Bilingualism in the labour market

Joanna Clifton-Sprigg and Kerry L. Papps*

Previous research has found that among the native-born population, bilingual people earn less in the U.S. labour market. We examine whether a similar pattern exists in the U.K. and attempt to provide an explanation. We find that bilingual men do no worse than monolingual men, but that bilingual women earn significantly less than monolingual women. This is not explained by differences in cultural background, parental education or other family background variables. The result also holds when we control for various degrees of bias in unobserved characteristics. Instead, the result appears to be driven by differences across areas in the prevalence of bilingualism, with the negative earnings effects restricted to bilingual women living in areas with relatively low proportions of English speakers. The negative effects of bilingualism on women are also concentrated among speakers of South Asian languages and relatively uncommon languages.

1. Introduction

Bilingualism is increasingly common across the developed world. Although reliable statistics are scarce due to varying definitions of bilingualism, for example about 6% of students taking the PISA 2009 assessment in the O.E.C.D. countries spoke another language at home (O.E.C.D.: Social Policy Division, Directorate of Employment, Labour and Social Affairs, 2012). The phenomenon has triggered a debate about its implications for individuals, the economy and society as a whole. Foreign, as well as European, languages feature in many aspects of the U.K. and E.U. education and labour market policies. The United Nations, UNESCO, the Council of Europe and the E.U. all emphasise the right of an individual to maintain their heritage language and culture and highlight the role of linguistic proficiency in achieving labour market success (European Commission 1995). This reflects a conviction that linguistic ability plays a role in shaping one's identity, social integration, employability and productivity, which directly influence the well-being and economic performance of a society as a whole.

* Clifton-Sprigg: J.M.Clifton-Sprigg@bath.ac.uk, Department of Economics, University of Bath, Bath, BA2 7AY, United Kingdom, tel: +44-1225-384581. Papps: K.L.Papps@bath.ac.uk, Department of Economics, University of Bath, Bath, BA2 7AY, United Kingdom, tel: +44-1225-384338. The authors would like to thank participants at the CReAM/RWI Workshop on the Economics of Migration (September 2017), the IFO Workshop on Labour Market and Social Policy (April 2017), the European Society of Population Economics Annual Conference 2017, the Colloquium for Socio-economic Value and Impacts of Languages, the EEA Annual Congress 2018, SOFI Brown Bag seminar and University of Western Australia Business School seminar for their helpful suggestions.

Within the U.K. there has been concern about a lack of foreign language skills in the labour market at a time when these are becoming more valuable in the global economy (Tinsley 2013). According to the 2011 U.K. Census, 7.7% of the population of England and Wales reported speaking a language other than English as their main language (Office for National Statistics 2011). However, the number of students studying foreign languages at school and university has been falling, as students increasingly choose science, technology, engineering and mathematics subjects instead.

Despite this, there is little empirical evidence on the labour market returns to foreign language skills, whether learned at school or at home. Research to date has focused predominantly on the schooling and labour market performance of first- and second-generation immigrants.[1] These studies focus on establishing the existence of gaps in labour market performance between native and immigrant populations and analysing factors which contribute to closing these disparities. Knowledge of the host country language has been identified by previous studies as a significant determinant of immigrant labour market outcomes, but few authors have examined the value of speaking an additional language.

Given the number of factors that may affect a person's labour market outcomes, the role of bilingualism is far from clear. Acquisition of more than one language early on in life may be seen as an investment in one's general human capital (Chiswick and Miller 2016), leading to an increase in cognitive and non-cognitive skills (Carneiro et al. 2013; Bak et al. 2014). The returns to human capital investments in early childhood are particularly high and persistent and are expected to significantly contribute to an adult's performance in the labour market (Heckman 2008). In this case, one would expect bilinguals to earn more than comparable monolinguals on average. However, bilingualism is likely to be correlated with many other characteristics that affect wages, such as family socio-economic status, cultural background and where a person lives. Many of these are typically unobservable by the researcher.

In this paper, we use data from the Understanding Society survey to compare the earnings of U.K.-born individuals who spoke a language other than English at home as children with those of their monolingual counterparts and examine possible

[1] For example Dustmann and Fabbri (2003), Bleakley and Chin (2004), Rooth and Saarela (2007), and Yao and van Ours (2015) consider the role of fluency in the language of the destination country; Blau et al. (2011) and Blau et al. (2013) analyse differences by country of origin, gender and intergenerational transmissions; and Bisin et al. (2011) investigate the influence of ethnicity.

explanations for this relationship. We are able to control for a richer set of control variables than most previous studies, allowing us to eliminate the confounding effect of family and cultural background, which may be correlated with language skills. In addition, we use a recently-developed technique to adjust for the potential bias caused by selection on unobservable characteristics. We consider men and women separately, acknowledging that bilingualism is known to have different effects on the learning opportunities of boys and girls (Lee and Hatteberg 2015) and that each gender faces a very different labour market.

