insignificant 0.018 log points in favour of immigrants. This implies that all remaining wage gap is explained by English deficiency. When we further include a dummy for migrating from a developing country[7] and dummies for age-at-arrival in the UK for immigrants (column 5), the immigrant coefficient becomes positive and statistically significant at 0.132 log points, while the EAL effect remains significant and of the same magnitude as the raw wage gap. We include age, age square and age-at-arrival in bands of 0-9, 10-15, 16-29 and 30+ to disentangle the effect of assimilation and effects of language (note that there is perfect multicollinearity between age, age-at-arrival and years living in the UK).[8] The fact that all age-at-arrival dummies are strongly negative suggests that for white immigrants arriving in the UK before 10 (the omitted category), there is no disadvantage associated with being an immigrant, holding all else constant.

Table 2B shows that accounting for differences in age profiles and region of residence also increases the native-immigrant gap by 0.10 log points for women. Further controlling for highest qualifications in column 2 makes virtually no difference. Additionally controlling for ethnicity in column 3 renders the female native-immigrant

---

6 The fact that the native-immigrant wage gap for both men and women is reduced by about two-thirds when ethnicity control is added is consistent with labour market disadvantage (discrimination) associated with being minority ethnicity. We choose to condition on race as we want to disentangle the effect of English deficiency from other characteristics.

7 The developing country dummy might partly capture the differences in quality of education.

8 However, our identification only relies on the ***interaction*** between born in a non-English speaking country and age-at-arrival greater than 9 (following e.g. Bleakley & Chin (2004, 2010)).

statistically insignificant again. Adding EAL control in column 4 also completely drives away any remaining native-immigrant wage gap, which is now slightly in favour of immigrants (although statistically insignificant). Finally, including a dummy for being born in a developing country and age-at-arrival dummies in column 5 leaves the language effect about one-third smaller but still statistically significant at the 10% level.

To sum up, we find strong evidence of a large and statistically significant native-immigrant wage gap for both men and women, after accounting for effect of age profile, region of residence and highest qualifications. Further controlling for ethnicity reduces the gap by about two-thirds for both genders, but the gap remains statistically significant for men. Moreover, English deficiency as measured by EAL is capable of explaining the entire remaining regression-adjusted native-immigrant wage gap. While the native-immigrant wage gap responds to the different sets of controls in a very similar pattern across genders, the size of the effect of EAL for women is only about 40% as large as that for men.

3.2 Two Stage Least Squares (2SLS)

In Table 3 we address the potential endogeneity of EAL, by estimating the *Two Stage Least Squares (2SLS)* model with the same set of control variables on the pooled natives and immigrants sample. To facilitate comparison between OLS and 2SLS estimates, we copy the final column of Table 2 into the first column of Table 3.

In column 2, we instrument EAL using born in a non-English-speaking country as well as its interaction with a dummy for age-at-arrival greater than 9. Figure 1A and 1B show the regression-adjusted means of probability of EAL, with 95% confidence intervals, by age-at-arrival and language of home country for men and women separately.[9] Immigrants from non-English-speaking countries who arrived before the age of 10-14 for men are, statistically, as likely to be EAL as immigrants from English-speaking countries. In contrast, if immigration occurred after age 10-14 the two groups are statistically different. In the case of women, the critical age is before and after age 5-9. These finding are consistent with Bleakley & Chin (2010) who use an age-at-arrival cut-off at 10 in their preferred specification of English proficiency. Therefore, in line with previous work, we use age 10 as the critical cut-off point to implement our IV estimator.[10]

Following the theory of the critical period for second language acquisition Bleakley and Chin (2004, 2010) argue that, after controlling for educational attainment and other background variables, differences in English proficiency between immigrants from English-speaking and non-English-speaking countries before and after the critical age are uncorrelated with current wages because any non-language age-at-arrival wage effects are the same for all immigrants regardless of their home country language. If this hypothesis is correct, as it is our view, the interaction term between language of country of origin and age-at-arrival is a valid instrument for EAL in the wage equations because it is correlated with current wages only through the channel of English deficiency as measured by the EAL status. As a consequence, the IV estimator is consistent. Notice that under the postulated identification strategy the IV estimator is analogous to a difference-

9 These are effectively predicted probabilities based on a linear probability model of EAL on age-at-arrival dummies interacted with a born in non-English-speaking country dummy and controls for age, age squared, region of residence, highest qualification and ethnicity. The patterns are robust to the exclusion of controls.

