The Returns to English-Language Skills in India

Mehtabul Azam

World Bank & IZA[*]

Aimee Chin

University of Houston & NBER[†]

Nishith Prakash

Cornell University, IZA & CReAM[‡]

July 2010[§]

Abstract

India's colonial legacy and linguistic diversity give English an important role in its economy, and this role has expanded due to globalization in recent decades. It is widely believed that there are sizable economic returns to English-language skills in India, but the extent of these returns is unknown due to lack of a microdata set containing measures of both earnings and English ability. In this paper, we use a newly available data set–the India Human Development Survey, 2005–to quantify the effects of English-speaking ability on wages. We find that being fluent in English (compared to not speaking any English) increases hourly wages of men by 34%, which is as much as the return to completing secondary school and half as much as the return to completing a Bachelor's degree. Being able to speak a little English significantly increases male hourly wages 13%. There is considerable heterogeneity in returns to English. More experienced and more educated workers receive higher returns to English. The complementarity between English skills and education appears to have strengthened over time. Only the more educated among young workers earn a premium for English skill, whereas older workers across all education groups do.

JEL Classification: J31, J24, O15
Keywords: English Language, Human Capital, India

---

[*]Consultant, World Bank, Washington, DC 20433 USA (e-mail: mazam1@worldbank.org).

[†]Associate Professor, Department of Economics, University of Houston, Houston, TX 77204-5019 USA (e-mail: achin@uh.edu).

[‡]Post Doctoral Research Associate, Charles H. Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY 14853-7801 USA (e-mail: nishith.prakash@cornell.edu).

1 Introduction

India's colonial legacy and linguistic diversity give English an important role in its economy, and this role has expanded due to globalization in recent decades. It is widely believed that there are sizable economic returns to English-language skills in India, but the extent of these returns is unknown due to lack of a microdata set containing measures of both earnings and English ability.[1] We take advantage of a newly available nationally representative individual-level data set, the India Human Development Survey (IHDS), 2005, to provide the first estimates of the returns to English-language skills in India. A second contribution of this paper is to provide new descriptive information about the prevalence of English ability in India. Based on the 1991 Census, 11% of the Indian population reported some English ability. It would be useful not only to have more recent figures, but also to examine English ability along various dimensions such as education, age and sex.

A major challenge to estimating the returns to English is the likely endogeneity of language skills in the earnings equation. A priori, omitted variables bias seems to be a serious concern; for example, omitted ability or local labor market conditions can be correlated with both language skills and earnings. We exploit the richness of IHDS data to address these concerns. For example, the IHDS has data on individual performance on the secondary school leaving certificate board examination, which provides a credible proxy for ability in the Indian context. As well, the large sample size and detailed geographic identifiers permit controlling for labor market characteristics via district fixed effects.

Our main findings are as follows. On average, wages are 34% higher for men who speak fluent English and 13% higher for men who speak a little English relative to men who speak no English. There is considerable heterogeneity in returns to English. More experienced and more educated workers receive higher returns to English. The complementarity between

---

[1]Munshi and Rosenzweig (2006) and Chakraborty and Kapur (2008)estimate the returns to attending a school using English as the medium of instruction. We explain in subsection 2.2 that this is not the same as the returns to English-language skills.

English skills and education appears to have strengthened over time. Only the more educated among young workers earn a premium for English skill, whereas older workers across all education groups do.

This study is of interest for several reasons. Foremost, knowing the returns to English would help individuals and policymakers in India make decisions about how much to invest in English skills. Language skills are costly to acquire, and it is difficult to make optimal choices without knowledge about the expected benefits of English-language skills. Additionally, this study informs on the more general question of the value of English in a context where English is not a prevalent language. English is often used as a lingua franca–the language of communication among two parties who do not share a common native language–and many countries, even ones that are not former British or American colonies, invest in English skills.

The remainder of this paper is organized as follows. Section 2 provides background on English in India and discusses the related literature. Section 3 presents the empirical framework. Section 4 describes the data. Section 5 presents the results on returns to English-language skills in India. Section 6 concludes.

2 Background and Related Literature

2.1 English in India

India is a linguistically diverse country–it has thousands of languages, of which 122 have over 10,000 native speakers according to the 2001 Census. English is only 44th on the list of languages in India with the most native speakers, belying its important role in India since the arrival of the British East India Company in the 1600s. India was formally ruled by the British Empire from 1757-1947 (by the British East India Company from 1757-1857, and by the British Crown from 1858-1947). During this time, English became the language of power and prestige. It was associated with the ruling British, the law was in English, and government administration, at least at the higher levels, was conducted in English.

Additionally, it became the medium of instruction in public schools.[2]

After India gained independence from the British in 1947, debate ensued over the role of the colonial language in the country. There were calls to replace English with a native Indian language as the official language of India to reinforce national identity. A natural candidate was Hindi, which is by far the most dominant mother tongue in India.[3] However, it was politically infeasible to make Hindi the sole official language of India as it was thought to be disadvantageous to states where Hindi was not prevalent–Hindi is spoken by most in the north, by few in the south. Thus, the Constitution of India names both Hindi and English as the official languages of India. Individual states legislate their own official languages, but communication among states and in the federal government would take place in Hindi or English.

From an individual's perspective, there are several economic incentives to learn English. On the one hand, English has value as a lingua franca. A knowledge of a common language facilitates communication. A common language is especially useful in linguistically diverse places, where the chances of meeting someone with the same native language is relatively low. In India, there is considerable variation in languages spoken even within narrowly defined regions, such as the district.[4] A common language is also useful for international trade. While English is not the only possible lingua franca, it is a natural one given India's colonial past and given the influence of the United States in the world economy. On the other hand, the use of English is firmly entrenched in government and schools due to the colonial past. To be a government official or teacher (other than at low levels), one needs to be proficient in English. These occupations are considered attractive in India because they are white-collar jobs providing secure employment and good benefits. In contrast, most jobs in the India are

---

[2]Under British rule, India established a system of public education; before, there were few schools and only the elite received schooling. It was decided after much debate that English would be the medium of instruction in this new system of public schools.

[3]In 2001, 40% of the population named Hindi as their mother tongue; the next language with most native speakers, Telugu, claimed only 8%.

[4]See Shastry (2008) for more on within-district variation in languages spoken. India is comprised of 35 states/union territories which are subdivided into 628 districts.

on household farms or in casual labor, which tend to provide uncertain means of livelihood and involve strenuous physical labor.

Though only 0.2% of the Indian population reported English as their mother tongue in the 2001 Census, considerably more know it as a second or third language. According to the 1991 Census, 11% of the Indian population reports English as a second or third language. It is widely believed that English knowledge has grown since 1991, but there has been no data to substantiate these claims until now, with the release of the India Human Development Survey (IHDS), 2005 (we describe these data in Section 4). Table 1 reports the mean English ability among individuals aged 18-65 in the IHDS along various dimensions. One in five Indians report having the ability to speak English, comprised of 4% who can converse fluently in English and 16% who can converse a little in English. English ability is higher among men–approximately 26% of men report having the ability to speak English compared to 14% of women–and this is likely due to the differences in educational attainment, which we discuss below. English ability is higher among younger people–25% of people aged 18-35 speak English compared to 13% for people aged 51-65. These differences may be due to differences in educational attainment, greater incentives to learn English due to globalization in recent decades, or depreciation of English skills with time since leaving school.

The ability to speak English increases dramatically with educational attainment in India. Almost 89% of individuals who have at least a Bachelor's degree can speak English as compared to 56% for those who have completed secondary schooling (10-14 years of schooling completed), 11% for those who have completed 5-9 years, and virtually nil for those who have less schooling. The positive relationship between English ability and educational attainment is not surprising since English is not the native language of 99.8% of the Indian population, and thus the main exposure to English for children would be in schools. In India, many public schools follow the "Three Language Formula" recommended by the central government, which generally leads to teaching in English by middle school.[5] According

---

[5]This calls for the teaching in the mother tongue or regional language during primary school. After

to the 1986 All-India Education Survey, which is a census of schools, 1.3% of schools with grades 1-5 used English as the medium of instruction, and 15% reported teaching English as a first or second language. In schools with grades 6-8, these figures rise to 3.6% and 63%, respectively. In secondary schools (covering grades 9 and 10), 8.2% use English as the medium of instruction, and 65% teach English as a first or second language. In higher secondary schools, colleges and universities, English is often used though it should be pointed out that it is possible to graduate from secondary school and college without being proficient in English; except in the science and engineering fields, many courses are offered in Hindi or the state language, and exams may be written in English, Hindi or the state language.

