Final Data Project - Data Management with R

Kris Best

15 December 2017

Introduction

This project seeks to identify factors that contributed to individuals being found guilty of and/or being executed for practicing witchcraft in early modern Scotland. It makes use of the *Scottish Witchcraft Database*, which was compiled by four researchers at the University of Edinburgh in 2003 and contains demographic and situational data on the cases of 3,219 individuals (both male and female) who were accused and tried for witchcraft in Scotland between 1563 and 1736. This paper uses logit and linear probability regression models to investigate the probability that an individual would be found guilty of and ultimately executed for witchcraft as a function of their sex, marital status and the specific motives for their accusation.

The public Github repository for this project can be found at: https://github.com/krisbest/hertie-dataproject

Data Preparation

Structure of the dataset

The *Scottish Witchcraft Database* is organised at three levels. The top *'accused'* level contains the characteristics of the individual accused persons and their family members, including brief demographic information where available. The second *'case'* level contains details regarding the case brought against the accused, e.g. the date and content of the accusations, number of accusers, and particular elements of witchcraft involved. The first and second level are linked by the key *accusedref*, which refers to the individual accused. The third *'trial'* level of the database contains information related to the trial process, including the proceedings and outcome. As an individual accused may have multiple trials, each observation at the third level is assigned a *trialref* and is linked to the second level of the database through the key *caseref*.

Preparing data and creating the database

The original data was provided in a zipped file that contained 38 separate tables in CSV format. These tables were first unzipped and read into R as dataframes. These 38 dataframes were reviewed, and any tables and columns that were not useful for the analysis were dropped from future steps. Three tables were used going forward that correspond to the 'overview' tables used at each level of the database: *accused, case,* and *trial*.

In addition to dropping unused tables and variables, some of the existing factor variables had to be recoded at this stage into simpler categories in order to make them more amenable for future stages of the analysis. For example, marital status and socioeconomic status were recoded from five and seven categories respectively into three categories each. More substantive recoding was undertaken for the indicator variables corresponding to the researchers' assessment of the *primary motive or characterisation* of the accusation of witchcraft.

Recoded motive	Description	Original motives (estimated by researchers)
Association	Individual was primarily accused for their association with another witch	Consulting a witch, being implicated by another accused witch

Recoded motive	Description	Original motives (estimated by researchers)
Evil Magic	Individual was primarily accused for engaging in 'evil' or malevolent forms of magic	Demonic elements, demonic possession, maleficium
Good Magic	Individual was primarily accused for engaging in 'good' forms of magic	Fairies, folk healing, white magic
Non-Magic	Individual was primarily accused for suspected non-magic-related reasons	Neighbourhood dispute, property motive, political motive, treason, refused charity

The three 'overview' tables were then imported into a SQLite database. The three tables were joined by a database query using the relevant keys at each level (*accusedref*,*caseref*) and then pulled into a set of tibbles for local use in R.