10 Figure A2A and A2B show the corresponding regression-adjusted means of probability of any difficulty in English by age-at-arrival and language of home country, for men and women respectively. While the overall pattern is the same as for EAL, there is a lack of precision, presumably due to the greater noise with this self-reported measure.

in-differences estimator that calculates language wage effects net of age-at-arrival wage effects.[11] Hence, we are able to disentangle language and age-at-arrival wage effects.

The first-stage estimates in the bottom panels of Table 3A shows that the instruments are strong predictors of EAL status both individually and jointly (F-stat of 249). Being born in a non-English-speaking country increases the incidence of EAL for male immigrants by 33%. However, this is dominated by the effect of arriving in the UK after age 9 from a non-English-speaking country (i.e. the interaction term), which increases the probability by another 49%. Moreover, the model also easily passes the over-identification test, with a p-value of 0.89. We find that the EAL effect becomes larger when we allow it to be endogenous but remains statistically significant at the 5% level, although we won't be able to reject the null of equality with the OLS estimate.

Column 3 shows that instrumenting using the interaction term only makes virtually no difference for male immigrants. The implied causal effect of EAL in column 3 is -0.253 log points (or a -22.4%), compared to -0.251 log points (or -22.2%) in column 2. A male immigrant from a non-English-speaking country arriving in the UK after age 9 is 76 percentage points more likely to declare EAL than another male who immigrated either by age 9 or from an English-speaking country (or both). Note that the effect of the interaction term alone on EAL almost captures the entire combined effect of the interaction and the language of home country in the over-identified model. Apart from a much higher F-stat (of 360) compared to the over-identified model, the just-identified IV is also unlikely to be subject to a weak-instrument critique. GMM and LIML estimates

11 Basically there are four groups: (a) immigrants from English-speaking countries arrived to the UK before age 10-14, (b) immigrants from English-speaking countries arrived to the UK after age 10-14, (c) immigrants from non-English-speaking countries arrived to the UK before age 10-14, and (d) immigrants from English-speaking countries arrived to the UK after age 10-14. The language wage effect, net of age-of-arrival wage effects, is the wage DiD between groups ((d)-(c))-((b)-(a)).

(available upon request) also come out very similar, giving further support to the robustness of our IV results (Angrist and Pischke 2009).

Table 3B reports the corresponding 2SLS estimates for women. Again the results are similar to those for men, but weaker in both magnitude and statistical significance. The 2SLS estimate in the over-identified model is a statistical insignificant -0.071 log points, which is of the same magnitude as the OLS effect. The just-identified IV effect is a much larger -0.124 log points, which is also significant at the 10% level. However, a low p-value of 0.028 for the over-identification test casts some doubt on the validity of the instruments. It could be due to the smaller proportion of immigrants in the female subsample. Another likely cause is the selectivity of women into salaried employment, which might differ between white natives and immigrants as well as within immigrants by ethnicity or culture.[12]

In Table 4, we rerun OLS and 2SLS using the subsample of immigrants only. The EAL coefficient in the OLS specification for males is −0.164, virtually unchanged from its full sample counterpart of -0.158, whereas the coefficients on born in developing country dummy and age-at-arrival dummies all maintain their magnitude and statistical significance. This implies that the EAL effect is not driven by systematic differences in characteristics between natives and immigrants (the composition effect). Moreover, the IV estimates are virtually indistinguishable across the two tables, as the 2SLS estimates for both the first and second stage in both the over-identified model and the just identified model differ by at most 0.02 log points. This fits well with our story that the causal effect is identified by variation ***within*** the sub-population of immigrants in English deficiency

12 This issue, although important, is beyond the scope of this paper.

induced by age-at-arrival between immigrants from English-speaking and non-English-speaking countries.

