Di Paolo, Antonio; Tansel, Aysit

Working Paper

English skills, labour market status and earnings of Turkish women

GLO Discussion Paper, No. 319

Provided in Cooperation with:
Global Labor Organization (GLO)

This Version is available at:
https://hdl.handle.net/10419/191947

English skills, labour market status and earnings of Turkish women*

Antonio Di Paolo, AQR-IREA, University of Barcelona

Aysit Tansel, Middle East Technical University, ERF and IZA

January 29, 2019

Abstract

In this paper, we investigate the effect of the level of English skills on the labour market outcomes of Turkish women, using data from the Adult Education Survey of 2007. By adopting a bivariate equation framework, we jointly model the effect of English skills on labour market status and, conditional on being a wage earner, on monthly earnings and occupational status. The multinomial equation that explains labour market status allows for a different effect of language knowledge on the probability of being employed, unemployed but actively looking for a job, an unpaid family worker or involved in household tasks. The results indicate that being proficient in English is conditionally associated with a higher probability of being employed as a wage earner and, to a lesser extent, unemployed but looking for a job, whereas it decreases the likelihood of being involved in household tasks. Moreover, there is a significant conditional correlation between having a high level of skills in English and earnings, which is only modestly reduced when job-related variables and (especially) occupation dummies are included as additional controls. Indeed, being proficient in English barely affects occupational status when selection into employment status is controlled for. Therefore, the knowledge of foreign languages (in this case English) seems to stimulate labour market participation and earnings capacity, but does not substantially affect the occupational position of women in the Turkish labour market.

JEL Codes: J16, J24, J31, O15, Z13

Keywords: English skills, females, labour market status, earnings, occupation.

1 Introduction

In this paper, we investigate the relationship between the knowledge of English as a foreign language and the labour market performance of women in Turkey. More specifically, we analyse whether knowing English correlated with females' labour market status and, taking into account self-selection into paid employment, affects monthly earnings. The knowledge of foreign languages is becoming more and more important in an increasingly globalized world, where economic and social interactions between agents with different linguistic repertoires are very frequent (Fidrmuc 2011). To begin with, sharing a common language but also knowing foreign languages stimulate trade and economic transactions between different countries (Frankel and Rose 2002; Melitz 2008; Egger and Lassmann 2012). In addition, the recent literature has highlighted the importance of language proficiency for economic and social outcomes at the individual level. This is especially relevant for migrants, since being fluent in the host country's language contributes to their social and economic integration and success (Chiswick and Miller 2007, 2014; Isphording 2015).

Moreover, there is a wide consensus about the fact that the knowledge of foreign languages among the natives also matters for labour market performance.[1] This is especially true for the case of English, which represents the most common foreign language spoken in the European Union as well as in other non-English-speaking countries and nowadays represents the *lingua franca* for international exchanges (European Union 2012). There are several channels through which proficiency in a foreign language, as a form of human capital, might affect labour market performance. First, fluency in a language other than the mother tongue might improve earnings potential not only because it increases communicative opportunities with co-workers and customers but also because of the potential effects of bilingualism on cognitive skills (Adeosope et al. 2010; Bialystok et al. 2009; Costa and Sebastián-Gallés 2014). Second, the knowledge of a foreign language (e.g. English in a non-English-speaking country) enhances the potential in job searches and might provide access to more prestigious and better remunerated occupations. Third, foreign language skills might be rewarded even if not directly used or requested in the workplace since they might represent a positive signal regarding other productive skills from the employer's perspective. Fourth, the observed

---

[1] Notice that there are also other related papers that investigate the effects of speaking dialects (Yao and van Ours 2018) or local languages (Rendon 2007; Di Paolo and Raymond 2012) among the natives.

language skills variable might also capture other unobserved inputs, such as the quality of education.

In line with these arguments, a growing number of papers have reported positive wage effects of the knowledge of foreign languages on natives' wages. For example, Saiz and Zoido (2005) analysed the wage return to the knowledge of foreign languages (other than English) among US college graduates. They highlighted a positive return to speaking a second language (around 2–3% using a variety of methods), but lower in size with respect to other estimates from non-English-speaking countries. There are two cross-country studies that investigated the return to different foreign languages in several European countries. Williams (2011) analysed the effect of using different foreign languages in the workplace. His findings indicate substantial returns to the use of second languages at work, which vary across countries and seem to be correlated to the extent of tourism in the country. Similarly, Ginsburgh and Prieto-Rodriguez (2011) studied the return to foreign languages used in the workplace for native male workers in nine European countries. Their evidence suggests that the wage return to the knowledge of foreign languages varies substantially across countries and depends on their relative scarcity. Moreover, they also highlighted the role of English as *lingua franca* since it is the foreign language most commonly used by firms for commerce and trade. In a subsequent paper, Ginsburgh and Prieto-Rodriguez (2013) considered gender differences in the return to foreign language. They found that the earnings effects of using English at work are generally higher for males than for females, but to varying degrees for different countries. Moreover, Stöhr (2015) investigated more deeply the return to the use of English in the workplace in Germany, reporting substantial wage returns to language proficiency that are higher for migrants than for natives and are concentrated in more language-intensive occupations (especially those related to trade).[2]

The return to the knowledge of English has been also investigated in several developing countries. The knowledge of foreign languages (English in particular) is especially relevant for the process of economic development (Seargeant and Erling 2011), since the language skills endowment of the workforce is likely to affect trading opportunities with developed countries, which in turn stimulate economic growth and development. In line with this theoretical argument, the existing research reports sizeable

---

[2] There is also specific evidence on the wage return to English skills in post-communist countries (Toomet 2011 for Estonia and Latvia; Fabo et al. 2017 for Czechia, Hungary, Poland and Slovakia), which confirms the economic value of English in the labour market.

earnings effects of English proficiency for several emerging countries. Specifically, Casale and Posel (2011) estimated wage equations for African male workers in South Africa and obtained a positive return to reading and writing English very well, which is especially pronounced for those who hold a tertiary education degree. A study by Di Paolo and Tansel (2015) focused on the case of Turkey and contains estimates of economic return to the knowledge of several foreign languages (English, French, German, Russian and Arabic) among male wage earners. They found important effects from the knowledge of English and Russian, whereas German and French appear to be relevant languages in the Turkish labour market only in specific occupations. Azam et al. (2013) investigated the same issue for the case of India and confirmed the existence of positive wage effects of proficiency in English. They also performed separate analyses for males and females, indicating that the return to English skills is higher for the former than for the latter (34% higher wages for being fluent in English among male workers, 22% for their female counterparts). Finally, Wang et al. (2017) estimated the effect of being proficient in English for male and female workers in China using a variety of methods. Their results confirmed the existence of a wage premium attached to the command of English as a foreign language, but also highlight differences with respect to the Indian case since the estimates are higher for women than for men.