Descriptive Statistics

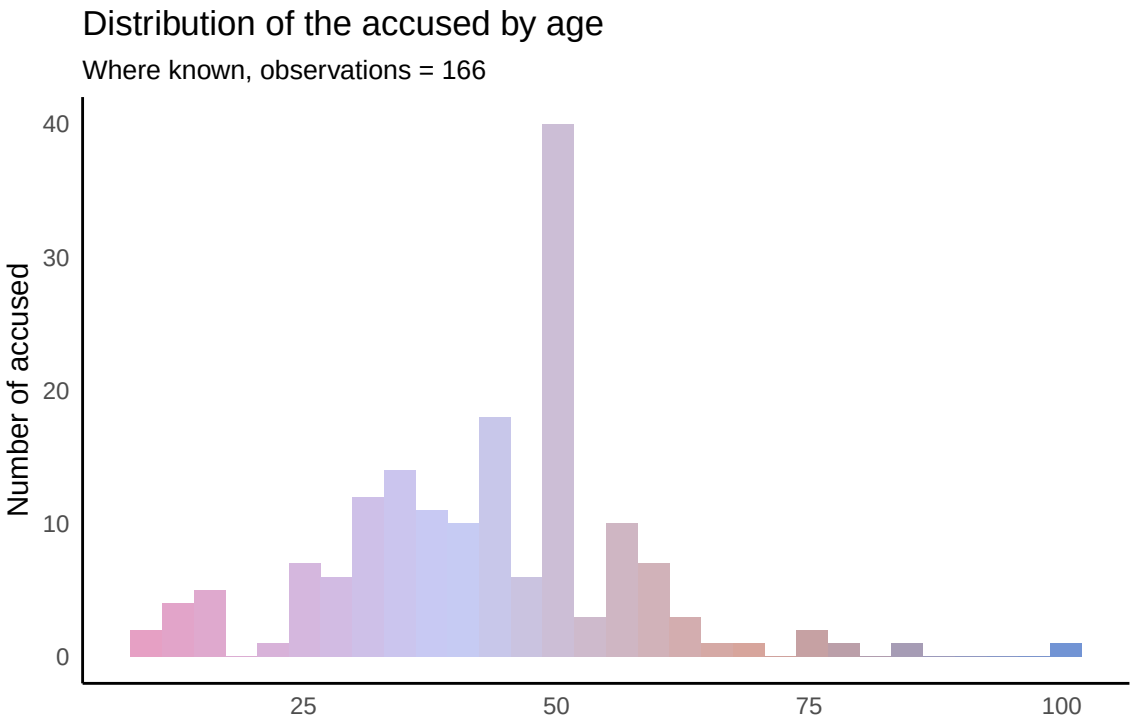
The joined dataset contained 3,594 observations on 3,219 individuals accused of witchcraft, with each observation corresponding to an accused individual and a trial. Some accused individuals were tried mutiple times and are therefore subject to multiple observations; in contrast, some accused individuals were not tried at all, and these observations relate only to the accusation. Note that dispute the high number of observations, only a small number (often a few hundred or less) actually contained detailed data on the variables of interest in this analysis.

Demographic statistics

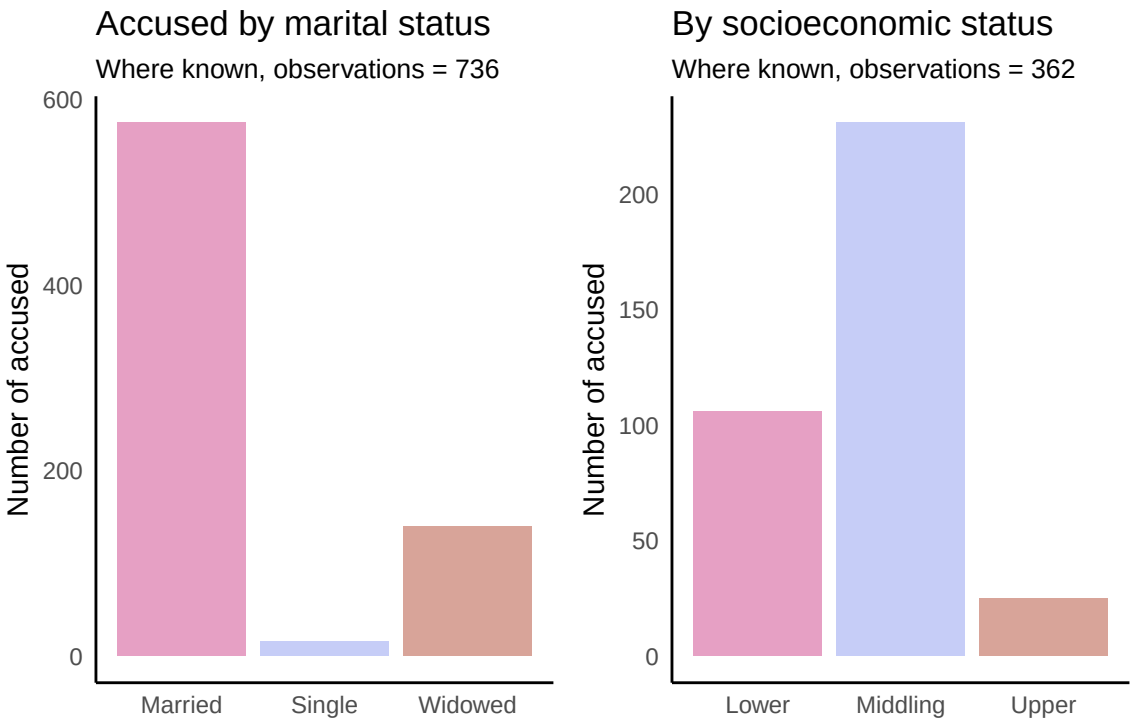
As expected from common stereotypes of witchhunts in early modern Europe, the gender balance of the accused individuals is highly skewed: 85% of the accused were female, with the counts by sex shown in the table below.

