

IZA DP No. 2664

Occupational Language Requirements and the Value of English in the U.S. Labor Market

Barry R. Chiswick
Paul W. Miller

March 2007

Occupational Language Requirements and the Value of English in the U.S. Labor Market

Barry R. Chiswick

*University of Illinois at Chicago
and IZA*

Paul W. Miller

*University of Western Australia
and IZA*

Discussion Paper No. 2664
March 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Occupational Language Requirements and the Value of English in the U.S. Labor Market^{*}

This paper is concerned with the English language requirements (both level and importance) of occupations in the United States, as measured by the O*NET database. These scores are linked to microdata on employed adult (aged 25 to 64) males, both native born and foreign born, as reported in the 2000 Census, one percent sample. Working in an occupation that requires greater English language skills, whether measured by the level of these skills or the importance of English for performing the job, has a large effect on earnings among the native born, and an even larger effect among the foreign born. This effect is reduced by 50 percent, but is still large, when worker characteristics, including their own English language skills, are held constant. Earnings increase with the respondent's own proficiency in English, with the English proficiency required for the occupation, and when those with high levels of proficiency work in jobs requiring English language skills (interaction effect). There is, therefore, a strong economic incentive for the matching of worker's English skills and the occupation's requirements, and this matching does tend to occur in the labor market.

JEL Classification: J24, J31, J62, F22

Keywords: English language, earnings, occupation, immigrants, schooling

Corresponding author:

Barry R. Chiswick
Department of Economics
University of Illinois at Chicago
601 South Morgan Street
Chicago, IL 60607-7121
USA
E-mail: brchis@uic.edu

^{*} We thank Derby Voon for research assistance. Chiswick acknowledges research support from the Institute of Government and Public Affairs, University of Illinois. Miller acknowledges financial assistance from the Australian Research Council.

**“Occupational Language Requirements and the Value of English
in the U.S. Labor Market”**

I. INTRODUCTION

Jobs in the U.S. labor market differ in many ways. They obviously differ in terms of the levels of education and training they require. But they also differ in a host of other requirements, covering, among other aspects, job-specific skills (*e.g.*, time management, negotiation, instructing), abilities (*e.g.*, deductive reasoning, memorization, arm-hand steadiness), and knowledge (*e.g.*, of economics, the English language, a foreign language). The focus of this study is on the extent to which jobs in the U.S. labor market require, or offer a premium for, English language proficiency.

Information on the skill requirements of jobs can be obtained from official agencies in many countries. A well-known source is the Dictionary of Occupational Titles (DOT) developed by the U.S. Department of Labor (Rumberger 1981). The DOT has now been superseded by the Occupational Information Network, or O*NET, database.¹ This is a comprehensive database of worker attributes and job characteristics. Of primary interest to the current study is the information on work-related areas of knowledge. Knowledge areas are, according to the O*NET Knowledge Questionnaire, “sets of facts and principles needed to address problems and issues that are part of a job”.

One knowledge area is the English language. The O*NET database has information on occupational requirements concerning “Knowledge of the structure and content of the English language, including the meaning and spelling of words, rules of composition, and grammar”. Two sets of information were collected. The first is about

¹ The National O*NET Consortium was organized to develop O*NET and its related products for the US Department of Labor, Employment and Training Administration (ETA). The Consortium currently comprises the O*NET Management Partnership (the North Carolina Employment Security Commission, the Center for Employment Security Education and Research, and the ETA) and the National O*NET Support Group (composed primarily of the National Center for O*NET Development, the Research Triangle Institute, and the Human Resources Research Organization). Web address of O*NET data: <http://online.onetcenter.org>.

“How important is knowledge of the ENGLISH LANGUAGE to the performance of *your current job*?” (emphasis in O*NET Knowledge Questionnaire). The second is “What level of ENGLISH LANGUAGE is needed to perform *your current job*?” (emphasis in original).

Section II provides an overview of the O*NET data on the English language requirements of jobs in the U.S. labor market. Section III then links this information to data from the 2000 U.S. Census for an analysis of the relationship between earnings and English language requirements. This analysis is conducted within a framework similar to that used by Johnson and Solon (1986). Hence, it first explores the unstandardized relationship between individual earnings and the English language requirements of the respondent’s occupation. This is followed by the study of the standardized relationship between individual earnings and the language requirements of the occupation, where the control variables are based on the human capital earnings equation. The differences in the estimated impact of language requirements on earnings from these two approaches are then examined using the omitted variables formula. These analyses are conducted separately for native-born men and for foreign-born men. The reliability of the findings is examined using the two-step procedure for analysis of samples that combine aggregate (in this instance occupation-level) data with micro-level data proposed by Dickens and Katz (1987). Section IV provides concluding comments.

II. THE O*NET INFORMATION ON ENGLISH LANGUAGE REQUIREMENTS

As noted above, two sets of information on English language requirements are presented in the O*NET database. The first relates to the importance of the English language to performance in an individual’s job, and the second relates to the level of English language skills needed for the occupation.

The information on the importance of the English language was collected on a five-point scale: (1) Not important; (2) Somewhat important; (3) Important; (4) Very important; and (5) Extremely important. The information on the level of English language proficiency needed to perform in the current job was collected only among those who felt that English was somewhat or more important to the performance of their current job. A

seven-point scale was used, with three benchmark descriptors offered as a guide: 2 = write a thank you note; 4 = edit a feature article in a local newspaper; and 6 = teach a college English class. Individuals who did not feel that English was important to the performance of their current job were coded as zero on the scale for level of English. Hence, it is an eight-point scale (0-7).

To make the O*NET data more intuitively understandable to users, descriptor average ratings were standardized to a scale ranging from 0 to 100. This is accomplished using the formula:

$$S = ((O - L) / (H - L)) * 100$$

where S is the standardized score, O is the original rating score, L is the lowest possible score on the rating scale used, and H is the highest possible score on the rating scale used. The standardization procedure assumes an equal distance between points on the 5-point and 8-point scales. Hence, the original scores on the five-point importance of English scale become 1 = 0; 2 = 25; 3 = 50; 4 = 75; and 5 = 100. The scores on the eight-point level of English scale become 0 = 0; 1 = 14.3; 2 = 28.6; 3 = 42.9; 4 = 57.1; 5 = 71.4; 6 = 85.7; 7 = 100.

When the O*NET database was first developed, job analysts relied on information from the Dictionary of Occupational Titles, and modified this to suit the set of occupational codes (Standard Occupational Classification System²) used in O*NET. From June 2001, data have been collected from workers in targeted subsets of the occupations identified in O*NET, using a two-stage sampling design based on random samples of workers in targeted occupations within a random sample of businesses. These survey data have been progressively integrated into the initial O*NET database. About one-half of occupations have new survey information in Version 8 of the database used in this study. Hence, while these data provide information on job requirements, the method of collection cannot be categorized neatly into one of the “Job Analyst”, “Worker Self-assessment”, or “Realized Matches” approaches used in the undereducation/overeducation literature (see Hartog, 2000). The job requirements obtained from the O*NET database should therefore be viewed as having been compiled

² This is based around 23 major groups. In much of the analysis below, however, broader groupings are used for simplicity of presentation.

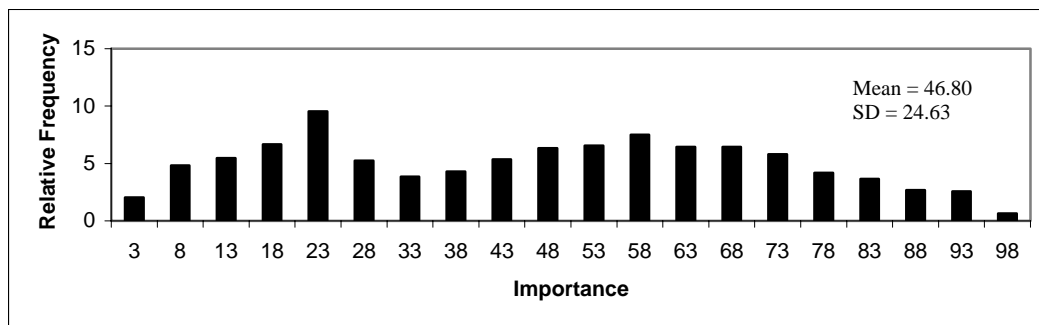
using a hybrid of the Job Analyst and Worker Self Assessment approaches used in the undereducation/overeducation literature.

Importance of English

There is considerable variation in the importance of knowledge of the English language to job performance. The mean standardized score is 46.8, which is close to the mid-point (“Important”) of the scale used in data collection. The standard deviation is 24.6, which is the equivalent of a change in one category in the underlying five-point scale. Moreover, the importance varies from minimal amounts (standardized scores around zero) in some occupations, to occupations where knowledge of English is very important. Occupations where English is not important include “Maids and Housekeeping Cleaners” (score of 0), “Bakers, Bread and Pastry” (4) and “Slaughterers and Meat Packers” (8). Examples of occupations where English is very important are “Public Relation Specialists” (90), “Government Service Workers” (92) and “Judges, Magistrate Judges and Magistrates” (95).

