**WRITING SAMPLE**

**The following document along with the do.file and any input and output in the folder is one of my Stata exercises at Georgetown. It uses a household-level dataset with 38797 observations to clean, process, and analyze data. The document, codes, and outputs are my own work. Please do not distribute it without my permission. The file can also be accessed via GitHub at** [**https://github.com/GGGGUKIM/Coding\_sample\_Stata.git**](https://github.com/GGGGUKIM/Coding_sample_Stata.git)**.**

**Table 1[[1]](#footnote-1):** Please compute the household composition by the following age groups, and report for **each country** in the LiTS sample, the average (weighted) household age composition.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Country | Share of children under age 15 | Share of youth (age 15-24) | Share of adults (age 25-69) | Share of elderly (70+) |
| Albania | 5.22% | 18.21% | 70.72% | 5.84% |
| Armenia | 15.77% | 18.20% | 55.30% | 10.73% |
| Azerbaijan | 17.76% | 21.92% | 55.90% | 4.42% |
| Belarus | 11.09% | 15.43% | 67.42% | 6.06% |
| Bosnia | 9.08% | 12.00% | 65.97% | 12.95% |
| Bulgaria | 5.53% | 7.17% | 64.91% | 22.38% |
| Croatia | 9.13% | 7.37% | 63.47% | 20.03% |
| Czech Republic | 11.16% | 10.52% | 72.22% | 6.11% |
| Estonia | 9.57% | 13.41% | 54.79% | 22.23% |
| France | 13.75% | 8.25% | 60.01% | 18.00% |
| Georgia | 13.25% | 12.61% | 56.11% | 18.03% |
| Germany | 5.64% | 6.97% | 71.28% | 16.12% |
| Great Britain | 7.43% | 5.85% | 54.66% | 32.06% |
| Hungary | 6.32% | 7.25% | 61.76% | 24.67% |
| Italy | 10.52% | 8.02% | 66.04% | 15.42% |
| Kazakhstan | 11.41% | 18.70% | 65.90% | 3.99% |
| Kosovo | 15.26% | 22.58% | 59.22% | 2.94% |
| Kyrgyzstan | 16.38% | 24.12% | 55.84% | 3.66% |
| Latvia | 8.17% | 10.31% | 56.63% | 24.89% |
| Lithuania | 7.02% | 10.00% | 57.60% | 25.38% |
| Macedonia | 13.25% | 13.79% | 62.99% | 9.97% |
| Moldova | 12.70% | 12.69% | 61.34% | 13.26% |
| Mongolia | 16.02% | 23.02% | 57.95% | 3.01% |
| Montenegro | 12.19% | 16.08% | 63.35% | 8.38% |
| Poland | 10.22% | 9.38% | 66.67% | 13.72% |
| Romania | 8.96% | 8.74% | 63.14% | 19.16% |
| Russia | 11.20% | 12.06% | 61.04% | 15.70% |
| Serbia | 8.49% | 9.24% | 62.80% | 19.47% |
| Slovakia | 12.38% | 16.01% | 69.47% | 2.14% |
| Slovenia | 8.35% | 11.25% | 66.05% | 14.35% |
| Sweden | 6.89% | 5.63% | 69.87% | 17.60% |
| Tajikistan | 21.27% | 25.76% | 50.37% | 2.59% |
| Turkey | 20.48% | 15.22% | 60.61% | 3.69% |
| Ukraine | 11.55% | 9.45% | 67.70% | 11.30% |
| Uzbekistan | 24.75% | 20.64% | 51.40% | 3.20% |

