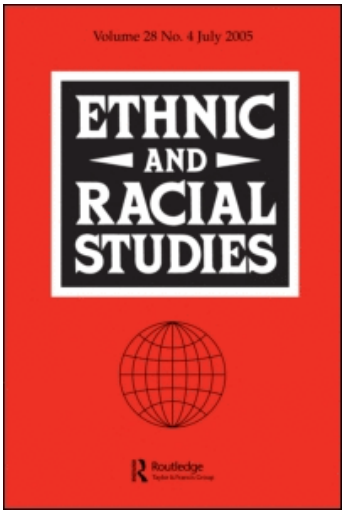


Ethnic and Racial Studies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713685087

The conditional relationship between English language proficiency and earnings among US immigrants

Sean-Shong Hwang ; Juan Xi ;Yue Cao

First published on: 29 March 2010

To cite this Article Hwang, Sean-Shong , Xi, Juan andCao, Yue(2010) 'The conditional relationship between English language proficiency and earnings among US immigrants', Ethnic and Racial Studies,, First published on: 29 March 2010 (iFirst)
To link to this Article: DOI: 10.1080/01419871003642375
URL: http://dx.doi.org/10.1080/01419871003642375

Ethnic and Racial Studies 2010 pp. 1–28, iFirst article

The conditional relationship between English language proficiency and earnings among US immigrants

Sean-Shong Hwang, Juan Xi and Yue Cao

(*First submission April 2009; First published March 2010*)

Abstract

Using the 2000 US census data for immigrants of twenty language groups resided in metropolitan areas, we test the hypothesis that the rate of returns (in earnings) to English proficiency is not constant but varies with the language environment (as defined by group size, segregation, linguistic heterogeneity and inequality) in which immigrants are embedded. Results from our hierarchical model indicate that while an increase in the size and segregation of the language group diminishes returns to English proficiency, a rise in linguistic heterogeneity and inequality in the metropolitan area has the opposite effects. This study expands the scope of the previous studies by identifying conditions under which returns to English proficiency among immigrants are modified by a set of contextual factors often overlooked.

Keywords: Immigrants; English proficiency; earnings; heterogeneity; segregation; linguistic environment.

Being able to speak the language of the host society affects not only an immigrant's ability to interact with members of other groups, but also his/her economic viability in the labour market. Sociological theories which see host society's language as a means of intergroup interactions should, therefore, be useful for understanding host society's language as a human capital. Because there is little overlap between the two literatures, the potential usefulness of sociological insights about intergroup relations for labour economic studies has not been fully realized. Building on past studies of intergroup relations and economic

ISSN 0141-9870 print/1466-4356 online
DOI: 10.1080/01419871003642375

Routledge
Taylor & Francis Group

assimilation, this study demonstrates the utility of a structural theory of intergroup relations for a major area of study in labour economics.

One dominant view in labour economics is that the success of immigrants in the host society's labour market hinges on their ability to speak its language (e.g. McManus, Gould and Welch 1983; Grenier 1984; Tainer 1988; Carliner 2000; Dustmann and Van Soest 2002). Increasing evidence, however, suggests that the positive association between immigrants' English proficiency and economic success is anything but unconditional (Portes and Bach 1985; Jasso and Rosenzweig 1990; McManus 1990; Chiswick and Miller 1995, 2002).

Building on Blau's (1977) structural theory, a recent study by Hwang and Xi (2008) demonstrated that US immigrants' English-speaking ability was affected by their size and segregation, and by the heterogeneity and inequality of the metropolitan area in which they lived. We extend their work by arguing that economic returns to English proficiency among immigrants are contingent upon the size and segregation of a language group, which tend to reduce the returns to English proficiency; and upon heterogeneity and inequality of the community in which immigrants reside, which tend to augment the returns. The empirical validity of these hypotheses is examined using the 2000 US census for twenty language groups resided in metropolitan areas.

English-language advantage hypothesis

Ability to speak the language of the host society has been conceptualized as a human capital reflecting the employability and earnings potentials of immigrants in the host society's labour market (Mincer 1958). Learning English is considered a worthwhile investment because it promises to improve immigrant workers' marketability (Chiswick 1991). Based on this conceptualization, immigrants who possess the language skills of the host society are expected to be more successful in the labour market because of their greater ability to convert past education and work experiences into better economic returns (Chiswick and Miller 1995).

Being able to speak English fluently is said to open up new job opportunities that are otherwise inaccessible to immigrants (Borjas 1990). Portes and Bach's (1985) study of Cuban workers in Miami found that knowledge of English facilitated the entry of Cuban refugees into the primary labour market where pay was generally better. Because the productivity of workers in the primary labour market depends on their ability to communicate effectively with supervisors, co-workers and the clientele they serve, workers with better English skills are better rewarded (McManus, Gould and Welch 1983; Tainer 1988). Workers who possess the language skills required

for the workplace are also preferred by employers because they reduce the costs of job training and supervision.

Overall, it seems that these arguments make a convincing case for the *English-language advantage* hypothesis. English proficiency promoted acculturation, served as a catalyst for other human capital resources, increased the number of jobs available to immigrants and increased work performance.

The validity of *English-language advantage* hypothesis, however, rests on the assumption that English language proficiency is an indispensable human capital in the labour market. However, the assumption's validity seems to depend on the *language environment* of where immigrant workers live and work. Being able to speak English fluently is crucial for immigrant workers in labour markets where English is the primary language to conduct business; the value of English proficiency is likely to be discounted in environments where non-English languages are accepted or even preferred (Jasso and Rosenzweig 1990).

Challenges to the English advantage hypothesis

Not denying the importance of English proficiency as a means to economic success in mainstream economy for immigrants, increasing evidence, however, points to the conditional nature of the association between the two. Wilson and Portes (1980), for example, pointed out that knowledge of English *failed* to reliably predict income of Cuban immigrant workers in Miami's enclave economy. Jasso and Rosenzweig (1990) indicated that although English deficiency significantly lowered the economic returns to German and Hispanic immigrants to the US in the past, these immigrants could overcome their disadvantages by moving to areas where a higher proportion of German and Spanish speakers resided. Similar findings were reported by Chiswick and Miller (2002), indicating that although fluency in English gave immigrants to the US an advantage, such advantage was significantly reduced in communities with greater concentration of non-English speakers.

McManus (1990) indicated that Hispanic men who work in ethnic enclaves were able to find jobs despite their English deficiency. Although English-deficient Hispanic male workers in integrated communities tended to crowd into menial, non-verbal jobs, their counterparts in communities with large Hispanic population enjoyed greatly expanded occupational choices. Immigrants who live in ethnic communities find little need for English when their native tongue is the primary means of communication (Borjas 1990). When employers, co-workers and customers share the same national origin and speak a common non-English language, learning English becomes superfluous.

In sum, the evidence suggests that the presumed positive effects of English proficiency on immigrants' economic success seem to be contingent upon the language environment. Although the importance of English is unquestionable in communities where English speakers dominate, such importance is likely to reduce in communities where the ethnic population is large and residentially segregated from English speakers and other groups (Carliner 2000).

A multilevel conceptual model and derived hypotheses

The human capital model (Mincer 1958) has been the dominant conceptual framework guiding empirical analyses of immigrants' economic performance. Initially proposed to explain variations in performance of workers in the labour market (Becker 1993), the model was later applied to examine differences in earnings between native- and foreign-born workers and the process through which the gap is narrowed over time (Chiswick 1978). According to this model, the success of workers in the labour market depends on the possession of knowledge, skills and experiences (or human capital, for short) needed for the job they do. The gap in earnings between immigrants and native workers exists not because of immigrants' inferior human capital but because human capital acquired in one country is often not transferable to a different country (Chiswick 1978; Friedberg 2000; Zeng and Xie 2004). Although early formulations of human capital models focused on education and work experiences (see Chiswick 1978; Becker 1993), proficiency in the host society's language was later recognized because it enables immigrants to convert their previous schooling and work experiences into full market values in the host society (Grenier 1984; Carliner 2000).

