

The labor market benefits of bilingualism: The case of Costa Rica

Alejandro Abarca

Observatory of Development, University
of Costa Rica

Suráyabi Ramírez

Observatory of Development, University
of Costa Rica

June, 2017

Abstract

This paper estimates the job market benefits of bilingualism in Costa Rica for the 2005-2017 period. Our estimates show that for all wage earners, English speakers have a monthly wage that is 20.6% higher than their non-English-speaking counterparts. While for private sector workers, the premium estimated is 28.4%. Furthermore, this wage premium does not have a declining trend overtime, as one would expect in competitive labor markets. This suggests that the Costa Rican labor market has not automatically adjusted to the relative scarcity of English speaking labor. In addition, we found that English speakers work less hours per week, have a higher chance of getting paid vacations, paid sick leave and a thirteenth wage at the end of the year with respect to non-English speakers.

Introduction

In this paper, we study the job market benefits of bilingualism for Costa Rica, an open small economy with high development indicators, and one of the most politically stable countries in Latin America. Over recent decades, the country has changed its growth strategy from an import substitution model in the sixties and seventies, to an open economy that has made foreign direct investment (FDI) the principal engine of the country's economic dynamism. As a result of this new development path, all governments had invariably committed to attracting knowledge-intensive firms and to fostering the tourism sector that mostly demand English-speaking labor¹. Hence, these relatively recent labor demands have required training of the labor force through different educational public policies (OECD, 2012). In addition, the main commercial partners of Costa Rica have not significantly changed over time, but the requirement of foreign language skill has. Thus, the demand of foreign language -especially English- coincides with the period when the country made foreign direct investment (FDI) the principal engine of the country's economic dynamism and fostered its integration into the world economy through the attraction of knowledge-intensive firms in the manufacturing, services and tourism sectors.

Furthermore, contrary to Europe or the United States, most of the immigration that Costa Rica experiences is from Central American countries -that share Spanish as an official language-, so the cultural diversity that migration creates, does not contribute significantly to the generation of bilingual abilities within the population. Therefore, any job market benefit associated with proficiency in English is mainly due to labor market demands and skill compensations. In addition, we show that being fluent in English is not an endogenous variable with respect to the short-term labor market dynamics or the state of the economy. These features make Costa Rica a suitable country to measure the labor market benefits of speaking English fluently as a second language in a small developing open economy.

The bilingualism job market premium research has covered topics such as i) measuring the returns to immigrants of learning the host-country language (Chiswick, 1998; Dustmann and van Soest, 2002; Fry and Lowell, 2003; Bleakley and Chin, 2004; Aldashev et al., 2009) ii) the value for local workers of speaking a foreign language (Saiz and Zoido, 2005; Stöhr, 2015), iii) and the returns to different languages in multilingual countries and regions (Rendon, 2007; Cattaneo and Winkelmann, 2003; Vaillancourt et al., 2007).

For the United States, Fry and Lowell (2003) found no returns on English skills once education attainment is considered, Bleakley and Chin (2004) estimated pos-

¹English is considered the global language for international relations, science, communications and international commerce (Crystal, 2003; Ku and Zussman, 2010): in the early 2000s a third part of the world was exposed routinely to English and a quarter of the world -approximately 1.5 billion people- was competent in the language (Crystal, 2003).

itive effects of English proficiency on wages among adults who immigrated to the United States as children, while Saiz and Zoido (2005) found small salary returns to speaking a second language for native college graduates (Saiz and Zoido, 2005). Along similar lines, Vaillancourt et al. (2007) show positive effects on earnings of speaking both English and French in Quebec, while Chiswick (1998) found for Israel that speaking English is associated with higher earnings regardless of the country of origin.

For Europe, the literature has found that bilingualism accrues substantial positive effects on labor market variables for Germany (Dustmann and van Soest, 2002, Aldashev et al., 2009; Stöhr, 2015), Catalonia (Rendon, 2007), Estonia and Latvia (Toomet, 2011); and various European countries (Ginsburgh and Prieto-Rodriguez, 2011). In developing countries, a positive English skill premium has been found for Turkey (Di Paolo and Tansel, 2015), India (Azam et al., 2013), and South Africa (Casale and Posel, 2011). Along similar lines, for the indigenous Spanish speaking population in Bolivia (Chiswick et al., 2000), in spite of their bilingualism, it has been found that they are discriminated in the labor market, thus accruing a negative wage premium.

We extend on the existing literature firstly by measuring the benefits of bilingualism in the labor market from a wider perspective. To the best of our knowledge, all studies dedicated to measuring the bilingualism premium focus on estimating the effect on wages and employability, whereas we -in addition to the wage premium- estimate the change in job conditions through variables such as the total hours worked per week, whether the worker is covered by social security, receives paid vacations and sick paid leave.

Secondly, our dataset is comprised of thirteen annual household surveys that are representative at the national and regional level, which also includes a non-censored sample of workers of Costa Rica. Therefore, we can determine how these premiums have changed over time for more than a decade, while allowing us to control for various potentially confounding factors and to split samples by sex, area and levels of education attained. In a nutshell, our dataset allows for a more rigorous analysis of the bilingualism premium, since not all labor benefits are given as a salary.

Our results show that there is a highly significant wage premium associated with fluently speaking English in Costa Rica. For the whole period of study (2005-2017), we estimate the English-speaking premium to be 20.6% for all wage earners and of 28.4% for private sector wage earners. This premium ranges anywhere between a 14.3% and 25.1% for all wage earners and between 22.1% and 36.2% for private sector wage earners for any year considered in this study. Furthermore, this premium does not have a declining trend over time, as one would expect in competitive labor markets. This result is robust and significant for different model specifications and sample splits. English proficiency also lowers weekly worked hours, and raises the probability of getting paid vacations and paid sick leave.

The paper is structured as follows. The first section provides a background of English teaching in Costa Rica. Section 2 presents the data and econometric models used. Section 3 presents our results. The fourth section discusses our results and public policy implications, and also provides evidence of the exogeneity of our results. Finally, section 5 concludes.

1 A brief history of English teaching in Costa Rica

English teaching in Costa Rica was formally established in 1901, around the time when the economy was booming and many American and British companies had operations in the country. English teachers were scarce and the language was only formally taught in a few educational institutions in the country. It was not until the 1940s, that teaching of English as a second language took a significant boost. Many cultural and academic exchanges between Costa Rica and the United States took place, and English teaching schools were created. In 1944 it was enacted that schools in the country should teach English as part of their basic curriculum (Arroyo, 2013). English was primarily taught by native speakers or by Costa Ricans that had lived in an English speaking country. It was not until 1957 that the University of Costa Rica created the English teaching major and the country started to formally train English teachers (Córdoba et al., 2005).

