NBER WORKING PAPER SERIES

THE LANGUAGE ABILITY OF U.S. IMMIGRANTS: ASSIMILATION AND COHORT EFFECTS

Geoffrey Carliner

Working Paper No. 5222

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1995

George Borjas, Philip Levine, and Suzanne Reyes provided useful comments on an earlier draft, and Peronet Despeignes helped with research assistance. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #5222
August 1995

THE LANGUAGE ABILITY OF U.S. IMMIGRANTS: ASSIMILATION AND COHORT EFFECTS

ABSTRACT

This paper uses data from the 1980 and 1990 U.S. Census of Population to examine the English language skills of natives and immigrants. The first main finding is that lack of fluency in spoken English is rare among native-born Americans. In 1990, 98.4 percent of natives aged 18 to 64 reported speaking only English or speaking it very well. Among native-born children of ethnic groups who have come to the U.S. in large numbers during the past 30 years, such as Hispanics and East Asians, a substantial fraction were not fluent when they entered grade school, but at most 3 to 5 percent of teenagers and adults in these groups reported speaking English poorly or not at all.

Second, the vast majority of immigrants speak English well. In 1990, only a quarter of immigrants reported speaking English poorly or not at all, though more than half of Mexicans and one third of immigrants from other non-English speaking western hemisphere countries could not speak English well. Although English skills improve with length of residence, after 30 or more years in the U.S. over a quarter of Mexican immigrants spoke English poorly or not at all.

Third, since the 1950s there has been a trend decrease in the probability of fluency (speaking only English or speaking it very well) among new immigrants of about 0.1 percentage points per year, caused by the shift from European immigrants with strong English skills to Latin American and East Asian immigrants who arrive speaking less English. On average, each additional year of residence in the U.S. increases the probability of fluency by 1.1 percentage points. An additional year of schooling increases the probability of fluency by about 5 percentage points. Overall, women are slightly more likely to be fluent than men, especially East Asian and European women. Even after controlling for differences in education, years since arrival, and other factors, large differences in English skills by region of origin remain. These differences seem to be more associated with geographic distance from the U.S. than with the source country's per capita income or linguistic distance from English.

Geoffrey Carliner
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138

THE LANGUAGE ABILILITY OF U.S. IMMIGRANTS: ASSIMILATION AND COHORT EFFECTS

Geoffrey Carliner

I. Introduction

Speaking English well is important to success in the U.S. labor market and to full participation in American society. Immigrants who speak little or no English have greater difficulty finding jobs, especially well paid jobs outside immigrant enclaves. Jobs which involve contact with native-born customers or co-workers usually require the ability to speak English well. Therefore people who plan to migrate to the U.S. have an incentive to learn English before they come. Immigrants who arrive without speaking it, or speaking it poorly, have even greater incentives to invest in English language skills.

Native-born Americans also have an interest in immigrants speaking English well. Without strong English skills immigrants are more likely to be in poverty, more likely to raise children who do poorly in school, and less able to become citizens and to join in the political process. In some areas, concern over the language abilities of immigrants has expressed itself in legal steps to require the use of English in certain state activities.

This paper examines the English speaking abilities of the U.S. population, as measured in the 1980 and 1990 Censuses of Population. I use pooled data from these two large cross sections to estimate the effects of education, sex, country of origin, age at entry, years since entry, and year of entry on language skills. Most of the previous studies which examine the language skills of immigrants, surveyed in Borjas (1994b), use it to explain earnings and do not try to analyze its determinants. The studies which have focussed on

immigrants' language skills have used cross section data and have not been able to separate the effect of years since entry (assimilation effect) from changes over time in skills which immigrants have when they arrive (cohort effect) Depending on the data used, these studies vary in their measures of language skills, the characteristics of the respondents, and their analytical approach. In general they find that education, age at entry, and years since entry have large effects on these skills.

Veltman (1988) examines the shift from Spanish to English in the language spoken at home by Hispanic immigrants. He finds that school-age children make this shift soon after arrival, but the speed of the shift falls rapidly with age. Similarly, Portes and Schauffler (1994) report that 86 percent of native-born Cuban-American children prefer speaking English to Spanish. Chiswick (1991) analyzes a survey of illegal immigrants, mostly Mexican, arrested in California in 1986-87. He finds that the vast majority spoke no English when they arrived in the U.S., but most learned at least a little within a few years.

Other studies examine national samples of immigrants. Chiswick and Miller (1992), (1995), and Dustmann (1994) analyze data from the U.S., Canada, Australia, and Germany respectively. They all find large differences in language skills by country of origin and by education level, after adjusting for other characteristics. Boyd (1992) presents average skill levels by country of origin, education, and years since arrival to focus on male-female differences among immigrants in the U.S. and Canada. Finally, Carliner (1981), Robinson (1988), and Chiswick and Miller (1994) analyze the choices made by immigrants and natives in speaking French and/or English in Canada.

The next section of this paper discusses the determinants of English skill levels.

Section III describes the data used in this analysis. Section IV presents simple statistics on the language skills of immigrants and native-born Americans. Section V analyzes regression results for immigrants from nonEnglish speaking countries. A conclusion summarizes the main findings.

II. Investing in English Language Skills

Native-born American children learn to speak fluent English in the normal course of growing up. In contrast, immigrants from nonEnglish speaking countries must usually invest time, effort, and often money to learn English, either in school before they immigrate or perhaps informally after they arrive. Like other investments in human capital, the decision to invest in English language skills depends on the costs and benefits of doing so.

In many nonEnglish speaking countries children often learn English in school. People who anticipate emigrating to the United States may be more likely to study English, or people who have studied English may be more likely to emigrate. In either case, a large fraction of immigrants arrive in the U.S. speaking some English. However, both the costs and benefits of investing in English skills change sharply after they arrive. The returns to these skills, both in labor markets and in consuming leisure, are surely higher in the U.S. than for most immigrants had they chosen not to immigrate. The costs of learning English via television, informal contact with English speakers, or formal schooling are likely to be lower in the U.S. English language skills should therefore increase with years since migration (YSM) to the U.S. for immigrants from nonEnglish speaking countries.

