LISTENING TO WHAT THE WORLD SAYS: BILINGUALISM AND EARNINGS IN THE UNITED STATES

Albert Saiz and Elena Zoido*

Abstract—Is there a shortage of critical foreign language skills in the United States? Recent concerns about national security and economic globalization suggest increased demand and wage premia for foreign language speakers. The use of English as the international language, however, suggests a decrease in demand for foreign language skills in the United States. To address this question, we study a representative sample of U.S. college graduates. Ordinary least squares regressions with controls for cognitive ability, nonparametric methods based on the propensity score, and panel data methods suggest a 2%–3% wage premium for college graduates who can speak a second language.

I. Introduction

Is the knowledge of a foreign language important in today’s globalized economy? Is bilingualism bound to become less relevant for countries in which English, the de facto international common language, is spoken? Do labor markets reward bilingualism? These are important questions for those pondering in which skills their children should invest.

Anecdotal evidence seems to suggest that speaking a foreign language is a valuable skill. Learning a second language is a possibility many consider at some point. Language courses are an option or a requirement in elementary, secondary, and college education. Private foreign-language institutions, educational materials, and distance courses are regularly advertised in newspapers and magazines. Reports in the popular press suggest that speaking a second language has become a “desirable skill in the work force, one employers are willing to pay for.” In 2001, a U.S. Congress resolution argued that “research indicates that the United States is failing to graduate enough students with expertise in foreign languages, cultures and policies to fill the demands of business, government, and universities.” If these statements are accurate, one would expect substantial wage premia for those with the ability to speak foreign languages.

Recent concerns about national security also highlight the key role of knowledge of foreign languages. On September 20, 2001, only nine days after the attack on the World Trade Center, the New York Times reported that according to U.S. intelligence experts, “the government is short of seasoned analysts and new ones lack such essential skills as knowledge of foreign languages.” On October 10, 2001, an article published in the same newspaper argued that the deficiencies in the U.S. national security system had “been most obvious on the linguistic front.” These deficiencies had been denounced well in advance. The House Permanent Select Committee on Intelligence argued in 1997 that workers in the intelligence services were “often unable to speak foreign languages and unfamiliar with the countries they analyzed.”

Diana Schemo, a journalist at the New York Times, wrote on April 16, 2001, that “diplomatic and intelligence officials are warning of critical shortages in their ability to understand the languages of other nations, and so unravel their secrets. [· · ·] The reasons are many. With English increasingly becoming the world’s lingua franca, the study of foreign language has suffered.” The implication of this argument is that the returns to bilingualism in the U.S. labor market should be low or nil.

Several reasons suggest that the knowledge of a foreign language might represent a useful skill in the American labor market. First, American companies export and import products to and from the rest of the world. Second, offering a service in the language of the prospective customers may provide a competitive advantage for firms, especially in areas where the foreign-born tend to settle. Hence, knowledge of a second language may be a valuable asset for companies doing business abroad or for companies catering to immigrants or people of foreign background within the United States. At the same time, access to foreign media and literature may help foster innovation and the adoption of best practices from abroad, thus improving workers’ productivity. In addition, second-language skills are in demand from several government agencies that deal with foreign affairs: the diplomatic service, the CIA, military intelligence, and the like.

Furthermore, an extensive literature suggests that learning a second language may help people develop their cognitive and communicative abilities, and underlines the possible advantages of bilingualism in terms of intellectual and academic achievement. Cooper (1987) finds that SAT math and verbal scores increase with each additional year of foreign language study. Olsen and Brown (1992) show that “students who had completed a foreign language course in high school tended to have higher scores on the ACT exams in English and math regardless of their ability level.” Learning a foreign language may help develop analytic and interpretative capacities. Cook (1997) reports that “increased metalinguistic awareness of phonology, syntax, and the arbitrary nature of meaning, and gains in cognitive...
flexibility” are established outcomes from learning a second language. If speaking a second language is important for improving cognitive capabilities, we should find that the individuals who speak a second language are more productive and earn higher wages.

All these arguments suggest that there might exist wage premia associated with the knowledge of a second language, whether individuals actually use the second language on the job or not. Despite the importance of this question, little research has been undertaken on the topic. The literature on language acquisition and labor market outcomes has so far concentrated on the immigrants’ returns to learning the official language of their destination countries. In this paper we thus turn to the question of the expected labor market returns to bilingualism. What are, on average, the returns to speaking a second language?

We focus on the estimation of the returns to speaking a second language for native college graduates in the United States. Available data allow us to try to identify a causal treatment effect for this sample of individuals. We find that college graduates with conversational knowledge of a second language earn, on average, wages that are 2%–3% higher than those without. We include a complete set of controls for general ability, using information on grades and college admission tests in order to reduce concerns about omitted ability bias. We obtain similar results if we use nonparametric methods based on the propensity score and if we exploit the temporal variation in the knowledge of a second language. We thus conclude that the estimates are not driven by observable differences in the composition of the pools of bilinguals and monolinguals, by the linear functional form that we impose in OLS regressions, or by constant unobserved heterogeneity. Finally, we discuss instrumental variable (IV) estimates obtained using high school and college admission and graduation requirements as instruments that provide very imprecise estimates, but suggest a positive causal effect of foreign languages on earnings.

Our results contribute to the explanation of individual decisions on whether to learn a second language. Under the tenets of human capital theory, one should invest in the acquisition of a foreign language if the present value of the future returns for doing so exceeds the costs. Some of the returns from learning a second language consist of the direct consumption of services produced by the individual’s knowledge of the language. Speaking a second language while traveling abroad, asking an immigrant shopkeeper for a product in her native language, and relating to foreign friends are all examples of these. This paper cannot address them. The expected value of the labor market skills that learning a foreign language provides is the other important component of the returns to learning the language, and this is the focus of the paper. Our estimates of the effect of bilingualism on earnings are relatively small (2%–3%) and compare unfavorably with recent estimates on the returns to one extra year of general schooling (8%–14%), which may help explain the current second-language investment decisions of monolingual English-speakers in the United States.

The paper proceeds as follows. In section II we discuss the existing literature on the returns to learning a second language and show some suggestive evidence on the demand for foreign-language skills. Section III addresses the potential methodological problems of estimating the returns to a second language. In section IV we introduce the data sources and describe the characteristics of individuals speaking a second language in our sample. Section V presents the results of the ordinary least squares (OLS), propensity scores, panel data, and IV approaches. Section VI concludes.

II. Background: Second Language and Earnings

Economic analysis has been recently applied to improving our understanding of the individual decision to invest in a second language. Grenier and Vaillancourt (1983) were the first to identify foreign languages as an element in the human capital portfolio of individuals. They use the framework laid out by Becker (1964) and Mincer (1974) to describe how individual characteristics affect the gross costs, the gross returns, and the information on the value of the returns to this investment. Similarly, Ridler and Pons-Ridler (1984) analyze the decision to learn a second language as an investment decision. These authors emphasize the importance of second languages as consumption goods and discuss the economic costs of forced “francisation” in Quebec as an economic problem.

So far, the empirical literature on the returns to speaking a second language in the United States has concentrated on the incentives and wage premia that immigrants receive for mastering English. Carliner (2000) finds that most immigrants in the United States are proficient in English. He shows that each additional year of residence in the United States increases the probability of being proficient in English by 1.1 percentage points. Chiswick and Miller (1998a, 1998b) show that the probability of speaking English for immigrants in the United States responds to the costs and benefits of doing so. In particular, fluency rates are higher for those with more schooling, who immigrated at a younger age, who have lived in the United States longer, who live in areas with fewer origin-language speakers, and, among women, who have fewer and younger children. Fluency rates are also higher for those with less access to media in their mother tongue, with a lower probability of returning to the origin country, whose country of origin is geographically further from the United States, and whose origin language is linguistically closer to English. McManus, Gould, and Welch (1983), Kossoudji (1988), Rivera-Batiz (1991), and Chiswick and Miller (1997) study the labor market returns to English acquisition among immigrants in the United States. They find that immigrants whose native language is not English but who are fluent in it earn higher
wages and are more likely to work in high-wage positions than other immigrants. Mora and Dávila (1998) find that the English-language premium for immigrants in the U.S. labor market is bigger for males. Gonzalez (2000) establishes that the returns on oral proficiency are greater than the returns on literacy skills, and that writing skills are more valuable than reading skills.