We focus exclusively on the native population of the U.K. and assess the labour market returns to fluency in languages other than English. The native-born population who are fluent in English and another language ("native bilinguals") is not comparable to first-generation immigrants. Although many native bilinguals have an immigrant background, they were born in the U.K. Therefore, they have been exposed to the same cultural and institutional environment as their monolingual counterparts. In particular, both groups received education in the U.K. This is particularly useful for two reasons. Firstly, unlike the majority of immigrants, we can assume that all respondents are fluent in English. Thus, our focus is on the gains from the ability to speak a second language, rather than the penalty associated with an insufficient knowledge of English. Secondly, we can eliminate different institutional or educational systems as factors potentially confounding the relationship studied.[2]

We find that bilingual women have lower earnings than comparable monolingual women, but that there are no significant differences among men. This pattern is not explained by differences in the country of birth of a person's parents, differences in parental education or other measures of family background. However, the negative effects of bilingualism on earnings among women are found to be concentrated among women who live in areas with a high fraction of non-English speakers. One explanation for this is that there may be limited employment opportunities in such areas. A negative

[2] When analysing a relationship between one's ability to speak several languages and his/her performance in the labour market, it is highly likely that many other factors (e.g. culture, different quality of education, social norms) which are correlated with bilingualism also affect one's employability and income. If they are not explicitly accounted for, the alleged relationship we are investigating may also be capturing these other influences, preventing any conclusions about the sole role of language. By comparing individuals born and brought up in the same country, we control for a wide range of such associated factors, which may otherwise confound the relationship.

earnings effect is only found among bilingual women who speak South Asian or less-common languages.

In the next section, we review the relevant literature on bilingualism in economics and other disciplines. We describe the dataset we use in Section 3, before presenting our results in Section 4, starting with estimates of the overall wage effect of being bilingual, before considering heterogeneity between different groups of people. Some concluding comments are given in Section 5.

2. Background

The role of language in labour market outcomes has long been acknowledged in the economic literature. Most studies focus on the economic performance of first- and second-generation immigrants, who often lack proficiency in the language of the host country. Analyses for first-generation immigrants demonstrate the existence of an earnings gap between natives and immigrants, which closes with the duration of stay in the host country. Some studies (Bleakley and Chin 2004; Chiswick and Miller 1995; Dustmann and van Soest 2001; Yao and van Ours 2015) focus more specifically on linguistic skills and find that they are associated with better labour market performance. There is also evidence of differences in language acquisition, probability of employment and earnings depending on the ethnic origins of immigrants and the strength of their ethnic identity (Dustmann and Fabbri 2003; Bisin et al. 2011). This literature also acknowledges the role of age at arrival and argues that a significant part of the positive effect found among young immigrants is channelled through schooling (Bleakley and Chin 2004).

Measurement error inherent in people's assessments of their linguistic skills and endogeneity in the relationship between language and earnings pose the main estimation challenge, which is often addressed by an instrumental variables approach (Chiswick and Miller 1995; Dustmann and van Soest 2002; Bleakley and Chin 2004). Rooth and Saarela (2007) propose an alternative way of eliminating bias by considering the labour market outcomes of Finns in Sweden. In doing so they are comparing outcomes of immigrants of the same nationality, some of whom are native speakers of Swedish (the host country language) and others are native speakers of Finnish. This allows them to minimise the concerns about measurement error in linguistic ability and eliminate the role country of origin may play in outcomes. The approach is similar in spirit to ours in that the two groups studied differ solely in terms of language spoken and are otherwise comparable. Their results provide even higher estimates of the positive effect of immigrants' language proficiency on earnings.

The research focusing on native-born individuals, to which our paper contributes, has predominantly focused on second-generation immigrants in the U.S. (Chiswick and Miller 2018; Fry and Lowell 2003) or on native populations in countries with several official languages or dialects (e.g. Carliner (1981) for Canada, Paolo and Raymond (2012) and Rendon (2007) for Catalonia, Chiswick et al. (2000) for Bolivia, Yao and van Ours (2019a, 2019b) for the Netherlands, Henley and Jones (2005) for Wales). These studies find mixed results with no overall returns found in the US and Welsh cases, negative estimates for the Netherlands and positive associations for Catalonia.

Our work is closest in spirit to that of Chiswick and Miller (2018). They investigate the labour market returns to bilingualism among natives using U.S. data, focusing solely on males. They conclude that, although overall bilingualism is negatively associated with earnings, there is significant heterogeneity within the group of bilinguals and across quantiles of income, with the returns to some languages being positive.

Work in linguistics and cognitive psychology points towards skill enhancement due to language acquisition as a major driver of differences in earnings and employability. Baker (1999) provides an extensive overview of the impacts of bilingualism on cognitive outcomes in children, which typically translate into working life performance. Bilinguals seem to have an advantage in certain thinking dimensions, such as divergent thinking, creativity and metalinguistic awareness (Blumenfeld and Marian 2009), selective attention and inhibitory control (Bialystok et al. 2009). The ability to speak several languages may also delay onset of dementia (Bak et al. 2014). At the same time, however, it has been found that bilinguals may have a slower response time and make more errors in vocabulary tests focused on word retrieval. This may be reflected in speech production and is thought to be related to the processing of two languages at the same time (Bialystok et al. 2009). The research so far has found no correlation between bilingualism and I.Q. (Kaushanskaya and Marian 2007).