Female	Male
2702	468

Age statistics were available from only 166 of the accused individuals and ranged from a minimum of 9 to a maximum of 100, with a mean age of 43. As the following figure shows, witchcraft accusations primarily affected those in their mid-20s or older, with a distinct peak in middle age. Note that the spike at precisely age 50 reflects the fact that many of the age statistics were estimations on the part of the researchers.

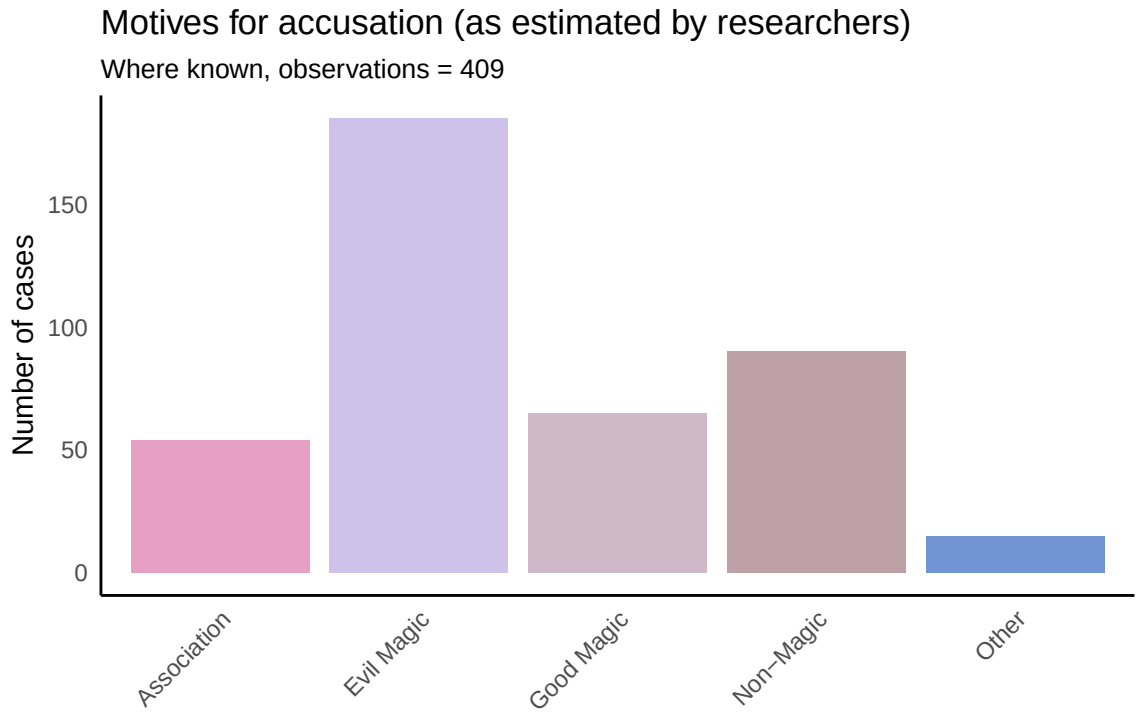


The following figures show the distribution of the accused individuals by marital status and socioeconomic status, where these statistics were available. Individuals accused of witchcraft were overwhelmingly married, with most hailing from the middle or lower classes.



