J Popul Econ (2002) 15: 137–160

Journal of
Population
Economics

The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas

Michael A. Shields, Stephen Wheatley Price

Public Sector Economics Research Centre, Department of Economics, University of Leicester, Leicester, LE1 7RH, UK (Fax: +44-116-2522908; e-mail: mas9@le.ac.uk, swp2@le.ac.uk)

Received: 26 November 1999/Accepted: 03 August 2000

Abstract. This paper examines two crucial aspects of the assimilation experience of ethnic minority immigrants in the United Kingdom. It explores the determinants of their English language (speaking) fluency and the key role such skills play in their occupational success. Our sample is derived from the Fourth National Survey of Ethnic Minorities undertaken in 1994. Uniquely this data contains an interviewer-assessed measure of English language fluency. Importantly, we also attempt to control for possible endogeneity bias in the estimates of the effect of language fluency on occupational success. We find that fluency is associated with significantly higher mean hourly occupational wages.

JEL classification: J15, J24, J61

Key words: Language, occupational success, immigrants and ethnic minorities

1. Introduction

The United Kingdom (UK) was a major source of international migration flows over the last two centuries. Only relatively recently has it become a

All correspondence to Stephen Wheatley Price. The authors are grateful for comments received from participants at the Centre for Economic Policy Research conference "Marginal Labour Markets in Metropolitan Areas" held in Dublin between 10–12 October 1999 where an earlier version of this paper was presented. In addition Barry R. Chiswick and three anonymous referees have made valuable suggestions which have greatly improved the final version of this paper. The Fourth National Survey of Ethnic Minorities (deposited by the Policy Studies Institute) and the Quarterly Labour Force Survey of the United Kingdom (Crown Copyright; deposited by the Office for National Statistics) were used with permission. The authors are grateful to the Data Archive at the University of Essex for supplying the data and documentation. *Responsible editor:* Alan Barrett.

country of net immigration, mainly due to substantial inflows from South Asian countries where English is not widely spoken (see Hatton and Wheatley Price 1999 for an extended discussion). According to the 1991 Census the total stock of immigrants (i.e. those born outside the UK) numbered nearly 4 million people (or 7.4% of the UK population) whilst approaching 3 million people (or 5.5% of the total) belonged to the ethnic minorities (i.e. report an ethnicity other than White), the majority of whom (2.95% of the UK population) were born abroad. Both ethnic minority and immigrant groups are highly concentrated in the metropolitan areas of England. In particular, 45% of Britains' ethnic minorities reside in Greater London, together with 37% of all immigrants (Owen 1992, 1993), and the majority of the remainder live in an urban environment.

The labour market disadvantages of Britain's ethnic minorities are well recognised (e.g., Modood et al. 1997) and a number of recent papers have investigated the extent of racial discrimination in their employment, promotion and earnings performance (e.g., Blackaby et al. 1994, 1997, 1998; Pudney and Shields 2000). However, the fact that the majority of these individuals are immigrants has been largely ignored. Recently, factors such as country of birth, years since migration and the transferability of human capital have been shown to be important in determining their employment and unemployment propensities (Wheatley Price 2001a, 2001b), employer-funded training outcomes (Shields and Wheatley Price 1999a, 1999b) and earnings (Shields and Wheatley Price 1998) using 1990s data. However, the role of English language skills in the labour market outcomes experienced by the vast majority of Britain's ethnic minorities has yet to be examined using recent data.[1]

In this paper we focus on male ethnic minority immigrants living in English metropolitan areas since this is where the vast majority are concentrated, and where their disadvantage is greatest. For example, ethnic minority men in Greater London are twice as likely to be unemployed than white men (Modood et al. 1997). Nevertheless, this study concerns those in a relatively favourable labour market position, namely paid employees. If we can gain an insight into the causes of their success, especially what role fluency in the English language plays in helping them climb the occupational ladder, then we may be able to suggest policies to help others out of the labour market difficulties they face and make some recommendations about future UK immigration policy.

We explore two aspects of their assimilation experience, namely their English language (speaking) fluency and occupational success, using data from the Fourth National Survey of Ethnic Minorities, undertaken in 1994, by the Policy Studies Institute in London (Modood et al. 1997). Uniquely this data provides interviewer-assessed measures of English language speaking fluency, thus avoiding the form of measurement error endemic in studies that use self-reported measures (Dustmann and van Soest 1998a, 1998b). However, the interviewers themselves may be a source of a different form of measurement error. In this paper occupational success is defined as the mean gross hourly wage associated with each 3-digit Standard Occupational Classification type of employment, using information on average earnings from the 1993–1995 Quarterly Labour Force Surveys of the United Kingdom.[2] Therefore we are concerned, in this study, with comparisons across occupations rather than within them. Due to the continuous nature of this variable we can use similar econometric techniques to those employed in recent studies of immigrant earnings.

We estimate determinants of language fluency models, following Chiswick and Miller (1995), and then attempt to capture its affect on occupational success, along with other immigrant-related characteristics. As has been demonstrated by Chiswick and Miller (1992, 1995) and Chiswick (1998) Ordinary Least Squares (OLS) estimates, of the coefficient on language fluency, may be biased due to the presence of unobserved heterogeneity affecting both language fluency and measures of earnings. Dustmann and van Soest (1998a, 1998b) demonstrate that errors in the measurement of linguistic ability also lead to biased OLS estimates. In an attempt to investigate these possibilities we use Instrumental Variables (IV). Our results using OLS appear to underestimate the importance of English language fluency on the occupational success of ethnic minority male immigrants living in English conurbations.

The paper is set out as follows. Section 2 reviews the theoretical hypotheses concerning language acquisition, proposed by Chiswick and Miller (1992, 1995), and labour market success, suggested by Chiswick (1978). We introduce our data source and describe the specific sample in Sect. 3. Section 4 discusses the determinants of English language fluency whilst Sect. 5 presents our results concerning occupational success and, in particular, the partial effect of English language speaking fluency. Section 6 concludes the paper and discusses some policy implications.

2. Theoretical considerations

Cross-sectional studies of the earnings of immigrants in the United States' labour market have revealed much about their economic assimilation, following the seminal paper by Chiswick (1978). Since labour market experience gained in the destination country is valued more highly than that gained in the source country, due to the necessity of acquiring location-specific human capital, immigrant earnings growth is so rapid that the wages of comparable natives are exceeded after approximately 10–15 years (Chiswick 1978).

One of the most important forms of location-specific human capital is the ability to communicate in the host country's language. These skills are embodied in the person, productive in the labour market and/or in consumption, and are costly to acquire, both in terms of time and other resources (Chiswick and Miller 1992, 1995). For immigrants the acquisition of this form of human capital has been shown to be crucial to their labour market success in a number of different countries (see Chiswick and Miller 1995 and Chiswick 1998 for summaries of this literature).

Chiswick and Miller (1992, 1995) argued that language fluency was determined by economic incentives, exposure to the language and the efficiency of acquisition. We now briefly summarise their main hypotheses. Economic incentives include the expected economic benefit from fluency (giving rise to possible endogeneity between measures of economic success, such as earnings, and language fluency) and the expected future duration in the destination (see Dustmann 1999 for evidence). Clearly those immigrants who are primarily labour migrants have a greater incentive to acquire language skills than those who are refugees.

Exposure may occur before immigration through formal education or through the use of the language in everyday life (e.g., English is an official language in many of countries of origin for United Kingdom immigrants).

Post-immigration exposure may occur with time spent in the destination country, through education, through marriage to a native born person or through specific language training. Exposure may be lessened if the immigrant lives and works mainly amongst their own ethnic group, is married to an immigrant who shares the same first language (Chiswick and Miller 1996) or if they have children who act as translators for them. Alternatively, it may be the case that the presence of children encourages the development of linguistic skills through informal learning processes. The efficiency of acquisition of language skills depends on age, particularly the age at which acquisition begins (often upon immigration), the individuals learning ability (which is related to their level of education) and the linguistic distance between the immigrant's mother tongue and English (Chiswick and Miller 1998).

Empirically this model predicts that fluency in the destination country's language would be positively related to the expected wage increase arising from fluency, the expected future duration in the country, the number of years since migration, formal language training, informal learning from children and the individual's level of education. However, being married to a fellow immigrant, having children who act as translators, living amongst other members of the same ethnic group, increasing age of immigration and linguistic distance would be expected to be associated with reduced language proficiency (Chiswick and Miller 1995).