The demand for particular language skills in the labour market is an alternative explanation for any positive association. Certain languages may be in particular demand, either due to the high frequency of their use locally or due to the nature of the business one is employed in. For example, an individual who fluently speaks a foreign language may be rewarded by his/her employer if many of the person's customers speak that language. Conversely, bilingual workers may suffer a wage penalty if discrimination against foreign language speakers occurs. This may take the form of direct pay discrimination against bilinguals in areas where a lot of workers (and potentially a lot of employers) speak other languages. Alternatively, it may be indirect and due to firms choosing not to locate in areas with higher fractions of non-English speakers, resulting in low levels of labour demand.

Identifying whether one or more of the aforementioned mechanisms explain any differences in outcomes between bilingual and monolingual individuals is complicated by the fact that language is also highly correlated with culture. Language may influence behaviour (Hicks et al. 2015), including labour force participation (Alesina et al. 2013), through its links to culture and impact on cognitive processes, rather than through an increased skill set or labour market demand.

The ability to speak more than one language may, through cultural links, affect labour market prospects of females differently than those of males. For instance, exposure to language and culture of a country with traditional gender roles may lead to lower labour market participation by females (Gay et al., 2018). In case of bilinguals, the knowledge of a second language is usually a consequence of having strong ties with other countries, typically the countries of birth of a person's parents or grandparents. This family background may reinforce traditional gender roles or embrace modern attitudes to female labour market participation.

For these reasons it is important to differentiate between genders and to control for cultural factors which may be correlated with bilingualism and directly affect labour market outcomes. We do so by running separate regressions for males and females and controlling for mothers' country of birth dummies in the regressions.

3. Data

We use data from Understanding Society, which is an annual longitudinal study of 40,000 households in the U.K., capturing information about respondents' demographic characteristics, socio-economic circumstances, attitudes, behaviour and health. Our data come from waves 1 to 5 of the survey, which were conducted between 2009 and 2014.

Information on the language a person spoke at home during childhood was collected during wave 2 of Understanding Society. Speaking a language other than English is strongly related to immigration. Figure 1 shows the fraction of Understanding Society respondents who spoke another language as a child, by immigrant generation. 69% of first generation immigrants (those who were born outside the U.K.) spoke another language, with little difference in rates between men and women. However, among the second generation (those who were born in the U.K. but whose parent(s) were born elsewhere), the fraction drops to 29%. Among the third generation (those born in the U.K. with U.K.-born parents but a grandparent born elsewhere) and fourth generation (where the respondent and his/her parents and grandparents were all born in the U.K.), only 8% and 9%, respectively, spoke a language other than English as a child.

Since there is no information in Understanding Society on English proficiency, we exclude the first generation from our analysis, because they may have poor English skills. However, immigrants born in the U.K. should have been educated in English and be native speakers. Therefore, an individual is identified as a bilingual in the sample if he/she reported speaking a language other than English at home during childhood. We further focus on individuals aged 16-65 and therefore part of the labour force. This leaves us with a sample of 10,239 female and 8,393 male respondents who are observed over the five survey waves, once we drop observations with missing values for the main regression specification used in the next section. In total, we have 26,566 and 21,277 person-year observations on women and men, respectively.

We use gross monthly pay, which is a derived variable, explicitly provided in the survey. It measures income from all sources of employment. We adjust for inflation using annual C.P.I. data, so that everything is expressed in 2009 pounds. In addition to Understanding Society data, we merge in data on the proportion of people (aged 3 and over) who speak English in the local area (specifically, the middle-layer super output area or M.S.O.A.) from the 2011 U.K. Census.

Some summary statistics can be found in Table 1. Bilingual women and men earn slightly less than their monolingual counterparts. However, the two groups differ systematically in terms of their demographic characteristics. The bilingual sample is younger than their monolingual counterparts, has more children and is more likely to be married. Bilinguals also tend to live in areas with much lower fractions of people who speak English.

4. Analysis

To begin with, we estimate the following specification (which is similar to that used by Chiswick and Miller), separately for employed men and women:

$$\ln EARNINGS_{it} = \alpha BILINGUAL_{i} + \mathbf{X}_{it}\boldsymbol{\gamma} + u_{it}, \tag{1}$$

where ln *EARNINGS* is the log of monthly gross pay of person *i* in year *t*, *BILINGUAL* is a dummy variable for whether the person spoke a language other than English at home as a child and **X** is a vector of control variables. We weight by the inverse of the number of observations for each person in the sample.

The results of estimating equation 1 are presented in Table 2. Initially, we use a similar set of control variables to Chiswick and Miller (2018); specifically, age and age squared, education dummies (completed A-level equivalent, undergraduate degree, postgraduate degree), a dummy for being married, number of children, race dummies (Asian, black, other, mixed), a London dummy and year dummies. As seen in the first and third columns, respectively, bilingualism is associated with a 5.2% reduction in earnings among women but has no significant relationship with earnings among men, in contrast to the findings of Chiswick and Miller.

The results for women cannot be interpreted as a causal effect, however, because bilingualism reflects differences in the choices made by the respondents' parents regarding which language(s) to expose the respondents to when they are young. If the choice to raise a child as bilingual is correlated with other family characteristics that affect the child's future employment outcomes, this will introduce bias to the estimates. Bilingualism might be more common among immigrants from less developed countries, who would do worse in the U.K. labour market anyway. Furthermore, even within immigrant groups, bilingualism might be more common among less educated families, where the parents may not be able to speak English well and the children might be expected to acquire less human capital than children with educated parents.