The relative frequency distribution in Figure 1 shows that the occupations in the U.S. labor market cover a full range of values on the standardized measure of the importance of the English language.³

Figure 1
Relative Frequency of Occupations on Standardized Score of Importance of English



³ For Figure 1 and related relative frequency distributions, the standardized scores have been collapsed into bands of width five (*e.g.*, 46-50, 51-55, 56-60). A mid-point is used on the horizontal axis. The 933 occupations for which details are available in the O*NET database are used in the compilation of the data for Figure 1. They were not weighted by the number of workers in the occupation, and hence the distribution is sensitive to how jobs are grouped into occupations.

Analysis by major occupational categories illustrates further the variation in the importance of knowledge of the English language to job performance (Figure 2).⁴ The “Management, Professional and Related” occupations have the highest mean score, of 67. Moreover, as the graph of the relative frequencies for this occupational group in Figure 2 shows, while the occupations are typically in the upper half of the range of possible scores, there is still considerable variation in the importance of knowledge of the English language across this more homogeneous grouping. The standard deviation of the scores for the “Managerial, Professional and Related Occupations” is 17 (compared to 25 for all occupations in Figure 1). The distribution is skewed to the left.

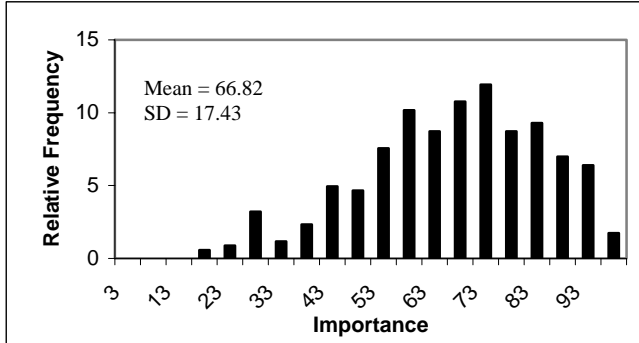
“Sales and Office” occupations also have a relatively high mean score, of 56 for the importance of English. This group of occupations, however, is mainly bunched around the middle of the range of possible scores. On the other hand, “Farming, Fishing and Forestry”, together with “Construction, Extraction and Maintenance” and “Production, Transportation and Material Moving” occupations, all have relatively low mean scores on the importance of English to performing the job (of 30, 28 and 26, respectively). Moreover, the scores for these particular occupations are bunched towards the bottom of the range of possible standardized scores, with only a few of their component occupations having a higher score for the importance of English (skewed to the right).

⁴ The number of occupations in the six broad groups employed in this analysis is: “Management, Professional and Related” (344 groups); “Service” (94); “Sales and Office” (93); “Farming, Fishing and Forestry” (19); “Construction, Extraction and Maintenance” (133); and “Production, Transportation and Material Moving” (250).

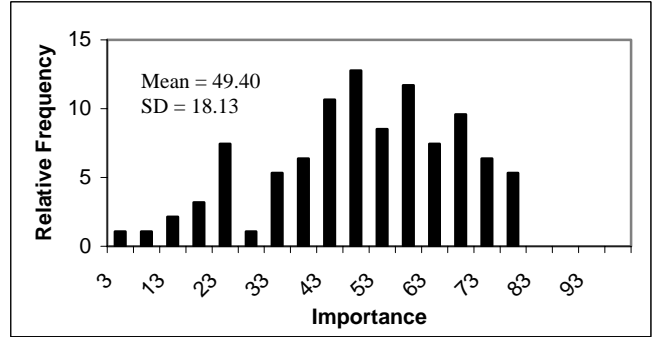
Figure 2

Relative Frequency of Occupations on Standardized Score of Importance of English

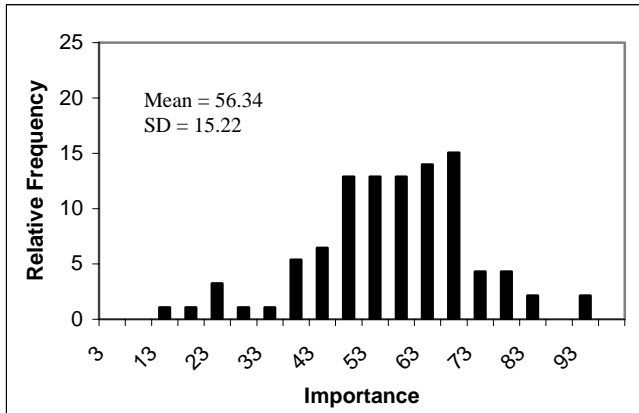
Management, Professional and Related Occupations



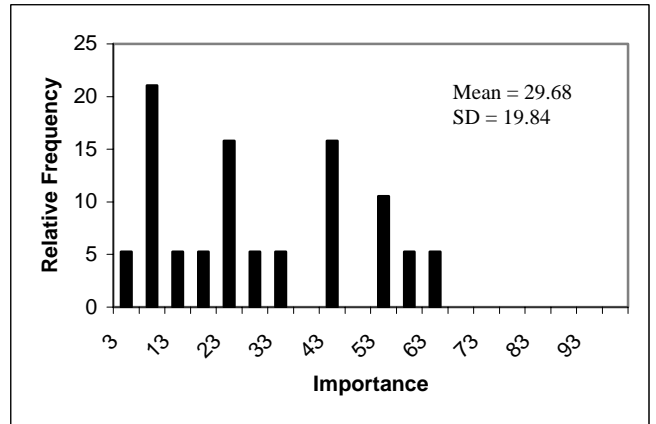
Service Occupations



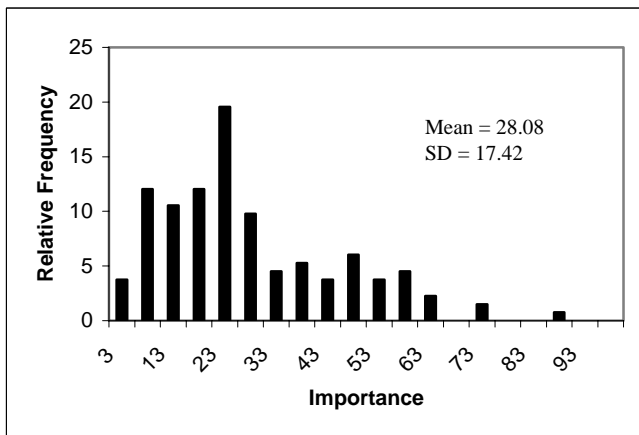
Sales and Office Occupations



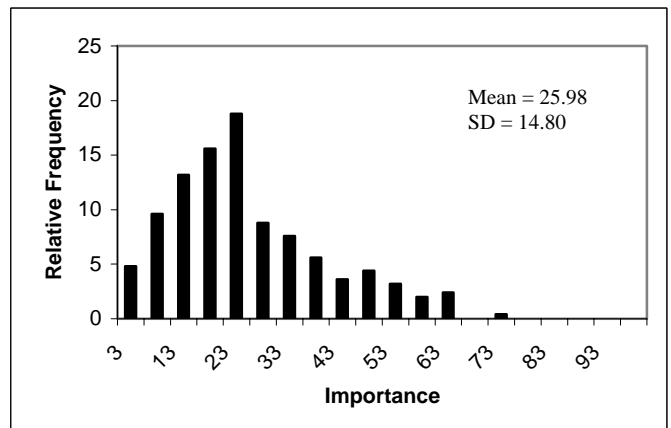
Farming, Fishing and Forestry Occupations



Construction, Extraction and Maintenance Occupations



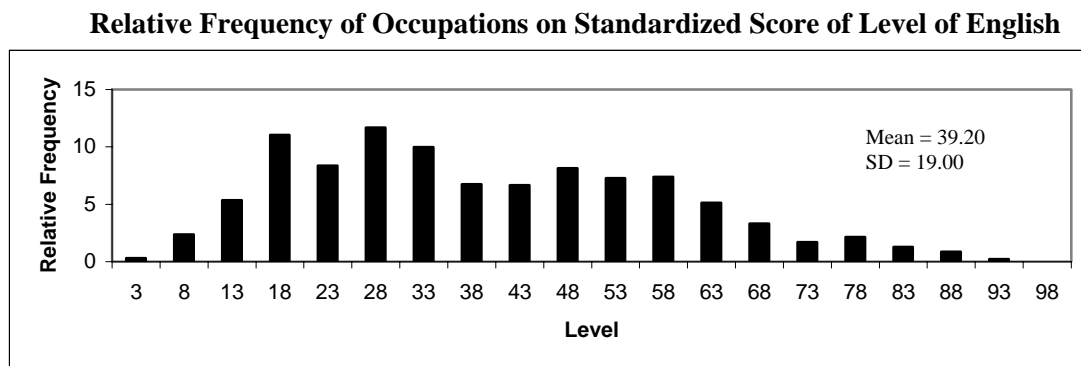
Production, Transportation and Material Moving Occupations



Level of English

The data on the level of English for all occupations are illustrated in Figure 3. The mean for all occupations is 39.2, which is midway between levels 29 and 57 which had the benchmark descriptors of “write a thank you note” and “edit a feature article in a local newspaper”. The standard deviation of the standardized score is 19. The occupations in the U.S. labor market cover a wide range of the standardized scores, although there is limited representation above scores of 85. Compared to the frequency distribution for the importance dimension, the data for the level of English needed to perform the job tend to be bunched more in the bottom one-third of the standardized scores. Nevertheless, there is a very high correlation (0.93) between the scores for the importance of English and the level of English. That is, occupations where knowledge of the English language is held to be important to job performance are occupations where a relatively high level of English language proficiency is needed to perform in the job.