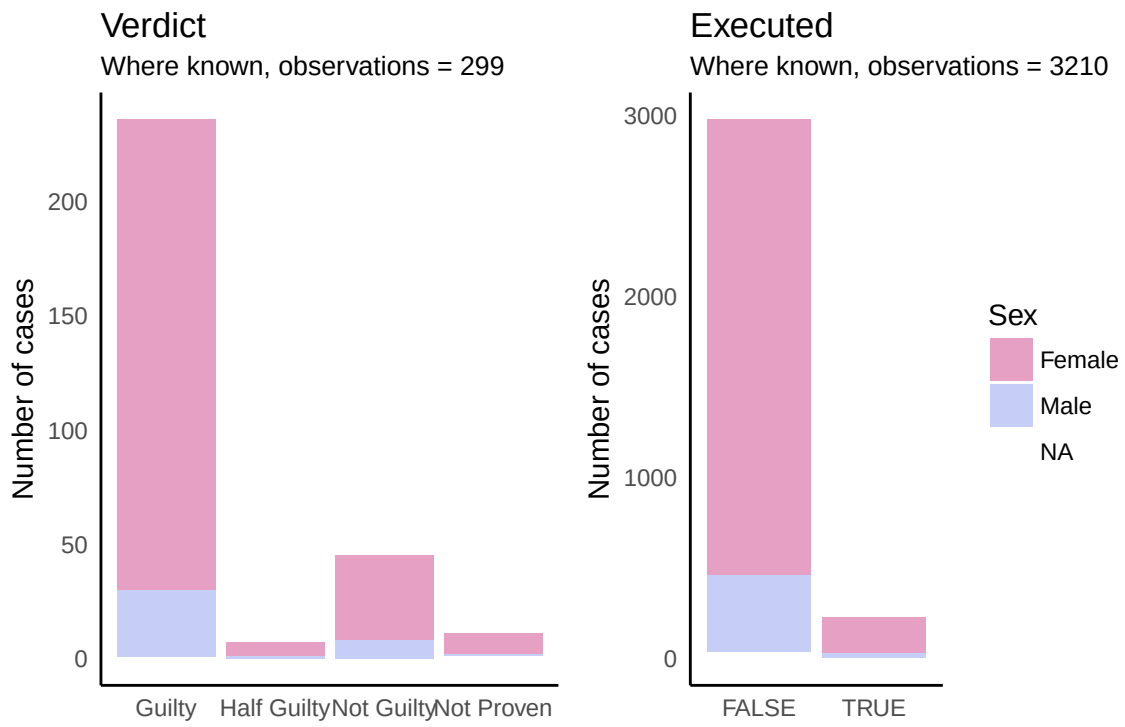
Descriptive statistics of the case and outcome

Individuals were accused of witchcraft for a variety of different reasons. The following figure shows the distribution of primary motives or characteristics of the accusations where this was known to or could be accurately estimated by the researchers (409 cases).



As the figure above indicates, the most common characteristics of a witchcraft accusation related to the practice of 'evil' magic, i.e. to demonic possession, demonic activity, or maleficium (malevolent sorcery). However, 90 accusations were categorised by the researchers as having non-magic-related motives. Of the non-magic-related accusations, the researchers estimated that 65 related to a neighbourhood dispute and 10 related to 'refused charity', e.g. that the individual refused to share their food, shelter or wealth. Among the few cases that listed 'other' motives, the researchers listed the commission of crimes (notably murder and theft) as well as 'sexual misconduct', incest and rape. In the majority of cases (more than 3000), the primary motivation of the accusation could not be determined.

The following figures show the ultimate verdict of the trial (where known) as well as whether or not the accused individual was executed as a result. Both figures are further broken down according to the sex of the accused individual.



The accused individual was found guilty in 236 out of the 299 cases where verdict data exists, i.e. nearly 80% of the time. A total of 230 accused individuals were executed out of 3210 accused for which execution data exists. Note that this does not mean that almost all of the 236 individuals found guilty were executed: in many cases the data indicated that an execution took place where there was no corresponding data for the verdict (i.e. the verdict appeared as a missing value).