To test whether bilingualism captures differences in the labour market outcomes of immigrants from different countries, we add fixed effects for a person's mother's country of birth to **X**.[3] To control for the influence of parental education, we add dummies for whether the respondent's mother and father had a high school or university education. In this specification, the effects of bilingualism are estimated from differences in outcomes between bilingual and monolingual people whose parents had the same education level and came from the same country. As seen in the second column of Table 2, rather than being wiped out, the coefficient on the bilingual dummy for women becomes stronger and is now significant at the 1% level. The coefficient for men (in the fourth column) remains insignificant. For women, mother's education, but

[3] Very similar results are found when the father's country of birth is controlled for instead.

not father's education, has a significant positive effect on earnings. For men, both mother's education and father's education (at least as far as high school) increases earnings.

There are additional characteristics that might be correlated with the number of languages a person speaks. For example, socially-conservative women may be more likely to speak a second language, but also be likely to earn lower wages. To control for this, we added dummies for whether a person was Christian or of another religion in the first column of Table A1. The sample size drops because information on religion is not available for all respondents. For women, being of a non-Christian religion raises wages by 8%. The wage penalty associated with bilingualism is slightly bigger than in Table 2 and remains significant. For men, religion has no effect on earnings and the coefficient on being bilingual remains insignificant.

A person's social group might also be important, given evidence on the importance of ethnic networks in labour market outcomes (Damm 2012). To capture this, we control for the fraction of a person's friends who are of same ethnic group. Possible answers are "all the same", "more than half", "about half" or "less than half". In the second and fourth columns of Table A1, we add a dummy variable for whether a person responded with the third or fourth of these categories. This is associated with significantly lower earnings among both men and women. However, its inclusion makes little difference to the coefficient on bilingualism.

To examine the effects bilingualism has at the extensive margin, we repeat the specifications from Table 2 using a dummy variable for whether a person is employed as the dependent variable. As shown in Table A2, bilingualism is associated with a lower likelihood of a woman being employed and a higher likelihood of a man being employed. This is robust to the inclusion of controls for mother's country of birth and parents' education.

Heterogeneity

Next, we consider what factors influence the magnitude of the wage penalty found for bilingual women in Table 2. As noted in the previous section, bilinguals can be divided into second generation immigrants and third or fourth generation immigrants, according to whether at least one parent was born in the U.K. In the first and third columns of Table 3, we allow the effect of bilingualism to vary according to a person's generation. The wage penalty for second generation bilingual women is 8.8%. Third or fourth generation bilingual women earn slightly more, but the difference is insignificant (most likely due to the small number of bilinguals who are third or higher generation). The earnings of bilingual men of either generation are not statistically significantly different from those of monolingual men.

The wage effects associated with bilingualism might also depend on the characteristics of the local labour market in which the person lives. As noted in Section 2, this may lead to either upward or downward bias in the bilingualism coefficient. If a lot of other people speak the same language in the local area, a bilingual person may have more opportunities in the labour market to exploit his/her language skills. However, if the presence of a lot of non-English-speakers leads to pay discrimination, either directly or by crowding out better-paying jobs because companies choose not to locate in the area, bilingualism might be associated with lower wages.

To investigate this possibility, the fraction of people speaking English in the local area was added to the regression in the first and third columns of Table 3, alongside the interaction of this variable and the bilingual dummy. For women, the fraction speaking English has no significant effect, but the interaction term does have a significant positive effect. This indicates that bilingual women experience lower wages if they live in areas with relatively few English speakers, but this effect dissipates as the fraction speaking English increases. For women in areas at the 10^{th} percentile of English-speaking fraction (81%), the coefficient on bilingualism is -0.078; whereas for women at the 90^{th} percentile (99%), the coefficient is 0.008 and is not significantly different from zero. For men, neither the uninteracted nor the interacted bilingualism terms are significant. However, the fraction speaking English has a significant positive effect, indicating that all men do better if they live in areas with many English speakers.[4]

Correction for selection

The results in Tables 2 and A1 indicate that the effects of bilingualism on earnings are robust to the inclusion of a variety of controls for family and cultural background. However, it is still possible that being bilingual reflects selection on some unobserved characteristics which also affect the labour market performance of individuals. For example, being bilingual may be a proxy for a person's preference for a traditional

[4] A similar pattern of results was found if the fraction of people in the local area speaking the respondent's own language was used instead.

division of labour in the household, even among others of the same religion, with similar ethnic networks and with parents of the same education level and from the same country. To shed light on the extent to which selection on unobservable characteristics poses a concern we consider the extent to which it would alter the OLS results. We follow the work of Altonji et al. (2005) and Oster (2019) and make assumptions about the degree of selection on unobservables relative to the degree of selection on observed variables in the regression to produce alternative, bias-adjusted estimates. Altonji et al. (2005) provide a method to correct for selection on unobservables, using the degree of selection on observed variables in the regression as a guide. They, and subsequently Oster (2019), argue that it is plausible that degrees of selection on unobservable and observable variables in a regression are equal (δ = 1) provided that the observable individual characteristics are just a random subset of a greater group of variables important for the outcome. The point estimates based on this assumption constitute an upper bound on OLS.

