Returns to Foreign Language Skills in a Developing Country: The Case of Turkey

Abstract:
This paper analyzes the economic value of foreign languages in Turkey, which is a country characterized by fast economic and social development. We find positive earnings returns to proficiency in English and Russian, which increase with the level of competence. French and German skills are also positively rewarded, although their return seems mostly linked to the likelihood to hold specific occupations. On the contrary, knowing Arabic does not generate an earnings premium. Focusing on English, we check for heterogeneous returns along the conditional earnings distribution. The results are qualitatively invariant when we account for misclassification errors in self-reported English skills.
Key Words: Foreign Languages, Earnings, Returns to Skills, Middle East, Turkey
JEL Codes: I25, J24, J31, O15, O53

1. **Introduction**

Foreign language (FL henceforth) skills represent a form of human capital that can be rewarded in the labour market. Several papers highlight the economic value of FL knowledge among the native populations of developed countries. Any existence of positive returns to FL competences is expected to be even more relevant in developing countries. Indeed, fostering widespread FL knowledge of the population, alongside formal schooling and other cognitive skills (Hanushek & Woessmann, 2008), might represent a stepping stone for economic development in the globalised world (Seargeant & Erling, 2011). However, there are relatively few studies of this topic in developing countries, mainly due to data limitations.

This paper investigates returns to FL competences in the Turkish labour market, defined as the *ceteris paribus* increase in earnings associated with FL knowledge[1]. Turkey provides an interesting case for several reasons. First, the economic value of FL proficiency in Turkey has not previously been investigated. Second, during recent decades, Turkey has experienced impressive growth rates as well as increases in international trade, tourist arrivals, foreign direct investment (FDI) and research and development (R&D) activities, all of which have contributed to the country's social and economic development. Along with these accomplishments, the globalisation process has stimulated demand for FLs (Fidrmuc & Fidrmuc, 2009; Fidrmuc, 2011; Hoon et al., 2011). Therefore, fostering FL skills would be especially important for a mid-sized, middle-income, emerging economy such as Turkey, since this would improve national performance in the global knowledge economy and stimulate the potential for further economic growth and development. Rising demand for FLs, combined with their relatively scarce supply among Turkish workers, generates the potential for important economic rewards. Therefore, our aim is to analyse the existence and amount of earnings returns to FL competences in the Turkish labour market. Another significant contribution of this work is that our rich dataset enables analysing the economic value of several FLs, without constraining the focus only to English — as was previously done for other developing countries.

---

[1] We focus on pecuniary returns, although there might be important non-monetary returns to FL proficiency (which are difficult to quantify). These might be, for example, the ability to read books and to watch movies in original languages or the increased facility to travel and to establish relationships with individuals from other countries, all of which are likely to generate positive gains in terms of individual utility. In addition, we implicitly consider "gross monetary benefits" from FL knowledge, as we are unable to capture the individual opportunity cost of learning one or more FLs.

In order to keep the empirical analysis tractable, we focus on male wage-earners between the ages of 25 and 65. We estimate earnings regressions augmented by indicators of FL competences. We also control for occupation fixed-effects in order to account for the indirect link between language and earnings through occupation. Parental education is used as an additional control, thus reducing the impact of unobserved factors such as unobserved ability and social networks. Indeed, obtaining unbiased estimates represents the main concern in the econometric analysis of returns to language skills, as it is for other forms of human capital such as schooling. Unobserved individual heterogeneity (e.g. innate ability) and the resulting omitted variable bias represent the main concerns. Unfortunately, we are unable to explicitly deal with this potential issue due to a lack of either panel data or exogenous sources of variation in FL knowledge. Nevertheless, the robustness of our results to the inclusion of control variables and to several sensitivity analyses as well as their similarity to the findings of previous studies allow us to argue that returns to FL proficiency should not be mainly driven by unobserved ability.

In addition, the self-reported nature of our measure of skills in FLs might represent another source of bias, since individuals tend to overstate their competences. Therefore, by focusing on English, we seek to control for possible misclassification errors in the level of skills by using information about the frequency of English use for leisure activities as a proxy for skills. This allows estimating a model in which the observed skills variable is replaced by its predicted value, which should be less affected by misclassification errors. The results obtained from this alternative methodology are qualitatively robust and highlight again that FLs (and especially English) skills are positively rewarded in the Turkish labour market.

2. Background

2.1 The Demand for Competences in Foreign Languages in Turkey

Turkey is considered to be a middle-income country. It is the world's 18th largest economy. The country's per-capita income, which has nearly tripled during the past decade, currently exceeds 10,000 US dollars. Several researchers, such as Adak (2010), suggest that the expansion of international trade seems to be one of the most important factors driving economic growth and development in Turkey over the past decade. The 1988 financial liberalisation fostered both exports and imports (Cebeci & Fernandes, 2013). As a result, trade openness was almost 50% in 2012. In 1996, Turkey entered into a Customs Union with the European Union (EU), which increased competitive pressure on the domestic economy. EU countries are Turkey's main trade partners, with Germany leading among them. In 2012, the largest share of exports went to Germany. In 2012, Russia was Turkey's leading import supplier, followed by Germany and China (TURKSTAT, 2013a). Increasing trade openness has boosted demand for FL competences, since speaking a common (foreign) language is likely to reduce transaction costs with trade partners[2].

---

[2] There are numerous studies documenting that language barriers represent a barrier to international trade flows. Hutchinson (2005) showed that trade flows among non-English-speaking countries decrease with an increase in the linguistic distance from English. In fact, linguistic distance (relative to mother tongue) is likely to increase the cost of FL learning and reduce learning efficiency, thus diminishing the volume of bilateral trade (Isphording & Otten, 2013). Melitz (2008) found that direct communication in a common language is three times more effective than indirect communication in promoting trade. Ku and Zussmann (2010) and Fidrmuc and Fidrmuc (2011) estimated gravity models augmented by

FDI brings financial resources as well as technological and managerial know-how to recipient countries and thus contributes to their economic growth. FDI flows to and from Turkey have increased substantially during recent decades, from only 18 million US dollars in 1980 to 12 billion US dollars in 2012 (reaching a peak of 22 billion US dollars in 2007; UNCTAD, 2013). In FDI activities, FL skills enable communication and interactions with foreign counterparts, authorities or customers in order to convey information about the functioning of foreign markets and reduce linguistic and cultural barriers.

FL skills are also strongly connected with R&D activities as suggested by Fidrmuc (2011). Command of FLs enables R&D personnel to follow new scientific and technological developments and to interact with international researchers and institutions. Improving competences in FLs would thus increase the country's research potential, leading to more innovation and other productive investments and promoting economic growth. However, Turkey's R&D expenditures and R&D personnel are rather low and it plans to increase these domains (SPO, 2006). Further, Turkey has not been able to attract foreign R&D investments in different key sectors. Several policies have been implemented at both the national and the local level in order to increase the country's attractiveness to foreign investors. Overall, these changes are likely to increase the need for FL skills in future years, since R&D personnel will have to be proficient to perform scientific and innovative activities as well as to attract more international R&D investments, which in turn will enhance economic growth.

Moreover, during recent decades, a rapid structural transformation has taken place in the Turkish labour market: while agricultural employment has declined, service sector

FLs. They suggested that significant gains can be realised by improving linguistic skills, highlighting the role of English as the *lingua franca* for commerce and trade.

employment (including tourism[3]) has increased substantially since the early 1980s (see TURKSTAT, 2013b). The growing importance of foreign tourism in the country also fosters demand for FLs.

Taken together, the recent high growth rates, increasing trade openness and economic internationalisation, the phenomenal growth of the tourism sector and the ongoing intensification of FDI and R&D activities in the Turkish economy represent the main factors contributing to increased demand for FL skills. Moreover, the 1999 announcement of the candidacy of Turkey for full membership of the EU and the accession negotiations to the EU since October 2005 have also stimulated demand for FLs, particularly for English, because of its role as the international *lingua franca* for commerce and trade (Ku & Zussmann, 2010; Fidrmuc, 2011).

2.2 The Supply of Foreign Language Skills among the Turkish Labor Force

The corollary of this growing demand for FL skills is their supply in the Turkish labour market[4]. In Turkey, competences in FLs are mainly acquired either at schools or in private language centres[5], which are common across the country. During the 2010–2011 academic

---

[3] Regarding the importance of FL skills for services and tourism, see Tucci and Wagner (2004) and Leslie et al. (2002), respectively.

[4] The mother tongue of most people in Turkey is Turkish, which is not an Indo-European language but belongs to the Altay-Uralic language family. Turkish is the only official language. However, there are many other native languages spoken in Turkey, Kurdish and Arabic being the most common (Yağmur, 2001).

[5] Foreign language knowledge in Turkey may also be related to migration background. For example, there have been several important waves of migration of ethnic Turks (i.e. individuals with direct or indirect Turkish origins) from

year, 34,905 individuals, about half of which were women, completed a course in an FL at a private language centre. Of these, 92% of students studied English and 5% studied German. The remaining 3% completed courses in Arabic, French, Italian, Japanese, Spanish and Russian (TURKSTAT, 2012).

FL instruction in the Turkish education system has changed significantly over time. French was the most common FL studied in schools before the 1950s. However, English has replaced French and is now the most widely studied FL in the schooling system. Until the 1997 educational reform[6], FL instruction in public schools started in the sixth grade and continued throughout high school. Moreover, so-called 'Anatolian' high schools, which are highly selective public high schools, offer more intensive English instruction and a few of them French or German, too. There are also private schools at all education levels in Turkey, where the medium of instruction is an FL, usually English. An increasing number of public universities have adopted English as medium of instruction, either only for some degrees or for the whole university. Turkish medium-of-instruction universities have elective FL courses, predominantly English. Finally, Arabic is taught in religious vocation high schools. Before the educational reform of 1997, Arabic instruction started in the sixth grade but since then it has begun in the ninth grade. At Anatolian religious vocation high schools, both Arabic and English are taught intensively.