Of the 230 individuals who were executed, details about the method of execution exist in 175 cases. The most common method of execution noted in 125 cases was 'strangle and burn', followed by burning alone in 46 cases. The remaining few execution cases are divided between beheadings (3) and hangings (1).

Model Specification

The regression analysis comprises two different regression models - a logistic model and a linear probability model (LPM) - each carried out for two different dependent variables for a total of four models. In both sets of models the dependent variable is a binary variable that relates to the outcome of the witchcraft case.

- The binary dependent variable *Guilty* is defined as 1 if the verdict was "Guilty" and 0 otherwise.
- The binary dependent variable *Executed* is defined as 1 if the individual was executed and 0 otherwise.

Although of course the coefficient interpretation and estimators vary significantly between the logistic and LPM models (with the logistic model using a maximum likelihood estimator and the LPM relying on ordinary least squares), the generic form of the model in both cases is the same and takes after the following:

$$Guilty = \alpha + \beta_1 female + \beta_2 maritalstatus + \beta_3 association + \beta_4 evilmagic + \beta_5 goodmagic + \beta_6 nonmagic + \epsilon$$

Where marital status is a factor variable (default category = married) and female, association, evilmagic, goodmagic, and nonmagic are all indicator variables that take on values of 1 or 0.

Various forms of the above model were tested and discarded due to lack of variation among the independent variables, including models that included other standard demographic variables such as age and socioeconomic

status. Although the joined data set is very large (containing 3,594 observations), the number of observations that contain data on the dependent variables are much smaller, with only 299 observations containing verdict data.[1] The data set also contains a very high number of missing values in variables related to demographic data or characteristics of the case. Very often, observations that contain data on the independent variables of interest do not contain verdict data and *vice versa*. Variables that were otherwise interesting but contained too many missing values (e.g. age, where 3,053 out of 3,219 observations at the accused level had missing values) were therefore dropped from the general model.

The problem of low variation among the dependent variables in the model also led to the recoding of several factor or indicator variables as described in the section on data preparation above, in an attempt to consolidate the data into fewer but larger categories for analysis (e.g. recoding the 12 original accusation motives into 4 overarching categories of association, evil magic, good magic, and non-magic; recoding marital status from five into three categories).

Model Results

Logistic model results

The table below displays the output from the two logistic regression models, with the coefficients presented in the form of odds ratios.

Table 3:

	Dependent variable:	
	Guilty	Executed
	(1)	(2)
Female	5.972** (2.673)	1.653 (1.665)
Single	0.238 (4.365)	1.094 (2.277)
Widowed	0.431 (1.740)	1.155 (1.342)
Association	0.362 (2.939)	2.408 (2.169)
Evil Magic	243,365,267.000 (Inf.000)	6.390*** (1.373)
Good Magic	136,051,281.000 (Inf.000)	5.911*** (1.728)
Non-Magic	0.760 (1.735)	9.447*** (1.399)
Constant	0.703 (2.510)	0.035 (1.653)
Observations	116	828
Log Likelihood	−48.507	−234.342
Akaike Inf. Crit.	113.013	484.685
Note:		*p<0.1; **p<0.05; ***p<0.01