Although a useful model for explaining *individual* variations in earnings, previous applications of the model have not paid enough attention to *group* and *contextual* variations in the rate of returns to English proficiency. Several studies have shown that the returns to English language proficiency are contingent upon the size of non-English speakers in the community (Evans 1989; McManus 1990; Jasso and Rosenzweig 1990; Chiswick and Miller 1995; 2002). Even in a host society where English dominates, non-English language gains premiums in communities where the non-English population size is relatively large (Breton 1964). When supervisors and shopkeepers speak the same non-English language as their workers and consumers, the job performance of workers and shopkeepers would not be affected by their poor English. Thus, ability to speak English becomes irrelevant as a criterion in determining how workers should be rewarded.

Besides group size, returns to English fluency are expected to be contingent upon other structural attributes which are known to affect immigrants' ability to speak English (Hwang and Xi 2008). For example, while an increase in the size of a single minority language group in the community may reduce the use of English by members of the group, simultaneous increases in the size of multiple language groups can have an opposite effect. The presence of multiple language groups in the community increases the opportunities of fortuitous encounters between members of different groups and necessitates the use of English as a *lingua franca*. Thus, linguistic heterogeneity is expected to exert a pressure on non-English speakers to speak English, an effect that is opposite to that of group size (Hwang and Xi 2008).

Peter Blau (Blau 1977; Blau, Blum and Schwartz 1982) has made the greatest contribution in advancing this line of reasoning, arguing that intergroup relations are affected by such macro-structural variables as group size, heterogeneity, segregation and inequality. With regard to group size, he argues that despite the prevailing preferences for people to associate with others sharing the same traits, group size imposes a structural constraint on the availability of the same group members with whom one can interact. Thus, while members of a large group can easily find others who speak the same language in the community, the likelihood of being able to do so is small for members of smaller groups. Being able to speak the language of the host society as a means of interaction is therefore more crucial for members of smaller than larger groups.

In addition, Blau (1977) suggests that the potential for members of one language group to associate with outsiders depends on heterogeneity, or 'the chance expectation that two randomly chosen persons do not belong to the same group' (Blau, Blum and Schwartz 1982, p. 46). The chance that two persons randomly chosen from a community do not speak the same language rises as the number of language groups increases and the size of these groups are equal.

Blau theorizes that heterogeneity promotes intergroup associations, including interactions in workplace and business exchanges. Accordingly, the need for English as a *lingua franca* is increased in communities where linguistic heterogeneity is high. Although the negative association between group size and returns to English proficiency has been noted (Chiswick and Miller1995), little is known about the implications of simultaneous increase in the size of multiple groups. We argue that simultaneous increases in the size of multiple minority groups should have a positive effect. As fortuitous encounters between persons who speak different languages rise, the need for English as a *lingua franca* is expected to be heightened, thus heightening the importance of English language skills as human capital.

The heterogeneity hypothesis questions the validity of a simplistic explanation focusing solely on group size. Thus, an increase in Spanish speakers in a community, for example, is expected to boost the importance of Spanish and lower that of English only when there are no other minority language groups of significant size in the community. Living in a highly heterogeneous community significantly increases the likelihood of out-group contacts in workplace and in business transactions; immigrants who speak English fluently are therefore more advantageous and should be better rewarded.

Segregation is yet another aspect of community structure that affects the importance of English as means of intergroup relations. Blau (1977) maintains that interactions between two persons depend on opportunities for social contacts and physical propinquity increases such opportunities. Segregation has been found to affect different aspects of intergroup coexistence such as interracial marriages (Hwang, Saenz and Aguirre 1997), viability of ethnic enterprises (Aldrich et al. 1985) and retention of native language among immigrants (Stevens 1992; Hwang and Xi 2008). The clustering of non-English-speaking immigrants in segregated neighbourhoods is bound to suppress intergroup interactions (Massey and Denton 1988) and the use of English. Thus, the usefulness of English language skills in workplace is expected to decline in segregated neighbourhoods. A segregated immigrant community promotes in-group interactions, which in turn lowers the likelihood that English language proficiency is used as a criterion for rewarding immigrant workers.

Finally, Blau (1977) maintains that intergroup relation is affected by inequality, although the exact mechanism has been debated (Rytina et al.1988). While income inequality between immigrants and native speakers is expected to suppress intergroup relations because of homophily (Hwang and Murdock 1998), the inequality may paradoxically enhance random encounters between natives and immigrants in complementary roles or symbiotic relations (cf. Rytina et al. 1988) and thus necessitates the learning of English. Thus, being able to speak English fluently is more essential for immigrant workers in unequal settings than in equitable ones. Although derived from a different theory, the hypothesis is consistent with Chiswick and Miller's (1995) argument that immigrants are more motivated to learn English when they noticed that native speakers earned a lot more than themselves.

In sum, Blau's theory makes clear that the extent to which immigrants adopt the host society's language is determined by group size, linguistic heterogeneity, segregation and inequality, which together define the language environment. Although the theory has largely been applied to study intergroup relations, it is our contention that it is also useful for predicting the importance of English as a human capital. In both cases, English is a communication tool for

other ends. As such, Blau's theory which specifies conditions under which the importance of English is varied provides useful qualifications for the English-language advantage hypothesis.

Based on Blau's theoretical arguments, we hypothesize that the importance of English, as measured by the rate of return in earnings to English proficiency, is conditioned by the same language environment factors that affect immigrants' English proficiency. It is not our intention to challenge the premise that English proficiency is a determining factor of immigrants' earnings; we are simply making a logical extension of the premise by suggesting that the effects of English proficiency on earnings are conditioned by language environment. Thus, we predict that rate of return to English proficiency would be lower for members of larger and more segregated ethnic groups because large ethnic group size and high level of segregation reduce the importance of English. On the other hand, the rate of return to English would be higher for members of ethnic groups who resided in metropolitan areas characterized by higher degrees of linguistic heterogeneity and income equality because such conditions heighten English's importance.

Data and methods

The data for this study come from two sources: the 5 per cent Public-Use Micro Data Samples (PUMS) of the 2000 US census, which supplied data for our outcome and individual-level independent variables; and the Summary File 3 (SF3) (US Bureau of the Census 2002), which we used to compute group- and MSA-specific structural variables. From PUMS, we identified all foreign-born individuals as potential subjects for the analysis. Because the outcome variable of interest is earnings, we restricted our sample to those individuals aged between twenty-five and sixty-four who had positive earnings in 1999. Since our primary interest is the variations of English language proficiency as a predictor of earnings across groups and MSAs, we further restrict the sample to members of the twenty largest non-English language groups resided in metropolitan US.[1] The twenty language groups together represent over 90 per cent of all people who reported speaking a non-English language at home. Our individual level data include 576,381 observations.

Measurement

Our outcome variable of interest is the total earnings of the immigrant in 1999. Earnings include both wages and salary incomes and income from self-employment. The variable is log transformed in our analysis to reduce skewness.

Our independent variables at individual level are those used in human capital models (Chiswick and Miller 1995). They include English language proficiency, years of education, years of experience and years in US. English language proficiency is the variable of our primary interest. Census 2000 contains three questions related to language use. The first (question 11.a) asked respondents whether they spoke a language other than English at home. Those who answered 'yes' to the question were asked to name the language (question 11.b) and to assess how well they spoke English using an ordinal scale (question 11.c). To avoid selection biases, we also include respondents who speak only English at home in the analysis and assume that they speak English very well. Because the language group membership of these individuals is unknown, we assign them to a group based on the dominant language spoken in their country of origin. Although past analyses typically transformed English language proficiency into a dummy variable (Chiswick and Miller 2002), we use the variable in its original ordinal scale. Following Chiswick (1978), years of EDUCATION is measured as an interval variable. The census provides no direct measure of years of EXPERIENCE; we approximate it by the equation: EXPERIENCE = (AGE – EDUCATION – 5) or 0, whichever is bigger. Years in US (YEARINUS) = (2000) – year of entry).