Bulgaria who would know Bulgarian and Russian as well. Moreover, more recently, many Turkish immigrants (first and second generations) residing in Germany have returned to Turkey (Aydın, 2012) and they could be proficient in German.

[6] The Turkish education system has experienced an increase in the emphasis given to FL teaching since the educational reform of 1997, which also extended compulsory schooling from five to eight years. However, our data do not cover individuals affected by the 1997 reform. It also seems to be worth mentioning that the recent (and gradually implemented) educational reform of 2012 established 12 years of compulsory education and a further increase in foreign language instruction at schools.

Despite efforts to increase competences in FLs, Turkey is characterised by a significant English language deficiency (OECD, 2012). According to Education First (2011), Turkey is a "very low proficiency" country. The special 2006 Euro-Barometer Survey about languages in Europe provides a similar picture (EU, 2006). It indicates that Turkey has the highest percentage of the population declaring an inability to converse in a language different from the mother tongue among the EU-25 and four candidate countries. Thus, growing demand for FLs, together with their scarce supply among Turkish workers, generates the potential for economic rewards to FL skills. Quantifying this economic return and finding out the FLs that matter most in the labour market are the main aims of this paper.

3. Related Literature

Conceiving language competences as a form of human capital has gained consensus in the literature during recent decades. Language proficiency, similar to formal schooling, is a costly asset that is embodied in the individual and is likely to be productive in the labour market. Most of the literature concerns immigrants, because competences in a host country's language are fundamental for their economic and social integration (for a recent overview, see Chiswick & Miller, 2014). The same framework can be applied to explain skills in FLs among the native population. There are several reasons to expect a positive relationship between language proficiency and earnings. First, language might directly affect productivity by enhancing efficiency in communication among co-workers and managers. Second, language itself may enable achieving more prestigious occupations that

are also likely to be better remunerated (Chiswick & Miller, 2009), as workers obtain a premium if their language skills match the linguistic requirements in the workplace (Chiswick & Miller, 2010, 2013). This means that a substantial part of the positive relationship between language competences and earnings is indirect, operating through the occupational channel. Third, language competences might also be remunerated when not directly used/relevant in the workplace, since this asset represents a positive signal of other cognitive skills from the employer's perspective (Grin et al., 2010). Indeed, there is substantial evidence in the literature on the improved cognitive skills of bilinguals or those who have studied an FL. According to Adesope et al. (2010), among others, bilinguals have a generalised cognitive advantage over monolinguals in several executive functions. Thus, in this context, language competences could increase earnings directly by improving cognitive skills.

Indeed, a growing number of papers have reported significant returns to FL skills among the native population. Some studies focus on developed countries. Saiz and Zoido (2005) analysed returns to FLs by using a sample of US college graduates. Williams (2011) reported significant earnings premiums for FL usage at work in 12 European countries. Ginsburgh and Prieto-Rodriguez (2011) confirmed substantial returns to English proficiency in several European countries. Lang and Siniver (2009) analysed the case of English in Israel. They showed that English knowledge is significantly remunerated in the Israeli labour market for both immigrants and natives, although the return is heterogeneous for different groups of workers.

The return to English proficiency has been analysed in a few developing countries. Toomet (2011) found that local languages are not remunerated in Latvia and Estonia, while English proficiency produces a significant earnings premium. Levinsohn (2007) and Casale

and Posel (2011) reported high reward for English competences in South Africa and Azam et al. (2013) obtained substantial returns to English skills in India. This evidence confirms that English proficiency is also a valuable asset in developing countries. Our study resembles the last group concerned with developing countries, since we investigate the return to FL skills in Turkey. However, in both South Africa and India, English is the former colonial language[7] and currently one of the official languages, whereas this is not the case in Turkey, Latvia or Estonia, where English is a non-native and unofficial language. In this sense, our study is close to the paper by Toomet (2011); however, he considered the case of a Russian minority, whereas we consider the whole native population of Turkey. Further, unlike previous studies of developing economies, we consider returns to several FLs spoken in Turkey.

4. Data and Descriptive Statistics

The empirical analysis is based on nationally representative Turkish data from the Adult Education Survey (AES). The details on the data and sample used in the empirical analysis are reported in Section A1 of the Appendix. Our analysis is based on a subsample of male wage-earners. The AES asks an individual's knowledge of up to seven languages different from the mother tongue. In the case of having some knowledge of at least one FL,

---

[7] Angrist and Lavy (1997) estimated the return to proficiency in French in Morocco, which is also the colonial language and was used as a language of instruction until 1983 (when it was replaced by Arabic). They also found that returns to education were substantially lower for those schooled in Arabic, mostly because of their reduced French skills.

individuals report detailed information about the two FLs they know best, the way they learnt those languages and their frequency of use at work and for leisure.

Table 1 shows that roughly 64% of the individuals in the sample do not speak any FLs. Over one in three individuals are able to speak at least one FL (most speak just one). FL knowledge is more common among the younger cohort, those with greater educational attainment and white collar employees (especially if highly skilled). Knowing at least one FL is slightly more common in urban than in rural areas.

[TABLE 1]

Table 2 reports that English is the most widely known FL, with about 82% of those who possess some knowledge of FLs declaring that English is one of the languages they know. German represents the second most frequent language known by 11%, considerably less common than English. The largest share of Turkey's exports go to Germany that, as noted above, is also the traditional destination for Turkish emigrants. Arabic is the third most frequent language (8.6%), which is taught in religious vocation high schools and might be common among the indigenous population in the south-southeast of the country as well as among people with some migration experience of the Middle East and North Africa (which were alternative migration destinations during the 1980s). Arabic is followed by French (6.7%), which was widely taught as part of the oldest generation's schooling. Less common is Russian (2.3%), which is not taught in the school system. However, Russian is likely to be known by returning Turkish workers from the migration wave towards Russia and Central Asia that occurred in the 1990s (Tansel & Yaşar, 2010) as well as migrants from Bulgaria.

[TABLE 2]

Further, English is relatively more common within the younger cohort, as is Russian, while knowledge of German, Arabic and French is somewhat higher among older groups. Russian, German and Arabic are significantly more common among the less educated. In particular, almost 50% of the Arabic speakers sampled have five or fewer years of schooling. Those who know German among the less educated may be return migrants from Germany, but, unfortunately, we do not have information about previous migration experiences. English is more frequently known among white collar employees, while blue collar employees who declare they know FLs are relatively more likely to know Arabic. Finally, the incidence of Arabic knowledge and ― to a lesser extent ― German knowledge seems to be relatively high in rural areas. In the case of German, this evidence might reflect a previous (direct or indirect) migration background from Germany. Regarding Arabic, its incidence among low educated individuals residing in rural areas might mirror ethnic identities with Arabic roots[8].

In Table 3, we focus more deeply on the FLs individuals know better. English represents the primary FL for about four-fifths of FL speakers, followed by German (7.4%), Arabic (5.9%) and French (4.9%). Next, we consider the way in which people learnt the best FL they know (not shown here). Most people acquired English skills at school (79% among those who affirm English to be the best FL they know), on a private course (10%) or by self-learning (7%). On the contrary, 94% of French speakers learnt this language at school. The share of individuals who declare they learnt German and Russian abroad is significantly higher than those for other languages. Finally, 40% of those that consider

---

[8] Unfortunately, it is not possible to discern this point since the Turkish questionnaire of the AES does not include specific questions about mother tongue (which are indeed included for other countries). Therefore, information about Arabic knowledge should be taken with caution, since its distinction with ethnic background is somewhat subtle.

Arabic to be the best FL they know learnt Arabic at (religious vocation) schools and 35% in the family.

[TABLE 3]

We can obtain a better insight into the quality of FL skills by considering basic, regular and advanced levels of competences. Among those who declare English to be their first FL, 55% report having a basic level, about 31% have regular skills and only 14% have advanced skills. The distribution of German skills follows a similar pattern, whereas French skills are mostly concentrated in the basic level and those who claim Arabic to be their first FL are likely to have advanced skills.

Finally, unconditional earnings differentials by FL knowledge are reported in Table 4. The AES includes net monthly earnings from the main job (in Turkish liras). Tabulating monthly earnings by deciles shows that the incidence of top-coded earnings is significantly higher among those who speak at least one FL, but lower in the bottom deciles. Similarly, average monthly earnings are markedly higher among FL speakers. However, not all FLs are positively related to earnings.

[TABLE 4]

Indeed, knowledge of German, English, Russian or French is clearly associated with higher earnings. However, this is clearly not the case for Arabic, which instead seems to be associated with lower earnings. Nevertheless, the relationship between FL knowledge and earnings that we observe in the raw data might be confounded by other individual and labour market characteristics that are likely to co-vary with both FL knowledge and earnings. Therefore, in the next section, we analyse the return to FL knowledge in a regression framework, which would provide the *ceteris paribus* or conditional association between FL knowledge and skills and labour market earnings. The complete list of

explanatory variables used in the empirical analysis is provided in Table A1 in the Appendix together with some descriptive statistics.

5. Empirical Results

5.1 Foreign Language Knowledge and Earnings

The starting point of our empirical analysis consists of comparing the results from several OLS regressions of log earnings ($ln(Y_i)$) on typical human capital and labour market variables (included in the vector X_i of control variables), plus different indicators of FL knowledge ($g(FL)$), that is:

$$\ln(Y_i) = \alpha + \beta' X_i + g(FL) + \varepsilon_i \qquad (1)$$

First, we include dummies for the number of FLs known by each individual, which means that equation (1) is specified as:

$$\ln(Y_i) = \alpha + \beta' X_i + \sum_{k=1}^{3} \delta_k NFL_k + \varepsilon_i, \qquad (2)$$

where NLF_k represents dummy variables and δ_k are the coefficients that capture the conditional earnings differentials associated with knowing one, two or three (or more) different FLs — relative to individuals who do not know any language other than their mother tongue.