As noted above, in this paper we extend the work by Di Paolo and Tansel (2015) by considering the relevance of English skills for labour market outcomes of female workers in Turkey, using data from the Adult Education Survey of 2007. In their work, the authors focused on males in order to minimize problems of endogenous self-selection into employment – possibly related to the knowledge of foreign languages as well. This is (implicitly or explicitly) the same strategy followed by the majority of existing studies on the topic, for both developed and developing countries. However, the recent contribution by Donado (2017), based on European data, indicates that skills in foreign languages matter not only for wages but also for employment status, reporting significant negative effects on the probability of being unemployed (which also vary between languages and across countries). This means that, to some extent, the knowledge of foreign languages might stimulate employment and labour market participation, but also that the estimates of wage returns to language skills that are obtained without controlling for self-selection into employment status might be biased due to selectivity and its potential relationship with language proficiency.

Therefore, with this paper we contribute to the literature by estimating a joint model for labour market status and earnings of Turkish women, in which the level of skills in English appears as an explanatory variable of both outcomes. More precisely, the labour market participation status is modelled by a multinomial probit equation.[3] The four alternatives considered are being involved in household tasks, unemployed but actively looking for a job, an unpaid family worker or employed as a wage earner. This multinomial equation is jointly estimated with the earnings equation for female wage earners, which provides estimates of the earnings effects of English skills that account for selection into employment. Moreover, the estimates of the multinomial model for labour market status are of independent interest since they provide evidence regarding the effect of this alternative form of human capital on the labour market participation and employment of women in Turkey. As better explained below, investigating the case of Turkish women is particularly interesting because of their very low levels of labour force participation according to international standards and lower earnings compared to men. Increasing women's participation in the labour market is a desirable goal from an equity perspective since it would improve their relative economic position. It is also desirable from the point of view of efficiency considerations, which in turn would foster the development potential of the country.

Self-selection into employment is, however, not the only source of bias and inconsistency in the estimation of the wage effect of English proficiency. Another is the potential issue of unobserved individual heterogeneity such as ability. We acknowledge that we are not able to address this issue directly.[4] This is because the survey data we are using is only a cross-section of individuals rather than a panel, which would enable controlling for time-invariant individual heterogeneity. In addition, we are unable to identify an exogenous source of variation in English proficiency. In this sense, the ideal setting to identify the causal effect of English skills on labour market outcomes would be a randomized experiment, in which individuals are assigned to English training programmes by lottery. Alternatively, a quasi-natural experiment such as the implementation of a reform that changes foreign language instruction during compulsory

---

[3] A relatively similar methodology has been also applied by Aldashev et al. (2009) to estimate the return to language knowledge among migrants in Germany, accounting for selection into labour market participation and occupation/sector.

[4] Therefore, the reader must bear in mind that when we describe the results, we refer to conditional correlations, and the wording referring to effects, impacts or returns should not be strictly interpreted in terms of causal relationships.

education would enable obtaining causal estimates. Unfortunately, these two empirical strategies are inapplicable, at least for the Turkish case.[5]

Another econometric problem is due to the self-reported nature of the English skills variable. It is well known that survey respondents tend to overstate their capabilities, including those for language skills. Thus, there may be possible misclassification errors in the levels of English skills we are using, which tend to introduce a bias in the estimates (Dustmann and Van Soest 2004).[6] Overall, these issues imply that we are forced to rely on conditional correlations, rather than causal effects, due to the lack of a credible identification strategy (as in all the existing papers on the return to foreign languages, to the best of our knowledge). Nevertheless, we consider that the estimates we obtain are qualitatively reliable and, to a large extent, capture the real positive impact of English proficiency on females' labour market outcomes. Indeed, the evidence emerging from the most credible identification strategies used in the parallel literature on the return to language skills among migrants indicates that there exists a positive (causal) effect of language knowledge on socio-economic outcomes (Bleakley and Chin 2004, 2010). Moreover, the two main sources of bias tend to operate in opposite directions, since the existence of omitted variables is likely to introduce an upward bias, whereas misreporting appears to bias the estimate of language proficiency downwards.

Our results indicate that being proficient in English is conditionally associated with a higher probability of being employed as a wage earner and decreases the likelihood of being involved in household tasks. Basic English skills are also associated with a somewhat higher probability of working as an unpaid family worker, whereas high English skills tend to increase the propensity to be actively looking for a job if not working (i.e. unemployed). Moreover, there is a significant conditional correlation between having a high level of English skills and earnings, which is only somewhat reduced when job-related variables and (especially) occupation dummies are included as additional controls. The earnings return to advanced English skills is also robust to the estimation of the joint model that controls for selection into salaried employment, which

---

[5] As commented in Di Paolo and Tansel (2015), the Turkish compulsory school reforms of 1997 and 2012 also increased the presence of English at school. However, our data do not cover individuals affected by these education policies. Moreover, isolating the foreign language component of the reforms from the general changes they produced in the school system would be a challenging task from the econometric point of view.

[6] As explained by the authors, the categorical nature of the variable capturing language proficiency implies that the direction of the bias due to misreporting is ex ante unpredictable. This is in contrast with the case of classical measurement errors in continuous explanatory variables, which bias the associated estimates towards zero (i.e. the attenuation bias).

also reveals a positive and significant effect of basic English skills. In order to shed light on the occupational channel as a mechanism behind the language–earnings relationship, we also estimate a model in which the final outcome is a measure of occupational status. This alternative estimation indicates that being proficient in English barely affects the occupational status of women when selection into employment status is controlled for. Therefore, the knowledge of foreign languages (in this case English) seems to stimulate labour market participation and earnings capacity, but does not substantially affect the occupational position of women in the Turkish labour market. This evidence suggests that although occupational sorting matters, this is not the main mechanism through which the knowledge of foreign languages affects the earnings of working women in Turkey. We consider that the relationship is mostly driven by the combined effect of other mechanisms such as increased productivity, improved cognitive skills and, in general, a better endowment of human capital (besides formal schooling) of women who are fluent in English.

The rest of the paper proceeds as follows: in Section 2, we present the background to the conditions of women in the Turkish labour market. Section 3 describes the data and presents some descriptive evidence, while Section 4 explains the methodology. Section 5 reports the results and Section 6 contains the conclusions.

2 Women in the Turkish Labour Market

This section provides information about the labour market participation of women in Turkey. As noted above, the labour market participation rate of women in Turkey is very low, well below the OECD average and international standards. Women's (aged 15 and over) participation rate was 32.5% in 2016 (TURKSTAT 2018), while the OECD average during the same year was almost twice as large with 63.6% (OECD 2018). The women's participation rate was 23.6% in 2007 (TURKSTAT 2018), increasing by almost 10 percentage points between 2007 and 2016. Regional disparities in terms of participation rates are notable. The highest participation rate is in Istanbul and the lowest participation rate is in Southeast Anatolia.

The unemployment rate of women was 14% in Turkey in 2016 while the OECD average was 7%. Unpaid family work is a widespread status among women. Approximately 24.1% of working women and, in contrast, only 4% of working men are

unpaid family workers (TURKSTAT 2018). The quality of jobs is poor for most women. About 43% of working women work in the informal sector without social security, in contrast to about 24% of men working in the informal sector (TURKSTAT, 2018). Among women in the age group 15–24, about 32.6% are not in education, employment or training (NEET), while the fraction men in NEET from the same age cohort is just 13.5%. Overall, this is equivalent to about 3.5 million young individuals.