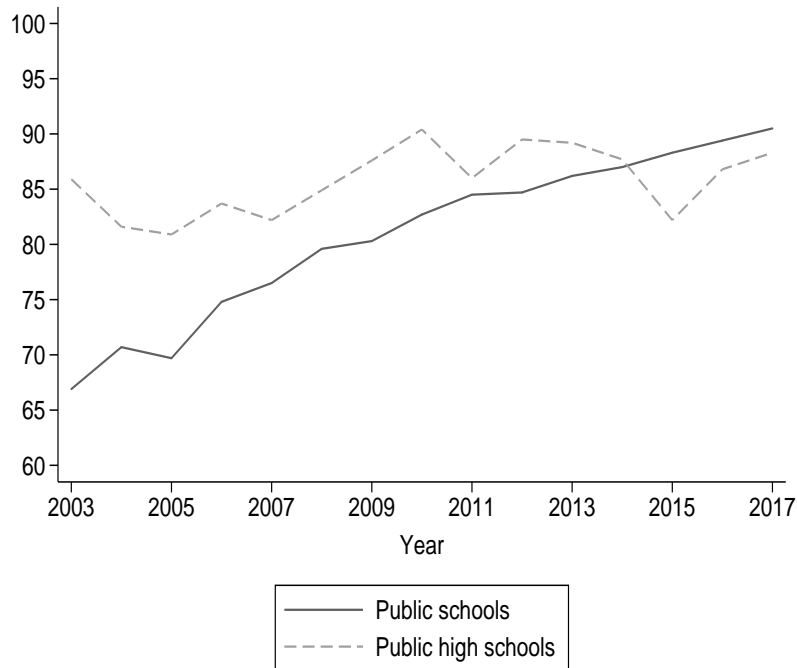
During the 1960s and 1970s, Costa Rica implemented an import substitution growth strategy which fostered economic growth, exports and industrialized the economy. Under this context, the country devoted more resources into education and enacted various educational reforms. As for English lessons, during the 1960s saw the number of hours English that was taught in schools per month changed from 15 to 18 (Arroyo, 2013). In addition, a new curriculum for English classes was created in 1972, which defined objectives, activities and evaluation templates for English teachers (Córdoba et al., 2005).

During the 1980s, educational spending was severely reduced due to the economic crisis at the beginning of this decade. Consequently, no important educational reforms were undertaken until 1995, when it was set as a policy objective that by the time students finished secondary school, they should fluently speak one or two languages other than Spanish. In 1997, English was set as a mandatory class in primary education and in 1998 it is included in the regular curriculum of preschools (Córdoba et al., 2005; Jiménez and Arce, 2014).

Unfortunately, there are no official statistics of English coverage that allow us to analyze the scope of these policies before 2003. Figure 1 below shows the percentage

of students that are taught English in the public educational system of Costa Rica. While coverage in schools has had a consistent increasing trend since 2003, coverage in high schools² has lingered around 85% for the past 15 years.

Figure 1
Percentage of students that are taught English, 2003-2017



Source: Ministry of Education, Department of Statistical Analysis (2018)

The efforts of promoting bilingualism are consistent with the economic changes the country has gone through since the 1990s. In particular, basing its economic growth and development on the promotion of international trade and the attraction of foreign direct investment (FDI) focused in the manufacturing, services and tourism sectors.

2 Data

The data comes from Costa Rica's household surveys conducted by the National Institute of Statistics and Census (INEC in Spanish) from 2005 to 2017³, which are the years where each member of the household was asked whether they speak English. Data is collected through in person interviews, and it is focused on collecting

²Up to 2008, there is only official coverage data for 10th, 11th and 12th for public high schools.

³*Encuesta de Hogares de Propósitos Múltiples* from 2005 to 2009, and *Encuesta Nacional de Hogares* from 2010 to 2017.

-among other characteristics- for all the members of the household: income, access to education and social insurance of its occupants, employment situation and working conditions. These surveys are conducted in July of each year, with a sample of over 10,000 households per year that is representative at national and regional levels. For this particular study, we focus on the household and employment characteristics of the surveys, such as income earned after taxes, wage after taxes, age, industry where each person works, whether a person’s job is in the public or private sector, level of education, zone of residence, number of members in the household, level of poverty, possession of durable consumer goods, access to public services and government programs, among many other variables.

Table 1 displays the number of total original observations and the observations used from each household survey in our samples. To construct the samples, we dropped observations out of the 25 to 65 age range, observations that have no education, workers that are firm owners, work in domestic chores or were not paid, and unemployed people that do not work because they are either retired or have a disability that prevented them from working.

Table 1
Observations per household survey

Year	Total	Wage earners	Private sector wage earners
2005	43.682	6.037	4.347
2006	45.139	6.285	4.451
2007	46.278	6.652	4.791
2008	46.101	6.647	4.882
2009	48.071	7.307	5.157
2010	41.184	6.056	4.228
2011	40.860	6.357	4.446
2012	39.390	6.253	4.435
2013	38.779	6.090	4.195
2014	38.399	6.484	4.532
2015	37.291	6.827	4.941
2016	37.006	6.848	4.987
2017	34.843	6.511	4.744
Total	501.623	84.354	60.136

Source: Authors’ estimates using household surveys 2005-2017.

Additionally, we dropped all of the workers that worked less than 30 and more than 60 hours per week in their main occupation. This range of hours corresponds to the people that are working at least three quarters of a full time and at most full time and a half per week. By limiting workers in this range we exclude people that might decide to work half time or less, and not necessarily because their job requires them to work this amount of hours. This also eliminates the possibility that some people might work very few hours because in their main occupation they have no choice but

to take a job in such conditions. Similarly, one does not know if someone who works more than 60 hours per week because said person chooses to do so or because his or her job demands it. Said people could either be exploited workers or workaholics. Therefore by bounding the hours per week worked, we minimize non-controllable biases by comparing people with more similar characteristics and incentives.

The “Wage earners” column comprises of people that earn a salary; thus, self-employed and unemployed people are not included. The last column, only considers the wage earning population that works in the private sector of the economy. This separation is analytically relevant, since private sector wages reflect more closely the market value of the language premium. Conversely, most of the salary compensations of the Costa Rican public sector are defined by strict set of rules that do not vary significantly through time.

2.1 Dependent variables

We employ 5 dependent variables on which the effect of bilingualism is estimated: wage, hours worked per week, whether the person receives a thirteenth wage⁴ at the end of the year, whether the person receives paid vacations from their work and whether the person can get sick paid leave from their work. The wage used in this paper refers to the remuneration received from the main occupation after paying for social security and personal income tax, expressed in constant US dollars of June of 2015⁵.

As for hours worked per week, each household survey reports the usual amount of hours a person dedicates to his or her primary and secondary occupation. The estimations only include the weekly hours dedicated to the main occupation. Regarding the rest of the variables, the household surveys explicitly indicate whether or not each worker gets paid vacations, sick leave and the thirteenth wage in his or her main occupation.

⁴The “thirteenth wage” (known in Spanish as *aguinaldo*), is an inalienable economic benefit that the Costa Rican legislation grants to all workers. It consists of a remuneration given once a year which is equivalent to one month’s salary. All regular employees, are entitle to receive this benefit in full if they have accumulated a year of service. While employees that have not accumulated a full labored year are entitled to receive the benefit partially, according to the months they have labored.

⁵In each household survey, this wage is presented in Costa Rican colones. We deflated each wage to express them in constant Costa Rican colones of June of 2015, which is the base year and month of the latest estimates of Costa Rica’s price consumer index. Then, we converted these constant colones to constant US dollars by using the exchange rate at the end of June of 2015.

2.2 Independent variables

The independent variable of interest is comprehensive English competence. An individual is competent exclusively when they speak, read and write English thoroughly. If the person is not proficient in one of the characteristics mentioned above, then is not considered bilingual.