Second, we estimate several separate equations containing a dummy for each specific and non-mutually exclusive FL (FL_l, with l = 1 for English, 2 for French, 3 for German, 4 for Arabic and 5 for Russian). Therefore, equation (1) now becomes,

$$\ln(Y_i) = \alpha + \beta' X_i + \gamma_l FL_l + \varepsilon_i, \qquad (3)$$

where γ_l represents the parameters to be estimated, which reflect the earnings premium associated with a general knowledge of each FL. Third, the five different FL indicators are jointly included in the regression in an additive way, that is:

$$\ln(Y_i) = \alpha + \beta' X_i + \sum_{l=1}^{5} \varphi_l FL_l + \varepsilon_i. \tag{4}$$

Finally, we add occupation fixed-effects and dummies for parental education to the vector of control variables (X_i). Occupation fixed-effects (two-digit ISCO88 classification) inform us about the extent to which the relationship between FL knowledge and earnings is indirect, working through the occupational channel — i.e. individuals who know FLs earn more because they are attracted into better paid occupations. Furthermore, controlling for parental education should help pick up unobservable characteristics such as cognitive and non-cognitive ability and social networks.

The estimated coefficients for the control variables are quite standard and are not discussed for brevity reasons (the estimates are displayed in Table A2 in the Appendix). The main coefficients of interest from equations (2)–(4) are reported in Table 5. The first column of Table 5 shows that having some knowledge of one FL is associated with 8% higher earnings, which increases to 15% in the case of knowing two FLs and up to 38% in the (infrequent) case of three or more FLs. However, not all languages are equally rewarded in the labour market, as shown by the estimates of equation (3) in columns 2–6 of the same table. In fact, knowing English has a clearly significant return (around 11%) and the estimates for French and German are also statistically significant with about 8% returns, while Arabic is associated with lower earnings[9]. Finally, we obtain a noticeably high and

---

[9] The negative sign for Arabic knowledge is because the FL knowledge indicators are not mutually exclusive (i.e. individuals may have some knowledge of more than one FL), which means that those who know Arabic are likely to earn

significant return to Russian knowledge (probably due to the low number of observations), which is associated with 20% higher earnings. When we simultaneously include all FL dummies (eq. (4), reported in column 7), the point estimates of English and German remain unchanged, while that of French becomes significant and slightly higher. The return to Russian is somewhat reduced in this full specification, indicating that it might be driven by other "language-related" unobservable characteristics.

[TABLE 5]

After controlling for occupation fixed-effects[10], we obtain a lower return to English knowledge (see column 8 of Table 5). Occupation accounts for about 30% of the return to English, which is still substantial and significant even among employees within the same occupation. A similar pattern is observed for French, German and Russian. The results regarding Arabic knowledge remain invariant even when controlling for occupation. In the

___

less than those who either do not know any FLs or have some knowledge of English, French, German, Russian or another language. Following the suggestions of the referees, we further investigated this issue. We performed additional regressions in which we eliminated observations of individuals who know FLs other than Arabic. In this case, the point estimate for Arabic was close to zero and statistically insignificant, i.e. Arabic speakers do not earn significantly different (conditional on other covariates) than those who do not have any knowledge of FLs. This result might be due to a confounding effect of the unobserved variables related to Arabic knowledge, which may hide the labour market value of having some command of this language. Further, we have a small number of observations of individuals with Arabic knowledge, which does not allow in depth analysis. We believe that specific information about mother tongue and place of birth and current residence (which is not available in the Turkish AES data) is needed to provide a better understanding of the economic value of Arabic knowledge.

[10] It might be argued that the inclusion of occupation fixed-effects represents "bad controls" (following Angrist & Pischke, 2009) in the sense that the estimation of the treatment effect (i.e. FL return) is confounded by the inclusion of controls that depend on the treatment itself (i.e. occupation). Therefore, under positive occupational sorting, the mediating impact of occupation in the language–earnings relationship is likely to represent a lower bound of the whole relevance of occupation.

last column of Table 5, we include dummies for parental education, which are significant predictors of earnings, implying a certain degree of social segmentation in the Turkish labour market. Moreover, we observe a very modest reduction in the point estimates of returns to FL knowledge.

It should be noted that the OLS estimates of returns to FL knowledge might be affected by omitted variable bias. This concern comes from the fact that our FL knowledge indicators in equations (2)–(4) might be correlated with the error term (ε_i) due to the existence of unobserved variables (e.g. cognitive ability) that affect both the propensity to know one or more FLs and earnings potential. However, the last evidence can be taken as suggestive of the robustness of our results to potential omitted variable bias. Indeed, Lang and Siniver (2009) also obtained similar estimates when they simultaneously included two different languages in the earnings equation. Since the ability to learn two or more different FLs should be similarly correlated with general unobservable skills, they argued that if knowing different languages mainly depends on unobserved ability, a significant change should be observed when dummies for different languages are simultaneously included. This conclusion was also confirmed by their panel estimates of the return to English proficiency. In our case, the coefficients of English and German knowledge are virtually unaffected by the inclusion of other FL indicators in the regression (while only minor changes are observed for other languages). This supports the fact that the effect of unobserved individual heterogeneity is limited. A more convincing argument was proposed by Saiz and Zoido (2005), who estimated the return to FLs among US college graduates using two-period panel data. They argued that if unobserved ability is the main driver of the return to language skills, one should observe similar earnings returns for those who currently speak an FL and those who were able to do it only in the past but not in the

current period. Indeed, their results go in the opposite direction, suggesting that the impact of omitted ability should be rather limited. Unfortunately, we only dispose of cross-section data that only contain information about current language knowledge. Moreover, we do not dispose of (credible) exogenous sources of variation in FL knowledge, which means that we are unable to perform instrumental variable estimations (as, for example, Bleakley & Chin, 2004). Therefore, the results should be taken as conditional correlations (rather than causal effects), although the reported evidence seems to indicate that the return to FLs knowledge should not be just a mirror of (omitted) ability bias.

5.2   Returns to Skill Levels in FLs

The results in the previous section point out that having some knowledge of languages other than the mother tongue generally has a market value in Turkey. However, if the labour market pays a different price for different levels of the command of a language, general levels of FL knowledge might be just a partial picture of the earnings return to this human capital asset. Hence, in this section we exploit the available information about the different skill levels of the best FL known. The language-augmented earnings equation (1) can be easily adapted to the inclusion of dummy variables that capture the level of skills in FLs. As in the case of general FL knowledge, we first estimate several OLS equations that include indicators for skills in each single FL[11] (SFL^l, with $l = 1$ for English, 2 for French, 3 for German and 4 for Russian), that is,

---

[11] Note that we excluded Arabic from the list of FLs considered in the estimation of returns to skills, as it seemed to be irrelevant for labour market outcomes.

$$\ln(Y_i) = \alpha + \beta' X_i + \sum_{j=1}^{3} \phi_j^l SFL_j^l + \varepsilon_i, \forall \, l = 1, 2, 3, 4, \tag{5}$$

where ϕ_j^l represents the return to having a skill level j (with $j = 1$ for basic, 2 for regular and 3 for advanced) in language l, relative to those who do not claim that language l is the FL they know better (holding the other control variables constant). Moreover, in line with the analysis of general FL knowledge, we also estimate the earnings equation that simultaneously controls for skills in each FL, which takes the form,

$$\ln(Y_i) = \alpha + \beta' X_i + \sum_{l=1}^{4} \sum_{j=1}^{3} \phi_j^l SFL_j^l + \varepsilon_i, \forall \, l = 1, 2, 3, 4. \tag{6}$$

Table 6 reports the results of the earnings regressions with skills dummies for each FL individually (eq. (5)) and simultaneously (eq. (6)). The results reveal that returns to different levels of proficiency in English are positive and increasing — ranging between 20% and 46%. Basic French skills are negatively associated with earnings and the coefficient for regular French skills is positive but imprecisely estimated, while having advanced French skills has a remarkably high return of 68%. By contrast, returns to German skills are insignificant. Russian skills are significantly remunerated only at the advanced level, with 27% increased earnings.

[TABLE 6]

Next, we include all the dummies for all the skill levels for all the FLs. In this case, the return to FL skills will be conditioned on competences in each of the (mutually exclusive) possible best FLs known and the reference category will be those individuals with no FL command. The results show that the return to English skills is almost unaffected by this exercise. However, basic French skills are no longer negative, while returns to regular and advanced French skills are significant and higher. Returns to basic and advanced German skills are both significant, and no change is observed in the return to advanced Russian

skills. Finally, adding occupation fixed-effects (and parental education) to this model produces modest reductions in returns to FL skills, indicating again that FL knowledge also affects earnings indirectly through occupational attainment.

Our analysis is complemented by several sensitivity exercises in which we focus more deeply on the economic value of English competences, since English is the most widely spoken FL in Turkey ― as well as in other non-English-speaking countries in Europe (see EU, 2006) ― and is commonly used as the *lingua franca* for commerce and trade. First, we consider the existence of heterogeneous returns according to observed and unobserved individual characteristics. Following the previous literature (among others, Saiz & Zoido, 2005; Lang & Siniver, 2009; Casale & Posel, 2011; Toomet, 2011; Azam et al., 2013), we allow for heterogeneous returns to English skills by several observed characteristics (i.e. frequency of language use at work, birth cohort, education, occupation, urban/rural location). The results (available upon request) reveal that returns to English skills in the Turkish labour market are positive and substantial for several subgroups of workers (i.e. we detect scarce heterogeneity in returns to English skills). Moreover, our results concerning the complementarity between English skills and formal schooling are somewhat at odds with those in the literature, since we obtain slightly higher returns for low educated workers (while previous evidence usually reported higher returns for more educated workers). Nevertheless, the main conclusion that can be drawn from the analysis of heterogeneous returns according to observable workers' characteristics is that the conditional positive correlation between English skills and earnings is in general stable and robust for individuals with different characteristics. As a further check regarding the existence of heterogeneous returns to English skills with respect to unobservable characteristics, we also allow the estimates to be different along the earnings distribution by applying quantile

regression. The results are fully described in section A2 of the Appendix and, in general, reinforce the robustness of the overall finding of this work. In the same Appendix (section A3), we also discuss the potential source of bias and limitations of our empirical strategy, together with an additional robustness check for the presence of misclassification errors in the language skills variable.