Education is associated with both lower costs and greater benefits to learning English.

Whatever personal characteristics lead individuals to invest in human capital in the form of years of schooling -- a lower discount rate or greater facility in acquiring new skills -- are also likely to apply to human capital in the form of English language skills. Furthermore, immigrants with little education can often find jobs within ethnic enclaves that do not require a knowledge of English, but most jobs for well educated workers involve extensive contact with natives. Therefore the wage premium for speaking English well rises with education, according to Kossoudji (1988), McManus (1990) and Carliner (1995). Because of both cost and benefit considerations, English skills should therefore increase with formal years of schooling among immigrants from nonEnglish speaking countries. Both the English skills of entering immigrants and the improvement in these skills with years of residence should be positively correlated with education.

There may also be a difference between men and women in English skills. If women are more likely than men to immigrate to join their spouses rather than for economic reasons, they may speak less English on entry. Because women are less likely to work in the labor market than men, on average they receive smaller rewards to learning English. They may also face higher costs in learning English, because immigrants who remain at home raising children come in contact with English speakers less often than immigrants in the workplace. Therefore English skills may show less improvement over time for women than for men.

Age at entry (Agem) will also affect English skills, since the ability to learn new languages declines with age. In addition, immigrants who arrive at a young age have more years in which to recoup their investment in English skills and therefore a greater incentive to invest. Other things equal, immigrants who come to the U.S. from nonEnglish speaking

countries when they are young will learn more English than immigrants who arrive when they are older. Including age at entry in the analysis seems more intuitive than using age at the time of the survey, as some studies do, since it is not clear why the latter would affect language skills. Including both is not possible, since by definition

(1) $$Age = Agem + YSM$$

for all individuals in all time periods, in pooled as well as in cross section data.

Finally, English skills also vary by country of origin. For nearby countries where information about U.S. labor markets is widely available, barriers to entry may be low. People with weaker English skills may choose to migrate than in more distance countries where only the best educated know about job opportunities in the U.S. Countries also vary in the use of English among people who speak different domestic languages, the emphasis given to teaching English in school, the number of television shows or movies which use English, and in general on the importance placed on learning English. Chiswick and Miller (1995) suggest that preimmigration exposure to English may depend on the linguistic distance of the source country's language from English. Borjas (1987), (1994b) suggests that immigrants from developing countries with less equal income distributions should have lower skills than immigrants from richer countries with more equal distributions. If this hypothesis is correct, then English skills as well as other forms of human capital should vary inversely with the per capita income of the source country. For all these reasons, the average English skills of immigrants on arrival will vary by country of origin. In addition, country of origin

may also affect the increase in English skills that occurs after arrival. Immigrants who expect to remain in the U.S. for the rest of their lives, for instance political refugees, may be more likely to invest in English skills than economic immigrants who anticipate returning to their home countries after a few years.

Lazear (1995) and others have suggested that social and cultural factors, for instance the size and cohesion of immigrant enclaves in the U.S., may also affect the amount of improvement in English abilities that occurs after entry. However, the effect of language skills is likely to be even more important in immigrants' choice of where to live. Immigrants with little or no English will be more likely to live in ethnic enclaves than fluent speakers. Bartel (1989) finds that education, which is highly correlated with English language skills, decreases the probability that a newly arrived immigrant will settle in cities with a high concentration of fellow countrymen. Borjas (1994a) reports that immigrants to the U.S. from Canada or the U.K. are far less likely to live in ethnic enclaves than Mexican immigrants, whose average English skills are low when they arrive.

A priori, it seems likely that speaking English poorly would have a larger effect on the probability of living in an ethnic enclave than living in an enclave would have on the probability of learning to speak well. If so, including measures of neighborhood effects in language regressions, as Chiswick and Miller (1992), (1995) do, will capture the effect of language on neighborhood choice more than the effect of neighborhood on language skills. Separating these two effects requires data on individuals who do not choose where they live. For instance, a study which compared children of immigrants with poor English skills raised in ethnic enclaves with similar children raised elsewhere could avoid the simultaneity

problem. Borjas (1994a) uses such an approach to estimate neighborhood effects on the education and wage rates of the children of immigrants, but not on their English language skills.

There is also simultaneity between language ability and the country of origin of a spouse. Marrying a native English speaker may improve the language skills of immigrants, but being married to a native is more likely the result than the cause of proficiency in English. Sharing a mutual language is usually a prerequisite to a successful courtship. Therefore the language regressions estimated below do not include measures of marital status or residence in an ethnic enclave.

In sum, the English skill (L_i) of an immigrant in the U.S. includes the skill level at entry (E_i) and the amount learned with increased residence (A_i*YSM_i).

$$(2) \qquad L_i = E_i + A_i*YSM_i$$

Both the entering skill level and the improvement in English skill with increased residence -- the assimilation effect -- may vary with the individual's education, age at entry, sex, and country of origin (C), and other unobserved characteristics (u, v).

$$(3) \qquad E_i = f(Ed_i, Agem_i, Sex_i, C_i, u_i)$$

$$(4) \qquad A_i = g(Ed_i, Agem_i, Sex_i, C_i, v_i)$$

If the characteristics of the immigrant flow changes over calendar time, as occurred during

the past 25 years, both the entering skill level and the assimilation effect may change for different cohorts of immigrants.

III. Data

The data used in this study come from the 1980 and 1990 Census Public Use Samples, which asked respondents if they spoke English only, very well, well, not well, or not at all. These data are far better than variables available in earlier Censuses on mother tongue or ethnic origin. They are also better than the 1976 Survey of Income and Education used in other research because of the far larger sample size. They also provide the advantage of snapshots from two points in time, though individuals are not matched across Census years.

Census data are not perfect, however. Although precise year of age in the Census year is reported, only a range of years is given for year of entry into the U.S. I therefore assigned the midpoint of each range as the value of year of migration (YRM), the difference between the Census year and YRM as years since migration (YSM), and the difference between age in the Census year and YSM as age at migration (Agem). Immigrants who arrived before 1950 were excluded from any analysis which used these variables, since this range was open ended in Census coding. Because YRM, YSM, and Agem are all measured with error, their coefficients in the language regressions reported below will be biased toward zero.