A set of independent variables that control for confounding factors that affect the dependent variables of interest are used. Particularly, the sex, age, urban or rural area of residence, maximum level of education attained, the current region of residence, the region of birth and the household survey from which each observation was extracted. In addition, each model takes into account the migrant condition of each individual, where each observation is either one of the three following: i) Non-migrant, which is a person that lives in the region where he or she was born, ii) Internal migrant, which is a person that was born in a different region than the one where he or she currently resides, and iii) International migrant, the person was born in a different country. Table 2 displays the descriptive statistics for the two samples used in this research. For each sample, we show the mean value of each dependent and independent variable for bilinguals and non-bilinguals, and the difference of means for each variable along with its significance.

Table 2
Summary statistics

	All wage earners			Private sector wage earners		
	English speakers	Non-English speakers	Difference	English speakers	Non-English speakers	Difference
Dependent Variables						
Wage	1.538,37	810,50	727,87***	1.427,47	628,19	799,28***
Hourly wage	8,57	4,41	4,15***	7,80	3,30	4,51***
Hours worked per week	45,74	47,30	-1,56***	46,49	48,28	-1,79***
Thirteenth wage	94,7%	86,1%	8,5%***	92,6%	81,0%	11,6%***
Sick leave	92,5%	81,0%	11,5%***	90,2%	74,5%	15,7%***
Vacations	93,5%	82,4%	11,1%***	91,3%	76,1%	15,2%***
Personal characteristics						
Women	40,1%	33,6%	6,4%***	35,2%	27,3%	7,8%***
Age	35,89	39,40	-3,51***	34,74	38,37	-3,63***
Rural area	23,7%	46,2%	-22,5%***	23,1%	49,9%	-26,9%***
Private sector worker	68,7%	71,7%	-3,0%***			
Highest level of education attained						
Primary						
Incomplete	0,4%	11,0%	-10,5%***	0,6%	14,8%	-14,1%***
Complete	2,6%	29,9%	-27,3%***	3,4%	36,3%	-32,9%***
Secondary						
Incomplete academic	5,9%	17,2%	-11,3%***	7,9%	20,1%	-12,2%***
Incomplete technical	13,7%	13,7%	0,0%	16,8%	13,3%	3,5%***
Completed academic	0,3%	0,6%	-0,3%***	0,5%	0,7%	-0,2%**
Completed technical	2,1%	2,1%	0,0%	2,3%	1,9%	0,4%**
Tertiary						
Postsecondary education	61,8%	23,0%	38,8%***	59,6%	12,3%	47,3%***
Graduate education	13,1%	2,6%	10,5%***	8,9%	0,6%	8,3%***
Migrant condition						
Non-migrant	48,3%	52,7%	-4,4%***	46,4%	50,7%	-4,3%***
Internal migrant	44,3%	38,0%	6,3%***	44,3%	37,0%	7,2%***
International migrant	7,4%	9,3%	-1,9%***	9,3%	12,2%	-2,9%***
Region of residence						
Central	72,6%	53,1%	19,4%***	76,5%	53,6%	22,9%***
Chorotega	6,0%	10,0%	-4,0%***	5,8%	9,1%	-3,3%***
Central pacific	5,6%	8,7%	-3,2%***	5,7%	8,9%	-3,2%***
Brunca	4,4%	9,7%	-5,3%***	3,0%	8,8%	-5,8%***
Atlantic huetar	6,4%	10,3%	-3,9%***	4,4%	11,5%	-7,1%***
North huetar	5,0%	8,1%	-3,1%***	4,7%	8,1%	-3,4%***
Year of survey						
2005	6,5%	7,3%	-0,8%***	6,1%	7,4%	-1,3%***
2006	6,7%	7,6%	-0,9%***	6,3%	7,6%	-1,3%***
2007	7,4%	8,0%	-0,6%**	7,3%	8,1%	-0,8%**
2008	6,9%	8,1%	-1,2%***	6,8%	8,3%	-1,5%***
2009	8,2%	8,8%	-0,6%**	8,4%	8,7%	-0,3%
2010	6,7%	7,2%	-0,6%**	6,4%	7,1%	-0,7%**
2011	6,8%	7,6%	-0,8%***	6,5%	7,5%	-1,0%***
2012	6,4%	7,5%	-1,1%***	6,4%	7,5%	-1,1%***
2013	6,4%	7,3%	-0,9%***	6,0%	7,1%	-1,0%***
2014	8,2%	7,6%	0,6%**	8,2%	7,4%	0,8%**
2015	10,1%	7,8%	2,3%***	10,7%	7,9%	2,8%***
2016	10,2%	7,8%	2,4%***	10,8%	7,9%	2,8%***
2017	9,6%	7,4%	2,1%***	10,1%	7,6%	2,5%***

Source: Authors' estimates using household surveys 2005-2017.

Notes: ***, ** and * represent significance at 1%, 5% and 10% respectively.

In comparison to non-English speakers, English speakers have higher salaries and hourly wages, they work less hours per week, a larger percentage of them receive a thirteenth wage, and are offered sick leave and vacations in their work. With respect

to non-English speakers, bilinguals are on average younger, a higher percentage of them are women and live in urban areas, and a lower percentage of them work in the private sector.

As for education, most English speakers have attained some form of higher education, whereas most of non-English speakers have attained at most primary education. The composition of the migrant population on each sample is similar, where the majority of observations are people that live in the region where they were born. This proportion of the observations is closely followed by internal migrants, which are people that at the time of the survey live in a different region from the one where they were born. In the case of international migrants, the proportion is lower for English speakers than for non-English speakers in the two samples considered in this research. The fifth panel displays the region of residence at the time of the survey for each observation, where for each sample and group of observations, most of the observations resided in the central region of Costa Rica at the time of the survey. Lastly, the final panel displays the distribution of all of the observations by the survey from which they were extracted.

2.3 Econometric model

Our estimates are based on the following econometric model:

$$Y_{isg} = \alpha + \beta English_i + \delta X_i + \rho Migrant_i + \tau Education_i + \lambda Region_i + \psi BirthRegion_i + \kappa_s + \pi_g + \varepsilon \quad (1)$$

Where Y_{isg} are the set of dependent variables for each observation i , extracted from the household survey s and was born in the year g . Namely, our dependent variables are: the natural logarithm of the after tax wage, the number of hours worked per week, a dummy variable indicating whether or not the observation received a thirteenth wage at the end of the year, a dummy variable indicating whether or not the worker receives sick paid leave from his or her job and a dummy variable indicating whether or not his or her work grants paid vacations. $English_i$ is a dummy variable indicating whether or not the person speaks English fluently, hence β is our estimate for the benefit of knowing English.

X_i is the set of personal characteristics for each observation shown in the second panel of the summary statistics table: sex, age, type of area where the person currently lives and a dummy variable that indicates whether the person is a private sector worker or not. $Migrant_i$ is a dummy variable that indicates the migrant condition, i.e., if the person currently resides in a different region or country from they were born in. $Education_i$ are dummy variables indicating the highest level of education attained. $Region_i$ are dummy variables of the region of residence of the

observation at the time of the survey. $BirthRegion_i$ are dummy variables of the region of birth for each person. Finally, s are household survey fixed effects and g are year of birth fixed effects.