5.3 Discussion

Our analysis highlights that knowledge of FLs has a positive economic value in the Turkish labour market, which seems to be in part related to the occupational channel (i.e. those who master FLs are likely to be attracted into better paid occupations)[12]. The results are robust to the inclusion of controls for parental education. English competences clearly represent the most valuable asset. Advanced knowledge of Russian seems to be highly rewarded, although this evidence should be taken with caution because it is based on the few observations of Russian speakers in our sample. There are also positive rewards for mastering either French or German, although their economic value seems to be mostly linked to occupation rather than productivity within occupations and they seem to be complementary to English proficiency. On the contrary, knowing Arabic seems not to be rewarded in the labour market.

In general, the earnings premiums for FL knowledge are comparable, but somewhat lower, than the returns to different levels of education in Turkey provided in Tansel (1994, 2010). The return to English skills obtained for Turkey is consistent with those for other

---

[12] Modelling the complex relationship between English knowledge, occupation and earnings represents an interesting extension of the current work, which will be the object of future research.

developed and developing countries. The European evidence[13] obtained by Ginsburgh and Prieto-Rodríguez (2011) suggests that returns to English knowledge vary from 11% in Denmark, where English is widely spoken, to 49% in Spain, where speaking English as an FL is significantly less common. Regarding developing countries, Toomet (2011) reported a return to English skills of about 45% in Estonia and 62% in Latvia. Studies for South Africa report returns to English proficiency that range between 18% and 25% (Levinshon, 2007) and 41% and 44% (Casale & Posel, 2011). Finally, the results reported by Azam et al. (2013) show a 35% premium for advanced English skills for Indian men.

Moreover, we checked for the heterogeneity of returns to skills in English. They seem to be stable for different subgroups of workers. Further, the quantile regression estimates show increasing returns at higher quantiles of the conditional earnings distribution, but also a surprisingly high coefficient in the left tail, which is mostly due to language-related occupational selection. The estimates obtained when controlling for occupation seem to be almost flat along the earnings distribution. Therefore, our analysis with respect to observable and unobservable heterogeneity reveals that a positive economic value of English skills exists for several subgroups of workers and along the entire conditional distribution of earnings, which points to the overall significance and robustness of our results. The conclusions on the economic reward to English skills are that it is qualitatively unaffected when using a complementary econometric strategy that seeks to deal with the presence of errors in self-reported language skills variables and using the frequency of English use for leisure as a proxy for the level of skills.

---

[13] Williams (2011) also obtained significant returns to the use of English at work in several European countries. He highlighted that the use of other languages — especially French and German — is relevant in some countries, which can be considered to be in part consistent with our results regarding these two languages.

Even though there are reasons to consider that the positive return to an FL (and especially English skills) should not be entirely driven by omitted variable bias, this remains a relevant and unresolved issue. In fact, the unavailability of neither panel data nor plausibly exogenous instruments for FL knowledge and skills prevents the use of more adequate econometric techniques aimed at obtaining causal estimates in the presence of unobserved individual heterogeneity (correlated with FLs). Therefore, the results reported in this paper are, and should be taken as, "meaningful" conditional correlations ― just like those reported by Azam et al. (2013) for the case of India ― which highlight the positive economic value of FLs (and especially English skills) in the Turkish labour market.

6. Conclusions and Policy implications

Knowledge of FLs represents a form of human capital. Drawing on data from the 2007 AES, this is the first study to estimate the earnings returns to FL skills in Turkey — a country characterised by recent rapid economic and social development. The ongoing changes in the Turkish economy have fostered the relevance of and demand for FL competences in the labour market. However, the endowment of FL skills among the Turkish labour force seems to be rather scarce. Overall, this situation points to the existence of substantial pecuniary premiums for the command of FLs. Quantifying such returns represents the main purpose of this paper.

Examining returns to FLs is important since it will guide policymakers and individuals about how much to invest in fostering competences in FLs among current and future generations of workers. The results from our study suggest that acquiring competences in

FLs generally represents a profitable investment in the Turkish labour market. Returns to this investment are clearly positive at the individual level. Indeed, becoming proficient in English and, to a lesser extent French and German provides significant potential for higher earnings and, more generally, for better labour market performance, as FL knowledge seems to increase the chances of obtaining a better and more remunerated job. The high potential reward for Russian knowledge should be better investigated in the future due to the low number of cases in our sample. Therefore, proficiency in FLs has important implications in terms of labour market outcomes, since it improves employability, occupational prospects and earnings potential. Moreover, it seems to be plausible that the economic value of FL knowledge would be positive not only at the individual level, but also at the societal level.

The low level of human capital of (current and future) Turkish workers has been identified as one of the main concerns for the future development of the country (OECD, 2012). The performance of Turkish 15-year-olds in the PISA test is rather poor, as the country has ranked at the bottom of the test score distribution in the past two waves of the survey. The average 15-year-old student in Turkey is one full year behind the OECD average (World Bank, 2013). Enhancing human capital, the endowment of education and its equitable distribution among different socio-economic groups present current challenges for achieving and maintaining a sustainable path of growth and development in the mid- to long-term in Turkey and in other developing countries (Hanushek & Woessmann, 2008). Our results suggest that fostering FL skills should be taken as an additional challenge for Turkish policymakers. Increasing FL competences among the Turkish population would further promote international trade, internationalisation and openness in the Turkish economy as well as R&D activities and innovation. In turn, this would generate greater

potential for growth and socio-economic development, improving the nation's position in the global knowledge economy.

Indeed, improving English skills among the population would be especially beneficial for a mid-sized developing country such as Turkey, since it may help reduce existing disparities in global competition among emerging economies for international trade and attracting new FDI. This is extremely relevant in light of the significant scale and resource advantages of the two leading Asian emerging countries, India and China. In fact, in the former, English represents a former colonial language that is co-official and widely spoken among the population, especially among the highly educated, and the latter has the largest English-learner population in the world (Crystal, 2008; He & Li, 2009). Moreover, fostering competences in English as well as in other relevant European languages might be especially important in Turkey for two additional reasons. First, given the geographical location of the country, this could favour its historical role as a "bridge" for commodities trading between Asia and Europe. Second, reducing language barriers would be especially relevant for further attenuating the already reduced cultural barriers between Turkey and EU countries, which might represent an additional stimulus for commerce and trade.

Several policy implications can be directly advocated in light of our results, which can be reasonably extrapolated to other developing countries as well as to developed countries with an insufficient endowment of FL skills in their labour forces. First, policymakers should emphasise the teaching of English at schools in order to increase the English proficiency of future generations of workers. This would be especially important due to growing demand for FL competences in the Turkish labour market in the near future, with the prospect of further economic growth and development and possible accession to the EU. The 1997 Turkish education reform increased the amount of FL teaching during the

schooling process. However, there is no evidence for the effectiveness of these reforms in improving the FL proficiency of students from different grades. Therefore, evaluating the effect of the 1997 reform on English proficiency represents an object of our future research. Moreover, the government should also foster English teachers' training and professional requirements, since teachers play a fundamental role in guaranteeing the effectiveness of the above-mentioned educational reforms. Second, for the current generation of workers, future public policies should encourage and subsidise their attendance at private FL centres. This is a sensible approach as our findings point out a certain, albeit low, substitutability between English skills and general schooling for the young. In fact, beyond earnings, FL skills may also enhance employability and labour market opportunities for young, low educated individuals who may come from disadvantaged socio-economic and family backgrounds.

Here, we emphasise English as the FL to be taught, not only because of its international value and relevance compared with other languages, but also because there currently is a substantial stock of English language teachers, albeit still less than the demand for them. Teaching Russian in schools would take time to accomplish, because of the need to train teachers first. The current demand for Russian-speaking workers could be met by teaching Russian in special schools such as tourism schools or at FL centres. There may also be some space for policies aimed at improving competences in German and French. However, owing to our less conclusive results regarding these two languages, and given the hegemony of English as the *lingua franca*, less priority should be given to investments in these languages.

By contrast, our findings suggest that there is no earnings premium for knowledge of Arabic in the Turkish labour market. However, we consider that more detailed data are

26

needed to further disentangle this puzzling evidence, which should enable separating the human capital component from the ethnicity-related attributes that the Arabic knowledge variable (as collected in Turkish AES data) might be picking up. Also noteworthy is the absence of Chinese language instruction in Turkey, except a couple of recently introduced university programs. Chinese language instruction could be important given the recent increases in the volume of trade with China[14].

Further, since most productive potential of FL skills is expected to be allocated to the private sector, especially among firms exposed to English-intensive activities such as international trade, R&D, ICT and tourism, private businesses should contribute to financing FL training for their workforce and complement the government's public investment. Finally, Turkey should be able to benefit more from the language competences of its citizens with immigrant backgrounds, such as the growing population of return migrants from Germany attracted by the current economic development path of the country.

---

[14] China was Turkey's 14th largest export partner and the 3rd largest import partner in 2012 and Chinese FDI stock in Turkey is becoming substantial (TURKSTAT, 2013a). Therefore, Chinese proficiency might have a significant labour market value in Turkey. Unfortunately, we were unable to quantify the return to Chinese knowledge, since just one Chinese speaker appeared in our sample.

References

Adak, M. (2010). Foreign trade and economic growth: The case of Turkey. *Middle Eastern Finance and Economics*, 8, 137-145.

Adesope, O. O., Lavin, T., Thompson, T., Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, 80, 207-245.

Angrist, J. D., & Lavy, V. (1997). The effect of a change in language of instruction on the returns to schooling in Morocco. *Journal of Labor Economics*, 15, 48-76.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. New Jersey: Princeton University Press.