The results for women reported in Table 4B are broadly similar to those in Table 3B, and again with larger and statistically more significant IV estimate in the just-identified model. Another thing worth emphasizing is that the p-value for the identification is now above the critical value of 0.05, suggesting that the use of the instruments are more appropriate when the female sample is restricted to first-generation immigrants.

4 Conclusions

We start by showing that the raw native-immigrant wage gap in the UK is a highly significant 15.9% for men and an insignificant 1.7% for women. Controlling for differences in age profile and region of residence increases this gap by 10 percentage points for both men and women, making it statistically significant at any conventional level for both genders. Interestingly, further controlling for the highest qualification makes little difference to the wage gap for either men or women.

In order to focus on the effect of language deficiency, we further condition on ethnicity. We find a composition-adjusted native-immigrant wage gap for male employees in the UK of 11.7%, slightly below the raw wage differential. For women, the adjusted native-immigrant wage gap is 3.8%, which is above the raw wage gap but still statistically insignificant. However, both gaps virtually disappear after controlling for the EAL indicator.

We address the potential endogeneity of EAL with an IV strategy and use born in non-English-speaking country and age-at-arrival as instruments for identifying a LATE that is straightforward to interpret for the subpopulation of first-generation immigrants affected by the instrument and offers a meaningful control group. Our IV regressions indicate that EAL has a causal effect of -24% on wages for male immigrants, which is significant at 5%, and a marginally significant -12% for female immigrants. The causal effect of EAL on the native-immigrant wage gap is robust to various specifications.

The size of the effect of English deficiency we find in our more recent data is comparable to studies based on surveys conducted in the early 1990s. This implies that the large inflow of immigrants following the EU expansion in 2004 has not significantly affected the returns to English proficiency in the UK labour market.

It is worth noting that the effect of English deficiency is conditional on the highest educational qualification, which is often attained by the immigrant after arriving in the UK. Recently Dustmann et al. (2010) singled out improved English proficiency as the most important factor why ethnic minority pupils improve relative to White British pupils in the compulsory education stage which ends at age 16, using the National Pupil Database (NPD) and the Millennium Cohort Studies (MCS). To the extent that late arrival from a non-English-speaking country (i.e. our IV) will have an adverse effect on educational attainment, our IV estimate can be regarded as a lower bound (i.e. biased towards zero) of the gross effect of language deficiency. Further research is needed before we can have a better understanding of all the channels through which language deficiency impacts labour market outcomes.

References

Angrist, J.D., Pischke, J.S. 2009. Mostly Harmless Econometrics – An Empiricist's Companion. Princeton University Press, Princeton.

Blackaby, D.H., Leslie, D.G., Murphy, P.D., O'Leary, N.C., 2005. Born in Britain: How are Native Ethnic Minorities Faring in the British Labour Market? Economics Letters 88, 370-375.

Bleakley, H., Chin, A., 2004. Language skills and earnings: evidence from childhood immigrants. Review of Economics and Statistics 86, 481 – 496.

Bleakley, H., Chin, A., 2010. Age-at-arrival, English proficiency, and social assimilation among US immigrants. American Economic Journal: Applied Economics 2, 165–192.

Chiswick, B.R., 1991. Speaking, reading, and earnings among low-skilled immigrants. Journal of Labor Economics 9, 149-70.

Chiswick, B.R., Miller, P.W., 1999. Language skills and earnings among legalised aliens. Journal of Population Economics 12, 63-89.

Dustmann, C., 1994. Speaking fluency, writing Fluency and earnings of migrants. Journal of Population Economics 7, 133-56.