Figure 3

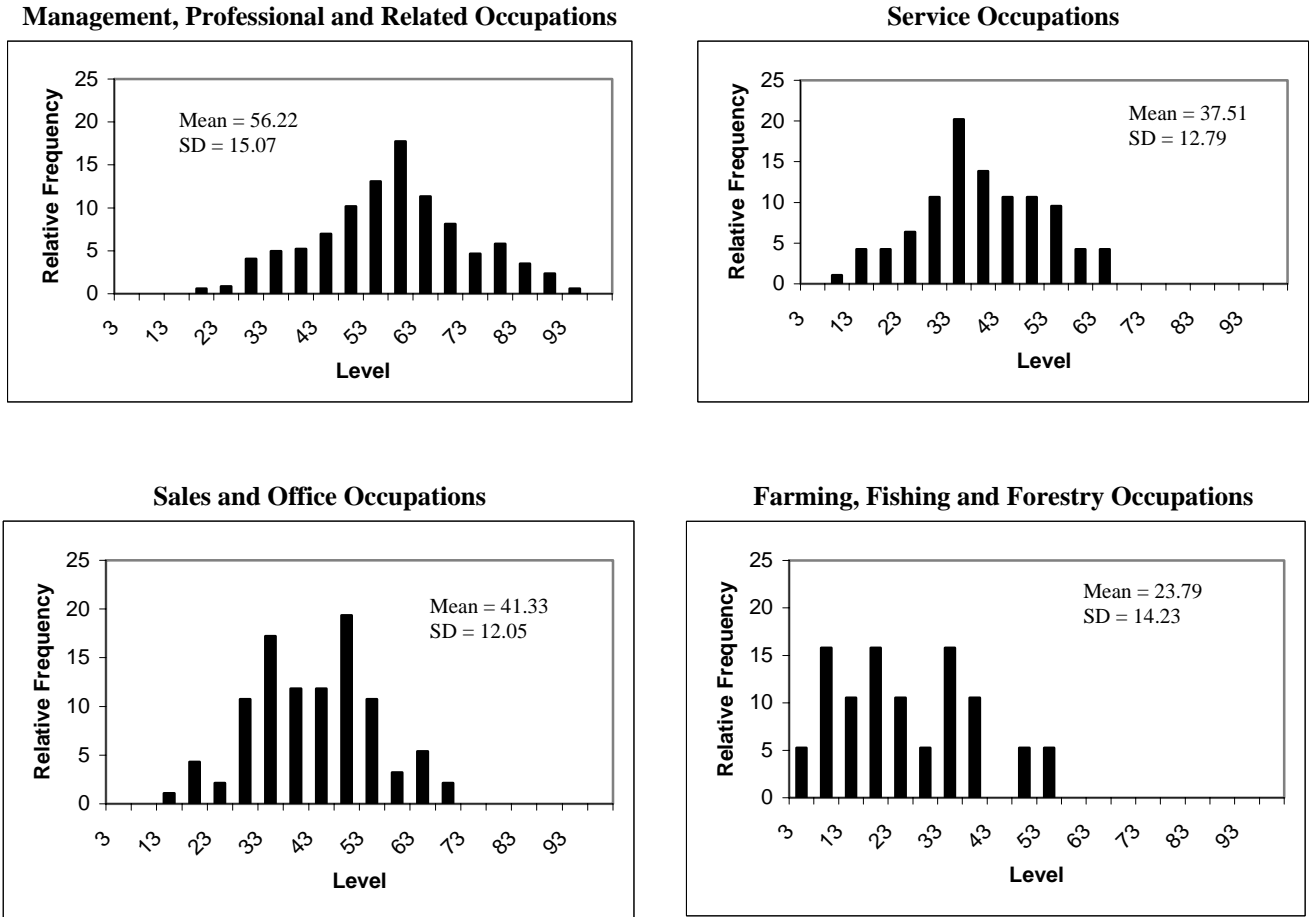


Examination of the standardized scores for the level of English needed to perform the job in the six major groups of occupations reveals a similarity with the data relating to the importance of English. Hence the ranking of the groups of occupations is the same, other than for a lower position of “Farming, Fishing and Forestry” on the level dimension. This occupational category is ranked four in terms of the importance of English, but it is ranked six in terms of level of English. Thus, it is important to have at least a basic understanding of English in these occupations. Within each broad major occupational category there is a high correlation between the level and importance scores,

ranging from a low of 0.80 for “Production, Transportation and Material Moving” to a high of 0.97 for the small number (19) of “Farming, Fishing and Forestry” occupations.⁵

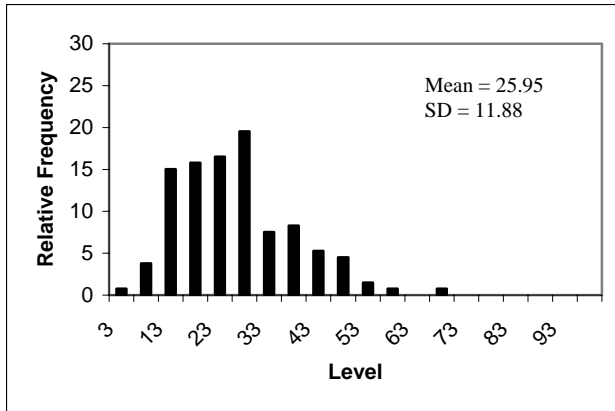
Figure 4

Relative Frequency of Occupations on Standardized Score of Level of English

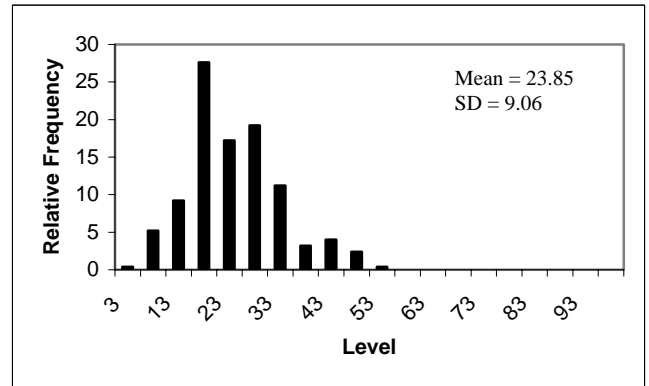


⁵ Tests show that this high correlation for “Farming, Fishing and Forestry” is not due to the inclusion of farm managers.

Construction, Extraction and Maintenance Occupations



Production, Transportation and Material Moving Occupations



III. ENGLISH LANGUAGE REQUIREMENTS AND EARNINGS

There is a growing empirical literature which shows that individuals, particularly immigrants, who are proficient in English, earn more in the U.S. labor market than those who have limited English skills (see, for example, Chiswick and Miller, 2002). The premium for English skill, between those who speak English well or better versus those with poorer spoken English language skills, is typically in the range of 10 to 15 percent, or the equivalent of one to two additional years of schooling. This literature has noted that there is a potential endogeneity issue associated with the inclusion of a measure of the individual’s English proficiency in an earnings equation. Attempts at accommodating this issue using Instrumental Variables (IV) have produced much higher estimates of the premium for English proficiency (see Chiswick and Miller, 1995). For example, Table 8 of Chiswick and Miller (1995) shows an IV estimate of the English language earnings premium among immigrants in the U.S. in 1979 of 57 percent, compared to an OLS estimate of 17 percent, with the IV estimate being described in the Chiswick and Miller (1995, p. 277) study as “surprisingly large”. The analyses reported below suggest that English language skills may be more valuable than this comment implies.

The mapping between the Census occupation codes and the Standard Occupational Classification System (SOC) codes in the O*NET database requires a number of approximations. Where English requirements were provided in the O*NET

database for sub-groups of a Census occupational category, the simple average of the scores for these sub-groups was used. For example, the English score for the Census occupational category Chief Executives was computed using the average of the O*NET categories of Government Service Executives and Private Sector Executives. In a small number of cases, data on the English requirements were not available for Census occupations, and the score for a similar occupation was used (*e.g.*, the score for Government Service Executives was used for Legislators). Finally, averages of occupations in the relevant categories were used for the Census “all other” categories.