Aydın, Y. (2012). Emigration of highly qualified Turks: A critical review of the societal discourses and social scientific research. In S., Paçacı-Elitok & T. Straubhaar (Eds), *Turkey, Migration and the EU: Potentials, Challenges and Opportunities* (pp. 199-227). Hamburg: Edition HWWI, Institute of International Economics (HWWI).

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

Bleakley, H., & Chin, A. (2004). Language Skills and Earnings: Evidence from Childhood Immigrants. *The Review of Economics and Statistics*, 86, 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, 385-393.

Cebeci, T., & Fernandes, A. M. (2013). *Micro dynamics of Turkey's export boom in the 2000s* (Working Paper No. 6452). Washington DC: The World Bank.

Chiswick, B. R., & Miller, P. W. (2009). Earnings and occupational attainment among immigrants. *Industrial Relations*, 48, 454-465.

Chiswick, B. R., Miller, P. W., (2010). Occupational language requirements and the value of English in the US labor market. *Journal of Population Economics*, 23, 353-372.

Chiswick, B. R., & Miller, P. W. (2013). The impact of surplus skills on earnings: Extending the over-education model to language proficiency. *Economics of Education Review*, 36, 263-275.

Chiswick, B. R., & Miller, P. W. (2014). International Migration and the Economics of Language. In B. R.

Chiswick & P. W. Miller (Eds.). *Handbook of the Economics of International Migration 1A* (in press). Amsterdam: North Holland.

Crystal, D. (2008). Two thousand million? *English Today*, 24, 3-6.

Education First, (2011). Education First English Proficiency Index. Retrieved from www.ef.com/epi.

European Union, (2006). Europeans and their languages. Special Euro-Barometer 386/Wave EB77.1. Retrieved from http://ec.europa.eu/languages/documents/2006-special-eurobarometer-survey-64.3europeans-and-languages-report_en.pdf.

Fidrmuc, J. (2011). The economics of multilingualism in the EU. In T. Eger & H. B. Schäfer (Eds.). *Research Handbook on the Economics of European Union Law* (pp. 331-354). Cheltenham : Edward Elgar.

Fidrmuc, J., & Fidrmuc, J. (2009). *Foreign languages and trade* (Working Paper No. 7228). London: Centre for Economic Policy Research**.**

Ginsburgh, V. A., & Prieto-Rodriguez, J. (2011). Returns to foreign languages of native workers in the EU, *Industrial and Labor Relations Review*, 64, 599-617.

Grin, F., Sfreddo, C. & Vaillancourt, F. (2010). *The Economics of the Multilingual Workplace*. New York: Routledge.

Hanushek, E.A., & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, 46, 607-68.

He, D., & Li, D. C. S. (2009). Language attitudes and linguistic features in the 'China English' debate. *World Englishes*, 28, 70-89.

Hoon, C. O., Selmier, W. T., & Lien, D. (2011). International trade, foreign direct investment, and transaction costs in languages. *The Journal of Socio-Economics*, 40, 732-735.

Hutchinson, W. K., & William K. (2005). "Linguistic Distance" as a Determinant of Bilateral Trade. *Southern Economic Journal*, 72, 1–15.

Isphording, I., & Otten, S. (2013). The costs of Babylon – Linguistic distance in applied economics. *Review of International Economics*, 21, 354-369.

Koenker, R. (2005). *Quantile Regression*. Cambridge: Cambridge University Press.

Ku, H., & Zussmann, A. (2010). Lingua Franca: The role of English in international trade. *Journal of Economic Behavior and Organization*, 75, 250-260.

Lang, K., & Siniver, E. (2009). The return to English in a non-English speaking country: Russian immigrants and native Israelis in Israel. *The B.E. Journal of Economic Analysis and Policy*, 9, 1-30.

Leslie, D., Russell, H., & Forbes, A. (2002). Foreign language skills and tourism management courses in the UK. *Industry and Higher Education*, 16, 403-414.

Levinsohn, J. (2007). Globalization and the returns to speaking English in South Africa. In A. E. Harrison (Ed.), Globalization and Poverty (pp. 629-646) Chicago: University of Chicago Press.

Melitz, J. (2008). Language and foreign trade. *European Economic Review*, 52, 667-699.

Organization for Economic Co-Operation and Development (2012). *OECD Economic Surveys: Turkey.* Retrieved from http://www.oecd-ilibrary.org/economics/oecd-economic-surveys-turkey-2012_eco_surveys-tur-2012-en.

Saiz, A., & Zoido, E. (2005). Listening to what the world says: Bilingualism and earnings in the United States. *The Review of Economics and Statistics*, 87, 523-538.

Seargeant, P., & Erling, E. (2011). The discourse of English as a language for international development: policy assumptions and practical challenges. In H. Coleman (Ed.), *Dreams and Realities: Developing Countries and the English Language* (pp. 2-21). London : British Council.

State Planning Organization (2006). *Ninth Five-Year Development Plan 2007-2013*. Retrieved from http://www.maliye.gov.tr/Lists/TabMenuIcerik/Attachments/106/9developmentplan.pdf.

Tansel, A. (1994). Wage employment, earnings and returns to schooling for men and women in Turkey. *Economics of Education Review*, 13, 305-320.

Tansel, A. (2010). *Changing returns to education for men and women in a developing country: Turkey, 1994-2005*. Paper presented at the European Society for Population Economics Conference (London) and at the Middle East Economic Association Conference (Nice).

Tansel, A., & Yaşar, P. (2010). Macroeconomic impact of remittances on output growth: Evidence from Turkey. *Migration Letters*, 7, 132-143.

Toomet, O. (2011). Learn English, Not the local language! Ethnic Russians in the Baltic States. *American Economic Review: Papers and Proceedings*, 101, 526-531.

Tucci, I., & Wagner, G. (2004). Foreign language skills- an important additional qualification in the services sector. *Economic Bulletin*, 41, 43-46.

Turkish Statistical Institute (TURKSTAT) (2010). *Adult Education Survey 2007* (Report No. 3433). Retrieved from http://www.tuik.gov.tr/Kitap.do?metod=KitapDetay&KT_ID=5&KITAP_ID=218.

Turkish Statistical Institute (TURKSTAT) (2012). *Informal Education in Turkey*. Ankara: Turkish Statistical Institute Publications.

Turkish Statistical Institute (TURKSTAT), (2013a) Foreign Trade Statistics. Retrieved from www.tuik.gov.tr.

Turkish Statistical Institute (TURKSTAT) (2013b). Tourism Statistics. Retrieved from www.tuik.gov.tr.

United Nations Conference on Trade and Development (2013). World Investment Report. Retrieved from http://unctad.org/en/PublicationsLibrary/wir2013_en.pdf.

Williams, D. (2011). The economic returns to multiple language usage in Western Europe. *International Journal of Manpower*, 32, 372-393.

World Bank, (2013). *Promoting excellence in Turkey's schools. Human Development Sector Unit Europe and Central Asia Region* (Report No: 77722). Washington DC: The World Bank.

Yağmur, K. (2001). Languages in Turkey. In G. Extra & D. Gorter (Eds.), *The Other Languages of Europe* (pp. 407-428). Clevedon: European Cultural Foundation.

TABLES

Table 1: Foreign language knowledge among Turkish male employees

	no Foreign Languages	one Foreign Language	two Foreign Languages	three or more FLs	TOTAL (%)
BIRTH COHORT					
cohort 25-39	61.72	33.78	3.75	0.74	100
cohort 40-65	67.96	27.56	4.12	0.36	100
COMPLETED EDUCATION					
low-education	92.91	6.26	0.67	0.16	100
medium-education	56.48	39.4	3.59	0.53	100
high-education	18.19	69.09	11.09	1.63	100
OCCUPATION					
high-skilled white collars	33.92	56.78	8.09	1.22	100
low-skilled white collars	60.81	34.18	4.16	0.85	100
high-skilled blue collars	83.27	15.4	1.18	0.15	100
low-skilled blue collars	80.77	17.36	1.7	0.17	100
URBAN/RURAL AREAS					
urban	63.04	32.63	3.74	0.58	100
rural	68.02	26.84	4.47	0.66	100
TOTAL (%)	64.04	31.47	3.89	0.6	100
NUMBER OF OBSERVATIONS	3854	1894	234	36	6018

Table 2: Foreign languages spoken (among foreign language speakers, N = 2164)

	English	French	German	Arabic	Russian	Others
BIRTH COHORT						
cohort 25-39	86.5	4.2	10.4	7.2	2.6	1.4
cohort 40-65	73.9	11.6	13.1	11.6	1.5	2.0
COMPLETED EDUCATION						
low-education	47.8	3.3	14.4	28.3	5.6	11.1
medium-education	82.0	6.8	9.9	7.4	2.8	1.3
high-education	88.9	7.1	12.2	6.3	1.2	0.1
OCCUPATION						
high-skilled white collars	85.4	7.4	11.8	7.8	1.7	0.6
low-skilled white collars	85.7	5.4	11.5	6.7	3.0	0.9
high-skilled blue collars	76.2	5.7	9.7	9.7	2.2	3.5
low-skilled blue collars	71.7	6.5	10.6	13.3	3.2	4.4
URBAN/RURAL AREAS						
urban	84.0	6.7	11.0	6.7	2.3	1.2
rural	74.6	6.2	12.7	17.4	2.1	3.4
TOTAL (%)	82.3	6.7	11.3	8.6	2.3	1.6

Table 3: Foreign language skills (among foreign language speakers, N = 2164)

	% over FL speakers	*basic skills*	*regular skills*	*advanced skills*
first FL = English	79.44			
		55.03	*31.47*	*13.50*
first FL = French	4.85			
		73.33	*20.00*	*6.67*
first FL = German	7.44			
		50.93	*31.06*	*18.01*
first FL = Arabic	5.87			
		20.47	*38.58*	*40.94*
first FL = Russian	0.74			
		12.50	*75.00*	*12.50*
first FL = Other	1.66			
		27.78	*19.44*	*52.78*