Oster (2019) provides an extension to this work, noting that the unexplained variation in the outcome can be decomposed into two elements – idiosyncratic and one driven by unobserved characteristics. One can thus impose a further upper bound (R_{max}) on the R^2 from the regression of the dependent variable on all observed and non-idiosyncratic unobserved variables, effectively narrowing the bound on the OLS estimates of interest.[5] Both approaches allow also for estimation of the degree of selection on unobservables required to eliminate the treatment effect.

We present estimates of β obtained using methods by Oster (2019) assuming R_{max} = 0.3 and considering a range of values of δ.[6] We use the regression specification in the second and fourth columns of Table 2. The results can be found in Table 4. They indicate that, regardless of the choice of δ, a negative effect of bilingualism is found for women. However, among men, when δ is above 0.3, the bilingual coefficient becomes positive. This implies that selection on unobserved characteristics larger than 30% of the degree of selection on observed characteristics in the regression would be sufficient to mask the negative effect of bilingualism on male earnings found earlier. Hence, the

[5] Note that in Altonji et al. (2005) the assumption was that if all variables were observed, the model would be fully explained and R_{max} would be equal to 1.

[6] The choice of the maximum explanatory power assumed in the regression is justified by the fact that the highest R^2 obtained in the regressions here does not exceed 0.25. Therefore, the assumption that the techniques correcting for bias can explain as much as 30% of the variation in the dependent variable is generous and more realistic than assuming 100% of the variation would have been explained.

negative earnings effect for women found in Table 2 appears to be robust to selection on unobservable characteristics. However, under a relatively small degree of selection on unobservable characteristics, the earnings effect for men might in fact become positive.

Results for separate languages

The results presented so far provide evidence of the overall labour market effects of bilingualism. However, it is reasonable to imagine that there will be substantial heterogeneity in this by language group, reflecting the characteristics of different immigrant communities and the different levels of demand for their languages. In Table 5, the bilingual dummy is replaced by separate dummies for which language a person spoke as a child. These dummies capture whether a person spoke include a native U.K. language (Welsh or Gaelic), any European language, Arabic, Chinese, a South Asian language (Gujarati, Bengali, Punjabi or Urdu) or any other language.

The first and third columns include these dummies alongside the basic control variables used in equation 1. The results indicate that the overall negative effect of bilingualism among women is driven by women speaking South Asian and other languages. No significant effects of bilingualism are found for any language among men. Adding controls for family background (in the second and fourth columns) makes little difference to the results and the negative earnings effect associated with speaking a South Asian language or other language persists among women. These results indicate that, for example, among women with well-educated Pakistani-born parents, Urdu speakers fare particularly badly in the labour market, compared to women who speak only English.

Given the results in Table 3 that the negative effects of bilingualism are concentrated among those who live in areas with relatively low fractions of non-English speakers, the results in Table 5 suggest that there may be linguistic “enclaves” where bilingual women do particularly poorly in the labour market. This does not appear to be due to the presence of unobserved person-specific characteristics (as seen in Table 4). Instead, it may be due to differential patterns of labour demand or of discrimination across local labour markets.

5. Conclusion

In this paper, we have used data from the Understanding Society survey in the U.K.

for 2009-2014 to examine the effects bilingualism has on a person's labour market outcomes. We find no evidence that bilingualism leads to higher earnings, as would be expected if it contributed to a person's human capital. Instead, bilingualism is associated with lower earnings among females and has no effect on males. The negative effect for women does not appear to be driven by either observable or unobservable differences in a person's family background. However, neighbourhood effects are found to be important. Specifically, bilingual women do significantly worse if they live in areas with few English speakers, perhaps due to discrimination. There is also significant variation in the returns to speaking specific languages. Among women, speaking South Asian or less common languages is associated with significantly lower earnings, while speakers of other languages do not earn significantly less than monolinguals. Among men, no language has a significant effect on earnings.

References

Adsera, A., and B. R. Chiswick (2007): "Are there gender and country of origin differences in immigrant labor market outcomes across European destinations?", *Journal of Population Economics*, 20, 495-526.

Alesina, A., P. Giuliano, and N. Nunn (2013): "On the Origins of Gender Roles: Women and the Plough", *The Quarterly Journal of Economics*, 128, 469-530.

Altonji, J. G., et al. (2005): "Selection on Observed and Unobserved Variables: Assessing the effectiveness of Catholic Schools", *Journal of Political Economy*, 12, 151-184.

Bak, T. H., J. J. Nissan, M. M. Allerhand, and I. J. Deary (2014): "Does bilingualism influence cognitive ageing?", *Annals of Neurology*, 75, 959-963.

Baker, C. (1999): *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters, Bristol.

Bialystok, E., et al. (2009): "Bilingual minds", *Psychological Science of Public Interest*, 10, 89-129.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2011): "Ethnic identity and labour market outcomes of immigrants in Europe", *Economic Policy*, 26, 57-92.

Blau, F. D., L. M. Kahn, A.-H. Liu, and K. L. Papps (2013): "The transmission of women's fertility, human capital and work orientation across immigrant generations", *Journal of Population Economics*, 26, 405-435.

Blau, F. D., L. M. Kahn, and K. L. Papps (2011): "Gender, source country characteristics and labour market assimilation among immigrants", *Review of Economics and Statistics*, 98, 43-58.

