

# Why do people want to get married ?

Colineaux Yvanne - Marine Dufau

ENAC

May 24, 2013

# Introduction

What are the individual characteristics that impact on their desire to get married?

# Composition of the database

## Answers from a survey

- ENAC students and workers (about 80%)
- Social network (Facebook)
- 1 week
- 423 observations

## First data treatment

- Desire to marry: 0 to 5
- People already married removed from data
- Geographical origins studied: Paris, Toulouse and Brittany (more representatives data)

# What would we expect?

<b>Explanatory variables</b>	<b>Expected effects</b>
Age	positive
Woman	positive
Level of study	positive
Homosexual	negative
Paris	negative
Toulouse	?
Brittany	positive
Separated parents	negative
Monoparental	negative
Siblings number	positive
Entourage already married	positive
Couple	positive
Believer	positive
Churchgoer	positive
Romantic	positive

# Linear regression

Using the least squares method, we have the following equation:

$$\begin{aligned} \widehat{MARRIAGE}_i = & \beta_0 + \beta_1 AGE_i + \beta_2 BELIEVER_i + \beta_3 BRITTANY_i \\ & + \beta_4 CHURCHGOER_i + \beta_5 COUPLE_i + \\ & + \beta_6 ENTOURAGE_i + \beta_7 HOMOSEXUAL_i \\ & + \beta_8 MONOPARENTAL_i + \beta_9 PARIS_i \\ & + \beta_{10} ROMANTIC_i + \beta_{11} SEPARATED_i + \beta_{12} SIBLINGS_i \\ & + \beta_{13} STUDY_i + \beta_{14} TOULOUSE_i + \beta_{15} WOMAN_i \quad (1) \end{aligned}$$

# EViews results

Dependent Variable: MARRIAGE  
 Method: Least Squares  
 Date: 05/17/13 Time: 17:45  
 Sample: 1 423  
 Included observations: 415

	Coefficient	Std. Error	t-Statistic	Prob.
C	2.203919	0.388388	5.674532	0.0000
AGE	-0.003899	0.007062	-0.552122	0.5812
BELIEVER	0.648757	0.148915	4.356589	0.0000
BRITTANY	-0.262278	0.195722	-1.340056	0.1810
CHURCHGOER	0.738282	0.215487	3.426106	0.0007
COUPLE	0.302487	0.129399	2.337635	0.0199
ENTOURAGE_MARRIED	0.273410	0.156767	1.744054	0.0819
HOMOSEXUAL	-0.667152	0.318229	-2.096453	0.0367
MONOPARENTAL	-0.426283	0.382754	-1.113727	0.2661
PARIS	0.108915	0.170948	0.637122	0.5244
ROMANTIC	0.367382	0.202086	1.817945	0.0698
SEPARATED_PARENTS	-0.190591	0.218690	-0.871513	0.3840
SIBLINGS	-0.045246	0.047594	-0.950660	0.3424
STUDY_LEVEL	0.706204	0.280786	2.515100	0.0123
TOULOUSE	-0.141496	0.185697	-0.761973	0.4465
WOMAN	0.002328	0.145241	0.016031	0.9872
R-squared	0.192179	Mean dependent var	3.248193	
Adjusted R-squared	0.161810	S.D. dependent var	1.365054	
S.E. of regression	1.249744	Akaike info criterion	3.321546	
Sum squared resid	623.1821	Schwarz criterion	3.476853	
Log likelihood	-673.2207	Hannan-Quinn criter.	3.382960	
F-statistic	6.328101	Durbin-Watson stat	1.922806	
Prob(F-statistic)	0.000000			

Figure : First model results

## First equation

With the estimated values, the model is:

$$\begin{aligned} \widehat{MARRIAGE}_i = & 2.20 - 0.004AGE_i + 0.649BELIEVER_i \\ & - 0.262BRITTANY_i + 0.738CHURCHGOER_i \\ & + 0.302COUPLE_i + 0.273ENTOURAGE_i \\ & - 0.667HOMOSEXUAL_i - 0.426MONOPARENTAL_i \\ & + 0.109PARIS_i + 0.367ROMANTIC_i \\ & - 0.191SEPARATED_i - 0.045SIBLINGS_i \\ & + 0.706STUDY_i - 0.141TOULOUSE_i \\ & + 0.002WOMAN_i \end{aligned} \quad (2)$$

# Interpretation

Explanatory variables	Expected effects	Real Effect
Age	+	-
Woman	+	+
Level of study	+	+
Homosexual	-	-
Paris	-	+
Toulouse	?	-
Brittany	+	-
Separated parents	-	-
Monoparental	-	-
Siblings number	+	-
Entourage already married	+	+
Couple	+	+
Churchgoer	+	+
Romantic	+	+