Table 4: Distribution of foreign languages and net monthly earnings (in Turkish liras)

Income deciles	All the sample	at least one FL	English	French	German	Arabic	Russian	Other FL
% Q$_1$ (top decile)	9.39	4.0	3.4	0.69	2.86	10.7	4.08	20.59
% Q$_2$	10.14	5.5	4.9	6.25	5.71	8.56	2.04	8.82
% Q$_3$	12.38	6.9	6.5	3.47	4.9	8.02	10.2	17.65
% Q$_4$	7.76	5.4	5.4	3.47	5.31	5.88	4.08	5.88
% Q$_5$	12.4	9.2	9.7	5.56	9.8	6.42	10.2	2.94
% Q$_6$	7.68	6.8	6.3	6.25	8.16	4.81	8.16	23.53
% Q$_7$	13.78	15.7	15.4	13.19	13.88	25.67	18.37	2.94
% Q$_8$	7.98	12.6	12.2	19.44	11.02	15.51	14.29	5.88
% Q$_9$	10.5	17.9	18.3	21.53	17.55	10.16	10.2	11.76
% Q$_{10}$ (bottom decile)	8.01	16.1	17.9	20.14	20.82	4.28	18.37	--
Total (%)	100	100	100	100	100	100	100	100
average monthly earnings (in Turkish Liras)	923.9	1255.9	1280.2	1543.6	1332.9	962.1	1230.8	713.2

Table 5: foreign language augmented earnings regressions

	Number of foreign languages	English	French	German	Arabic	Russian	all FLs	all FLs	all FLs
no foreign languages	*ref. cat.*								
one foreign language	0.075^a								
	(0.016)								
two foreign languages	0.147^a								
	(0.037)								
three or more foreign languages	0.376^a								
	(0.096)								
knows English		0.107^a					0.111^a	0.086^a	0.072^a
		(0.016)					*(0.016)*	*(0.015)*	*(0.015)*
knows French			0.084^c				0.103^b	0.090^b	0.089^b
			(0.044)				*(0.043)*	*(0.041)*	*(0.040)*
knows German				0.082^b			0.085^a	0.065^b	0.056^c
				(0.032)			*(0.032)*	*(0.031)*	*(0.030)*
knows Arabic					-0.066^b		-0.066^b	-0.088^b	-0.067^c
					(0.033)		*(0.034)*	*(0.034)*	*(0.034)*
knows Russian						0.199^a	0.141^b	0.136^b	0.129^c
						(0.069)	*(0.070)*	*(0.068)*	*(0.067)*
occupation fixed effects	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*yes*	*yes*
number of observations	6018	6018	6018	6018	6018	6018	6018	6018	6018
Adjusted R^2	0.351	0.351	0.346	0.347	0.346	0.347	0.353	0.410	0.417

Note: OLS regressions with robust standard errors within parenthesis in italic. [c] Significant at p<0.1, [b] significant at p<0.05, [a] significant at p<0.01. All the models include controls for years of schooling, potential experience (quadratic), type of contract and a dummy for urban area (see table 2A in the Appendix for the estimates of other control variables). Indicators for knowledge of each FL are not mutually exclusive (i.e. individuals may declare to know more than one language).

Table 6: Return to foreign language skills

	English skills	French skills	German skills	Russian skills	all FLs skills	all FLs skills	all FLs skills
reference category:	*English no first FL*	*French no first FL*	*German no first FL*	*Russian no first FL*	*No foreign languages*		
basic English skills	0.005				0.019	0.015	0.011
	(0.017)				*(0.018)*	*(0.017)*	*(0.017)*
regular English skills	0.196[a]				0.213[a]	0.175[a]	0.160[a]
	(0.025)				*(0.026)*	*(0.025)*	*(0.025)*
advanced English skills	0.456[a]				0.475[a]	0.396[a]	0.366[a]
	(0.034)				*(0.035)*	*(0.035)*	*(0.036)*
basic French skills		-0.079[b]			0.001	0.011	0.014
		(0.039)			*(0.040)*	*(0.038)*	*(0.038)*
regular French skills		0.171			0.259[b]	0.192[c]	0.208[b]
		(0.115)			*(0.117)*	*(0.104)*	*(0.100)*
advanced French skills		0.679[c]			0.792[c]	0.640	0.575
		(0.403)			*(0.404)*	*(0.402)*	*(0.385)*
basic German skills			0.004		0.078[c]	0.067[c]	0.063
			(0.040)		*(0.041)*	*(0.038)*	*(0.038)*
regular German skills			-0.078		0.002	-0.035	-0.034
			(0.065)		*(0.065)*	*(0.063)*	*(0.063)*
advanced German skills			0.113		0.189[b]	0.174[b]	0.160[c]
			(0.085)		*(0.088)*	*(0.087)*	*(0.088)*
basic Russian skills				0.626	0.636	0.597	0.578
				(0.464)	*(0.452)*	*(0.434)*	*(0.437)*
regular Russian skills				0.174	0.213	0.234[c]	0.225[c]
				(0.145)	*(0.140)*	*(0.135)*	*(0.136)*
advanced Russian skills				0.272[a]	0.280[a]	0.228[a]	0.252[a]
				(0.030)	*(0.017)*	*(0.039)*	*(0.068)*
parental education	*no*	*no*	*no*	*no*	*no*	*no*	*yes*
occupation fixed effects	*no*	*no*	*no*	*no*	*no*	*yes*	*yes*
number of observations	6018	6018	6018	6018	6018	6018	6018
Adjusted R^2	0.370	0.348	0.346	0.346	0.374	0.423	0.428

*Note: OLS regressions with robust standard errors within parenthesis in italic. [c] Significant at p<0.1, [b] significant at p<0.05, [a] significant at p<0.01. All the models include controls for years of schooling, potential experience (quadratic), type of contract and a dummy for urban area. FL language skills refer to the (mutually exclusive) FL the individual knows better. *No foreign languages (reference category in the last three columns) means that neither English, French, German or Russian are the first FL that the individual knows.*

Appendix

A1 Data and Sample for the Empirical Analysis

The empirical analysis is based on nationally representative Turkish data from the AES. This survey contains detailed information on FL knowledge, skill levels and use together with socio-demographic and labour market characteristics. The sample includes 39,478 individuals aged 18 and over. We restrict the sample to men aged 25–65 who were regularly employed as wage-earners at the time of the survey. Women are excluded from the analysis in order to avoid problems of endogenous selection into labour market participation and employment. We consider individuals aged between 25 and 65 because regular schooling is usually completed before 25 years of age and can therefore be taken as exogenous, which helps limit the potential endogeneity of schooling in the earnings regressions.

We retain wage-earners since this facilitates the interpretation of the earnings regressions, even though the results obtained by also considering employees and self-employed workers are virtually the same. Selection into wage employment among men could also be an issue. For this reason, we controlled for endogenous selection into wage employment among men and the results were virtually unchanged (full results are available upon request). Therefore, we decided to focus on male wage-earners, implying that we aim to provide evidence that is consistent for the selected sample. Part-time employees and migrants were excluded, as were observations with missing information. The final sample was made of 6,018 male employees.

A2 Quantile Regression Estimates

The empirical analysis of this work considered earnings differences associated with FL skills at the mean of the conditional wage distribution. Therefore, with the aim of going a step further, we now investigate returns to English skills along the conditional wage distribution using quantile

regression analysis[15], as previously carried out by Ginsburgh and Prieto-Rodríguez (2011) and Toomet (2011). This is especially important since acquiring English skills might change not only the location but also the shape of the conditional earnings distribution. Further, as conditional earnings differentials proxy for unobservable earnings potential, quantile regression sheds light on the differential returns by unobserved individual heterogeneity. The corresponding representation of equation (5) for each quantile θ would be,

$$\ln(Y_i) = \alpha^\theta + \beta^{\theta\prime}X_i + \sum_{j=1}^{3} \phi_j^\theta SFL_j + \varepsilon_i^\theta, \tag{7a}$$

$$Q_\theta(\ln(Y_i)\,|X_i, SFL) = \alpha_\theta + \beta'_\theta X_i + \sum_{j=1}^{3} \phi_j^\theta SFL_j \ , \forall\, \theta \in (0,1) \tag{7b}$$

where $Q_\theta(\cdot)$ represents the conditional earnings distribution (assuming that the θth conditional quantile of ε_i is equal to zero) and ϕ_j^θ are the coefficients associated with each level of skills in English, which are allowed to be different for each quantile.

The quantile regression estimates are reported in Table A3 of this Appendix. Because the heteroskedasticity test rejects the null hypothesis of constant residual variances in all cases, we therefore report robust standard errors (Machado et al., 2011). The first panel contains the results for the baseline set of controls. The return to basic English skills is virtually zero in any quantile, while returns to regular and advanced English competences are positive and increasing across different quantiles. The F tests reject the null hypothesis of constant coefficients along the earnings distribution. The results show that workers with higher earnings potential obtain higher returns to English proficiency than those who are in the middle of the earnings distribution, which is in line with the results obtained in several European countries (Ginsburg & Prieto-Rodríguez, 2011) as well as for Estonia and Latvia (Toomet, 2011). However, the estimated return to advanced English skills in the first decile of the conditional earnings distribution represents a noticeable exception, since the coefficient at the lowest decile of the distribution is virtually the same as the return estimated at the top of the distribution.

---

[15] See Koenker (2005) for additional details about quantile regression methods.

Nevertheless, when occupation and parental education are controlled for, the return to regular English becomes constant along the wage distribution and the increasing returns to advanced competences are clearly attenuated. The F tests for coefficient equality along the earnings distribution for any skill level are thus not rejected. The modest heterogeneity in returns according to unobservable earnings potential seems to be mostly linked to occupation-related selection based on unobservable traits. This evidence reconfirms that mastering FLs is especially important in specific occupations, mostly at the upper percentiles of the earnings distribution. Disentangling the effect of selection into occupation is, however, a complicated task that requires more detailed and complete data, and we defer a deeper analysis of this issue to future research. In general, quantile regression analysis reveals that the earnings premium associated with a command of English is positive and substantial across the entire conditional distribution of earnings, reinforcing the robustness of the overall findings from this research.