Similar wage premia are observed in other countries. Chiswick (1997), for example, analyzes the wage differential for immigrants who speak Hebrew in Israel. Chiswick and Miller (2000) find that immigrants in Canada who can speak one of the official languages earn wages that are 10% to 12% higher. They find that people who learn an official language have higher earnings than those who do not, for reasons other than their language skills, implying a positive association between the propensity to learn a second language and other determinants of earnings. For the case of Germany, Dustmann and van Soest (2001) also find a positive effect of fluency in German on earnings.4

Despite the extensive literature on the cognitive and developmental effects of bilingualism, there is little research on the returns to speaking a second language for nonimmigrants. Saiz and Zoido (2002) show evidence of a positive association between using a second language at work and higher earnings in the European Union. This association is present in such English-speaking countries as Ireland and the United Kingdom. In the Canadian context, Shapiro and Stelcner (1997), Pendakur and Pendakur (1998), and Christofides and Swidinsky (1998) find that speaking a nonofficial language is associated with a negative earnings differential.

The interpretation of these results is subject to several problems. In particular, selection into learning a language different from the official ones may affect these estimates dramatically. For example, if immigrants speak other languages in a much greater proportion than the general population, it is difficult to disentangle the effects of assimilation from the pure effects of speaking more languages. This finding points to the fact that selection into learning a second language can be correlated with individual attributes that decrease earnings, which suggests the importance of controlling for potentially confounding factors.

An important and motivating finding for us is the fact that taking foreign-language courses in high school is associated with higher wages and wage growth in the United States. Altonji (1995) estimates that an additional year of foreign-language courses leads to a wage increase of approximately 1.7%. Indeed, Altonji’s results suggest higher marginal returns for foreign-language courses than for courses in mathematics, science, and English. Because foreign-language skills in the United States are perceived to be of little importance, these results were unexpected by the author. One possibility is that the results are driven by the correlation between taking second-language courses and omitted variables, such as high school quality or family background, or by selection of those individuals with higher returns to speaking a second language into foreign-language classes. One could argue, for example, that foreign-language courses are more sensitive to the tastes for education of the families whose children attend a school than math, science, or English language courses are. Altonji’s study controls for a wealth of socioeconomic variables and includes high-school effects, but other relevant earnings determinants may remain omitted. A second possibility, one that we are interested in exploring, is that speaking a second language (as opposed to simply having taken foreign-language courses) is a valuable skill in the labor market or contributes to the development of other work-related skills.

Lopez (1999) studies the wage effect of speaking a second language in the United States. He uses observations from the National Adult Literacy Survey (NALS) of 1992. The main shortcoming of this data set is that only individuals who belong to a language minority were questioned regarding their language skills. English speakers who have subsequently learned a foreign language, the main treatment group of interest for us, are excluded from the questionnaire on second languages. Thus, Lopez’s basic test is whether language retention among language minorities is associated with higher or lower labor market earnings. Using OLS regressions, Lopez finds that those who speak a second language proficiently earn wages that are 14.2% higher than those earned by those in language minorities who do not speak the minority language well. He also finds that minority-language people who speak their mother tongue well earn 13.5% more than English monolinguals, but this difference is not statistically significant.

In this paper, we focus on the ex ante (expected) returns to speaking a second language for a sample of college graduates, for which such data are available. Whereas advantages of speaking a second language may exist for those who do not use their language skills at work, a number of jobs require direct knowledge of foreign languages. To obtain a sense of the order of magnitude of foreign language demand, we performed a search in two major Internet job-search Web sites: Monster.com and careerbuilder.com (details in appendix table A1). We searched for all jobs that required a bachelor’s degree (the population that we will study in the paper). Given the large number of hits, we constrained the sample to the jobs posted in the last 24 hours. We found 98 postings that included a foreign language as a requirement, after making sure that we did not duplicate the postings, as some of them mentioned two languages or more. These represented 2.62% of the positions for college graduates in the United States. If one accepts the representativeness of these samples, this would be a lower bound for usage, for companies that do not require second-language skills may find the option value

4 These authors argue that omitted ability biases OLS estimates downward, whereas measurement error introduces the opposite bias. For the immigrants in the German Socioeconomic Panel Sample, these authors find the latter downward bias to be more important quantitatively.
desirable when considering alternative resumes. In as much as only 34% of college graduates claim to speak a second language, a nontrivial 7.72% of our foreign-language sample (2.62/34) can expect to derive a direct labor market advantage from their second-language skills.

III. Empirical Strategy

Measuring the returns to learning a foreign language presents considerable methodological challenges, in common with the work on the returns to schooling. Think of running a Mincerian regression like the following:

\[
\log w_i = \alpha + \beta X_i + \gamma SL_i + \epsilon_i,
\]

where \(\log w\) is the log of earnings, \(X\) is a vector of personal characteristics, and \(SL\) is an indicator of whether the individual \(i\) speaks a second language.

Obviously, speaking a second language is not randomly assigned in the population. Individuals choose whether to learn a foreign language according to their potential earnings. Thus, a selection problem arises. Consider the distribution of the benefits of speaking a second language among the population. Assume that the costs of learning it are the same for all individuals. Then, those individuals with higher returns to speaking a second language will have greater incentives to invest in learning it. Empirically, we may find that the individuals who speak a second language obtain higher earnings. If the second-language skills were randomly assigned, however, the average returns to speaking it would be smaller. Selection by earnings will thus bias conventional estimates upward. This problem is akin to the selection problem in the college attendance decision (Willis & Rosen, 1979).

Additionally, the costs of learning a second language may be smaller for individuals with higher cognitive competence. If this is the case, more able individuals will be more likely to learn foreign languages. If we do not correct for this differential selection with respect to ability, our estimates will contain an ability bias, as the coefficient on the foreign language variable will capture part of the effect of ability on earnings (Griliches, 1977). A similar problem has been pointed out by Borjas (1995) in the context of the returns to speaking English for immigrants in the United States. Again, this causes the estimates of the returns to a second language to be biased upward.

We also need to take into account that there may be other unobservable characteristics of the individuals that simultaneously determine earnings and whether an individual speaks a second language. People who study foreign languages may have different preferences regarding the types of jobs they like and the kind of studies they choose. Consider the population of college graduates. People with humanities majors study foreign languages with higher likelihood and are more likely to choose teaching jobs, for example. This selection problem would again bias our estimates, although the direction of the bias is not clear a priori.

An additional problem with the OLS specification is that it imposes an additively separable relationship between speaking a second language and other variables in the model. The returns to learning a second language, however, may vary markedly among different groups. Simple OLS regressions may not yield unbiased estimates of the average effect of learning a foreign language for those who do so if this effect is contingent on observable individual characteristics. We can think of this as a problem of model specification. The coefficient of the foreign-language indicator cannot be interpreted independently of the other characteristics of the individual. A fully interacted model may not be the solution, because we do not know ex ante the relevant interactions and the functional form of the treatment effect. Moreover, if the treatment group (the people who speak a second language) and the control group (those who do not) are different in their observable characteristics, the coefficient of a treatment indicator may be very sensitive to the specification (Lalonde, 1986) and capture nonlinear effects of the included controls.

We address all these concerns through several complementary empirical approaches. First, we include a very complete set of control variables in our OLS regressions to reduce the concern that selection may bias our estimates. We try to mitigate the ability bias by including test scores as ability proxies (SAT/ACT scores and GPA),\(^5\) parental education, and indicators of the quality of the college attended in our regressions. We include controls for the academic major chosen by the individuals to capture unobserved individuals’ characteristics, as the choice of a college major may be viewed as an indicator of career preferences. Second, we address the possibility that the linearity imposed on the OLS regression biases our estimates of the average effect of speaking a second language on earnings. We examine the sensitivity of the results to this assumption by using nonparametric methods based on matching on the propensity score.

