Logistic & Multinomial Regression for The Perception of Corruption in Germany

Yahia M. Quolquela (5200901)

Statistics Department, Faculty of Economics & Political Science, Cairo University

SE402: Categorical Data Analysis

Dr. Sally Farid Abdelfatah

January 9, 2024

Contents

Key points

- Before conducting the binary logistic regression, our dependent variable need to be binary (perception of corruption). Perception of corruption is a categorical variable with 10 levels. The first five levels were recoded as no corruption & the last five were corruption. For multinomial regression the first three categories became low level of corruption & the last 3 became high levels of corruption while the middle 4 became moderate levels of corruption.
- The reference category for the explanatory variables are the first category (default in stata)

Table(1): variables & their categories

Variables	Categories
State Authorities, Business executives, Civil service & Journalists media	1- None of them 2- Few of them 3- Most of them 4- All of them
People pay bribe	1- Never 2- Rarely 3- Frequently 4- Always
Educational level	1-primary 2-Secondary 3-Higher
Worries Losing job	1- Very Much 2- A Good Deal 3- Not Much 4- Not at All
Confidence in the government	1- A Great Deal 2- Quite a Lot 3- Not Very Much 4- Not at All
Justifiable political violence	1- Never justifiable 2- Rarely justifiable 3- Sometimes justifiable 4- Often justifiable 5- Always justifiable
Importance of god	1- Not at all important 2- Moderately important 3- Very important
Income level	1- Lower class 2- Middle class 3- Upper class

Binary Logistic Regression

Table(2): binary logistic regression

Corruption level	**Odds Ratio**	**Std. Err.**	**z**	**P>|z|**	**95% Conf.Interval**	
State Authorities						
Few of them	2.329289	0.6790212	2.9	0.004	1.315485	4.1244
Most of them	8.381679	3.192595	5.58	0	3.972894	17.68296
All of them	25.91076	38.15079	2.21	0.027	1.446032	464.2825
People pay bribe						
Rarely	1.393664	0.1991902	2.32	0.02	1.053173	1.844236
Frequently	2.082378	0.9471071	1.61	0.107	0.8539181	5.078121
Business executives						
Few of them	1.276621	0.5330805	0.58	0.559	0.5631503	2.894006
Most of them	2.083947	0.8979085	1.7	0.088	0.8956332	4.848899
All of them	1.34342	1.200776	0.33	0.741	0.2330187	7.745203
Worries Losing job						
A Good Deal	0.6049018	0.2022233	1.5	0.133	0.3141389	1.164791
Not Much	0.964364	0.2802418	0.12	0.901	0.545612	1.704504
Not at All	0.9925938	0.278027	0.03	0.979	0.5732562	1.718677
Civil service						
Few of them	1.668972	0.2846951	3	0.003	1.194676	2.331568
Most of them	4.247533	1.791933	3.43	0.001	1.857954	9.710428
All of them	0.5372105	0.8557867	0.39	0.696	0.0236681	12.1934
Confidence in the government						
Quite a Lot	1.335332	0.5212965	0.74	0.459	0.6212855	2.870037
Not Very Much	2.908609	1.140665	2.72	0.006	1.348562	6.273356
Not at All	3.43163	1.490173	2.84	0.005	1.465106	8.037699
Journalists media						
Few of them	0.774842	0.2312365	0.85	0.393	0.4317065	1.390713
Most of them	1.227422	0.3899022	0.65	0.519	0.65857	2.287632
All of them	1.225793	0.7255566	0.34	0.731	0.384224	3.91066
cons	0.0859744	0.047232	4.47	0	0.0292915	0.252346

Logistic regression

Number of obs = 1,182
Log likelihood = -676.0572
LR chi2(20) = 286.21
Prob > chi2 = 0.0000
Pseudo R2 = 0.1747

From the previous output we can see that:

- The model is statistically significant as Prob > chi2 = 0.0000 which is less than 0.05.

- Some variables in the model are not statistically significant so we will not interpret them. Such as: business executives, worries losing job & journalists media