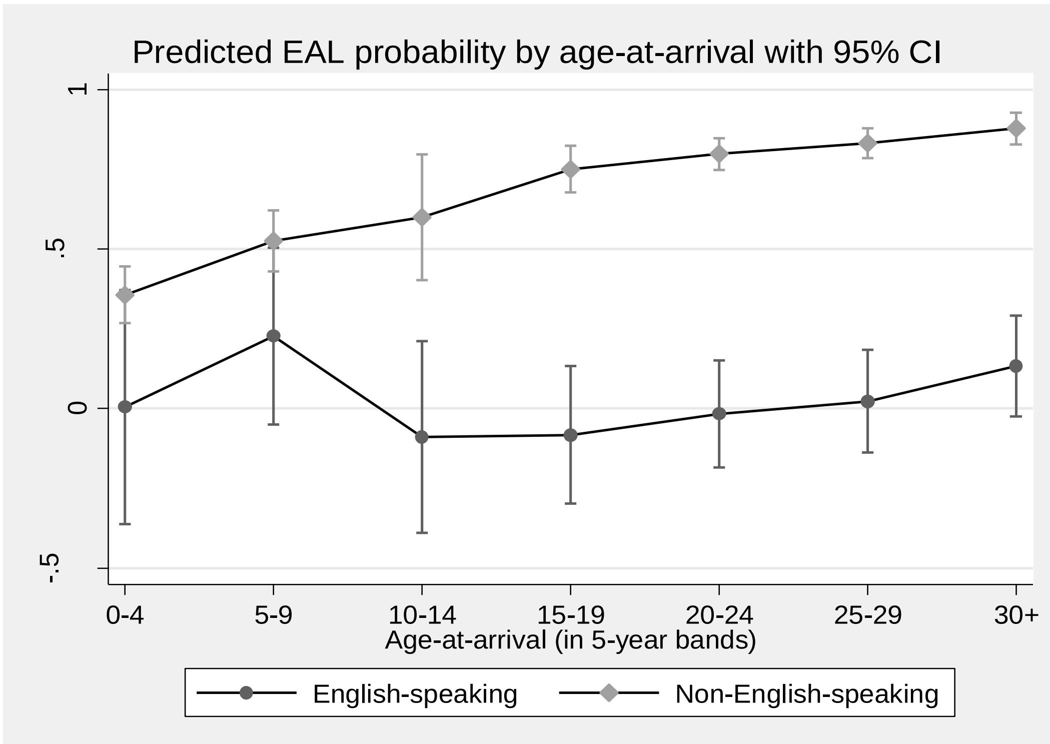
Dustmann, C., Fabbri, F. 2003. Language proficiency and labour market performance in the UK. Economic Journal 113, 695-717.

Dustmann, C., Machin, S. & Schönberg, U. 2010. Ethnicity and Educational Achievement in Compulsory Schooling. Economic Journal 120, F272-F297.

Van Ours, J. C., Veenman, J., 2006. Age at immigration and educational attainment of young immigrants. Economics Letters 90, 310-316.

Figure 1: Regression-adjusted EAL probability by age-at-arrival and home country language

