Problem Set #6: Regression

Using the Turkey Demographic and Health Survey (TDHS) micro-level data explored in the previous problem set, I estimate a series of probit regressions in order to analyze the changes in differences in the probability of being a college graduate among the individuals who wear headscarf and who do not, before and after the lifting of the headscarf ban. The empirical specification is

$$C_i^* = \beta_0 + \beta_1 Headscarf_i + \beta_2(Headscarf_i \times d_i) + \beta_3 d_i + \mathbf{\Gamma X_i} + \epsilon_i \tag{1}$$

where C_i is a dummy variable indicating that college level education is the highest level of education attained by an individual

$$C_i = \begin{cases} 1, & \text{if } C_i^* > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Variable *Headscarf* is a dummy variable indicating whether the respondent wears a headscarf when she goes out. d_i is a dummy variable that equals one if the respondent's birth year is greater than or equal to 1990. These individuals were at least eighteen years old in 2008, the year the ban was lifted. Hence their college education demand is not affected by the ban. $\mathbf{X_i}$ is a vector of control variables with corresponding coefficients contained in $\mathbf{\Gamma}$. These control variables include; age, being married, whether the mother is a college graduate, whether the father is a college graduate, proxy variable for conservatism that indicates whether domestic violence in the household is considered normal. Throughout the analysis of college education, the sample is restricted to those who are at least 18 years old at the time of the survey. This eliminates the individuals who are not old enough to

get college education, which would otherwise interfere with our analysis. The results are,

Table 1: Regression results.

	Model 1	Model 2	Model 3
headscarf	-0.16^{***}	-0.13^{***}	-0.19^{***}
	(0.01)	(0.01)	(0.02)
headscarf $\times$ d9094	0.11^{***}	0.09^{***}	0.13^{***}
	(0.02)	(0.02)	(0.03)
d9094	-0.03^{***}	-0.02^{***}	0.02
	(0.01)	(0.00)	(0.02)
age	-0.00^{***}	-0.00^{***}	0.03^{***}
	(0.00)	(0.00)	(0.00)
married	-0.09^{***}	-0.08^{***}	-0.23^{***}
	(0.01)	(0.01)	(0.02)
mother-college	0.06^{*}	0.05^{*}	0.03
	(0.03)	(0.02)	(0.04)
father-college	0.24^{***}	0.21^{***}	0.26^{***}
	(0.02)	(0.02)	(0.05)
domest-violence		-0.06^{***}	-0.07^{***}
		(0.00)	(0.01)
Num. obs.	17518	17518	3231
Log Likelihood	-4063.13	-3966.10	-1031.94
Deviance	8126.27	7932.20	2063.88
AIC	8142.27	7950.20	2081.88
BIC	8204.43	8020.13	2136.61
Baseline controls	+	+	+
Conservatism	−	+	+
Closer cohorts	−	−	+

Average marginal effects are reported. Standard errors are shown in parentheses.

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Model 1 is a preliminary regression where only the baseline controls are applied. We can see that the interaction term is significantly positive, which indicates that once the ban was lifted, the probability of individuals who wear headscarf obtaining college level education increased by 11% (Note that average marginal effects are reported). The gap before the ban was lifted was approximately 16%. Summing these two coefficients show that even

when the ban was lifted, females who wear headscarf has a lower probability of obtaining college level education compared to females who don't wear headscarf. In addition, being married decreases the probability of obtaining college level education while parents being college graduates positively impacts the probability of the individual obtaining college level education. Model 2 includes the baseline controls as well as a proxy variable to control for the individual's degree of conservatism. Once the controls are added, the magnitude of the interaction term becomes smaller. This makes sense as there are more variables that can explain the probability of an individual obtaining college level education in this setup. The log-likelihood has also decreased in absolute terms, indicating a better fit. We can also see that the conservatism-proxy variable decreases the individual obtaining college level education by 6%. Model 3 uses a restricted sample where only the individuals who were born after 1984 are included in the analysis. This allows for a comparison of closer cohorts. We can observe that the interaction term has a stronger magnitude now. It implies that the impact of the ban is stronger for closer cohorts, further indicating that the ban was a significant intervention altering individuals' college level education patterns.