Immigrants were grouped into seven country-of-origin categories: English speaking countries, Mexico, other western hemisphere countries, other (continental) European

countries, Africa and the Middle East, South Asia, and East Asia. English speaking countries were defined as all countries from which at least half the immigrants in the 1990 Census reported speaking only English. This category included Canada, Bermuda, Jamaica, Anguilla, Antigua, Aruba, Bahamas, Barbados, Cayman Islands, Grenada, Montserrat, St. Barts, St. Kitts, St. Lucia. St. Vincent, Trinidad, Turks and Caicos, Belize, Guyana, United Kingdom, Ireland, Gibraltar, Liberia, Zimbabwe, South Africa, Australia, and New Zealand. Countries of the former Soviet Union were included with continental European countries. The Middle East was defined to extend from Turkey to Iran. South Asia includes all countries from Afghanistan to Bangladesh. East Asia includes China, Burma (Myanmar) and countries eastward into the Pacific Ocean. This classification is necessary in the absence of information on mother tongue in the Census data. Unfortunately, it results in the anomaly that a small number of immigrants from English speaking countries (Canada) do not speak English and other immigrants from nonEnglish speaking countries (India) were raised speaking only English.

IV. Language Skills in the United States

Before examining immigrants, let us review briefly the English skills of natives. As Table 1 indicates, in 1990 93.9 percent of natives aged 18 to 64 reported that they spoke only English, and an additional 4.7 percent said that they spoke it very well. Only 0.4 percent reported speaking English not well or not at all. Thus a lack of English skills, at least at the basic level measured in Census data, is not a serious problem for U.S. natives of working age.

The 1980 and 1990 Censuses do not identify natives whose parents were immigrants, but it does identify natives of Hispanic descent, many of whom are the children of immigrants. Table 1 presents the percentage distribution of English language skills in 1990 for Hispanic natives of Mexican descent aged 5-8, 9-12, 13-17, and 18-64. Among the adults, 31 percent spoke only English, and an additional 52 percent spoke it very well. Only 4.5 percent reported speaking English not well or not at all.

Among the children, English ability improves with age, presumably because of increased contact with people outside the family, especially in school. Less than 68 percent of natives of Mexican ancestry who were 5 to 8 years old spoke English only or very well, but over 83 percent of 13 to 17 year olds spoke it only or very well, and just 4 percent spoke English not well or not at all. Natives of other recent immigrant groups have similar age patterns but slightly higher levels of English skills. For instance, the percentage of native adults who spoke English not well or not at all was 1.1 percent for Cubans, 3.6 percent for Central Americans, 2 percent for Chinese, and 0.9 percent for Filipinos. Thus even among ethnic groups who have come to the U.S. in large numbers within the last generation, lack of English fluency does not seem to be a significant problem for teenagers or adult natives.

Table 2 presents data on language skills for immigrants aged 18 to 64 in 1980 and 1990 by region of origin. Several patterns are worth mentioning. First, in both 1980 and 1990 the vast majority of immigrants of working age spoke English well or very well, though speaking ability declined slightly during the decade. In 1980 23 percent of immigrants aged 18 to 64 spoke English poorly or not at all, versus 25 percent in 1990. However, differences vary widely by region of origin. About 97 percent of immigrants from English speaking

Table 1
English Speaking Ability of U.S. Natives, 1990
by Age and Ethnicity
(percent)

	5-8	9-12	13-17	18-64
All Natives				
Only English	89.1	90.4	89.2	93.9
Very Well	6.2	6.8	8.0	4.7
Well	3.0	2.0	1.8	1.0
Not Well	1.6	0.8	0.9	0.4
Not at All	0.2	0.0	0.0	0.0
Mexicans				
Only English	36.4	36.6	33.9	30.7
Very Well	31.2	42.0	49.4	51.8
Well	20.5	16.4	12.7	13.0
Not Well	10.5	4.6	3.7	3.8
Not at All	1.4	0.4	0.3	0.7
Cubans				
Only English	26.1	18.8	21.6	30.1
Very Well	48.5	63.0	70.3	64.0
Well	14.1	14.1	6.6	4.8
Not Well	9.6	4.1	1.5	1.1
Not at All	1.7	0.0	0.0	0.0
Central Americans				
Only English	18.8	17.7	21.4	43.1
Very Well	41.3	61.4	69.4	46.2
Well	25.3	17.2	8.8	7.1
Not Well	12.6	3.1	0.4	3.4
Not at All	2.0	0.6	0.0	0.2
Chinese				
Only English	33.2	43.3	42.4	64.4
Very Well	32.7	40.3	47.7	29.7
Well	23.0	12.9	6.7	3.9
Not Well	10.6	3.5	3.2	1.9
Not at All	0.5	0.0	0.0	0.1
Filipinos				
Only English	74.1	82.7	82.6	84.6
Very Well	18.6	13.4	14.4	12.0
Well	4.2	2.2	2.1	2.5
Not Well	3.1	1.7	0.9	0.9
Not at All	0.0	0.0	0.0	0.0