Next, we examine percent speaking English by social group. In India, the two most disadvantaged social groups are the schedule tribes (STs) and scheduled castes (SCs).[6] The Other Backward Classes (OBCs) are above the SCs in ritual standing, but are also much worse off than the higher castes. English ability is greater among members of higher castes than members of lower castes or the scheduled tribes. This is likely related to the lower educational attainment, and in the case of the scheduled tribes, of their geographic isolation.

The bottom of Table 1 presents mean English ability by geography. There is a large difference in English ability by urban status: 35% of individuals living in urban areas report to have ability to speak English as compared to only 14% living in rural areas. There is considerable variation in English ability across the regions as well. This is further illustrated in Figure 1, which displays mean English ability by state.

Having described the prevalence of English proficiency in India using the IHDS, we proceed to estimate the returns to English proficiency in India. Before we do this, we discuss the related literature and our empirical framework.

---

primary school, introduce a second language–this might be Hindi (in states where Hindi is not the dominant language) or English or some other modern Indian language. After middle school, introduce a third language.

[6]STs are distinguished by their tribal culture and physical isolation. SCs are groups with low social and ritual standing. In the British era, these were called the depressed classes, and colloquially they have also been called the untouchables and backward classes though these terms are out of favor. The Constitution (Scheduled Castes) Order of 1950 and the Constitution (Scheduled Tribes) Order of 1950 lists which castes, races and tribes are designated SCs and STs, respectively.

2.2 Previous Literature

We are aware of two previous studies on the relationship between English-language skills and earnings in India: Munshi and Rosenzweig (2006) and Chakraborty and Kapur (2008), where the latter is an unpublished manuscript. Both estimate the returns to attending a school with English (as opposed to some native language) as the medium of instruction. Munshi and Rosenzweig collected their own data on Maharashtrians living in Dadar, which is located in Mumbai, Maharashtra, India. Using data on parents' income histories and the language of instruction in their secondary school (Marathi or English), they estimate significant positive returns to an English-medium education.[7] Attending an English-medium school increased both women's and men's income by about 25% in 2000. Chakraborty and Kapur use National Sample Survey data to estimate the impact of a 1983 policy in West Bengal which eliminated English as the medium of instruction in primary schools. They find that switching from English to Bengali medium of instruction significantly reduced wages. Simple comparisons of cohorts attending primary school before and after the policy change suggest that English-medium schooling raised wages about 15% in the 2000s.[8]

Our study differs from the two aforementioned studies in two key respects. First, the "returns to English" that we are estimating is the returns to English-language skills as opposed to the returns to English-medium education. In general, we might think that being taught in English would increase one's English-language skills relative to being taught in some other language, so the latter estimates just need to be scaled up by some factor to obtain the former.[9] Angrist and Lavy (1997), for example, find that French-language skills significantly deteriorated in Morocco as a result of a policy that changed the language of instruction

---

[7]These returns are described in greater detail in Munshi and Rosenzweig (2003).

[8]Estimates controlling for secular cohort trends suggest somewhat larger effects.

[9]That is, we might think of the returns to English-medium schooling as a reduced-form relationship between English-language skills and earnings. In order to recover the returns to English-language skills, one needs to know the "first-stage" effect of English-medium schooling on English-language skills. In practice, there might be complications since English-medium schooling might impact earnings through mechanisms other than English-language skills. For example, Roy (2004) finds that the West Bengal policy that changed the medium of instruction from English to Bengali increased educational attainment.

in post-primary grades from French to Arabic. However, Angrist, Chin and Godoy (2008) find that in Puerto Rico, switching the medium of instruction from English to Spanish in Puerto Rico had no impact on the English-speaking proficiency of Puerto Ricans; thus, it is not a foregone conclusion that instruction in a foreign language will lead to greater proficiency in that foreign language. In fact, the premise of He, Linden and MacLeod (2008) is that Indian primary schools are ineffective at teaching English.[10] A second difference is that our study uses a large, nationally representative data set, which enables us to explore potential heterogeneity in returns to English-language skills along various dimensions (below, we will allow returns to vary by sex, age, education, urban/rural residence and social group). Munshi and Rosenzweig's findings come from one community in Mumbai, and Chakraborty and Kapur's findings come from a policy change in one state, West Bengal.

There is a large literature on the effects of language skills on wages using data from other countries. However, most of these studies estimate the returns to the host-country language for immigrants to that host country, such as the returns to English for U.S. immigrants. Bleakley and Chin (2004) provide a brief overview of these studies. Fewer studies estimate the return to a language that is not the country's dominant language. Two studies that estimate the effect of a colonial language are Angrist and Lavy (1997), who estimate the return to French-language skills in Morocco, and Levinsohn (2007), who estimates the returns to speaking English in South Africa. Two that estimate the effect of foreign languages that do not have a colonial past in the country are Saiz and Zoido (2005) and Lang and Siniver (2006). Saiz and Zoido estimate the returns to Spanish, French and other foreign languages among U.S. college graduates. Lang and Siniver estimate the returns to English proficiency in Israel, a country where English is neither a dominant nor official language. The latter two studies suggest that proficiency in an international language such as English is rewarded

---

[10]They perform a randomized evaluation of a new methodology for teaching English in primary schools. At the outset of the experiment, they found that only 10% of second and third graders could identify the picture of the correct object when given the object's English name even though these words were part of the official English curriculum.

more in the labor market than proficiency in some other foreign language.

Our study is also related to a couple of recent working papers on the connection between language and economic development in India. Shastry (2008), Oster and Millett (2010) and Jensen (2010) find using different data sets and methodologies that school enrollment increases with the expansion of the information technology sector. The results are consistent with households perceiving a higher returns to English skills though these returns are not estimated. Clingingsmith (2008) finds that districts that had greater increases in manufacturing employment experienced greater increases in the proportion of minority-language speakers becoming bilingual (where the second language is a regional or national language).

This study makes several contributions. It is the first to estimate the returns to English-language skills in India. A major hurdle in the past has been the availability of a data set measuring both language skills and earnings of individuals. The recent availability of data from the India Human Development Survey (IHDS), 2005 helps us clear this hurdle. Additionally, it adds to the handful of studies that estimates the effect of proficiency in a foreign language that has colonial roots or serves as an international language. Finally, it adds to the few studies on the topic using data from a developing country.

3 Empirical Framework

English-language skills are a form of human capital. Individuals, or parents acting on their behalf, weigh the marginal costs and marginal benefits of investing in English-language skills. There could be both monetary and non-monetary costs associated with acquiring English-language skills. Non-monetary costs include the effort to learn English, which is not the native language of 99.8% of the Indian population. They might also include weakened ties to one's tradition social network because in order to learn English well, one might have to attend different schools or study the native language less relative to other members of the network (Munshi and Rosenzweig 2006). Extra monetary costs are incurred to the extent

that parents enroll their child in private schools, hire tutors, or invest in more years of schooling than they otherwise would in order to help the child learn English.[11]

Among the benefits of having English-language skills is getting a better job. Though it is widely acknowledged that English-language skills are valuable in India, due to lack of data, the returns to English-language skills have not been quantified.[12] This skill price is determined by the supply of and demand for English-proficient workers. The recent availability of data from the India Human Development Survey (IHDS), 2005, which contains measures of English-speaking ability and earnings, enables us to estimate it.

The relationship between English-language skills and earnings might be approximated by the following equation:

$$y_i = \alpha + \beta English_i + \pi X_i + e_i \tag{1}$$

where y_i is the log hourly earnings of individual i, $English_i$ is a measure of English-language skills, and X_i is a set of demographic controls (e.g., age, sex, social group). The coefficient of primary interest is β, which gives the returns to English-language skills.

Omitted variables bias will likely be a concern for several reasons. One omitted variable that ex ante we thought would be important is years of schooling. Early grades are more likely to take place in the native language, with the regional or national language used as the medium of language in later grades. This generates a positive relationship between English proficiency and years of schooling, leading the ordinary least squares estimate of β in Equation 1 to be upward biased; some of the estimated effect of English proficiency is in fact due to schooling. To address this issue, we will control for years of schooling.

A second important omitted variable in Equation 1 is geographic characteristics. Places

---

[11]Roy (2004) finds that as a result of the change in medium of instruction in West Bengal public primary schools from English to Bengali in 1983, parents spent more on private tutors (presumably to provide English lessons). There was no estimated impact on private school attendance but Roy suggests that perhaps there were supply constraints in the short run.