## Validity of the model

### R-squared value

- R-squared measures the proportion of total variation of outcomes carried by the explanatory variables
- $R=0.19$
- So we have a weak model

### Significant variables

- t-statistic tests at the 10% level
- The following parameters are relevant: BELIEVER, CHURCHGOER, COUPLE, ENTOURAGE, HOMOSEXUAL, ROMANTIC, STUDY

# Improving the model

## Correlation

- Objective: see if there are redundant explanatory variables in the model
- Correlation matrix

# Improving the model

	AGE	BELIEVER	BRITTANY	CHURCHGOER	COUPLE	ENTOURAGE	HOMOSEXUAL	MONOPARENTAL	PARIS	ROMANTIC	SEPARATED	SIBLINGS	STUDY	TOULOUSE	WOMAN
AGE	1.00000	0.063374	-0.064213	0.023905	-0.060695	-0.033228	-0.013837	0.062160	-0.018660	-0.016950	0.042446	0.013473	-0.037468	-0.061826	0.024428
BELIEVER	0.063374	1.00000	0.011617	0.485434	-0.171789	0.089941	-0.062052	-0.059993	0.028411	0.031785	-0.105000	0.189673	-0.124633	-0.037522	-0.204016
BRITTANY	-0.064213	0.011617	1.00000	-0.165359	-0.021242	-0.069739	0.021438	-0.025037	-0.180684	0.082252	-0.094836	0.002291	-0.066111	-0.162253	0.112850
CHURCHGOER	0.023905	0.485434	-0.165359	1.00000	-0.180150	0.246468	-0.024149	-0.032928	0.104379	-0.017954	-0.130000	0.120927	-0.019798	-0.042864	-0.333302
COUPLE	-0.060695	-0.171789	-0.021242	-0.180150	1.00000	-0.055831	-0.015071	0.114469	-0.001240	0.038354	0.003912	0.048100	-0.022133	-0.000167	0.054947
ENTOURAGE	-0.033228	0.089941	-0.069739	0.246468	-0.055831	1.00000	0.058697	-0.056696	-0.030703	-0.034079	-0.097510	0.103983	-0.056657	0.092500	-0.081081
HOMOSEXUAL	-0.013837	-0.062052	0.021438	-0.024149	-0.015071	0.058697	1.00000	-0.036743	0.056674	0.039880	0.259882	-0.089890	0.050380	-0.088390	0.002528
MONOPARENTAL	0.062160	-0.059993	-0.025037	-0.032928	0.114469	-0.056696	-0.036743	1.00000	-0.041955	-0.013930	-0.061698	0.012673	-0.023352	-0.031021	0.195245
PARIS	-0.018660	0.028411	-0.180684	0.104379	-0.001240	-0.030703	0.056674	-0.041955	1.00000	-0.079702	0.114485	0.038838	-0.024571	-0.191794	-0.102220
ROMANTIC	-0.016950	0.031785	0.082252	-0.017954	0.038354	-0.034079	0.039880	-0.013930	-0.079702	1.00000	-0.002357	-0.020476	-0.055842	-0.013448	0.144702
SEPARATED	0.042446	-0.105000	-0.094836	-0.130000	0.003912	-0.097510	0.259882	-0.061698	0.114485	-0.002357	1.00000	-0.098110	0.016982	0.001986	0.287084
SIBLINGS	0.013473	0.189673	0.002291	0.120927	0.048100	0.103983	-0.089890	0.012673	0.038636	-0.020476	-0.098110	1.00000	0.022598	-0.106102	-0.014510
STUDY	-0.037468	-0.124633	-0.066111	-0.019798	-0.022133	-0.056657	0.050380	-0.023352	0.024571	-0.055842	0.016982	0.022598	1.00000	-0.084017	-0.072645
TOULOUSE	-0.061826	-0.037522	-0.162253	-0.042864	-0.000167	0.092500	-0.088390	-0.031021	-0.191794	-0.013448	0.001986	-0.106102	-0.084017	1.00000	0.098403
WOMAN	0.024428	-0.204016	0.112850	-0.333302	0.054947	-0.081081	0.002528	0.195245	-0.102220	0.144702	0.287084	-0.014510	-0.072645	0.098403	1.00000

Figure : Correlation Matrix

# Improving the model

## Wald test

- Using Fisher-statistics
- Objective: joint hypothesis testing

$$H_0 : \beta_1 = 0, \beta_3 = 0, \beta_8 = 0, \beta_9 = 0, \beta_{11} = 0, \beta_{12} = 0, \beta_{14} = 0$$

Wald Test:  
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	0.717304	(8, 399)	0.6763
Chi-square	5.738429	8	0.6765

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(4)	-0.262278	0.195722
C(9)	-0.426283	0.382754
C(10)	0.108915	0.170948
C(2)	-0.003899	0.007062
C(12)	-0.190591	0.218690
C(13)	-0.045246	0.047594
C(15)	-0.141496	0.185697
C(16)	0.002328	0.145241

Restrictions are linear in coefficients.