Interpretation of significant variables

- For the variable State Authorities
  - e^{β} few of them = 2.329289: the estimated odds of the perception of corruption among Germans who think that few of the state authorities are involved in corruption is 132.9% greater than that of Germans who think none of the state authorities are involved in corruption holding other variables constant.
  - e^{β} most of them = 8.381679: the estimated odds of the perception of corruption among Germans who think that most of the state authorities are involved in corruption is 8.381679 times that of Germans who think none of the state authorities are involved in corruption holding other variables constant.
  - e^{β} All of them = 25.91076: the estimated odds of the perception of corruption among Germans who think that all of the state authorities are involved in corruption is 25.91076 times that of Germans who think none of the state authorities are involved in corruption holding other variables constant.
- For the variable people pay bribe
  - e^{β} Rarely = 1.393664: the estimated odds of the perception of corruption among Germans who think that people rarely pay a bribe to get services is 39.37% greater than that of Germans Who think that people never pay a bribe to get services holding other variables constant.
- For the variable Civil service
  - e^{β} Few of them = 1.668972: the estimated odds of the perception of corruption among Germans who think that few of the Civil services are involved in corruption is 66.9% greater than that of Germans who think none of the civil services are involved in corruption holding other variables constant.
  - e^{β} most of them = 4.247533: the estimated odds of the perception of corruption among Germans who think that most of the Civil services are involved in corruption is 4.247533 times that of Germans who think none of the civil services are involved in corruption holding other variables constant.
- For the variable confidence in the government
  - e^{β} not very much = 2.908609: the estimated odds of the perception of corruption among Germans whose confident in the government is quite a lot is 2.908609 times that of the Germans whose confident in the government is a great deal holding all other variables constant.
  - e^{β} not at all = 3.43163: the estimated odds of the perception of corruption among Germans whose confident in the government is not at all is 3.43163 times that of the Germans whose confident in the government is a great deal holding all other variables constant.

Testing the goodness of fit of the model

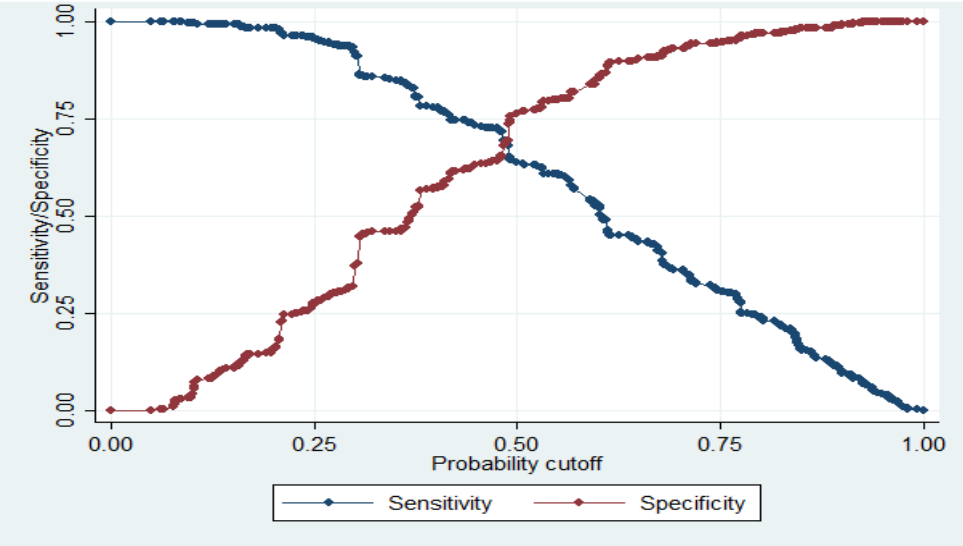
Table(3):Logistic model for Corruption level, goodness-of-fit test

number of observations =	**1182**
number of covariate patterns =	**411**
Pearson chi2(390) =	**405.07**
Prob > chi2 =	**0.2889**

From table(3) we can say that the model is a good fit for the data (p-value > 0.05)

The predictive power of the model

Figure (1): plot fitting sensitivity & specificity against cutoff point



From figure (1), we can determine the best cutoff point to make the classification table.

From the plot the best cutoff point is 0.484

Table(4): classification table

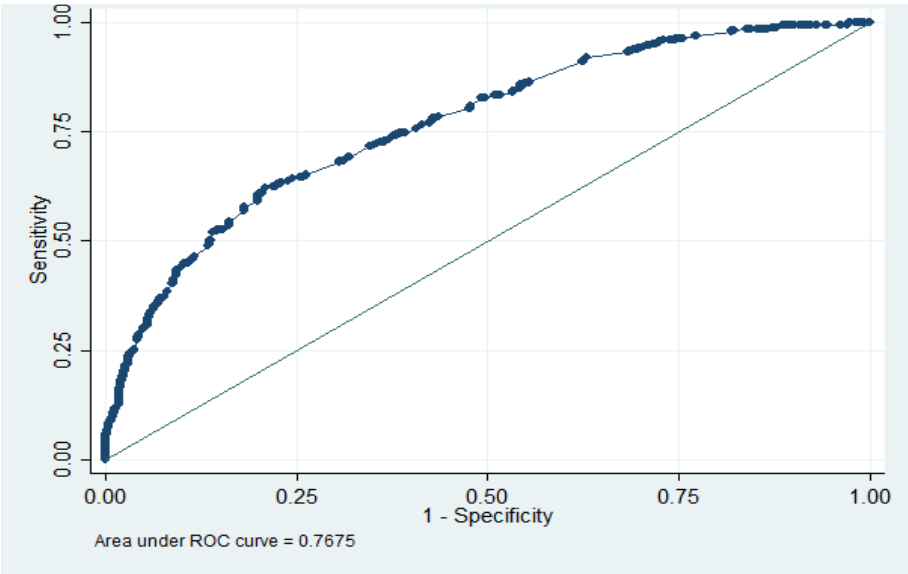
	TRUE		
Classified	D	~D	Total
+	402	190	592
-	180	410	590
Total	582	600	1182