Bleakley, H., and A. Chin (2004): "Language skills and earnings: evidence from childhood immigrants", *The Review of Economics and Statistics*, 86, 481-496.

Blumenfeld, H., and V. Marian (2009): "Language-cognition interactions during bilingual language development in children", in Recent Trends in Education, ed. by B. Kuzmanovic, and A. Cuevas, chap. 3. Nova Science Publishers.

Cannadine, D. (2019): "Brexit Britain cannot afford to be laissez-faire about its language crisis", *The Guardian*, accessed on 18.04.2019 via: https://www.theguardian.com/commentisfree/2019/mar/01/britain-learning-languages-brexit--education.

Carliner, G. (1981): "Wage Differences by Language Group and the Market for Language Skills in Canada", *The Journal of Human Resources*, 16, 384-399.

Carneiro, P., C. Meghir, and M. Parey (2013): "Maternal education, home environments, and the development of children and adolescents", *Journal of the European Economic Association*, 11, 123-160.

Chiswick, B. R., and P. W. Miller (1995): "Endogeneity between language and earnings: international analyses", *Journal of Labor Economics*, 13, 246-288.

Chiswick, B. R., and P. W. Miller (2018): "Do native-born bilinguals in the US earn more?" *Review of Economics of the Household*, 16, 563-583.

Chiswick, B. R., H. Patrinos, and M. Hurst (2000): "Indigenous language skills and the labour market in a developing economy: Bolivia", Economic Development and Cultural Change, 48, 349-367.

Damm, A. P. (2012): "Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence", in *Recent Developments in the Economics of International Migration: Volume 1: Immigration: Flows and Adjustment*, ed. by B. R. Chiswick and P. W. Miller. Cheltenham: Edward Elgar Publishing.

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 A. van Soest (2001): "Language fluency and earnings: estimation with misclassified language indicators", *The Review of Economics and Statistics*, 83, 663-674.

Dustmann, C., and A. van Soest (2002): "Language and the Earnings of Immigrants", *Industrial and Labor Relations Review*, 55.

European Commission (1995): "White paper on education and training", accessed on 20/05/20 via: https://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf.

Fry, R., and L. Lowell (2003): "The value of bilingualism in the US labor market," *Industrial and Labor Relations Review*, 57, 128-143.

Gay, V., D. L. Hicks, E. Santacreu-Vasut, and A. Shoham (2018): "Decomposing culture: Can gendered language influence women's economic engagement?", *Review of Economics of the Household*, 16(4), 879-909.

Heckman, J. (2008): "Schools, skills, and synapses", *Economic Inquiry*, 46, 289-324.

Henley, A., and R. E. Jones (2005): "Earnings and linguistic proficiency in a bilingual economy", The Manchester School Working Paper Series, 73, 300{320.

Hicks, D. L., E. Santacreu-Vasut, and A. Shohan (2015): "Does mother tongue make for women's work? Linguistics, household labor and gender identity", *Journal of Economic Behaviour and Organisation*, 110, 19-44.

Isphording, I., et al. (2016): "Speaking in Numbers: The Effect of Reading Performance on Math Performance Among Immigrants", *Economics Letters*, 139, 52-56.

Isphording, I., and S. Otten (2014): "Linguistic Barriers in Destination Language Acquisition of Immigrants", *Journal of Economic Behavior and Organization*, 105, 30-50.

Jeffreys (2019): "Language learning: German and French drop by half in UK schools", *BBC*, accessed on 18/04/19 via: https://www.bbc.com/news/education-47334374.

Kaushanskaya, M., and V. Marian (2007): "bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming", *Language Learning*, 57, 119-163.

Lee, J.C., and S.J. Hatteberg (2015): "Bilingualism and Status Attainment among Latinos", *The Sociological Quarterly*, 56(4):695-722.

O.E.C.D.: Social Policy Division, Directorate of Employment, Labour and Social Affairs (2012): "C03.6 Percentage of immigrant children and their educational outcomes", accessed on 19/07/16 via: www.oecd.org/els/soc/49295179.pdf.

Office for National Statistics (2011): "Language in England and Wales: 2011", accessed on 20/05/20 via: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/articles/languageinenglandandwales/2013-03-04.

Oster, E. (2019): "Unobservable selection and coefficient stability: theory and validation", *Journal of Business Economics and Statistics*, 37(2):187-204.

Paolo, A. D., and J. L. Raymond (2012): "Language knowledge and earnings in Catalonia", *Journal of Applied Economics*, 15, 89-118.

Pendakur, K., and R. Pendakur (2002): "Speaking in tongues: language knowledge as human capital and ethnicity", *International Migration Review*, 36, 147-178.

Rendon, S. (2007): "The Catalan premium: language and employment in Catalonia", *Journal of Population Economics*, 20, 669-686.

Rooth, D., and J. Saarela (2007): "Native language and immigrant labour market outcomes: an alternative approach to measuring the returns for language skills", *Journal of International Migration and Integration*, 8, 207-221.

Tinsley, T. (2013): "Languages: The State of the Nation", Report for British Academy.