We use the natural logarithm of the after tax wage as the dependent variable to express this premium in percentage terms. In particular, the interpretation of β holds the following relationship with the percent change on Y_{isg} :

$$\% \Delta Y_{isg} = 100(e^\beta - 1) \tag{2}$$

This is the difference in the expected geometric means of the log of the after tax wage between English Speakers and non-English speakers. Our tables of results present the β as it appears in the regression models, but we provide the exact interpretation in the text.

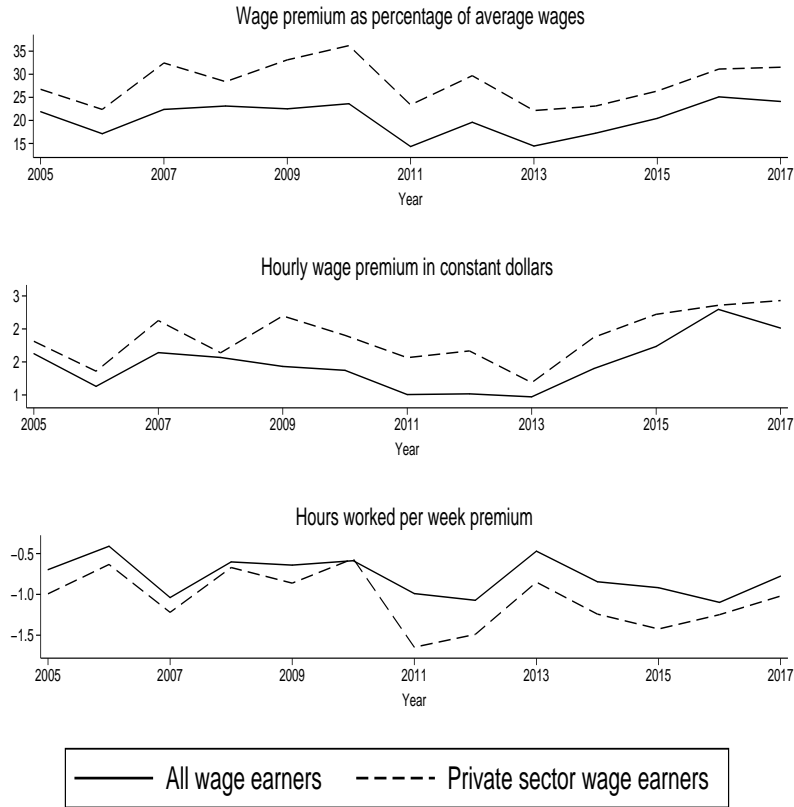
3 Results

Figure 1 displays the estimated job market premiums for each year from 2005 to 2017 for the sample of all workers that earned wages and the sample that only considers private sector wage earners for the three main dependent variables of this research. Each point in the figure represents the estimated coefficient for our dependent dichotomous variable that indicates whether a person fluently speaks English or not. All of the estimates include all the control variables discussed in the previous section and all regressions are estimated using clustered errors by household⁶.

⁶All estimates for the after tax wage premium and hourly wage are significant at a 1% significance level. For the hours worked per week premium estimates, most are also significant at a 1% significance level and others at the 5% and 10% significance levels.

Figure 2

Estimated job market premiums for bilinguals in Costa Rica, 2005-2017



Source: Author's estimates using Costa Rica's household surveys.

The estimates indicate that there are important and robust job market premiums over time. Moreover, the premium for private sector workers is higher than the premium for all workers for every year and for each variable. This comes as no surprise since civil servants are scantily rewarded for fluency in a second language in Costa Rica. When considering all wage earners, an English speaker earned on average from 14.3% to 25.1% more than a non-English speaker. For private sector workers, this premium ranged from a 22.1% to 36.2%. Along similar lines, a bilingual worker earned on average between 1 and 2.3 additional US dollars per hours with respect to non-English speakers when considering all wage earners, and between 1.1 and 2.4 additional US dollars when considering only private sector wage earners. Finally, a fluent English speaker worked on average anywhere from 0.4 to 1.1 less hours per week than their non-English speaking counterparts when considering the entire sample. Whereas for private sector workers, we found that English speaking ones worked from 0.5 to 1.6 less hours per week than non-English speakers.

In addition to the robustness and notable magnitude of the results, there is no clear

pattern for any of these premiums over time for either sample. While there are clear peaks and valleys on each series in the figure, there is no decreasing trend over time, as one would expect in competitive labor markets.

To put these wage premium results in perspective, we quantify these in US dollars in the Table 3 below. Bilinguals earn on average 68 to 221 more US dollars with respect to non-English speakers for any given year between 2005 and 2017. Furthermore, the English wage premium is between 13.6% and 44.3% of Costa Rica’s minimum wage for any given year and sample considered. This in turn, implies that the bilingual skill wage premium has been anywhere between 9% and 30.5% of average wages in Costa Rica.

Table 3
Geometric average wage and English premium in dollars by year

Year	Average wage		English premium		Premium as percentage of minimum wage	
	All sample	Private sector	All sample	Private sector	All sample	Private sector
2005	575.0	494.1	134.85	145.20	30.3%	32.7%
2006	593.9	507.0	79.35	106.51	17.8%	23.8%
2007	615.4	535.9	98.81	146.85	21.8%	32.4%
2008	623.6	543.4	153.18	204.86	33.5%	44.8%
2009	676.6	577.5	120.86	141.12	25.9%	30.2%
2010	691.8	564.8	160.57	168.02	33.2%	34.7%
2011	710.0	570.3	91.73	115.03	18.8%	23.6%
2012	723.1	589.2	221.07	209.09	44.3%	41.9%
2013	732.9	586.9	68.09	96.37	13.6%	19.2%
2014	734.1	601.1	78.78	125.03	15.3%	24.3%
2015	814.7	681.3	104.84	114.65	19.7%	21.6%
2016	816.3	676.9	107.45	161.08	20.1%	30.1%
2017	820.4	684.3	123.53	125.15	23.1%	23.4%

Source: Authors’ estimates using household surveys 2005-2017.
Notes: all wages have been converted to US dollars of June 2015.

Table 4 below shows our estimates for the entire period and all the dependent variables of interest. We find that for all wage earners, English speakers earn on average 20.6% more than non-English speakers. This premium is 28.4% for the case of private sector workers that earned a wage. By the same token, bilinguals earn more per labored hours and work less hours per week than non-bilinguals. More precisely, within the entire wage earners sample, bilinguals earn on average 1.5 more US dollars per hours and works almost an hour less per week. For private sector workers, English speakers earn on average 1.9 more US dollars per labored hour and work 1.1 less hours per week. In addition, fluent English speakers have a significantly higher probability of being granted paid vacations, paid sick leave and a thirteenth wage at the end of the year.

Table 4
General results

	Dependent Variables					
	Logarithm of wage	Hourly wage	Worked hours	Vacations	Paid sick leave	Thirteenth wage
All wage earners						
Effect	0.188***	1.540***	-0.808***	0.0256***	0.0278***	0.0148***
Robust standard error	(0.00650)	(0.0650)	(0.0735)	(0.00305)	(0.00324)	(0.00274)
Observations	84,354	84,354	84,354	84,354	84,354	84,354
Private sector workers						
Effect	0.250***	1.952***	-1.116***	0.0231***	0.0280***	0.0102**
Robust standard error	(0.00828)	(0.0727)	(0.0950)	(0.00452)	(0.00471)	(0.00412)
Observations	60,136	60,136	60,136	60,136	60,136	60,136
Controls						
Sex	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth	Yes	Yes	Yes	Yes	Yes	Yes
Type of area	Yes	Yes	Yes	Yes	Yes	Yes
Migrant condition	Yes	Yes	Yes	Yes	Yes	Yes
Private sector	Yes	Yes	Yes	Yes	Yes	Yes
Region of residence	Yes	Yes	Yes	Yes	Yes	Yes
Region of birth	Yes	Yes	Yes	Yes	Yes	Yes
Year of survey	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' estimates using household surveys 2005-2017.