Turkish women also still lag behind men in education. The average years of schooling is 7 for women as compared to 8.8 for men. The gender gap in schooling (1.8 years) is larger than it is in OECD countries (0.8 years) and the world (1.1 years) (UNDP 2016). There is also a large income gap. The per capita gross national income of women is 39.4% that of men in Turkey, while the OECD average is 59.6% and the world average is 55.5% (UNDP 2016).

Women's labour force participation increases with the level of education. Participation rates increase sharply with a vocational high school diploma and tertiary education. The highest participation rate is achieved among women who completed higher education (university or above), among which 71% are active in the labour market. Thus, raising women's educational attainments is very important for increasing their labour force participation. This is also verified by studies at the individual level (Tansel 1994; Tansel 2001; Tansel 2010). These studies provide joint estimates of women's participation in the labour force as wage earners and the selectivity-corrected estimates of wage equations for men and women in Turkey. They all find that the probability of participation of women in labour as wage earners increases significantly with the level of education, and that the highest probability of participation occurs at the university level. Further, these studies also indicate that returns to the education of women increase with the level of schooling, the highest returns occurring at the university level. The estimates of earnings equations for women are more responsive to education than those of men (Tansel 2001). In the current study, we aim to investigate how participation in wage employment is influenced by English skills for Turkish women, as well as their rewards from this alternative form of human capital. Therefore, we expect to find a positive relationship between language proficiency and earnings, as well as between language skills and labour market participation and salaried employment.

3 Data and Descriptive Statistics

We use data from the Turkish Adult Education Survey (AES) of 2007, which is a representative sample of the Turkish population aged 25 to 65 and contains several variables that are useful for our purposes. Specifically, it provides information on net monthly earnings (in Turkish Liras, TL),[7] labour market status and job-related variables, family and sociodemographic characteristics and, most importantly, self-reported information about the knowledge of and level of skills in up to seven different foreign languages. Here we focus on English, since it is the most common foreign language spoken in Turkey (as well as in other non-English-speaking countries) and has been reported to have an important economic value in the labour market (as discussed in the introduction). Individuals who claim that English is the foreign language they know best are also asked to rate their level of skills in this language. The language skills variable is recoded into three values, which represent no skills in English, regular skills in English and high skills in English.[8]

The estimation sample contains only females who were born in Turkey, who are not in education, retired or permanently disabled at the time of the survey. We also remove the few (380) cases of women who are either self-employed or employers, because of the different meaning of the earnings variables for these groups of workers, although the evidence presented below is unaffected by whether or not we include them in the estimation sample. Moreover, we exclude females who are working part-time (395 observations), but again the results are unaffected by this choice. After cleaning for the few missing values in relevant variables, we obtain a sample of 13,492 observations.

---

[7] Although the data from the 2012 and 2016 waves of the Adult Education Survey are also available for Turkey, we decided to use only the 2007 wave because the information on individual monthly earnings is not collected in the most recent version of the AES. Note also that no regional variables are provided in the AES data (there is only an urban/rural indicator referring to the place of residence).

[8] The information about self-reported skills in foreign languages is originally reported in four distinct levels: 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 2, 3 and 4 in the selected sample, we decided to group them into a single category, referring to high skills in English. This is because we wanted to avoid empty cells by labour market status, which would invalidate our empirical strategy (see next section). Indeed, the conditional association between earnings and the original levels of English knowledge is statistically equal for levels 2, 3 and 4, which justifies the grouping of the original variable (in addition to practical purposes).

Basic descriptive statistics for the variables used in the empirical analysis are displayed in Table 1, separately by labour market status. First, around 80% of individuals in the sample declare they are mostly busy with household tasks, reflecting the low participation rate of Turkish women in the labour market. Approximately 12% are employed as wage earners, and slightly less than 6% are unpaid family workers employed in the family business without receiving direct remuneration. Finally, only around 1.4% of our female sample declare being unemployed, that is, not regularly working but actively looking for a job. As expected, wage earners are more educated and younger than their counterparts who are either household or unpaid family workers. Consistently, the former are also less likely to be married and tend to have, on average, fewer children than the latter. However, the few women who are unemployed show a profile that is much more similar to wage earners rather than females in other situations. Also, regarding the level of skills in English, women who are neither regularly employed nor unemployed are less likely to be proficient in English, which may indicate that knowing English stimulates the female labour supply.

As shown in Table 1, the information on monthly earnings is available only for individuals who are regularly employed as wage earners. Female workers with regular English skills earn around 8% more than the average (approximately equal to 860 TL), whereas the earnings of females with advanced skills are 50% higher than the average. A more immediate way of visualizing (unconditional) differences in earnings by level of English skills can be achieved by estimating kernel densities for log earnings according to the knowledge of English, as displayed in Figure 1. It is evident that the distribution of log earnings is significantly shifted towards the right for women who know English, indicating the existence of a significant (unconditional) premium for the knowledge of English that operates at different parts of the earnings distribution. This is especially true for individuals with high skills in English (who also display a lower variability of earnings). The shift towards the right is also observed for women with basic skills in English, but their earnings distribution also shows a greater dispersion than for others. As for the case of labour market status, these differences in earnings might be due to the return to skills in English in the Turkish labour market, or to differences in (observable and unobservable) characteristics of the female workers. In order to control for differences in observable characteristics and obtain conditional correlations between language skills and the earnings of Turkish women, we adopt a regression framework in which we progressively include additional explanatory variables that may affect earnings

and correlate with English skills. Moreover, we also adopt a joint model that simultaneously estimates a multinomial equation that explains labour market status and, conditional on being employed, the earnings equation for the subsample of wage earners. Both equations include English skills as an explanatory variable. Therefore, the joint model enables appreciating the conditional correlation between language skills and both labour status and earnings, but controlling for potential self-selection into regular employment as wage earner (and the indirect impact of English on the probability of being a wage earner or in other situations).