In the absence of direct information concerning expected future duration in the destination country we suggest that immigrants who send remittances to family or friends in their home country have fewer incentives to invest in location-specific language capital.[3] Galor and Stark (1990) have argued that if the worker is a remitter they would be more likely to return migrate than if they were not sending remittances to their country of origin. Merkle and Zimmermann (1992) provide supporting empirical evidence for this hypothesis amongst German immigrants. As Dustmann (1999) has argued, an immigrant who is likely to return migrate has fewer incentives to invest in linguistic skills appropriate only to the country of temporary residence. Combining these two factors we hypothesise that workers who are sending remittances would be less likely to be fluent in the English language than those who are not. Furthermore, we suggest that if an individual has a long-term health problem, that limits the type of work that they can do, they are less likely to invest in language fluency due to both reduced economic incentives and lower efficiency in attaining fluency.

In addition to the standard human capital, familial and locational characteristics, which are widely perceived as being associated with higher wages and occupational success, aspects of being foreign born have been shown to be important in such models. Specifically, as mentioned earlier, Chiswick (1978) has shown the separate effect of time spent in the destination country, usually measured as years since migration, on the earnings of immigrants. Furthermore, numerous differences amongst immigrants are often captured by country of origin variables. These include differences in the quality and transferability of the education and experience they received abroad. In addition there may be systematic differences in unobserved factors associated with country of birth such as the political or economic climate at the time of emigration, the nature of the migration flows (e.g., refugees) and the likelihood of return migration (Chiswick 1978; Borjas 1985, 1987). Finally, there is some evidence to suggest that immigrants' success in the wage and salary sector

may be inversely related to the concentration of their own linguistic group in their residential area (Chiswick and Miller 1999b). We examine these hypotheses in Sect. 5.

3. Data

3.1. Data source

The data source we use in this paper is the Fourth National Survey of Ethnic Minorities, conducted by the Policy Studies Institute in 1994 (see Modood et al. 1997 for fuller details). The samples of ethnic minorities included in the survey were selected using data from the 1991 Census to divide all electoral wards in England and Wales into three bands (high, medium and low), according to the proportion of the population who reported being members of an ethnic minority. Random samples of wards were selected and, within each ward, addresses were randomly sampled. High band ethnic minority wards were over-sampled. Interviewers then visited the resulting 130,000 addresses to identify whether any members of the target ethnic minority groups (Black Caribbeans, Indians, Pakistanis, Bangladeshis, African Asians and Chinese) were living at each address. Since the survey deliberately over-samples high density ethnic minority areas and Black Africans are not included[4] the sample is not nationally representative of all ethnic minority groups.

At each household containing adults (aged 16 or over) from these target groups, up to two were randomly selected for interview. Where there were two respondents in one household each was asked one of two randomly assigned questionnaires, including the same core questions (which we use in this study), with different sets of secondary questions. Both individuals who were born in the UK and those who were born abroad were included in the sampling. Interviews were successfully undertaken in 3291 ethnic minority households, involving 5196 adults. The response rates were 61% for Black Caribbeans, 74% for Indians and African Asians, 73% for Pakistanis, 83% for Bangladeshis, 66% for Chinese and 71% for the comparison sample of 2867 Whites. Importantly, a member of the same broad ethnic group as the respondent, and who spoke both English and the respondents other main language, usually conducted the interview in order to maximise response rates and minimise misunderstandings. Uniquely, amongst national level sources of data in the United Kingdom interviews could be conducted wholly or partly in the interviewees' language of preference, as well as in English. This data therefore captures the substantial proportion of the ethnic minority population with poor language skills who are missed by other surveys that only interview in English.

Furthermore, the interviewer's assessment of the respondent's English language speaking fluency is recorded in the data, together with whether the interview was conducted wholly, partly or not at all in English. The relevant questions are the last on the questionnaire. Firstly the interviewer is asked to record whether the interview was conducted wholly in English, partly in English and partly in another language or wholly in another language. Next the other language used is specified and finally the interviewer is instructed to assess the respondent's "English language ability". The questionnaire specifically states "if interview not conducted in English, attempt a conversation to

assess ability''. The interviewer must assess the ability as ''Speaks English fluently'', ''Fairly Well'', ''Slightly'' or ''Not at all''.

Most other studies of this nature base their findings on self-reported measures of language fluency, which have been shown to systematically misclassify language ability. This results in under-estimates of the true importance of fluency on earnings (Dustmann and van Soest 1998a, 1998b). Our data is therefore free of self-reported measurement error, although the interviewers themselves may have incorrectly assessed the language ability of respondents. Unfortunately, our data only contains information on speaking fluency. Chiswick (1991) and Dustmann (1994) provide evidence to suggest that reading fluency and writing fluency, respectively, are even more important determinants of earnings than speaking fluency.[5]

The data records the earnings of employees, but only grouped in bands, and around 20% of responses are missing. Furthermore the non-response to this questions varies considerably by ethnic group (see Dustmann and Fabbri 2000, for details). However, by using the (complete) records of the individuals' 3-digit Standard Occupational Classification (899 categories) we can compute the mean gross hourly wage for each occupational category using data from the Quarterly Labour Force Survey of the United Kingdom (1993–1995). This information is derived from the 83777 full-time, male and female, white and ethnic minority, native born and foreign born, employees, aged 22–64, who reported wage information in one of these surveys.[6] This is the measure we use to rank occupational success in our analysis (following Nickell 1982 and Stewart 1983). Since the Fourth National Survey of Ethnic Minorities provides no direct information on years of schooling, and no simple way to accurately derive one, we use highest qualification as our measure of education.[7] For similar reasons, we use age, rather than years of potential experience, in our models. In addition, the data includes information, derived from the 1991 Census, on respondents' own ethnic group density at ward level (about 60,000 individuals).

3.2. Sample descriptive statistics

Selecting ethnic minority, foreign born male employees, aged between 22 and 64, who were living in the metropolitan areas of England in 1994, provides us with a sample of 565 individuals.[8] Table 1 provides the mean values of the dependent and independent variables used in our analyses of the determinants of language fluency and occupational success. Over half the sample (58.2%) have been assessed by their respective interviewees as being fluent in (speaking) the English language whilst the average gross hourly mean occupational wage is £6.14. The equivalent figure for white native born men is £7.13 clearly indicating that ethnic minority immigrant men are more likely to be employed in lower paying occupations.

The mean age of our sample is 40 years old and on average they have been in the UK for 21.5 years. This implies a mean arrival year of 1972 or 1973 aged 18–19 years old. Over 30% of our sample are Indian (born in India) with a further 20.5% being Pakistani (born in Pakistan) and 10.3% are Bangladeshi (born in Bangladesh). Of the 21.2% who were born in East Africa, but who have their historical roots in the Indian Sub-continent (referred to as African Asians), the vast majority are of Indian ethnicity but there are also a few

Table 1. Descriptive statistics (ethnic minority foreign born male employees, age 22–64, living in metropolitan areas of England in 1994)

Variable	Mean	S.E.
Gross hourly (mean occupational) wage	6.14	0.084
Fluent in (speaking) the English language	0.582	0.021
Age	39.94	0.418
$(\text{Age})^2/100$	16.94	0.355
Age at immigration	18.42	0.372
$(\text{Age at immigration})^2/100$	4.18	0.160
Years since immigration	21.52	0.387
Indian (born in India)	0.303	0.019
Bangladeshi (born in Bangladesh)	0.103	0.013
Pakistani (born in Pakistan)	0.205	0.017
African Asian (born in East Africa)	0.212	0.017
Black Caribbean (born in the Caribbean)	0.136	0.014
Chinese (foreign born of Chinese ethnicity)	0.041	0.008
Married	0.858	0.015
Wife UK born	0.108	0.013
Wife foreign born	0.750	0.018
Not married	0.142	0.015
No dependent children (aged < 16)	0.273	0.019
One dependent child (aged < 16)	0.219	0.017
Two dependent children (aged < 16)	0.262	0.019
Three dependent children (aged < 16)	0.142	0.015
$>$ Three dependent children (aged < 16)	0.104	0.013
Oldest child (in household) aged > 15	0.257	0.018
Oldest child (in household) aged 12–15	0.161	0.016
Oldest child (in household) aged 5–11	0.258	0.018
Oldest child (in household) aged 0–4	0.145	0.015
No children (in household)	0.179	0.016
Degree (or equivalent highest) qualification	0.223	0.018
A-level (or equivalent highest) qualification	0.074	0.011
Vocational (highest) qualification	0.202	0.017
O-level (or equivalent highest) qualification	0.145	0.015
No qualifications	0.356	0.020
Living in the Midlands (metropolitan area)	0.281	0.019
Living in the North (metropolitan area)	0.184	0.016
Living in Greater London (metropolitan area)	0.535	0.021
0–5% own ethnic density (in Census ward)	0.228	0.018
5–15% own ethnic density (in Census ward)	0.363	0.020
15–33% own ethnic density (in Census ward)	0.285	0.019
>33% own ethnic density (in Census ward)	0.124	0.014
Manufacturing (sector)	0.375	0.020
Non-financial services (sector)	0.223	0.018
Financial services (sector)	0.120	0.014
Public (sector)	0.103	0.013
Transport (sector)	0.080	0.011
Other industrial (sector)	0.099	0.013
Large firm (> 50 employees at the workplace)	0.457	0.021
Long term health problem (that limits work)	0.039	0.008
Remitter (of money to country of origin)	0.296	0.019
Interview conducted wholly in English	0.596	0.021
Interview conducted partly in English	0.264	0.019
Interview conducted wholly in another language	0.140	0.015
Sample Size	565	

Note: Authors own calculations using a sample derived from the Fourth National Survey of Ethnic Minorities. For dummy variables, the values shown are the proportion of the sample for which the value is one. S.E. stands for the standard error of the mean.

whose families originate from Pakistan or Bangladesh (Modood et al. 1997). A small number (4%) of ethnic Chinese immigrants are also included in our sample.