We will be able to give a causal interpretation to the estimates obtained by the OLS regressions and propensity-score methods only if, once we condition on the observable characteristics of the individuals in our sample, speaking a foreign language is independent of an individual’s potential earnings. This may not be the case. To address the problem of constant unobserved individual characteristics that may bias our results, we exploit the longitudinal dimension of the data set. We thus compare the evolution of earnings for people who learned a second language between 1993 and

---

\(^5\) The SAT is required for admission at the undergraduate level to many American colleges. It is a standardized test of verbal and mathematical reasoning abilities. ACT stands for American College Test. It is a national college admission examination that consists of tests in English, reading, mathematics, and science.
I997 with the evolution of earnings for other groups. Finally, we discuss estimates from an IV strategy.

IV. Data

Our main data source is the “Baccalaureate and Beyond Longitudinal Study” (B&B henceforth), from the National Center for Education Statistics, which contains unique information on second-language ability. This data set tracks the experiences of a nationally representative cohort of college graduates who received their bachelor’s degrees during the 1992–1993 academic year. It contains information on the characteristics of the universities the individuals attended (four-year institutions in all cases), from the students’ undergraduate transcripts, and from answers to several questions posed to the individuals in surveys carried out in 1993, 1994, and 1997. These questions include information on demographics, parents’ background, earnings and job characteristics, academic major, and courses taken in college and after graduation. The sample included initially 11,192 students. Of these, 9,274 responded to the initial interview and all follow-ups.

Respondents were asked in 1993 and 1997 whether they spoke a foreign language and to identify which language this was. The question was, “Do you have conversational knowledge of languages other than English?” If the answer was affirmative, the interviewer followed with the question “What are these languages?” This question is useful for us, as it is intended to capture knowledge of a second language and not whether an individual belongs to a language minority. This feature distinguishes this data set from others, which only contain questions on second language ability for linguistic minorities. Unfortunately, B&B does not contain information on the fluency of those who claim to speak a second language. Thus, the effect that we are estimating is an average over all possible fluency levels. In particular, our study is likely to underestimate the returns to speaking a second language fluently. Another issue with the question is measurement error. Different people will draw different lines when thinking about their ability to engage in a conversation in another language. Note, finally, that the data do not allow us to examine independently the different aspects of language fluency (as in Gonzalez, 2000).

Our choice of this data set is very much driven by the existence of a second-language question geared to native speakers. Other features of B&B, however, make it very suitable for this paper: It includes variables that can be considered cognitive ability proxies, namely parental education levels, SAT or ACT scores, and college grade point average (GPA). It provides information on where the respondent’s parents were born, whether she spoke English at home while growing up, and whether she is an American citizen. We will use these variables to condition on whether the individuals were born to parents born in the United States and on whether they are American citizens who spoke English as their mother tongue. B&B also contains information on academics, job careers, and earnings. The only shortcoming of our using this data source is that the results we obtain are representative only for the population of college graduates. Thus, we suggest caution with their generalization. It is not known, a priori, if one should expect greater or smaller effects on the general population. For example, Altonji (1995) points out that “one finds a much stronger relationship between foreign language [courses in high school] and wages for those who did not attend college than for those who did.”

We complement B&B with some additional information. We use an indicator of college quality compiled by Hoxby and Long (1999). This is a classification of colleges into six categories according to the average SAT of admitted students in the college and the average high school GPA of attending students (see Table A2). Per capita income in the states of residence in 1989, 1993, and 1997 are from the Bureau of Labor Statistics. The percentages of Hispanic residents in a state are obtained for 1989, 1993, and 1997 from the U.S. Census Population Estimates, available at www.census.gov.

In a previous working paper version of this research (Saiz & Zoido, 2003) we exploited variation in high school and college graduation requirements, and we briefly comment on the results in this paper. State high school minimum graduation requirements in 1989 are obtained from the Digest of Education Statistics. We gathered information on second-language college requirements from the Modern Language Association (MLA) Survey of Foreign Language Entrance and Degree Requirements (FLEDR) (Brod & Lapointe, 1989). As shown in Saiz and Zoido (2004), these variables are good predictors of whether a college graduate has knowledge of a second language.

The basic demographic characteristics of the individuals in our sample are presented in table 1. The table is based on the 1997 wave of the survey. Panel A shows that the average individual is around 30 years old in this second survey and that she has more than eight years of labor market experience. A third of the individuals are married; less than half of them are men. More than a quarter of the individuals have received some graduate degree, and most of them (67%) attended a public college to obtain their degree. Of the individuals in the sample, 34% claim to speak a foreign language or more. Panel B shows the regional distribution of the individuals in the sample.

A. Who Speaks a Foreign Language?

We start our analysis with a brief description of the characteristics of the individuals who speak one or more foreign languages, in tables 2 and 3. Table 2 compares individuals with a bachelor’s degree who speak and who do not speak foreign languages. Panel A reveals some differences between these two groups: the probability of working, having attended a public college, holding an MBA degree, and being married is higher among those who speak English
only. They have slightly more work experience. On the other hand, those who speak a foreign language are more likely to belong to a language minority, as reflected by the variable "spoke a language different than English at home," and to have parents born outside of the United States.

The proportion of residents who speak a foreign language varies by region, as table 2C shows. New England and the Pacific region have the highest proportion of residents who speak a second language (41%). The percentage of foreign-language speakers is lowest in East North Central and East South Central.

---

**Notes:**
1. Subsample of individuals who have hourly wages above $2.8 (1st percentile) and below $42.3 (99th percentile), who answer the question on whether they speak a foreign language and with complete data on age, experience, gender, marital status, race, state of residence, college GPA, and type of college attended.
2. Mean values weighted using sample weights.
3. New England: CT, MA, NH, RI, VT; Middle Atlantic: NJ, NY, PA; East North Central: IN, IL, MI, OH, WI; West North Central: IA, KS, MN, MO, ND, SD; South Atlantic: DE, DC, FL, GA, MD, NC, SC, VA, WV; East South Central: AL, KY, MS, TN; West South Central: AR, LA, OK, TX; Mountain: AZ, CO, ID, NM, MT, UT, NV, WY; Pacific: AK, CA, HI, OR, WA.
4. Experience is calculated by subtracting 22 from the reported age of the individual.

---

**Table 1.**—Descriptive Statistics, Baccalaureate & Beyond, 1997 Survey

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log hourly wage</td>
<td>2.55</td>
<td>0.420</td>
<td>1.03</td>
<td>3.73</td>
</tr>
<tr>
<td>Speaks FL</td>
<td>0.340</td>
<td>0.474</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age</td>
<td>30.102</td>
<td>6.543</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>Experience</td>
<td>8.102</td>
<td>6.543</td>
<td>0</td>
<td>51</td>
</tr>
<tr>
<td>Married</td>
<td>0.305</td>
<td>0.460</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Male</td>
<td>0.448</td>
<td>0.497</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Normalized GPA</td>
<td>3.063</td>
<td>0.498</td>
<td>1.17</td>
<td>4</td>
</tr>
<tr>
<td>Public college</td>
<td>0.677</td>
<td>0.467</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MBA</td>
<td>0.028</td>
<td>0.167</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PhD</td>
<td>0.029</td>
<td>0.168</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other master’s</td>
<td>0.256</td>
<td>0.436</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Observations: 7940

**Table 2.**—Descriptive Statistics: Comparison of Bilingual and Monolingual Working Individuals, Baccalaureate & Beyond, 1997 Survey