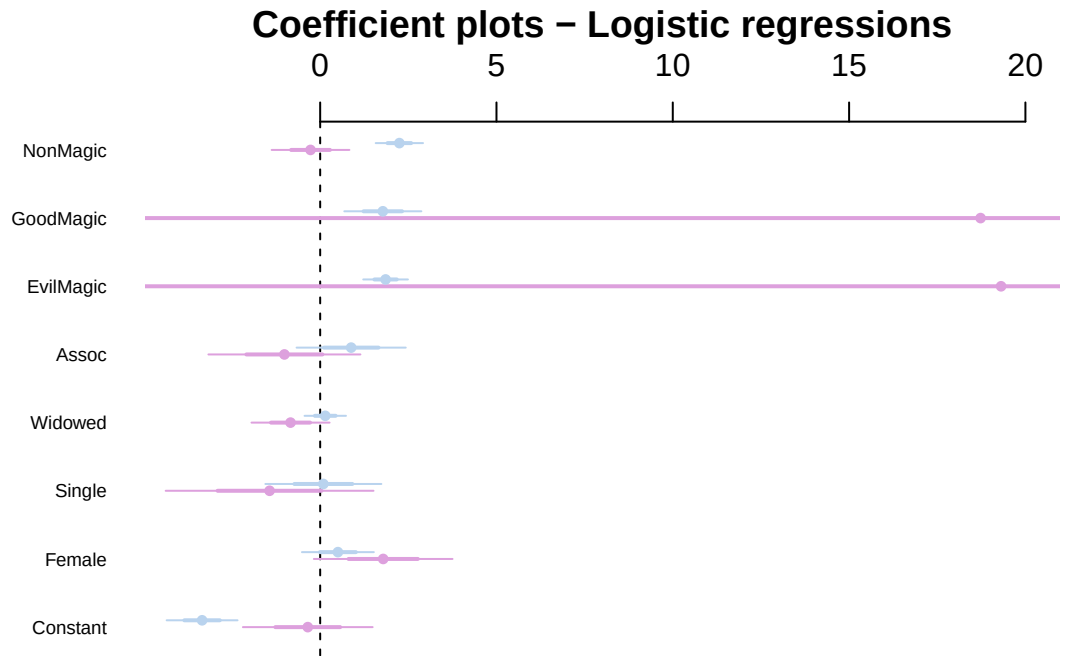
In the *Guilty* logistic model, only one coefficient, the female indicator variable, is found to be significant at the 5% level, suggesting that being a female accusee is associated with a nearly 6 times greater likelihood of being found guilty of witchcraft compared to their accused male counterparts. The evil magic and good magic coefficients took on rather extreme values in this model, but remain insignificant due to their even higher standard errors (too large to be displayed in the table).

In the *Executed* logistic model, the evil magic, good magic, and non-magic coefficients are all significantly associated with higher likelihoods of execution. In particular, being accused of witchcraft for a non-magic-related motive (such as a neighbourhood dispute or political motive) appears to be associated with a greater likelihood of execution than being accused for reasons related to evil magic or good magic. However, as the

[1] Although the number of observations with valid entries in the execution variable was much larger at 3210 observations, only 230 observations contain a "TRUE" value. Combined with the large number of missing values in the other variables described above, this means that there is little variation among "TRUE" cases for the regression analysis to be able to disentangle potential significant effects.

following plot of the logistic regression coefficients shows, the difference between the non-magic and good/evil magic coefficients does not appear to be significant.

Note that the coefficients in the plot below are shown in *log-odds*, not in odds ratios. The thicker lines represent the 50% confidence interval around the coefficient estimate while the thinner lines represent the 95% confidence interval. The *Guilty* model is shown in purple and the *Executed* model is shown in blue.



Linear probability model (LPM) results

The following table presents the regression output of the two linear probability models.

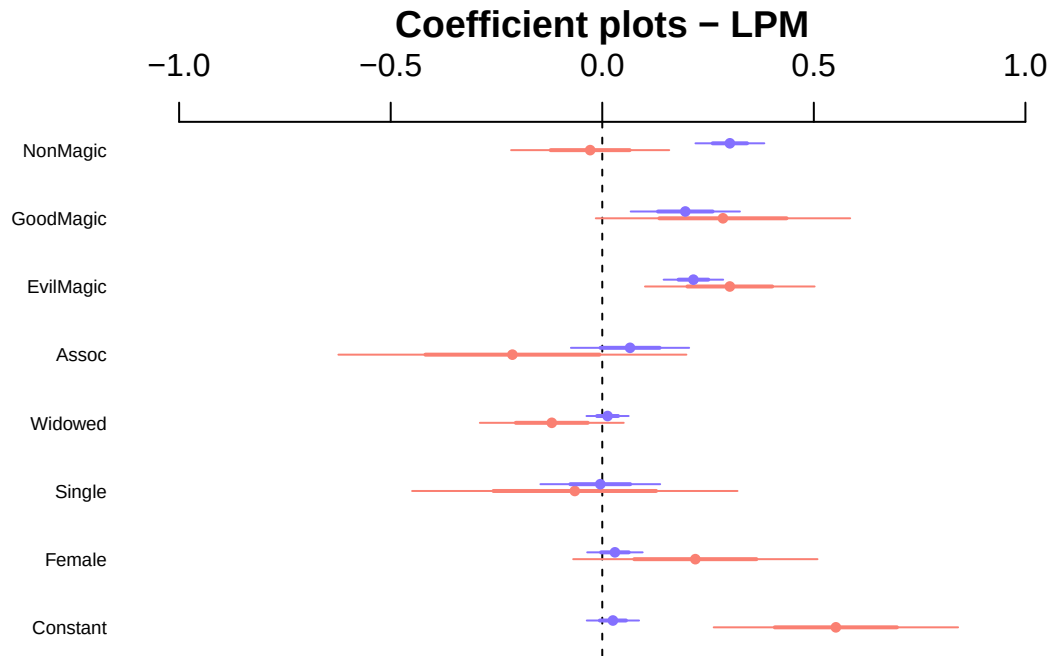
Table 4:

	Dependent variable:	
	Guilty	Executed
	(1)	(2)
Female	0.220 (0.144)	0.030 (0.033)
Single	−0.065 (0.192)	−0.005 (0.071)
Widowed	−0.119 (0.085)	0.013 (0.025)
Association	−0.212 (0.206)	0.066 (0.070)
Evil Magic	0.301*** (0.100)	0.215*** (0.035)
Good Magic	0.285* (0.150)	0.196*** (0.065)
Non-Magic	−0.029 (0.093)	0.302*** (0.041)
Constant	0.552*** (0.144)	0.025 (0.031)
Observations	116	828
R^2	0.141	0.103
Adjusted R^2	0.085	0.096
Residual Std. Error	0.395 (df = 108)	0.284 (df = 820)
F Statistic	2.530** (df = 7; 108)	13.510*** (df = 7; 820)
Note:		*p<0.1; **p<0.05; ***p<0.01

In the *Guilty* LPM, in contrast to the logistic regression, the female coefficient is no longer significant (although the effect remains positive). Significant and positive coefficients can now be found on the evil magic and good magic indicators. Interestingly, the magnitude of the estimated effects for good and evil magic are similiar, as both are associated with a roughly +0.30 increase in the probability of being found guilty of witchcraft.

In the *Executed* LPM, just as in the logistic regression, the coefficients on the evil magic, good magic, and non-magic indicators are all highly significant and positively associated with a higher probability of execution in the magnitude of about +0.20 (for evil and good magic) and +0.30 (for non-magic). This shows a similiar pattern to the results of the logistic model, in which non-magic motives for accusation were associated with a higher likelihood of execution than either evil magic or good magic. However, again similarly to the logistic model, the following coefficient plot indicates that this difference is potentially not significant (falling outside the 50% confidence intervals but within the 95% confidence intervals).

In the following figure, the orange lines represent the *Guilty* model and the blue lines represent the *Executed* model.



Summary

The results of the two *Guilty* models are inconclusive as to the factors that might lead an accused individual to be found guilty of witchcraft. While the logistic model suggests that being female might be positively associated with a higher likelihood of being found guilty of witchcraft with all else held equal, this result is not replicated in the parallel linear probability model.

However, the logistic and LPM models for the *Executed* variable show remarkable consistency, suggesting that being accused on the basis of evil magic, good magic or non-magic-related motives are all associated with a significantly higher probability of execution. An interesting result is that while it appears to make no difference whether one is accused on the basis of practicing evil magic versus 'good' magic, being accused on the basis of non-magic-related motives in both models seems to result in an even higher probability of execution. However, the greater probability of execution associated with being accused for non-magic-related reasons does not appear to be statistically significant at a 95% confidence level.

Data Source

Julian Goodare, Lauren Martin, Joyce Miller and Louise Yeoman, *'The Survey of Scottish Witchcraft'*, www.arts.ed.ac.uk/witches/ (archived January 2003, accessed 13 December 2017).