Nearly 86% of the ethnic minority foreign born men are married, most to fellow immigrants reflecting their cultural practices. Similarly, large families are not uncommon amongst these groups with 10% of men reporting to have four or more children and 14% having three dependent children. Over 25% of respondents have a child over the age of 15 still living in the household, with less than 15% having no children at home. The highest qualification variables record the prevalence of individuals at both ends of the educational achievement spectrum. More than 22% of these immigrants possess degree-level qualifications and yet over 35% have no qualifications whatsoever. Over 20% have vocational qualifications, 7.5% possess formal schooling certificates at A-level and 14.6% at O-level, or their equivalents, corresponding to the UK school-leaving ages of 18 and 16 years old, respectively.

As mentioned in the introduction to this paper, Greater London (53.5%) is the dominant locational choice amongst these groups. The (West) Midlands metropolitan area accounts for 28.1% of our sample with the remainder being distributed amongst the northern metropolitan areas of Merseyside, Greater Manchester, South and West Yorkshire and Tyneside. The concentration of specific ethnic groups in Census wards is highlighted by the fact that less than a quarter (22.8%) of our sample reside in a ward with a density of their own ethnic group under 5%. Over 40% live in a ward where their ethnic group accounts for at least 15% of the population.[9]

The substantial minority (37.5%) of ethnic minority immigrant male employees work in the manufacturing sector with the non-financial service sector (including retail, sales, personal, hotel and catering services) accounting for a further 22.3%. The remainder of the sample is fairly evenly distributed (about 10% each) amongst the financial, public, transport and other sectors of industry. Just under half (45.7%) of the workers are employed in large firms (> 50 employees at the workplace). Less than 4% of our sample reported a long-term health problem that limited the type of paid work that they could do, whilst nearly 30% remitted some money to family or friends in their country of origin. Lastly, nearly 60% of respondents were interviewed, for the survey, wholly in English, a further 26.4% partly in English and just 14% wholly in another language.

4. The determinants of English language fluency

Each respondent to the Fourth National Survey of Ethnic Minorities had their English language speaking fluency assessed by the interviewer, a member of the same ethnic group and fluent in the respondent's other main language. The categories of assessment were fluent, fair, poor or none. In order to investigate the determinants of language fluency we constructed a dichotomous variable taking the value one if the individual was assessed as fluent in speaking the English language and zero otherwise. Table A1 (in the Appendix) presents some descriptive statistics concerning the fluency of our sample according to a number of characteristics that the theory highlighted earlier suggested as potentially important determining factors. It also shows whether the mean fluency rate of the sample, when each characteristic holds, is statistically significantly different from that when the characteristic does not hold

(i.e. when the respective dummy variables take the value 1 and 0). Since these bivariate comparisons take no account of other confounding factors we will not discuss them here.[10]

However, the fluency rates according to the linguistic nature of the interview are particularly interesting. Evidently being assessed as fluent in speaking the English language is significantly and positively correlated to the likelihood of being interviewed wholly in English, whilst the opposite is true for those interviewed only partly in English or wholly in another language. Since the question on the linguistic nature of the interview comes just before that on interviewer-assessed language fluency in the questionnaire (see note 3) it is surprising that only 88.1% of those interviewed wholly in English are recorded as being fluent. Even more intriguing is the finding that 21.5% of those interviewed partly in English and partly in another language, and 2.5% of those interviewed wholly in another language, were actually capable of being interviewed wholly in English, since they were assessed as fluent. It may be the case that other factors than merely English language speaking fluency determined the joint linguistic decision by interviewer and interviewee. Alternatively, the divergence between the interviewers' assessment of English speaking fluency and the linguistic nature of the interview may be indicative of some form of interviewer measurement error. This may be the case especially amongst those who were interviewed wholly in another language since their English language ability was determined on the basis of an attempted (and necessarily much shorter) conversation than if the interview was conducted wholly in English.

Following Chiswick and Miller (1995) we estimate the determinants of English language (speaking) fluency with independent variables attempting to investigate most of the hypotheses outlined in Sect. 2. This multivariate approach allows us to estimate the separate effect, of each of the explanatory variables, on language fluency. The first model we estimate (model A) includes variables capturing the age of immigration (and square), years since immigration,[11] country of birth, level of highest qualification and region of residence. Unfortunately, in common with most other studies of this nature, we do not observe an individual-specific expected wage premium from fluency or whether the individual has received any formal language training. Neither do we have a direct measure of the expected future duration in the country or linguistic distance nor do we know the reason for immigration. Country of birth variables may capture these latter three effects together with the effect of whether English was spoken (or not) in the country of origin.

In model B we add variables to capture the impact of other members of the household on language fluency. We include a dummy variable indicating whether the immigrant is married to a UK born or foreign born wife, the number of dependent children and indicators of the age of the oldest child (corresponding to their levels of compulsory schooling). In addition we control for the density of own ethnic group at the Census ward level. Since the own ethnic densities are calculated for all ethnic minority residents, both immigrants and native born, they will over-estimate the extent of linguistic compatibility. The resulting coefficients may therefore underestimate their true effect on language fluency. Finally we investigate whether the individual's long-term health or whether they are sending remittances to their home country are statistically associated with language skill acquisition.

Table 2 reports the coefficients, their standard errors and the marginal effects[12] of our two probit models of the determinants of English language

Table 2. Determinants of fluency in (speaking) the English language: probit estimates (sample as Table 1)

Variable	Model A			Model B		
	Coeff.	S.E.	M.E.	Coeff.	S.E.	M.E.
Constant	−0.5893	0.4072+	–	−0.6722	0.5508	–
Age at immigration	−0.1063	0.0263*	−0.040	−0.1033	0.0291*	−0.039
(Age at immigration)^2/100	0.1764	0.0585*	0.067	0.1893	0.0632*	0.071
Years since immigration	0.0330	0.0088*	0.013	0.0458	0.0117*	0.017
Indian	∼	–	–	∼	–	–
Bangladeshi	0.1223	0.2402	0.046	−0.0394	0.2627	0.015
Pakistani	−0.1510	0.1907	−0.058	−0.0744	0.2125	−0.028
African Asian	0.4734	0.1842*	0.169	0.2880	0.1961+	0.104
Black Caribbean	1.480	0.2528*	0.398	1.510	0.2943*	0.394
Chinese	0.4465	0.3474+	0.154	0.1705	0.3927	0.062
Wife UK born	∼	–	–	0.1428	0.3253	0.052
Wife foreign born	∼	–	–	0.3674	0.2559+	0.145
Not married	∼	–	–	∼	–	–
No dependent children	∼	–	–	∼	–	–
One dependent child	∼	–	–	0.5199	0.2364*	0.182
Two dependent children	∼	–	–	0.5052	0.2376*	0.178
Three dependent children	∼	–	–	0.2237	0.2687	0.081
> Three dependent children	∼	–	–	−0.2382	0.3265	−0.092
Oldest child aged > 15	∼	–	–	−0.6915	0.2781*	−0.266
Oldest child aged 12–15	∼	–	–	−0.5428	0.3287#	−0.211
Oldest child aged 5–11	∼	–	–	−0.5735	0.3004#	−0.221
Oldest child aged 0–4	∼	–	–	−0.3805	0.3155	−0.147
No children	∼	–	–	∼	–	–
Degree qualification	2.002	0.2074*	0.520	2.013	0.2210*	0.512
A-level qualification	1.366	0.2563*	0.357	1.441	2717*	0.358
Vocational qualification	1.286	0.1935*	0.387	1.379	0.2090*	0.398
O-level qualification	0.7275	0.1968*	0.242	0.8597	0.2077*	0.272
No qualifications	∼	–	–	∼	–	–
Living in the Midlands	∼	–	–	∼	–	–
Living in the North	0.0691	0.1961	0.026	0.0450	0.2091	0.017
Living in Greater London	0.3619	0.1549*	0.137	0.4052	0.1681*	0.152
0–5% own ethnic density	∼	–	–	∼	–	–
5–15% own ethnic density	∼	–	–	−0.2077	0.2045	−0.079
15–33% own ethnic density	∼	–	–	−0.4858	0.2137*	−0.186
>33% own ethnic density	∼	–	–	−0.2984	0.2540	−0.115
Long term health problem	∼	–	–	−0.5237	0.3601+	−0.205
Remitter	∼	–	–	−0.3340	0.1551*	−0.128
Actual probability of fluency		0.582			0.582	
Predicted probability of fluency		0.626			0.637	
Restricted Log-Likelihood (Slopes = 0)		−383.94			−383.94	
Unrestricted Log-Likelihood		−255.65			−238.95	
Model χ^2		256.59*			289.98*	
Degrees of Freedom (χ^2 test)		14			29	
Pseudo-R^2_{ANN}		0.542			0.589	
Sample Size		565			565	