[12]As discussed in the previous section, Munshi and Rosenzweig (2006) and Chakraborty and Kapur (2008) estimate the returns to English-medium schooling which is not the same as the returns to English-language skills. Moreover, it is of interest to explore heterogeneity in returns to English, which we can do because of our large, diverse sample.

where English is more prevalent are different from places where English is less prevalent. For example, English is more prevalent in major cities and outside the Hindi belt (states in central and northern India where Hindi is the mother tongue), but these are also places where wages are on average higher. There could be a causal relationship between wages and English-language skills, or a correlation could exist through some third factor. A causal story going from language skills to industrial development is that the availability of a more skilled (e.g., English proficient) labor force in a particular region attract firms to locate there. Shastry (2008) finds that districts with a lower cost of acquiring English (and therefore, where English was more prevalent) had greater growth in information technology jobs. A causal story in the reverse direction is that higher returns for a particular skill (e.g., English proficiency) motivates individuals to invest more to develop that skill. Clingingsmith (2008) finds that Indian districts with more factory employment experience greater growth in bilingualism. Third factors which might be behind the observed correlation include the place's institutional environment, economic conditions, climate and natural resources. The foregoing considerations suggest that it is essential to control for local conditions. We therefore control for urban residence and state of residence. In some regression models, we control for geographic conditions more finely by including district of residence fixed effects, i.e., we are comparing individuals who speak English and those who do not *within* the same district.

A final important source of omitted variable bias in Equation 1 is the standard "ability" bias. Higher-ability individuals (e.g., those with greater aptitude or more advantaged family background) are more likely to have better English-language skills as well as better jobs, leading to an upward biased estimate of the return to English since ability is omitted. Controlling for years of schooling mitigates the ability bias somewhat because those with greater ability tend to complete more schooling, but this is imperfect. We take advantage of the richness of the IHDS data to address the ability bias in additional ways. One way is to control for father's education; controlling for parental education has been a traditional approach in the returns to education literature to address ability bias (see Card(1999) for a

review). Another way is to control for the individual's performance on the secondary school leaving certificate (SSLC) examination. In India, students must pass a standardized exam developed by the board of education under whose jurisdiction their school falls in order to receive a SSLC. This exam is typically taken at the end of 10th grade, and the passing categories, from highest to lowest level of distinction, are Division I, II and III. Performance on the SSLC exam–even if measured in a only a few categories–seems like a credible proxy for ability in the Indian context, and is akin to controlling for aptitude test scores to address the ability bias when estimating the returns to schooling in the U.S. context. It is unusual, and fortunate for us, that the IHDS collected this SSLC exam performance data. However, it should be recognized that this information is available only for individuals who have attended 10th grade, something that only a fifth of individuals aged 18-65 in the data set has done. Thus we regard SSLC exam performance as a proxy of ability among those with more schooling. As a proxy for ability among the less educated, we use the response to the question of whether the individual has failed or repeated a grade.

Given the forgoing discussion on likely sources of omitted variables bias, we amend Equation 1 as follows:

$$y_i = \alpha_r + \beta English_i + \delta Schooling_i + \gamma Ability_i + \lambda Urban_i + \pi X_i + e_i \qquad (2)$$

where $Urban_i$ indicates whether individual i lives in an urban area, α_r is either state or district fixed effects, $Schooling_i$ is years of schooling completed and $Ability_i$ are proxies for ability (i.e., father's education, SSLC exam performance, and failing or repeating a grade).

Though we believe Equation 2 addresses the main sources of bias, nonetheless there could be concerns about remaining omitted variables and measurement error. Measurement error is potentially serious because respondents self-report their English ability and English ability is measured in a few categories. Previous work addressing the endogeneity of language-skills measures in the earnings equation suggests that the downward bias due to classical

measurement error may be much more severe than the upward bias due to ability bias (Bleakley and Chin (2004), Dustmann and van Soest (2002)), and consequently instrumental-variables estimates of the returns to English ability tend to be larger than OLS estimates.[13] Thus, we might regard the estimates of the returns to English in this paper as lower bounds.

4 Data

We use data from the 2005 India Human Development Survey, a nationally representative household data set collected by the National Council of Applied Economic Research in New Delhi and the University of Maryland (Desai, Reeve and NCAER 2009). IHDS covers 41,554 households located throughout India.[14] This data set became available for public use only recently, and has not yet been exploited by economists. The survey contained many questions that are not asked in the larger and more commonly used Indian household survey, the National Sample Survey. Most relevant for us is that information about each household member's ability to converse in English is collected. We are not aware of any other large-scale individual-level data set in India that contains a measure of English-language skills.[15]

Since the outcome of interest is earnings, we restrict our sample to individuals aged 18 to 65. Our main analysis will use individuals who report working for a wage or salary last year. In our sample of 125,170 people aged 18-65, 38% participated in wage employment (the figure is 54% for men and 21% for women). This raises concerns sample selection bias. This problem could be serious in the context of India, where over 70% of the population is rural, and family farms and non-farm businesses continue to absorb much of the labor force. To address this, we will use two alternative measures of earnings that are observed regardless

---

[13]The OLS estimate suffers from both attenuation bias and ability bias whereas the IV estimate is consistent.

[14]The survey covered all the states and union territories of India except Andaman and Nicobar and Lakshadweep, two union territories which together account for less than .05% of India's population.

[15]The census does collect information on language knowledge of the population, however these data are reported only in aggregate form. Additionally, the census does not collect data on wages, income or consumption.

of an individual's employment status: household income and household consumption. We perform this latter analysis using male heads of households to avoid counting the same household multiple times. Table 2 presents the descriptive statistics for the male and female wage samples, and the male household head sample.

5 The Returns to English-Language Skills

5.1 Main Results

Using the sample of male wage earners aged 18 to 65, we estimate Equation 2. These results are shown in Table 3, with each column adding more controls.[16] Column 1 presents the raw difference in log hourly wage by English proficiency. Compared to log wages of 2.07 for men who have no English ability, men who speak a little English earn 0.69 log points more, and men who speak English fluently earn 1.37 log points more.[17] These differences are large, but likely overstate the true returns to English; for example, in Table 1, we observed that English ability is correlated with numerous variables that are known to be correlated with earnings, such as age and education.

In Column 2 of Table 3, we control for age and social group. In Column 3, we add a full set of dummies for educational attainment. Controlling for education cuts the estimated coefficients for English ability by more than half. Adding state fixed effects and a dummy for urban residence, which helps control for local institutions and labor market conditions, also leads to sizable reductions in the estimated coefficients for English ability (Column 4). In Column 5, we add dummies for SSLC exam performance and grade repetition/failure to control for ability bias (on top of what controlling for educational attainment accomplishes). Although performing in the top category (Class I) on the SSLC exam significantly increases wages, and having ever failed or repeated a grade significantly decreases wages, we find a

---

[16]We estimate all models separately for men and women. The results for women are reported in the appendix.

[17]In rupee terms, hourly wages are 10, 23 and 42 for men with no, little and fluent English.

relatively modest decline in the estimated coefficients. Finally in Column 6, we add district fixed effects, which is by far more detailed a geographic control than the state (India has 35 states/union territories which are subdivided into 628 districts) and this does not make much difference. The estimates in Column 6 suggest that, compared to having no English ability, being fluent in English increases male hourly wages by 34%, and being able to speak a little English increases wages by 13%. These results are economically meaningful effects. For example, the return to being fluent is as large as the return to completing secondary school, and half as large as the return to completing a Bachelor's degree. However, these returns are smaller than what has been estimated for the returns to English for U.S. immigrants.[18] Perhaps it is not surprising that returns to English would be higher in the U.S. than India. After all, English is the predominant language of America and necessary for participation in schools and the mainstream economy. In contrast in India, English is a minority language that is used in only a narrow segment of society.

In Appendix Table 1, we present the parallel results for women. The raw difference in wages by English proficiency are greater for women compared to men (Column 1), but once we add all the control variables, the difference becomes smaller for women (Column 6).[19] Column 6 suggests that for women, the return to speaking fluent English is 22% and the return to speaking a little English is 10%. Though the point estimates of the returns to English are lower for women, it should be noted that these estimates are imprecise. Based on pooled regressions in which we allow returns to English to vary by sex, we cannot reject that the returns for women and men are equal.

An issue with our analysis thus far is that it is restricted to people engaged in wage employment, which may be a non-random sample of the population. In particular, we may be concerned that English proficiency affects both participation in wage employment and

---

[18]Bleakley and Chin (2004) find that immigrants (both male and female) who speak English very well earn 67% (100%) more, and those who speak English not well earn 22% (33%), compared to immigrants who do not speak English at all according to their OLS (2SLS) estimates.