Figure : Wald test

## First model conclusion

We kept the following parameters :

- Believer
- Going to church
- In couple
- Number of friends already married
- Homosexual
- Romantic movie lover
- Level of study

## EViews results

Dependent Variable: MARRIAGE  
 Method: Least Squares  
 Date: 05/17/13 Time: 19:07  
 Sample: 1 423  
 Included observations: 423

	Coefficient	Std. Error	t-Statistic	Prob.
C	1.943922	0.301566	6.446096	0.0000
BELIEVER	0.579761	0.144917	4.000640	0.0001
CHURCHGOER	0.824766	0.203958	4.043807	0.0001
COUPLE	0.290266	0.126731	2.290412	0.0225
ENTOURAGE_MARRIED	0.265210	0.152578	1.738193	0.0829
HOMOSEXUAL	-0.705359	0.303808	-2.321726	0.0207
ROMANTIC	0.329506	0.198683	1.658450	0.0980
STUDY_LEVEL	0.759727	0.278388	2.729020	0.0066
R-squared	0.170413	Mean dependent var		3.262411
Adjusted R-squared	0.156420	S.D. dependent var		1.364645
S.E. of regression	1.253380	Akaike info criterion		3.308297
Sum squared resid	651.9495	Schwarz criterion		3.384843
Log likelihood	-691.7048	Hannan-Quinn criter.		3.338543
F-statistic	12.17842	Durbin-Watson stat		1.886049
Prob(F-statistic)	0.000000			

Figure : Eview results for the second equation

## Second equation

$$\widehat{MARRIAGE}_i = 1.944 + 0.58BELIEVER_i + 0.82CHURCHGOER_i + 0.29COUPLE_i + 0.26ENTOURAGE_i - 0.705HOMOSEXUAL_i + 0.33ROMANTIC_i + 0.763STUDY_i \quad (3)$$

# Interpretation

Not surprising! The same type of effects than with the first model.



# Validity of the model

## R-squared

- We have a lower R-squared than before ( $R^2=0.17$ )
- Some significant parameters may be missing

## Improving the model

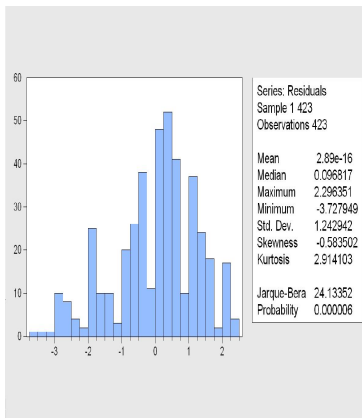


Figure : Residual weights versus normal distribution

## Improving the model

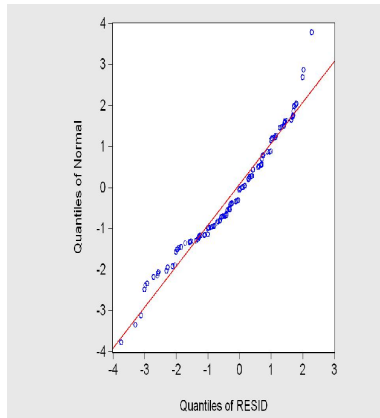


Figure : Quantile-quantile plot of the error against the normal distribution

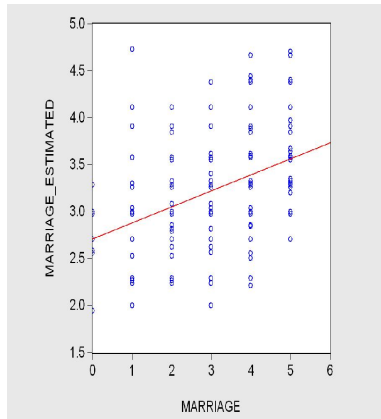


Figure : Estimated model versus real observations

## Second model conclusion

Why do people want to get married ?

- Religion, high studies
- Romantic feelings, suggestion in a couple, entourage,
- Things could change for homosexual population

## Second model conclusion

Limits of the model:

- More explanatory variables needed
- Nature of the survey: heterogeneity? level of reflexion?  
Self-evaluation? Change of opinion?
- Nature of the topic

# Conclusion

Thank you for your attention!