Note: ∼ indicates an omitted variable. Coeff. is an abbreviation for the estimated coefficient. S.E. stands for standard error. *, # and + indicate significance at the 5%, 10% and 20% levels, respectively. M.E. indicates the marginal effect on the predicted probability of fluency in (speaking) the English language, calculated for an individual with sample mean characteristics. For continuous variables the marginal effect is calculated for an increase of 1 year. For the dummy variables it represents an average person with that particular characteristic as compared to the base characteristic. Pseudo-R^2_{ANN} (the Aldrich and Nelson (1984) measure normalised) was proposed by Veall and Zimmermann (1992) and, amongst the significance-of-fit class of pseudo-R^2s, most closely corresponds to the OLS-R^2 (see Veall and Zimmermann, 1996).

(speaking) fluency. Since the inclusion of the extra variables significantly improves the maximum log-likelihood of the extended model we only discuss the results of model B.[13] Compared to the actual probability of fluency of 0.582 the model predicts the probability of fluency for a male immigrant, holding all characteristics at their sample means, to be 0.637. As Chiswick and Miller's (1992, 1995) model anticipates English language speaking fluency is statistically associated with age at immigration. Increasing years since immigration significantly increases the probability of language fluency (by 0.017 per year). Holding age constant, there is a double benefit from immigrating young. Not only is the person more efficient at acquiring language skills but also they are subject to greater exposure to the English language through more years spent in the United Kingdom after immigration.

Once other factors are controlled for only those ethnic minority immigrant men who are Black (and born in the Caribbean), and those who are African Asian (though only at the 20% level), are associated with significantly increased probabilities of fluency, when compared to Indians. The marginal effects are 0.394 and 0.104, respectively. As noted earlier these variables may capture a number of effects, perhaps most importantly the effect of being exposed to spoken English in the country of origin. Interestingly, being married to someone born in the United Kingdom is associated with a much lower probability of fluency (0.05) in the English language than being married to a fellow immigrant (0.145), but this difference is not statistically significant. Nevertheless it is surprising that being married to a UK born individual conveys no linguistic benefit. However, since the majority of our sample are from South Asia, and second generation South Asians are often brought up in homes where their parents first language is widely used and marriages are often arranged within linguistic groups, it may be the case that the marital partners communicate in this other language. The UK born partner may therefore acts as a translator into English, rather than as an informal teacher of English within South Asian marriages.

However, having one or two dependent children is associated with a significant increase in the probability of fluency (by 0.182 and 0.178, respectively) compared to an equivalent immigrant with no dependent children. Controlling for the number of children, those men with school age or older children are significantly less likely to be fluent than those with none. The marginal effects are large (e.g., -0.266 for the oldest child aged > 15). It may be the case that school-aged children, who are likely to be fluent in the English language, act as translators for their parents reducing the parents' need to attain fluency. This effect outweighs any linguistic benefit the parents may gain by learning from their children.

Educational attainment is clearly a closely associated with English language speaking ability. Any qualification significantly improves the probability of being assessed as fluent by the interviewer over (an otherwise identical person with) no qualifications. Possession of a degree (or equivalent) highest qualification has a marginal effect of 0.512, whilst the effects of A-levels (0.358), vocational qualifications (0.398) and O-levels (0.272) are also large. Evidently learning skills gained through education increase the efficiency of language acquisition and, given that many immigrants completed their education in the UK, higher levels of education will increase exposure to the English language. The economic incentives may also be larger for the more highly educated, as the returns to fluency may be much greater in professional

occupations than in skilled manual jobs. We examine this hypothesis in Sect. 5. Furthermore, English language skills may complement existing human capital by improving its transferability.

With regard to location, only residing in Greater London is significantly associated with improved fluency, when compared to living in the Midlands. The marginal effect is 0.152. It may be the case that the greatest economic benefits to fluency are to be found in Greater London, holding other factors constant. Compared to a Census ward level own ethnic density of 0–5%, those living amongst 15–33% of their own ethnicity are significantly less likely to be fluent (marginal effect = 0.186). Evidently decreased exposure to English speaking people may be a cause.[14] However, those immigrants living in a ward with >33% own ethnic density are not significantly less likely to be fluent than those in the base category (and their marginal effect is much lower than that of the 15–33% group). It may be the case that English language training opportunities, either publicly (e.g., through local councils) or privately funded, are concentrated in these areas. However, since the densities include native born as well as immigrant members of these ethnic groups, and we do not know their relative distributions, we must be cautious about these findings. Furthermore, these variables are potentially endogeneous in such models as the locational choice may be partly determined by linguistic ability (Dustmann 1997).

We find the expected negative coefficient for individuals who have a long-term health problem but, whilst the marginal effect of −0.205 is large, the coefficient is only marginally significant (at the 20% level). Finally, being a remitter is statistically associated with a lower probability of fluency (−0.128). This finding provides some evidence to support the contention that remitters are more attached to their country of origin, more likely to return migrate and therefore less likely to invest in location-specific human capital such as language fluency.

5. The determinants of occupational success

5.1. OLS estimates

Using our measure of occupational success, the (natural logarithm of the) mean gross hourly wage according to the 3-digit Standard Occupational Classification (derived from the Quarterly Labour Force Surveys between 1993 and 1995), we estimate our models using Ordinary Least Squares (OLS). The independent variables in model 1 are the standard human capital measures of experience (proxied by age and its square), education (our highest qualification measures), marriage and locational dummies. Following Chiswick (1978) we also include years since immigration[15] and country of birth variables. Due to the cross-sectional nature of our data our estimates are subject to Borjas' (1985, 1987) critiques of this methodology. Our estimates of the coefficient on the years since immigration variable may be biased if the average unobserved heterogeneity of immigrants remaining in England varies with time spent in the United Kingdom. This problem will be diminished to the extent that country of birth dummy variables can account for this variation.

Following Chiswick and Miller (1999b), the Census ward-level own ethnic minority group densities variables are also added to the empirical model.

Finally, the long term health problem dummy variable and our measure of English language (speaking) fluency are also included as explanatory variables. Model 2 adds the work-related characteristics of sector of employment and firm size to model 1. Since an F-test indicates that the null hypothesis (that the coefficients on the additional variables in the extended model are jointly zero) can be rejected at the 5% level (F-statistic = 7.06, $F(6, 538)$ critical value = 2.10) we will only discuss the results from model 2 below. It is straightforward to see from Table 3 that our findings are reasonably robust across both models.

Controlling for years since immigration, age has the expected non-linear effect on occupational success. Time spent in the United Kingdom clearly significantly increases occupational success, holding all other characteristics constant. The effect is an increase in the mean occupational wage of about four and a half percentage points for an additional ten years since immigration. Only Bangladeshis and African Asians have significantly different occupational attainments from Indians. The former group have a 14% lower mean occupational wage, other things being equal, whilst African Asians are more successful than Indians (6.2% higher mean occupational wages) even after controlling for linguistic ability.

Those ethnic minority immigrants with degree or equivalent highest qualifications are in occupations that, holding other characteristics constant, are paid 29% higher gross hourly wages than the jobs occupied by individuals with no qualifications. Furthermore, possession of A-levels or vocational qualifications significantly raises the mean occupational wage, by about 10%, above the base group. However, there is no significant occupational reward to those with just O-level or equivalent highest qualifications, over those with none. Neither are the married, locational or long term health problem dummy variables significantly different from their respective base groups. The coefficients associated with the own ethnic group density variables are only marginally statistically significant and are indicative of a 4.5% occupational success penalty to living in high ethnic minority density (15%–33%) census ward.