[19]In rupee terms, hourly wages are 5.5, 19 and 32 for women with no, little and fluent English.

wages such that our estimates may not give the unconditional effect of English proficiency on wages. We address this issue by looking at household income and consumption as outcomes; we have measures of income and consumption for all households regardless of members' participation in wage employment. We present this analysis in Table 4, which is restricted to male household heads. In Column 1, we repeat estimate the same specification as in Table 3, Column 6 to demonstrate that very similar effects on wages are found when we use the sample of male household heads. In Column 3, we use household income as the outcome. We find that in households where the head speaks English fluently, annual income is 34% higher. In households where the head speaks a little English, annual income is 11% higher. In Column 5, we use household consumption as the outcome and also find comparable, though generally smaller, effects; the smaller effects could be due to richer households saving more. These results using income and consumption thus are consistent with the wage results, and suggest that sample selection bias does not seriously bias our wage analysis.

In the even columns of Table 4, we control for head's father's education as a further method to address concerns about ability bias.[20] This does not change the estimated returns to English ability much.

5.2   Heterogeneity in Returns to English

Thus far, we have focused on estimating the average effect of English skills on the wages of men and women. In Table 5, we allow the effect English skills to vary by various observable characteristics for men (Appendix Table 2 shows parallel results for women).

First, we allow the effects to vary by urban/rural residence. To the extent that jobs rewarding English tend to be located in cities (where, say, higher levels of government, multinational firms, or information-technology firms are located), all else equal, we might expect higher returns in urban areas. However, depending on labor mobility and the relative

---

[20]Father's education is only asked of household heads so specifications controlling for this variable cannot be estimated using the broader sample.

supply of and demand for workers with English skill in urban areas, urban areas may not have a higher return to English. We find that there is no differential return to English by sector of residence for males (Table 5, Column 2). In contrast, we find a stark differential return by sector for females (Appendix Table 2, Column 2). Women living in rural areas do not receive a premium for their English-language skills while women living in urban areas command premiums that are at least as large as men's. A possible explanation is that women have much less mobility than men. A woman's location of residence is typically decided by her parents (for single women) or husband (for married women), and household responsibilities and social norms constrain women from commuting long distances for work. Men, on the other hand, can relocate to where an attractive job exists or commute there (e.g., live in the village and work in town). In urban areas, given the transportation and density of employers, women with English skills are able to find a job that matches their skills. In rural areas, though, a good job match is less possible for women.

Second, we allow effects to vary by social group. In Table 5, Column 3, we find that the returns to English are significantly lower for SCs relative to higher castes, and no different for STs and OBCs relative to higher castes.[21] Speaking English fluently increases the wages of high-caste men by 39% and SC men by only 21%. The wage premium for speaking a little English is 15% for high-caste men but only 9% for SC men. Lower caste women also receive a significantly lower return on English skill (Appendix Table 2, Column 3). In fact, women who are members of SCs and OBCs have returns to English skill that are not statistically different from zero. In contrast, high-caste women earn 30% more when they speak a little English and 33% when they speak fluent English. These results suggest that upward mobility does not come automatically with English skills in India; some obstacles, which likely include long-rooted discrimination against low castes, impede low caste group members even when they have a skill that is valued by the modern labor market.

---

[21]The interactions with STs are imprecisely estimated, though, because STs are only a small share (about 10%) of the population and therefore our sample.

Third, we allow effects to vary by age/cohort; given that we have a single cross-section, we cannot distinguish between age and cohort effects. We find that older cohorts receive significantly higher returns for English skills. This holds for both men (Table 5, Column 4) and women (Appendix Table 2, Column 4). The youngest workers in the sample, who are aged 18-35, receive only a modest increase in wages when they are fluent in English (11% for men and 6% for women, with the latter estimate insignificant at conventional levels), and actually a reduction in wages when they speak a little English relative to speaking no English at all. Men aged 36-50 have returns of 49% for fluent English and 30% for a little English. Men aged 51-65 have returns of 68% for fluent English and 54% for a little English. Although demand for workers with English skill has expanded in recent decades due to the growth of trade and outsourcing, the supply of workers with such skill must also have grown rapidly lest we would have observed a rising English skill price. Of course the question arises as to why the skill price is not equalized across the cohorts, i.e., with the influx of young workers speaking English into the labor market, why aren't the wages of older workers bid down? Consistent with these empirical observations is a story where there are language-skill complementarities. For example, English skill enables workers to enter more lucrative career tracks, i.e., these are jobs that have better promotion (or career progress) and raise opportunities. During his tenure, an older vintage worker would have acquired significant work experience. Although a younger worker possesses the English skills that would have been necessary to obtain the entry-level position decades ago, he lacks the work experience, making him unable to substitute for the older vintage worker today.

Fourth, we allow the returns to English to vary by educational attainment. There is evidence of higher returns to English for the more educated. For men who have not completed their secondary schooling (i.e., those with 0-9 years of completed schooling), being fluent in English raises wages 24% and speaking a little English raises earnings 7% (Table 5, Column 5). More educated men earn 9% more for speaking a little English, and over 12% more for being fluent in English. For women, there is also a positive interaction between educational

attainment and returns to English, however it is interesting to note that women who have not completed their secondary schooling do not earn any premium for their English skill (Appendix Table 2, Column 5). Rather, only educated women earn positive returns on their English skills, with those possessing a Bachelor's degree receiving the highest returns. These results by education are consistent with English being valued in jobs that are relatively high-skilled.

Finally we allow the returns to vary by both cohort and education. Understandably, the results are imprecise given the number of interaction terms in the model. To make the model more tractable, we just use two age categories instead of three: young (age 18-35) and older (age 36-65). The results are summarized in Figure 2 (which is derived from the coefficients reported in Table 5, Column 6). Perhaps the most striking feature of Figure 2 is that the education profile for returns to English is upward-sloping one for younger men but basically flat for older men. Men aged 18-35 who have not completed secondary school do not receive any premium for their English skill. More educated young men, though, do earn significant returns to English: those who complete secondary school have a 26% return for being fluent in English and a 5% return for speaking a little English, and those who complete a Bachelor's degree earn 40% and 17% more, respectively. For older cohorts, even low educated men aged 36-65 receive a sizable English premium–it is 43% for being fluent in English, and 22% for speaking a little English (which are not so different from the premia for college-educated young men!). For women, the results are quite a bit noisier due to the number observations in some cells, but we can also see that the premium rises more with education for younger cohorts (Appendix Table 2, Column 6).[22]

These findings of Column 6 help put the earlier results where we allow effects to vary by age alone and by education alone into greater perspective. Earlier (in Column 4), we

---

[22]In Column 7, we present the results of estimating a less demanding model that combines the two lower education groups, leaving us with only two education categories: Bachelor's degree and less. These results show more readily than Column 6 that the older women, and young women with more education, earn sizable returns to English.

had found larger returns to English for the older cohorts, with the youngest cohort earning only modest returns. We found in Column 5 that the return to English was increasing in educational attainment, though often the interaction effects were weak. It is possible, though, that some of the age interactions are due to differences in education by age and vice versa. Column 5 addresses this possibility, and reveals that returns are higher for older cohorts, and the subset of the young who are more educated.

Berman, Lang and Siniver (2003) and Lang and Siniver (2006) find evidence of language-skill complementarity in the Israeli context. Immigrant workers in high-skilled occupations received larger wage increases when they learned Hebrew and English (on top of their own native language, Russian). In India, we also find evidence of language-skill complementarity. On the one hand, more experienced workers receive a higher return for English skill; this is based on the observation that older workers have a higher English premium than younger workers. On the other hand, for recent cohorts, more educated workers receive a higher return for English skill. It is interesting that such a language-education complementarity was not found for older cohorts. This is consistent with increasing complementarity between education and English skills over time. For example, at the entry level, workers with English skill may have been able to find a good job decades ago but now only the subset with more education would find a good job. This could be because it has become more competitive to get good jobs (because the supply of educated workers has expanded so much), or because there are new jobs that require both higher education as well as English skill to perform.

6   Conclusion

In India, the raw difference in earnings between people who speak English and people who do not is large, but this overstates the economic value of English because higher ability people are more likely to be proficient in English. In this paper, we take advantage of a rich data set to avoid this ability bias and find that there are large, statistically significant returns

20

to English-language skills in India. Wages are on average 32% higher for men who speak fluent English and 13% higher for men who speak a little English relative to men who speak no English. For women, the average returns is 22% for fluent English and 10% for a little English. There is considerable heterogeneity in returns to English-language skills. More experienced and more educated workers receive higher returns to English-language skills. The complementarity between English skills and education appears to have strengthened over time. Only the more educated among young workers earn a premium for English skill, whereas older workers across all education groups do.