Table 2
English Speaking Ability
Immigrants by Region of Birth

	English speaking	Europe	Mexico	Other W. Hemisphere	Africa and Mid East	South Asia	East Asia	Total
1980								
English Ability								
Only English	88.2	24.7	2.6	6.1	13.4	11.7	8.5	25.1
Very Well	8.7	38.2	19.7	32.8	45.0	62.3	34.6	29.5
Well	2.5	23.8	22.9	28.8	30.2	19.1	34.4	22.7
Not Well	0.5	10.6	30.5	21.6	9.2	5.7	17.9	15.1
Not at All	0.1	2.7	24.3	10.7	2.2	1.2	4.6	7.6
Education	12.7	11.5	7.4	11.7	13.8	15.8	12.9	11.3
YSM	18.8	20.2	12.4	12.5	9.2	7.9	9.6	16.6
Female (%)	58.6	53.3	46.2	53.7	37.8	42.4	56.3	52.4
Percent	16.6	27.3	18.3	14.4	5.0	2.4	16.0	100.0
1990								
English Ability								
Only English	88.4	32.8	4.0	7.6	16.3	10.7	11.6	22.6
Very Well	8.9	38.5	23.5	34.1	52.9	61.4	35.8	31.5
Well	2.1	18.5	22.0	25.1	22.2	19.9	30.6	21.2
Not Well	0.6	8.5	30.8	22.6	7.2	6.5	18.2	17.0
Not at All	0.0	1.7	19.7	10.6	1.4	1.5	3.8	7.7
Education	13.0	12.3	8.0	11.0	13.6	14.7	12.7	11.4
YSM	19.7	23.6	13.6	14.2	12.9	11.0	12.4	15.8
Female (%)	54.7	52.2	43.7	50.9	39.8	43.8	53.8	49.7
Percent	12.5	17.1	22.9	18.2	5.2	3.4	20.7	100.0

Note: Data are for immigrants 18-64.

countries (including Canada and other countries with other official languages) spoke English only or very well. Immigrants from South Asia, continental Europe, and Africa and the Middle East also had good English skills, with only a small fraction speaking English not well or not at all. However, more than half of Mexican immigrants were in these bottom two categories, as were about one third of immigrants from other western hemisphere countries and about one fifth from East Asia. The level of language skills by region of origin changed little between 1980 and 1990.

Table 2 also presents average education, years since migration, and percent female by region of origin for 1980 and 1990. Immigrant groups with the lowest English skills were also the groups with the least education. For instance, Mexican immigrants reported an average of 8 years of schooling in 1990, compared with 11 years for other western hemisphere immigrants, 12.3 years for continental Europeans, and 14.7 years for South Asians. Immigrant groups also differ in the average number of years since they entered the U.S. The shift from European to Latin American and Asian immigration is reflected in YSM averages, which are lower for the groups whose shares have increased in recent decades. In 1990 continental Europeans had been in the U.S. the longest, 23.6 years, and came when they were youngest, 18.7 years old. At the other extreme, South Asians on average immigrated 11 years before the 1990 Census, and came when they were 25.9 years old. More than half the immigrants from English speaking countries, Europe, other western hemisphere countries, and East Asia were women, compared to only 40 to 44 percent of immigrants from Africa and the Middle East, Mexico, and South Asia.

Many immigrants come to the U.S. speaking English, or learn to speak it soon after

arrival, but average skill varies widely by length of residence in the U.S. as well as by region of origin. As Table 3 indicates, among immigrants who entered the U.S. between 1987 and the Census survey date of April 1990, 45 percent reported speaking English not well or not at all. However, the share of recent immigrants with weak English skills was 16 percent of South Asians, 24 percent of Africans and Middle Easterners, 32 percent of continental Europeans, 39 percent of East Asians, 59 percent of other western hemisphere, and 74 percent of Mexicans.

Census data do not provide information on the exact date of entry into the U.S., or on language skills that immigrants had on arrival. To the extent that immigrants rapidly improve their ability to speak English during their first one or two years in the U.S., the percentages in Table 3 overstate language skills at entry, and results discussed below on the effect of living in the U.S. will understate skills acquired after migrating. Chiswick (1991) reports that among a group of illegal Mexican aliens whose average education was only 7 years, the percentage speaking no English fell from 80 percent at time of first entry to the U.S. to 41 percent at the time of the interview, which typically occurred within 2 years of first entry. Similar improvements in English skills occurred among men from other, mostly Latin American, countries.

Table 3 shows a clear pattern of improvement over time, especially during the first several years after entering the U.S. For instance, the percentage of South Asians who speak English only or very well increases from 57 percent among immigrants during 1987-90 to 64 percent for the 1985-86 cohort, to 70 percent for the 1980-84 cohort, etc. There is a similar improvement among groups who enter with weaker skills. The percentage of Mexicans who

Table 3
English Speaking Ability of Immigrants, 1990
by Region of Origin and Period of Arrival
(percent)

	pre-1950	1950s	1960s	1970s	1980-4	1985-6	1987-90	Total
English Speaking Countries								
Only English	91	90	88	90	88	85	84	88
Very Well	7	8	9	8	9	11	11	9
Well	2	1	2	1	2	3	4	2
Not Well	0	1	1	1	1	1	1	1
Not at All	0	0	0	0	0	0	0	0
Europe								
Only English	51	45	39	23	13	12	11	33
Very Well	36	39	38	41	44	40	32	38
Well	11	14	16	22	27	26	25	18
Not Well	2	2	6	12	13	19	23	9
Not at All	0	0	1	2	3	3	9	2
Mexico								
Only English	11	8	7	4	4	4	4	4
Very Well	44	40	36	28	18	12	11	23
Well	21	26	26	26	22	17	11	22
Not Well	17	19	23	30	36	37	32	31
Not at All	7	7	10	12	20	30	42	20
Other W. Hem.								
Only English	53	26	11	7	4	3	4	8
Very Well	31	45	49	40	28	22	18	34
Well	8	20	23	27	29	26	19	25
Not Well	7	8	13	20	28	32	32	22
Not at All	1	1	4	6	11	17	27	11
Africa & Middle East								
Only English	51	46	36	15	12	10	7	16
Very Well	41	40	45	60	59	51	40	53
Well	6	11	15	20	22	29	30	22
Not Well	2	3	4	5	6	8	18	7
Not at All	0	0	0	0	1	2	5	2
South Asia								
Only English	87	54	30	10	9	5	5	11
Very Well	3	38	58	70	61	59	52	61
Well	0	7	9	16	23	23	27	20
Not Well	10	1	3	3	6	11	13	6
Not at All	0	0	0	1	1	2	3	2
East Asia								
Only English	46	48	34	13	3	4	3	12
Very Well	29	31	41	44	34	32	23	36
Well	15	14	18	29	36	32	35	30
Not Well	10	7	6	12	22	26	31	18
Not at All	0	0	1	2	5	6	8	4
TOTAL								
Only English	57	48	35	19	12	11	12	23
Very Well	28	33	36	37	30	26	22	31
Well	10	13	17	23	26	23	21	21
Not Well	4	5	9	16	23	26	26	17
Not at All	1	1	3	5	9	14	19	8

speak no English falls from 42 percent in the 1987-90 cohort, to 30 percent for the 1985-86 cohort, to 20 percent for the 1980-84 cohort, etc. However, even after 30 or more years in the U.S., over a quarter of Mexican immigrants in 1990 spoke English not well or not at all.