Following the assignment of the occupational English scores outlined above, the mean level of English for native-born male workers is 40.8 and that for foreign-born male workers is 37.1. The mean importance of English for native born male workers is 50.1, and the mean importance for foreign-born male workers is 45.1. Hence, native-born male workers have 3.7 points (or nine percent) higher scores in the level of English needed to perform in their occupation. They have 5 points (or ten percent) higher scores for the importance of English in their jobs. These differences could help explain the lower mean earnings of the foreign born in the U.S. labor market (10.38 log points for the native born, compared to 10.12 log points for the foreign born, a difference of 0.26 log points, or approximately 26 percent). However, prior to investigating this matter, the relationships between these measures of the English requirements of the occupation and earnings need to be established.

A. Simple Regression Analysis

How do the English language requirements in occupations translate into earnings? To address this question, a simple regression that relates the earnings of individual i to the English requirements was estimated, namely:

$$(1) \quad \ln Y_i = \alpha_0 + \alpha_s ER_i + \varepsilon_i$$

where income (Y) is a measure of annual earnings from wage and salaried employment or self-employment, ER_i is a variable for the English requirements of the occupation in which the individual was employed (either the level or importance, as described in Section II), and ε_i is a stochastic disturbance term. α_s provides the assessment of the gross (or simple regression) link between earnings and the O*NET measure of English

requirements. This simple regression equation is estimated using data from the 2000 U.S. Census one percent Public Use Microdata Sample. The focus is on men aged 25 to 64 years who worked in paid employment in 1999. Separate equations are estimated for native-born workers and for foreign-born workers.

The estimates of the simple regression outlined in equation (1) are presented in column (i) of Table 1 for the level of English, and in column (i) of Table 2 for the importance of English. They are illustrated in Figure 5 for native-born men, and Figure 6 for foreign-born men. The first graph for each birthplace group is for the level of English, and the second is for the importance of English. The estimates from the simple regression are portrayed in the upper line in each figure.

The estimated coefficient on the O*NET measure of the level of English required in the respondent's occupation is 0.017 for the native born, and 0.020 for the foreign born, both of which are highly significant. For the O*NET measure of the importance of English in the respondent's occupation, the estimated coefficients are 0.013 for the native born, and 0.015 for the foreign born, also highly significant.

These estimates indicate that there is a difference in predicted mean earnings between those at the extremes of the level of English scale (0 to 100.0) of 1.7 log points for the native born, and a difference of 2.0 log points for the foreign born. The differences between the predicted mean earnings of those at the extremes of the importance of English scale are slightly smaller, 1.3 log points and 1.5 log points for the native born and the foreign born, respectively. In comparison, the difference between the predicted mean earnings of those with zero education and those with 20 years of education (the extremes of the education variable), computed from a simple regression of log earnings on years of education, is 2.5 log points for the native born, and 1.5 log points for the foreign born. Thus, English language skills appear to have an effect on the earnings of the foreign born comparable to that of schooling. However, English language skills have a much weaker effect than schooling on the earnings of the native born, most likely reflecting the greater prevalence of English.

Table 1
Estimated Earnings Functions with Focus on Level of English Skill,
by Nativity, 2000 US Census

Variable	Native Born			Foreign Born		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Constant	9.677 (2680.38)	4.449 (197.99)	4.444 (197.59)	9.378 (1233.42)	5.670 (105.52)	5.483 (98.05)
Log Weeks Worked	(a)	1.000 (182.35)	1.000 (182.38)	(a)	0.869 (72.79)	0.867 (72.73)
Years of Education	(a)	0.078 (129.60)	0.078 (126.67)	(a)	0.037 (44.99)	0.037 (44.22)
Experience	(a)	0.035 (76.07)	0.035 (76.05)	(a)	0.015 (14.41)	0.016 (15.17)
Experience Squared/100	(a)	-0.061 (60.67)	-0.061 (60.65)	(a)	-0.024 (12.16)	-0.025 (12.77)
Married	(a)	0.260 (108.14)	0.260 (108.11)	(a)	0.202 (34.34)	0.201 (34.22)
South	(a)	-0.061 (26.94)	-0.061 (26.95)	(a)	-0.072 (11.99)	-0.074 (12.43)
Metropolitan Area	(a)	0.195 (33.08)	0.195 (33.04)	(a)	0.113 (4.24)	0.121 (4.56)
Veteran	(a)	-0.038 (14.17)	-0.038 (14.12)	(a)	-0.071 (5.46)	-0.064 (4.90)
Blacks	(a)	-0.143 (40.25)	-0.143 (40.11)	(a)	-0.178 (16.72)	-0.171 (15.94)
Years Since Migration (YSM)	(a)	(a)	(a)	(a)	0.012 (14.08)	0.012 (14.13)
YSM Squared/100	(a)	(a)	(a)	(a)	-0.013 (6.38)	-0.013 (6.61)
English Very Well	(a)	-0.065 (12.94)	-0.004 (0.30)	(a)	-0.078 (8.49)	-0.038 (1.58)
English Well	(a)	-0.103 (8.08)	-0.081 (2.45)	(a)	-0.218 (22.09)	0.062 (2.58)
English Not Well/Not at All	(a)	-0.004 (0.19)	0.096 (1.97)	(a)	-0.304 (28.61)	0.074 (3.13)
Level of English ^(b)	0.017 (212.71)	0.008 (93.32)	(a)	0.020 (101.22)	0.010 (47.17)	(a)
Level*English Only	(a)	(a)	0.008 (92.54)	(a)	(a)	0.014 (30.93)
Level*English Very Well	(a)	(a)	0.006 (21.04)	(a)	(a)	0.013 (42.85)
Level*English Well	(a)	(a)	0.007 (9.48)	(a)	(a)	0.007 (19.41)
Level*English Not Well/Not at All	(a)	(a)	0.005 (4.84)	(a)	(a)	0.003 (5.98)
Adjusted R ²	0.082	0.347	0.347	0.115	0.381	0.386
Sample Size	532,109	532,109	532,109	84,172	84,172	84,172

Note: (a) Variables not entered; (b) Level of English required in the occupation of employment according to the O*NET database.

Source: 2000 United States Census, 1% PUMS.

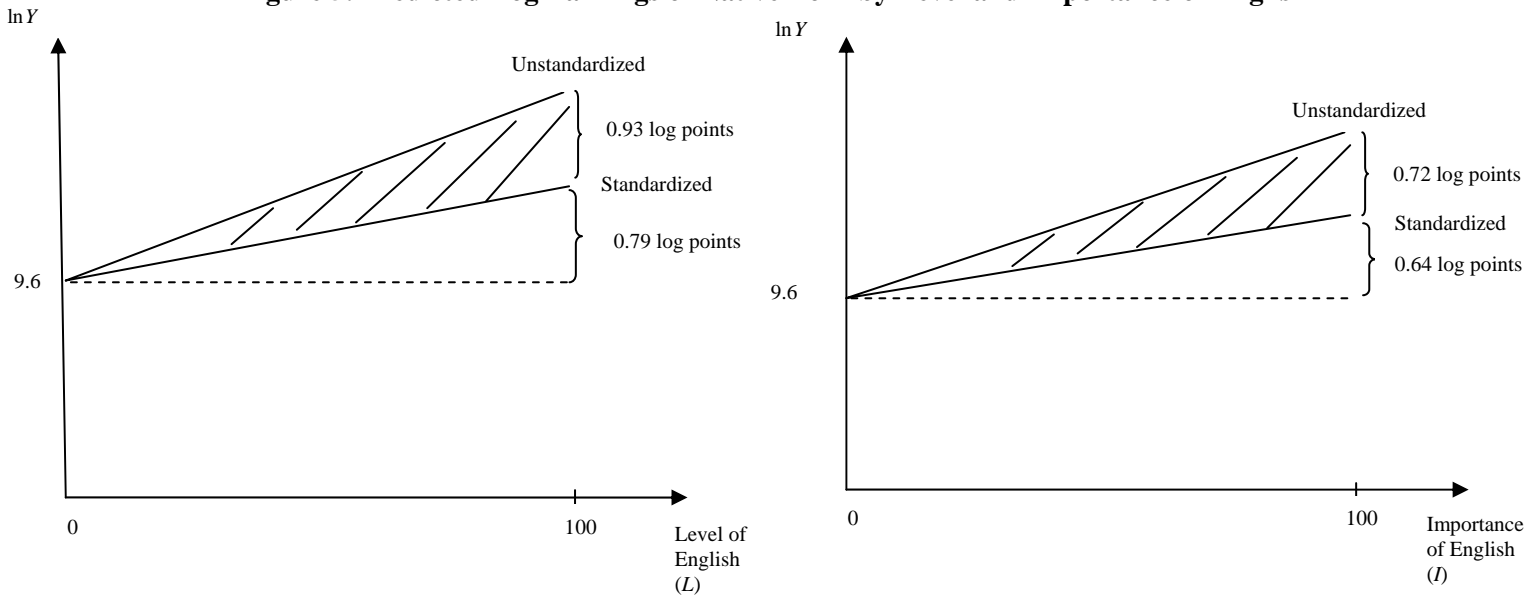
Table 2
Estimated Earnings Equations with Focus on the Importance of English Skill,
by Nativity, 2000 US Census