The incorporation of work-related characteristics adds important detail to the picture of occupation success for these immigrant employees. Compared to similarly endowed individuals in the manufacturing sector, workers in the financial services (8.6%), public sector (7.9%) and other industrial sectors (9.8%) earn significantly greater average wages, whilst those in the non-financial services (−7.5%) are rewarded significantly less. Furthermore, currently working for a large firm (> 50 employees) increases the mean occupational wage by around 4.5%, compared to employees in smaller firms.

Ethnic minority immigrant men who are assessed as fluent in speaking the English language, by their interviewer, are significantly more likely to have higher occupational attainment than comparable individuals who are not fluent. The effect of fluency is to increase the average hourly wage rate by about 8.9%. This is similar to the 9.4% penalty, in terms of occupational attainment, for poor speaking English ability found by Stewart (1983).[16]

5.2. Issues in IV estimation

As we mentioned earlier OLS estimates of the effect on language fluency on the earnings of immigrant workers are potentially biased due to the presence of unobserved heterogeneity affecting both language skills and outcomes. This

Table 3. Determinants of occupational success: OLS and IV estimates (sample as Table 1)

Variable	Model 1			
	OLS		IV	
	Coeff.	S.E.	Coeff.	S.E.
Constant	1.305	0.1674*	1.268	0.1696*
Age	0.0118	0.0081+	0.0136	0.0082#
$(\text{Age})^2/100$	−0.0169	0.0094#	−0.0184	0.0095#
Years since immigration	0.0053	0.0015*	0.0040	0.0016*
Bangladeshi	−0.1845	0.0379*	−0.1853	0.0382*
Pakistani	−0.0083	0.0301	−0.0048	0.0304
African Asian	0.0652	0.0286*	0.0565	0.0291#
Black Caribbean	−0.0402	0.0368	−0.0671	0.0391#
Chinese	−0.0468	0.0568	−0.0486	0.0573
Married	0.0203	0.0305	0.0173	0.0309
Degree qualification	0.3324	0.0327*	0.2901	0.0385*
A-level qualification	0.1286	0.0425*	0.0968	0.0453*
Vocational qualification	0.1263	0.0305*	0.0965	0.0337*
O-level qualification	0.0006	0.0325	−0.0151	0.0336
Living in the North	−0.0179	0.0309	−0.0186	0.0312
Living in Greater London	−0.0198	0.0255	−0.0279	0.0260
5–15% own ethnic density	0.0359	0.0286	0.0404	0.0289+
15–33% own ethnic density	−0.0288	0.0302	−0.0206	0.0307
>33% own ethnic density	0.0103	0.0375	0.0133	0.0378
Long term health problem	0.0070	0.0519	0.0188	0.0526
Non-financial services	~	–	~	–
Financial services	~	–	~	–
Public	~	–	~	–
Transport	~	–	~	–
Other industrial	~	–	~	–
Large firm	~	–	~	–
English language fluency	0.1084	0.0250*	0.1821	0.0422*
Adjusted R^2	0.381		0.393	
F statistic	18.3*		18.1*	
Sample Size	565		565	

Note: The dependent variable is the natural logarithm of the mean gross hourly wage according to the 3-digit standard occupational classification derived from the Quarterly Labour Force Survey of the United Kingdom (1993–1995). Coeff. is an abbreviation for the estimated coefficient. S.E. stands for standard error. *, # and + indicate significance at the 5%, 10% and 20% levels,

may be because individuals with higher overall ability are more likely to invest in language capital (Chiswick and Miller 1992, 1995) or arise from measurement error in the language fluency variable (Dustmann and van Soest 1998a). Similar concerns surround our estimates of the impact of English language speaking fluency on our measure of occupational success.

However, in our data any measurement error would arise from the interviewer, rather than the individual respondent, miss-classifying language ability. In other words, interviewers might assess a survey respondent as fluent in speaking the English language whom an employer might consider less than fluent (or visa versa). Furthermore, because overall English language skills are likely to be important in many occupations, employers may take into account reading and writing skills, as well as speaking skills in assessing fluency and

Table 3. (continued)

Variable	Model 2			
	OLS		IV	
	Coeff.	S.E.	Coeff.	S.E.
Constant	1.394	0.1629*	1.338	0.1674*
Age	0.0094	0.0079	0.0115	0.0080+
$(\text{Age})^2/100$	−0.0149	0.0092+	−0.0166	0.0093#
Years since immigration	0.0044	0.0014*	0.0032	0.0015*
Bangladeshi	−0.1427	0.0383*	−0.1424	0.0388*
Pakistani	−0.0115	0.0293	−0.0086	0.0296
African Asian	0.0619	0.0281*	0.0545	0.0286#
Black Caribbean	−0.0301	0.0358	−0.0594	0.0382+
Chinese	−0.0349	0.0542	−0.0270	0.0567
Married	0.0166	0.0299	0.0155	0.0302
Degree qualification	0.2895	0.0337*	0.2476	0.0384*
A-level qualification	0.1030	0.0424*	0.0732	0.0448#
Vocational qualification	0.1087	0.0301*	0.0768	0.0328*
O-level qualification	0.0021	0.0317	−0.0141	0.0327
Living in the North	−0.0220	0.0300	−0.0223	0.0303
Living in Greater London	−0.0165	0.0255	−0.0232	0.0260
5–15% own ethnic density	0.0374	0.0272+	0.0411	0.0281+
15–33% own ethnic density	−0.0448	0.0249#	−0.0261	0.0298
>33% own ethnic density	−0.0070	0.0332	0.0066	0.0369
Long term health problem	0.0104	0.0504	0.0205	0.0512
Non-financial services	−0.0750	0.0294*	−0.0723	0.0298*
Financial services	0.0861	0.0346*	0.0751	0.0353*
Public	0.0786	0.0373*	0.0713	0.0377#
Transport	−0.0476	0.0390	−0.0468	0.0393
Other industrial	0.0977	0.0362*	0.0880	0.0368*
Large firm	0.0449	0.0214*	0.0433	0.0217*
English language fluency	0.0887	0.0245*	0.1651	0.0417*
Adjusted R^2	0.419		0.409	
F statistic	16.7*		16.5*	
Sample Size	565		565	

respectively. F-tests indicate that OLS model 2 is a significant improvement over OLS model 1 ($F(6, 538) = 7.06$, 5% critical value = 2.10). The instruments (set A) used in the IV estimation procedures are interviewed partly in English and partly in another language and interviewed wholly in another language.

rewarding such ability. Since these skills represent a higher level of language ability, than speaking fluency alone, and are more highly valued in the labour market (see Chiswick 1991; Dustmann 1994), classifying language skills according to speaking fluency alone is likely to overestimate the true level of language ability resulting in downwardly biased OLS estimates.[17]

According to Modood et al. 1997, the interviewers received specific training for this survey, the interviews took place face to face, usually between members of the same broad ethnic group, and the lasted on average 50.5 minutes. Therefore, it might be the case that the extent of any measurement error would be less than that in self-reported data. Unfortunately, are unable to identify individual interviewers in our dataset and so cannot control for interviewer-specific effects.

The use of instrumental variable estimation (IV) provides one method to account for these potential problems and thus allow us to more accurately assess the true impact of language fluency on occupational success. The practical difficulty with IV estimation is finding an instrument or set of instruments that are significant correlated with language fluency but also orthogonal to the residuals of the main equation (in our case occupational success). Moreover, a number of recent studies have questioned the interpretation that can be given to IV estimates, and their general usefulness for policy evaluation. For example, Heckman (1997) has shown that both OLS and IV techniques require very restrictive assumptions in order to provide estimates of the average effect of 'treatment on the treated'. Angrist et al. (1996) argue that the only effect that IV can consistently estimate is the average treatment effect for those who change treatment status (i.e. improve their English speaking fluency) because they comply with the assignment-to-treatment implied by the instrument(s). They refer to this parameter as the 'local average treatment effect' (LATE).[18]

One implication of this possible interpretation is that different instruments should provide very different estimates of the effect of language fluency on occupational success. In order to investigate the robustness of our results, we estimate separate two-stage-least-squares (2SLS) models using two alternative sets of instruments. The first instruments (A in Table 4) we use are based on the language of interview, which could be wholly in English, partly in English and partly in another language or wholly in a language other than English. Following Chiswick and Miller (1992, 1995) and Chiswick (1998), our second set of instruments (B in Table 4) are whether the individuals is married to a UK born spouse and the number of dependent children in the household.