4 Empirical Methodology

The starting point of our empirical analysis is the estimation of an OLS regression that explains the log of monthly earnings (Y_i) as a function of dummies for English skills and a set of control variables, that is:

$$\ln(Y_i) = \alpha + \beta' X_i + \delta_1 I(ES_i = 1) + \delta_2 I(ES_i = 2) + \varepsilon_i \qquad (1)$$

Here ES_i represents the level of skills in English (basic = 1, advanced = 2) and the δ coefficients capture the conditional difference in earnings relative to the base category (having no skills in English, $ES_i = 0$). The set of control variables X_i initially contains only years of schooling and quadratic potential experience. Then we proceed with a stepwise inclusion of additional control variables that might affect women's earnings, ranging from household characteristics (residence in urban areas, marital status, the number of children and the number of adults in the household) to job-related variables, namely having a fixed-term contract and occupation dummies for two-digit International Standard Classification of Occupations (ISCO) classification codes. In this way, we will be able to gauge whether the conditional correlation between the knowledge of English and earnings is, to some extent, mediated by the effect of these control variables.[9]

However, as long as language skills may also affect labour market status and employed women are not a random sample of females due to self-selection into

---

[9] Notice that job-related characteristics such as the type of contract or occupation dummies could be considered to be "bad controls", in the terminology of Angrist and Pischke (2009), since these variables might themselves be determined by the level of English skills. However, their inclusion still provides suggestive evidence about the role played by job characteristics as mediating factors in a setting of stepwise inclusion of additional controls, as commonly performed in the literature (e.g. Azam et al. 2013; Di Paolo and Tansel 2015).

employment, we adopt a joint estimation technique that enables modelling the probability of being in a given labour market status (L) as a function of a set of observable characteristics and English skills, together with the earnings equation. Specifically, we consider that the probability of being involved in household tasks ($L_i = 1$), unemployed but actively looking for a job ($L_i = 2$), an unpaid family worker ($L_i = 3$) or a wage earner ($L_i = 4$) can be modelled using a multinomial probit model. In this case, the utility of choosing alternative "j" ($j = 1 \dots 4$) depends on a set of observable characteristics (Z_i), with the corresponding alternative-specific coefficients (γ_j) and error terms (u_{ij}):

$$U_{ij} = \mu_j + \gamma_j' Z_i + u_{ij} \qquad (2)$$

Therefore, the probability of choosing alternative "j" will be equal to the probability that the utility obtained from choosing this alternative is greater than the utility derived from other alternatives, that is:

$$p_{ij} = \Pr(L_i = j | Z_i) = P(U_{ij} > U_{il}, \forall\, l \neq j\, | Z_i) \qquad (3)$$

We consider the error terms u_{ij} to be joint normally distributed but independent across alternatives (together with the normalization $u_{i1} = 0$ that is imposed to achieve identification), which corresponds to the multinomial probit specification.[10] However, we also allow the error terms u_{ij} ($j \neq 1$) to be correlated with ε_i, namely the error term of the earnings equation, in order to control for self-selection into labour market status. Notice that, in order to not rely only on identification due to non-linearities in the functional form, we include some exclusion restrictions in the vector of explanatory variables of the multinomial selection equation (Z_i), in addition to the control variables used in the earnings equation.

In order to find suitable exclusion restrictions for the multinomial selection equation for labour market status, we refer to the specific literature regarding the estimation of sample selection model applied to the issue of employment selectivity. Specifically, we exploit information contained in the AES dataset, which enables computing the share of children aged 0 to 3 years and 4 to 5 (similar to Martins 2001, Chang 2011, and Huber and Mellace 2014) over the total number of children aged 0–18 in the household, as well as the share of adults above 64 in the family. The underlying hypothesis is that conditional on the total number of children and adults (which are included as controls in the wage equation), as well as other observable characteristics, these variables affect the labour

---

[10] Therefore, the reader should bear in mind that we are assuming the validity of the Independence of Irrelevant Alternatives (IIA) hypothesis for the multinomial selection equation.

supply of women due to their potential impact on caregiving activities, but do not directly correlate with earnings received as salaried workers. We also experimented with other typical variables used as exclusion restrictions in similar settings, such as partner's education or wages, which provided similar results. However, we preferred to rely on family structure as exclusion restrictions, since the validity of a partner's characteristics has been strongly questioned in the recent literature (for a detailed discussion, see Huber and Mellace 2014).

5 Results

5.1 OLS Regressions

The results from the OLS estimations of Equation (1) are displayed in Table 2. In the first column, the vector of controls X_i contains a very parsimonious set of explanatory variables, namely years of schooling, years of (potential) labour market experience and its square. Keeping these variables fixed, the coefficient attached to having a basic knowledge of English is not statistically significant due to the low estimation precision, whereas the advanced command of this language is associated with 27% ($= 100[\exp(\hat{\delta}_2) - 1]$) higher monthly earnings. This is in line with the estimates obtained for Turkish males by Di Paolo and Tansel (2015), as well as from other studies conducted in either developed or developing countries. Considering that each additional year of schooling increases earnings by 10%, having an advanced knowledge of English generates approximately the same effect on earnings as 2.7 years of education, which is a substantial effect.[11] In the second column, we add some household characteristics, that is, dummies for living in urban areas and for marital status, plus the number of adults and children in the household, since these sociodemographic variables are likely to be conditioning factors of earnings (especially among women). We also include dummies for the highest level of education between the parents, which capture social network effects and other unobservable determinants of earnings related to parental investment.

---

[11] One may wonder about whether an individual could have learned English in the workplace, which would generate an issue of reverse causality. However, 83% of the selected sample who claim to have some skills in English state that they learned the language at school, and restricting the estimation sample to these cases does not alter the results presented in this paper (the complete results are available upon request).

Indeed, we can observe that, as expected, women with at least one parent with tertiary education obtain higher earnings. Married women earn slightly more than others (which is in line with the findings usually obtained for males), whereas having more children or living in larger households are not associated with earnings. Moreover, individuals residing in urban areas earn more than their counterparts living in rural areas. However, the estimated coefficients of the dummies of English skills are virtually unaffected by the inclusion of household controls.

In column (3), we also include two job-related control variables, which correspond to an indicator for having a fixed term and dummies for occupational classifications (two-digit ISCO 88) since, as common in these studies, occupation seems to be the main mechanism behind the relationship between the knowledge of foreign languages and wages. Indeed, we also observe for Turkish women a certain reduction in the coefficient attached to the dummy for advanced English skills from 0.22 to 0.194, corresponding to a 12% decrease. This means that, to some (although moderate) extent, female workers who are proficient in English tend to earn more also because they attain better remunerated occupations, as in the case of Turkish males and given other realities that have been previously investigated.[12] Nevertheless, the estimated return to high English skills remains sizable and significant, even conditioning on occupation. Therefore, with the aim of better understanding the role of occupation as a mediating factor for the relationship between language knowledge and earnings, in column (6) we consider as dependent variable the so-called International Socio-Economic Index (ISEI) of occupational status proposed by Ganzeboom (1992), which is a measure that is computed from ISCO occupational codes at two digits and reflects occupational quality (or the socio-economic status attached to each occupation). This additional estimation reveals that, consistent with the evidence from the earnings equation, basic skills in English are significantly associated with higher earnings, whereas advanced skills generate a substantial increase in occupational status, corresponding to around 62% of the effect of increasing years of schooling by one year.

Overall, it appears that the advanced knowledge of English improves the remuneration of Turkish women who are employed as wage earners, partially (but not

---

[12] As already noted, these results can be just taken as suggestive evidence of the role played by occupational sorting, due to the fact that is difficult to disentangle the true effect of language skills from potential unobservables affecting occupational choices. Indeed, as suggested by an anonymous referee, the modest reduction in the coefficient for English proficiency might be because better paid occupations could attract only the most able individuals with no knowledge of foreign languages.

exclusively) because this additional form of human capital favours their access to better paid jobs. However, it is also possible that language proficiency not only affects earnings but also stimulates labour supply, making women more likely to be active in the labour market (because they are either regularly employed or actively looking for a salaried job). Indeed, it also seems interesting to analyse the effect of English skills on the labour market status of women because the joint model for labour market status and earnings enables dealing with potential issues of self-selection into regular employment as wage earners, which could bias the estimates presented so far.