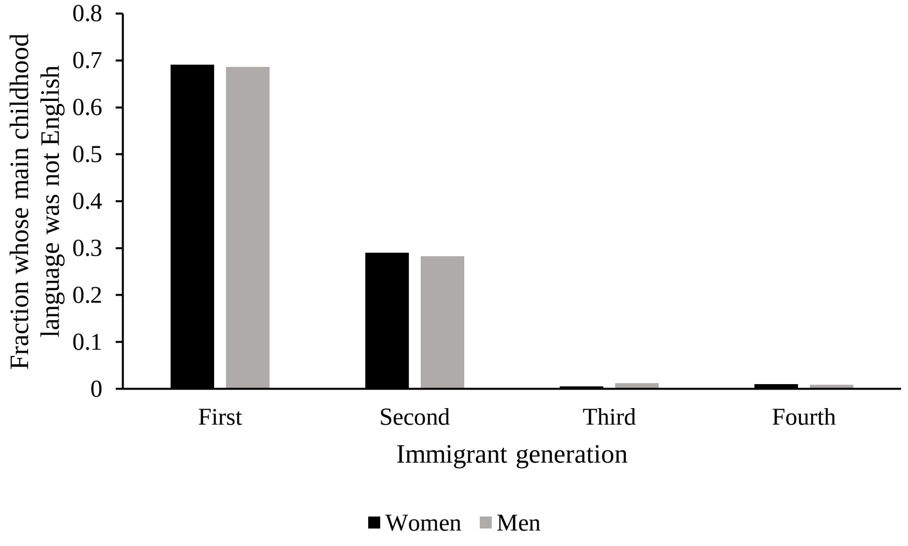
Wacziarg, R., and E. Spolaore (2016): "Ancestry, Language and Culture," in *The Palgrave Handbook of Economics and Language*, ed. by V. Ginsburgh, and S. Weber, chap. 6. Palgrave Macmillan, London.

Yao, Y., and J. C. van Ours (2015): "Language skills and labour market performance of immigrants in the Netherlands", *Labour Economics,* 34, 76-85.

Yao, Y., and J.C. van Ours (2019a): "Dialect speech and wages", *Economics Letters*, 177, 35-38.

Yao, Y., and J.C. van Ours (2019b): "Dialect speaking and wages among native Dutch speakers," *Empirica*, 46(4), 653-668.

Figure 1: Fraction of non-English speakers by immigrant generation



Note: First generation are those who were born outside the U.K.; second generation are those who were born in the U.K. with at least one foreign-born parent; third generation are those who were born in the U.K. with two U.K.-born parents but at least one foreign-born grandparent; fourth generation are those who were born in the U.K. with U.K.-born parents and grandparents.

Table 1: Descriptive statistics for the estimation sample

Variable	Women		Men	
	Monolinguals	Bilinguals	Monolinguals	Bilinguals
Gross hourly pay	15.143	15.031	15.139	14.548
Age	40.648	33.837	40.638	33.689
A-level education	0.114	0.173	0.123	0.152
Undergraduate education	0.311	0.332	0.235	0.295
Postgraduate education	0.085	0.124	0.097	0.130
Married	0.484	0.495	0.513	0.563
Number of children	0.585	0.778	0.568	0.817
Asian	0.018	0.681	0.024	0.743
Black	0.025	0.045	0.017	0.025
Other race	0.002	0.007	0.002	0.006
Mixed race	0.018	0.017	0.013	0.009
Hours worked	30.037	29.838	37.649	35.846
Lives in London	0.081	0.328	0.085	0.321
Second generation immigrant	0.120	0.854	0.116	0.833
Mother had high school education	0.446	0.309	0.426	0.282
Mother had university education	0.061	0.054	0.058	0.041
Father had high school education	0.405	0.317	0.426	0.310
Father had university education	0.076	0.079	0.078	0.110
Number of observations	25,597	969	20,455	822

Notes: Observations are weighted by the inverse of the total number of observations for each person in the sample.

Table 2: Results for log wage regressions

Variable	Women		Men	
	(1)	(2)	(3)	(4)
Bilingual	-0.052**	-0.072***	-0.017	-0.008
	(0.025)	(0.027)	(0.027)	(0.028)
Age	0.067***	0.067***	0.071***	0.071***
	(0.002)	(0.002)	(0.002)	(0.002)
Age squared	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Married	0.048***	0.047***	0.119***	0.116***
	(0.009)	(0.009)	(0.010)	(0.010)
Number of children	0.040***	0.036***	0.015***	0.016***
	(0.005)	(0.005)	(0.005)	(0.005)
London	0.152***	0.149***	0.160***	0.157***
	(0.015)	(0.015)	(0.015)	(0.016)
Hours worked	-0.007***	-0.007***	-0.011***	-0.011***
	(0.000)	(0.000)	(0.000)	(0.000)
Mother had high school education		0.058***		0.055***
		(0.010)		(0.011)
Mother had university education		0.071***		0.044**
		(0.019)		(0.021)
Father had high school education		0.002		0.028***
		(0.010)		(0.011)
Father had university education		0.010		0.024
		(0.017)		(0.018)
Mother's country of birth fixed effects	No	Yes	No	Yes
Observations	26,566	26,566	21,277	21,277
R-squared	0.156	0.160	0.218	0.222

Notes: All regressions include a full set of race (5 categories), education (4 categories) and year (4 categories) dummies.
Standard errors are presented in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Observations are weighted by the inverse of the total number of observations for each person in the sample.