Sensitivity **Pr(+| D) 69.07%**

Specificity	**Pr(-|~D)**	**68.33%**
Correctly classified		**68.70%**

From the classification table, the sensitivity is 69.07% which is the probability of correctly specifying Germans who have perception of corruption while the specificity is 68.33% which is the probability of correctly specifying Germans who do not have perception of corruption. The percentage of correct classification is 68.7%. since the sensitivity, specificity & correct classification are all greater than 60%, therefore the model has acceptable predictive power.

Figure (2): ROC Curve



The ROC curve is more informative than classification table since it plots against all possible cutoff points. The Area Under the Curve(AUC) for the ROC curve is 0.7675. this indicates that the model is better than the intercept-only model & the model has good predictive power. Also AUC is called concordance index, which estimates the probability that the predictions and outcomes are concordant.

Multicollinearity Analysis

From table (5) the VIF values for all the variables are less than 10, hence the variables are not correlated (no multicollinearity)

Table(5):Collinearity Diagnostics

Variable	VIF	SQRT VIF	Tolerance	R-Squared
_IState_Aut_2	2.73	1.65	0.3662	0.6338
_IState_Aut_3	2.81	1.68	0.3563	0.6437
_IState_Aut_4	1.52	1.23	0.6567	0.3433
_Ipeople_pa_2	1.12	1.06	0.8918	0.1082
_Ipeople_pa_3	1.13	1.06	0.8854	0.1146
_IBusiness__2	7.29	2.7	0.1372	0.8628
_IBusiness__3	7.45	2.73	0.1342	0.8658
_IBusiness__4	1.42	1.19	0.7022	0.2978
_IWorries_L_2	2.25	1.5	0.4439	0.5561
_IWorries_L_3	3.7	1.92	0.2704	0.7296
_IWorries_L_4	4.07	2.02	0.2456	0.7544
_ICivil_ser_2	1.47	1.21	0.6804	0.3196
_ICivil_ser_3	1.52	1.23	0.6589	0.3411
_ICivil_ser_4	1.38	1.17	0.7249	0.2751
_IConfidenc_2	6.63	2.58	0.1508	0.8492
_IConfidenc_3	6.91	2.63	0.1447	0.8553
_IConfidenc_4	3.61	1.9	0.277	0.723
_IJournalis_2	4.54	2.13	0.2201	0.7799
_IJournalis_3	4.45	2.11	0.2249	0.7751
_IJournalis_4	1.77	1.33	0.5636	0.4364
Mean VIF	3.39			

Multicategory Logit Models

Ordinal Logistic Regression

Since our dependent variable (perception of corruption) is ordinal, we should apply Ordinal Logistic Regression. First we need to check if the proportional odds assumption is satisfied.

Table(6):Approximate likelihood-ratio test of proportionality of odds across response categories:

chi2(20) =	**35.77**
Prob > chi2 =	**0.0164**

The proportional odds assumption is violated since P-value < 0.05, hence we will not use ordinal logistic regression. We will use multinomial logistic regression.

