Economics and Business Administration

Department of Economics

Applied MicroEconometrics

Proffesor: Giannakopoulos Nicholas

Project 8

Peter Pan

Employment Status and Working hours per week and factors affecting them

Panou Panagiotis

This article analyzes the factors influencing the employment status and working hours per week of the sampled individuals. For our analysis, three regression models were used which are the Multinomial logit model, Ordered probit model and the Ordered logit model. As for the data, it is worth noting that the country of interest is Greece and for the study of these factors data from the European Social Survey Round 10 was used. The above research was carried out in Greece between 21/11/2021 and 19/4/2022. The empirical findings are revealing that the work status is mainly influenced by the gender of the person, with men being the gender where it is favored, and by the access the person can have to the binary system. In other words, the person's familiarity with the internet, while the weekly hours of employment seem to be positively influenced by both the aforementioned factor and whether the person holds a university degree or a higher level of education.

Introduction

The purpose of this article is to study the factors that affect the professional status of the individual as well as the hours of his weekly employment. The current employment situation includes many categories of both workers and people who are not working for various reasons. For example some of these are when the person is actively working as an employee, is the employer, is associated with the employer in a business relationship (eg as a supplier or contractor dealing with the employer) is not working at all but is actively looking for work (unemployed) or even if the person is not even looking for work and is classified as inactive. From the aforementioned categories, the ones we will deal with in this article are the employed, the unemployed and the inactive. In particular, we will see how the individual's employment status is affected by the factors of his age, his gender, his marital status, whether he is native or not, his level of education, the use of the internet that the individual makes, but and the place from which he accesses the internet his political beliefs as well as how often he attends religious services. The exact same variables will be used to see how the individual's weekly working hours are affected.

Data

The data for our research came from the European Social Survey Round 10 database. Initially, as we mentioned in the introduction, the country of interest in this article is Greece and the above research was carried out in Greece between 21/11/2021 and 19/4/ 2022.

In the first stage, the variables of interest that will be used in our analysis from the sample are the following: agea (the variable wich shows the age of respondent), gndr (individual’s gender), brncntr (if

the respondent has born in Greece), rshpsts(the relationship with husband,wife or partner currently living with together), edulvlb (the highest level of education that the respondent has), mnactic (the main activity at last 7 days that the respondent had), wkhct (the total contracted hours per week in main job overtime excluded), netusoft (how often the individual make use of internet), chldo12 (the number of children aged 12 or over), acchome (if home is the location where the respondent is able to access to internet), accwrk (if work is the location where the respondent is able to access to internet), nuts2 (the individual's region), rlgatnd (this variable shows us how often the respondent attends religious services except on special occasions), lrscale (this variable shows us the political beliefs of the individual starting from the left and ending after a scale from 1 to 9 in the right position).

We create a few categorical variables which are: empstatus (which shows the professional status of the person: 1 if the person is employed, 3,4 if the person in sample is unemployed, and 2,5,6,7,8,9 if the person's professional status is inactive), hours (this variable take the price 0 if the person don't work at all-have zero working hours during the week, price 1 if the person work from 1 to 29 hours during the week, and the price 2 if the person work from 30 to 1000 hours during the week), variable female which takes the price 0 if the individual is man and 1 if the individuals gender is woman, variable married which takes the price 0 if the individual is married and the price 0 of the individual has not yet made his relationship official, variable foreign takes the price 0 if the person was born in Greece and price 1 if he was born in another country, variable university which takes the price 0 if the person's education level is till the post-secondary but non tertiary education and the price 1 if the person has education level bigger than this which referred, internetuse_high which shows us the internet use that the person makes and takes the value 0 if the person uses the internet up to a few times a week and the value 1 if the use is more frequent than this value, variable internetaccess which shows us if the individual has internet access in his house and in this occasion our variable takes the price 1 but if he hasn't access in internet from his home variable takes the price 0, variable religattend which show us how often the respondent attends religious services and takes the price 0 if the person attends religious events up to once a month and the price 1 if he attends religious events at a lower frequency.

Then we limit the sample, we will deal with by setting some criteria on the variables mentioned. Initially, in terms of age, the sample I will use is between the ages of 25 and 54. Then we keep data on people who have education from the minimum - they have not completed primary school and have up to those who hold a doctorate (ISCED 0 = Early childhood education , ISCED 1 = Primary Education, ISCED 2 = Lower Secondary Education, ISCED 3 = Upper Secondary Education, ISCED 4 = Post-secondary non-Tertiary Education, ISCED 5 = Short-cycle tertiary education, ISCED 6 = Bachelors degree or equivalent tertiary education level, ISCED 7 = Masters degree or equivalent tertiary education level, ISCED 8 = Doctoral degree or equivalent tertiary education level). We set the value 0 for people who do not hold a university degree and 1 for people who are graduates of a university school or higher. Summary statistics and other information's of the variables mentioned are shown in the following tables 1,23,4.