<table>
<thead>
<tr>
<th>Variable</th>
<th>Speaks FL</th>
<th>Does Not Speak FL</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working</td>
<td>0.955</td>
<td>0.967</td>
<td>−0.012***</td>
</tr>
<tr>
<td>Log hourly wage</td>
<td>2.562</td>
<td>2.544</td>
<td>0.018*</td>
</tr>
<tr>
<td>Speaks FL</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Experience</td>
<td>7.611</td>
<td>8.356</td>
<td>−0.744***</td>
</tr>
<tr>
<td>Married</td>
<td>0.244</td>
<td>0.337</td>
<td>−0.093***</td>
</tr>
<tr>
<td>Black</td>
<td>0.090</td>
<td>0.104</td>
<td>0.014**</td>
</tr>
<tr>
<td>Male</td>
<td>0.445</td>
<td>0.44</td>
<td>−0.003</td>
</tr>
<tr>
<td>Normalized GPA</td>
<td>3.062</td>
<td>3.063</td>
<td>0.001</td>
</tr>
<tr>
<td>Public college</td>
<td>0.653</td>
<td>0.689</td>
<td>−0.035***</td>
</tr>
<tr>
<td>MBA</td>
<td>0.023</td>
<td>0.031</td>
<td>−0.008**</td>
</tr>
<tr>
<td>PhD</td>
<td>0.043</td>
<td>0.021</td>
<td>0.021***</td>
</tr>
<tr>
<td>Other master’s</td>
<td>0.269</td>
<td>0.249</td>
<td>0.020**</td>
</tr>
<tr>
<td>Spoke English at home</td>
<td>0.792</td>
<td>0.940</td>
<td>−0.147***</td>
</tr>
<tr>
<td>Father born in the U.S.</td>
<td>0.783</td>
<td>0.962</td>
<td>−0.179***</td>
</tr>
<tr>
<td>Mother born in the U.S.</td>
<td>0.770</td>
<td>0.952</td>
<td>−0.182***</td>
</tr>
</tbody>
</table>

Observations: 2756 (34.71%) 5184 (65.28%)

**Notes:**
1. Subsample of individuals who have hourly wages above $2.8 (1st percentile) and below $42.3 (99th percentile), who answer the question on whether they speak a foreign language and with complete data on age, experience, gender, marital status, race, state of residence, college GPA, and type of college attended.
2. Sample means weighted using sample weights.
3. Sample standard deviations shown in parentheses below sample means and standard errors of the difference in parentheses under the differences in means in panels A and B.
4. ***Statistically significant at the 1% level, **at the 5% level, *at the 10% level.
5. New England: CT, ME, MA, NH, RI, VT; Middle Atlantic: NJ, NY, PA, East North Central: IN, IL, MI, OH, WI; West North Central: IA, KS, MN, MO, ND, SD; South Atlantic: DE, DC, FL, GA, MD, NC, SC, VA, WV; East South Central: AL, KY, MS, TN; West South Central: AR, LA, OK, TX; Mountain: AZ, CO, ID, NM, MT, UT, NV, WY, Pacific: AK, CA, HI, OR, WA.
More than 15% of those who speak a foreign language actually speak more than one language other than English (table 3A). Around 14% speak two foreign languages, and almost 2% speak three. We find no differences by gender in the number of languages spoken. Most bilingual individuals speak Spanish (58%), followed by French and German (table 3B). Of the individuals in our sample, 4% speak an eastern or southeastern Asian language, and 2% speak Chinese. Spanish remains the most popular language in all regions and for both men and women. In relative terms, a higher percentage of women choose French and Italian as their second language, whereas men prefer German (the differences are statistically significant). Different languages are spoken by residents of different regions: Spanish is more popular as a foreign language in the West South Central, East South Central, and Mountain regions; residents in the South Atlantic region and New England speak French more frequently, and German speakers are disproportionately located in the West North Central region (table 3C).

V. The Returns to Speaking a Second Language

A. OLS Results

We begin our investigation of the returns to speaking a second language with a variety of conventional earnings functions estimated by OLS. Table 4 reports these estimates. The dependent variable is the log of hourly earnings in 1997 (3 to 4 years after graduation). We construct this variable by...
dividing total earnings by reported working hours.\textsuperscript{8} The main explanatory variable of interest is a dummy variable that equals 1 if the individual claims to have conversational knowledge of a second language in 1997 and 0 otherwise. We include quadratic functions of experience, and race, gender, marital status, and parental education controls. We add controls for the income per capita in the state of residence, for college quality, for whether the individual holds a graduate degree (MBA, PhD, and other master’s), and for ability proxies (normalized GPA and SAT quartile). To avoid omitted variable bias from choice of career we also control for the major in college: in appendix table A2 we illustrate how the majors in which the percentage of graduates who report speaking a second language is higher are also the majors in which the average wages are lower. The treatment of interest is a hypothetical random assignment of language skills: we are interested in the effect of speaking a second language conditional on the current career distribution, rather than considering what would happen if we shifted more people toward languages and humanities (where second-language courses are more prevalent).

The estimated foreign-language coefficient implies a 2.8% wage premium associated with speaking a foreign language for the average individual in the sample (column 1). In column 2, we restrict our sample to those individuals whose native language is English.\textsuperscript{9} Column 3 focuses on

\begin{table}[ht]
\centering
\caption{OLS Estimates} 
\begin{tabular}{lcccc}
\hline
Sample & (1) & (2) & (3) & (4) \\
\hline
Speaks FL & .028*** & .022*** & .022** & .025*** \\
 & (.010) & (.010) & (.010) & (.010) \\
Experience & .017*** & .017*** & .018*** & .016*** \\
 & (.003) & (.002) & (.003) & (.002) \\
Experience\textsuperscript{2} & -.0003*** & -.0003*** & -.0003*** & -.0003 \\
 & (.0001) & (.0001) & (.0001) & (.0001) \\
Male & .072*** & .078*** & .077*** & .065*** \\
 & (.010) & (.010) & (.010) & (.010) \\
Married & .042*** & .045*** & .050*** & .039*** \\
 & (.010) & (.011) & (.011) & (.010) \\
Black & .049*** & .005 & .005 & .045** \\
 & (.023) & (.032) & (.059) & (.023) \\
Log state income & .496*** & .50*** & .471*** & .497*** \\
 & (.035) & (.037) & (.039) & (.086) \\
Public college & -.0002 & -.0009 & -.008 & -.0002 \\
 & (.009) & (.011) & (.011) & (.010) \\
Normalized college GPA & .015* & .011 & .011 & .018** \\
 & (.008) & (.009) & (.009) & (.009) \\
Spoke FL in 1983 only & & & & \\
 & & & & \\
Parents' education & Yes & Yes & Yes & Yes \\
Major & Yes & Yes & Yes & Yes \\
Graduate degree & Yes & Yes & Yes & Yes \\
SAT-ACT quartile & Yes & Yes & Yes & Yes \\
College quality & Yes & Yes & Yes & Yes \\
State HS requirements & Yes & Yes & Yes & Yes \\
Adj. $R^2$ & .187 & .184 & .184 & .227 \\
Observations: & 7887 & 7075 & 6480 & 7444 \\
\hline
\end{tabular}
\textsuperscript{Notes:}
1. Subsample of individuals who have hourly wages above $2.8 (1st percentile) and below $42.3 (99th percentile). Each regression is performed for the maximum number of observations for which all the covariates were nonmissing.
2. Observations are weighted using sample weights.
3. Standard errors are reported in parentheses.
4. Log state income is the log of the per capita income in the state of residence in 1997 (BEA estimation).
5. Normalized GPA on a 0–4 scale for all respondents.
6. Parents’ education is a set of dummies that capture the education level of the individual’s mother and father.
7. Major is a detailed set of indicators for the student’s major and degree in college (100 major categories and 6 degree categories).
8. Quality-of-college dummies were provided by Bridget Terry Long. See appendix A for a description, and Hoxby and Long (1999) for details.
9. Graduate degree dummies are three variables that indicate if the individual has a PhD, an MBA, or a master’s.
10. ***Statistically significant at the 1% level, **at the 5% level, *at the 10% level.}
\end{table}

\textsuperscript{8} We obtain hourly wages, the dependent variable of interest to us, by dividing yearly, monthly, weekly, or daily wages by the total number of hours the individual reports having worked in the relevant period. Data on the number of hours worked are usually less reliable than earnings data (Angrist & Krueger, 1999), and this is likely to introduce measurement error. Indeed, we find some extreme values in earnings and adjust the sample by windsorizing at 1%. That is, we drop from the sample the observations in the bottom and upper 1% tails of the distribution. See Angrist and Krueger (1999) for an overview of these trimming methods as a way to reduce the biases due to measurement error.