We also control for gender (MALE), marital status (MARRIED), white-collar occupations (WCOCC) and self employment status (SELFEMPLOY) owing to their known associations with both English proficiency and earnings (Spener and Bean 1999).

To test hypotheses derived from Blau's theory require measures of group and community characteristics. Group size and segregation are measured for each group in each MSA instead of using national-level measures which ignore between-MSA variations in group characteristics. Heterogeneity and inequality, on the other hand, are characteristics of the community shared by all groups in the same community.

The Summary File 3 of the census 2000 is used to measure these macro-level variables. Our choices of 'community' units are limited to two available options Public Use Microdata Area (PUMA) and Metropolitan Statistical Area (MSA) in the PUMS data. We choose MSA because MSAs are larger units which take functional integration of all residents in the MSA into consideration in their definitions (Siegel and Swanson 2004) and therefore better reflect the language environment. In addition, metropolitan area is by far the most widely used approximation of urban community in the literature (Massey and Denton 1993) and the use is consistent with Blau's previous work on intermarriage (e.g. Blau, Blum and Schwartz 1982).

For each of the language groups in each MSA, a measure of group size and a measure of segregation are obtained.[2] The group size for language group *j* reflects the number of people who spoke language *j* at home in the MSA. We log transformed SIZE to reduce skewness.

SEGREGATION for language group j is measured by White's (1983) *average proximity index* (P_{xx}) using tract-level data for language spoken at home:

$$p_{xx} = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{x_i x_j c_{ij}}{X^2}$$

The index measures average proximity between members of the same group x in different census tracts i and j given the total number of X members in the MSA. The c_{ij} in the numerator is the negative exponential function of distance between census tracts i and j.[3] Among the many possible measures of segregation (Massey and Denton 1988), we considered spatial proximity index most appropriate. Unlike other measures which compare group differences in spatial distribution across areal units, this index focuses on spatial clustering of minority areas (White 1983). Because of the latter attribute, the measure has been touted as a proxy for ethnic enclave and was recommended as an indicator of potentials for intergroup associations (Massey and Denton 1988; Siegel and Swanson 2004).

Of the two community characteristics shared by all groups in the MSA, linguistic HETEROGENEITY is measured by $1-\Sigma P_i^2$. P_i stands for the proportion of the MSA's population who speak language i. It is a widely used measure of diversity in social science literature (Blau, Blum and Schwartz 1982). INEQUALITY is measured by the *ratio of the median income* for natives vs. that for foreign born individuals in the MSA.

In addition, we also control for two labour market variables: MSA's total POPULATION and its median earnings (MEDEARN). These variables are expected to affect immigrants' earnings by affecting the overall wage rate and costs of living in the MSA (Tienda and Wilson 1992). Controlling MSA's median earnings also serves the purpose of minimizing the risk of selectivity because immigrants with better English skills may incline to move to areas where wages are higher. Total MSA population is log transformed to remove skewness.

Analytic strategy

We will first examine group- and MSA-variations in returns to English language proficiency by fitting the human-capital model to each language group in each MSA where the group has a minimum of twenty-five persons who meet the inclusion criteria in terms of age and earnings. Our next task is to examine the extent to which such variations can be explained by group- and community-level variables derived from Blau's model.

Our data have a three-level hierarchical structure in which individuals are nested within groups and groups are, in turn, nested within MSAs. The nesting of lower-level units under upper-level units makes it clear that those lower-units in the same upper-level unit are not independent. Thus, we use hierarchical linear modelling (HLM), which takes into consideration the *intraclass correlation* between units and adjusts for its effects accordingly (Raudenbush and Bryk 2002).

Our hierarchical model consists of one level 1 equation specified according to the human capital model, and a set of level 2 and level 3 equations which use group- and MSA-level variables to explain variations in the slope of English language proficiency and the intercept. Following the convention of HLM, coefficients for different levels are distinguished by different symbols. Our level 1 model takes the following form:

$$\begin{aligned} Y_{ijk} = {} & \pi_{0jk} + \pi_{1jk}\mathrm{ENGLISH}_{ijk} + \pi_{2jk}\mathrm{MALE}_{ijk} + \pi_{3jk}\mathrm{EDUCATION}_{ijk} \\ & + \pi_{4jk}\mathrm{WCOCC}_{\mathrm{ijk}} + \pi_{5\mathrm{jk}}\mathrm{MARRIED}_{\mathrm{ijk}} + \pi_{6\mathrm{jk}}\mathrm{SELFEMPLOY}_{\mathrm{ijk}} \\ & + \pi_{7\mathrm{jk}}\mathrm{EXPERIENCE}_{\mathrm{ijk}} + \pi_{8\mathrm{jk}}(\mathrm{EXPERENCE}_{\mathrm{ijk}})^2 \\ & + \pi_{9\mathrm{jk}}\mathrm{YEARINUS}_{\mathrm{ijk}} + \pi_{10\mathrm{jk}}(\mathrm{YEARINUS}_{\mathrm{ijk}})^2 + \mathrm{e}_{\mathrm{ijk}} \end{aligned} \tag{1}$$

The $\boldsymbol{Y}_{ijk}$ in the equation is the expected log earnings for individual i in group j and in MSA k. The π_{0jk} is the intercept or the average log earnings for group j in MSA k after adjusting for first-level predictors. This interpretation of intercept is justified because all of the first-level independent variables have been grand mean-centred (Hofmann and Gavin 1998). Of all the π coefficients, the one that we are most interested in is π_{1jk}, or the slope of English language proficiency. We included the quadratic terms of EXPERIENCE and YEARINUS in the equation because past studies have shown curvilinear associations between earnings and these variables (Chiswick 1978). The last term in the equation, e_{ijk} represents the individual level residual.

Our level 2 model consists of eleven equations:

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}\mathrm{SIZE}_{jk} + \beta_{02k}\mathrm{SEGREGATION}_{jk} + r_{0jk} \tag{2.1}$$

$$\pi_{1jk} = \beta_{10k} + \beta_{11k}\mathrm{SIZE}_{jk} + \beta_{12k}\mathrm{SEGREGATION}_{jk} + r_{1jk} \tag{2.2}$$

$$\pi_{pjk} = \beta_{p0k}, \text{ Where } p = 2, 3, 4, 5, 6, 7, 8, 9, 10 \tag{2.3–2.11}$$

Of the eleven level 2 equations, equation (2.2) is our focus. It tests the hypotheses that variations in the rate of return to English language proficiency (π_{1jk}) across 1,475 MSA groups are associated with group size and segregation. Equation (2.2) is equivalent to the more familiar *interaction* hypotheses: English proficiency interact with group size

and segregation (Raudenbush and Bryk 2002).[4] While theoretically justified, it is unconventional to include interaction terms but not their main effects in a model. We evaluate the main effects of group size and segregation on group earnings using equation (2.1).[5] Because the slopes for other human capital variables are not the focus of this study, they are assumed to be fixed.

Our level 3 model consists of fifteen equations:

$$\beta_{00k} = \gamma_{000} + \gamma_{001}\text{HETERROGENEITY}_k + \gamma_{002}\text{INEQUALITY}_k + \gamma_{003}\text{POPULATION}_k + \gamma_{004}\text{MEDEARAN}_k + u_{00k} \quad (3.1)$$

$$\beta_{10k} = \gamma_{100} + \gamma_{101}\text{HETEROGENEITY}_k + \gamma_{102}\text{INEQUALITY}_k + u_{10k} \quad (3.2)$$

$$\beta_{p0k} = \gamma_{p00}, \quad \text{While } p = 2, 3, 4, 5, 6, 7, 8, 9, 10 \quad (3.3\text{–}3.11)$$

$$\beta_{0qk} = \gamma_{0q0}, \text{ and } \beta_{1qk} = \gamma_{1q0}, \text{ While q refers to SIZE \& SEGREGATION} \quad (3.12\text{–}3.15)$$

Our focus is on equation (3.2), which explains variations in the rate of return to English proficiency (β_{10k}) across MSAs using as predictors the MSA-level measures of heterogeneity and inequality. The equation asserts that there are significant *interactions* between English proficiency and the two MSA-level predictors. We also include equation (3.1) because variation in the average earnings across MSAs (β_{00k}) is likely to be affected by MSA characteristics.[6] The coefficients associated with other human capital factors are assumed fixed at level 3.