In India and many other developing countries, there is active debate over whether to promote the local language or a more globally accepted language like English in schools. While promoting the local language might make primary schooling more accessible and strengthen national identity, it may reduce economic opportunities because of the special role of English in the global economy. We do find that English proficiency raises wages in India, but the returns are considerably lower for more recent entrants into the labor market. For the marginal worker deciding whether to invest in English skill, the relevant return would be the one estimated for the recent entrants. Thus the implication is that providing English classes to adults may not necessarily raise their wages. For recent entrants, English skills help increase wages only when coupled with high education; those who have not completed their secondary schooling would not see wage increases due to acquisition of English-language skills. Policymakers should be aware of these language-skill complementarities when designing policies. For example, English programs for children in schools–which would be in time to influence their educational attainment–would be more effective than adult English classes.

This study is the first to use a nationally representative sample to estimate the economic returns to English-language skills. Besides providing estimates of the average wage differences by English proficiency, we uncovered some patterns in the returns that were previously undocumented for India and worth further exploration, such as the language-skill complementarity, lower returns for scheduled castes compared to other social groups, and lower

returns for women in rural areas compared to urban areas. Another area for future work is the impact of English-language skills on social outcomes; the effects of English-language skills can extend beyond the labor market, and in ways that dramatically affect society.

References

Angrist, J., A. Chin and R. Godoy. "Is Spanish-Only Schooling Responsible for the Puerto Rican Language Gap?." *Journal of Development Economics*, February 2008, *85*, pp. 105–108.

Angrist, J. and V. Lavy. "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco." *Journal of Labor Economics*, January 1997, *15* (1), pp. S48–76.

Berman, E., K. Lang, and E. Siniver. "Language-Skill Complementarity: Returns to Immigrant Language Acquisition." *Labour Economics*, 2003, (10), pp. 265–290.

Bleakley, H. and A. Chin. "Language Skills and Earnings: Evidence from Childhood Immigrants." *Review of Economics and Statistics*, May 2004, *86* (2) pp. 481–496.

Card, D. "Chapter 30: The Causal Effect of Education on Earnings." *Handbook of Labor Economics, Volume 3*, in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*. Amsterdam: Elsevier Science, North-Holland, 1999, pp. 1801–1863.

Clingingsmith, D. "Industrialization, Bilingualism, and Linguistic Heterogeneity in Mid-20th Century India." Case Western Reserve University unpublished paper, November 2008.

Desai, S, R. Vanneman, and National Council of Applied Economic Research, New Delhi . India Human Development Survey (IHDS), 2005 [Computer file]. ICPSR22626-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-06-22.

Dustmann, C. and A. van Soest. "Language and the Earnings of Immigrants." *Industrial and Labor Relations Review*, 2002, *55* (3), pp. 473–492.

He, F., L. Linden and M. MacLeod. "How to Teach English in India: Testing the Relative Productivity of Instruction Methods within the Pratham English Language Education Program." Columbia University unpublished paper, July 2008.

Hohenthal, A. "English in India: Loyalty and Attitudes." *Language in India*, May 2003, *3.*

Jensen, Robert T. "Economic Opportunities and Gender Differences in Human Capital: Experimental Evidence for India." *NBER Working Paper 16021*, 2010, National Bureau of Economic Research, Cambridge, MA.

Kapur, S. and T. Chakraborty "English Language Premium: Evidence from a Policy Experiment in India." Washington University in St. Louis unpublished paper, September 2008.

Lang, K. and E. Siniver. "The Return to English in a Non-English Speaking Country: Russian Immigrants and Native Israelis in Israel." *NBER Working Paper 12464*, 2006, National Bureau of Economic Research, Cambridge, MA.

Levinsohn, J. "Globalization and the Returns to Speaking English in South Africa," in Ann Harrison, ed., *Globalization and Poverty.* Chicago: University of Chicago Press, 2007, pp. 629–646.

Munshi, K. and M. Rosenzweig. "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy." *BREAD Working Paper 038*, 2003, Bureau for Research in Economic Analysis of Development.

Munshi, K. and M. Rosenzweig. "Why is Mobility in India So Low? Social Insurance, Inequality and Growth." Brown University unpublished paper, March 2009.

Munshi, K. and M. Rosenzweig. "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy." *American Economic Review*, September 2006, *96* (4), pp. 1225–1252.

Oster, E. and M.B. Millett. "Do Call Centers Promote School Enrollment? Evidence from India." *NBER Working Paper 15922*, 2010, National Bureau of Economic Research, Cambridge, MA.

Roy, J. "Can Changing Academic Standards Affect Educational Outcomes? Evidence from a Policy Experiment in India." Georgetown University unpublished paper, 2004.

Saiz, A. and Zoido, E. "Listening to What the World Says: Bilingualism and Earnings in the United States." *Review of Economics and Statistics*, January 2005, *87* (3), pp. 523–538

Shastry, G.K. "Human Capital Response to Globalization: Education and Information Technology in India." University of Virginia unpublished paper, October 2008.

TABLE 1

Adult English-Speaking Ability in India

Variable Name	Converses fluently in English	Converses a little in English (but not fluent)	Total with English ability
All Individuals Aged 18-65	3.79	16.23	20.03
	(0.19)	(0.36)	(0.40)
By Sex:			
Male	4.96	20.96	25.92
	(0.21)	(0.40)	(0.43)
Female	2.62	11.53	14.16
	(0.15)	(0.31)	(0.34)
By Age:			
18-35	4.35	20.21	24.56
	(0.25)	(0.40)	(0.43)
36-50	3.25	12.80	16.05
	(0.17)	(0.33)	(0.36)
51-65	2.95	9.70	12.65
	(0.16)	(0.29)	(0.33)
By Educational Attainment:			
No Completed Schooling	0.00	0.03	0.03
	(0.00)	(0.01)	(0.01)
Some Primary (1-4)	0.07	1.14	1.21
	(0.02)	(0.10)	(0.10)
Primary (5-9)	0.47	10.33	10.81
	(0.06)	(0.30)	(0.31)
Secondary (10-14)	5.98	49.52	55.50
	(0.23)	(0.49)	(0.49)
College Graduate (>= 15)	35.34	53.20	88.54
	(0.47)	(0.49)	(0.31)
By Social Group:			
Scheduled Tribes	2.53	8.37	10.91
	(0.15)	(0.27)	(0.31)
Scheduled Castes	1.54	11.37	12.91
	(0.12)	(0.31)	(0.33)
Other Backward Castes	2.93	14.69	17.63
	(0.16)	(0.35)	(0.38)
Higher Castes	7.57	25.98	33.55
	(0.26)	(0.43)	(0.47)
By Geography:			
Urban	9.00	25.83	34.84
	(0.28)	(0.43)	(0.47)
Rural	1.68	12.37	14.06
	(0.12)	(0.32)	(0.34)

Notes: The sample consists of individuals aged 18-65 from the India Human Development Survey, 2005.
Standard deviations are shown in parentheses below the mean percent of given subpopulation with English ability.