\textsuperscript{9} The sample of nonnative English speakers is interesting in its own right, as it is likely that those who have conversational knowledge of a second language never enrolled in formal classes or needed to make a substantial investment to become proficient. We estimate a premium of
individuals who spoke English at home while growing up, are American citizens, and whose parents were born in the United States (more than 80% of the sample). When we condition on being an English native speaker with American parents, the point estimates of the foreign-language effect on earnings are only slightly lower: 2.2%. Unreported regressions controlling for occupation and industry dummies suggest a bigger coefficient of around 0.035.

The OLS results should be interpreted with caution. In our regressions, we include a complete set of controls for ability (GPA and SAT scores), marital status, college quality, and parental background, to try to reduce the concern of omitted variable biases. Still, we acknowledge the possibility that unobservables correlated with wages and the ability to speak foreign languages may bias our OLS estimates. We therefore use information on past (1993) responses to the second-language question to examine whether the estimates from our baseline earnings equations simply capture an ability bias. Our strategy is the following. If learning a second language requires cognitive skills that are not controlled for in our baseline specifications, the coefficient on the second-language indicator may be capturing the effect of some of these on earnings. We can exploit the differences between the responses regarding second-language knowledge between the 1993 and 1997 interviews to examine whether people with the necessary ability to learn a second language, but who could not speak it in 1997, earn the same wage premium as second-language speakers.10

In column 4 of table 4 we replicate the previous OLS regressions, but we now add to the control variables a dummy that equals 1 if the individual could speak a second language in 1993 but not in 1997. If the cognitive ability necessary to learn a second language is the only determinant of earnings, we should see a similar coefficient for those who could speak a second language in 1997 and those who could speak it in 1993 (during the last year at college or right after college). In fact, the coefficient of the indicator for speaking a second language only in 1993 is not very different from 0 and is statistically insignificant. The results are consistent with the idea that the returns to a second language are not completely driven by an ability bias.

1.7% associated with speaking a second language for these individuals, but the coefficient is not statistically significant at conventional significance levels.

10 The transition matrix in and out of speaking a second language between 1993 and 1997 is the following:

<table>
<thead>
<tr>
<th>Speaks FL in 1993</th>
<th>Does not speak FL in 1993</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaks FL in 1997</td>
<td>2127</td>
<td>543</td>
</tr>
<tr>
<td>Does not speak FL in 1997</td>
<td>926</td>
<td>4090</td>
</tr>
<tr>
<td>Observations</td>
<td>3053</td>
<td>4633</td>
</tr>
</tbody>
</table>

This transition matrix corresponds to the subsample of individuals who have hourly wages above $2.8 (1st percentile) and below $42.3 (99th percentile), who answer the question on whether they speak a foreign language in both surveys, and with complete data on age, experience, gender, marital status, race, state of residence, college GPA, and type of college attended. Sample means are weighted using sample weights.

These estimates exploit the fact that some individuals decide to learn a language and do not subsequently exert an effort in keeping it alive. The returns to speaking the language for the latter may have been very low. An unbiased estimate of the ex ante returns to speaking a second language, taking into account that some forget it, is given by the difference between the coefficient of those who speak minus the coefficient of those who forgot, multiplied by the fraction of people who speak now out of those who ever spoke a second language (see appendix B for an extensive discussion). Such estimates yield an approximate return of 3%.

Are the returns to speaking a foreign language homogeneous? In table 5A, we estimate OLS regressions following the previous specifications but allowing the coefficient to vary by language (Spanish, French, German, and other). In our sample, German and “other language” speakers are those who obtain the highest rewards in the labor market. The returns to speaking German or “other languages” are 4% percent; to speaking French, 2.7%; and to speaking
Spanish, 1.7%. The results indicate that those who speak languages known by a smaller number of people obtain higher rewards in the labor market. By occupation (table 5B), individuals in personal services, business support, and management positions are the ones who are more highly rewarded in the labor market for their foreign-language skills. In table 5C, we examine the returns to speaking Spanish by the percentage of Hispanics in the state of residence (for those whose second language is Spanish). The results reveal that speaking Spanish pays off less in states in which larger shares of the population are of Hispanic origin (which is consistent with a supply effect). We also examined differences by gender, race, grades in college, and graduate degree. We did not find any statistically significant differences for the other individual characteristics analyzed.

B. Panel Data Methods

In this subsection we continue to use the information on earnings and self-reported foreign-language skills from the two consecutive surveys and examine if the observed wage profiles of those who learn and forget a second language between the two surveys are consistent with the results obtained with OLS. B&B contains information on second-language ability in 1993 and 1997 and on wages in its 1994 and 1997 waves. The 1994 interview took place when college graduates had just started their first job after graduation.

We exploit the longitudinal aspect of the data set to address the omitted ability problem inherent in the previous estimation methods. Consider the following explanatory model for the logarithm of wages:

\[ \ln w_{it} = \alpha_i + \gamma S_{it} + \beta'X_{it} + \rho t + \epsilon_{it}, \]

where \( i \) and \( t \) are subscripts for individuals and time, respectively. \( S_{it} \) is a dummy that equals 1 if individual \( i \) speaks a second language at time \( t \), and \( X_{it} \) is a vector of explanatory variables. \( \gamma \) is the coefficient of interest, \( \rho \) can be interpreted as a time effect, common to all individuals, \( \alpha_i \) is an individual fixed effect; one can broadly interpret this coefficient as individual unobserved constant heterogeneity-like ability. The problem with our OLS specification is that we cannot observe \( \alpha_i \). If \( S_{it} \) and \( \alpha_i \) are correlated, the coefficient on the second-language indicator may be partially capturing the effect of ability on earnings. This suggests using the differences between the values of the variables in the two periods for which we have information:

\[ \ln w_{i1} - \ln w_{i0} = \gamma (S_{i1} - S_{i0}) + \beta'(X_{i1} - X_{i0}) + \rho + (\epsilon_{i1} - \epsilon_{i0}). \]

Table 6 shows the results of the estimation of this model. Note that only those variables that experience any change between both surveys appear in differenced form. The results that we obtain in column 1, are very similar to our OLS estimates. The wages of people who learn (forget) a second language tend to increase by almost 2% more (less) than those of other people.

One may worry that people who have the ability to learn a second language differ in their wage level and also in their wage profile over time. If this were the case, our estimate of \( \gamma \) in this first specification would be capturing the steeper wage profile over time for higher-ability individuals. One would then expect that those who speak a second language in both periods (call them *always speakers*, AS) should also experience higher wage increases. In column 2 of table 6, we include an AS indicator to allow for this possibility. As the results show, people who spoke a second language in both 1993 and 1997 do not experience higher wage increases over time. In column 3, we repeat this exercise for the subsample of individuals who spoke English at home while they were growing up, are American citizens, and have American-born parents, as we want to render the groups of those who learn a second language later in life and the AS as similar as possible. Our estimate of \( \gamma \) varies little across these specifications. These results thus suggest that the intrinsic knowledge of a second language is the explanation for the wage hike captured in the panel estimate of \( \gamma \), which is close to the estimate in the OLS regressions: the results are consistent with labor market returns to learning a second language of around 2%.