Multinomial Logistic Regression

Corruption_level	RRR	Std. Err.	z	P>z	95% Conf.Interval	
low_levels_of_corruption						
State_Authorities						
Few of them	0.371641	0.099085	-3.71	0	0.220384	0.626711
Most of them	0.148343	0.08135	-3.48	0.001	0.050638	0.43457
All of them	1.73E-07	0.000521	-0.01	0.996	0	0
Business_executives						
Few of them	0.41426	0.169461	-2.15	0.031	0.185811	0.923578
Most of them	0.263763	0.118084	-2.98	0.003	0.109683	0.634293
All of them	1.02E-07	0.000235	-0.01	0.994	0	0
Civil_service						
Few of them	0.731511	0.139446	-1.64	0.101	0.503453	1.062877
Most of them	0.292652	0.232918	-1.54	0.123	0.061502	1.392555
All of them	1.478663	8199.134	0	1	0	0
Journalists_media						
Few of them	0.949358	0.307784	-0.16	0.873	0.502884	1.792221
Most of them	0.516564	0.19105	-1.79	0.074	0.250212	1.066447
All of them	1.63E-07	0.000245	-0.01	0.992	0	0
people_pay_bribe						
Rarely	0.717199	0.134562	-1.77	0.076	0.496519	1.035963
Frequently	0.459034	0.321553	-1.11	0.266	0.1163	1.811803
Educational_level						
Secondary Education	0.382596	0.208555	-1.76	0.078	0.131446	1.113612
Higher Education	0.916108	0.629402	-0.13	0.899	0.238307	3.521728
Worries_Losing_job						
A Good Deal	1.076963	0.450051	0.18	0.859	0.474781	2.442914
Not Much	1.334927	0.490353	0.79	0.432	0.649815	2.742368
Not at All	1.154873	0.408846	0.41	0.684	0.577023	2.3114
Confidence_in_the_government						
Quite a Lot	0.25084	0.094623	-3.67	0	0.119758	0.525399
Not Very Much	0.186207	0.07303	-4.29	0	0.086329	0.401637
Not at All	0.309279	0.149917	-2.42	0.015	0.119605	0.79975
justifiablepoliticalviolence						
Rarely justifiable	1.414937	0.64722	0.76	0.448	0.577272	3.468114
Sometimes justifiable	0.423347	0.365252	-1	0.319	0.078037	2.296627
Often justifiable	8.635731	8.699546	2.14	0.032	1.198968	62.20005
Always justifiable	4.531872	6.695292	1.02	0.306	0.250455	82.00207
importanceofgod						
Moderately important	0.861226	0.17614	-0.73	0.465	0.576802	1.2859

Very important	1.046874	0.206948	0.23	0.817	0.710604	1.542273
incomelevel						
Middle class	0.296991	0.170603	-2.11	0.035	0.096334	0.915603
Upper class	3.47E-08	0.000112	-0.01	0.996	0	0
edu_level_2_inc_level_2	2.551187	1.664739	1.44	0.151	0.710074	9.166024
edu_level_2_inc_level_3	2.84E+07	9.14E+10	0.01	0.996	0	0
edu_level_3_inc_level_2	2.01246	1.563937	0.9	0.368	0.438768	9.230368
edu_level_3_inc_level_3	3.36E+07	1.08E+11	0.01	0.996	0	0
cons	33.88782	25.26052	4.73	0	7.86221	146.0638
Moderate_levels_of_corruption	(base outcome)					
High_levels_of_corruption						
State_Authorities						
Few of them	1.962307	1.036219	1.28	0.202	0.697076	5.524001
Most of them	5.664561	3.168774	3.1	0.002	1.89235	16.9563
All of them	8.463173	8.598825	2.1	0.036	1.155282	61.99808
Business_executives						
Few of them	0.297751	0.17421	-2.07	0.038	0.094586	0.937299
Most of them	0.543457	0.321533	-1.03	0.303	0.170434	1.7329
All of them	0.658574	0.588372	-0.47	0.64	0.114324	3.793767
Civil_service						
Few of them	1.299647	0.322828	1.06	0.291	0.798712	2.11476
Most of them	2.012382	0.778845	1.81	0.071	0.942482	4.296828
All of them	1.291626	1.818802	0.18	0.856	0.081756	20.40587
Journalists_media						
Few of them	1.174558	0.501018	0.38	0.706	0.509081	2.709952
Most of them	1.28333	0.55945	0.57	0.567	0.546095	3.015837
All of them	1.374072	0.810059	0.54	0.59	0.432713	4.363347
people_pay_bribe						
Rarely	1.175505	0.213303	0.89	0.373	0.823698	1.67757
Frequently	1.252022	0.538511	0.52	0.601	0.538889	2.908873
Educational_level						
Secondary Education	0.598955	0.326426	-0.94	0.347	0.205823	1.742983
Higher Education	1.028088	0.740733	0.04	0.969	0.250463	4.220047
Worries_Losing_job						
A Good Deal	0.724106	0.282838	-0.83	0.409	0.33676	1.556982
Not Much	0.871657	0.299521	-0.4	0.689	0.444481	1.709376
Not at All	0.995145	0.326291	-0.01	0.988	0.52335	1.892261
Confidence_in_the_government						
Quite a Lot	0.287666	0.170092	-2.11	0.035	0.090279	0.916623
Not Very Much	0.886638	0.515576	-0.21	0.836	0.283646	2.771511
Not at All	1.426391	0.861731	0.59	0.557	0.436515	4.66099
justifiablepoliticalviolence						