The most important conclusions we can draw from table 1 are the following. Initially we notice that the majority of the sample consists of people who work, and more specifically 79.48% of the people from the total sample of 1238 people are working, 9.05% are actively looking for work while the remaining 11.47 % belong to the inactive category. Also from the hours variable we observe that 3.46% of the employed work up to 29 hours per week, while the absolute majority, the remaining 51.93%, work more than 29 hours per week. The absolute majority of the sample consists of people who are married, and more specifically this percentage amounts to 93.56%, as well as regarding the place of birth we notice that 96.78% are Greek. 60.06% are made up of people who have completed up to secondary education, while only 39.94% have higher levels of education. We also notice that the majority of the sample and more

specifically 93.48% of the sample often use the internet while regarding the religious beliefs of the individuals 67.55% of the individuals attend at least once a month a sacrament. Lastly we can observe that the majority of the sample and more specifically the 72.38% they don't have any children.

Table 1: Categorical Variables

Categorical Variables		Freq.	Percent	Cum.
empstatus	1 (employed)	984	79.48	79.48
	2 (unemployed)	112	9.05	88.53
	3 (inactive)	142	11.47	100
	Total	1238	100	
hours	0 (0 hours)	554	44.61	44.61
	1 (1-29 hours)	43	3.46	48.07
	2 (30+ hours)	645	51.93	100
	Total	1242	100	
female	0 (men)	606	48.79	48.79
	1 (women)	636	51.21	100
	Total	1242	100	
married	0 (married)	1162	93.56	93.56
	1	80	6.44	100
	Total	1242	100	
foreign	0 (native)	1202	96.78	96.78
	1	40	3.22	100
	Total	1242	100	
university	0	746	60.06	60.06
	1 (Bachelor's and over degree)	496	39.94	100
	Total	1242	100	
internetuse_high	0	81	6.52	6.52
	1 (most days/ everyday)	1.16	93.48	100
	Total	1242	100	
chldo12	0	899	72.38	72.38
	1	163	13.12	85.51
	2	151	12.16	97.67
	3	24	1.93	99.60
	4	5	0.40	100
	Total	1242	100	
internetaccess	1	41	3.30	3.30

	2	13	1.05	4.35
	3	468	37.68	42.03
	4	720	57.97	100
	Total	1242	100	
religattend	0	839	67.55	67.55
	1	403	32.45	100
	Total	1242	100	

Source: European Social Survey Round 10

In table 2 we observe the range of political beliefs of the individuals in the sample. Where at the two extremes of the right and left quite small percentages of 2.82% and 0.32% are observed respectively. The majority is exactly in the middle and corresponds to 39.29% of the total sample. in general, we notice that the closer we get to the center (value 5), the higher the percentages in our sample, while the closer we get to the extremes, the percentages are quite small.

Table 2: Range of political beliefs

Variable		Freq.	Percent	Cum.
lrscale	Left	4	0.32	0.32
	1	19	1.53	1.85
	2	50	4.03	5.88
	3	144	11.59	17.47
	4	148	11.92	29.39
	5	488	39.29	68.68
	6	110	8.86	77.54
	7	130	10.47	88
	8	88	7.09	95.09
	9	26	2.09	97.18
	Right	35	2.82	100
	Total	1242	100	

Source: European Social Survey Round 10

Table 3 provides us information about the regions that the individuals reside. Those regions are East Macedonia and Thrace, Center Macedonia, West Macedonia, Epirus, Thessaly, Ionian islands, West Greece, Central Greece, Attica, Peloponnese, North Aegean, South Aegean and Crete. As we can observe the regions with the highest proportion are Attica with a percentage of 33.98%, Central Macedonia 20.77% and Thessaly 6.76%, leading us to the conclusion that more than 50% of the people in our sample they live in two regions, Attica and Central Macedonia.

Table 3: Regions

Variable		Freq.	Percent	Cum.
Region, [country]	Attica	422	33.98	33.98
	North Aegean	20	1.61	35.59
	South Aegean	47	3.78	39.37
	Crete	81	6.52	45.89

	East Macedonia and Thrace	79	6.36	52.25
	Central Macedonia	258	20.77	73.03
	Western Macedonia	30	2.42	75.44
	Epirus	42	3.38	78.82
	Thessaly	84	6.76	85.59
	Ionian islands	19	1.53	87.12
	West Greece	61	4.91	92.03
	Central Greece	45	3.62	95.65
	Peloponnese	54	4.35	100.00
	Total	1036	100.00	

Source: European Social Survey Round 10

As can be seen in Table 3, which contains statistics on the ages of individuals, the sample contains 1242 observations, and the mean age of the individuals is 40.92, indicating a distribution close to the mean. The standard deviation that captures the spread of is equal to 8.25, and the age range which we have defined is from 25 to 54 years old.