V. Regression Results

Interpreting the changes in language skills shown in Table 3 as assimilation effects, ie, as the result of time spent in the U.S., implicitly assumes that cohort effects are zero. However, as Borjas (1985) and others have observed in connection with earnings regressions, English skills may appear to rise over time in cross section data in part because earlier immigrants spoke better English than recent immigrants at the time of entry, not because their skills have improved with length of stay.

Previous studies of language skills which used cross section data, including Chiswick and Miller (1994), (1995), and Dustmann (1994), had no choice but to assume that there are no cohort effects. In cross section data

$$T - YRM = YSM \tag{5}$$

The year in which the data were collected (T) minus the year of arrival in the U.S. (YRM) always equals the number of years since arrival (YSM) when T is the same for all observations. Thus it is not possible to tell whether English skill improves as YSM increases, or whether it has increased over time because the skills of recent immigrants (large YRM) are lower than the skills of earlier immigrants. Only when T varies, as in pooled

data, can the effects of YRM and YSM be separately identified.

To estimate assimilation and cohort effects, I began by running the following equation on pooled data from the 1980 and 1990 Censuses of Population.

$$(6) \qquad L_{it} = \alpha_0 + \alpha_1 YRM_i + \alpha_2 YSM_{it} + \alpha_3 YSMSQ_{it} + \alpha_4 AGEM_i + \alpha_5 AGEMSQ_i + \epsilon_{it}$$

The sample was restricted to immigrants from non-English speaking countries who were 18-64 at the time of the census. Immigrants who arrived before 1950 (30 or 40 years prior to the survey dates) were excluded from the regression sample because YSM and YRM were poorly identified in this open-ended category. The independent variables included year of entry (YRM), years since entry (YSM) and its square (YSMSQ).

In addition, age at entry (AGEM) and its square (AGEMSQ) were included, since limiting the sample to working age individuals excludes immigrants who arrived a long time ago at a relatively old age. For instance, immigrants under 65 in 1990 who arrived before 1960 must have entered when they were under 35, but immigrants in the sample who arrived during the 1980s could be as old as 64. Since English skills fall with AGEM, failing to control for it will result in a spurious negative estimate of the trend in cohort effects.

Equation (6) constrains trends in the English skills of entering immigrants and their subsequent assimilation rates to be the same across source countries, sexes, and education groups. It also assumes that the learning process after migration has not changed over time, eg, that immigrants were just as likely to improve their English skills in the 1980s as in the

1970s. These assumptions will be relaxed below.

Two measures of English speaking skills were used. The first measure (Fluent), used by Chiswick (1991) and Chiswick and Miller (1995), equals one for respondents who report speaking only English or speaking it very well, and zero for others. The second (Lang) measures language skill on a five point scale similar to a grade point average, with 4 for speaking only English and 0 for speaking no English. "English only" has a higher score than "very well" because the English skills of people who say they speak only English seem to be higher than of people who say they speak it very well. Carliner (1995) reports that in earnings regressions estimated for a sample of native-born workers, "English only" speakers earn more than "very well" speakers who are otherwise similar.

Table 4 presents coefficients, standard errors, and derivatives estimated at the means of the independent variables. Fluent coefficients were estimated by logit and Lang coefficients were estimated by OLS. Coefficients from an ordered logit regression on Lang are also shown. Logit estimation of the probability of speaking only English or speaking it very well has the advantage of easy interpretation and comparability with most other studies. However, it does not use all the information about language skills provided in the Census data. For instance, it ignores differences between immigrants who speak no English and those who speak it poorly. OLS regressions with language score as the dependent variable use this additional information but provide results with less intuitive meaning. Moreover, the results may be sensitive to the use of cardinal scores for an ordinal variable. Assigning different scores for the five categories could conceivably yield different regression estimates. Ordered logit regressions with language score as the dependent variable solve this problem

Table 4
Language Regressions for Immigrants
Cohort and Assimilation Effects

	Fluent (logit)		Lang (OLS)		Lang (ordered logit)
	Coefficients	Derivatives	Coefficients	Derivatives	Coefficients
Constant	11.386		16.649		
	(2.127)		(1.009)		
YRM	-.005	-.001	-.007	-.007	-.010
	(.001)		(.001)		(.001)
YSM	.053	.011	.037	-.026	.053
	(.002)		(.001)		(.002)
YSMsq/100	-.038		-.040		-.029
	(.005)		(.000)		(.004)
Agem	-.113	-.014	-.046	-.029	-.001
	(.002)		(.001)		(.001)
Agemsq/100	.127		.038		.097
	(.003)		(.001)		(.002)
R^2	.12		.20		.08
NOBS	187,936		187,936		187,936

Note: Coefficients and standard errors are shown in columns 1, 3, and 5. Derivatives calculated at the sample means are shown in columns 2 and 4. See text for an explanation of the dependent and independent variables. The R^2s in the logit columns are pseudo R^2s.

but are even more difficult to interpret. Dustmann (1994) uses this technique but discusses the results for only one of his three ordered categories.

Fortunately, regressions estimated by all three approaches show similar results. Cohort effects are negative and significant but very small. Arriving a year later lowers the probability of being fluent by 0.1 percent and lowers an immigrant's score by 0.007 points. The cumulative trend in fluency over 30 years is thus a decline of about 3 percentage points in fluency and 0.2 points in English scores. The coefficient on YRM in the ordered logit regression is also negative, significant, and small.