Our hypotheses about the effects of group size and segregation on rate of return to English proficiency would be confirmed if β_{11k}, β_{12k}, in equation (2.2) indicate a weaker association between English language proficiency and earnings as size and segregation increase. And our hypotheses about linguistic heterogeneity and inequality would be confirmed if γ_{101} and γ_{102} in equation (3.2) indicate a stronger association between English language fluency and earnings in communities with greater degrees of linguistic heterogeneity and inequality.

Results

Because this study is based on the premise that there are noticeable variations in the rate of return to English proficiency across groups and MSAs, it seems necessary to examine the empirical validity of the assumption first. We did so by fitting the human capital model – i.e.

equation (1) – with OLS regression to each of the 1,474 MSA-specific language groups.

Table 1 summarizes the results of the analyses. Because the focus of this study is variations in the slope of English proficiency, the effects of other variables are not shown. It is apparent that there are variations in the slope of interest both across groups and across MSAs. For example, for a total of 241 models (MSAs) we have examined for Spanish speakers, the slope varies between −0.46 and 0.73, with a mean of 0.07 and a SD of 0.14. Only 43.6 per cent of the 241 estimated slopes are positive and statistically significant. Between-MSA variations in the slope are also evident for other groups. For the twelve MSAs with a large enough Armenian population, for example, the slope ranges from −0.60 to 0.85, with a mean 0f 0.16 and a standard deviation of 0.36. And only three of the twelve slopes are positive and statistically significant.

In addition, there are noticeable variations across groups in the average value of the slopes. For example, the mean slope ranges from 0 for French speakers to 0.28 for Persians. Group differences are also evident in terms of the percentage of slopes that are positive and significant. Using this criterion, it is evident that while 40 per cent or more of the slopes are positive and significant for Spanish and Chinese speakers, the percentage of slopes that met the criterion is much smaller for such groups as French, Italians, Germen, Greeks, Gujarati and Japanese. As a whole, less than one-quarter of the 1,475 slopes are positive and statistically significant. Such findings are clearly at odds with the English advantage hypothesis.[7]

Aside from variations across MSAs, there are also variations across language groups within individual MSAs. Figure 1a–1c provides such information for the three largest MSAs (Chicago, Los Angeles–Long Beach and New York) where all twenty language groups are found. Together, these graphs clearly show that even within the same metropolitan area there are significant differences in the rate of return to English proficiency for the twenty groups.

Our next task is to explain the variations in the slope of English language proficiency, as evident in Table 1 and Figure 1a–1c, using group and MSA-level variables as independent variables. Table 2 presents the findings of the multilevel analyses.

The results for model 1 indicate that the sample as a whole has an average (log) earnings of 10.03. Although 92 per cent ($=0.93/(0.93+0.055+0.023)$) of the total variance in earnings occurred within group, there are significant variations in the average earnings across groups and MSAs. The results of model 2 indicate that the between-group and between-MSA variations in average (log) earnings are partly a result of differences in human capital traits for different groups and for different MSAs. Adding human capital variables to our model reduces

Table 1. *The effect of immigrants' English language proficiency on earnings in metropolitan US by groups*

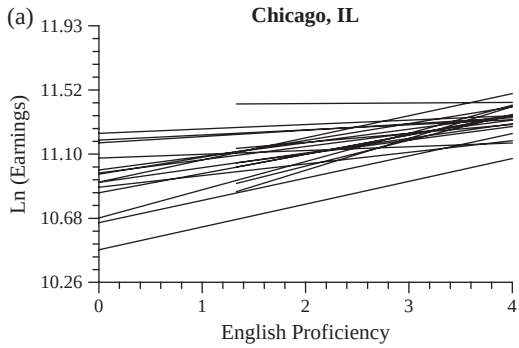
	Number of MSAs	Mean Effect	SD	Min.	Max.	% pos. & sig.	Largest MSA	Effect	Smallest MSA	Effect
Spanish	241	0.07	0.14	−0.46	0.73	43.57	Los Angeles − Long Beach	0.08*	Decatur, AL	0.32*
French	86	0.00	0.40	−2.05	0.68	15.12	New York	0.10*	Charleston, SC	−2.05
F. Creole	28	0.07	0.13	−0.19	0.34	21.43	New York	0.12*	Houston, TX	−0.18
Italian	58	0.08	0.32	−0.78	1.11	13.79	New York	0.05	Sarasota − Bradenton, FL	0.21
Portuguese	56	0.07	0.20	−0.47	0.64	21.43	Boston	0.01	Vallejo–Fairfied–Napa, CA	0.12
German	160	0.08	0.59	−1.76	2.29	11.25	Washington	−0.16	Gainesville, FL	0.96*
Greek	36	0.01	0.45	−1.37	1.34	8.33	New York	0.06	Charlotte–Gastonia–Rock Hill, NC–SC	−0.61
Russian	60	0.11	0.23	−0.48	0.66	28.33	New York	0.19*	Syracuse, NY	−0.48
Polish	46	0.08	0.20	−0.59	0.50	21.74	Chicago	0.06*	Sarasota − Bradenton, FL	0.06
Armenian	12	0.16	0.36	−0.60	0.85	25.00	Los Angeles–Long Beach	0.21*	Philadelphia, PA–NJ	0.25
Persian	37	0.28	0.41	−1.15	1.14	37.84	Los Angeles Long Beach	0.19*	Nashville,TN	0.22
Gujarathi	37	0.11	0.22	−0.48	0.69	13.51	Chicago	0.02	Norfolk–Virginia Beach–Newport News, NC–VA	−0.29
Hindi	50	0.19	0.58	−1.28	1.99	20.00	New York	0.18*	Kansas City MO–KS	0.99
Urdu	34	0.05	0.36	−1.01	1.00	17.65	New York	−0.01	Denver, CO	0.08
Chinese	119	0.17	0.23	−0.34	0.84	39.50	New York	0.15*	Bakersfield, CA	0.71
Japanese	68	0.05	0.32	−0.79	0.90	10.29	Los Angeles–Long Beach	0.13*	Bremerton, WA	0.02
Korean	75	0.09	0.23	−1.21	0.69	17.33	Los Angeles–Long Beach	0.11*	Clarksville–Hopkinsville, TN–KY	0.02

Table 1 (*Continued*)

	Number of MSAs	Mean Effect	SD	Min.	Max.	% pos. & sig.	Largest MSA	Effect	Smallest MSA	Effect
Vietnamese	103	0.06	0.18	−0.49	0.54	21.36	Orange County	0.11*	Lawrence, MA–NH	0.02
Tagalog	96	0.08	0.33	−1.26	1.00	20.83	Los Angeles–Long Beach	0.13*	Albany–Schenectady–Troy, NY	0.21
Arabic	73	0.11	0.27	−0.57	0.84	19.18	New York	0.17*	Trenton, NJ	0.27
Total	1,475	0.09	0.33	−2.05	2.29	23.93				

*$p < 0.05$, one-tailed t test

Figure 1a. *Earnings on English proficiency* for twenty language groups resided in Chicago, IL, MSA***
*Controlled for human capital variables specified in equation (1). All independent variables, including English language proficiency, have been grand-mean centred
**The slopes range from −0.02 for Armenian speakers and 0.43 for Persian speakers



the between-group (level 2) variance from 0.055 to 0.013, or a 76 per cent reduction; and the between-MSA (level 3) variance from 0.023 to 0.013, or a 43 per cent reduction. Despite the reductions, between-group and between-MSA variances remain statistically significant. In addition, the results indicate significant variations in the slope of English language proficiency across language groups (0.007) and across MSAs (0.002).