Another possibility is that the wage profile (the steepness of wages with respect to seniority) may differ according to the initial observable characteristics of the individuals:

\[ \ln w_{it} = \alpha_i + \gamma S_{it} + \beta'X_{it} + \rho t + \delta_i X_{i0} (1 - t) + \delta_i' X_{i0} t + \epsilon_{it}. \]

If the characteristics of the individuals who learn and forget a second language between the two surveys are different, our estimate of \( \gamma \) will capture only such differences. Taking differences, we have

\[ \Delta \ln w_{it} = \gamma \Delta S_{it} + \beta' \Delta X_{it} + \rho + \delta_i X_{i0} + \epsilon_{it}, \]

where \( \delta \) is \( \delta_1 - \delta_0 \). In columns 4 and 5 of table 6 we estimate this model. We control for the 1994 values of the individual characteristics that we used in our baseline specification as regressors. Column 4 includes all individuals in the sample, and column 5 is limited to those English-speaking natives who are third-generation Americans. The results are in line with those of previous specifications.

We have two caveats in the interpretation of the previous estimates. The first is the common problem of false transitions. Many people who spoke a second language in 1993 and declare they couldn’t speak it in 1997 may actually be capable of doing so after some short training. Some of them

\[ \text{(11)} \]

This is consistent with anecdotal evidence. See, for example, *Wall Street Journal*, November 13, 1999: “Where bilingual workers are in short supply, the employees may command a premium or land a job they wouldn’t otherwise have gotten.”
may not have been capable of speaking it in the first period. People who declare they have learned a second language may have actually spoken it in the first wave but not been sure about their answer. It is likely that the measurement error in the question “Do you have conversational knowledge of languages other than English?” is larger for the group who do not give a consistent answer in 1993 and 1997 than for the group that did not change its response. This amounts to bigger measurement error in the variable capturing whether the individual’s knowledge of a foreign language changes between 1993 and 1997. It should bias our estimates of γ downward, so that we obtain a lower bound. In fact, Dustmann and van Soest (2001) find that “measurement error corrections are more important than the correction for correlated unobserved heterogeneity” in the case of German-language ability for a sample of immigrants in Germany.

A second problem arises from the selection of individuals into learning and forgetting a second language. If our results are not driven by ability bias, learning a second language is, indeed, a profitable investment. Thus, we should expect people with lower returns to speaking a second language not to make the effort toward its maintenance. Conversely, those learning a second language after college are most likely the people with higher labor market returns to speaking it. These facts make the interpretation of our results in table 6 difficult. More specifically, the returns for those who learn a second language should be higher than the returns for those who forget it. Our specification imposes symmetry on the returns for the two groups. These facts make the interpretation of our results in table 6 difficult.

### Table 6.—Exploiting Information from First Survey

<table>
<thead>
<tr>
<th>Sample</th>
<th>All</th>
<th>All English Speaker</th>
<th>U.S. Citizen</th>
<th>All English Speaker</th>
<th>U.S. Born Parents</th>
<th>All English Speaker</th>
<th>U.S. Born Parents</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔSpeaks FL</td>
<td>.018** (.009)</td>
<td>.019** (.009)</td>
<td>.023** (.009)</td>
<td>.013 (.009)</td>
<td>.018* (.010)</td>
<td>.037** (.015)</td>
<td>.034** (.016)</td>
</tr>
<tr>
<td>Speaks FL 1997 only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoke FL 1993 only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaks FL in both 1997 and 1993</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔExperience2</td>
<td>.0012*** (.0001)</td>
<td>.0011*** (.0001)</td>
<td>.001*** (.0001)</td>
<td>.002*** (.0008)</td>
<td>.002*** (.0003)</td>
<td>.001*** (.0001)</td>
<td>.002*** (.0003)</td>
</tr>
<tr>
<td>Δlog state income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δmarried</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Experience in 1994)2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married in 1994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log state income 1994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public college</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normalized college GPA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College quality</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SAT-ACT quartile</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Major</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Parents’ education</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>State HS requirements</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Graduate degree</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>.028</td>
<td>.029</td>
<td>.025</td>
<td>.071</td>
<td>.068</td>
<td>.028</td>
<td>.070</td>
</tr>
<tr>
<td>Observations:</td>
<td>7686</td>
<td>7686</td>
<td>6329</td>
<td>7187</td>
<td>5929</td>
<td>7686</td>
<td>7248</td>
</tr>
</tbody>
</table>

C. Propensity-Score Methods

If the returns to speaking a second language depend on individual characteristics, or if the people who speak and do not speak a second language are very different, as we saw in table 4, then our estimates of the returns to speaking a second language may be biased because of the linearity imposed in OLS regressions. To explore how sensitive our results are to the linearity assumption, we opt for propensity-score methods that give flexibility to the functional form of the effect on earnings of speaking a second language across groups.

The propensity score is the probability of being assigned to a treatment, conditional on a set of covariates. The treatment, in our setting, is speaking a foreign language. The estimated propensity score allows us to control for differences between treatment and control groups when the treatment is not randomly assigned, the number of predetermined variables is large, and the groups are not very similar. It conveniently summarizes all individual characteristics in a single variable in the unit interval. Under selection on observables, conditioning on the propensity score is enough to yield independence between the treatment indicator and the potential outcomes (Rosenbaum & Rubin, 1983).

Though the assumption of selection on observables is strong in this setting, Dehejia and Wahba (1999) show that propensity-score methods yield estimates of treatment effects that are closer to experimental benchmark estimates than do traditional econometric methods for nonexperimental data. Thus, although the propensity score is unknown and has to be estimated, inference for average treatment effects seems to be less sensitive to specifications of the propensity score than to the specification of the conditional expectation of potential outcomes implicit in OLS regressions.

We first estimate the propensity score for our sample using a logistic probability model. We start by including only those variables that are truly predetermined to the individual and cannot thus be affected by the treatment (we do not know when the individual learnt to speak the second language): age and age squared, gender, race, whether she spoke English at home, whether parents are foreign-born, and parental education dummies. As seen in table 2, the characteristics of the treatment and control groups are very dissimilar. In particular, foreign-language speakers are much more likely to have foreign-born, foreign-speaking families. After estimating the propensity score, we use matching on the propensity score: we match each individual who speaks a foreign language with the control individual with the closest propensity score. Note that the distributions of the characteristics of both treated and untreated matched samples will correspond to that of the treatment population. In order to achieve a common support for the matched treatment and control groups, we discard treated observations for which the difference from the closest control’s propensity score is more than 0.05 (caliper matching). We then drop from our sample the observations that correspond to individuals who do not speak a second language and were not matched to “treated” individuals.

Panel B in table 7 presents the characteristics of the matched sample for those individuals who speak a foreign language and those who do not. The matched samples are remarkably close in terms of the variables used for the matching: we cannot reject that all means are identical.

We then calculate the difference between hourly wages of the individuals in the matched treated and control groups. This is an estimate of the average treatment effect on the treated observations. We find that those who speak a second language earn a wage premium of approximately 2.1% (Table 7A.I). Row (a) shows the standard error of the difference between the treatment and control matched sample. Row (b) shows the standard error of the estimate if we bootstrap the whole procedure (that is, estimating the propensity score, matching, and obtaining the difference between matched control and treatment samples). In panel A.I, column 2, we obtain an adjusted estimate of the returns to speaking a second language by controlling for the rest of covariates used in table 4, to take account of any remaining observable differences between treated and control individuals. The point estimate does not change much.

Are the results sensitive to using only the set of predetermined variables? In table 7A.II, we repeat the exercise using all covariates in the OLS regression (table 4). Some of these covariates may have been affected by the treatment (Rosenbaum, 1984). Consider, for instance, the case in which the individual learned the second language early in life. Assume that early bilingualism helps develop cognitive abilities later in life (and thus increases SAT scores, for instance). In this context conditioning on SAT scores would yield the effect of foreign language on earnings net of any effect through increased cognitive ability. This is, nevertheless, a conditional treatment effect of considerable interest. Moreover, as suggested in Rosenbaum (1984), adjustment for some posttreatment characteristics may help yield unbiased estimates if the

12 These authors compare a randomized evaluation of the National Supported Work (NSW) Demonstration with nonexperimental comparison units from survey data sets, following Lalonde’s analysis of this issue in 1986. They conclude that “there may be important unobservable covariates for which the propensity score method cannot account. However, [...,] there are substantial rewards in exploring first the information contained in the variables that are observed. Propensity score methods can offer both a diagnostic on the quality of the comparison and a means to estimate the treatment impact.” Smith and Todd (2000) have recently expressed skepticism about the effectiveness of propensity-score methods. Such methods are widely used in biology, epidemiology, and medicine. It is our view that, conditional on the relevant identification assumptions (most importantly, the ignorability in the treatment), propensity-score methods are reliable as long as the final matched samples are balanced. This depends, of course, on whether there is enough common support between the treated and control samples. Thus, it is an empirical matter.