Table 2. Summary Statistics

	Men in wage employment (1)	Male household heads (2)	Women in wage employment (3)
Log hourly wage last year	2.250	2.308	1.620
	(0.763)	(0.800)	(0.679)
Log household income last year		10.322	
		(0.980)	
Log monthly per capita consumption last year		6.497	
		(0.678)	
Fluent English	0.046	0.040	0.026
	(0.210)	(0.197)	(0.159)
Little English	0.165	0.167	0.060
	(0.372)	(0.373)	(0.238)
Urban residence	0.292	0.291	0.178
	(0.455)	(0.454)	(0.382)
Age	36.748	44.157	36.366
	(11.730)	(10.998)	(11.286)
Social group: scheduled tribes	0.095	0.083	0.152
	(0.293)	(0.276)	(0.359)
Social group: scheduled castes	0.271	0.223	0.313
	(0.445)	(0.416)	(0.464)
Social group: other backward castes	0.338	0.354	0.355
	(0.473)	(0.478)	(0.479)
Social group: Muslims	0.110	0.115	0.058
	(0.313)	(0.319)	(0.233)
Years of schooling completed: 0	0.297	0.316	0.677
	(0.457)	(0.465)	(0.468)
Years of schooling completed: 1	0.006	0.006	0.004
	(0.079)	(0.075)	(0.061)
Years of schooling completed: 2	0.023	0.026	0.018
	(0.151)	(0.160)	(0.133)
Years of schooling completed: 3	0.030	0.034	0.021
	(0.170)	(0.181)	(0.143)
Years of schooling completed: 4	0.044	0.049	0.027
	(0.206)	(0.216)	(0.163)
Years of schooling completed: 5	0.085	0.089	0.055
(Primary school completion)	(0.279)	(0.285)	(0.227)
Years of schooling completed: 6	0.032	0.031	0.016
	(0.177)	(0.173)	(0.126)
Years of schooling completed: 7	0.055	0.051	0.032
	(0.227)	(0.220)	(0.177)
Years of schooling completed: 8	0.069	0.067	0.023
(Middle school completion)	(0.253)	(0.250)	(0.149)
Years of schooling completed: 9	0.086	0.071	0.023
	(0.280)	(0.257)	(0.151)
Years of schooling completed: 10	0.106	0.109	0.030
(Secondary school completion)	(0.308)	(0.312)	(0.170)
Years of schooling completed: 11	0.017	0.017	0.005
	(0.130)	(0.128)	(0.071)
Years of schooling completed: 12	0.058	0.054	0.023
(Higher secondary school completion)	(0.234)	(0.225)	(0.150)
Years of schooling completed: 13	0.002	0.002	0.002
	(0.046)	(0.047)	(0.040)
Years of schooling completed: 14	0.003	0.003	0.002
	(0.056)	(0.051)	(0.041)
Years of schooling completed: 15	0.086	0.075	0.043
(Bachelor's degree or more)	(0.280)	(0.263)	(0.202)
Performance on SSLC exam: Class I	0.056	0.048	0.033
	(0.231)	(0.214)	(0.178)
Performance on SSLC exam: Class II	0.141	0.139	0.048
	(0.348)	(0.346)	(0.214)
Performance on SSLC exam: Class III	0.049	0.046	0.014
	(0.215)	(0.210)	(0.119)
Has repeated or failed a grade	0.170	0.143	0.063
	(0.376)	(0.350)	(0.243)
Number of observations	32445	33376	12205

Notes: The sample consists of men aged 18-65 from the India Human Development Survey, 2005.
   Standard deviations are shown in parentheses below the means.

Table 3. Effect of English-Speaking Ability on Wages, Men

	No Controls (1)	Demographic Controls (2)	Education Controls (3)	Geographic Controls (4)	'Ability' Controls (5)	District Fixed Effects (6)
Fluent English	1.370***	1.201***	0.603***	0.393***	0.339***	0.345***
	(0.0230)	(0.0228)	(0.0301)	(0.0302)	(0.0317)	(0.0301)
Little English	0.690***	0.607***	0.224***	0.137***	0.125***	0.129***
	(0.0168)	(0.0160)	(0.0186)	(0.0201)	(0.0214)	(0.0186)
Age		0.041***	0.040***	0.032***	0.032***	0.031***
		(0.00242)	(0.00235)	(0.00221)	(0.00222)	(0.00212)
Age squared/100		-0.039***	-0.035***	-0.027***	-0.027***	-0.025***
		(0.00313)	(0.00304)	(0.00289)	(0.00290)	(0.00275)
Social group: scheduled castes		-0.335***	-0.210***	-0.125***	-0.118***	-0.102***
		(0.0146)	(0.0143)	(0.0133)	(0.0131)	(0.0130)
Social group: other backward castes		-0.303***	-0.228***	-0.119***	-0.116***	-0.085***
		(0.0146)	(0.0142)	(0.0136)	(0.0137)	(0.0132)
Social group: scheduled tribes		-0.529***	-0.385***	-0.237***	-0.231***	-0.167***
		(0.0172)	(0.0173)	(0.0180)	(0.0177)	(0.0175)
Social group: Muslims		-0.236***	-0.102***	-0.158***	-0.156***	-0.112***
		(0.0172)	(0.0166)	(0.0163)	(0.0162)	(0.0164)
Years of schooling completed: 1			0.069	0.0606	0.066*	0.069*
			(0.0425)	(0.0396)	(0.0397)	(0.0410)
Years of schooling completed: 2			0.041	0.037	0.041	0.056**
			(0.0268)	(0.0251)	(0.0252)	(0.0230)
Years of schooling completed: 3			0.047*	0.017	0.027	0.044*
			(0.0267)	(0.0268)	(0.0263)	(0.0231)
Years of schooling completed: 4			0.092***	0.048***	0.057***	0.075***
			(0.0188)	(0.0174)	(0.0176)	(0.0166)
Years of schooling completed: 5 (Primary school completion)			0.187***	0.123***	0.133***	0.122***
			(0.0161)	(0.0143)	(0.0142)	(0.0145)
Years of schooling completed: 6			0.228***	0.128***	0.140***	0.148***
			(0.0242)	(0.0218)	(0.0218)	(0.0210)
Years of schooling completed: 7			0.266***	0.185***	0.200***	0.191***
			(0.0209)	(0.0201)	(0.0202)	(0.0193)
Years of schooling completed: 8 (Middle school completion)			0.336***	0.217***	0.234***	0.232***
			(0.0188)	(0.0180)	(0.0185)	(0.0175)
Years of schooling completed: 9			0.352***	0.239***	0.266***	0.254***
			(0.0181)	(0.0173)	(0.0179)	(0.0167)
Years of schooling completed: 10 (Secondary school completion)			0.520***	0.377***	0.373***	0.334***
			(0.0194)	(0.0206)	(0.0605)	(0.0408)
Years of schooling completed: 11			0.360***	0.334***	0.325***	0.292***
			(0.0387)	(0.0364)	(0.0638)	(0.0513)
Years of schooling completed: 12 (Higher secondary school completion)			0.579***	0.467***	0.435***	0.410***
			(0.0265)	(0.0264)	(0.0593)	(0.0453)
Years of schooling completed: 13			0.621***	0.529***	0.496***	0.518***
			(0.105)	(0.0997)	(0.107)	(0.106)
Years of schooling completed: 14			0.605***	0.536***	0.490***	0.474***
			(0.106)	(0.109)	(0.122)	(0.110)
Years of schooling completed: 15 (Bachelor's degree or more)			0.913***	0.796***	0.720***	0.679***
			(0.0286)	(0.0283)	(0.0607)	(0.0457)
Performance on SSLC exam: Class I					0.227***	0.236***
					(0.0687)	(0.0494)
Performance on SSLC exam: Class II					0.065	0.070
					(0.0640)	(0.0456)
Performance on SSLC exam: Class III					-0.094	-0.063
					(0.0646)	(0.0465)
Has repeated or failed a grade					-0.061***	-0.069***
					(0.0155)	(0.0128)
Urban residence				0.413***	0.405***	0.341***
				(0.0110)	(0.0108)	(0.0134)
State fixed effects	No	No	No	Yes	Yes	No
District fixed effects	No	No	No	No	No	Yes
R-squared	0.231	0.299	0.360	0.455	0.459	0.516
Number of observations	32445	32445	32445	32445	32445	32445

Notes: The sample consists of men aged 18-65 who reported wage and salary work from the India Human Development
Survey, 2005. The dependent variable is log hourly wage. The omitted social group are the high castes (88% of this omitted group),
Christians, Sikhs and Jains. The omitted educational attainment group is zero years of schooling completed.
Robust standard errors are shown in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01).

Table 4. Effects of English-Speaking Ability for Male-Headed Households

	Dependent Variable:					
	Household Head's Log Hourly Wage		Household's Log Total Income Last Year		Household's Log Monthly Per Capita Expenditure	
	District Fixed Effects (1)	Add Father's Education (2)	District Fixed Effects (3)	Add Father's Education (4)	District Fixed Effects (5)	Add Father's Education (6)
Fluent English	0.330***	0.312***	0.337***	0.320***	0.257***	0.235***
	(0.0378)	(0.0380)	(0.0365)	(0.0364)	(0.0231)	(0.0232)
Little English	0.163***	0.160***	0.113***	0.106***	0.103***	0.098***
	(0.0238)	(0.0237)	(0.0240)	(0.0239)	(0.0140)	(0.0142)
R-squared	0.556	0.557	0.369	0.373	0.468	0.473
Number of observations	21057	20672	32916	32272	33340	32688

Notes: The sample consists of male heads of household aged 18-65 from the India Human Development Survey, 2005. Each column reports the results of a separate regression that also controls for an age quadratic, social group dummies, education dummies, urban dummy, district dummies, exam performance dummies, and dummy for failing or repeating a grade (i.e., the same specification as in Table 3, Column 6). The even columns also control for a full set of dummies for father's years of schooling completed. Robust standard errors are shown in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01).