Notes: ***, ** and * represent significance at 1%, 5% and 10% respectively. Clustered errors by household are shown in parentheses.

3.1 Results by zone of residence

In Table 5 we display the results by splitting the sample between individuals that currently reside in urban and rural areas. Our results are consistent with previous general results. Namely, speaking English fluently entails a significant wage premium, working less hours per week, earning more per hour labored and having better chances of getting paid vacations, paid sick leave and a thirteenth wage at the end of the year. Nonetheless, there is no pattern of the comparative magnitude of effects across dependent variables and between the types of area considered.

Table 5
Results by area

	Dependent Variables					
	Logarithm of wage	Hourly wage	Worked hours	Vacations	Paid sick leave	Thirteenth wage
All wage earners						
Urban	0.196***	1.655***	-0.735***	0.0223***	0.0255***	0.0110***
Robust standard error	(0.00757)	(0.0780)	(0.0849)	(0.00346)	(0.00369)	(0.00310)
Observations	47,690	47,690	47,690	47,690	47,690	47,690
Rural	0.184***	1.258***	-0.695***	0.0426***	0.0429***	0.0336***
Robust standard error	(0.0129)	(0.116)	(0.148)	(0.00637)	(0.00671)	(0.00576)
Observations	36,664	36,664	36,664	36,664	36,664	36,664
Private sector workers						
Urban	0.253***	2.028***	-0.972***	0.0189***	0.0262***	0.00626
Robust standard error	(0.00963)	(0.0873)	(0.109)	(0.00516)	(0.00540)	(0.00468)
Observations	32,028	32,028	32,028	32,028	32,028	32,028
Rural	0.260***	1.748***	-1.083***	0.0464***	0.0449***	0.0333***
Robust standard error	(0.0164)	(0.132)	(0.190)	(0.00936)	(0.00971)	(0.00861)
Observations	28,108	28,108	28,108	28,108	28,108	28,108
Controls						
Sex	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth	Yes	Yes	Yes	Yes	Yes	Yes
Type of area	Yes	Yes	Yes	Yes	Yes	Yes
Migrant condition	Yes	Yes	Yes	Yes	Yes	Yes
Private sector	Yes	Yes	Yes	Yes	Yes	Yes
Region of residence	Yes	Yes	Yes	Yes	Yes	Yes
Region of birth	Yes	Yes	Yes	Yes	Yes	Yes
Year of survey	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' estimates using household surveys 2005-2017.

Notes: ***, ** and * represent significance at 1%, 5% and 10% respectively. Clustered errors by household are shown in parentheses.

The wage premium is higher for urban workers than rural workers when considering all wage earners, whereas for private sector workers, it is slightly higher for rural workers than for urban workers. However, consistent with the general results, the wage premium for speaking English is notably higher for private sector workers than for all wage earners. The same pattern is found with the wage per hour and the hours labored in a week. We also find that English-speaking wage earners have higher chances of getting paid vacations, paid sick leave and a thirteenth wage. These results however, lose robustness when considering only private sector workers.

3.2 Results by sex

When we split our samples by sex, we find for all wage earners the English premium for men is higher for all variables except worked hours per week. Whereas for private sector workers, women have larger wage, hourly wage and the probability of having

paid sick leave premiums.

Table 6
Results by sex

	Dependent Variables					
	Logarithm of wage	Hourly wage	Worked hours	Vacations	Paid sick leave	Thirteenth wage
Men	0.197***	1.623***	-0.745***	0.0298***	0.0295***	0.0169***
Robust standard error	(0.00860)	(0.0870)	(0.0969)	(0.00427)	(0.00449)	(0.00387)
Observations	55,344	55,344	55,344	55,344	55,344	55,344
Women	0.177***	1.333***	-0.871***	0.0152***	0.0225***	0.00927**
Robust standard error	(0.00951)	(0.0918)	(0.111)	(0.00410)	(0.00451)	(0.00363)
Observations	29,010	29,010	29,010	29,010	29,010	29,010
Private sector workers						
Men	0.231***	1.912***	-1.124***	0.0236***	0.0271***	0.0106**
Robust standard error	(0.0105)	(0.0940)	(0.120)	(0.00590)	(0.00612)	(0.00537)
Observations	43,170	43,170	43,170	43,170	43,170	43,170
Women	0.280***	1.934***	-1.061***	0.0194***	0.0288***	0.00876
Robust standard error	(0.0129)	(0.106)	(0.153)	(0.00677)	(0.00725)	(0.00612)
Observations	16,966	16,966	16,966	16,966	16,966	16,966
Controls						
Age	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth	Yes	Yes	Yes	Yes	Yes	Yes
Type of area	Yes	Yes	Yes	Yes	Yes	Yes
Migrant condition	Yes	Yes	Yes	Yes	Yes	Yes
Private sector	Yes	Yes	Yes	Yes	Yes	Yes
Region of residence	Yes	Yes	Yes	Yes	Yes	Yes
Region of birth	Yes	Yes	Yes	Yes	Yes	Yes
Year of survey	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' estimates using household surveys 2005-2017.

Notes: ***, ** and * represent significance at 1%, 5% and 10% respectively. Clustered errors by household are shown in parentheses.

3.3 Results by highest level of education attained

Lastly, Table 7 shows our results by each level of education considered in this research. There are high and significant wage premiums for each level of education considered. On the other hand, most of the significant benefits of bilingualism are concentrated in the people with the highest levels of education and people who completed primary school.