5.2 Multinomial Selection Equation

As explained in Section 4, in order to control for selection into paid employment, we jointly estimate a multinomial selection equation (with normally distributed error terms) and the outcome equation, allowing the error terms of the latter equation to be correlated with the errors of the latent equation for the probability of being in each labour market status.[13] The estimates from the multinomial selection equation are of independent interest, especially because this model enables investigating the extent to which English proficiency represents a conditioning factor of the labour market status of Turkish women. The (average) marginal effects on the probability of being in each possible status (i.e. household work, unpaid family worker, unemployed and employed as wage earner) from the multinomial selection equation are reported in Table 3 (the coefficients' estimates are not reported but are available upon request).

We start by examining the results regarding the coefficients of control variables. It appears that schooling makes women less likely to be (exclusively) involved in household tasks, and more likely to be employed as salaried workers. However, parental education increases the likelihood of carrying out household tasks and reduces the probability of being either unemployed or an unpaid family worker, possibly reflecting a family wealth effect that prevents women from actively participating in the labour market. Residing in urban areas increases the chances of being a household worker and, to a significantly lower extent, a wage earner, but decreases the likelihood of being an unpaid family worker (which is a more common condition in rural areas). Marital status is an important predictor of labour market status. Indeed, relative to single women who were never

---

[13] The estimations are carried out with Stata, using the user-written routine "cmp" (Roodman 2018).

married, their married counterparts are more likely to be household workers and less likely to be unemployed or wage earners. A similar pattern is observed for widowed individuals, with the additional negative effect on the probability of working in a family business without receiving a monetary remuneration. Similarly, divorced females are less likely to be unpaid family workers than married or never married women. An increase in the total number of children is associated with a higher probability of being an unpaid family worker, but decreases the probability of being involved in household tasks or being a wage earner. One tentative explanation for this finding is that, on the one hand, women in families with more children might prefer collaborating in the family business through unpaid work than being a salaried worker because it is more compatible with childbearing. However, on the other hand, they are more likely to provide unpaid work rather than being house workers due to more demanding economic needs in the family (conditional on other characteristics).

On the contrary, living in a household with more adults makes females more likely to be devoted to household work than in other statuses (especially a salaried worker). Conditional on the number of children, an increase in the share of children aged 0–3 and 5–6 raises the likelihood of performing household work, and decreases the likelihood of working in a remunerated job. A similar pattern is observed for the share of adults aged 65 or more in the household, with the additional positive effect on the probability of being an unpaid family worker and the negative impact (but modest in size) on the probability of being unemployed but actively looking for a job.

Most importantly, the estimates from the multinomial selection equation also reveal the conditional correlation between English skills and labour market status. Specifically, from Table 3 it can be seen that the increase in language skills is negatively associated with the probability of being involved in housework and positively with the probability of being a salaried worker. Moreover, having advanced skills in English also slightly increases the likelihood of being unemployed, and having basic skills is positively related to the chances of being an unpaid family worker.

5.3 Selectivity-Corrected Estimations

The coefficients of the earnings equation that are jointly estimated with the selection equation to correct for self-selection into salaried employment are displayed in Table 4. As long as we exploit the share of children and adults of given age groups as exclusion

restrictions (while controlling for the total number of adults and children in the earnings equation), we consider the starting model to be the one that includes the vector of family characteristics (as in the OLS estimations reported in column (2) of Table 2).

After controlling for selection into paid employment,[14] the coefficient associated with basic English skills increases in size and becomes significant, whereas the impact of having advanced skills is just slightly higher in the selectivity-corrected estimation (from 0.22 to 0.267) and remains strongly significant. After controlling for job-related controls (fixed-term contract and occupation dummies), again we observe a certain reduction in the estimates of English skills (both basic and high), which indicates that having basic skills in English raises earnings by 5.6% $(= 100[\exp(\hat{\delta}_1) - 1])$ and advanced skills by 28%. However, these conditional correlations are high enough to say that the relationship between language proficiency and income is not mostly driven by occupational selection. In fact, when we change the dependent variable and use the ISEI occupational score, but still control for selection into salaried employment through our joint model, the impact of English skills becomes much lower than when using simple OLS, and is no longer statistically different from zero due to the relatively low precision with which we estimate the coefficient of advanced language skills.

It is also worth pointing out that the estimates of the correlation coefficients between the error term of the earnings equation and those from the latent equations explain the probability of being in each possible employment status. Controlling for selection into employment status matters,[15] since we obtain a statistically significant correlation coefficient between unobserved earnings potential and the error term of the equation that explain the chances of being unemployed. The positive and significant correlation indicated that unobserved factors that drive the decisions of remaining unemployed are positively correlated with the error term of the earnings equation, possibly capturing the higher reservation wage of women with higher wage potential. However, no statistically significant correlation coefficient for self-selection is found when the dependent variable is occupational status rather than earnings, also indicating that the OLS estimates for this

---

[14] The estimates of control variables are virtually the same as those obtained by OLS, at least in qualitative terms, and are not discussed for brevity reasons. The only exception is the coefficient of the dummy for being married that is no longer significant when estimated with the joint model.

[15] We tested for the joint significance of the correlation coefficients between the error term of the earnings equation and the latent equations, capturing the difference in expected utility between being in each status and the base category (being involved in household work) through a likelihood ratio test. The corresponding null hypothesis is soundly rejected, which indicates that overall controlling for selection into working status is relevant.

outcome appear to be consistent. Overall, we can conclude that the knowledge of English stimulates the labour supply of Turkish women and that being proficient in English is also positively related to earnings, even accounting for a relatively large set of observable characteristics and controlling for self-selection into employment status.

6 Conclusions

We analysed the effects of skills in English as a foreign language on the labour market performance of women in Turkey. Existing evidence for Turkey reported high earnings returns to the knowledge of foreign languages, especially to English skills, but was limited to male wage earners. The results from other developed and developing countries point to different effects of language fluency by gender, although the size and the sign of this gap also vary by country. The estimates reported in this paper are in line with the results obtained from other studies from developing countries (i.e. India and China) in terms of the magnitude of the effects (i.e. around 25–30% higher earnings for the advanced command of English, depending on the specification) and are quite similar to those obtained for Turkish male workers reported in the literature. The conditional correlation between a high level of fluency in English and earnings is stable to the inclusion of family and job characteristics and is just slightly reduced when occupation fixed effects are added to the regression. These results are unchanged when using a joint model for earnings and labour market outcomes that account for self-selection of women into paid employment based on observable and unobservable characteristics, which also highlight a modest positive effect of basic English skills. Moreover, the estimates from the multinomial selection equation indicate that the knowledge of English also stimulates labour market participation (as does formal schooling), since basic and especially advanced skills English proficiency increase the likelihood of being a wage earner and, to a lesser extent, of being unemployed, but rises the probability of being a household worker. Finally, we also use a measure of occupational status as outcome, with the aim of better understanding the role of occupation as mediator of the language–earnings relationship. The results from the joint model of earnings and labour market status show that English skills are barely associated with occupational status when selection into paid employment is controlled for.