13 We perform matching with replacement, that is, allowing the same observation for a control individual to provide the match for several treatment units. See Dehejia and Wahba (1999).

posttreatment variables are not affected by the treatment and are good proxies for relevant omitted pretreatment variables. In practice, the results (table 7A.II) suggest a slightly higher language premium of 4.1% to 4.3% higher earnings.

All propensity-score methods confirm our findings from the OLS and panel regressions: speaking a foreign language is a valuable skill in the job market. Returns are around 2%–4%. The fact that all the estimates point in the same direction and are consistent across specifications and methods is reassuring.

D. Instrumental Variable Methods

In an earlier working-paper version of this research (Saiz & Zoido, 2002), we exploited exogenous variation in the decision to learn a foreign language. We used high school and college graduation requirements, and college admission requirements by state as our instruments. All these types of academic requirements significantly increase the probability of speaking a foreign language (Saiz & Zoido, 2004). The use of the instruments yielded statistically significant estimates of the returns to speaking a foreign language of around 20%, much higher than the OLS estimates. However, the standard errors were very large. Kane, Rouse, and Staiger (1999) and Black, Berger, and Scott (2000) show that in the presence of measurement error that is negatively correlated with the actual value of the explanatory variable of interest, IV estimates are biased away from 0 in the direction of the causal effect (so the causal effect is multiplied by a positive constant). This will always happen when a dummy variable is the endogenous explanatory variable of interest. The implication is that our IV results overestimate the actual effect of speaking a second language on earnings. Although the IV estimates in Saiz and Zoido (2002) are not very useful in a quantitative sense, they do provide supporting evidence on the positive effect of bilingualism on labor market earnings.

VI. Conclusions

Recent developments highlight the key strategic importance of foreign languages and the lack of critical language skills in sectors of the U.S. intelligence community. But is there really a generalized shortage of foreign language skills in the United States? One would expect economic globalization to increase demand for foreign languages in the United States, whereas the establishment of English as an international lingua franca may reduce demand. Given the current supply of foreign-language skills produced by the U.S. educational system, it is important to examine the labor market returns to speaking a second language in order to investigate this question.

In this paper we have explored this question using a variety of empirical strategies. Although none of them provides an ideal approach to the problem, all of them point in the same direction: speaking a foreign language is rewarded in the labor market. The earnings of college graduates who speak a foreign language are higher than the earnings of those who don’t. Our estimates of the impact of bilingualism on earnings are relatively small (2%–3%) and compare unfavorably with recent estimates on the returns to one extra year of general schooling (8%–14%), which may help explain current second-language investment decisions of monolingual English-speakers in the United States. The returns may be higher for individuals in management and the business services occupations, and they seem to be higher ex post for individuals who learn a second language later in life. Note that the results in this paper are estimates of the gross returns to learning a second language. Individual decisions on whether to study a foreign language will

| Table 7.—Estimated Effect of Speaking a Second Language, Matching on the Propensity Score |
|---------------------------------------------|----------|----------|
| (A) Results                                | (1) Unadjusted | (2) Adjusted |
| I. Predetermined Variables                |          |          |
| Speaks FL                                 | .021     | .020     |
| (a)                                       | (.013)   | (.013)   |
| (b)                                       | (.022)   | (.023)   |
| Observations: 4201                        |          |          |
| II. All Variables (Table 4)               |          |          |
| Speaks FL                                 | .043     | .041     |
| (c)                                       | (.013)***| (.011)***|
| (d)                                       | (.021)***| (.019)***|
| Observations: 4535                        |          |          |

Notes:
1. Sample of individuals who have hourly wages above the 1st percentile and below the 99th percentile, who answer the question on whether they speak a foreign language, and for whom the variables included in the regressions were complete. The sample is further restricted to individuals who speak a foreign language and to those who don’t with the closest propensity score.
2. Matching with repetition. Panel A.I uses the propensity scores estimated using a logit model for matching (only predetermined variables). Panel A.II uses a logit model that includes all control variables.
3. Adjusted results include the set of controls presented in table 3.
4. Standard errors in parentheses. (a) and (c): standard errors from differences within the final matched sample. (b) and (d): bootstrapped errors for propensity-score estimation and matching procedure. Although the results replicate some observations from the control group, all standard errors take into account the actual number of observations.
5. ***Statistically significant at the 1% level, **at the 5% level, *at the 10% level.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Speak</th>
<th>Not Speak</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>29.470</td>
<td>29.137</td>
<td>0.333</td>
</tr>
<tr>
<td>Black</td>
<td>0.073</td>
<td>0.074</td>
<td>−0.001</td>
</tr>
<tr>
<td>Male</td>
<td>0.423</td>
<td>0.410</td>
<td>0.013</td>
</tr>
<tr>
<td>Spoke English at Home</td>
<td>0.784</td>
<td>0.799</td>
<td>−0.015</td>
</tr>
<tr>
<td>Father Born in U.S.</td>
<td>0.789</td>
<td>0.798</td>
<td>−0.009</td>
</tr>
<tr>
<td>Mother Born in U.S.</td>
<td>0.780</td>
<td>0.787</td>
<td>−0.007</td>
</tr>
</tbody>
</table>

(1) Unadjusted (2) Adjusted
depend on (among other things) the opportunity cost of the time devoted to learning it and its nonmonetary rewards.

REFERENCES


APPENDIX A

Data Appendix: Quality-of-Schooling Variables

This is a set of dummies that indicate the quality of the college. The classification (table A1) was provided by Bridget Terry Long and is defined in detail in Hoxby and Long (1999). Colleges were grouped according to criteria taken from the 1997 edition of Barron’s *Profiles of American Colleges*. These criteria include raw test scores, percentage of the class scoring over a certain level, and high school rank. Schools
LISTENING TO WHAT THE WORLD SAYS

TABLE A1.—SECOND-LANGUAGE REQUIREMENTS FOR COLLEGE GRADUATES: INTERNET JOBS

<table>
<thead>
<tr>
<th>Category</th>
<th>GPA</th>
<th>SAT</th>
<th>ACT</th>
<th>Sample Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most competitive</td>
<td>3.75–4.50</td>
<td>1280–1600</td>
<td>31–36</td>
<td>Princeton, Harvard, UVA</td>
</tr>
<tr>
<td>Highly competitive</td>
<td>3.40–3.74</td>
<td>1120–1279</td>
<td>26–30</td>
<td>UC Berkeley, University of Michigan</td>
</tr>
<tr>
<td>Very competitive</td>
<td>3.00–3.39</td>
<td>1000–1119</td>
<td>21–25</td>
<td>University of Connecticut</td>
</tr>
<tr>
<td>Competitive</td>
<td>2.50–2.99</td>
<td>850–999</td>
<td>16–20</td>
<td>University of Massachusetts</td>
</tr>
<tr>
<td>Less competitive</td>
<td>2.00–2.49</td>
<td>650–849</td>
<td>12–15</td>
<td>University of Mississippi</td>
</tr>
<tr>
<td>Noncompetitive</td>
<td>1.00–1.99</td>
<td>400–649</td>
<td>4–11</td>
<td>Ohio State University—Mansfield Campus</td>
</tr>
</tbody>
</table>

Notes: The Careerbuilder.com search was performed on April 14, 2003, from 6:30 to 7:30 p.m. (approximately). The Monster.com search was on April 15, 2003 from 10:40 to 11:35 a.m. We selected those jobs that contained the word bachelor’s and had been posted in the last 24 hours. We then searched within those jobs for the language names and checked each posting to make sure that the second language was, indeed, required and to avoid duplications (some postings mention two or more languages).