Table 4: Summary Statistics

Variable	Mean	Std. dev.	Min	Max	Obs
Age	40.92	8.25	25	54	1242

Source: European Social Survey Round 10

Empirical model

For our analysis, a total of three different regression models were used, where depending on the type of dependent variable, the corresponding model was used. These three models are the Multinomial logit model, Ordered probit model and the Ordered logit model.

Where Multinomial logistic regression is a classification method that generalizes logistic regression to multiclass problems, i.e. with more than two possible discrete outcomes. That is, it is a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables (which may be real-valued, binary-valued, or categorical-valued). As the regression coefficients of covariates in the multinomial logit model are not interpretable substantively, a supplementary procedure is to use the fixed-effect estimates to predict the probabilities marginalized at certain covariate values. An alternative to least-squares regression that guarantees the fitted probabilities will be between 0 and 1 is the method of multinomial logistic regression. In the multinomial logit model. So we have:

$$log\frac{Pr(y=k|x)}{Pr(y=K|x)} = {\beta_o}^k + {\beta_1}^k x_1 + \cdots + {\beta_p}^k x_p$$

for k = 1, …, K – 1

In the classification problem, a new observation x_0 is classified as belonging to the group k for which $\widehat{Pr}(y = k|x_0)$ is maximized.

Conditional probability since we have an ordered response model where dependent variable is not strictly continuous nor binary. Ordered response y, y = [0, 1, 2, ..., J]

$$Pr(y_1 = j|x_i)$$

Odds ratios in logistic regression can be interpreted as the effect of a one unit of change in X in the predicted odds ratio with the other variables in the model held constant.

$$\frac{odds\ (if\ the\ corresponding\ variable\ is\ incremented\ by\ 1)}{odds\ (if\ variable\ not\ increamented)}$$

$$\frac{P(event|x+1)/(1-P(event|x+1)}{P(event|x)/P(event|x)}$$

The ordered logit model is an ordinal regression model—that is, a regression model for ordinal dependent variables. For example, if one question on a survey is to be answered by a choice among "poor", "fair", "good", "very good" and "excellent", and the purpose of the analysis is to see how well that response can be predicted by the responses to other questions, some of which may be quantitative, then ordered logistic regression may be used. It can be thought of as an extension of the logistic regression model that applies to dichotomous dependent variables, allowing for more than two (ordered) response categories.

Ordered probit is a generalization of the widely used probit analysis to the case of more than two outcomes of an ordinal dependent variable (a dependent variable for which the potential values have a natural ordering, as in poor, fair, good, excellent). Similarly, the widely used logit method also has a counterpart ordered logit. Ordered probit, like ordered logit, is a particular method of ordinal regression. The Probit model uses the standard normal cumulative function ensuring that the estimated probabilities will be in the range [0,1]. The standard normal cumulative function:

$$\mathrm{G}(z) = \Phi(z) = \int_{-\infty}^{z} \varphi(v)dv$$

In an Ordered Probit Model, we do not observe the latent variable, y*. Instead, we observe outcomes (choices)

$$y = 0\ if\ y^* \leq\ a_1$$

$$y = 1\ if\ a_1 <\ y^* \leq\ a_2$$

$$y = 2\ if\ a_2 <\ y^* \leq\ a_3$$

...

$$y = J\ if\ a_J <\ y^*$$

We estimate the probability of one of the ordered choices of y to happen, when the x variable increases by 1.

