The influence of a person's religion on his income worldwide/in Ukraine

Nazar Liubas

May 2021

Introduction

Man is born to be happy. In the process of learning about the world and forming a worldview, a person is influenced by many factors, including family upbringing, education at school or higher education institutions, traditions, customs, religion. Everyone will say that to be happy, the main thing is to be in good health, but most people will also believe that happiness requires material goods and money. And most people will get an education, work hard, organize various projects, etc. to earn an income that will provide a comfortable environment for maintaining physical and intellectual condition. Since religion plays an important role in human life, I was interested in studying the problem of the impact of religion creed on people's incomes.

Literature review

There is not much research about this topic, but L. Bettendorf and E. Dijkgraaf[1] investigated how the relationship between income of people and their religion differs in low- and high-income countries. They found that church membership has a positive effect on income for high-income countries, this effect is negative for low-income countries.

---

[1] L. Bettendorf, E. Dijkgraaf, Religion and income: Heterogeneity between countries,
Journal of Economic Behavior & Organization,
Volume 74, Issues 1–2, 2010, Pages 12-29, ISSN 0167-2681,
https://doi.org/10.1016/j.jebo.2010.02.003.
(https://www.sciencedirect.com/science/article/pii/S0167268110000235)

Data description

Data is taken from joint European and World Values Survey[2]. Surveys were conducted at 2017-2020 in 79 countries. There is data about 127,358 people that responded to 231 questions. However, not all questions were asked to each person, as, for example, there is only about 50% of data at variable "Level of income".

Methodology explanation

As the aim of this research is to find the relationship between a person's religion and his income, the first thought is to build a model that predicts income with "religious" variables. Variable "level of income" in the dataset is represented as a person's subjective assessment of the household income on a 10-points scale with 1 indicating the lowest income group in the country, and 10 - the highest. That means that the dependent variable has categorical data, so we can't use a linear model. But we can use the probit model, by creating binary dependent variables that equal 1 if a person's income is bigger than 5, and 0 otherwise.

As explanatory variables, we take dummy variables for 10 denominations of people (Do not belong to a denomination (ateists), Roman Catholic, Protestant, Orthodox (Russian/Greek/etc.), Jew, Muslim, Hindu, Buddhist, Other Christian (Evangelical/Pentecostal/Free church/etc.), Other). That means, if a person is Jew, there will be value 1 for variable Jew, and 0 for others.

Remark, that if there are N dummy variables, we can put in a model only N-1 of them to avoid breaking the multicollinearity assumption. Also, we should remember that a variable that isn't in the model is a basis group, so estimates of the rest of dummy variables measure the proportionate difference in regressand relative to the basis group. Also, even changing a basis dummy variables, such data about models as deviances, AIC, Number of Fisher Scoring iterations will remain the same.

First, we consider Ukraine. Remark that there is no data about buddhists in Ukraine, so we will include in model only 8 dummies. Also, we run 9 models, each time choosing a different basis variable.

Then we conduct all the same procedures, but for all world (all dataset), and including buddhists, so 9 dummies in result.

---

[2] https://www.worldvaluessurvey.org/WVSEVSjoint2017.jsp

Results

Interpreting estimated slopes of binary models, we look only for signs of slopes, as to get size of actual change in dependent variable from change in independent variables, we should take derivatives of logistic function.

Ukraine results

| | | Basis Dummy | | | | | | | |
| | | roman | | | | other | | | |
	ateist	catholic	orthodox	muslim	jew	christian	protestant	other	hindu
ateist	na	-	-	-	-	+	-	-	+
roman_catholic	+	na	+, *	+	-	+, *	+	-	+
orthodox	+	-, *	na	-	-	+	-	-	+
muslim	+	-	+	na	-	+	+	-	+
jew	+	+	+	+	na	+	+	+	+
other_christian	-	-, *	-	-	-	na	-	-, .	+
protestant	+	-	+	-	-	+	na	-	+
other	+	+	+	+	-	+, .	+	na	+
hindu	-	-	-	-	-	-	-	-	na

So, there is a table, where only signs of estimated slopes near dummies were taken. If slopes have some statistical significance, it is given after a comma. Significant codes are: '***' 0.001; '**' 0.01; '*' 0.05; '.' 0.1.
Interpretation is next: these signs are given, related to the basis dummy. For example, if the basis dummy is roman catholic, the estimated value opposite to orthodox is with the sign "-". So, we may say that at significance level 0.05 income of orthodox is less than income of roman catholic.
Talking about goodness-of-fit of model, we have such results:

```
    Null deviance: 1186.9  on 974  degrees of freedom
Residual deviance: 1175.6  on 966  degrees of freedom
AIC: 1193.6

Number of Fisher Scoring iterations: 12
```

From here we can calculate pseudo R^2 by formula $R^2 = 1 - \frac{L_{ur}}{L_r} = 1 - \frac{Residual\ deviance}{Null\ deviance}$.

For this model, it equals 0.00950936. It is quite a poor result, as only 1% of variation in data is explained by this model.

World results

	Basis Dummy									
	ateist	roman catholic	orthodox	muslim	jew	other christian	protestant	other	hindu	buddhist
ateist	na	-	-, .	-	-	-	-, **	+, ***	-, **	+, ***
roman_catholic	+	na	-	+	-	-	-, .	+, ***	-, **	+, ***
orthodox	+, .	+	na	+	-	+	-	+, ***	-, *	+, ***
muslim	+	-	-	na	-	-	-, *	+, ***	-, **	+, ***
jew	+	+	+	+	na	+	+	+, *	-	+, **
other_christian	+	+	-	+	-	na	-	+, ***	-, *	+, ***
protestant	+, **	+	+	+, *	-	+	na	+, ***	-, .	+, ***
other	-, ***	-, ***	-, ***	-, ***	-, *	-, ***	-, ***	na	-, ***	+
hindu	+, **	+, **	+, *	+, **	+	+, *	+, .	+, ***	na	+, ***
buddhist	-, ***	-, ***	-, ***	-, ***	-, **	-, ***	-, ***	-	-, ***	na

In this case interpretation is the same, but the results are different. There we can see that representatives of buddhist, hindu and other, not listed religions have income levels lower than other religions from the list at the significance level 0.01.

Results of model are next:

```
    Null deviance: 87865  on 68274  degrees of freedom
Residual deviance: 87739  on 68265  degrees of freedom
AIC: 87759
```

Pseudo R^2 for this model equals 0.001442151 and AIC is too high, which is again a poor result.

Conclusion/limitations and next steps

There were some limitations that could make results worse. First, the level of person's income in this data is a subjective feeling that could not represent the real state of affairs. Also, in Ukraine there were quite a small samples of people that represent religious groups:

```
Number of representatives:
ateists:  32
roman_cath:  91
protestants:  9
orthodox:  772
jews:  3
muslims:  3
hindus:  2
others:  5
other_christs:  58
```
.

But, having data that is available, we can make a conclusion that the religious denomination isn't a crucial factor that implies income level, as it shows our models with poor goodness of fit.

However, we can see that for some reasons some religious groups have a bigger level of income than others, e.g. roman catholic against orthodox and other christian in Ukraine.

In the future, it could be interesting to explore deeply dependence of income level and other "religious" variables from the dataset, such as frequency of praying, active/non-active membership in church, etc.