Table 7
Results by highest education level attained

	Dependent Variables					
	Logarithm of wage	Hourly wage	Worked hours	Vacations	Paid sick leave	Thirteenth wage
All wage earners						
Incomplete primary	0.359***	1.161***	1.544	0.0211	0.0280	0.0768
Robust standard error	(0.0690)	(0.437)	(1.049)	(0.0730)	(0.0723)	(0.0648)
Observations	8,180	8,180	8,180	8,180	8,180	8,180
Complete primary	0.260***	1.105***	-0.541	0.0963***	0.0840***	0.0562**
Robust standard error	(0.0329)	(0.184)	(0.449)	(0.0235)	(0.0244)	(0.0226)
Observations	22,417	22,417	22,417	22,417	22,417	22,417
Incomplete secondary	0.159***	0.790***	-0.741***	-0.00280	0.0112	-0.00116
Robust standard error	(0.0210)	(0.106)	(0.280)	(0.0161)	(0.0163)	(0.0145)
Observations	13,833	13,833	13,833	13,833	13,833	13,833
Complete secondary	0.217***	1.220***	-0.894***	0.0222***	0.0266***	0.0163**
Robust standard error	(0.0144)	(0.0857)	(0.168)	(0.00799)	(0.00828)	(0.00705)
Observations	13,340	13,340	13,340	13,340	13,340	13,340
Higher education	0.201***	1.943***	-0.653***	0.0112***	0.0134***	0.00338
Robust standard error	(0.00806)	(0.0881)	(0.0885)	(0.00299)	(0.00336)	(0.00267)
Observations	26,584	26,584	26,584	26,584	26,584	26,584
Private sector workers						
Incomplete primary	0.361***	1.176***	1.523	0.0204	0.0268	0.0796
Robust standard error	(0.0700)	(0.447)	(1.069)	(0.0747)	(0.0739)	(0.0659)
Observations	7,886	7,886	7,886	7,886	7,886	7,886
Complete primary	0.262***	1.081***	-0.458	0.108***	0.0951***	0.0608**
Robust standard error	(0.0337)	(0.181)	(0.480)	(0.0253)	(0.0264)	(0.0248)
Observations	19,543	19,543	19,543	19,543	19,543	19,543
Incomplete secondary	0.174***	0.852***	-0.853***	-0.00185	0.0156	-0.00157
Robust standard error	(0.0218)	(0.110)	(0.294)	(0.0174)	(0.0174)	(0.0157)
Observations	11,677	11,677	11,677	11,677	11,677	11,677
Complete secondary	0.242***	1.342***	-0.910***	0.0245**	0.0278***	0.0170**
Robust standard error	(0.0161)	(0.0949)	(0.188)	(0.00960)	(0.00997)	(0.00851)
Observations	9,445	9,445	9,445	9,445	9,445	9,445
Higher education	0.277***	2.465***	-0.920***	0.0196***	0.0265***	0.00603
Robust standard error	(0.0114)	(0.113)	(0.123)	(0.00510)	(0.00554)	(0.00463)
Observations	11,585	11,585	11,585	11,585	11,585	11,585
Controls						
Sex	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth	Yes	Yes	Yes	Yes	Yes	Yes
Type of area	Yes	Yes	Yes	Yes	Yes	Yes
Migrant condition	Yes	Yes	Yes	Yes	Yes	Yes
Private sector	Yes	Yes	Yes	Yes	Yes	Yes
Region of residence	Yes	Yes	Yes	Yes	Yes	Yes
Region of birth	Yes	Yes	Yes	Yes	Yes	Yes
Year of survey	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' estimates using household surveys 2005-2017.

Notes: ***, ** and * represent significance at 1%, 5% and 10% respectively. Clustered errors by household are shown in parentheses.

In addition, every educational level does have a significant gain from learning English in their hourly wage. The same cannot be said for the hours worked per week, where most of the gains are accrued by the wage earners that have more than primary school studies. In other words, even if the benefits of speaking English do not have similar effects across educational levels, each of those levels can have significant gains from speaking English.

4 Discussion

We have shown that there are important job market premiums associated with bilingualism in Costa Rica. The most notable are on the wage and the hourly wage of workers, which are very large and robust for every specification and sample considered. Along similar lines, bilingualism not only provides higher wages, but also more job benefits. In addition to earning more, English speaking workers work less hours per week and are more likely to receive paid vacations, paid sick leave and, a thirteenth wage.

A possible explanation for these results, is the existence of very high added value sectors in the Costa Rican economy that explain these premiums. In Table 8 we show the percentage of workers in each economic activity considered in Costa Rica's household surveys that speak English fluently, where we indeed find that there are some activities with few or no workers that are fluent English speakers and others were more than a quarter of its wage earners are fluent English speakers.

Table 8
Percentage of workers that speak English fluently by area of economic activity

Economic activity	All wage earners	Private sector wage earners
Agriculture, cattle raising and fishing	1,70%	1,70%
Minery	2,76%	2,76%
Manufacturing industries	9,36%	9,37%
Electricity supply	10,91%	10,87%
Water supply	4,21%	2,97%
Construction	4,51%	4,49%
Commerce	10,38%	10,40%
Transportation and storage services	12,12%	11,70%
Hotels and restaurants	13,36%	13,36%
Information and communications	40,15%	49,82%
Finances	22,49%	27,16%
Real state	18,55%	18,59%
Professionals and scientists	30,30%	30,54%
Administration services	22,31%	22,29%
Public administration	11,33%	10,31%
Teaching	16,23%	27,72%
Health	12,54%	16,24%
Arts	20,29%	20,48%
Other services	12,34%	11,12%
NGOs and other organizations	48,09%	48,06%

Source: Authors' estimates using household surveys 2005-2017.

As a possibility, the results might be driven by a certain group of workers that happen to be in industries that have greater job market benefits in comparison to the rest of the Costa Rican economy. One way to test for this is to control for economic activity in the regressions; however, doing so might not be an ideal econometric strategy due to the potentially endogenous nature of these controls with respect to our independent variable of interest. In other words, workers in certain economic activities might self-select themselves to work in said activities because they speak English. Hence, the beta estimate would not be capturing the premium associated with speaking English in the economy but rather the job market premium associated to this potential self-selection bias. Nonetheless, if it is the case that certain sectors explain the job market premiums found in the previous section, one would expect that our estimated beta would notably lower its magnitude and perhaps significance with respect to our previous results. Table 9 displays the general results when adding economic activity dummies to our regressions.

Table 9
Results by economic activity

	Dependent Variables					
	Logarithm of wage	Hourly wage	Worked hours	Vacations	Paid sick leave	Thirteenth wage
All wage earners						
Effect	0.183***	1.533***	-0.699***	0.0177***	0.0218***	0.00792***
Robust standard error	(0.00653)	(0.0658)	(0.0729)	(0.00306)	(0.00326)	(0.00276)
Observations	84,314	84,314	84,314	84,314	84,314	84,314
Private sector workers						
Effect	0.248***	1.947***	-0.922***	0.0159***	0.0238***	0.00369
Robust standard error	(0.00838)	(0.0743)	(0.0951)	(0.00460)	(0.00481)	(0.00421)
Observations	60,097	60,097	60,097	60,097	60,097	60,097
Controls						
Sex	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes
Educational level	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth	Yes	Yes	Yes	Yes	Yes	Yes
Type of area	Yes	Yes	Yes	Yes	Yes	Yes
Migrant condition	Yes	Yes	Yes	Yes	Yes	Yes
Private sector	Yes	Yes	Yes	Yes	Yes	Yes
Region of residence	Yes	Yes	Yes	Yes	Yes	Yes
Region of birth	Yes	Yes	Yes	Yes	Yes	Yes
Year of survey	Yes	Yes	Yes	Yes	Yes	Yes
Economic activity	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' estimates using household surveys 2005-2017.

Notes: ***, ** and * represent significance at 1%, 5% and 10% respectively.

As it can be appreciated, there is indeed a minor decrease in the magnitude of the estimated job market premiums. Nonetheless, this decrease is marginal and the results remain very significant, large and robust⁷. The econometric evidence indicates

⁷The results found in the different sample splits hold after adding these controls as well, they are not shown in this document to save space but are available upon request.

that the results are not being driven by the agglomeration of workers in certain economic activities in Costa Rica. This is an encouraging result, because it entails that every worker in Costa Rica regardless of the economic activity where he or she labors can gain job benefits by learning English.

However, it is interesting how in spite of all the benefits associated to speaking English, there is no decreasing trend over time for any of these benefits. This is inconsistent with the development strategy Costa Rica has been trying to adopt for more than two decades, which has focused on attracting high added value foreign investment and tourism, both sectors that usually operate in English. While from 2005 to 2016, FDI has almost doubled⁸, and tourism has more than doubled its value⁹, these premiums have not responded accordingly.