Table 5. Heterogeneity in Effect of English-Speaking Ability on Log Hourly Wages, Men

	Base	Urban/Rural	Social Group	Age	Education	Age-Education	Age-Education
Variable 1:		Urban	ST	36-50	BA	Age 36-65	Age 36-65
Variable 2:			SC	51-65	Secondary	BA	BA
Variable 3:			OBC			Secondary	
Omitted group:		*Rural*	*Higher castes*	*Age 18-35*	*< 10 yrs schooling*	*Lower educ & younger*	*Lower educ & younger*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fluent English	0.345***	0.361***	0.395***	0.113***	0.241**	0.041	0.134**
	(0.0301)	(0.0523)	(0.0345)	(0.0369)	(0.108)	(0.172)	(0.0523)
Little English	0.129***	0.135***	0.150***	-0.074***	0.069***	-0.039	-0.060***
	(0.0186)	(0.0256)	(0.0232)	(0.0180)	(0.0245)	(0.0249)	(0.0188)
Fluent English x Variable 1		-0.028	-0.042	0.376***	0.123	0.391*	0.401***
		(0.0537)	(0.113)	(0.0439)	(0.121)	(0.217)	(0.0719)
Little English x Variable 1		-0.015	0.057	0.369***	0.095	0.263***	0.406***
		(0.0286)	(0.0643)	(0.0267)	(0.0606)	(0.0453)	(0.0301)
Fluent English x Variable 2			-0.180**	0.565***	0.155	0.360*	0.267***
			(0.0708)	(0.0554)	(0.115)	(0.186)	(0.0881)
Little English x Variable 2			-0.094***	0.611***	0.091***	0.210***	0.231***
			(0.0322)	(0.0486)	(0.0352)	(0.0733)	(0.0714)
Fluent English x Variable 3			-0.075			0.223	
			(0.0500)			(0.179)	
Little English x Variable 3			-0.011			0.087**	
			(0.0351)			(0.0357)	
Fluent English x BA x Age 36-65						-0.511**	-0.523***
						(0.241)	(0.126)
Little English x BA x Age 36-65						-0.292**	-0.436***
						(0.114)	(0.109)
Fluent English x Secondary School x Age 36-65						-0.267	
						(0.231)	
Little English x Secondary School x Age 36-65						-0.0749	
						(0.0642)	

Notes: The sample consists of men aged 18-65 who reported wage and salary work from the India Human Development Survey, 2005 (32,445 observations). The dependent variable is log hourly wage. Each column reports the results of a separate regression that also controls for an age quadratic, social group dummies, education dummies, urban dummy, district dummies, exam performance dummies, and dummy for failing or repeating a grade (i.e., the same specification as in Table 3, Column 6). Columns 4, 6 and 7 additionally control for a dummy for being age 36-50 and one for being 51-65. Columns 6 and 7 also allow effects of education to vary by the age groups. Robust standard errors are shown in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01).

Appendix Table 1. Effect of English-Speaking Ability on Wages, Women

|	No Controls (1)	Demographic Controls (2)	Education Controls (3)	Geographic Controls (4)	'Ability' Controls (5)	District Fixed Effects (6)
Fluent English	1.619***	1.526***	0.549***	0.322***	0.259***	0.223***
	(0.0705)	(0.0693)	(0.110)	(0.0849)	(0.0803)	(0.0775)
Little English	0.916***	0.892***	0.217**	0.092	0.081	0.100**
	(0.117)	(0.0989)	(0.0953)	(0.0670)	(0.0560)	(0.0511)
Age		0.021***	0.024***	0.015***	0.0136***	0.0129***
		(0.00681)	(0.00419)	(0.00348)	(0.00331)	(0.00294)
Age squared/100		-0.019**	-0.019***	-0.011**	-0.010**	-0.010***
		(0.00818)	(0.00523)	(0.00443)	(0.00425)	(0.00382)
Social group: scheduled castes		-0.241***	-0.140***	-0.077***	-0.0747***	-0.066***
		(0.0390)	(0.0274)	(0.0227)	(0.0225)	(0.0219)
Social group: other backward castes		-0.282***	-0.211***	-0.114***	-0.113***	-0.077***
		(0.0279)	(0.0260)	(0.0220)	(0.0219)	(0.0217)
Social group: scheduled tribes		-0.283***	-0.183***	-0.093***	-0.090***	-0.061**
		(0.0301)	(0.0278)	(0.0239)	(0.0238)	(0.0242)
Social group: Muslims		-0.220***	-0.149***	-0.204***	-0.203***	-0.183***
		(0.0351)	(0.0339)	(0.0312)	(0.0311)	(0.0314)
Years of schooling completed: 1			0.225**	0.159*	0.173*	0.116
			(0.101)	(0.0958)	(0.0954)	(0.0962)
Years of schooling completed: 2			0.086*	0.073*	0.079*	0.069*
			(0.0486)	(0.0420)	(0.0423)	(0.0386)
Years of schooling completed: 3			0.076**	0.02	0.038	0.039
			(0.0358)	(0.0268)	(0.0284)	(0.0249)
Years of schooling completed: 4			-0.022	-0.069**	-0.043	-0.044
			(0.0352)	(0.0315)	(0.0327)	(0.0296)
Years of schooling completed: 5 (Primary school completion)			0.127***	0.080**	0.104***	0.082***
			(0.0433)	(0.0315)	(0.0274)	(0.0275)
Years of schooling completed: 6			0.165***	0.078**	0.099**	0.113***
			(0.0409)	(0.0390)	(0.0398)	(0.0369)
Years of schooling completed: 7			0.156***	0.093***	0.114***	0.081***
			(0.0370)	(0.0323)	(0.0330)	(0.0307)
Years of schooling completed: 8 (Middle school completion)			0.275***	0.142***	0.173***	0.142***
			(0.0431)	(0.0385)	(0.0381)	(0.0358)
Years of schooling completed: 9			0.208***	0.124***	0.178***	0.166***
			(0.0395)	(0.0347)	(0.0380)	(0.0372)
Years of schooling completed: 10 (Secondary school completion)			0.568***	0.404***	0.385***	0.287***
			(0.0669)	(0.0570)	(0.0822)	(0.0769)
Years of schooling completed: 11			0.456***	0.407***	0.390***	0.389***
			(0.124)	(0.112)	(0.132)	(0.122)
Years of schooling completed: 12 (Higher secondary school completion)			0.806***	0.627***	0.562***	0.513***
			(0.0926)	(0.0723)	(0.103)	(0.0970)
Years of schooling completed: 13			0.579***	0.397***	0.378**	0.309*
			(0.189)	(0.142)	(0.166)	(0.168)
Years of schooling completed: 14			0.941***	0.818***	0.733***	0.672***
			(0.193)	(0.188)	(0.198)	(0.178)
Years of schooling completed: 15 (Bachelor's degree or more)			1.198***	1.005***	0.898***	0.841***
			(0.102)	(0.0770)	(0.103)	(0.0989)
Performance on SSLC exam: Class I					0.248***	0.239***
					(0.0958)	(0.0925)
Performance on SSLC exam: Class II					0.058	0.040
					(0.0897)	(0.0847)
Performance on SSLC exam: Class III					-0.059	-0.031
					(0.0930)	(0.0914)
Has repeated or failed a grade					-0.115***	-0.080***
					(0.0297)	(0.0275)
Urban residence				0.390***	0.385***	0.300***
				(0.0190)	(0.0189)	(0.0242)
State fixed effects	No	No	No	Yes	Yes	No
District fixed effects	No	No	No	No	No	Yes
R-squared	0.237	0.269	0.325	0.437	0.440	0.550
Number of observations	12205	12205	12205	12205	12205	12205

Notes: The sample consists of women aged 18-65 who reported wage and salary work from the India Human Development Survey, 2005. The dependent variable is log hourly wage. The omitted social group are the high castes (89% of this omitted group), Christians, Sikhs and Jains. The omitted educational attainment group is zero years of schooling completed. Robust standard errors are shown in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01).