The rate of return to English proficiency (0.09) is positive and statistically significant ($p \leq 0.001$). It means that earnings are expected

Figure 1b. *Earnings on English proficiency* for twenty language groups resided in Los Angeles–Long Beach, CA, MSA***
*Controlled for human capital variables specified in equation (1). All independent variables, including English language proficiency, have been grand-mean centred
**The slopes range from 0.03 for German speakers and 0.34 for Polish speakers

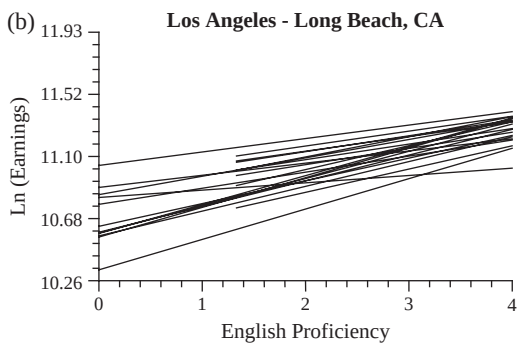
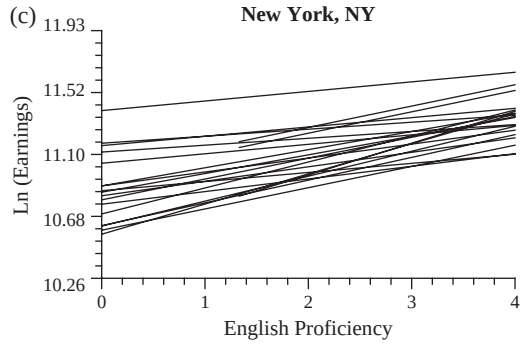


Figure 1c. *Earnings on English proficiency* for twenty language groups resided in New York MSA***
*Controlled for human capital variables specified in equation (1). All independent variables, including English language proficiency, have been grand-mean centred
**The slopes range from 0.01 for Urdu speakers and 0.42 for Persian speakers



to improve by 9.4 per cent ($=100*(e^{.09}-1)$) for one unit improvement in English language fluency, assuming that the ordinal measure can be approximated by an interval scale.[8] The earnings of the immigrant are also affected positively by educational attainment, having a white-collar occupation, being married and being a man; and negatively by self-employment status. Labour market experience and the number of years an immigrant has spent in the US are both found to improve earnings but with declining rates. All coefficients are statistically significant at the 0.001 level.

Model 3 tests, among other things, the hypotheses that the effects of English language fluency on earnings are contingent upon the size and segregation of a non-English language group. An examination of the two cross-level interactions makes clear that both of them are negative and statistically significant. These findings suggest that an increase in size and segregation of a non-English language group weakens the positive association between English language fluency and earnings. For example, one standard deviation increase in log group size (1.95) will reduce the rate of return to English proficiency from 0.10 to 0.08 ($=0.10+(-0.01)*1.95$) through the interaction effect. Similarly, one standard deviation increase in segregation (0.03) will reduce the rate by the same magnitude given a negative coefficient (i.e. -0.61) associated with the interaction term. While the main effects of group-level variables on earnings are not the primary focus of the study, we note in passing that the two group-level variables both exert significant negative effects on earnings. A comparison of the *deviance* statistics for models 2 and 3 clearly indicates a significant improvement in the fit of the model when the two group-specific variables are added.

Table 2. *Multilevel analysis of immigrants' earnings*

	Model 1 Coefficients	Model 2 Coefficients	Model 3 Coefficients	Model 4 Coefficients
Fixed effects				
Grand average log earnings	10.03***	9.86***	9.88***	9.86***
Individual-level predictors from human-capital model				
English proficiency		0.09***	0.10***	0.10***
Male		0.52***	0.52***	0.52***
Education		0.04***	0.04***	0.04***
White-collar occupations		0.30***	0.30***	0.30***
Married		0.09***	0.09***	0.09***
Self-employment		−0.22***	−0.22***	−0.22***
Experience		0.02***	0.02***	0.02***
Experience^2/100		−0.03***	−0.03***	−0.03***
Years in US		0.03***	0.03***	0.03***
Years in US^2		−0.04***	−0.04***	−0.04***
Group-level predictors				
Ln (group size)			−0.02***	−0.02***
Segregation			−0.42*	−0.09
MSA-level predictors				
Heterogeneity				0.10***
Inequality				−0.09***
Ln (total population)				0.02***
Median earning/1000				0.03***
Cross level interactions				
*Group-level variables*English proficiency*				
Ln (Group Size)			−0.01***	−0.01***
Segregation			−0.61***	−0.68***

Table 2 (*Continued*)

	Model 1 Coefficients	Model 2 Coefficients	Model 3 Coefficients	Model 4 Coefficients
*MSA-level variables*English proficiency*				
Heterogeneity				0.23***
Inequality				0.05*
Random Effects				
Level 1 residual	0.93	0.75	0.75	0.75
Variance components for intercept				
Level 2	0.055***	0.013***	0.012***	0.011***
Level 3	0.023***	0.013***	0.015***	0.001**
Variance components for English proficiency				
Level 2		0.007***	0.007***	0.007***
Level 3		0.002***	0.002***	0.001***
Comparing models				
Model deviance	1596060.61	1474541.61	1474375.84	1473956.80
Chi-square		121519.00***	165.77***	419.04***
df		14	4	6

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

The main and interaction (with English proficiency) terms of linguistic heterogeneity and inequality are added in model 4. The addition of the interaction terms is necessary for the testing of our hypotheses that the rate of return to English proficiency varies with heterogeneity and inequality of the community. The 'main' effects of these variables are included as controls rather than for theoretical reasons. We also control for MSAs' population and median earnings because wage rates often correlate positively with these factors.

The results of the analysis support our hypotheses. Both of the cross-level interaction terms are positive and statistically significant, confirming our predictions that the rate of return to English proficiency is greater in MSAs with higher levels of heterogeneity and inequality. Thus, an increase of linguistic heterogeneity by one standard deviation (0.14) raises the slope of English proficiency by 0.032 ($=0.23*0.14$); and an increase of inequality by one standard deviation (0.34), in contrast, raises the slope by 0.017 ($=0.05*0.34$).

Besides their effects on the slope of English proficiency, heterogeneity and inequality also exert significant effects on the intercept (average MSA earnings). The average earnings of immigrants are significantly higher in more heterogeneous communities but significantly lower in areas where higher levels of inequality prevail. As expected, immigrants who reside in metropolitan areas with larger population and higher median income also enjoy significantly higher earnings than their counterparts who live elsewhere. Finally, it is noteworthy that the addition of the four MSA-level variables has significantly reduced the main effect of segregation on earnings. This finding seems to suggest that the low average earnings of certain immigrant groups is a result not of segregation *per se*, but because they happen to reside in MSAs with a combination of traits (i.e. higher inequality and homogeneity, and lower median income and small population size) which suppress wage rates.

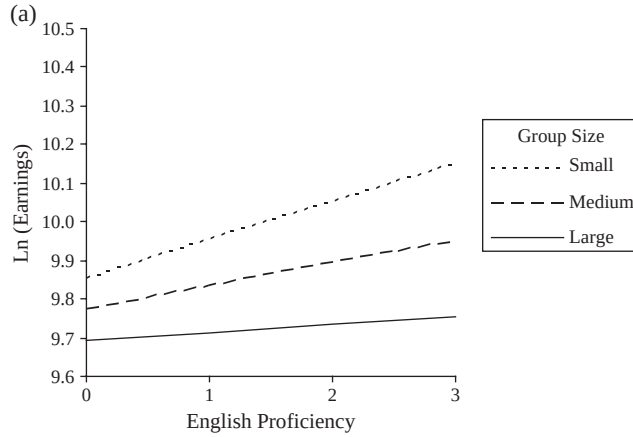
Finally, adding these variables to our model reduces the model *deviance* by 419, a significant improvement in the fit of the model using the LR-χ^2 test with six degrees of freedom. The improvement is also evident in a 93 per cent reduction in the between-MSA (level 3) variance in average earnings or intercept and a 50 per cent reduction in the between-MSA variance of the rate of return to English proficiency.