Before using IV estimation, Bound et al. (1995) suggest that the quality of the instruments should be checked in two ways as weak instruments may result in a large bias in the estimates. Firstly, any potential instruments should significantly improve the first-stage model determining the endogenous variable. In our case, the results from likelihood ratio tests suggest that the

Table 4. The partial effect of fluency in (speaking) the English language: IV estimates (sample as Table 1)

Variable	Model 2			
	OLS		IV	
	Coeff.	S.E.	Coeff.	S.E.
Whole sample (instrument set B; $n = 565$)	0.0887	0.0245*	0.1142	0.1159
Sub-samples (instrument set A)				
Some qualifications ($n = 364$)	0.1033	0.0346*	0.1964	0.0597*
UK highest qualification ($n = 223$)	0.1172	0.0625#	0.2804	0.1129*
Non-UK highest qualification ($n = 141$)	0.0499	0.0502	0.1723	0.0935#
No qualifications ($n = 201$)	0.0761	0.0311*	0.1447	0.0533*

Note: *, # and + indicate significance at the 5%, 10% and 20% levels, respectively. Instrument set A are interviewed partly in English and partly in another language and interviewed wholly in another language. Instrument set B are wife UK born, one, two, three and more than three dependent children (aged < 16).

inclusion of both instrument sets A and B (separately included) leads to a significant improvement in the explanatory power of the language fluency probit model at the 5% level of significance ($\chi^2(2) = 168.6, \chi^2(5) = 13.9$).[19] The second indicator suggested by Bound et al. (1995) is the increase in the adjusted-R^2 measure when the above exercise is carried out. The resulting partial R^2s suggest that adding instrument set B explains 2% more of the variation in language model A, but instrument set A appears much more powerful since it appears to explain more than 20% of the variation.[20]

In addition to the above tests, since each of our IV models is over-identified, with the number of exogenous instruments being greater than endogenous variables, we are able to compute the Sargan χ^2 statistics to test the general validity of the instrument sets. The results suggest that the IV estimates based on the language of interview instruments are valid ($\chi^2(1) = 0.894$, critical value = 3.84 at the 5% level), but that some caution should be given to the estimates based on our family-related instruments ($\chi^2(4) = 13.3$, critical value = 9.48 at the 5% level).

5.3. *IV estimates*

Table 3 provides the IV results for both the basic and extended models of occupational success, based on the language of interview instruments (set A), which is our preferred set of instruments given the results above. The results reveal a lot about the biases in our single-equation models. Whereas the coefficients on the exogenous variables remain virtually unchanged from our OLS results, in both IV model 1 and 2, the coefficient on fluency in (speaking) the English language increases by nearly a factor of two and the estimates retain statistical significance at the 5% level. In the extended model, for example, fluency is now found to increase the mean occupational wage, over an identical person who is not assessed as fluent, by 16.5%. This is roughly comparable to the extra return to possessing a degree level qualification rather than an A-level or its equivalent. The corresponding model 2 results (see Table 4) using our less powerful sets of instruments based on whether the wife is UK born and the number of dependant children in the household, confirm this downward bias in the OLS estimate, with the 'return' to English language fluency estimated at 11%.[21] However, given the weaker power of these instruments the coefficients are imprecisely estimated and not statistically significant at the 5% level.

Due to the nature of our occupational success variable, it is not possible to directly compare our results with those of previous studies that have examined the effect of language fluency on wages. It is interesting, however, to note that our estimates of the returns to speaking English fluently are similar in size to those estimated previously in wage equations for other countries.[22] Furthermore, these findings have recently been confirmed using the actual (banded) wage data from the Fourth National Survey of Ethnic Minorities (Dustmann and Fabbri 2000).

The above IV results may be taken to represent the average effect of English language speaking fluency on occupational success. However, it might also be the case that the returns to fluency are heterogenous amongst our sample and, in particular, that those for whom the reward to fluency is greatest self-select

into fluency whilst those who would not gain substantially from fluency choose not to acquire it. Chiswick and Miller (1999b) suggest using a selectivity correction procedure, with a binary logit fluency model as the first stage, to explore whether this is the case. Separate OLS log wage models, including as an additional explanatory variable the inverse mills-ratio from the logistic language model, are estimated for those who are fluent in English and those who are not. The coefficient on this additional variable attempts to capture unobservable heterogeneity between the fluent and non-fluent. They find, using the US 1990 Census of Population, evidence to suggest that the distribution of individuals according to fluency is non-random. Their estimates show a positive selection into each state indicating that English-language skills are endogenously determined in their earnings' models. We have followed this procedure using our measure of occupational success but find no statistically significant evidence of self-selection.[23]

An alternative approach is to utilise the fact that there are significant complementarities between language skills and educational skills (see Chiswick and Miller 1999b, for extensive tests of this hypothesis) and also between linguistic ability and the country where qualifications were obtained. Therefore we explore whether the returns to fluency vary by educational attainment and by whether qualifications were gained in the UK or not. The model 2 results of the OLS and IV (using instrument set A) estimates of the partial effects on language fluency for a number of sub-samples sorted by the highest qualification variables are reported in Table 4.[24] This process also has the benefit of providing direct estimates of the various returns to fluency that the selectivity-correction method does not. These results confirm that the OLS estimates are biased downwards, regardless of the sub-sample we choose. The coefficient on fluency increases in every case and is statistically significant. There are some differences in the IV estimates of the returns to fluency between those individuals with some qualifications and those with none, but they are not statistically significant (t-statistic $= 0.65$). Neither do we find a significant difference between immigrants possessing UK qualifications and those with foreign qualifications (t-statistic $= 0.74$) despite the 11% gap in estimated returns.

Overall, we have found considerable evidence to suggest that OLS provides downwardly biased estimates of the effect of English language fluency on occupational success. The additional experiments we have undertaken suggest that this is more likely to be the result of measurement error than the existence of unobserved heterogeneity affecting both occupational success and English language fluency. As we mentioned earlier, we believe that this could be the result of interviewers systematically over-estimating 'overall' language fluency, since they evaluate only English speaking fluency rather than reading and writing skills as well. Employers are likely to be concerned about overall language ability in determining progression up the occupational ladder.

6. Conclusions

In this paper we have estimated the determinants of English language speaking fluency for ethnic minority immigrant men, aged 22–64 years old, who live

in the metropolitan areas of England. Our sample derived from the Fourth National Survey of Ethnic minorities, conducted by the Policy Studies Institute in 1994, is not nationally representative. However, this data source is advantageous in that a member of the same ethnic group as the respondent conducted the interviews and that they could be undertaken wholly or partly in the respondents' preferred language. Furthermore, the interviewer assessed the language fluency of the individual, thus avoiding the self-reported measurement error endemic in similar studies.

Our results broadly confirm Chiswick and Miller's (1995) hypotheses for these immigrants in England. Increasing age at immigration reduces, and more years since immigration increases, language fluency. Black Caribbeans and African Asians are the most likely to be fluent, other things being equal, whilst Pakistanis, Indians and Bangladeshis have the lowest predicted probabilities of fluency. There are clearly synergies between education and fluency and linguistic benefits from residing in Greater London or in a Census ward with low own ethnic group density (0–5%). However, remitting money to the country of origin is significantly associated with lower probabilities of language fluency.

We found that English language speaking fluency is the second most important determinant of occupational success, after possession of a degree or equivalent highest qualification, amongst the immigrants in our sample. Using the method of instrumental variables we have shown that the ordinary least squares results under-estimate the importance of fluency to occupational success irrespective of the model, instruments or sub-sample used. Our most statistically reliable estimate suggests that being fluent in speaking the English language raises the mean occupational wage by approximately 16.5% compared to similar individuals who are not fluent. However, these results are based on a sample of men in the most favourable labour market positions, namely paid employment. If these individuals are more able or motivated than those not in paid employment it is likely that we have underestimated the effect of language fluency for these groups as whole.

Attaining fluency in speaking the English language may be one route out of the low-paid jobs currently occupied by many immigrants in English metropolitan areas. Furthermore, as Chiswick (1991) and Dustmann (1994) have shown for other countries, reading and writing language skills are likely to even more important than speaking skills for the labour market success of immigrants in Britain. The provision of comprehensive English language training for these groups could dramatically improve their current and future labour occupational attainment. One specific method would be to encourage the acquisition of vocational or formal qualifications in the United Kingdom, particularly amongst those with no qualifications. This would exploit the double benefit to labour market success from both more education and gaining English fluency, particularly amongst those with degree-level qualifications. Additionally, the United Kingdom could introduce an English language fluency requirement into its immigration policy or require immigrants to undertake compulsory English language training as a condition of residence. Since UK law currently requires all resident children to spend at least 10 years of their lives (in education) acquiring the skills necessary for labour market success, requiring labour market entrants from abroad to invest in the language skills essential for their occupational success should not be too controversial.