In these circumstances, it is worth asking if the population is actually learning more English. Figure 2 shows this variation between the years 2005-2017. In 13 years, the percentage of the population between 25 and 65 years of age that speak English has increased from 9.5% in 2005 to 9.8% in 2017. Thus, showing that proficiency in English in Costa Rica's labor force has stagnated in the period of study. This result is not due to generational differences. Even when comparing only the population between 25 to 35 years of age -which are on average more educated and have had more exposure to English-, the change has been from 13% in 2005 to 14.9% in 2017. Furthermore, this stagnation is also present regardless of the type of area. In urban areas less than 20% of this population speak English and in rural areas this percentage has been around 5% over the whole 13 year period.

These results reinforce two very important elements: neither has the labor market self-adjusted to reduce the relative scarcity of second language competence, nor have the educational policies been effective in correcting this. This is a warning for revisiting and designing better public policies regarding second language skills and -in a broader sense- development. Furthermore, this lack of labor market adjustment seems to be founded in structural problems of the educational system and the economy rather than on short-term labor market dynamics or imperfections. We address these claims in Figure 4 and Figure 5.

⁸From 1,364.13 millions of dollars in 2005 to 2,681.01 millions of dollars in 2016.

⁹From 1,839.9 millions of dollars in 2005 to 3,708 millions of dollars in 2016.

Figure 3
 Percentage of population that speak English by ranges of age,
 2005-2017

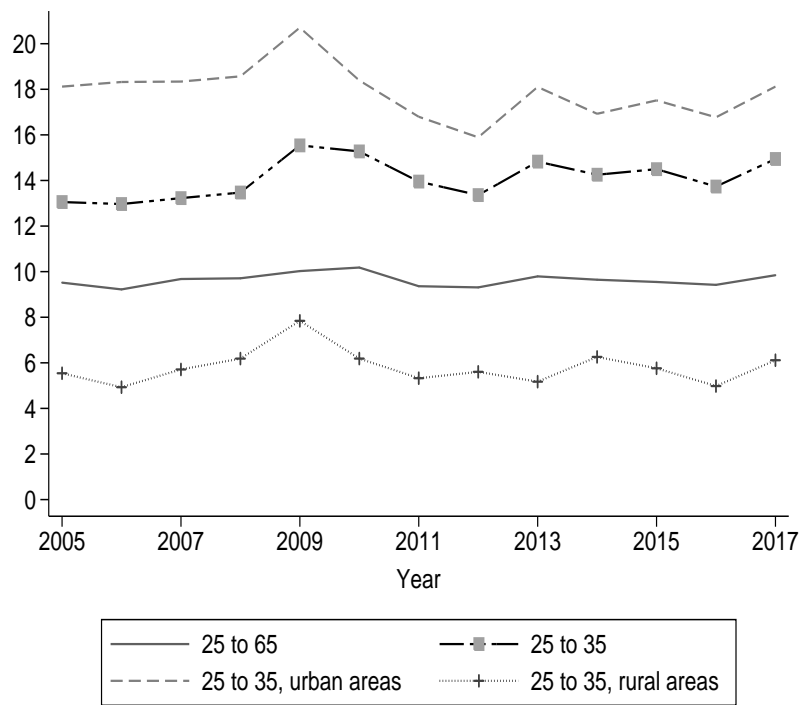
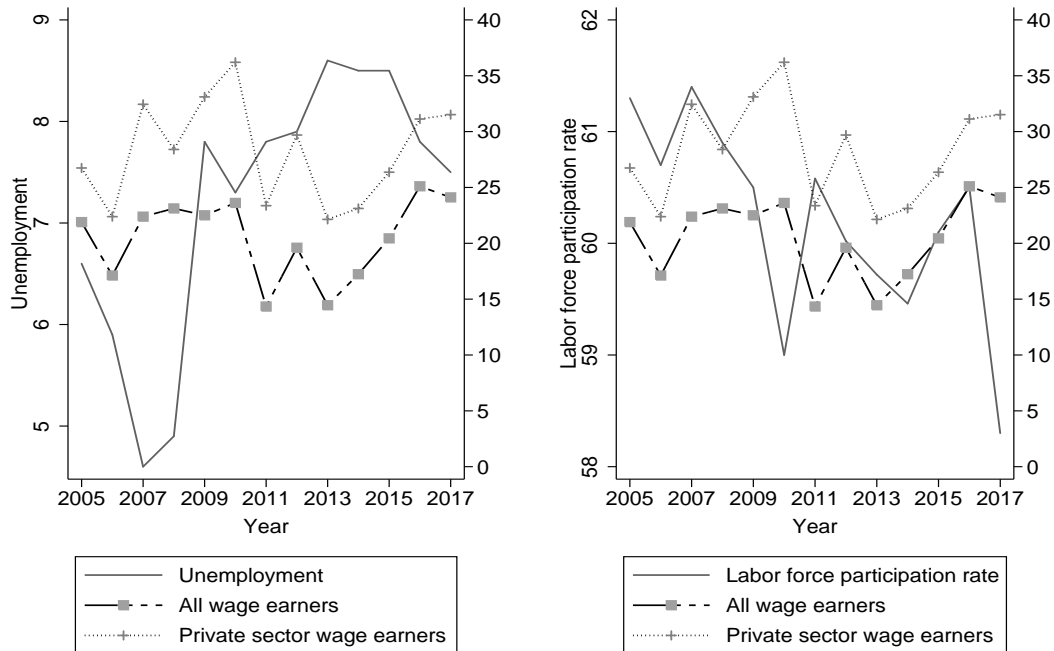


Figure 4

Unemployment, labor force participation and after tax wage premiums, 2005-2017



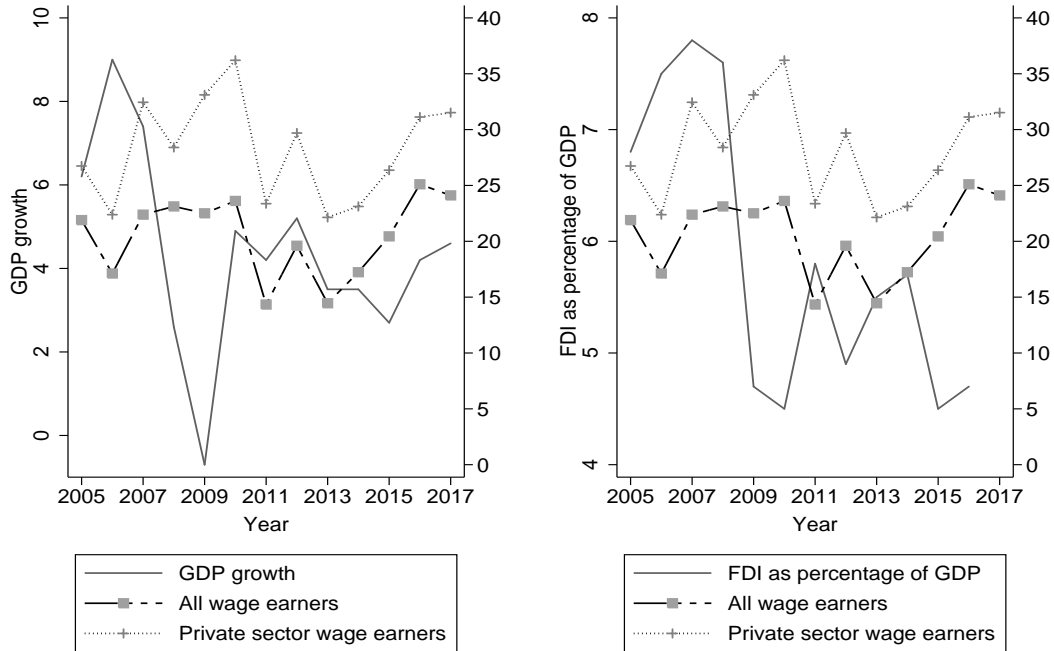
Source: Authors' estimates using household surveys 2005–2017.