TABLE A2.—QUALITY GROUPING CRITERIA

<table>
<thead>
<tr>
<th>Category</th>
<th>GPA</th>
<th>SAT</th>
<th>ACT</th>
<th>Sample Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most competitive</td>
<td>3.75–4.50</td>
<td>1280–1600</td>
<td>31–36</td>
<td>Princeton, Harvard, UVA</td>
</tr>
<tr>
<td>Highly competitive</td>
<td>3.40–3.74</td>
<td>1120–1279</td>
<td>26–30</td>
<td>UC Berkeley, University of Michigan</td>
</tr>
<tr>
<td>Very competitive</td>
<td>3.00–3.39</td>
<td>1000–1119</td>
<td>21–25</td>
<td>University of Connecticut</td>
</tr>
<tr>
<td>Competitive</td>
<td>2.50–2.99</td>
<td>850–999</td>
<td>16–20</td>
<td>University of Massachusetts</td>
</tr>
<tr>
<td>Less competitive</td>
<td>2.00–2.49</td>
<td>650–849</td>
<td>12–15</td>
<td>University of Mississippi</td>
</tr>
<tr>
<td>Noncompetitive</td>
<td>1.00–1.99</td>
<td>400–649</td>
<td>4–11</td>
<td>Ohio State University—Mansfield Campus</td>
</tr>
</tbody>
</table>


Notes: The categories were created referring to Barron’s criteria for school groupings, which are based on the characteristics of the freshman class entering 1995–1996 (before SAT scores were recentered).

TABLE A3.—PERCENTAGE OF SECOND-LANGUAGE SPEAKERS AND AVERAGE EARNINGS BY MAJOR, 1997 SURVEY

<table>
<thead>
<tr>
<th>Category</th>
<th>% Speak a FL</th>
<th>Average Log(Wage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>0.214 (.021)</td>
<td>2.818 (.019)</td>
</tr>
<tr>
<td>Engineering</td>
<td>0.216 (.022)</td>
<td>2.844 (.020)</td>
</tr>
<tr>
<td>Computer sciences</td>
<td>0.219 (.036)</td>
<td>2.774 (.032)</td>
</tr>
<tr>
<td>Vocational/technical</td>
<td>0.226 (.036)</td>
<td>2.524 (.032)</td>
</tr>
<tr>
<td>Business</td>
<td>0.227 (.011)</td>
<td>2.599 (.010)</td>
</tr>
<tr>
<td>Education</td>
<td>0.271 (.014)</td>
<td>2.384 (.012)</td>
</tr>
<tr>
<td>Life sciences</td>
<td>0.277 (.021)</td>
<td>2.404 (.019)</td>
</tr>
<tr>
<td>Other technical/professional</td>
<td>0.318 (.018)</td>
<td>2.530 (.016)</td>
</tr>
<tr>
<td>Social sciences</td>
<td>0.350 (.014)</td>
<td>2.489 (.012)</td>
</tr>
<tr>
<td>Mathematics</td>
<td>0.346 (.044)</td>
<td>2.519 (.039)</td>
</tr>
<tr>
<td>Physical sciences</td>
<td>0.382 (.047)</td>
<td>2.496 (.041)</td>
</tr>
<tr>
<td>Humanities</td>
<td>0.450 (.017)</td>
<td>2.428 (.015)</td>
</tr>
</tbody>
</table>

That had not been ranked were categorized according to these criteria. These have remained fairly constant over the last 15 years, and thus the use of the 1997 list is valid. Scores were reported as nonrecentered. The SAT scores were converted into percentiles.

APPENDIX B

A Simple Model

The following simple model illustrates the relation between ability and the returns to learning a second language. It provides a justification for the strategy followed in the regressions in table 7. Assume that the decision on learning a second language (SL hereafter), which involves a period of training, is undertaken before full participation in the labor market. Assume, furthermore, that there are two kinds of individuals. A fraction $\alpha$ of the population has low costs of learning an SL, $C_L$ (which corresponds to high ability). The rest of the population has costs $C_H$ with $C_L < C_H$. Ability is not observable by the econometrician. Assume that if the SL is learned, it will turn out to be useful in the labor market (producing extra earnings $V_A$) with probability $p$ and turn out not to be useful with probability $1 - p$. Assume that there are no consumption advantages from speaking a foreign language. Gross labor market earnings are also a function of ability:

$$w_i = \begin{cases} V_A + A - \beta C_H & \text{if the individual speaks an SL} \\ A - \beta C_L & \text{if the individual does not speak an SL} \end{cases}$$

For $i = H, L$, low-learning-cost (high-ability) individuals will decide to learn the SL if

$$pV_U \geq C_L$$

For high-learning-cost (low-ability) individuals, the condition is

$$pV_U \geq C_H$$

Consider the case when

$$C_H \geq pV_U \geq C_L$$

Then all high-ability individuals learn an SL, and low-ability individuals do not.

Assume that individuals can maintain their SL human capital after the realization of the labor market shock with a very small cost $\epsilon$. High-ability individuals for which the SL turned out to be useful will decide to maintain their stock of SL capital. High-ability individuals for which the SL did not turn out to be useful will decide not to maintain it. Ex post, only individuals with high ability and a positive realization will speak an SL. The pool of those who do not speak an SL will be formed by (1) low-ability individuals and (2) high-ability individuals with negative labor market realizations.

We know that the expected return of learning an SL, corresponding to a complete randomization of the SL treatment, is equal to $pV_U$. We may try to estimate the returns to learning an SL using the difference between the wages of those who speak an SL and the rest ($w_i - w_{in}$). Let $f$ be an indicator that takes value 1 if the individual speaks an SL and 0 otherwise. The expectation of this estimator is
\[
E(w_s - w_{\text{忘了}}) = E(w|SL = 1) - E(w|SL = 0) \\
= (V_u + A - \beta C_L) \\
- \left[ \frac{(1 - p)\alpha}{1 - \alpha p} (A - \beta C_L) + \frac{1 - \alpha}{1 - \alpha p} (A - \beta C_H) \right] \\
= V_u + \frac{\beta(1 - \alpha)}{1 - \alpha p} (C_H - C_L) > V_u > pV_u.
\]

This estimate is too high, for two reasons. First, there is an ability bias: individuals who speak an SL have higher average unmeasured ability. Second, there is a problem of selection by earnings: those who speak an SL ex post tend to have greater returns than the rest.

These biases suggest the use of the alternative estimator: \( p \cdot (w_s - w_{\text{忘了}}) \), where \( w_{\text{忘了}} \) is the wage for those individuals who learned the SL and did not invest in its maintenance ex post. This is an unbiased estimator of the ex ante return to speaking an SL:

\[
E(w_s - w_{\text{忘了}}) = E(w|SL = 1) - E(w|\text{forgot}) \\
= (V_u + A - \beta C_L) - (A - \beta C_L) = V_u.
\]

Multiplying by the fraction of people who speak an SL out of the total who ever spoke it yields the ex ante expected returns to learning an SL. From column 1 in table 4 we can derive

\[
w_s - w_{\text{忘了}} = 0.029 - (-0.013) = 0.042.
\]

The fraction of people who speak an SL out of the fraction who ever spoke it is equal to the number of people who speak it divided by the number of people who speak it plus the number of people who forgot it:

\[
p = \frac{N_s}{N_s + N_{\text{忘了}}} = \frac{2562}{2562 + 887} = 0.743.
\]

Thus, our estimate of the ex ante expected returns of learning an SL is

\[
p \cdot (w_s - w_{\text{忘了}}) = 0.031.
\]

Note how this is a conservative estimate of the expected returns to learning an SL, as we are assuming that the returns for those who forgot the SL is 0. Furthermore, note that the fact that the coefficient for those who forgot an SL is close to 0 suggests that the ability bias is small.