Variable	Native Born			Foreign Born		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Constant	9.695 (2750.01)	4.451 (198.20)	4.446 (197.77)	9.433 (1313.98)	5.707 (106.35)	5.504 (98.94)
Log Weeks Worked	(a)	0.997 (181.77)	0.997 (181.80)	(a)	0.869 (72.79)	0.866 (72.69)
Years of Education	(a)	0.079 (133.13)	0.080 (133.23)	(a)	0.039 (48.19)	0.039 (47.07)
Experience	(a)	0.035 (75.70)	0.035 (75.68)	(a)	0.014 (13.20)	0.014 (13.83)
Experience Squared/100	(a)	-0.061 (60.34)	-0.061 (60.32)	(a)	-0.021 (10.93)	-0.022 (11.40)
Married	(a)	0.259 (107.43)	0.258 (107.39)	(a)	0.204 (34.58)	0.202 (34.35)
South	(a)	-0.061 (26.82)	-0.061 (26.83)	(a)	-0.071 (11.89)	-0.074 (12.37)
Metropolitan Area	(a)	0.193 (32.70)	0.193 (32.66)	(a)	0.111 (4.15)	0.117 (4.43)
Veteran	(a)	-0.041 (15.26)	-0.041 (15.19)	(a)	-0.076 (5.84)	-0.070 (5.39)
Blacks	(a)	-0.145 (40.77)	-0.145 (40.62)	(a)	-0.183 (17.12)	-0.176 (16.40)
Years Since Migration (YSM)	(a)	(a)	(a)	(a)	0.012 (13.85)	0.012 (13.88)
YSM Squared/100	(a)	(a)	(a)	(a)	-0.012 (6.25)	-0.013 (6.53)
English Very Well	(a)	-0.066 (13.15)	-0.001 (0.06)	(a)	-0.078 (8.52)	-0.016 (0.68)
English Well	(a)	-0.101 (7.94)	-0.055 (1.76)	(a)	-0.221 (22.44)	0.080 (3.41)
English Not Well/Not at All	(a)	-0.003 (0.18)	0.114 (2.46)	(a)	-0.307 (28.80)	0.074 (3.26)
Importance of English (Impt) ^(b)	0.014 (212.65)	0.006 (100.30)	(a)	0.015 (100.11)	0.007 (44.61)	(a)
Impt*English Only	(a)	(a)	0.007 (99.42)	(a)	(a)	0.011 (30.77)
Impt*English Very Well	(a)	(a)	0.005 (21.40)	(a)	(a)	0.010 (40.63)
Impt*English Well	(a)	(a)	0.006 (9.20)	(a)	(a)	0.005 (17.79)
Impt*English Not Well/Not at All	(a)	(a)	0.004 (4.72)	(a)	(a)	0.002 (5.64)
Adjusted R ²	0.083	0.349	0.349	0.109	0.379	0.384
Sample Size	532,109	532,109	532,109	84,172	84,172	84,172

Note: (a) Variables not entered; (b) Importance of English required in the occupation of employment according to the O*NET database.

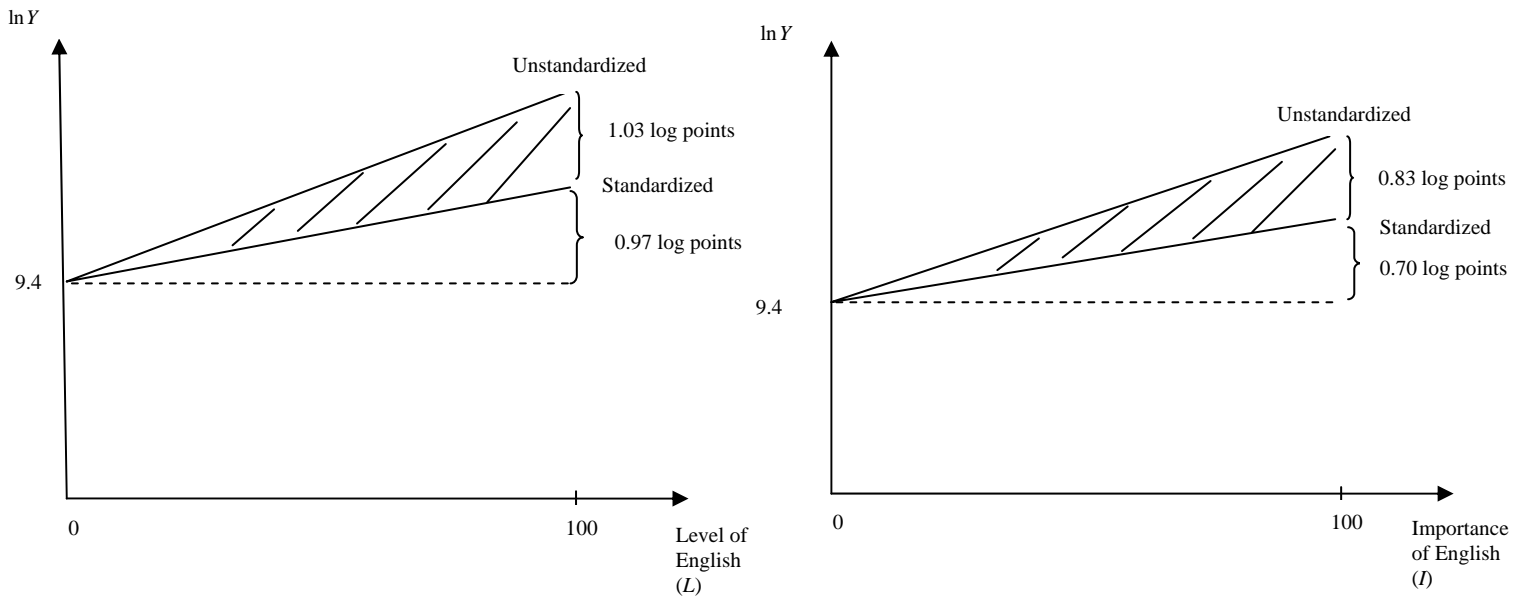
Source: 2000 United States Census, 1% PUMS.

Figure 5: Predicted Log Earnings of Native Born by Level and Importance of English



Source: Authors' calculations based on Tables 1 (Level) and 2 (Importance).

Figure 6: Predicted Log Earnings of Foreign Born by Level and Importance of English



Source: Authors' calculations based on Tables 1 (Level) and 2 (Importance).

B. Multiple Regression Analysis

Part of the difference in mean earnings across the occupational language requirements is due to the fact that workers in jobs that have low English requirements have fewer human capital skills than workers in jobs that have high English requirements. To establish the impact, *ceteris paribus*, on earnings of the English requirements of jobs, the earnings equation is augmented with a set of human capital and demographic standardizing variables (X). Hence it becomes:

$$\ln Y_i = \alpha_0 + X_i\beta + \alpha_m ER_i + \varepsilon_i$$

where the X_i vector contains the personal and other job-related characteristics that affect the earnings of individual i (including educational attainment, potential labor market experience, location, race, marital status, birthplace, the respondent's self-reported proficiency in English, and among the foreign born, duration of residence in the U.S.), and α_m is the partial effect of the English requirements variable on earnings.

Results from this specification are presented in column (ii) of Tables 1 and 2. The coefficients on the standardizing variables in the vector X are reasonably close to the evidence reported in recent studies. Thus, the payoff to one additional year of education for the native born is about 8 percent and the payoff for the foreign born is only one-half of this (*i.e.*, about 4 percent).⁶ The relative magnitudes of these payoffs for the native born and foreign born are consistent with evidence reported in Chiswick and Miller (2006). Reflecting the use of information on occupational English requirements in the current study, the payoffs to education in Tables 1 and 2 are about 20 percent less than those in the comparison study.

The continuous variable that records the natural logarithm of weeks worked in 1999 indicates an elasticity of earnings with respect to weeks worked close to unity for the native born, and around 0.87 for the foreign born. The marital status variable shows that married men earn more than their non-married counterparts, with their earnings

⁶ The payoff to potential labor market experience is 2.3 percent for the native born and 1 percent for the foreign born, when evaluated at 10 years. Similar results are documented by Chiswick and Miller (2006), where the payoff to potential labor market experience in the absence of control for the occupational English requirements was reported to be 2.3 percent for the native born and 0.8 percent for the foreign born.

advantage being 26 percent for the native born and 20 percent for the foreign born. These weeks worked and marital status effects on earnings are virtually identical to those reported where occupational English requirements are not held constant.