Conditional returns to English language proficiency: a graphic summary

The above detected conditional nature of English proficiency effects can be made more explicit when the regression coefficients presented in model 4 of Table 2 are illustrated in graphics.

Figure 2a–2c show the effects of group size on the rate of return to English proficiency (and on average earnings) using the regression

Figure 2a. *Earnings on English proficiency by group size* (Spanish)*
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred



coefficients of model 4 while holding other variables constant. Since space does not allow us to show the graphs for all twenty groups, we have chosen to show only graphs for selected groups of greater interest. For purposes of contrast, we identify MSAs with minimum, medium, and maximum group size for each group and juxtapose their expected intercepts and slopes while holding other variables constant. These

Figure 2b. *Earnings on English proficiency by group size* (Chinese)*
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred

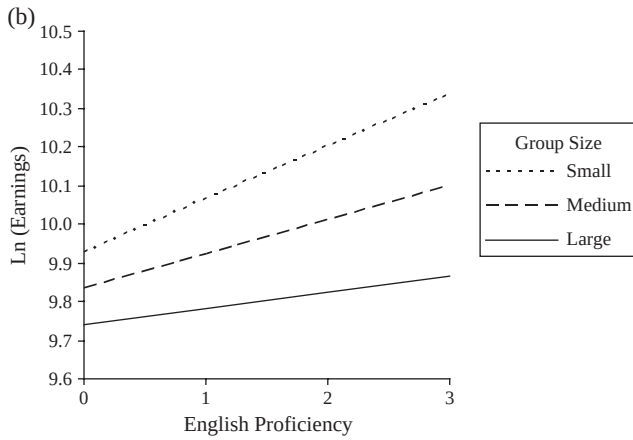
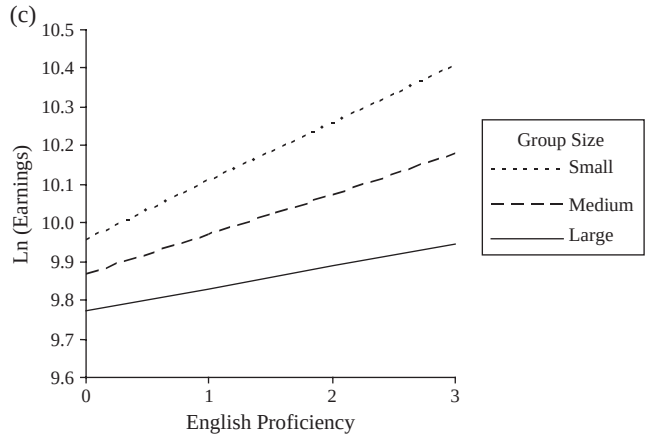


Figure 2c. *Earnings on English proficiency by group size* (Arabic)*
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred

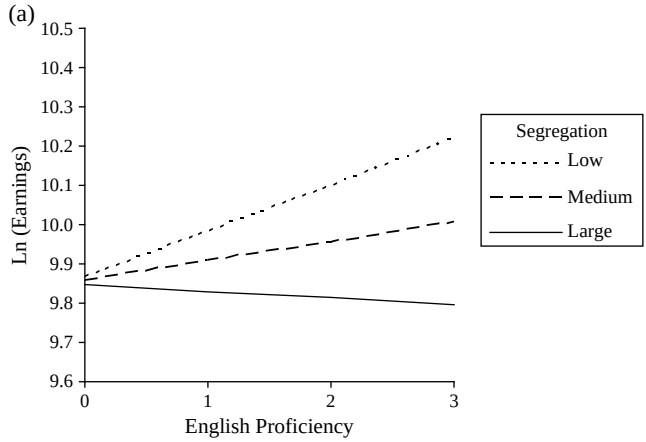


graphs clearly show that the slope is steepest in the MSA where the size of the group is smallest and flattest in the MSA where the group is largest for all three groups. These graphs also show a negative correlation between group size and intercept (or the average earnings) for each of these groups. Together, immigrants who resided in MSAs with a large number of compatriots are jeopardized not only by a smaller return to their English language skills but also by a lower average wage.

A similar analysis examines the conditioning role of segregation in predicting the rate of return to English proficiency (see Figure 3a– 3c). The negative effects of segregation on the slope are again evident. Thus, immigrant workers who live in least-segregated MSAs enjoy greater return to their English proficiency than their counterparts in more segregated MSAs. It is noteworthy that the slopes associated with MSAs with highest degree of segregation are either close to zero or slightly negative. This indicates that immigrants who speak fluent English enjoy no earnings advantage if they live in a very segregated community. Despite the presumed social benefits of living among compatriots, immigrants who speak English fluently but live in segregated ethnic communities pay a large economic cost. Segregation, however, has no effect on intercept (or starting salary).

The conditioning roles of linguistic heterogeneity and inequality are examined in Figure 4 and 5 respectively. Because heterogeneity and inequality are characteristics of MSA shared by all groups, the same graph applies to all language groups. For the purpose of juxtaposition, we again identify MSAs with values equal to minimum, medium and

Figure 3a. *Earnings on English proficiency by segregation* (Spanish)*
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred



maximum for each of these two variables and plot the expected regression lines while holding other factors constant. Figure 4 clearly shows that linguistic heterogeneity exerts a positive effect on the slope. The rate of return to English proficiency is steepest in the most heterogeneous MSA and least so in the least heterogeneous community. Heterogeneity also exerts a positive effect on the average earnings. The rate of return to English proficiency is also affected by inequality

Figure 3b. *Earnings on English proficiency by segregation* (Chinese)*
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred

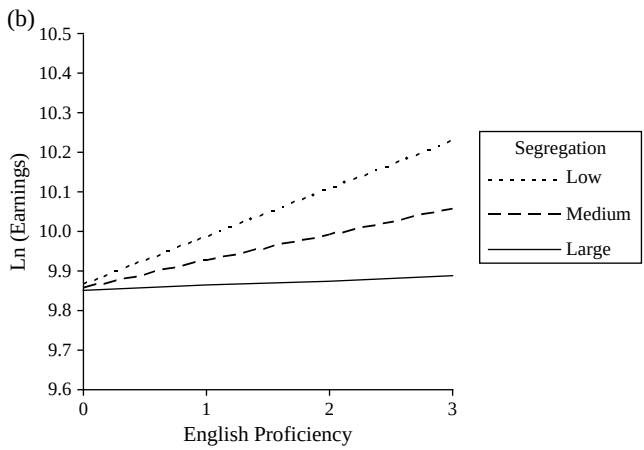
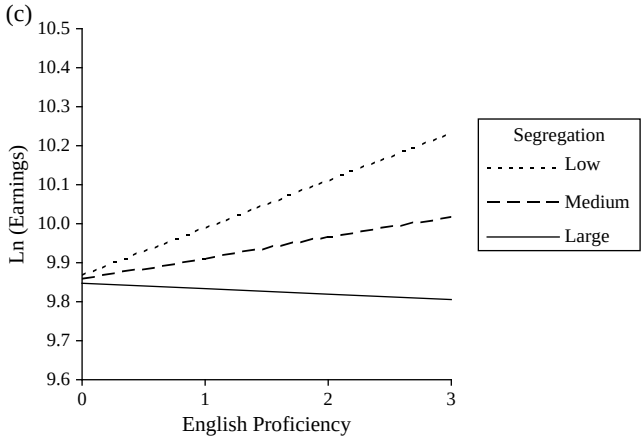


Figure 3c. *Earnings on English proficiency by segregation* (Arabic)*
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred



between immigrants and natives in the MSA in a positive and significant manner. In addition, inequality exerts a negative impact on intercept, or the average earnings immigrants received.