Estimation results

Table 5 shows the results of applying the Multinomial Logit model and having Employment Status as a dependent variable. The reason why we use this model has been discussed in the above section (empirical model). The main conclusions we draw from table 5 are the following. Initially we notice that women in our sample are 11.8% less likely to work than men, while they are likely not working and belong to the unemployed category with a probability of 6.2%. Finally, regarding the comparison based on the person's gender, we notice that women again have 5.7% more chances of not looking for work and belonging to the inactive category. All percentages reported are statistically significant at all significance levels. We also notice that if a person has access to the internet at work but not at home, he is 22.6% more likely to work than another person who does not have access to the internet either at work or at home (statistically significant at all significance levels). A person who has access to the Internet from work but not at home is 13% less likely to belong to the inactive category than a person who has no access to the Internet from the aforementioned places (statistically significant at a significance level of 5 %). We notice that if a person has access to the internet both at work and at home, he is 24.4% more likely to work compared to another person who does not have access to the internet either at work or at home. (statistically significant at all significance levels). If the same person who has access to the internet from work but also at home has a 12.2% less chance of belonging to the unemployed category (statistically significant at the 10% significance level), but also a 12.3% less chance (statistically significant at the 5% level of significance), to belong to the category of inactive compared to someone who has no access to the Internet (neither from home nor from work). Finally, we observe that people who attend a religious service less than once a month are 1.5% more likely to belong to the category of inactive people, i.e. people who are not even looking for work, compared to someone who attends to a religious sacrament up to once a month (statistically significant at the 10% level of significance).

Table 5: Multinomial Logit Model Employment Status

Variables	Employed	Unemployed	Inactives
age	0.001 (0.001)	-0.000 (0.001)	0.001 (0.00)
female	-0.118*** (0.017)	0.062*** (0.012)	0.057*** (0.012)
married	0.013 (0.026)	-0.001 (0.020)	-0.012 (0.010)
foreign	0.016 (0.026)	-0.009 (0.022)	-0.006 (0.010)
university	0.021 (0.013)	-0.011 (0.010)	-0.009 (0.007)
internetuse_high	0.006 (0.021)	0.014 (0.014)	-0.019 (0.015)
chldo12	0.005 (0.009)	-0.009 (0.007)	0.004 (0.004)
internetaccess=work	0.226*** (0.082)	-0.096 (0.067)	-0.130** (0.055)
internetaccess=home	-0.066 (0.086)	-0.016 (0.068)	0.082 (0.060)
internetaccess=home & work	0.244*** (0.084)	-0.122* (0.068)	-0.123** (0.055)
religattend	-0.003 (0.014)	-0.012 (0.010)	0.015* (0.008)
lrscale	0.003 (0.004)	-0.003 (0.003)	0.000 (0.002)
Region dummies	Yes	Yes	Yes
Pseudo R2	0.310	0.310	0.310

Source: European Social Survey Round 10

Notes: Standard errors on parenthesis. *** denotes 1% significance, ** 5% significance, and * 10% significance

Table 6 presents the results from a different model, the Ordered Probit. In our analysis the dependent variable is the individual's weekly working hours. First we observe that individuals who hold a university degree are 12.4% less likely to have zero weekly work hours compared to someone who holds a lower level of education (statistically significant at the 1% level of significance). They are also 0.2% less likely to work up to 29 hours per week (statistically significant at the 5% significance level), while they are 12.6% more likely to work more than 30 hours per week than someone without a university degree (statistically significant at all significance levels). We also observe that people who have children over 12 years old are 3.8% less likely to work zero hours per week (statistically significant at the 10% significance level), as well as 3.8% more likely to be in love for more than 30 hours weekly (statistically significant at the 10% significance level). Finally, the access the person has to the internet both from work and from home has a great influence on the weekly working hours. More specifically, people who have internet access from home and work are 26.4% less likely to work zero weekly hours than people who have no internet access from both home and work (statistically significant at all significance levels), while correspondingly the first aforementioned category of people has 26.5% more chances of working more than 30 weekly hours compared to the second category we mentioned (statistically significant at all significance levels).

Table 6: Ordered Probit Model Working hours per week

Variables	0 hours	1-29 hours	30 + hours
age	0.001 (0.002)	0.000 (0.000)	-0.001 (0.002)
female	0.046 (0.030)	0.000 (0.000)	-0.047 (0.031)
married	0.033 (0.064)	0.000 (0.000)	-0.033 (0.064)
foreign	- 0.069 (0.083)	- 0.001 (0.002)	0.071 (0.085)
university	- 0.124*** (0.031)	- 0.002** (0.001)	0.126*** (0.032)
internetuse_high	-0.008 (0.061)	-0.000 (0.001)	0.008 (0.062)
chldo12	-0.038* (0.021)	0.000 (0.000)	0.038* (0.021)
internetaccess=work	-0.162 (0.125)	0.001 (0.003)	0.160 (0.125)
internetaccess=home	0,001 (0.078)	0.000 (0.002)	- 0.001 (0.076)
internetaccess=work & home	- 0.264*** (0.078)	- 0.001 (0.002)	0.265*** (0.076)
religattend	0.020 (0.033)	0.000 (0.000)	- 0.020 (0.033)
lrscale	- 0.001 (0.009)	0.000 (0.000)	0.001 (0.009)
Region dummies	Yes	Yes	Yes
Pseudo R2	0.112	0.112	0.112