The estimated partial effect on earnings of the individual's proficiency in English, however, is greater in absolute value when the occupational English requirements are held constant than in the absence of this standardization. From the standard omitted variables formula, given that occupational English requirements have a positive correlation with the respondent's English skills, this change implies an intuitively reasonable negative correlation between the individual's self-reported proficiency in English variables (which record limited English skills) and the English requirements of the occupations in which they work. Nevertheless, there is only a modest change in the estimated effects on earnings of English language proficiency following standardization for the occupational English requirements. The estimated partial effects show that native-born men who speak a language other than English at home and who speak English very well or well have earnings 7 to 10 percent lower than those who only speak English at home. Among foreign-born men, those who speak a language other than English at home and who speak English very well earn 8 percent less than those who speak only English at home, while those who speak English well have about 22 percent lower earnings and those who speak English not well or not at all have earnings 31 percent less than the reference group that speaks only English at home.

The estimated impact of the English language requirements of the occupation falls by slightly more than one-half once account is taken of the differences across workers in other productivity-related (human capital and demographic) characteristics. Thus, the coefficient on the O*NET level of English variable falls from 0.017 to 0.008 for the native born, and it falls from 0.200 to 0.010 for the foreign born. The coefficient on the importance of English variable for the native born falls from 0.013 to 0.006, while that for the foreign born falls from 0.015 to 0.007. The relationships between earnings and the level and importance of English, *ceteris paribus*, are illustrated as the lower solid line in Figures 5-6.

C. Johnson and Solon (1986) Decomposition

The reasons for the diminution of the estimated coefficient on the English requirements variables as the focus moves from the simple regression (α_s) to the multiple regression that includes control variables (α_m) can be examined using the standard omitted variable formula (see Johnson and Solon, 1986):

$$\alpha_s - \alpha_m = \sum_{k=1}^K \alpha_k b_{kc},$$

where α_k is the estimated coefficient on the k^{th} control variable, and b_{kc} is the coefficient from a simple regression of the k^{th} control variable on the English requirements variable. In other words, this decomposition provides an explanation for the shaded areas in Figures 5 and 6. The results from this decomposition are presented in Table 3.

Table 3
Decomposition of the Influence of Regressors on the Estimations of the Level and Importance of English Skills Effects

	Native Born		Foreign Born	
	Level	Importance	Level	Importance
Difference between Impacts Estimated in Simple and Multiple Regressions	0.0093	0.0072	0.0103	0.0083
<i>Derives from:</i>				
Log Weeks Worked	0.0020	0.0017	0.0018	0.0015
Years of Education	0.0064	0.0048	0.0058	0.0046
Experience ^(b)	-0.0003	-0.0002	-0.0005	-0.0003
Married	0.0007	0.0006	0.0004	0.0004
South	0.0000	0.0000	0.0000	0.0000
Metropolitan Area	0.0001	0.0001	0.0000	0.0000
Veteran	0.0000	0.0000	-0.0000	-0.0000
Blacks	0.0002	0.0002	-0.0001	-0.0001
Years Since Migration (YSM) ^(b)	(a)	(a)	0.0003	0.0002
English Proficiency	0.0000	0.0000	0.0026	0.0020

Note: (a) Variables not relevant; (b) Includes effect of squared term.

Source: Authors' calculations based on Tables 1 and 2.

Table 3 shows the contribution of each set of variables to the diminution (see Figures 5 and 6) in the coefficients of the Level and Importance of English skills variables in the multiple regression setting compared to the simple regression. For the native born, all variables in the multivariate model, except for experience and veteran status, contribute to the diminution. Among the foreign born, weeks worked, educational attainment, marital status, years since migration and English proficiency contribute to the lessening of the impact of the occupational language requirements variables on earnings in the multivariate framework.

It is apparent from Table 3, however, that educational attainment is the chief contributor to the shaded areas of Figures 5 and 6. It accounts for about two-thirds of the change in the estimated impact of the occupational English requirements variables for the native born, and around 55 percent of the change for the foreign born. This is a reflection of better educated workers being matched with jobs that have higher English requirements.

The weeks worked variable accounts for a further 21 to 23 percent of the change in the estimated impact for the native born, and 17 percent of the change for the foreign born. Workers with a stronger attachment to the labor force are more likely to be found in occupations that require a higher level of English.

The English proficiency variables, which do not contribute to the diminution of the impact of the English requirements variable for the native born, account for around 24 percent of the diminution for the foreign born. The absence of an effect for the native born is associated with their nearly universal proficiency in English. The reasonably strong finding for the foreign born indicates a positive sorting of workers with better English skills into occupations that have more demanding English language requirements. The effects of all other variables are inconsequential.

D. Dickens and Katz (1987) Robustness Analysis

A potential problem with the analysis above is that it combines variables that are measured at two different levels of aggregation. Thus all data other than the O*NET occupation data are individual-level data, whereas the O*NET information is grouped data, collected at the level of the occupation. Dickens and Katz (1987) and Moulton

(1986) draw attention to the possibility that combining aggregate data with individual data may bias the estimates and also lead to incorrect standard errors.

Dickens and Katz (1987) outline an expedient approach that can be followed in the current analysis to check the reliability of the results in the face of this potential problem. They propose that the earnings equation in column (ii) of Tables 1 and 2 be estimated without the variables for the O*NET English requirements but allowing for occupation fixed effects. The occupation fixed effects (*i.e.*, the coefficients from occupation dummies included in this modified specification of the earnings equation) are then regressed on the O*NET English requirements variables in a subsequent regression, estimated at the level of the occupation.

This two-step approach should provide reliable estimates of the links between earnings and the occupational English requirements. It also provides information on the extent to which the across-occupation variation in earnings (after standardization for the workers' characteristics) can be accounted for using information on the occupational English requirements.

Table 4 lists the results of the earnings equations without the occupational English requirements but both with and without the dichotomous occupation variables.

The incorporation of occupation dummy variables into the regression equation (509 dichotomous variables for the native born and 491 for the foreign born) raises the adjusted R-squared by 8 percentage points for the native born (a 23 percent increase in the explanatory power of the equation) and by 9 percentage points for the foreign born (a 24 percent increase in the explanatory power of the equation). The change in specification is generally associated with a reduction (in absolute value) of the estimated impacts on earnings of the individual-level variables. For example, the estimated effect of years of education on earnings falls by about one-half once the occupational fixed effects are included in the model. This is a consequence of variables that are associated with higher earnings also being associated with above average representation in high-paying occupations.

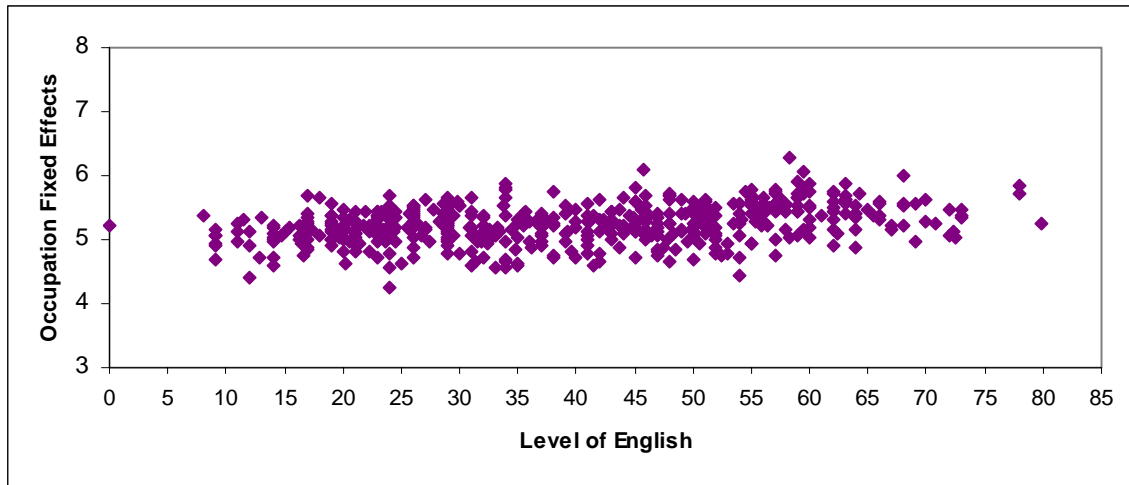
Table 4
Estimates of Earnings Function With and Without Occupation Dummy Variables,
by Nativity, 2000 US Census