1A) Males



1B) Females

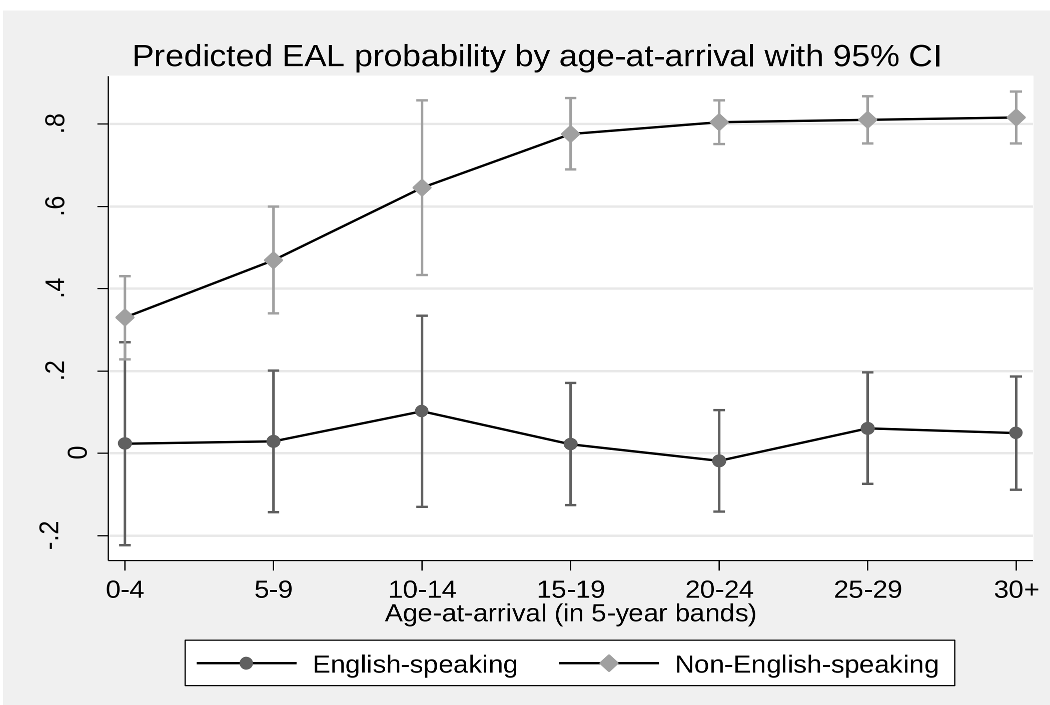


Table 1: Summary statistics, by immigrant status

1A) Males (N=6959)

	Natives (N=5756)	Immigrants (N=1203)	Native-immigrant gap
Log real hourly wage	2.504	2.345	0.159**
EAL	0	0.710	-0.710**
Born in developing country	0	0.874	-0.874**
No qualification	0.141	0.162	-0.022*
Below GCSE/O-Level	0.094	0.076	0.017**
GCSE/O-Level	0.286	0.081	0.206**
A-Level	0.119	0.097	0.022**
Higher Education Diploma	0.089	0.086	0.004
First Degree	0.173	0.212	-0.039**
Higher Degree	0.097	0.285	-0.187**
Highest qualification is foreign	0.002	0.392	-0.389**
Age	40.3	37.5	2.8**
White	1.000	0.185	0.815**
Mixed	0	0.013	-0.013**
Asian	0	0.552	-0.552**
Black	0	0.140	-0.140**
Other Ethnicity	0	0.111	-0.111**
London	0.061	0.485	-0.424**
Southeast	0.132	0.069	0.063**
Rest of England	0.615	0.391	0.223**
Wales	0.051	0.009	0.042**
Scotland	0.098	0.030	0.068**
Northern Ireland	0.044	0.016	0.028**

1B) Females (N=8423)

	Natives (N=7469)	Immigrants (N=954)	Native-immigrant gap
Log real hourly wage	2.289	2.272	0.017
EAL	0	0.616	-0.616**
Born in developing country	0	0.743	-0.743**
No qualification	0.129	0.142	-0.012
Below GCSE/O-Level	0.091	0.070	0.021**
GCSE/O-Level	0.273	0.116	0.157**
A-Level	0.108	0.089	0.018*
Higher Education Diploma	0.139	0.151	-0.011
First Degree	0.180	0.211	-0.031**
Higher Degree	0.080	0.221	-0.141**
Highest qualification is foreign	0.003	0.383	-0.379**
Age	41.0	38.7	2.3**
White	1.000	0.280	0.720**
Mixed	0	0.020	-0.020**
Asian	0	0.357	-0.357**
Black	0	0.238	-0.238**
Other Ethnicity	0	0.105	-0.105**
London	0.048	0.480	-0.432**
Southeast	0.129	0.104	0.025**
Rest of England	0.619	0.349	0.270**
Wales	0.053	0.018	0.035**
Scotland	0.095	0.025	0.070**
Northern Ireland	0.054	0.024	0.030**

Note: **(*) = significant at 5% (10%) level based on Welch's t-test.