To see if there are cohort effects which are not captured by this simple trend term, I also ran regressions with dummy variables identifying immigrants who entered the U.S.during the 1950s, 1960-64, 1965-69, 1970-74, 1975-79, 1980-81, 1982-84, 1985-86, and 1987-90. The coefficients on these variables indicate that the probability of fluency fell by about 4 percentage points from the pre-1965 period to 1965-80. During 1980-81, the probability of fluency fell by over 5 additional percentage points, perhaps as a result of the Mariel boatlift and the admission of refugees from Southeast Asia (see below for further discussion of this point). Immigrants who arrived after 1982 had average fluency above the levels of the 1965-74 cohort, though not as high as pre-1965 immigrants.

The assimilation effects implied by the YSM coefficients shown in Table 4 are considerably larger in absolute value than these cohort effects. The estimated difference in the probability of speaking English fluently between an immigrant with 5 years of residency and one with 15 years in the U.S. is 11 percentage points. The estimated difference between a newly arrived immigrant and a 20-year resident is 21 percentage points. As Table 4

indicates, the marginal effect of an additional year of residence in the U.S. is to increase the probability of fluency by 1.1 percentage points and to increase the language score by .026 points.

To see if shifts in the characteristics of immigrants since the 1950s have affected the amount of English learned after arrival, in addition to affecting skill level on entry, I ran an additional logit regression with an interaction term, YRM*YSM. The estimated coefficients imply that the probability of fluency at entry has increased by 0.2 percentage points annually but the trend assimilation effect has been falling. For instance the estimates imply that the increase in the probability of fluency from an additional year of residence fell by 0.005 percentage points from 1965 to 1985. This amounts to over 40 percent of the assimilation effect experienced by the cohort of immigrants who arrived in 1965, a substantial decline.

These estimates of assimilation and cohort effects take no account of shifts over time in country of origin or education levels of immigrants. However, the decrease in the percentage of immigrants from Europe and the increase from Latin America and Asia over the past 25 years is well known Borjas (1995) has documented improvements in the average education of immigrants between 1970 and 1990. There have also been smaller changes in the ratio of male to female immigrants over time. As discussed above in section II, all these changes may have affected immigrants' English skills. To examine their effects, I added education, its square, and dummy variables for women and for region of origin to the regressions discussed above. Coefficients, standard errors, and derivatives taken at the means of the sample are shown in Table 5.

Adding education, sex, and place of birth variables to the language regressions

Table 5
Language Regressions for Immigrants
Full Model

	Fluent (logit)		Lang (OLS)		Lang (ordered logit)
	Coefficient	Derivatives	Coefficient	Derivatives	Coefficients
Constant	-26.476		-2.629		
	(2.440)		(.867)		
YRM	.013	.003	.002	.002	.005
	(.001)		(.000)		(.001)
YSM	.097	.016	.052	.031	.105
	(.003)		(.001)		(.002)
YSMsq/100	-.121		-.077		-.141
	(.006)		(.002)		(.005)
Agem	-.134	-.017	-.041	-.027	-.114
	(.002)		(.001)		(.001)
Agemsq/100	.148		.032		.105
	(.003)		(.001)		(.002)
Ed	.038	.049	.063	.086	.138
	(.005)		(.001)		(.003)
Edsq/100	.709		.104		.215
	(.024)		(.007)		(.016)
Female	.063	.016	-.042	-.042	-.107
	(.011)		(.004)		(.009)
Mexico	-1.273	-.316	-.726	-.726	-1.619
	(.019)		(.007)		(.015)
Other W Hem.	-.829	-.206	-.467	-.467	-1.058
	(.018)		(.006)		(.014)
Africa and Middle East	.106	.026	.032	.032	.008
	(.026)		(.009)		(.020)
South Asia	.374	.093	.058	.058	.067
	(.033)		(.012)		(.025)
East Asia	-.694	-.172	-.241	-.241	-.579
	(.018)		(.006)		(.014)
R^2	.27		.44		.20

Note: European immigrants from nonEnglish speaking countries were the reference group in the regressions, which did not include immigrants from English speaking countries. Coefficients and standard errors are shown in columns 1,3, and 5. Derivatives calculated at the sample means are shown in columns 2 and 4. See text for an explanation of the dependent and independent variables. The R^2s in the logit columns are pseudo R^2s.

increases the estimated assimilation effect slightly. The marginal effect on fluency, estimated in the logit regression, increases from 1.1 percentage points to 1.6 percentage points. This estimate is also slightly larger than Chiswick and Miller's (1992) estimate from a regression which included several personal characteristics but not year of arrival. The marginal effect on language score, estimated in the OLS regression, rises from .026 points to .030 points. On the other hand, the estimated cohort effect reverses sign. In particular, the shift in country of origin from Europe to Latin America and East Asia fully accounts for the small trend decline in English skills of entering immigrants during the past 30 years. In fact, once shifts over time in education, sex, and country of origin are taken into account, more recent immigrants have higher language skills than earlier ones. This finding could be the result ofthe spread of English as an international language since the 1950s, including both its use in television and movie theaters of source countries and in their schools.

The coefficients presented in Table 5 also indicate that education has a large and significant effect on language skills. At the mean of the sample (11.1 years of schooling), an additional year of education is estimated to increase the probability of fluency by 4.9 percentage points. Other characteristics equal, a high school graduate is about 20 percentage points more likely to be fluent than an immigrant with only a grade school education. The difference between a college graduate and a high school graduate is about 24 percentage points. Similarly, the difference in language scores between a high school graduate and a grade school graduate is 0.34 points. These estimates are somewhat larger than those reported in previous studies such as Chiswick and Miller (1992) and (1995) which used cross section data.

To see if education affects both entering skill level and assimilation, I ran additional regressions with an interaction term Ed*YSM. The results imply that when immigrants arrive in the U.S., each additional year of education is associated with an additional 1.5 percentage points in the probability of fluency. In addition, education significantly increases the amount of English which immigrants learn with time in the U.S. Each added year of residence increases the probability of fluency by 1.2 percentage points for immigrants with 8 years of school, but by 2.0 percentage points for college graduates.