Rarely justifiable	0.902257	0.444845	-0.21	0.835	0.343287	2.37139
Sometimes justifiable	0.811792	0.575067	-0.29	0.768	0.202517	3.254084
Often justifiable	9.10E-08	0.000374	0	0.997	0	0
Always justifiable	4.006873	5.064531	1.1	0.272	0.33645	47.71897
importanceofgod						
Moderately important	0.821002	0.16521	-0.98	0.327	0.553423	1.217957
Very important	0.799489	0.171313	-1.04	0.296	0.525314	1.216763
incomelevel						
Middle class	0.337327	0.210606	-1.74	0.082	0.099224	1.146802
Upper class	3.005007	2.920827	1.13	0.258	0.447187	20.19306
edu_level_2_inc_level_2	2.937788	2.004937	1.58	0.114	0.77109	11.19272
edu_level_2_inc_level_3	0.457336	0.521012	-0.69	0.492	0.049037	4.265308
edu_level_3_inc_level_2	1.596239	1.334811	0.56	0.576	0.309959	8.220379
edu_level_3_inc_level_3	0.198483	0.238136	-1.35	0.178	0.0189	2.084377
cons	0.666697	0.664266	-0.41	0.684	0.094587	4.699229

Interpretation of significant variables

For low levels of corruption

- **State Authorities**
  - e^{β} few of them = 0.371641: the estimated relative risk of Germans with low levels of perception of corruption who think that few of the state authorities are involved in corruption compared to those who think that none of the state authorities are involved in corruption is about 62.84 % less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.
  - e^{β} most of them = 0.148343: the estimated relative risk of Germans with low levels of perception of corruption who think that most of the state authorities are involved in corruption compared to those who think that none of the state authorities are involved in corruption is about 85.17% less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.
- **Business executives**
  - e^{β} few of them = 0.41426: the estimated relative risk of Germans with low levels of perception of corruption who think that few of the business executives are involved in corruption compared to those who think that none of the business executives are involved in corruption is about 58.57 % less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.
  - e^{β} most of them = 0.263763: the estimated relative risk of Germans with low levels of perception of corruption who think that most of the business executives

are involved in corruption compared to those who think that none of the business executives are involved in corruption is about 73.62% less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.

- **Confidence in the government**
  - e^{β} quite a lot = 0.25084: the estimated relative risk of Germans with low levels of perception of corruption whose confidence in the government is quite a lot compared to those whose confidence in the government is a great deal is about 74.91% less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.
  - e^{β} not very much = 0.186207: the estimated relative risk of Germans with low levels of perception of corruption whose confidence in the government is not very much compared to those whose confidence in the government is a great deal is about 81.38% less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.
  - e^{β} not at all = 0.309279: the estimated relative risk of Germans with low levels of perception of corruption whose confidence in the government is not at all compared to those whose confidence in the government is a great deal is about 69.07% less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.
- **Income level**
  - e^{β} middle class = 0.296991: the estimated relative risk of Germans with low levels of perception of corruption who belong in the middle class compared to those who belong in the lower class is about 70.3% less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.

For high levels of corruption

- **State Authorities**
  - e^{β} most of them = 5.664561: the estimated relative risk of Germans with high levels of perception of corruption who think that most of the state authorities are involved in corruption compared to those who think that none of the state authorities are involved in corruption is 5.664561 times the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.
  - e^{β} all of them = 8.463173: the estimated relative risk of Germans with high levels of perception of corruption who think that all of the state authorities are

involved in corruption compared to those who think that none of the state authorities are involved in corruption is 8.463173 times the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.

- **Business executives**
  - e^{β} few of them = 0.297751: the estimated relative risk of Germans with high levels of perception of corruption who think that few of the business executives are involved in corruption compared to those who think that none of the business executives are involved in corruption is about 70.22 % less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.

- **Confidence in the government**
  - e^{β} quite a lot = 0.287666: the estimated relative risk of Germans with high levels of perception of corruption whose confidence in the government is quite a lot compared to those whose confidence in the government is a great deal is about 71.23% less than the corresponding relative risk of having moderate levels of perception of corruption, holding all the other variables constant.

Table (7): Goodness-of-fit test for a multinomial logistic regression model
Dependent variable: Corruption_level

number of observations = 1182
number of outcome values = 3
base outcome value = 1
number of groups = 10
chi-squared statistic = 13.394
degrees of freedom = 16
Prob > chi-squared = 0.644

From the table, we can say that the model is good fit for the data(p-value>0.644)