THE IMPACT OF SURPLUS SKILLS ON EARNINGS:
Extending the Over-Education Model to Language Proficiency

Barry R. Chiswick
Department of Economics
George Washington University
and
IZA-Institute for the Study of Labor

Paul W. Miller
School of Economics and Finance
Curtin University

March 2013 (R)
ABSTRACT

• To study the effect on earnings of the matching of English language skills to occupational requirements or occupational norms for adult male immigrants.

• Data from the Occupational Information Network (O*NET) database (Worker Self Assessment) and a “Realized Matches” procedure to quantify expected levels and importance of English skills in each of over 500 occupations in the US Census.

• Earnings data from the 2000 US Census for foreign-born adult male workers are then examined in relation to these occupational English requirements or norms using the Over/Required/Under (or ORU) technique developed for the study of schooling. Analysis of native-born men for English requirements.

• The analyses show that earnings are related to language norms in the occupation (FB and NB) and a “correct” matching of individual language skill and the occupation norm (FB).

• Mismatches among immigrants have a smaller effect on earnings – positive for extra proficiency and negative for deficits in proficiency, relative to the norm in the occupation.

• The findings are robust with respect to a range of measurement and specification issues.
ORU Model: Over-, Required, or Under-Qualified

Education:

- Hartog (*EER* 2000)
  Immigrants: Chiswick & Miller (*LE* 2008)

Findings:

1. Earnings of adult male respondents increase with required educational attainment.

2. Over/Under-education effects are smaller in absolute value than the effects of required education.
   - Over-education: Positive effect on earnings
   - Under-education: Negative effect on earnings

3. In simple OLS regression,
   - Coefficient on actual schooling is smaller for immigrants than for the native born.
   - Coefficient on required schooling (controlling for over/under-education) is the same (about 15 percent) for immigrants and natives.

Other Human Capital = Language ??
Hypotheses: ORU Model for Language

1. Earnings rise with required English language proficiency (level or importance) in occupation.

2. Positive effect on earnings if proficiency is greater than required.

3. Negative effect on earnings if proficiency is less than required.

What are the magnitudes of these effects?
The Data

1. Worker self-assessment -- O*NET (Occupational Information Network)
   – Reports wide range of characteristics of narrowly-defined occupations

   “How important is knowledge of the English language to the performance of your current job?” (5-point scale)

   “What level of English language proficiency is needed to perform your current job?” (8-point scale)

   - Standardized (normed) to 100-point scale
   - High correlation (R=0.92) between Level and Importance (801 occupations, O*NET data)
   - Focus of Analysis: Level of English

2. Realized Matches
   - Uses mean English language proficiency of foreign-born adult workers in each occupation in the 2000 Census.
Importance of English:
Relative Frequency of Occupations
on Standardized Score

Mean = 59.84,  SD = 18.19,  No. of Occupations = 801

Low Importance
Paperhanger (8)
Precious Metal Workers (13)
Logging Equipment Operators (20)
Models (20)

High Importance
Judges (95)
Proofreaders (95)
Economists (91)
Sociologists (84)
Public Relations Managers (96)
Required Level of English:
Relative Frequency of Occupations on Standardized Score

Mean = 49.44,  SD = 15.60,  No. of Occupations = 801

Low Level
- Glaziers (20)
- School Crossing Guards (20)
- Postal Service Clerks (40)

High Level
- Economists (73)
- Sociologists (78)
- Environmental Science Teachers (Post-secondary) (80)

Correlation between required level and importance: \( R = 0.92 \)
(unweighted)
The Workers
(2000 Census of Population, PUMS, 1% sample)

• The Sample:
  – Men ages 25-64
  – Foreign-born and Native-born
  – Non-zero earnings in 1999

• The Language Question:
  – Is a language other than English spoken by the respondent in the home?
  – If so, what is it? How well do you speak English?

• Self-reported proficiency:
  5 = speaks only English *
  4 = speaks English Very Well
  3 = speaks English Well
  2 = speaks English Not Well
  1 = speaks English Not at all
  (Proficiency converted to a 100-point scale.)

* Nearly all native-born (≈95%) speak only English at home. Most of the others report speaking English “very well.”
The Equation and Hypotheses

\[
\ln Y_i = \beta X_i + \gamma_o ENG_{oi} + \gamma_r ENG_{ri} + \gamma_u ENG_{ui} + \eta_i
\]

where

- \( ENG_o = ENG_{\alpha} - ENG_r \) if \( ENG_{\alpha} > ENG_r \)
  \( = 0 \) otherwise

- \( ENG_u = ENG_r - ENG_{\alpha} \) if \( ENG_r > ENG_{\alpha} \)
  \( = 0 \) otherwise

and \( ENG_{\alpha} = ENG_r + ENG_o - ENG_u \)

Hypotheses:

- \( \gamma_r \) and \( \gamma_o > 0 \), \( \gamma_u < 0 \)
- \( \gamma_r > \gamma_o \), \( \gamma_r > \vert \gamma_u \vert \)

\( \alpha = \) Observed English proficiency
<table>
<thead>
<tr>
<th>Level of English</th>
<th>Foreign Born</th>
<th>Native Born</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required</td>
<td>0.017</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(62.1)</td>
<td>(67.9)</td>
</tr>
<tr>
<td>Over-qualified</td>
<td>--</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(20.86)</td>
</tr>
<tr>
<td>Under-qualified</td>
<td>--</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(12.5)</td>
</tr>
</tbody>
</table>

`t-ratios in parentheses.``

Required Level: Postal Service Clerks 40
Glaziers 20

Foreign Born: $20 \times 0.017 = 0.34$ log points

Native Born: $20 \times 0.013 = 0.26$ log points
### Table B

<table>
<thead>
<tr>
<th>Required Level of English</th>
<th>Actual Level of English</th>
<th>Skill Classification</th>
<th>ln Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>40</td>
<td>40</td>
<td>Correct Match</td>
</tr>
<tr>
<td>B</td>
<td>50</td>
<td>50</td>
<td>Correct Match</td>
</tr>
<tr>
<td>C</td>
<td>60</td>
<td>60</td>
<td>Correct Match</td>
</tr>
<tr>
<td>D</td>
<td>50</td>
<td>60</td>
<td>Overqualified</td>
</tr>
<tr>
<td>E</td>
<td>50</td>
<td>40</td>
<td>Underqualified</td>
</tr>
</tbody>
</table>

#### Earnings of the Five Hypothetical Workers Described Above

[Graph showing earnings of the five hypothetical workers with regression line and labels for underqualified and overqualified workers]
Robustness Checks

1. Analyses of Importance in addition to Level

2. Alternative measures of scoring or scaling English proficiency

3. Replace O*NET(self-assessment) with English scores by occupation in the 2000 Census for immigrant men (Realized matches)

4. Quadratic specification of O*NET English language requirements

Findings: Similar patterns.
Model is highly robust
Summary and Findings

• Analysis of effects on adult male earnings
  – English language requirements of occupation (natives and foreign born)
  – Mismatch between occupational requirements and own proficiency (foreign born)

• Better English skills improve earnings:
  – Qualifies for higher-earning occupations (major)
  – Improves earnings within occupations (minor)
  – Earnings penalty for underqualified workers

• Better matching of workers’ language skills to occupation improves earnings.

• Usefulness of ORU Approach to studying language capital.
  – Possible application to other forms of Human Capital?
    (e.g., health, stamina, collegiality, etc.)