Arriving in the U.S. at a young age also has a large impact on English skills, other things equal. An immigrant who arrives at age 15 is 25 percentage points more likely to speak English fluently than one who arrives at age 30. For immigrants who arrive in their 20s, each additional year of age decreases the probability of speaking English very well or only by about 1.7 percentage points and lowers English scores by 0.03 points. Adjusting for other differences, women are 1.6 percentage points more likely to be fluent than men. These differences are discussed below in more detail.

Finally, large differences in English skills by place of birth remain, even after taking into account differences in education, sex, years in the U.S., year of entry, and age at entry. Other things equal, Mexican immigrants are 32 percentage points less likely than continental Europeans (the reference group), to speak English fluently. Immigrants from other western hemisphere nonEnglish speaking countries are about 21 percentage points less likely to be fluent, and East Asians are 17 percentage points less likely. However, immigrants from Africa and the Middle East, and especially from South Asia, are more likely than otherwise similar continental Europeans to speak only English or to speak it well. Differences in

language scores by country of origin show a similar pattern.

These results do not support the hypothesis that linguistic distance is as important as other factors in explaining English skills, since Spanish is closer to English than Arabic or Chinese, and other European languages are closer than Hindi or Urdu. The results also do not support the hypothesis that immigrants from poor countries have lower skills than immigrants from richer countries. Immigrants from South Asia, Africa, and the Middle East speak better English than immigrants from Europe. The results do lend support to the hypothesis that nearby countries send immigrants with weaker language skills than more distant countries. Only the most skilled people in poor and distant countries with weak links to the U.S. (South Asia, Africa, and the Middle East) have the money and information to immigrate, but people in neighboring countries with efficient information flows often choose to immigrate even if they have only weak English skills.

To examine the large residual differences by country of origin further, I ran two sets of logit regressions for the six regional groups, one with variables for cohort and assimilation effects and age at entry, the second with all the variables shown in Table 5. Derivatives evaluated at the means of the independent variables for the entire sample (not at the means for each regression sample) for the first set of regressions are shown at the top of Table 6. They indicate that simple assimilation effects, not holding education and percent female constant, are remarkably similar for the different country groups, and similar to the results discussed above for the entire sample. In particular, these results suggest that economic immigrants who might expect to return home (Mexicans) do not learn less English with increased U.S. residence than political refugees (Cubans and Central Americans in the Other

Table 6
Effects on English Fluency
Immigrants by Region of Birth
(Derivatives from Logit Regressions)

	Europe	Mexico	Other W. Hemisphere	Africa and Middle East	South Asia	East Asia
Cohort and Assimilation Effects						
YRM	.005	.006	-.001	.007	-.007	-.003
YSM	.013	.014	.011	.018	.010	.012
Agem	-.020	-.016	-.021	-.013	-.001	-.015
Full Model						
YRM	.003	.006	.004	.009	.000	.000
YSM	.017	.015	.015	.021	.013	.015
Agem	-.018	-.012	-.020	-.014	-.012	-.016
Ed	.059	.046	.049	.052	.053	.045
Female	.047	-.006	-.020	-.035	-.085	.060

Note: Estimates of the marginal effect on the probability of being fluent, evaluated at the means of the sample, from logit regressions. All coefficients in the simple model were significant at the one percent level except for YRM for Other Western Hemisphere and Agem for South Asia. All coefficients in the full model were significant at the one percent level except for YRM for South and East Asia.

Western Hemisphere category).

Although assimilation effects are similar across source countries, the trend cohort effects are not. For continental Europeans, Mexicans, and Africans and Middle Easterners, there has been a trend improvement in English skills, but for immigrants from other western hemisphere countries, South Asia, and East Asia, the trend has been down. Results from regressions which included dummy variables for period of immigration suggest that a linear trend is a fair approximation for all the groups except immigrants from other western hemisphere countries and East Asia. Both these groups have very large negative coefficients for 1980-82, perhaps reflecting the large influx of refugees from Cuba and Southeast Asia who arrived during that period. For other western hemisphere immigrants, cohort effectsduring 1983-84 recovered somewhat, and then fell again during 1985-90. For East Asian immigrants, coefficients on period dummies rose after 1980-81 and did not fall during the rest of the decade.

The marginal effects for the second set of regressions are shown at the bottom of Table 6. Adjusting for education and other factors, trend cohort effects are positive for all six country of origin groups, though very small for East and South Asians. The marginal effect of an additional year of residence in the U.S, other characteristics constant, ranges from 1.3 percentage points for South Asians to 2.1 points for Africans and Middle Easterners. Similarly, the effects of additional years of education and of age at arrival are roughly similar for the six groups. However, male-female differences vary widely among the regional groups. East Asian women are 6 percentage points more likely to be fluent in English than otherwise similar East Asian men, and European women have a 4.7 point

advantage. In contrast, South Asian women are 8.5 percentage points less likely to be fluent than similar men. Women from the other three regional groups are slightly less likely to be fluent than comparable men.

To explore male-female differences further, Table 7 presents averages for year of arrival, years since arrival, age at arrival, education, and region of origin by sex for observations from the 1990 Census. Derivatives from logit regressions with Fluent as the dependent variable for the pooled 1980-1990 sample are also shown. The average woman in the 1990 sample arrived about one year earlier than the average man, has been in the U.S. one year longer, and arrived when she was 6 months older. (All of these differences may be due at least in part to higher mortality among men.) Women in this sample also have on average 0.3 years less of schooling. Immigrant women are also less likely to be from Mexico, African and the Middle East, and South Asia, and more likely to be from East Asia, continental Europe, and other western hemisphere countries than men.