**Table 2[[2]](#footnote-2):** Please compute the total monthly per capita household expenditures based on items in questions 2.22 and 2.24 of the questionnaire. In the LiTS survey household expenditures are recorded in national currencies. Based on this welfare aggregate please construct (i) population quintiles within each country; and (ii) a poverty indicator for each household that equals 1 if the household’s expenditures per capita are below 60% of the median household expenditures per capita in that country, and zero otherwise. Then please report, for each country, the average per capita expenditures for each quintile, as well as the poverty rate (share of population below the poverty line), as in the table below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Country | Average monthly per capita household expenditures | | | | | Poverty rate (%) |
| Quantile 1 | Quantile 2 | Quantile 3 | Quantile 4 | Quantile 5 |
| Albania | 2591.16 | 12649.80 | 22323.57 | 37881.84 | 98444.48 | 31.09 |
| Armenia | 19103.26 | 41769.02 | 69573.57 | 115009.70 | 237659.90 | 29.38 |
| Azerbaijan | 95.91 | 183.00 | 272.44 | 414.48 | 889.20 | 24.58 |
| Belarus | 33665.41 | 436306.50 | 908519.10 | 1620962.00 | 3806343.00 | 34.00 |
| Bosnia | 125.90 | 327.72 | 509.80 | 790.53 | 1758.57 | 26.84 |
| Bulgaria | 135.19 | 372.33 | 556.50 | 819.93 | 1712.84 | 26.11 |
| Croatia | 950.35 | 2314.55 | 3576.16 | 5345.64 | 12447.89 | 25.58 |
| Czech Republic | 3718.15 | 8899.80 | 13504.77 | 19538.67 | 47335.56 | 25.70 |
| Estonia | 2789.81 | 5427.68 | 7881.82 | 11250.48 | 23817.88 | 24.01 |
| France | 381.59 | 898.15 | 1370.08 | 2115.49 | 5655.86 | 25.99 |
| Georgia | 55.72 | 159.11 | 278.68 | 490.79 | 1174.47 | 31.14 |
| Germany | 462.91 | 965.81 | 1618.34 | 2790.85 | 7148.05 | 29.71 |
| Great Britain | 201.57 | 426.53 | 707.46 | 1221.39 | 4255.64 | 30.08 |
| Hungary | 27610.36 | 62982.41 | 93540.77 | 136261.40 | 248732.00 | 25.55 |
| Italy | 380.95 | 698.49 | 1007.68 | 1490.06 | 3171.55 | 22.07 |
| Kazakhstan | 9351.17 | 26027.14 | 43588.64 | 75935.29 | 169525.40 | 30.06 |
| Kosovo | 60.36 | 147.81 | 226.37 | 333.76 | 792.40 | 26.12 |
| Kyrgyzstan | 1948.68 | 5304.78 | 7615.29 | 11355.24 | 21327.73 | 23.47 |
| Latvia | 115.72 | 232.98 | 340.87 | 495.06 | 915.33 | 25.06 |
| Lithuania | 470.86 | 1032.64 | 1608.11 | 2295.17 | 4784.27 | 26.10 |
| Macedonia | 7083.28 | 13432.97 | 20300.76 | 31905.65 | 61609.45 | 27.96 |
| Moldova | 798.66 | 2042.50 | 3234.91 | 5056.90 | 10880.57 | 32.11 |
| Mongolia | 64814.73 | 197347.50 | 355269.00 | 638637.80 | 1541501.00 | 29.80 |
| Montenegro | 91.86 | 238.34 | 389.63 | 617.60 | 1312.79 | 31.89 |
| Poland | 59.64 | 541.93 | 1022.26 | 1722.67 | 3969.70 | 27.79 |
| Romania | 239.61 | 603.93 | 964.71 | 1505.91 | 3248.93 | 28.97 |
| Russia | 4559.26 | 11151.85 | 18142.11 | 28099.83 | 60470.70 | 26.88 |
| Serbia | 8743.62 | 20522.39 | 31475.61 | 47334.74 | 90190.98 | 32.31 |
| Slovakia | 84.51 | 253.28 | 441.97 | 700.17 | 1493.37 | 31.01 |
| Slovenia | 171.83 | 530.47 | 903.55 | 1457.00 | 3456.14 | 25.94 |
| Sweden | 5405.05 | 10431.62 | 15598.15 | 23797.44 | 69987.47 | 27.57 |
| Tajikistan | 97.56 | 264.33 | 406.06 | 630.54 | 1473.76 | 23.83 |
| Turkey | 261.35 | 493.51 | 724.63 | 1105.27 | 2476.48 | 23.83 |
| Ukraine | 815.06 | 1791.02 | 2747.35 | 4219.20 | 9067.73 | 25.99 |
| Uzbekistan | 70860.27 | 126004.00 | 172239.50 | 233091.00 | 377230.00 | 20.66 |

**Table 3:** Now using the poverty indicator you created above as the dependent variable and selected explanatory variables of your choice (4 variables max), please run a regression and interpret the results. Please don’t worry about reverse causality.