Therefore, occupational sorting and signalling effects are unlikely to be the main driver of the conditional relationship between English fluency and labour market outcomes. Indeed, other mechanisms related to productivity in the workplace, such as the cognitive effects of multilingualism, and the better endowment of human capital (besides formal schooling) possibly represent the channels through which the knowledge of foreign languages affects the earnings of working women in Turkey. Unfortunately, the available data do not allow disentangling the different factors behind the association between earnings (and labour market status) and the knowledge of foreign languages, and this represents a challenge for future research. Indeed, the availability of data including detailed information on cognitive and non-cognitive skills, as well as their use in the workplace, would be helpful to gauge the extent to which these cognitive and human capital gradients of the relationship between English skills and labour market outcomes.[16]

Overall, our results indicate that the knowledge of English as foreign language, as an alternative form of human capital, is beneficial for the labour market performance of women in Turkey, a country characterized by important gender disparities in the labour market and low labour market participation of women (especially among the low educated). Being proficient in English might help women to join the labour market and to catch up with their male counterparts, which would enhance women's empowerment in the labour market. These are important objectives to pursue, especially in the light of Turkey's aim to enter the European Union.

Although we were unable to obtain causal estimates due to data limitations, the findings reported in this paper have important policy implications about English (and foreign language) education in schools, including universities and language centres. The results are of interest to job seekers, employment agencies and governmental organizations such as the Ministry of Education and Council of Higher Education. Indeed, the Turkish Government and educational institutions should increase exposure to foreign languages (especially English) throughout school, which would be an effective measure for improving women's performance in the labour market. This is also due to potential complementarities between education and language skills, in addition to the fact that schooling itself represents one of the main instruments to foster employability and earnings potential. This need was recently acknowledged by the Turkish educational authorities, which implemented several reforms of the schooling system that also

---

[16] Alternatively, the implementation of a laboratory experiment could be a sensitive route to follow in order to understand the effect of cognitive skills induced by multilingualism on an individual's productivity.

involved an increase in English instruction at school. Specifically, the 1997 educational reform increased both compulsory schooling from five to eight years and the number of hours of foreign language instruction. The subsequent educational reform of 2012 increased compulsory schooling even more, from eight to 12 years, and lowered the grade level at which foreign language instruction starts to second grade. Indeed, evaluating the effect of the language component of these school reforms represents the objective for our future research.

References

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 (2): 207–245.

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.

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press. Princeton, New Jersey (USA).

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

Bialystok, E., Craik, F.I., Green, D.W., & Gollan, T.H. (2009). Bilingual minds. Psychological Science in the Public Interest, 10(3): 89–129.

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

Bleakley, H., & Chin, A. (2010). Age at arrival, English proficiency, and social assimilation among US immigrants. American Economic Journal: Applied Economics, 2(1): 165-92.

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.

Chang, S. K. (2011). Simulation estimation of two- tiered dynamic panel Tobit models with an application to the labor supply of married women. Journal of Applied Econometrics, 26(5): 854-871.

Chiswick, B. R., & Miller, P. W. (2007). The economics of language: International analyses. Routledge.

Chiswick, B. R., & Miller, P. W. (2014). International migration and the economics of language. Handbook of the Economics of Immigration, pp. 211-269.

Costa, A. & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? Nature Reviews Neuroscience, 15(5): 336-345.

Di Paolo, A., & Raymond, J. L., (2012). Language knowledge and earnings in Catalonia. Journal of Applied Economics, 15(1): 89-118.

Di Paolo, A. and A. Tansel (2015). Returns to Foreign Language Skills in a developing Country: The Case of Turkey. Journal of Development Studies, 51(4): 407-421.

Donado, A. (2017). Foreign Languages and their Impact on Unemployment. LABOUR, 31(3): 265-287.

Dustmann, C., & Van Soest, A. (2004). An analysis of speaking fluency of immigrants using ordered response models with classification errors. Journal of Business & Economic Statistics, 22(3): 312-321.

Egger, P. H., & Lassmann, A. (2012). The language effect in international trade: A meta-analysis. Economics Letters, 116(2): 221-224.

European Union, (2012). Europeans and their languages, Special Euro-Barometer 243/Wave 64.3 (http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf).

Fabo, B., Beblavý, M., & Lenaerts, K. (2017). The importance of foreign language skills in the labour markets of Central and Eastern Europe: assessment based on data from online job portals. Empirica, 44(3): 487-508.

Fidrmuc, J., (2011). The economics of multilingualism in the EU, in: Eger, T., Schäfer H-B. (Eds.), Research Handbook on the Economics of European Union Law. Edward Elgar, Cheltenham, pp. 331-352.

Frankel, J., and A. Rose (2002). An estimate of the effect of currencies on trade and income. Quarterly Journal of Economics 117: 437-466.

Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. Social Science Research. 21(1): 1-56.

Ginsburgh, V. A., & Prieto-Rodriguez, J., (2011). Returns to foreign languages of native workers in the EU. Industrial and Labor Relations Review, 64(3): 599-617.

Ginsburgh, V., & Prieto- Rodriguez, J. (2013). Is there a gender bias in the use of foreign languages in Europe? Kyklos, 66(4): 552-566.

Huber, M., & Mellace, G. (2014). Testing exclusion restrictions and additive separability in sample selection models. Empirical Economics, 47(1): 75-92.

Isphording, I. (2015). Language and labor market success. International Encyclopedia of Social and Behavioral Sciences, 2nd Edition, 2015, pp. 260-265.

Martins, M. F. O. (2001). Parametric and semiparametric estimation of sample selection models: an empirical application to the female labour force in Portugal. Journal of Applied Econometrics, 16(1): 23-39.

Melitz, J. (2008), Language and Foreign Trade. European Economic Review, 52(4): 667-699.

OECD (2018), OECD Labour Force Statistics 2017, OECD Publishing, Paris. http://dx.doi.org/10.1787/oecd_lfs-2017-en
http://stats.oecd.org/Index.aspx?DataSetCode=STLABOUR#

Rendon, S., (2007). The Catalan premium: Language and employment in Catalonia. Journal of Population Economics, 20(3): 669-686.

Roodman, D. (2018). CMP: Stata module to implement conditional (recursive) mixed process estimator. Available at:
https://econpapers.repec.org/software/bocbocode/s456882.htm

Saiz, A., & Zoido, E. (2005). Listening to what the world says: Bilingualism and earnings in the United States. Review of Economics and Statistics, 87(3): 523-538.

Seargeant, P. and E. Erling (2011) "The Discourse of English as a Language for International Development: Policy Assumptions and Practical Challenge," in Coleman, H. (ed.) *Dreams and Realities: Developing Countries and the English Language.* British Council, London, (pp. 255-274).

Stöhr, T. (2015). The returns to occupational foreign language use: Evidence from Germany. Labour Economics, 32: 86-98.