The regression results indicate that cohort, asimilation, education, and age effects are very similar for the two sexes. The marginal effects of arriving in the U.S. a year later, or a year older, or residing here a year longer are almost identical for men and women. An additional year of schooling raises the probability of fluency by 5.1 percentage points for women and 4.7 points for men. However, several of the estimated effects for country of origin are quite different for men and women. Relative to immigrants from continental Europe (the reference group in the regressions), men from Mexico, other western hemisphere countries, Africa and the Middle East, and especially from South Asia, are considerably more likely to be fluent than women from the same region. Only among East Asian

Table 7
English Fluency, Full Model
Immigrants by Sex
(Derivatives from Logit Regressions)

	Men		Women	
	1990 Means	Derivatives	1990 Means	Derivatives
Fluent (%)	46.9		47.3	
YRM	1976	.004	1975	.003
YSM	14.0	.016	15.0	.016
Agem	21.7	-.017	22.2	-.016
Ed	11.3	.047	11.0	.051
Engl. sp.	10.9%	*	13.2%	*
European	15.4	R	17.0	R
Mexico	26.0	-.297	20.5	-.337
Other W. Hemisphere	18.1	-.171	19.0	-.239
Africa and Middle East	6.3	.065	4.3	-.021
South Asia	3.9	.165	3.0	.016
East Asia	19.3	-.182	23.0	-.165

Note: Estimates of the marginal effect on the probability of being fluent, evaluated at the mean of the sample, from logit regressions. All coefficients were statistically significant at the one percent level. European immigrants from nonEnglish speaking countries were the reference group in the regressions, which did not include immigrants from English speaking countries.

immigrants do women have higher fluency than men.

VI. Conclusion

This paper uses data from the 1980 and 1990 U.S. Censuses of Population to examine the English language skills of natives and immigrants. The main findings are first, that lack of fluency in spoken English is rare among native-born Americans. In 1990 98.6 percent of natives aged 18 to 64 reported speaking only English or speaking it very well. Only 0.4 percent reported speaking English not well or not at all. Among native-born children of ethnic groups who have come to the U.S. in large numbers during the past 30 years, such as Hispanics and East Asians, a substantial fraction were not fluent when they entered grade school, but at most 3 to 5 percent of teenagers and adults in these groups reported speaking English poorly or not at all.

Second, the vast majority of immigrants speak English well. In 1990 only 25 percent of immigrants reported speaking English poorly or not at all, though more than half of Mexicans and one third of immigrants from other nonEnglish speaking western hemisphere countries could not speak good English. Among immigrants who came to the U.S. during the three years prior to the Census survey, 45 percent spoke English poorly or not at all, though again, the percentage varied widely by country of origin. Although English skills improve with length of residence, after 30 or more years in the U.S., over a quarter of Mexican immigrants spoke English poorly or not at all.

Third, since the 1950s there has been a trend decrease in the probability of fluency

(speaking only English or speaking it very well) among new immigrants of about 0.1 percentage point per year, caused by the shift from European immigrants with strong English skills to Latin American and East Asian immigrants who arrive speaking less English. Within ethnic groups, there has been a trend increase in fluency among continental Europeans, Mexicans, and Africans and Middle Easterners, and a trend decrease among South and East Asians. On average, each additional year of residence in the U.S. increases the probability of fluency by 1.1 percentage points.

Fourth, education, age at arrival, sex, and region of origin also affect English skills. An additional year of schooling increases the probability of fluency by about 5 percentage points. For immigrants in their 20s each additional year of age at arrival lowers the probability by about 1.7 percentage points. Overall, women are slightly more likely to be fluent than men, especially East Asian and European women, but women from South Asia, Africa, the Middle East, and Latin America are less likely. Men and women improve their English skills at the same rate with increased residence in the U.S. Even after controlling for differences in education, years since arrival, and other factor, large differences in English skills by region of origin remain. These differences seem to be more associated with geographic distance from the U.S. than with the source country's per capita income or linguistic distance from English.

References

Bartel, Ann, "Where Do the New U.S. Immigrants Live?" *Journal of Labor Economics* 7:4, October 1989.

Borjas, George, "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants," *Journal of Labor Economics* 3:4, October 1985.

--------, "Self-Selection and the Earnings of Immigrants," *American Economic Review* 77:4, September 1987.

--------, "Ethnicity, Neighborhoods, and Human Capital Externalities," Working Paper No. 4912, National Bureau of Economic Research, November 1994a.

--------, "The Economics of Immigration," *Journal of Economic Literature* 32:4, December 1994b.

--------, "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?" *Journal of Labor Economics* 13:2, April 1995.

Boyd, Monica, "Gender Issues in Immigration and Language Fluency." In *Immigration, Language and Ethnic Issues: Canada and the United States*, ed. Barry Chiswick, Washington, D.C.: American Enterprise Institute, 1992.

Carliner, Geoffrey, "Wage Differences by Language Group and the Market for Language Skills in Canada," *Journal of Human Resources* 16:3, Summer 1981.

--------, "Wages and Language Skills Among U.S. Immigrants," (mimeo) 1995.

Chiswick, Barry, "Speaking, Reading, and Earnings among Low-skilled Immigrants," *Journal of Labor Economics* 9:2, April 1991.

Chiswick, Barry and Paul Miller, "Language in the Labor Market: The Immigrant Experience in Canada and the United States." In *Immigration, Language and Ethnic Issues: Canada and the United States,* ed. Barry Chiswick, Washington, D.C.: American Enterprise Institute, 1992.

--------, "Language Choice among Immigrants in a Multilingual Destination," *Journal of Population Economics* 7:2, 1994.

--------, "The Endogeneity between Language and Earnings: International Analysis," *Journal of Labor Economics* 13:2, April 1995.

Dustmann, Christian, "Speaking Fluency, Writing Fluency and Earnings of Migrants," *Journal of Population Economics* 7:2, 1994.

Kossoudji, Sherrie, "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men," *Journal of Labor Economics* 6:2, April 1988.

Lazear, Edward, "Culture and Language," NBER Working Paper No. 52xx, June 1995.

McManus, Walter, "Labor Market Effects of Ethnic Enclaves," *Journal of Human Resources* 25:2, Spring 1990.

Portes, Alejandro and Richard Schauffler, "Language and the Second Generation: Bilingualism Yesterday and Today," *International Migration Review* 28:108, Winter 1994.

Robinson, Chris, "Language Choice: The Distribution of Language Skills and Earnings in a Dual Language Economy," *Research in Labor Economics* 9, 1988.

Veltman, Calvin, "Modelling the Language Shift Process of Hispanic Immigrants," *International Migration Review* 22:4, Winter 1988.