A3 Potential Sources of Bias in OLS Regressions: Omitted Variables and Misclassification Errors

The results reported in this paper show that, generally, FL knowledge is associated with higher labour market earnings. More specific results regarding the return to English skills are extremely robust, indicating that the conditional relationship between competences in English and earnings is positive and substantial. However, as in the case of the return to general FL knowledge, there might be some concern about the existence and extent of bias in these conditional correlations. In fact, the estimates of the return to language skills might be biased because of (uncontrolled) unobserved heterogeneity affecting language skills and earnings potential.

Previous evidence suggests that the return to FL skills is still positive when estimated by using either the individual fixed-effects or instrumental variable methods (see Saiz & Zoido, 2005; Lang & Siniver, 2009; Ginsburgh & Prieto-Rodríguez, 2011, among others). In our case, the stability of the results regarding returns to English skills supports our argument that the earnings premium associated with an increasing level of competences in this FL should not be entirely driven by omitted variable bias. However, as commented on above, we are unable to explicitly deal with this potential issue, since we do not dispose of suitable instrumental variables for English skills. In the case of a positive correlation between omitted ability and English skills, the OLS coefficients from eq. (5) would be upward biased. In this case, the estimates obtained in this work would represent an upper bound of the true causal parameters. However, omitted variable bias is not the only issue in this kind of estimation, since the self-reported nature of the FL skills variable might introduce additional bias due to possible misclassification errors coming from the tendency to over-report language competences. Dustmann and van Soest (2001, 2002, 2004) provided detailed discussions on this issue in the context of the earnings return to a host country's language proficiency among immigrants. Their results also suggested that misclassification errors due to over-reporting of language competences tend to introduce a downward bias in the OLS estimates of the return to language proficiency, which might overcome the potential upward bias due to omitted variables.

Based on this last concern, in what follows we apply a methodology to remove misclassification errors in self-reported language skills, which is close to the strategy implemented by Aldashev et al. (2009). Specifically, we replace the indicators of increasing levels of English skills in the earnings equation (eq. (8a)) with its predicted value ($\widetilde{SFL}^*$) from an Ordered Probit model (eqs. (8b) and (8c)) that estimates the probability of having a certain degree of (self-assessed) competences in English as a function of observed individual characteristics. Therefore, the model to be estimated[16] takes the form,

---

[16] We estimate the joint model by LIML using the STATA routing "cmp" developed by Roodman (2011). Note that the error terms of the two equations (ε_i and u_i) are allowed to be correlated and are assumed to follow a bivariate normal distribution.

$$\ln(Y_i) = \alpha + \beta' X_i + \rho \widetilde{SFL}^* + \varepsilon_i \tag{8a}$$

$$SFL^* = \pi' Z_i + u_i \tag{8b}$$

$$SFL = \begin{cases} 0 \ if \ SFL < \mu_1 \\ 1 \ if \ \mu_1 \le SFL < \mu_2 \\ 2 \ if \ \mu_2 \le SFL < \mu_3 \\ 3 \ if \ SFL \ge \mu_3 \end{cases} \tag{8c}$$

where the μ terms represent the cut-off points to be estimated and ρ is the parameter of interest associated with the "latent" propensity to have higher skills in English (i.e. the predicted latent index from the Ordered Probit for English skills) that, in principle, should not be affected by the presence of misclassification errors.

The explanatory variables considered in the vector Z are the same controls previously used for the earnings equations. In addition, we exploit information on the frequency of English use in leisure activities as a proxy for the level of English skills. We assume that the frequency of language use for leisure purposes 1) predicts the propensity of being more skilled in English and 2) is related to labour market earnings only through its direct effect on English competences (i.e. it is conditionally unrelated to the error term of the earnings equation). The validity of condition 1) can be directly inferred from the data. The second hypothesis cannot be directly tested. Moreover, it might be argued that using English more frequently in daily life would allow access to better social networks and thus increase the chances of obtaining a better remunerated job in which English represents a valuable asset. In this case, the second assumption would not be valid. However, we consider that controlling for the frequency of English use at work in the earnings equation would break this potential link between English use in leisure activities and unobserved earnings potential[17].

---

[17] In other words, if speaking English more frequently with friends, relatives and in daily life provides access to jobs in which English is more important (and in principle used more often), this would be mostly picked up by the included dummies for the frequency of English use at work.

The estimates from the joint model for self-reported English skills and (log)earnings (eqs. (8a) and (8b)) are reported in Table A4 of this Appendix. The results of the Ordered Probit for English skills are quite standard and not discussed here for space reasons. The earnings equation now contains as its right-hand side variable the predicted linear index from the Ordered Probit, which captures the latent propensity to have higher competences in English. The first specification does not control for the frequency of language use at work and, as expected, displays a significant and positive effect of latent English skills on earnings. However, this estimate might be picking up some of the indirect relationship between unobserved earnings potential (i.e. the error term of the earnings equation) and the frequency of English use in leisure activities. When we do control for the frequency of English use at work, with the aim of breaking this correlation, we still obtain a significant effect of latent English skills, but somewhat reduced (as expected). As observed for the original English skills variable, the return to latent English proficiency declines when occupation fixed-effects are included in the list of control variables and it presents a further reduction after controlling for parental education. However, from a qualitative point of view, the results obtained from this methodology are similar to those resulting from the OLS estimations, since their main message remains the same: English skills (observed or latent) are positively rewarded in the labour market (also) in the Turkish case.

[TABLE 4A]

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, 330-341.

Dustmann, C., & van Soest, A. (2001). Language fluency and earnings: Estimation with misclassified language indicators. *The Review of Economics and Statistics*, 83, 663-674.

Dustmann, C., & van Soest, A. (2002). Language and the earnings of immigrants. *Industrial and Labor Relations Review*, 55, 473-92.

Dustmann, C., & van Soest, A. (2004). An analysis of speaking fluency of immigrants using ordered response models with classification errors. *Journal of Business and Economic Statistics*, 22, 312-321.

Machado, J., Parente, P., & Santos Silva, J. (2011). *QREG2: Stata module to perform quantile regression with robust and clustered standard errors*. Retrieved from: http://ideas.repec.org/c/boc/bocode/s457369.html.

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. *Stata Journal*, 11, 159-206.

Tables for Appendix

Table 1A: Descriptive statistics and variables construction

	ALL THE SAMPLE		NO FL		AT LEAST ONE FL		BEST FL = ENGLISH	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
CONTROL VARIABLES								
years of schooling	8.83	3.86	7.06	3.11	12.00	2.95	12.37	2.67
age	37.15	8.47	37.60	8.48	36.35	8.40	35.47	8.23
potential experience (= years of schooling – age – 6)	22.32	9.71	24.54	9.45	18.35	8.85	17.10	8.49
TYPE OF CONTRACT								
employee/permanent contract	0.909	0.287	0.885	0.319	0.953	0.212	0.963	0.189
employee/fixed-term contract	0.091	0.287	0.115	0.319	0.047	0.212	0.037	0.189
GEOGRAPHICAL LOCATION								
rural area (< 20000 inhabitants)	0.201	0.400	0.213	0.409	0.178	0.383	0.157	0.364
urban area (≥ 20000 inhabitants)	0.799	0.400	0.787	0.409	0.822	0.383	0.843	0.364
PARENTAL EDUCATION (HIGHEST ATTAINMENT AMONG BOTH PARENTS)								
parental education = no education	0.286	0.452	0.151	0.358	0.525	0.499	0.540	0.499
parental education = primary or less	0.196	0.397	0.186	0.389	0.213	0.410	0.225	0.417
parental education = secondary	0.225	0.418	0.293	0.455	0.105	0.306	0.097	0.295
parental education = tertiary	0.293	0.455	0.369	0.483	0.157	0.364	0.138	0.345
OCCUPATION								
high-skilled white collars	0.286	0.452	0.355	0.479	0.164	0.370	0.126	0.332
low-skilled white collars	0.621	0.485	0.613	0.487	0.634	0.482	0.640	0.480
high-skilled blue collars	0.054	0.225	0.023	0.149	0.109	0.312	0.124	0.330
low-skilled blue collars	0.039	0.194	0.009	0.095	0.093	0.290	0.109	0.312
FREQUENCY OF ENGLISH USE AT WORK AND FOR LEISURE								
no English skills	0.714	0.452	--	--	0.206	0.404	--	--
English not used at work	0.187	0.390	--	--	0.519	0.500	0.654	0.476
English used less than once per month at work	0.020	0.141	--	--	0.056	0.231	0.071	0.257
English used at least once per month at work	0.021	0.144	--	--	0.059	0.236	0.074	0.263
English used at least once per week at work	0.024	0.153	--	--	0.067	0.249	0.084	0.277
English daily used at work	0.033	0.180	--	--	0.093	0.290	0.117	0.321
English used less than once per month for leisure	0.210	0.408	--	--	0.585	0.493	0.737	0.440
English used at least once per month for leisure	0.029	0.167	--	--	0.080	0.271	0.100	0.300
English used at least once per week for leisure	0.015	0.121	--	--	0.041	0.199	0.052	0.222
English daily used for leisure	0.018	0.133	--	--	0.050	0.219	0.063	0.244

Definition of years of schooling: illiterate = 0 years; literate with no formal education = 2; uncompleted primary education = 3.5; completed primary school = 5; uncompleted middle school = 6.5; completed middle school = 8; uncompleted high school = 9.5; completed high school = 11; uncompleted short college degree = 12; completed short college degree = 13; uncompleted college degree = 14; completed college degree = 15; uncompleted PhD = 17; completed PhD = 19.