Table 2: Log-wage equations, pooled natives and immigrants sample

2A) Males (N=6959)

	(1)	(2)	(3)	(4)	(5)
Immigrant	-0.258 (13.7)**	-0.276 (14.0)**	-0.117 (3.2)**	0.018 (0.4)	0.132 (2.3)**
EAL				-0.198 (6.0)**	-0.158 (4.4)**
Born in developing country					-0.092 (1.8)*
Age-at-arrival 10-15					-0.136 (2.2)**
Age-at-arrival 16-29					-0.090 (2.0)**
Age-at-arrival 30+					-0.172 (3.2)**
Highest qualification dummies	no	yes	yes	yes	yes
Ethnicity dummies	no	no	yes	yes	yes

2B) Females (N=8423)

	(1)	(2)	(3)	(4)	(5)
Immigrant	-0.119 (6.3)**	-0.118 (6.1)**	-0.038 (1.3)	0.018 (0.5)	0.064 (1.3)
EAL				-0.107 (3.5)**	-0.066 (1.9)*
Born in developing country					-0.130 (3.4)**
Age-at-arrival 10-15					0.025 (0.4)
Age-at-arrival 16-29					-0.015 (0.3)
Age-at-arrival 30+					-0.063 (1.1)
Highest qualification dummies	no	yes	yes	yes	yes
Ethnicity dummies	no	no	yes	yes	yes

Note: Absolute t-statistics based on robust standard errors in parentheses; **(*) = significant at 5% (10%) level. Other controls include age, age squared and region dummies.

Table 3: Two Stage Least Squares Estimates, pooled natives and immigrants sample

3A) Males (N=6959)

	OLS	IV (over-identified)	IV (just-identified)
Immigrant	0.132(2.3)**	0.151 (2.5)**	0.151 (2.5)**
EAL	-0.158 (4.4)**	-0.253 (2.4)**	-0.251 (2.3)**
Born in developing country	-0.092 (1.8)*	-0.061 (1.0)	-0.062 (1.0)
Age-at-arrival 10-15	-0.136 (2.2)**	-0.110 (1.6)	-0.111 (1.6)
Age-at-arrival 16-29	-0.090 (2.0)**	-0.056 (1.0)	-0.057 (1.0)
Age-at-arrival 30+	-0.172 (3.2)**	-0.135 (2.0)**	-0.136 (2.1)**
p-value for over-identification test		0.887	-
First stage (dependent variable = EAL)			
Born in non-English-speaking country		0.330 (2.7)**	
Born in non-English-speaking country * (age-at-arrival>9)		0.487 (4.1)**	0.764 (19.0)**
F-stat for exclusion restrictions (p-value)		248.9 (0.000)	359.7 (0.000)

3B) Females (N=8423)

	OLS	IV (over-identified)	IV (just-identified)
Immigrant	0.064(1.3)	0.065 (1.3)	0.068 (1.4)
EAL	-0.066 (1.9)*	-0.071 (1.0)	-0.124 (1.6)*
Born in developing country	-0.130 (3.4)*	-0.128 (2.9)**	-0.109 (2.4)**
Age-at-arrival 10-15	0.025 (0.4)	0.027 (0.4)	0.039 (0.6)
Age-at-arrival 16-29	-0.015 (0.3)	-0.013 (0.3)	0.007 (0.1)
Age-at-arrival 30+	-0.063 (1.2)	-0.061 (1.0)	-0.041 (0.7)
p-value for over-identification test		0.028	-
First stage (dependent variable = EAL)			
Born in non-English-speaking country		0.436 (6.8)**	
Born in non-English-speaking country * (age-at-arrival>9)		0.400 (6.9)**	0.735 (20.2)**
F-stat for exclusion restrictions (p-value)		256.6 (0.000)	408.9 (0.000)

Note: Absolute t-statistics based on robust standard errors in parentheses; **(*) = significant at 5% (10%) level. Other controls include age, age squared, region dummies, highest qualification dummies and ethnicity dummies.