Another way to quantify the interaction effects is to show the difference in earnings between poor and fluent English speakers in a different language environment. For example, Spanish speakers who speak English very well (group A) made $5,157 more than their counterparts who do not speak English at all (group B) in an MSA

Figure 4. *Earnings on English proficiency by linguistic heterogeneity**
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred

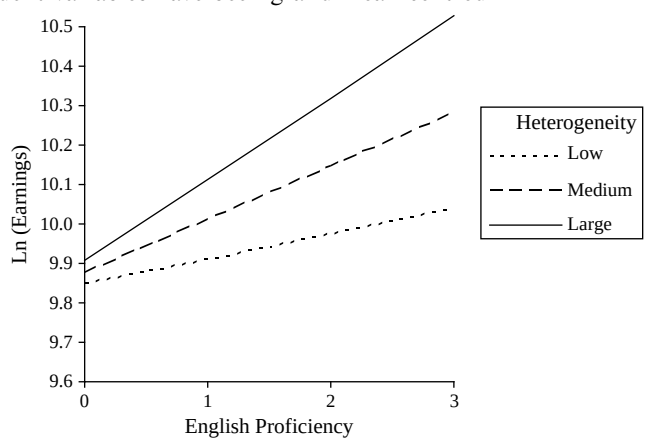
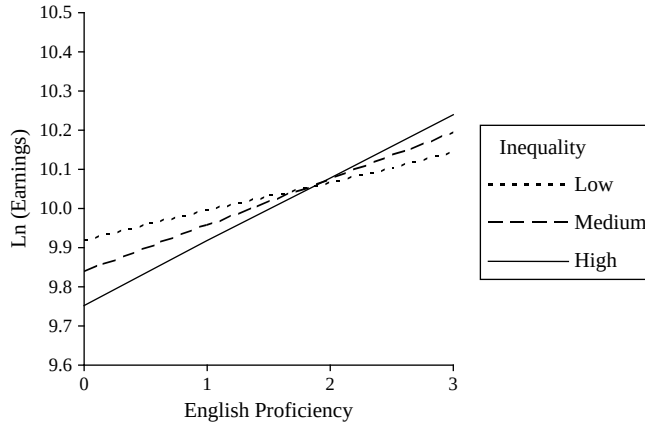


Figure 5. *Earnings on English proficiency by income inequality**
*Controlled for human capital variables specified in equation (1) and structural variables specified in equations (2.1), (2.2), (3.1) and (3.2). All independent variables have been grand-mean centred



where the size of Spanish speakers is smallest; the group A's advantage, on the other hand, is only $707 in an MSA where the size of the group is largest. In terms of segregation, group A has a $6,570 advantage in the least-segregated MSA but has a $4,578 *disadvantage* in the most-segregated MSA. On the contrary, group A's advantage is bigger in the most-heterogeneous MSA ($10,934) than in least-heterogeneous one ($3,323). Finally, group A enjoys a larger advantage in an MSA with highest income inequality ($7,664) than in an MSA with lowest inequality ($4,037).

Conclusions and discussions

Although the importance of English for immigrants in an English-speaking country has been well understood, little is known about whether or not such importance is diminished by the settlement patterns of immigrants and the community context in which they are embedded. Blau's structural theory suggests that intergroup relationships are determined by the characteristics of minority groups such as their size and segregation, and by community characteristics such as heterogeneity and inequality. In communities where the immigrant population is large and highly segregated from English speakers, it is conceivable that the need for intergroup interactions as well as English as a medium of interaction is likely to be reduced. However, small group size and heterogeneity of the community are expected to exert pressures for immigrants to learn English because interactions in such settings are predominantly inter-ethnic and therefore require a *lingua franca*. Thus, the importance of English as a mean for interactions and for making

a living is contingent upon factors which determine the need for intergroup interactions. In other words, Blau's structural theory makes clear that whether immigrants with a better English-speaking ability would actually have an advantage in the labour market, compared to their compatriots who speak English poorly, depends on the language environment in which these immigrants live.

Our findings demonstrate the utility of Blau's theory in understanding the conditional importance of English as a human capital in the US labour market. Because previous studies examining the relationship of English language proficiency and immigrants' earnings pooled immigrants of different groups resided in different MSAs together and observed the *average* association, group and MSA differences in such relationship were obscured. Blau's structural theory not only anticipates the conditional nature of the relationship but also suggests a set of testable hypotheses. Not denying the importance of English-language skills as a human capital in English-speaking societies, Blau's theory points out that English is important primarily because it is a tool for intergroup interactions; when there is little need for members of different language groups to interact with one another, the importance of English as a means for making a living inevitably reduced.

Since the findings reported here are based on cross-sectional analysis, the causal validity of our structural explanations can be challenged by invoking endogeneity as a competing hypothesis (Chiswick and Miller 1995). For example, because migration is known to be selective (Borjas 1990), it is possible to imagine a scenario in which immigrants with poor spoken English have a greater propensity to move to metropolitan areas where returns to English speaking ability is smaller; and those with better English are, instead, more inclined to move to areas where English speaking skills are better remunerated. If true, the structural characteristics which we used as independent variables are not exogenous to the sorting process but rather a result of it.

One way to fix the problem is to replace the contemporaneous measures of the structural variables with the time-lagged ones (Singer and Willett 2003); the rationale being that if reciprocal causality is due to the use of contemporaneous independent and dependent variables, lagging the independent variables would reduce such possibility. Analysis using lagged structural measures from the 1990 census yielded results similar to those reported here.

The study also has important policy implications. It shows that immigrants' economic adaptation in the US labour market is not necessarily hampered by their English language deficiency under certain conditions; the identification of such conditions may present alternative routes for immigrant incorporation. Due to the cross-sectional nature of our analyses, we are unable to directly address

issues related to the long-term relevance of English for immigrants already here and for future immigrants (e.g. Alba et al. 2002). However, our findings suggest that the answer to this question would depend not only on the language and assimilation policies of the land, but also on immigrant policies which determine the origin, volume and settlement patterns of future waves of immigrants to the country.

Notes

1. See Table 1 for the list of language groups included in the analysis.
2. Although there are 331 MSA/PMSAs (Primary Metropolitan Statistical Areas), only 297 can be uniquely identified. In addition, not all twenty groups are present in all MSAs. Thus, the number of language groups included in the analysis varies from one MSA to another. Although aggregate studies typically excluded smaller units from analyses to minimize their undue influences, no such restriction is needed here because the Hierarchical Linear Models (HLM) programme we use addresses the problem by weighting units according to relative size (Raudenbush and Bryk 2002).
3. Following White (1983), the distance between two persons living in the same census tract i is approximated by $0.6\sqrt{A}$, where A stands for the land area of census tract i.
4. The interpretation of equation (2.2) as a mean for testing cross-level interactions becomes obvious when the π_{1jk} in equation (1) is substituted by the identity in equation (2.2).
5. Although the exact effects of group size and segregation on immigrants' earnings remain equivocal, recent studies indicated that they tend to be negative (Borjas 1994; Bean and Stevens 2003). Such findings are consistent with Blalock's (1967) arguments that discrimination of minority group increases with the size of the minority group and the arguments that the enclave economy survives on low wages.
6. For example, large population size and greater heterogeneity are expected to benefit minority groups because they make discrimination of minority groups more difficult. The average earnings of minority groups is also expected to correlate positively with the MSA's median income and negatively with its inequality.
7. Although one can argue that the small sample size for most groups in a majority of MSAs may be the primary reason for such findings, our results indicate that many of the slopes were not significant even in MSAs where these groups have their largest presence.
8. The multilevel model becomes more parsimonious when English proficiency is treated as an interval-level measure. An ordinal measure of English proficiency would have generated three coefficients for the variable instead of just one and significantly complicated the modelling and the interpretations of higher-level model. See Zeng and Xie (2004) for a similar justification.