Tansel, A. (1994). Wage Employment, Earnings and Returns to Schooling for Men and Women in Turkey. Economics of Education Review, 13 (4): 305-320.

Tansel, A. (2001). "Self-Employment, Wage-Employment, and Returns to Schooling by Gender in Turkey," in Labor and Human Capital in the Middle East: Studies of Markets and Household Behavior, ed. by Djavad Salehi-Isfahani, Ithaca Press, 2001, Reading, UK. pp.637-667.

Tansel, A. (2010). Changing Returns to Education for Men and Women in a Developing Country: Turkey, 1994 and 2005. Paper presented at the ESPE conference, June 18-21, 2008, in London, ECOMOD conference, July 2-4, 2008 in Berlin, MEEA conference, March 20-23, 2009 in Nice, France and ICE-TEA conference, September 1-3, 2010 in Girne, Republic of Northern Cyprus.

Toomet, O. (2011). Learn English, Not the local language! Ethnic Russians in the Baltic States. American Economic Review: Papers and Proceedings, 101(3): 526-531.

TURKSTAT (2018) Turkish Labor Force Statistics, www.tuik.gov.tr

UNDP (United Nations Development Program) (2016) Human Development Report.

Wang, H., Smyth, R., & Cheng, Z. (2017). The economic returns to proficiency in English in China. China Economic Review, 43: 91-104.

Williams, D., (2011). The economic returns to multiple language usage in Western Europe. International Journal of Manpower, 32(4): 372-393.

Yao, Y., & van Ours, J. C. (2018). Daily dialect-speaking and wages among native Dutch speakers. Empirica (*on-line first*, DOI 0.1007/s10663-018-9420-z).

Tables and Figures

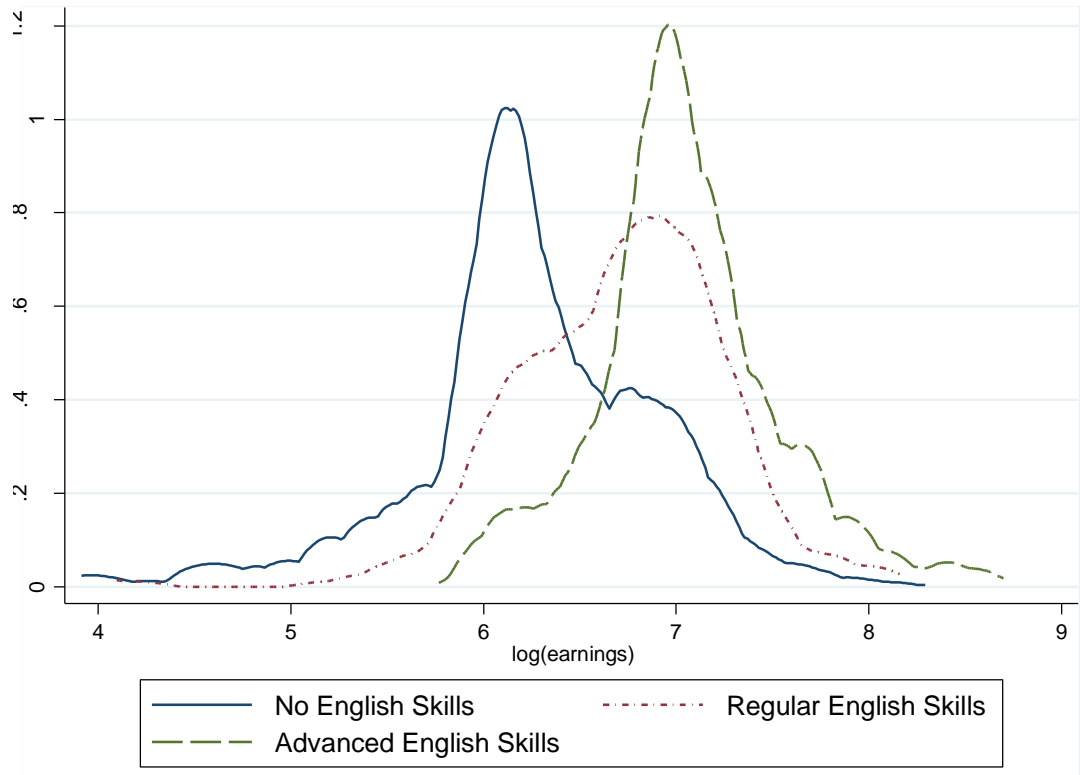Figure 1: Kernel Density of logged earnings by skills in English

Table 1: Descriptive Statistics by Labor Market Status

variable	wage earner		household work		unemployed		unpaid family worker	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
monthly earnings (TL)	860.9	614.0						
occupational status (ISEI)	48.04	17.94						
no skills in English	0.53	0.50	0.93	0.26	0.52	0.50	0.96	0.19
basic skills in English	0.21	0.41	0.05	0.23	0.20	0.40	0.03	0.17
advanced skills in English	0.27	0.44	0.02	0.14	0.28	0.45	0.01	0.09
years of schooling	10.29	4.47	4.86	3.32	9.98	4.46	3.89	2.93
potential experience (age – years of schooling – 6)	18.94	9.89	30.51	12.34	16.17	9.96	33.57	12.20
both parents with no education	0.15	0.36	0.43	0.50	0.20	0.40	0.57	0.50
parents with primary education	0.63	0.48	0.53	0.50	0.58	0.50	0.42	0.49
parents with secondary education	0.12	0.32	0.03	0.16	0.16	0.36	0.00	0.06
parents with tertiary education	0.11	0.31	0.01	0.11	0.07	0.26	0.01	0.07
urban area	0.86	0.34	0.72	0.45	0.84	0.36	0.13	0.34
never married	0.23	0.42	0.04	0.20	0.51	0.50	0.06	0.24
married	0.68	0.47	0.89	0.32	0.35	0.48	0.92	0.27
divorced	0.07	0.25	0.02	0.13	0.13	0.34	0.01	0.08
widowed	0.03	0.16	0.06	0.23	0.02	0.13	0.01	0.11
number of adults	2.51	1.07	2.67	1.14	3.08	1.33	3.11	1.36
number of children	0.69	0.90	1.12	1.33	0.41	0.68	1.05	1.39
% children 0-3	0.09	0.27	0.11	0.27	0.06	0.22	0.10	0.26
% children 4-5	0.07	0.23	0.08	0.22	0.03	0.15	0.06	0.17
% adults > 65	0.03	0.11	0.05	0.13	0.04	0.11	0.10	0.17
fixed-term contract	0.10	0.29						
number of observations	1674		10845		187		785	

Note: the variable occupational status refers to the International Socio-Economic Index of occupational status (ISEI) proposed by Ganzeboom et al. (1992).