Table 4: Two Stage Least Squares Estimates, immigrants sample only

4A) Males (N=1203)

	OLS	IV (over-identified)	IV (just-identified)
EAL	-0.164 (4.5)**	-0.232 (2.2)**	-0.242 (2.3)**
Born in developing country	-0.106 (2.0)**	-0.085 (1.4)	-0.081 (1.3)
Age-at-arrival 10-15	-0.164 (2.5)**	-0.145 (2.1)**	-0.142 (2.0)**
Age-at-arrival 16-29	-0.099 (2.1)**	-0.076 (1.3)	-0.072 (1.3)
Age-at-arrival 30+	-0.151 (2.8)**	-0.125 (1.8)*	-0.121 (1.8)*
p-value for over-identification test		0.612	-
First stage (dependent variable = EAL)			
Born in non-English-speaking country		0.332 (2.8)**	
Born in non-English-speaking country * (age-at-arrival>9)		0.490 (4.3)**	0.767 (18.8)**
F-stat for exclusion restrictions (p-value)		228.0 (0.000)	353.9 (0.000)

4B) Females (N=954)

	OLS	IV (over-identified)	IV (just-identified)
EAL	-0.075 (2.2)**	-0.080 (1.1)	-0.127 (1.7)*
Born in developing country	-0.141 (3.6)*	-0.140 (3.2)**	-0.124 (2.7)**
Age-at-arrival 10-15	0.056 (0.9)	0.057 (0.9)	0.068 (1.0)
Age-at-arrival 16-29	-0.012 (0.3)	-0.011 (0.2)	0.006 (0.1)
Age-at-arrival 30+	-0.065 (1.1)	-0.063 (1.0)	-0.046 (0.7)
p-value for over-identification test		0.061	-
First stage (dependent variable = EAL)			
Born in non-English-speaking country		0.452 (6.6)**	
Born in non-English-speaking country * (age-at-arrival>9)		0.384 (6.4)**	0.730 (19.4)**
F-stat for exclusion restrictions (p-value)		231.5 (0.000)	378.0 (0.000)

Note: Absolute t-statistics based on robust standard errors in parentheses; **(*) = significant at 5% (10%) level. Other controls include age, age squared, region dummies, highest qualification dummies and ethnicity dummies.