References

ALBA, RICHARD, *et al.* 2002 'Only English by the third generation? Loss and preservation of mother tongue among the grandchildren of contemporary immigrants', *Demography*, vol. 39, pp. 467–84

ALDRICH, HOWARD, *et al.* 1985 'Ethnic residential concentration and the protected market hypothesis', *Social Forces*, vol. 63, pp. 996–1009

BEAN, FRANK D. and STEVENS, GILLIAN 2003 *America's Newcomers and the Dynamics of Diversity*, New York: Russell Sage Foundation

BECKER, GARY S. 1993 *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, 3rd edn, Chicago, IL: University of Chicago Press

BLALOCK, HUBERT M. 1967 *Toward a General Theory of Minority Group Relations*, New York: Wiley

BLAU, PETER M. 1977 *Inequality and Heterogeneity*, New York: Free Press
BLAU, PETER M., BLUM, TERRY C. and SCHWARTZ, JOSEPH E. 1982 'Heterogeneity and intermarriage', *American Sociological Review*, vol. 47, pp. 45–62
BORJAS, GEORGE J. 1990 'Assimilation and the earnings of immigrants', in George J. Borjas, *Friends or Strangers: The Impact of Immigrants on the US Economy*, New York: Basic Books, pp. 97–115
——— 1994 'The economics of immigration', *Journal of Economic Literature*, vol. 32, pp. 1667–717
BRETON, RAYMOND 1964 'Institutional completeness of ethnic communities and personal relations to immigrants', *American Journal of Sociology*, vol. 70, pp. 193–205
CARLINER, GEOFFREY 2000 'The language ability of US immigrants: assimilation and cohort effects', *International Migration Review*, vol. 34, pp. 158–82
CHISWICK, BARRY R. 1978 'The effect of Americanization on the earnings foreign-born men', *Journal of Political Economy*, vol. 86, pp. 897–921
CHISWICK, BARRY R. 1991 'Speaking, reading, and earnings among low-skilled immigrants', *Journal of Labour Economics*, vol. 9, no. 2, pp. 149–70
CHISWICK, BARRY R. and MILLER, PAUL W. 1995 'The endogeneity between language and earnings: international analyses', *Journal of Labour Economics*, vol. 13, pp. 246–88
——— 2002 'Immigrant earnings: language skills, linguistic concentrations and the business cycle', *Population Economics*, vol. 15, pp. 31–57
DUSTMANN, CHRISTIAN and VAN SOEST, ARTHUR 2002 'Language and the earnings of immigrants', *Industrial and Labour Relations Review*, vol. 55, pp. 473–92
EVANS, M. D. R. 1989 'Immigrant entrepreneurship: effects of ethnic market size and isolated labour pool', *American Sociological Review*, vol. 54, pp. 950–62
FRIEDBERG, RACHEL M. 2000 'You can't take it with you? Immigrant assimilation and the portability of human capital', *Journal of Labour Economics*, vol. 18, pp. 221–51
GRENIER, GILLES 1984 'The effects of language characteristics on the wages of Hispanic American males', *Journal of Human Resource*, vol. 18, pp. 35–52
HOFMANN, DAVID A. and GAVIN, MARK B. 1998 'Centering decision in hierarchical linear models: implications for research in organizations', *Journal of Management*, vol. 24, pp. 623–41
HWANG, SEAN-SHONG and MURDOCK, STEVE H. 1998 'Racial attraction and racial avoidance in American suburbs', *Social Forces*, vol. 77, pp. 541–65
HWANG, SEAN-SHONG, SAENZ, ROGELIO and AGUIRRE, BENIGNO E. 1997 'Structural and assimilationist explanations of Asian American intermarriage', *Journal of Marriage and the Family*, vol. 59, pp. 758–72
HWANG, SEAN-SHONG and XI, JUAN 2008 'Structural and individual covariates of English language proficiency', *Social Forces*, vol. 86, pp. 1079–104
JASSO, GUILLERMINA and ROSENZWEIG, MARK 1990 *The New Chosen People: Immigrants in the United States*, New York: Russell Sage Foundation
MASSEY, DOUGLAS S. and DENTON, NANCY 1988 'The dimensions of residential segregation', *Social Forces*, vol. 67, pp. 281–315
——— 1993. *American Apartheid*, Cambridge, MA: Harvard University Press
MCMANUS, WALTER 1990 'Labour market effects of language enclaves: Hispanic men in the United States', *The Journal of Human Resources*, vol. 25, pp. 228–52
MCMANUS, WALTER, GOULD, WILLIAM and WELCH, FINIS 1983 'Earnings of Hispanic men: the role of English language proficiency', *The Journal of Labour Economics*, vol. 1, no. 2, pp. 101–30
MINCER, JACOB 1958 'Investment in human capital and personal income distribution', *The Journal of Political Economy*, vol. 66, no. 4, pp. 281–302
PORTES, ALEJANDRO and BACH, ROBERT 1985 *Latin Journey*, Berkeley, CA, Los Angeles, CA: University of California Press
RAUDENBUSH, STEPHEN W. and BRYK, ANTHONY S. 2002 *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd edn, Thousand Aaks, CA: Sage

RYTINA, STEVEN, *et al.* 1988 'Inequality and intermarriage: a paradox of motive and constraint', *Social Forces*, vol. 66, pp. 645–75
SIEGEL, JACOB S. and SWANSON, DAVID 2004 *The Methods and Materials of Demography*, Amsterdam, Boston, MA: Elsevier, Academic Press
SINGER, JUDITH D. and WILLETT, JOHN B. 2003 *Applied Longitudinal Data Analysis*, New York: Oxford University Press
SPENER, DAVID and BEAN, FRANK D. 1999 'Self-employment concentration and earnings among Mexican immigrants in the US', *Social Forces*, vol. 77, pp. 1021–47
STEVENS, GILLIAN 1992 'The social and demographic context of language use in the United States', *American Sociological Review*, vol. 57, pp. 171–85
TAINER, EVELINA 1988 'English language proficiency and the determination of earnings among foreign born men', *The Journal of Human Resources*, vol. 23, no. 1, pp. 108–22
TIENDA, MARTA and WILSON, FRANKLIN D. 1992 'Migration and the earnings of Hispanic men', *American Sociological Review*, vol. 57, pp. 661–78
U S BUREAU OF THE CENSUS 2002 *Technical Documentation: Census 2000 Summary File 3*, Washington, DC: US Government Printing Office
WHITE, MICHAEL J. 1983 'The measurement of spatial segregation', *American Journal of Sociology*, vol. 88, pp. 1008–18
WILSON, KENNETH L. and PORTES, ALEJANDRO 1980 'Immigrant enclave: an analysis of the labour market experience of Cuban in Miami', *American Journal of Sociology*, vol. 86, pp. 295–319
ZENG, ZHEN and XIE, YU 2004 'Asian-Americans' earnings disadvantage reexamined: the role of place of education', *American Journal of Sociology*, vol. 109, pp. 1075–108

SEAN-SHONG HWANG is Professor in the Department of Sociology at the University of Alabama at Birmingham.
ADDRESS: 460 Heritage Hall Building, 1401 University Blvd., Birmingham, AL 35294-1152, USA. Email: shwang@uab.edu

JUAN XI is Assistant Professor in the Department of Sociology at the University of Akron.
ADDRESS: Department of Sociology, University of Akron, Akron, OH 44325-1905, USA. Email: jx@uakron.edu

YUE CAO is a PhD candidate in the Department of Sociology at the University of Alabama at Birmingham.
ADDRESS: 460 Heritage Hall Building, 1401 University Blvd., Birmingham, AL 35294-1152, USA. Email: yuecao@uab.edu