Assignment 2 - Obama & Happiness - Discrete Choice Modelling

Marc Sparhuber

Table of contents

Loading in Packages & Data

Preparing Data & Descriptives

Table 1: Descriptives split by voting for Obama

		voted Obama				didn't vote Obama				combined			
		Mean	SD	N	Percent	Mean	SD	N	Percent	Mean	SD	N	Percent
Happiness	Very Happy			210	15.78			183	13.75			393	29.53
	Pretty Happy			483	36.29			301	22.61			784	58.90
	Not Too Happy			109	8.19			45	3.38			154	11.57
Bachelor or Graduate degrees of parents	1 or more			271	20.36			160	12.02			431	32.38
	0			531	39.89			369	27.72			900	67.62
Sex	Male			318	23.89			255	19.16			573	43.05
	Female			484	36.36			274	20.59			758	56.95
Age		52.84	16.67	802	60.26	55.48	16.10	529	39.74	53.89	16.49	1331	100.00
All				802	60.26			529	39.74			1331	100.00

Comments: Data from the General Social Survey R package.

```
[1] "0" "1"
```

Model Estimaton & Odd Ratios

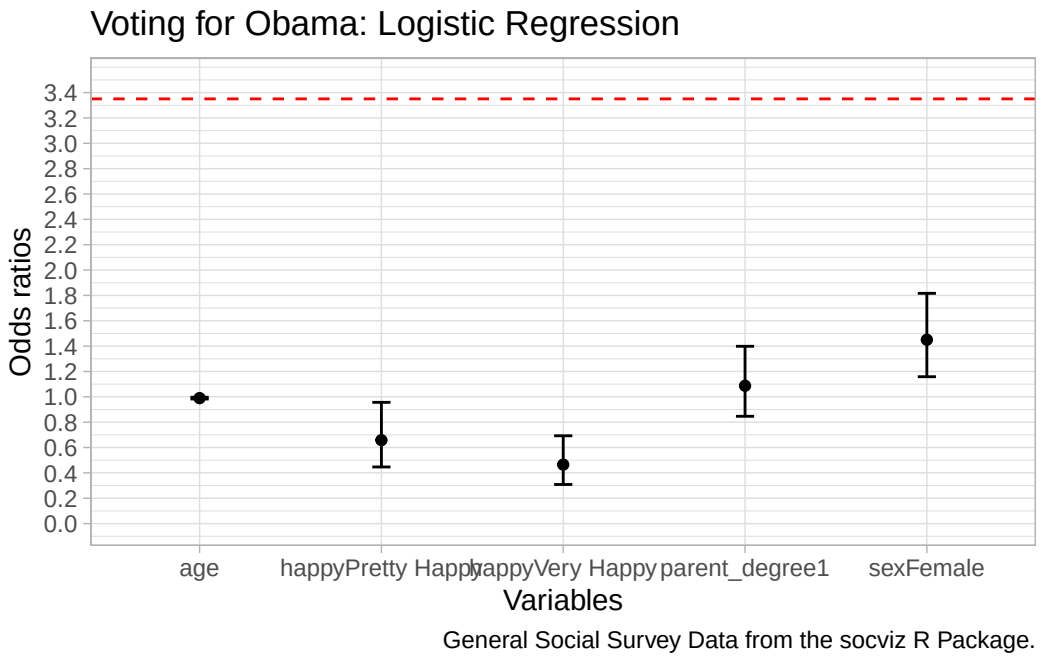
Table 2: Voting for Obama. Logistic probability models

	Model 1	Model 2	Model 3
(Intercept)	2.422***	2.316***	3.350***
	[1.724, 3.460]	[1.641, 3.322]	[1.922, 5.908]
happyVery Happy	0.474***	0.465***	0.465***
	[0.315, 0.703]	[0.309, 0.691]	[0.309, 0.693]
happyPretty Happy	0.662*	0.652*	0.658*
	[0.451, 0.959]	[0.444, 0.944]	[0.447, 0.956]
parent_degree1		1.207	1.087
		[0.952, 1.533]	[0.846, 1.399]
age			0.990**
			[0.983, 0.997]
sexFemale			1.450**
			[1.159, 1.816]
Num.Obs.	1331	1331	1331
AIC	1779.2	1778.8	1765.5
BIC	1794.8	1799.6	1796.7
Log.Lik.	−886.621	−885.416	−876.759

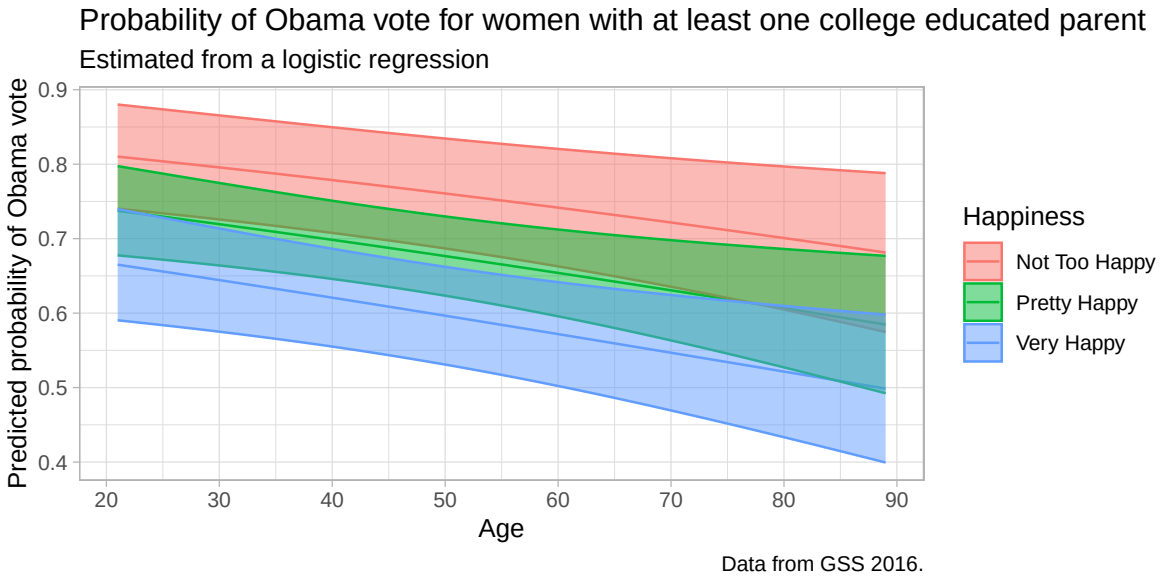
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Source: General Social Survey data from the socviz R package.

Comments: The reference cateogory for happy is 'Not Too Happy'.



Predicted Probabilities



Model Fit

Table 3: Voting for Obama: Model fit statistics

	Model 1	Model 2	Model 3
Nagelkerke's pseudo-R2	0.0156848	0.0181043	0.0353545
Share of correct predictions	0.6025545	0.6025545	0.6063110
Likelihood Ratio	NA	0.1205225	0.0001739

Note: A prediction is considered correct when its probability is greater than 0.5. The Likelihood ratio is always calculated with the model to the left. Data from GSS 2016.

Conclusion