Appendix

Table A1. Proportion fluent in (speaking) the English language by characteristic (sample as Table 1)

Variable	Proportion Fluent	S.E.	*T*-stat.
Whole sample	0.582	0.021	–
<10 years old at immigration	0.785	0.040	5.41*
10–15 years old at immigration	0.535	0.045	1.24
16–19 years old at immigration	0.571	0.054	0.22
20–24 years old at immigration	0.577	0.045	0.13
>24 years old at immigration	0.473	0.044	2.84*
<15 years since immigration	0.370	0.044	5.39*
15–19 years since immigration	0.593	0.052	0.23
20–24 years since immigration	0.644	0.044	1.56+
25–29 years since immigration	0.636	0.044	1.35+
>30 years since immigration	0.672	0.044	2.28*
Indian	0.567	0.038	0.48
Bangladeshi	0.431	0.066	2.44*
Pakistani	0.379	0.045	5.04*
African Asian	0.675	0.043	2.40*
Black Caribbean	0.870	0.039	7.46*
Chinese	0.652	0.102	0.70
Wife UK born	0.574	0.064	0.14
Wife foreign born	0.564	0.024	1.58+
Not married	0.688	0.052	2.16*
No dependent children	0.623	0.039	1.22
One dependent child	0.677	0.042	2.52*
Two dependent children	0.635	0.040	1.54+
Three dependent children	0.450	0.056	2.56*
> Three dependent children	0.322	0.061	4.47*
Oldest child aged > 15	0.524	0.042	1.63+
Oldest child aged 12–15	0.516	0.053	1.37+
Oldest child aged 5–11	0.596	0.041	0.38
Oldest child aged 0–4	0.537	0.055	0.89
No children	0.743	0.044	3.95*
Degree qualification	0.833	0.033	7.88*
A-level qualification	0.690	0.072	1.55
Vocational qualification	0.789	0.038	5.77*
O-level qualification	0.488	0.056	1.85*
No qualifications	0.323	0.033	9.92*
Living in the Midlands	0.447	0.040	4.09*
Living in the North	0.462	0.049	2.73*
Living in Greater London	0.695	0.027	5.98*
0–5% own ethnic density	0.682	0.041	2.72*
5–15% own ethnic density	0.580	0.035	0.07
15–33% own ethnic density	0.528	0.039	1.64+
>33% own ethnic density	0.529	0.060	0.96
Long term health problem	0.409	0.107	1.65#
Remitter	0.551	0.039	0.97
Interview conducted wholly in English	0.881	0.018	24.9*
Interview conducted partly in English	0.215	0.034	12.4*
Interview conducted wholly in another language	0.025	0.018	23.3*
Sample Size	565		

Note: Fluency in (speaking) the English language is assessed by the interviewer, who is a member of the same ethnic group as the respondent and fluent in the respondent's other main language. *T*-stat is the (absolute value of the) *T*-statistic which tests for the difference between the mean fluency rates of the samples when each characteristic holds and when it does not (i.e. when the dummy variables are 1 and 0). S.E. stands for standard error. *, # and + indicate significance at the 5%, 10% and 20% levels, respectively (critical values = 1.96, 1.645, 1.282).

Endnotes

[1] Stewart (1983) investigated the role of racial discrimination in the occupational attainment of non-white immigrants in 1975, using the National Training Survey. He includes a poor speaking English dummy variable, and experience before and after immigration (and their respective squares) variables in his estimations of the determinants of the log of average hourly occupational earnings. Gazioglu (1996) examines the impact of English language fluency on the earnings of 280 Turkish and Bangladeshi male immigrants in London. However, these latter groups account for less than 10% of Britain's ethnic minority population.

[2] This is a similar definition to that used by Nickell (1982) and Stewart (1983) who mapped average hourly earnings by Occupational Unit Group, from the General Household Survey, into the National Training Survey. More recently Harper and Haq (1997) found no difference in their results according to whether they used wages or occupations ranked by mean hourly wages in their study of occupational attainment amongst British men.

[3] Alternatively, as pointed out by a referee, it may be the case that those who send remittances to their country of origin may have an incentive to maximise earnings, perhaps through investing in relevant skills.

[4] This series of surveys began in 1964 when the main immigrant groups were Black Caribbeans and South Asians. As a result subsequent surveys have focused exclusively on these main groups.

[5] See also Gazioglu (1996) for similar findings amongst Turkish and Bangladeshi male immigrants in London.

[6] Unfortunately, due to the limited sample of ethnic minority workers in the Labour Force Survey (even after pooling consecutive surveys) we are not able to reliably estimate our occupational success measure separately for such groups (over the 899 categories) or even at a more aggregated level (e.g., two digit).

[7] The highest qualification variables used in this study are our own derivations from the raw data. The questionnaire asks individuals to report all their UK qualifications (33 categories) and all their qualifications gained abroad (9 categories). We have ranked both sets of qualifications and computed the individuals highest UK qualification and highest foreign qualification. This gives the highest qualification for the majority of cases, who have either UK qualifications or foreign qualifications. For those with qualifications obtained in the UK as well as abroad, the UK qualification is taken as the highest qualification. This seems reasonable given that any UK education will have been undertaken at an older age and is therefore likely to be of a higher level. Simple checks on the data confirm this. Finally five dummy variables were created for degree, A-level, O-level and vocational qualifications or their equivalents and for no qualifications. The vocational category is unable to be sub-divided since the foreign qualification categories do no distinguish between different levels.

[8] We exclude the 14% of the sample aged over 64, the 53% who are women, the 53% of working age men who are not employees, the 49% native born employees and the 25% of immigrant employees who don't live in a metropolitan area.

[9] To some degree these figures reflect the fact that high ethnic minority density wards were oversampled in the survey.

[10] However, note that only 87% of Black Caribbeans are assessed as fluent in the English language. It is for this reason that they are included in our sample.

[11] The inclusion of years since immigration squared did not significantly improve the maximum likelihood of any of the models and thus this variable was omitted.

[12] See note to Table 2 for details.

[13] Likelihood ratio tests (Greene 1993, p. 647) indicate that the null hypotheses that the coefficients on the additional variables in model B (compared to model A; Likelihood Ratio statistic = 33.4) are jointly zero can be rejected at the 1% level ($\chi^2(15)$ critical value = 30.6).

[14] Chiswick and Miller (1996) find that controlling for the presence of ethnic press, having relatives in the country and the origin language of the spouse eliminates the effect of ethnic concentration in Australian data. However, in our dataset are unable to explore the importance of these factors further.

[15] An F-test rejects the inclusion of a years since immigration squared term in both model 1 (F-statistic = 0.50, $F(1, 547)$ critical value = 3.84) and model 2 (F-statistic = 0.31, $F(1, 541) = 3.84$) at the 5% level.

[16] Gazioglu (1996) found a 10–13% earnings benefit for self-reported good or very good English speaking amongst Turks and Bangladeshis, though, of course, our study concerns occupational success.

[17] Chiswick and Miller (1999a) provide some evidence of a complementarity effect for being able to both speak and read English.

[18] In our context, IV estimates of the effect of language fluency on occupational success using say, number of children in household, would then be interpreted as the average 'return' to fluency for an individual whose fluency level has improved only because of having more (or less) children in the household, but would not have improved otherwise.

[19] Similarly, Staiger and Stock (1997) suggests that the results from an IV procedure should be considered of dubious quality when the *F*-statistics from the first-stage OLS regression fall below 10.

[20] Given that our first-stage is a probit model rather than an OLS this test is based on the Pseudo-R^2_{ANN} (the Aldrich and Nelson (1984) measure normalised) which was proposed by Veall and Zimmermann (1992). Amongst the significance-of-fit class of pseudo-R^2s this measure most closely corresponds to the OLS-R^2. In particular, it is distributed evenly over the 0–1 range (see Veall and Zimmermann, 1996). The previous language fluency-wages literature does not provide any information in order to assess the relative power of our instruments with respect to theirs. The only available comparisons are with the recent returns to education literature, using actual individual wage data. For example, Harmon and Walker (1995) report a partial R^2 of 0.0046, Harmon and Walker (1999) find values between 0.0025 and 0.0078 and Ichino and Winter-Ebmer (1999) report a range of 0.003–0.114.

[21] For brevity we do not provide the full IV results based on this set of instruments (set B) nor those for model 1 use both instrument sets. However, the coefficients for the exogenous variables remain largely unchanged from the OLS model. The results are available from the authors on request.