Table 2: OLS Regressions

	(1)	(2)	(3)	(4)
dependent variable:	(log) earnings	(log) earnings	(log) earnings	occupational status
constant	5.065***	4.929***	5.834***	10.973***
	(0.072)	(0.092)	(0.110)	(2.122)
no skills in English	*reference category*			
basic skills	0.054	0.052	0.030	-0.452
	(0.035)	(0.034)	(0.030)	(0.880)
advanced skills	0.240***	0.220***	0.194***	1.885**
	(0.035)	(0.035)	(0.033)	(0.939)
years of schooling	0.101***	0.094***	0.052***	3.033***
	(0.004)	(0.004)	(0.005)	(0.104)
potential experience	0.032***	0.027***	0.023***	0.035
	(0.005)	(0.005)	(0.005)	(0.111)
potential experience^2	-0.001***	-0.000***	-0.000***	0.002
	(0.000)	(0.000)	(0.000)	(0.002)
both parents with no education	*reference category*			
parents with primary education		0.020	-0.020	1.673*
		(0.039)	(0.036)	(0.870)
parents with secondary education		0.056	-0.011	3.483***
		(0.052)	(0.049)	(1.246)
parents with tertiary education		0.118**	0.041	4.523***
		(0.053)	(0.051)	(1.262)
urban area		0.271***	0.149***	2.721***
		(0.039)	(0.036)	(0.836)
never married	*reference category*			
married		0.079**	0.081***	0.693
		(0.033)	(0.030)	(0.808)
divorced		0.009	0.019	-1.713
		(0.063)	(0.059)	(1.386)
widowed		0.024	0.011	0.131
		(0.084)	(0.078)	(1.775)
number of adults		-0.017	-0.017	-0.260
		(0.011)	(0.011)	(0.287)
number of children		-0.008	0.015	-0.089
		(0.015)	(0.014)	(0.325)
fixed-term contract			-0.438***	
			(0.066)	
occupation FEs	*no*	*no*	*yes*	*no*
adjusted R-squared	0.475	0.497	0.599	0.620
number of observations	1644	1644	1644	1644

***, **, * denote significance at the 1, 5 and 10 percent level. Robust standard errors in parentheses.

Table 3: multinomial probit model for labor market status (average marginal effects)

	(1) ΔPr(household work)	(2) ΔPr(unemployed)	(3) ΔPr(unpaid family work)	(4) ΔPr(wage earner)
no skills in English		*reference category*		
basic skills	-0.060***	0.003	0.027**	0.029***
	(0.016)	*(0.003)*	*(0.014)*	*(0.009)*
advanced skills	-0.161***	0.014**	0.037	0.109***
	(0.026)	*(0.006)*	*(0.025)*	*(0.016)*
years of schooling	-0.016***	0.000	-0.000	0.016***
	(0.001)	*(0.000)*	*(0.001)*	*(0.001)*
potential experience	0.003***	-0.001***	0.000	-0.002***
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
both parents with no education		*reference category*		
parents with primary educ.	0.015**	-0.006**	-0.009**	0.000
	(0.007)	*(0.003)*	*(0.004)*	*(0.006)*
parents with secondary educ.	0.047***	-0.003	-0.049***	0.005
	(0.014)	*(0.005)*	*(0.008)*	*(0.011)*
parents with tertiary educ.	0.002	-0.012***	-0.007	0.017
	(0.025)	*(0.004)*	*(0.022)*	*(0.015)*
urban area	0.102***	0.000	-0.117***	0.016***
	(0.007)	*(0.002)*	*(0.004)*	*(0.006)*
never married		*reference category*		
married	0.122***	-0.035***	0.009	-0.096***
	(0.016)	*(0.007)*	*(0.008)*	*(0.013)*
divorced	-0.029	0.022	-0.031**	0.037
	(0.028)	*(0.014)*	*(0.012)*	*(0.024)*
widowed	0.120***	-0.030***	-0.042***	-0.047**
	(0.024)	*(0.010)*	*(0.009)*	*(0.022)*
number of children	-0.005*	0.001	0.010***	-0.005**
	(0.003)	*(0.001)*	*(0.001)*	*(0.003)*
number of adults	0.019***	-0.003**	-0.005***	-0.011***
	(0.004)	*(0.001)*	*(0.002)*	*(0.003)*
% children 0-3	0.058***	-0.005	-0.001	-0.053***
	(0.013)	*(0.004)*	*(0.008)*	*(0.011)*
% children 4-5	0.049***	-0.009	-0.017*	-0.023**
	(0.015)	*(0.006)*	*(0.010)*	*(0.011)*
% adults > 65	0.044*	-0.020**	0.046***	-0.069***
	(0.025)	*(0.008)*	*(0.012)*	*(0.022)*
number of observations	13491	13491	13491	13491

***, **, * denote significance at the 1, 5 and 10 percent level. Robust standard errors in parentheses.

Table 4: selectivity-corrected outcome(s) equations

dependent variable:	(1) (log) earnings	(2) (log) earnings	(3) occupational status
constant	4.914***	5.659***	18.556***
	(0.133)	*(0.158)*	*(4.986)*
no skills in English	*reference category*		
basic skills in English	0.070**	0.055*	-0.798
	(0.035)	*(0.033)*	*(0.968)*
advanced skills in English	0.267***	0.251***	1.085
	(0.043)	*(0.044)*	*(1.292)*
years of schooling	0.099***	0.062***	2.749***
	(0.006)	*(0.007)*	*(0.229)*
potential experience	0.025***	0.022***	-0.090
	(0.005)	*(0.005)*	*(0.112)*
potential experience^2	-0.000***	-0.000***	0.005**
	(0.000)	*(0.000)*	*(0.003)*
both parents with no education	*reference category*		
parents with primary education	0.010	-0.027	1.604*
	(0.038)	*(0.036)*	*(0.856)*
parents with secondary education	0.055	-0.011	3.409***
	(0.052)	*(0.049)*	*(1.231)*
parents with tertiary education	0.093*	0.026	4.219***
	(0.053)	*(0.051)*	*(1.265)*
urban area	0.259***	0.162***	1.623*
	(0.040)	*(0.035)*	*(0.855)*
never married	*reference category*		
married	0.004	0.003	1.441
	(0.039)	*(0.040)*	*(1.099)*
divorced	0.030	0.048	-2.151
	(0.064)	*(0.060)*	*(1.471)*
widowed	-0.040	-0.034	-0.489
	(0.084)	*(0.079)*	*(1.834)*
number of adults	-0.015	-0.018*	-0.074
	(0.011)	*(0.011)*	*(0.289)*
number of children	-0.021	-0.003	0.215
	(0.017)	*(0.017)*	*(0.411)*
fixed-term contract		-0.433***	
		(0.065)	
occupation FEs	*no*	*yes*	*no*
$\sigma_{\varepsilon u2}$	0.778***	0.726***	0.050
	(0.117)	*(0.156)*	*(0.146)*
$\sigma_{\varepsilon u3}$	0.167	0.111	0.225
	(0.102)	*(0.101)*	*(0.167)*
$\sigma_{\varepsilon u4}$	0.173	0.314*	-0.249
	(0.126)	*(0.161)*	*(0.222)*
number of observations	13491	13491	13491

***, **, * denote significance at the 1, 5 and 10 percent level. Robust standard errors in parentheses.