As the table shown below, I have run a probit model of the probability of a household being “poor” on monthly payment for mortgage (q209\_t1), length of mortgage (q213), children numbers in the household, and elder numbers in the household. Since the coefficients of a probit model cannot be directly interpreted, I also report the average marginal effects separately. As the table shows, the coefficient of the monthly payment for mortgage is not statistically significant at the conventional level. The coefficient of the length of mortgage is significantly negative at five percent level, and its marginal effect suggests that the probability of a household being “poor” decreases by about .1 percentage points per additional year of mortgage. The coefficients of children numbers and elder numbers are both significantly positive at five percent level. The marginal effect of children numbers means that the probability of a household being “poor” increases by about 5.7 percentage points with one more child in the household, while the one of elder numbers indicate that the likelihood of a household being “poor” increases by about 6.7 percentage points with one more elder in the household. The constant in the model is negative and statistically significant at the five percent level, but it is not practical since the probability cannot be negative.

**Table 3**. Probit results and average marginal effects

|  |  |  |
| --- | --- | --- |
| **Dependent variable** | (1) | **Poor**  (2) |
|  | Probit coefficients | Average Marginal Effects |
|  |  |  |
| Monthly payment for mortgage | .000 | .000 |
|  | (.000) | (.000) |
| Length of mortgage | -.003\*\*\* | -.001\*\*\* |
| Children numbers | (.000)  .175\*\*\* | (.000)  .057\*\*\* |
|  | (.013) | (.004) |
| Elder numbers | .205\*\*\* | .067\*\*\* |
| Constant | (.023)  -.973\*\*\* | (.007) |
| **Sensitivity**  **Specificity**  Observations | (0.035)  1.35%  99.32%  38797 | 38797 |

*Notes: The model has been weighted by using the democratic sampling weights designed by the questionnaire team. Robust standard errors are in parentheses. \*\*\* p<.01, \*\* p<.05, \* p<.1. All coefficients and average marginal effects are rounded up to three decimal places. Given that my dependent variable is dichotomized, I use sensitivity and specificity to describe the goodness of fit for my Probit models. Sensitivity (True Positive Rate) refers to the probability of a positive prediction, conditioned on truly being positive, while specificity (True Negative Rate) refers to the probability of a negative prediction, provided one truly being negative. If the predicted value is greater than 0.5, it is regarded as positive (1) and vice versa.*

1. All numbers reported in this table and used in the process of filling the table have been weighted by using the democratic sampling weights designed by the questionnaire team. All numbers are rounded up to two decimal places. [↑](#footnote-ref-1)
2. The household expenditures reported in this table include food consumptions (food, beverages, and tobacco), utilities (electricity, water, gas, heating, and fixed line phone), transportation (public transportation and fuel for car), education (tuition, books, kindergarten expenses), health (medicines and health insurance), clothing and footwear, and durable goods (e.g. furniture, household appliances, TV, car, etc). I have treated all missing values (-99, -98, and -97) in those expenditures variables as 0 in the process of doing calculations. The Quantile 1 groups represents households with monthly per capita household expenditures ranging from the lowest to the 20 percentiles in different countries. The Quantile 2 groups represents households with monthly per capita household expenditures ranging from the 20 percentiles to the 40 percentiles in different countries. The Quantile 3 groups represents households with monthly per capita household expenditures ranging from the 40 percentiles to the 60 percentiles in different countries. The Quantile 4 groups represents households with monthly per capita household expenditures ranging from the 60 percentiles to the 80 percentiles in different countries. And the Quantile 5 groups represents households with monthly per capita household expenditures ranging from 80 percentiles to the 99 percentiles in different countries. The expenditures in this table are recorded in national currencies. All numbers reported in this table and used in the process of filling the table have been weighted by using the democratic sampling weights designed by the questionnaire team, and all numbers are rounded up to two decimal places. [↑](#footnote-ref-2)