Appendix Table 2. Heterogeneity in Effect of English-Speaking Ability on Log Hourly Wages, Women

|	Base	Urban/Rural	Social Group	Age	Education	Age-Education	Age-Education
Variable 1:		Urban	ST	36-50	BA	Age 36-65	Age 36-65
Variable 2:			SC	51-65	Secondary	BA	BA
Variable 3:			OBC			Secondary	
Omitted group:		*Rural*	*Higher castes*	*Age 18-35*	*< 10 yrs schooling*	*Lower educ & younger*	*Lower educ & younger*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fluent English	0.223***	0.019	0.335***	0.061	-0.022	0.429**	-0.116
	(0.0775)	(0.134)	(0.0818)	(0.0857)	(0.458)	(0.175)	(0.106)
Little English	0.100**	0.002	0.301***	-0.099**	-0.013	-0.120	-0.093*
	(0.0511)	(0.0583)	(0.0662)	(0.0494)	(0.0771)	(0.0760)	(0.0523)
Fluent English x Variable 1		0.353**	-0.064	0.358***	0.463	-0.846	0.602***
		(0.140)	(0.144)	(0.0956)	(0.478)	(0.718)	(0.185)
Little English x Variable 1		0.244***	0.014	0.503***	0.325*	0.294*	0.533***
		(0.0670)	(0.122)	(0.0726)	(0.169)	(0.176)	(0.0893)
Fluent English x Variable 2			-0.255	0.753***	0.229	0.121	0.663***
			(0.167)	(0.135)	(0.466)	(0.239)	(0.194)
Little English x Variable 2			-0.382***	0.978***	0.161*	0.410**	0.381**
			(0.0836)	(0.114)	(0.0968)	(0.178)	(0.168)
Fluent English x Variable 3			-0.098			-0.408**	
			(0.132)			(0.197)	
Little English x Variable 3			-0.303***			0.187*	
			(0.0769)			(0.100)	
Fluent English x BA x Age 36-65						0.450	-0.997***
						(0.770)	(0.332)
Little English x BA x Age 36-65						-0.355	-0.591**
						(0.328)	(0.289)
Fluent English x Secondary School x Age 36-65						0.790***	
						(0.256)	
Little English x Secondary School x Age 36-65						1.001	
						(0.746)	

Notes: The sample consists of women aged 18-65 who reported wage and salary work from the India Human Development Survey, 2005 (12,205 observations).
The dependent variable is log hourly wage. Each column reports the results of a separate regression that also controls for an age quadratic,
social group dummies, education dummies, urban dummy, district dummies, exam performance dummies, and dummy for failing or repeating a grade
(i.e., the same specification as in Appendix Table 1, Column 6). Columns 4, 6 and 7 additionally control for a dummy for being age 36-50 and one for being 51-65.
Columns 6 and 7 also allow effects of education to vary by the age groups. Robust standard errors are shown in parentheses.
Asterisks denote significance levels (*=.10, **=.05, ***=.01).

Figure 1. English-Speaking Ability Across the Indian States
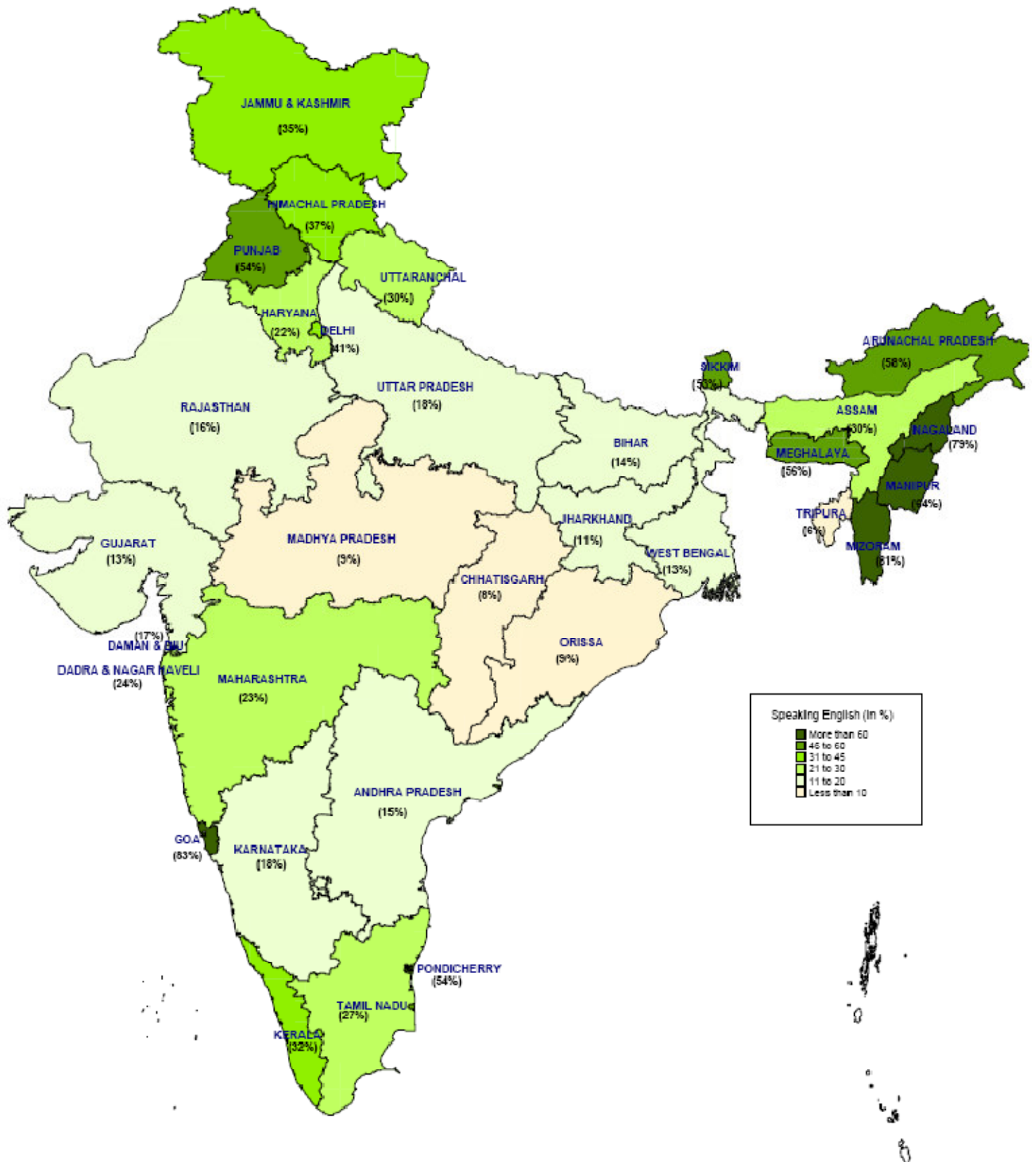
Notes: Reported in each state is the mean percentage speaking English (either fluently or a little). These means are calculated from data on individuals aged 18-65 from the India Human Development Survey, 2005.

Figure 2. Summary of Returns to English by Age and Education, Men
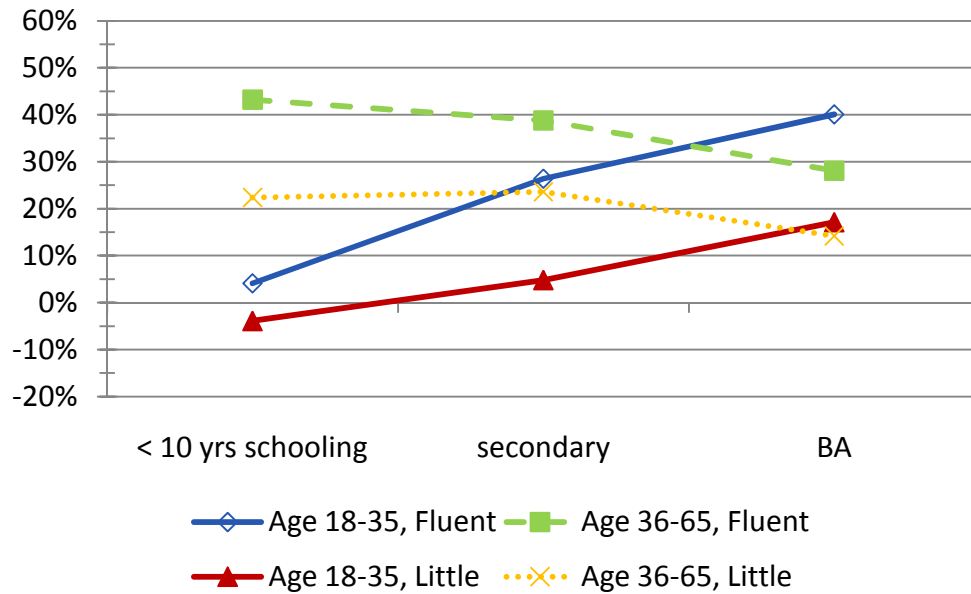
Notes: The returns are calculated based on the coefficients reported in Table 5, Column 6.