APPENDIX

Figure A1: Fractions of immigrants with difficulties in English, EAL=1

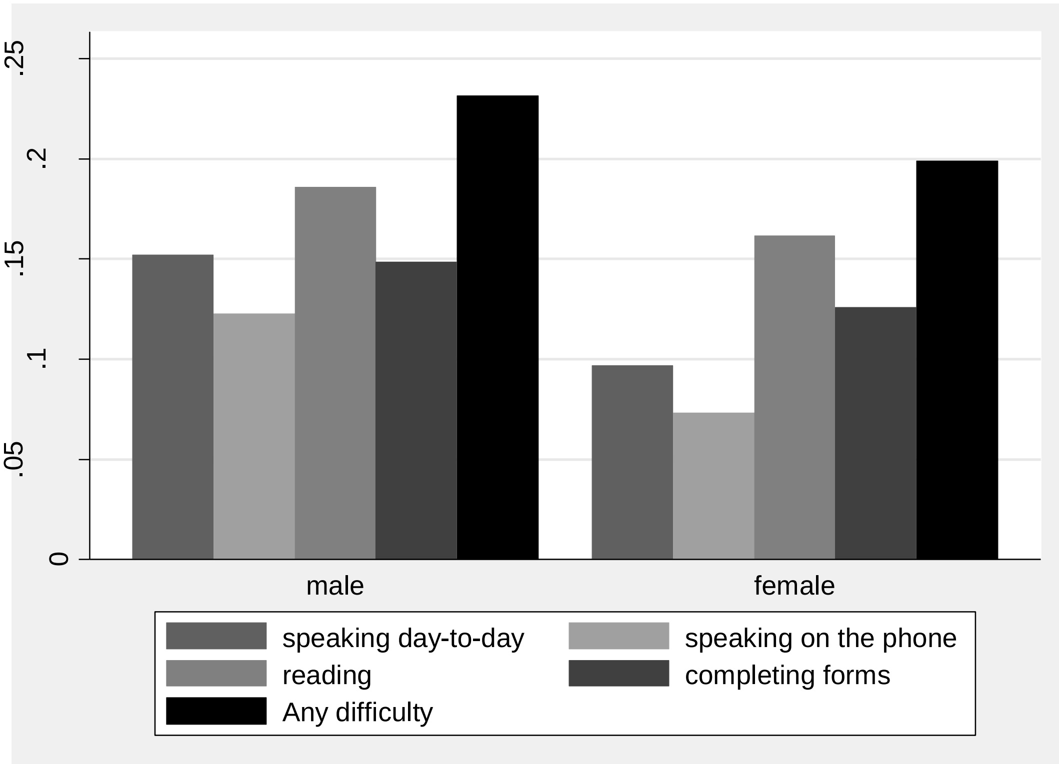
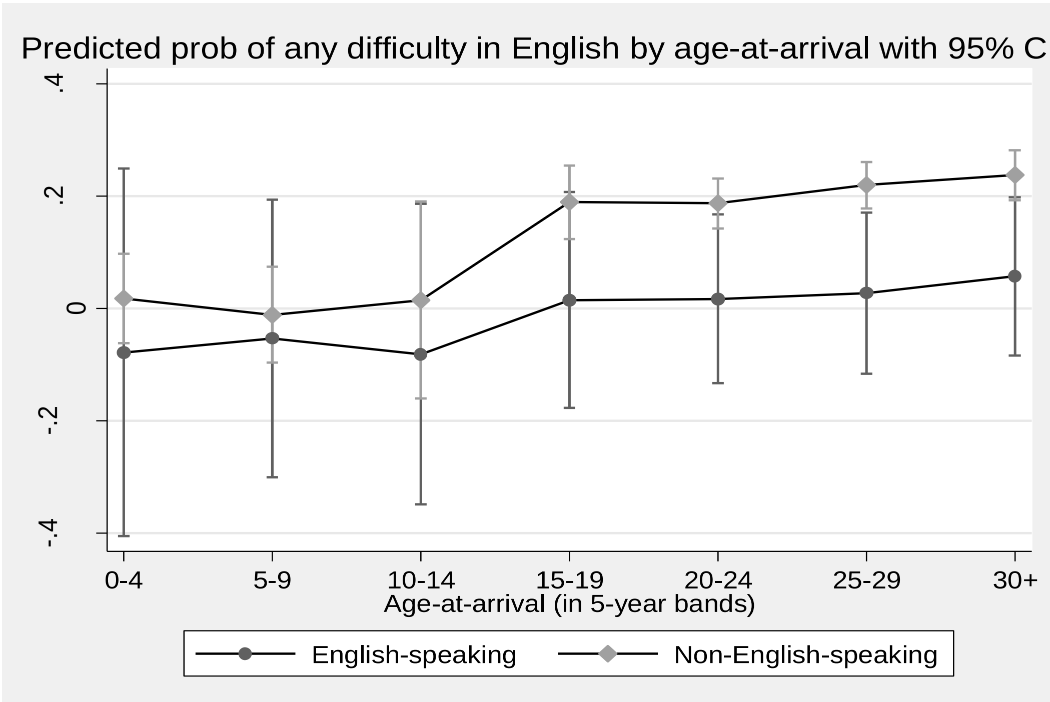


Figure A2: Regression-adjusted probability of any difficulty in English by age-at-arrival and home country language

A2A) Males



A2B) Females