Variable	Native Born		Foreign Born	
	Without Occupation Fixed Effects	With Occupation Fixed Effects	Without Occupation Fixed Effects	With Occupation Fixed Effects
Constant	4.346 (192.77)	(a)	5.850 (108.41)	(a)
Log Weeks Worked	1.010 (183.25)	0.957 (370.13)	0.875 (73.01)	0.842 (152.71)
Years of Education	0.106 (201.74)	0.058 (103.06)	0.053 (69.56)	0.023 (29.05)
Experience	0.033 (71.47)	0.032 (75.94)	0.011 (10.99)	0.019 (19.00)
Experience Squared/100	-0.056 (55.29)	-0.055 (63.35)	-0.016 (8.06)	-0.030 (16.31)
Married	0.270 (111.26)	0.203 (87.94)	0.214 (35.81)	0.159 (28.31)
South	-0.057 (24.71)	-0.075 (34.39)	-0.072 (11.75)	-0.089 (15.44)
Metropolitan Area	0.211 (35.57)	0.163 (30.63)	0.135 (5.02)	0.085 (3.25)
Veteran	-0.045 (16.52)	-0.045 (17.52)	-0.083 (6.28)	-0.046 (3.69)
Blacks	-0.153 (42.85)	-0.076 (21.28)	-0.185 (17.21)	-0.067 (6.53)
Years Since Migration (YSM)	(b)	(b)	0.011 (13.14)	0.012 (15.69)
YSM Squared/100	(b)	(b)	-0.011 (5.66)	-0.014 (8.18)
English Very Well	-0.059 (11.56)	-0.040 (8.26)	-0.080 (8.53)	-0.057 (6.95)
English Well	-0.097 (7.46)	-0.070 (5.93)	-0.261 (26.20)	-0.134 (15.05)
English Not Well/Not at All	0.013 (0.70)	0.010 (0.59)	-0.374 (34.99)	-0.224 (23.31)
Adjusted R^2	0.335	0.412	0.363	0.449
Sample Size	532,109	532,109	84,172	84,172

Notes: (a) Each occupation is assigned an intercept term; (b) Variable is not relevant.

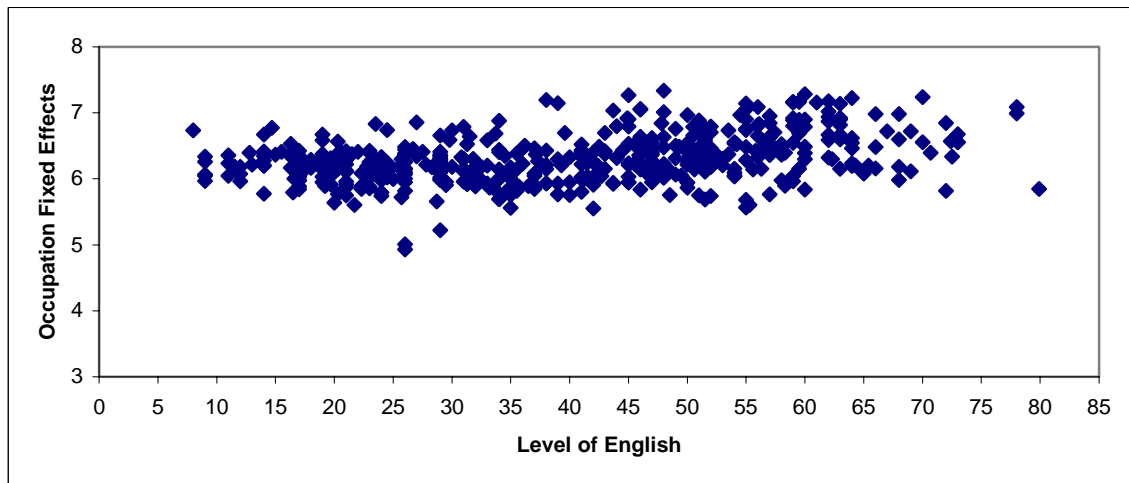
There is considerable variation in the magnitude of the coefficients of the estimated occupation fixed effects on earnings. For each birthplace group the coefficients of the fixed effects are plotted against the occupational English requirements, in Figure 7 for the native born and in Figure 8 for the foreign born. Given the similarity of the findings, only the plot for the level of English skill is presented.

Figure 7
Occupation Fixed Effects by Level of English, Native Born



Note: Authors' calculations based on Table 4.

Figure 8
Occupation Fixed Effects by Level of English, Foreign Born



Note: Authors' calculations based on Table 4.

The results from the analysis of the links between the coefficients of the occupation fixed effects on earnings and the occupation English requirements in Table 5 show that English requirements account for between 8 and 15 percent of the variation in the occupation fixed effects. The English requirements variable accounts for a higher proportion of the variation in the occupation fixed effects among the foreign born than among the native born, although the difference in the degree of explanation is small.

In each earnings regression, the English requirements variable is highly significant. However, reflecting the lower degrees of freedom when the data are analysed at the higher level of aggregation, the ‘t’ statistics are far smaller than those that were reported in Tables 1 and 2. Nevertheless, with the smallest ‘t’ being 7.17, the significance of the English requirements variables is not in doubt.

The estimated impact on earnings of the information on the English requirements, as per the analysis of the occupation fixed effects in Table 5, are all sizeable, though about 0.2 of a percentage point less than the effects estimated on the basis of mixing aggregate-level data with the micro level data (see Tables 1 and 2). This suggests merit to the robustness checks reported here. But it also shows that the fundamental theme of the study, that English language requirements are very important to the understanding of variations in earnings, carries across the alternative sets of analyses reported here.

Table 5
Estimated Effects of Occupation English Requirements on Occupation Fixed Effects^(a)

	Native Born		Foreign Born	
	Level of English	Importance of English	Level of English	Importance of English
Constant	4.993 (170.16)	5.053 (188.88)	5.961 (162.05)	6.045 (194.09)
English Requirements	0.006 (8.90)	0.004 (7.41)	0.009 (9.13)	0.005 (8.03)
\bar{R}^2	0.121	0.086	0.148	0.104
Sample Size	505	505	487	487

(a) Occupation Fixed Effects are from regressions as per columns (ii) and (iv) of Table 4.

E. English Skills of the Worker and of the Occupation

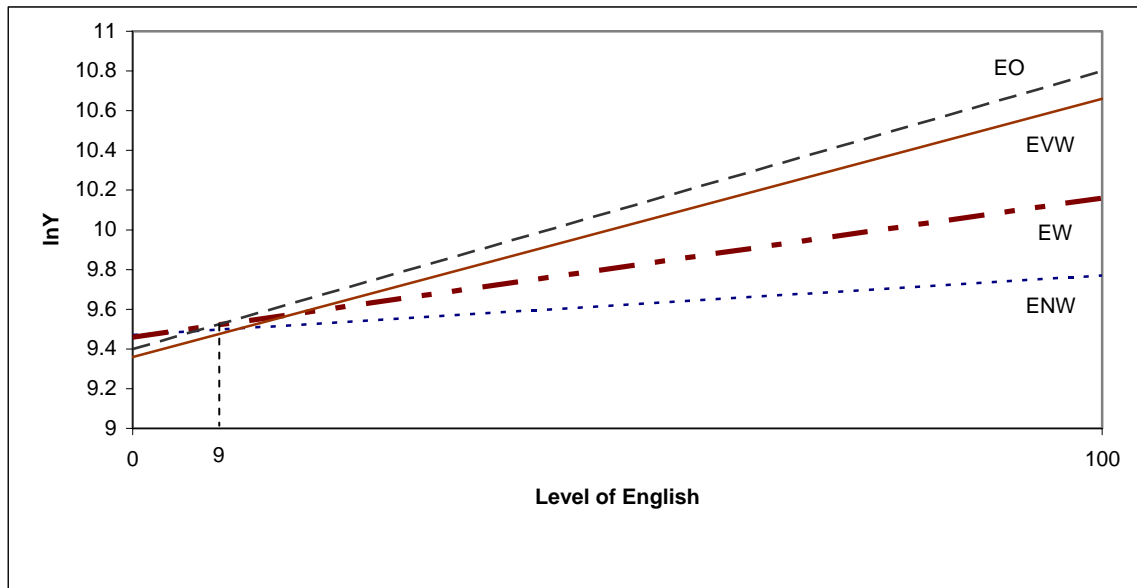
The Johnson and Solon (1986) decomposition reveals that workers with greater proficiency in English are more likely to be employed in occupations that require relatively high levels of English. This would indicate a matching of worker and job attributes. This matching is presumably also reflected in the relative rates of remuneration that workers with different levels of English can obtain in occupations requiring low and high levels of English language skills.

To examine this issue, interaction terms between the variables for O*NET English requirements and the workers' self-reported proficiency in English were included in the estimating equation. These results are reported in column (iii) of Tables 1 and 2. While the same sets of analyses are undertaken for the native born and for the foreign born, the discussion will concentrate on the findings for the foreign born because nearly all of the native born speak only English.

The patterns for the level and importance of English required for the occupation are similar, and the focus will be on the level of English. The earnings-English requirements profiles for the four levels of English proficiency are presented in Figure 9.