Definition of language skills: the AES questionnaire contemplate the following four self-reported levels of command of foreign languages: 1 "I can understand and speak a few words and sentences", 2 "I can understand and use the most general daily expressions", 3 "in the instances where the language is used in a clear fashion, I can understand the essence and express the experiences and events in a printed text" and 4 "I can understand and use the language in a flexible (fluent) manner in various subjects involving a series of difficult texts. I am almost completely competent in this language". Given the low number of cases for levels 3 and 4, these last two levels have been grouped into one single category. Therefore, in the empirical analysis we will use 3 separate levels of skills: 1) basic skills (corresponding the original level 1 in the survey), 2) regular skills (corresponding to level 2) and advanced skills (corresponding to either level 3 or 4 in the AES questionnaire). The level of skills refers to the specific language the individuals know better (if any).

Table 2A: Foreign language augmented earnings regressions (control variables estimates)

	Number of foreign languages	English	French	German	Arabic	Russian	all FLs	all FLs	all FLs
constant	5.293[a]	5.281[a]	5.285[a]	5.280[a]	5.272[a]	5.273[a]	5.301[a]	5.303[a]	5.236[a]
	(0.046)	*(0.046)*	*(0.047)*	*(0.047)*	*(0.046)*	*(0.046)*	*(0.046)*	*(0.107)*	*(0.105)*
years of schooling	0.079[a]	0.079[a]	0.086[a]	0.085[a]	0.087[a]	0.086[a]	0.078[a]	0.058[a]	0.055[a]
	(0.002)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*	*(0.003)*
potential experience	0.037[a]	0.038[a]	0.036[a]	0.036[a]	0.037[a]	0.036[a]	0.038[a]	0.035[a]	0.037[a]
	(0.003)	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*	*(0.003)*
(potential experience)^2	-0.001[a]	-0.001[a]	-0.001[a]	-0.001[a]	-0.001[a]	-0.001[a]	-0.001[a]	-0.001[a]	-0.001[a]
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
fixed term contract	-0.296[a]	-0.295[a]	-0.292[a]	-0.294[a]	-0.293[a]	-0.294[a]	-0.296[a]	-0.268[a]	-0.258[a]
	(0.026)	*(0.026)*	*(0.026)*	*(0.026)*	*(0.026)*	*(0.026)*	*(0.026)*	*(0.027)*	*(0.027)*
urban area	0.170[a]	0.166[a]	0.167[a]	0.168[a]	0.165[a]	0.167[a]	0.164[a]	0.127[a]	0.123[a]
	(0.015)	*(0.015)*	*(0.015)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.016)*	*(0.015)*	*(0.015)*
parental education = no education				*reference category*					
parental education = primary									0.078[a]
									(0.014)
parental education = secondary									0.122[a]
									(0.032)
parental education = tertiary									0.278[a]
									(0.040)
occupation fixed effects	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*yes*	*yes*
number of observations	6018	6018	6018	6018	6018	6018	6018	6018	6018
Adjusted R^2	0.351	0.351	0.346	0.347	0.346	0.347	0.353	0.410	0.417

Note: OLS regressions with robust standard errors within parenthesis in italic. [c] Significant at p<0.1, [b] significant at p<0.05, [a] significant at p<0.01.

Table 3A: Heterogeneous returns to foreign language skills — conditional wage distribution

	$Q_{0.1}$	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	$Q_{0.9}$	*F test (Prob.)*
no English skills			*reference category*			
basic English skills	-0.007	0.006	0.014	0.001	-0.013	*0.84*
	(0.025)	(0.026)	(0.018)	(0.020)	(0.029)	
regular English skills	0.156[a]	0.170[a]	0.148[a]	0.226[a]	0.294[a]	*0.04*
	(0.051)	(0.030)	(0.026)	(0.032)	(0.052)	
advanced English skills	0.520[a]	0.362[a]	0.423[a]	0.512[a]	0.527[a]	*0.02*
	(0.046)	(0.056)	(0.041)	(0.039)	(0.080)	
parental education	*no*	*no*	*no*	*no*	*no*	
occupation fixed effects	*no*	*no*	*no*	*no*	*no*	
pseudo R2	0.341	0.368	0.369	0.367	0.356	
no English skills			*reference category*			
basic English skills	-0.006	0.008	0.013	-0.025	0.011	*0.27*
	(0.023)	(0.018)	(0.018)	(0.020)	(0.026)	
regular English skills	0.137[a]	0.121[a]	0.173[a]	0.168[a]	0.205[a]	*0.3*
	(0.036)	(0.025)	(0.026)	(0.031)	(0.034)	
advanced English skills	0.324[a]	0.272[a]	0.384[a]	0.405[a]	0.422[a]	*0.05*
	(0.054)	(0.039)	(0.039)	(0.051)	(0.057)	
parental education	*no*	*no*	*no*	*no*	*no*	
occupation fixed effects	*yes*	*yes*	*yes*	*yes*	*yes*	
pseudo R2	0.381	0.406	0.416	0.407	0.382	
no English skills			*reference category*			
basic English skills	-0.011	0.013	0.011	-0.026	0.003	*0.26*
	(0.023)	(0.019)	(0.018)	(0.021)	(0.026)	
regular English skills	0.134[a]	0.107[a]	0.149[a]	0.165[a]	0.194[a]	*0.35*
	(0.037)	(0.027)	(0.026)	(0.031)	(0.037)	
advanced English skills	0.332[a]	0.271[a]	0.358[a]	0.364[a]	0.368[a]	*0.18*
	(0.051)	(0.040)	(0.040)	(0.055)	(0.062)	
parental education	*yes*	*yes*	*yes*	*yes*	*yes*	
occupation fixed effects	*yes*	*yes*	*yes*	*yes*	*yes*	
pseudo R2	0.393	0.413	0.421	0.415	0.386	
Number of Observations	6,018	6,018	6,018	6,018	6,018	

Note: Quantile Regressions with robust standard errors (Machado et al. 2011) within parenthesis in italic. [c] Significant at p<0.1, [b] significant at p<0.05, [a] significant at p<0.01. All the models include controls for years of schooling, potential experience (quadratic) type of contract and a dummy for urban area. The P-Vale from the F test for joint coefficients equality across quantiles is reported in the last column at the right of the table.

Table 4A: Self-Reported English Skills and earnings, joint LIML model

	specification 1		specification 2		specification 3		specification 4	
	ln(Y)	*SFL**	*ln(Y)*	*SFL**	*ln(Y)*	*SFL**	*ln(Y)*	*SFL**
constant	5.285[a]		5.298[a]		5.161[a]		5.100[a]	
	(0.046)		*(0.045)*		*(0.105)*		*(0.106)*	
$\widetilde{SFL}*$	0.138[a]		0.087[a]		0.076[a]		0.063[b]	
	(0.016)		*(0.029)*		*(0.028)*		*(0.028)*	
years of schooling	0.051[a]	0.222[a]	0.059[a]	0.212[a]	0.042[a]	0.214[a]	0.043[a]	0.209[a]
	(0.005)	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.009)*	*(0.007)*	*(0.009)*
potential experience	0.042[a]	-0.028[a]	0.042[a]	-0.030[a]	0.039[a]	-0.031[a]	0.040[a]	-0.027[a]
	(0.003)	*(0.009)*	*(0.003)*	*(0.009)*	*(0.003)*	*(0.010)*	*(0.003)*	*(0.010)*
(potential experience)2	-0.001[a]	0.000	-0.001[a]	0.000	-0.001[a]	0.000	-0.001[a]	0.000
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
fixed term contract	-0.297[a]	-0.008	-0.299[a]	0.002	-0.280[a]	0.125	-0.269[a]	0.138
	(0.028)	*(0.084)*	*(0.027)*	*(0.083)*	*(0.028)*	*(0.091)*	*(0.028)*	*(0.091)*
urban area	0.149[a]	0.104[b]	0.152[a]	0.099[c]	0.120[a]	0.074	0.117[a]	0.062
	(0.017)	*(0.050)*	*(0.016)*	*(0.051)*	*(0.015)*	*(0.052)*	*(0.015)*	*(0.052)*
no English or not used								
English used less than once per month		1.384[a]		0.863[a]		0.866[a]		0.842[a]
		(0.074)		*(0.088)*		*(0.088)*		*(0.088)*
English used at least once per month		1.653[a]		1.091[a]		1.120[a]		1.094[a]
		(0.121)		*(0.149)*		*(0.147)*		*(0.152)*
English used at least once per week		1.979[a]		1.356[a]		1.394[a]		1.362[a]
		(0.116)		*(0.132)*		*(0.131)*		*(0.131)*
English daily used		1.682[a]		0.740[a]		0.788[a]		0.792[a]
		(0.136)		*(0.160)*		*(0.164)*		*(0.166)*
parental education = no education								
parental education = primary							0.070[a]	0.135[b]
							(0.015)	*(0.053)*
parental education = secondary							0.086[b]	0.412[a]
							(0.034)	*(0.087)*
parental education = tertiary							0.213[a]	0.474[a]
							(0.042)	*(0.108)*
frequency of English use at work	no		yes		yes		yes	
occupation fixed effects	no		no		yes		yes	
mu1		2.500[a]		2.465[a]		2.524[a]		2.649[a]
		(0.135)		*(0.138)*		*(0.262)*		*(0.270)*
mu2		3.477[a]		3.564[a]		3.633[a]		3.767[a]
		(0.138)		*(0.142)*		*(0.264)*		*(0.273)*
mu3		4.632[a]		4.928[a]		5.006[a]		5.154[a]
		(0.146)		*(0.153)*		*(0.270)*		*(0.279)*
Corr(ε_i,u_i)	-0.184[a]		-0.131[b]		-0.125[c]		-0.104	
	(0.039)		*(0.066)*		*(0.067)*		*(0.069)*	

Note: robust standard errors within parenthesis in italic. [c] Significant at p<0.1, [b] significant at p<0.05, [a] significant at p<0.01.