Source: European Social Survey Round 10

Notes: Standard errors on parenthesis. *** denotes 1% significance, ** 5% significance, and * 10% significance

Table 7 presents the results from the last model we used for our analysis, the Ordered Logit Model. In our analysis the dependent variable is again the individual's weekly working hours. First, we observe here exactly the same results we found with the Ordered Probit Model regarding the effect of the individual's education on his weekly working hours. And more specifically we have noticed that individuals who hold a university degree are 12.4% less likely to have zero weekly work hours compared to someone who holds a lower level of education (statistically significant at the 1% level of significance). They are also 0.2% less likely to work up to 29 hours per week (statistically significant at the 5% significance level), while they are 12.6% more likely to work more than 30 hours per week than someone without a university degree (statistically significant at all significance levels). Little differences are again observed in these two models regarding the independent variable showing the number of children under 12 years of age. More specifically we observe that people who have children over 12 years old are 3.9% less likely to work zero hours per week (statistically significant at the 10% significance level), as well as 4% more likely to be in love for more than 30 hours weekly (statistically significant at the 10% significance level). Finally, again very small differences are found in the results of using the two models in the internetaccess=work & home variable. The access the person has to the internet both from work and from home has a great influence on the weekly working hours. More specifically, people who have internet access from home and work are 25.5% less likely to work zero weekly hours than people who have no internet access from both home and work (statistically significant at all significance levels), while correspondingly the first aforementioned category of people has 25.7% more chances of working more than 30 weekly hours compared to the second category we mentioned (statistically significant at all significance levels).

Table 7: Ordered Logit Model Working hours per week

Variables	0 hours	1-29 hours	30 + hours
age	0.001 (0.002)	0.000 (0.000)	-0.001 (0.002)
female	0.047 (0.031)	0.001 (0.001)	-0.048 (0.032)
married	0.038 (0.067)	0.000 (0.000)	-0.038 (0.068)
foreign	-0.072 (0.087)	-0.002 (0.003)	0.074 (0.090)
university	-0.124*** (0.032)	-0.002** (0.001)	0.126*** (0.033)
internetuse_high	-0.004 (0.063)	-0.000 (0.001)	0.004 (0.063)
chldo12	-0.039* (0.022)	-0.001 (0.000)	0.040* (0.022)
internetaccess=work	-0.156 (0.119)	0.001 (0.003)	0.154 (0.119)
internetaccess=home	0.016 (0.080)	-0.001 (0.003)	-0.015 (0.077)
internetaccess=work & home	-0.255*** (0.080)	-0.002 (0.003)	0.257*** (0.077)
religattend	0.022 (0.034)	0.000 (0.000)	-0.022 (0.035)
lrscale	-0.002 (0.009)	-0.000 (0.000)	0.002 (0.009)
Region dummies	Yes	Yes	Yes
Pseudo R2	0.111	0.111	0.111

Source: European Social Survey Round 10

Notes: Standard errors on parenthesis. *** denotes 1% significance, ** 5% significance, and * 10% significance

Conclusions

In this article we have studied the factors that can initially affect the employment status of the person as well as the hours of his weekly employment. For our analysis, three regression models were used which are the Multinomial logit model, Ordered probit model and the Ordered logit model. The Empirical model section analyzes precisely the methods used as well as the conditions of their use in order to arrive at the most accurate and unbiased results. As for the data, it is worth noting that the country of interest is Greece and for the study of these factors data from the European Social Survey Round 10 was used. The above research was carried out in Greece between 21/11/2021 and 19/4/2022. The most important conclusions we draw regarding the factors that can influence the employment status of the individual using the Multinomial logit model are the following. Firstly the gender of the individual plays a very important role in the probability that the individual has of being employed with men being in the advantaged category compared to women. In addition, the individual's access to the internet either only from work or both from work and from home increases the probability that the individual works compared to a person by 22.6% and 24.4% respectively which has no access to any of the sites mentioned and therefore no familiarity with the internet. Regarding the weekly working hours to derive the results we used the other two regression models the Ordered probit and the Ordered logit and the factors that seem to influence them are so common but also almost the same in their influence size for both of these models. In short we derive almost the same results from both these models. In this case having a university degree greatly increases the chances of the person working more than 30 hours as well as having access to the internet both at work and at home.

References

Cameron, A. C., & Trivedi, P. K. (2010). Microeconometrics using stata (Vol. 2). College Station, TX: Stata press.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data MIT press. Cambridge, MA, 108(2), 245-254.

https://ec.europa.eu/eurostat