Figure 9 makes it very clear that the earnings gains associated with the English requirements of the job depend on the individuals' proficiency in English. For English only speakers, there is a gain of 1.4 log points in earnings across the level of English scale. The gain for those who speak a language other than English at home and who speak English very well is slightly smaller, 1.3 log points. The difference in these effects is statistically significant at the 10 percent level. However, the increases in earnings across the O*NET level of English scale for those who speak English "Well" is only 0.7 log points – one-half the gains for English only speakers. Finally, the increases in earnings across the O*NET level of English scale for those who speak English "Not Well or Not at All" is even smaller, being only 0.3 log points, about 20 percent of the gains for English only speakers.

Figure 9
Predicted Earnings by Level of English for English Proficiency Groups, Foreign Born



EO = English Only;
 EVW = English Very Well;
 EW = English Well;
 ENW = English Not Well/Not at All.

Source: Authors' calculations based on Table 1.

Associated with the different slope effects of the English requirements variables are different intercept effects. While these differences in intercepts are small, they tell an interesting story. For those who speak English Very Well, the intercept shift effect is negative. This means that, compared to the English only benchmark group, those who speak a language other than English at home and speak English very well have lower earnings at all levels of English required in the job. The difference is 4 percentage points at a zero level of English required in the job, and this difference widens to 14 percentage points at (a hypothetical) 100 points level of English requirement in the job.

Those who speak a language other than English at home and who have more limited English skills (*i.e.*, they speak English Well, Not Well or Not at All) actually have higher earnings than the benchmark group of English only speakers if they are in jobs

with essentially no English requirements, (e.g., Maids and Housekeeping Cleaners have a score of zero.) However, once they are in jobs that require a level of English of 6 to 9 or more, those with limited English skills are at an earnings disadvantage (e.g., Slaughterers and Meat Packers have a score of 8). This disadvantage rises rapidly across the scale for required levels of English. This suggests a negative selectivity on other characteristics of English only speakers in jobs that require no English language proficiency.

The rewards to correct matching of the skills of workers and the requirements of jobs are very apparent in the U.S. labor market. This pattern is also evident in the analyses of the links between the English skills of foreign-born workers and the O*NET data on the importance of English reported in Table 2. The pattern also characterizes analyses for the native born, although these differences are compressed perhaps because only a very small percentage of the native born speak a language other than English.

Thus, there is an interaction effect on earnings between the respondent's proficiency in English and the level of English required for the occupation. Those with higher levels of English language proficiency earn more, but the increase in earnings is greater if they are in an occupation that requires greater proficiency in English. Those who are proficient in English but in occupations that do not require English language skills are not taking advantage of a skill that they have.

IV. CONCLUSION

This paper examines the English language occupational requirements in the U.S. labor market. It also estimates the value of proficiency in English by evaluating its impact on earnings for men aged 25 to 64 years who worked in 1999. It extends previous research by its use of measures of English language requirements in the O*NET database, and its exploration of the interactions between these measures, worker characteristics and earnings.

The O*NET database provides information pertaining to the standardized scores of the importance and level of English language skills in each of over 900 occupations. The importance and level scores display a high level of correlation, both overall and within each of the six broad occupational groups considered. These occupational English

scores were linked to data from the 2000 U.S. Census, and simple and multiple earnings regression analyses are conducted, separately for the native born and the foreign born.

The simple regression analyses show that earnings are strongly related to the English language requirements of the occupations. Among the native born, there is a difference of 1.7 log points across the level of English scale, and a difference of 1.3 log points across the importance of English scale from zero to 100. Among the foreign born, the English requirements variables have even stronger links with earnings, with the differences across the measurement scales being 2.0 log points for the level of English data and 1.5 log points for the importance of English data. However, these earnings effects are reduced by about one-half once other productivity-related characteristics, including the worker's own English language proficiency, are taken into account.

This result is robust to the alternative method of estimation proposed by Dickens and Katz (1987) that involves regressing the O*NET occupational English requirements on occupation fixed effects obtained from the earnings equation. The standard omitted variable formula (Johnson and Solon, 1986) is employed to explain this diminution of the estimated coefficient on the English requirements variables. The results indicate that educational attainment and weeks worked, and, for the foreign born, English language proficiency, are the main contributors to the diminution. This is a reflection of the sorting of workers with a higher level of human capital, including language capital, and a stronger attachment to the labor force, into jobs that require higher levels of English.

Interaction terms between the self-reported English language proficiency of workers and the English language requirements of the occupations in which they work reveal that there are rewards to correct matching of worker skills and job requirements in the U.S. labor market. Workers with poor English skills do relatively well when employed in jobs that have very low English language requirements. They do relatively poorly when employed in jobs that have high English language requirements. These results are very striking for the foreign born, but also characterize the determinants of earnings for the native born.

For both birthplace groups, labor markets appear to sort workers appropriately, with those with levels of English proficiency tending to work in jobs which require a high level of proficiency and in which English language proficiency is important.

APPENDIX A

DESCRIPTION OF VARIABLES

The variables used in the statistical analyses are defined below. Mnemonic names are also listed where relevant.

Data Source: 2000 United States Census of Population, 1 percent Public Use Microdata Sample; O*NET Occupational data from the O*NET Consortium described in footnote 1.

Definition of Population: Native-born and Foreign-born men aged twenty-five to sixty-four in the 50 States and the District of Columbia.

Dependent Variable:

Earnings: The natural logarithm of the individual's annual earnings from wage and salaried employment or self-employment for 1999.

Explanatory Variables:

Educational Attainment (EDUC): The total years of full-time equivalent education. The following values are assigned to the Census categories: completed less than fifth grade (2 years); completed fifth or sixth grade (5.5); completed seventh or eighth grade (7.5); completed ninth grade (9); completed tenth grade (10); completed 11th grade (11); completed 12th grade or high school (12); attended college for less than one year (12.5); attended college for more than one year or completed college (14); Bachelor's degree (16); Master's degree (17.5); Professional degree (18.5); Doctorate (20).

Labor Market Experience (EXP): A measure of potential labor market experience, computed as AGE – Years of Education – 6.

Log Weeks Worked (WEEKS): The natural logarithm of the number of weeks the individual worked in 1999.

Years Since Migration (YSM). This is computed from 2000 minus the year the foreign born person came to the United States to stay.

English Language Fluency: Three dichotomous variables that distinguish individuals who speak a language other than English in the home and who speak English either: (i) "Very Well"; (ii) "Well"; (iii) "Not Well"/"Not at All". The benchmark group is those who speak only English at home.

Race (BLACK): This is a dichotomous variable, set to one if the individual is Black, and set to zero for all other racial groups.

Marital Status (MARRIED): A dichotomous variable that distinguishes individuals who are married, spouse present (equal to 1) from all other marital states.

Location: The two location variables record residence in a non-metropolitan area (NON-MET) or in a Southern State (SOUTH). The states included in the latter are: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia.

Veteran (VETERAN): This is a dichotomous variable, set to one if the individual is veteran of the U.S. Armed Forces, and set to zero otherwise.

English Requirements (LEVEL and IMPORTANCE): These variables record the scores for the level and importance of English requirements for each occupation obtained from the O*NET database.

RERERENCES

- Chiswick, Barry R. and Miller, Paul W., (1995) “The Endogeneity Between Language and Earnings: International Analyses”, *Journal of Labor Economics*, Vol. 13, No. 2, pp.246-288.
- Chiswick, Barry R. and Miller, Paul W., (2002). “Immigrant Earnings: Language Skills, Linguistic Concentrations and the Business Cycle”, *Journal of Population Economics*, Vol. 15, No. 1, pp.31-57.
- Chiswick, Barry R. and Miller, Paul W., (2006). “Why is the Payoff to Schooling Smaller for Immigrants?”, mimeo, Department of Economics, The University of Illinois at Chicago.
- Dickens, W. T. and Katz, L. F., (1987). “Inter-Industry Wage Differences and Industry Characteristics”, in K. Lang and J.S. Leonard (eds.), *Unemployment and the Structure of Labor Markets*, Basil Blackwell, New York, USA.
- Hartog, J., (2000). “Over-education and Earnings: Where Are We, Where Should We Go?”, *Economics of Education Review*, Vol. 19, No. 2, pp. 131-147.
- Johnson, G. and Solon, G., (1986), ‘Estimates of the Direct Effects of Comparable Worth Policy’, *American Economic Review*, Vol. 76 (5), pp. 1117-1125.
- Moulton, B. R., (1986). “Random Group Effects and the Precision of Regression Estimates”, *Journal of Econometrics*, Vol. 32, No. 3, pp. 385-397.
- Rumberger, Russell, W., (1981). “The Rising Incidence of Overeducation in the U.S. Labor Market”, *Economics of Education Review*, Vol. 1 , No. 3, pp. 293-314.