[22] Chiswick and Miller (1992), for example, found an increase in the partial effect of language fluency on earnings (*t*-ratios in parentheses) from 0.169 (12.52, OLS) to 0.571 (5.43, IV) using 1980 United States data and veteran status, foreign marriage, children and minority language concentration measures as identifying instruments. They also noted a change from 0.122 (2.43, OLS) to 0.414 (1.34, IV) amongst immigrants in 1981 Canadian data with foreign marriage and minority language concentration measures as identifying instruments. In Australia the results changed from 0.052 (2.52, OLS) in 1981 and 0.083 (4.75, OLS) in 1986 to −0.243 (1.20, IV) and 0.043 (0.52, IV), respectively, with foreign marriage, number and age of children and minority language concentration measures as identifying instruments. Chiswick (1998) found an increase from 0.110 (12.66, OLS) to 0.351 (4.25, IV) using 1983 data from Israel using Tel Aviv, Jerusalem, foreign marriage, number of children, and minority language concentration measures as identifying instruments. Finally, Dustmann and van Soest (1998a), using German Socio-economic panel data between 1984–1993 found a language effect on earnings increase from 0.0538 (7.08, OLS) to 0.155 (2.28, IV) with father's education measures as identifying instruments.

[23] This may be a weak effect not picked up due to our small sample sizes (there are only 329 fluent and 239 non-fluent individuals in our data) or because the dependant variable is group-specific mean wages rather than actual individual wages. Alternatively, the power of our identifying variables may be weak – they are instrument set B plus the remitter, firm size and industrial sector variables. Unfortunately, we do not have the linguistic distance measure that Chiswick and Miller (1999b) use in our data.

[24] The results are very similar to those using model 1, which are therefore not reported but are available from the authors on request.

References

Aldrich JH, Nelson FD (1984) *Linear Probability, Logit and Probit Models*. Sage University Press, Beverly Hills

Angrist JD, Imbens GW, Rubin DB (1996) Identification of Casual Effects Using Instrumental Variables. *Journal of the American Statistical Association* 91:444–455

Blackaby DH, Clark K, Leslie DG, Murphy PD (1994) Black-White Male Earnings and Employment Prospects in the 1970s and 1980s. *Economics Letters* 46:273–279

Blackaby DH, Drinkwater S, Leslie DG, Murphy PD (1997) A Picture of Male and Female Unemployment Among Britain's Ethnic Minorities. *Scottish Journal of Political Economy* 44:82–197

Blackaby DH, Leslie DG, Murphy PD, O'Leary NC (1998) The Ethnic Wage Gap and Employment Differentials in the 1990s: Evidence for Britain. *Economics Letters* 58:97–103

Bound J, Jaeger DA, Baker R (1995) Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Variable is Weak. *Journal of the American Statistical Association* 90:443–450

Borjas GJ (1985) Assimilation, Changes in Cohort Quality and the Earnings of Immigrants. *Journal of Labor Economics* 3(4):321–342

Borjas GJ (1987) Self-Selection and the Earnings of Immigrants. *American Economic Review* 77:531–553

Chiswick BR (1978) The Effect of Americanisation on the Earnings of Foreign-Born Men. *Journal of Political Economy* 86:897–921

Chiswick BR (1991) Speaking, Reading and Earnings among Low-Skilled Immigrants. *Journal of Labor Economics* 9(2):149–170

Chiswick BR (1998) Hebrew Language Usage: Determinants and Effects on Earnings Among Immigrants in Israel. *Journal of Population Economics* 11(2):253–271

Chiswick BR, Miller PW (1992) Language in the Immigrant Labor Market. In: Chiswick BR (ed) *Immigration, Language and Ethnicity: Canada and the United States.* American Enterprise Institute, Washington

Chiswick BR, Miller PW (1995) The Endogeneity Between Language and Earnings: International Analyses. *Journal of Labor Economics* 13(2):246–288

Chiswick BR, Miller PW (1996) Ethnic Networks and Language Proficiency Among Immigrants. *Journal of Population Economics* 9(1):19–35

Chiswick BR, Miller PW (1998) English Language Fluency Among Immigrants in the United States. *Research in Labor Economics* 17:151–200

Chiswick BR, Miller PW (1999a) Language Skills and Earnings Among Legalized Aliens. *Journal of Population Economics* 12(1):63–89

Chiswick BR, Miller PW (1999b) Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle. Paper presented at the *Centre for Economic Policy Research Conference "Marginal Labour Markets in Metropolitan Areas"*. Dublin, October 10–12, 1999

Dustmann C (1994) Speaking Fluency, Writing Fluency and Earnings of Migrants. *Journal of Population Economics* 7:133–156

Dustmann C (1997) The Effects of Education, Parental Background and Ethnic Concentration on Language. *The Quarterly Review of Economics and Finance* 37 (Special Issue):245–262

Dustmann C (1999) Temporary Migration, Human Capital and Language Fluency of Migrants. *Scandinavian Journal of Economics* 101:297–314

Dustmann C, Fabbri F (2000) Language Proficiency and Labour Market Performance of Immigrants in the UK. *Institute for the Study of Labor (IZA)* Discussion paper no. 156

Dustmann C, van Soest A (1998a) Language and the Earnings of Immigrants. mimeo, University College London, September

Dustmann C, van Soest A (1998b) Language Fluency and Earnings: Estimation With Misclassified Language Indicators. mimeo, University College London, October

Galor O, Stark O (1990) Migrants' Savings, the Probability of Return Migration and Migrants' Performance. *International Economic Review* 31:463–467

Gazioglu S (1996) English Language Proficiency and the Earnings of Turkish and Bangladeshi Immigrants in London. In: Gazioglu S (ed) *Migrants in the European Labour Market*, J-Net, Aberdeen

Harmon C, Walker I (1995) Estimates of the Economic Return to Schooling for the United Kingdom. *American Economic Review* 85:1278–1286

Harmon C, Walker I (1999) The Marginal and Average Returns to Schooling in the UK. *European Economic Review* 43:879–887

Harper B, Haq M (1997) Occupational Attainment of Men in Britain. *Oxford Economic Papers* 47:638–650

Hatton TJ, Wheatley Price S (1999) Migration, Migrants and Policy in the United Kingdom. Institute for the Study of Labor (IZA) Discussion paper no. 81

Heckman JJ (1997) Instrumental Variables: A Study of Implicit Behavioural Assumptions Used in Making Program Evaluations. *Journal of Human Resources* 32:441–462

Ichino A, Winter-Ebmer R (1999) Lower Bounds and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation With Different Instruments. *European Economic Review* 43:889–901

McFadden D (1973) Conditional Logit Analysis of Qualitative Choice Behaviour. In: Zaremka P (ed) *Frontiers in Econometrics*. Academic Press, New York

Merkle L, Zimmermann ZF (1992) Savings, Remittances and Return Migration. *Economics Letters* 38:77–81

Modood T, Berthoud R et al. (1997) *Ethnic Minorities in Britain*. Policy Studies Institute, London

Nickell SJ (1982) The Determinants of Occupational Success in Britain. *Review of Economic Studies* 49:43–53

Owen D (1992) Ethnic Minorities in Great Britain: Settlement Patterns. Centre for Research in Ethnic Relations, 1991 Census Statistical Paper No. 1, University of Warwick, November

Owen D (1993) Country of Birth: Settlement Patterns. Centre for Research in Ethnic Relations, 1991 Census Statistical Paper No. 5, University of Warwick, December

Pudney SE, Shields MA (2000) Gender, Race, Pay and Promotion in the British Nursing Profession: Estimation of a Generalised Ordered Probit Model. *Journal of Applied Econometrics* 15:367–399

Sargan JD (1964) Wages and Prices in the United Kingdom: A Study in Econometric Methodology. In: Hart PE, Mills G, Whitaker JK (eds.) *Econometric Analysis for National Economic Planning*. Butterworth, London

Shields MA, Wheatley Price S (1998) The Earnings of Male Immigrants in England: Evidence from the Quarterly LFS. *Applied Economics* 30:1157–1168

Shields MA, Wheatley Price S (1999a) Ethnic Differences in British Employer-Funded on and Off-the-Job Training. *Applied Economics Letters* 6:421–429

Shields MA, Wheatley Price S (1999b) Ethnic Differences in the Incidence and Determinants of Employer-Funded Training in Britain. *Scottish Journal of Political Economy* 46:523–551

Staiger D, Stock Q (1997) Instrumental Variables Regression With Weak Instruments. *Econometrica* 65:557–586

Stewart MB (1983) Racial Discrimination and Occupational Attainment in Britain. *Economic Journal* 93:521–541

Veall MR, Zimmermann KF (1992) Pseudo-R^2's in the Ordinal Probit Model. *Journal of Mathematical Sociology* 16:332–342

Veall MR, Zimmermann KF (1996) Pseudo-R^2 Measure for Some Common Limited Dependent Variable Models. *Journal of Economic Surveys* 10:241–259

Wheatley Price S (2001a) The Employment Adjustment of Male Immigrants in England. *Journal of Population Economics* 14:193–220

Wheatley Price S (2001b) The Unemployment Experience of Male Immigrants in England. *Applied Economics* 33:201–215