Clearly there is no pattern between the additional compensation of knowing English and the unemployment rate or the labor market participation rate in Costa Rica. Unemployment increased from 4.6% in 2007 to 7.5% in 2017, while the labor market participation rate decreased from 61.4% to 58.3%. In this time frame, the wage premiums barely changed. The premium for all wage earners changed from 22.38% to 24.11% and from 32.44% to 31.52% for private sector wage earners.

Along similar lines, Figure 5 shows how growth of the economy and the share of the GDP that represents foreign direct investments have no clear relationship with the wage premiums. While both variables have had a decreasing trend over time, the labor market premiums have barely changed.

Figure 5

GDP growth, FDI as percentage of GDP and after tax wage premiums, 2005-2017



Source: Authors' estimates using household surveys and Costa Rica's central bank.

This shows that English labor market premiums hold no consistent relationship with labor market dynamics or the overall economy. Hence, there is evidence to claim that bilingualism is not an endogenous variable with respect to the labor market or the economy overall. This has two implications. First, without the endogeneity concerns, we can claim that our estimates do indeed approximate the labor market premium of bilingualism. Second, bilingualism in Costa Rica is determined by other factors rather than just the current state of the labor market or the economy.

Therefore, although Costa Rica has been successful in attracting FDI and promoting tourism, the job market premiums associated to speaking English have barely varied, evidencing that the job market has not self-adjusted to reduce the relative scarcity of second language competence. This is a warning for when revisiting and designing better public policies regarding second language skills and -in a broader sense- development.

5 Conclusions

We estimated the bilingualism job market premiums associated for Costa Rica, a small, open, developing economy. For this, we used the country's household surveys from 2005 to 2017, which provided samples of workers representative of the country's workforce.

We found for all wage earners a wage premium of 20.6%, an hourly wage premium of 1.5 dollars, a premium of 0.8 less hours worked per week, an increase of 2.5% in the probability of having paid vacations, an increase in the probability of having paid sick leave of 2.7% and an increase of the probability of receiving a thirteen wage of 1.4%. For private sector wage earners, the wage premium is of 28.4%, 1.95 dollars per hour worked, 1.11 less worked hours per week, and probability increases of having paid vacations, having paid sick leave and receiving a thirteen wage of 2.3%, 2.8% and 1.02% respectively. These results are robust to different models and over the period of study. We also find that these results hold for subsamples by sex, type of area and attained levels of education. This constitutes clear evidence that speaking English is a highly valuable skill in Costa Rica, and that every worker can accrue labor market benefits by becoming bilingual.

The results also show that these job market premiums do not fade over time. Denoting that Costa Rica's labor market has not adjusted to reduce the relative scarcity of second language competence, in spite of the fact that the country has focused in attracting foreign direct investment and fostering activities that are English speaking intensive in the past twenty years. Additionally, even though the coverage of English classes has increased over time and the country has devoted resources to enhance its teaching, this does not reflect in higher percentages of a bilingual working population. This constitutes a warning for revisiting and designing better public policies regarding second language skills and -in a broader sense- development in Costa Rica.

References

Aldashev, A., Gernandt, J., & Thomsen, S. L. (2009). Language usage, participation, employment and earnings: Evidence for foreigners in West Germany with multiple sources of selection. *Labour Economics*, 16(3), 330-341.

Arroyo, E. M. (2013). Cronología de la enseñanza del inglés en Costa Rica durante el Siglo XX. *Revista Comunicación*, 21(1 (2012)), 4-15.

Azam, M., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change*, 61(2), 335-367.

Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, 86(2), 481-496.

Casale, D., & Posel, D. (2011). English language proficiency and earnings in a developing country: The case of South Africa. *The Journal of Socio-Economics*, 40(4), 385-393

Cattaneo, A., & Winkelmann, R. (2003). Earning differentials between German and French speakers in Switzerland (No. 0309). Working Paper, Socioeconomic Institute, University of Zurich.

Chiswick, Barry R. (1998). Hebrew language usage: Determinants and effects on earnings among immigrants in Israel. *Journal of Population Economics*, 11: 253:271.

Chiswick, B. R., Patrinos, H. A., & Hurst, M. E. (2000). Indigenous language skills and the labor market in a developing economy: Bolivia. *Economic development and cultural change*, 48(2), 349-367.

Córdoba, P., Coto, R., & Ramírez, M. (2005). La enseñanza del inglés en Costa Rica y la destreza auditiva en el aula desde una perspectiva histórica. *Revista Electrónica” Actualidades Investigativas en Educación”*, 5(2).

Crystal, D. (2003). *English as a global language*. Cambridge university press.

Di Paolo, A., & Tansel, A. (2015). Returns to foreign language skills in a developing country: The case of Turkey. *The Journal of Development Studies*, 51(4), 407-421.

- Dustmann, C., & Van Soest, A. (2002). Language and the Earnings of Immigrants. *ILR Review*, 55(3), 473-492.
- Fry, R., & Lowell, B. L. (2003). The value of bilingualism in the US labor market. *ILR Review*, 57(1), 128-140.
- Grin, F. (2001). English as economic value: Facts and fallacies. *World Englishes*, 20(1), 65-78.
- Jiménez, R & Arce, G (2014). Los desafíos de la educación preescolar, primaria y media en Costa Rica. In Gregosz, D (Ed.) *Los Desafíos de Educación Preescolar, Básica y Media en América Latina* (pp. 147-176). Santiago, Chile: Konrad-Adenauer-Stiftung
- Ku, H., & Zussman, A. (2010). Lingua franca: The role of English in international trade. *Journal of Economic Behavior & Organization*, 75(2), 250-260.
- OECD. (2012). *Attracting Knowledge-Intensive FDI to Costa Rica: Challenges and Policy Options*. Making Development Happen Series No. 1, OECD Publishing, Paris.
- Programa Estado de la Educación (2017). *Sexto informe del estado de la educación*. Consejo Nacional de Rectores.
- Rendon, S. (2007). The Catalan premium: language and employment in Catalonia. *Journal of Population Economics*, 20(3), 669-686.
- Stöhr, T. (2015). "The returns to occupational foreign language use: Evidence from Germany." *Labour Economics* 32: 86-98.
- Toomet, O. (2011). Learn English, not the local language! Ethnic Russians in the Baltic states. *The American Economic Review*, 101(3), 526-531.
- Vaillancourt, F., Lemay, D., & Vaillancourt, L. (2007). Laggards no more: The changed socioeconomic status of Francophones in Quebec. Backgrounder-CD Howe Institute, (103), 0-1.