Table 3: Heterogeneity in the log wage regressions

Variable	Women		Men	
	(1)	(2)	(3)	(4)
Bilingual	-0.451***	-0.088***	-0.033	-0.030
	(0.124)	(0.031)	(0.127)	(0.033)
Fraction speaking English in M.S.O.A.	-0.052		0.259***	
	(0.070)		(0.070)	
Fraction speaking English in M.S.O.A. × bilingual	0.463***		0.017	
	(0.147)		(0.154)	
Third/fourth generation		0.005		-0.016
		(0.021)		(0.022)
Third/fourth generation × bilingual		0.059		0.075
		(0.060)		(0.061)
Observations	23,055	26,566	18,416	21,277
R-squared	0.153	0.160	0.220	0.222

Notes: All regressions include age and age squared, hours worked, number of children, married and London dummies and a full set of race (5 categories), education (4 categories), year (4 categories) and mother's country of birth dummies.
Standard errors are presented in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Observations are weighted by the inverse of the total number of observations for each person in the sample.

Table 4: Results for log wage regressions using Oster's (2019) approach

Sample	Assumed δ									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Women	-0.083	-0.096	-0.112	-0.131	-0.154	-0.185	-0.226	-0.284	-0.371	-0.519
Men	-0.005	-0.002	0.001	0.004	0.008	0.012	0.016	0.021	0.026	0.031

Notes: All regressions include age and age squared, hours worked, number of children, married, London and proud of mother's country dummies and a full set of race (5 categories), education (4 categories), year (4 categories), mother's country of birth (28 categories) and parental education (4 categories) dummies.

Observations are weighted by the inverse of the total number of observations for each person in the sample.

A value of $R_{\max} = 0.3$ is assumed.

Table 5: Results including separate language groups

Variable	Women		Men	
	(1)	(2)	(3)	(4)
Welsh or Gaelic	0.051	0.055	-0.003	0.023
	(0.054)	(0.056)	(0.055)	(0.057)
Any European language	-0.074	-0.126	0.044	0.039
	(0.085)	(0.092)	(0.102)	(0.116)
Arabic	-0.096	-0.164	0.096	0.039
	(0.161)	(0.162)	(0.375)	(0.375)
Chinese	0.126	-0.010	-0.055	0.150
	(0.149)	(0.190)	(0.125)	(0.151)
South Asian language	-0.068*	-0.085**	-0.047	-0.052
	(0.035)	(0.037)	(0.034)	(0.036)
Any other language	-0.139**	-0.179***	-0.071	0.087
	(0.058)	(0.062)	(0.070)	(0.072)
Parents' education fixed effects	No	Yes	No	Yes
Mother's country of birth fixed effects	No	Yes	No	Yes
Observations	26,566	26,566	21,277	21,277
R-squared	0.156	0.160	0.218	0.222

Notes: All regressions include age and age squared, hours worked, number of children, married and London dummies and a full set of race (5 categories), education (4 categories) and year (4 categories) dummies.
Standard errors are presented in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Observations are weighted by the inverse of the total number of observations for each person in the sample.

Appendix

Table A1: Robustness tests for log wage regressions

Variable	Women		Men	
	(1)	(2)	(3)	(4)
Bilingual	-0.101*	-0.053*	0.023	-0.020
	(0.054)	(0.029)	(0.053)	(0.030)
Christian	0.017		0.019	
	(0.012)		(0.012)	
Other religion	0.079**		-0.007	
	(0.031)		(0.029)	
Half or less of friends of same race		-0.070***		-0.057***
		(0.012)		(0.013)
Observations	12,312	23,459	11,543	18,031
R-squared	0.168	0.156	0.233	0.217

Notes: All regressions include age and age squared, hours worked, number of children, married and London dummies and a full set of race (5 categories), education (4 categories), year (4 categories), parents' education and mother's country of birth dummies..

Standard errors are presented in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.

Observations are weighted by the inverse of the total number of observations for each person in the sample.

Table A2: Results for employment regressions

Variable	Women		Men	
	(1)	(2)	(3)	(4)
Bilingual	-0.028***	-0.028***	0.032***	0.032***
	(0.006)	(0.006)	(0.007)	(0.007)
Age	0.014***	0.014***	0.013***	0.013***
	(0.001)	(0.001)	(0.001)	(0.001)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Married	0.001	0.001	0.005*	0.005*
	(0.002)	(0.002)	(0.003)	(0.003)
Number of children	0.003**	0.003**	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
London	-0.013***	-0.013***	0.007	0.008**
	(0.004)	(0.004)	(0.004)	(0.004)
Hours worked	0.001***	0.001***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Mother had high school education		-0.001		0.003
		(0.002)		(0.003)
Mother had university education		-0.006		0.000
		(0.005)		(0.006)
Father had high school education		-0.001		-0.003
		(0.002)		(0.003)
Father had university education		-0.003		-0.006
		(0.004)		(0.005)
Mother's country of birth fixed effects	No	Yes	No	Yes
Observations	26,566	26,566	21,277	21,277
R-squared	0.061	0.061	0.081	0.083

Notes: All regressions include age and age squared, hours worked, number of children, married and London dummies and a full set of race (5 categories), education (4 categories) and year (4 categories) dummies.
Standard errors are presented in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.
Observations are weighted by the inverse of the total number of observations for each person in the sample.