

Econometrics: Divorce rate

François Dubuisson & Laure Truchot

ENAC

May 9, 2012

Introduction

Which parameters can influence the divorce rate?

1 Introduction

2 First equation

- Parameters
- First equation

3 A better modelization

- Equations 1A & 1B
- Equation 2

4 Tests and interpretation

- Heteroscedasticity and normal tests
- Divorce rate modelization

5 Conclusion

First model: parameters

Dependent variable:

- Divorce rate 2010

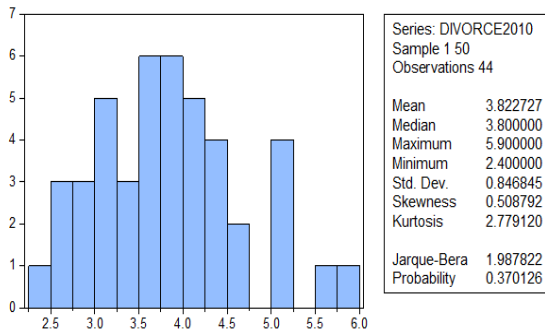


Figure : Divorce rate histogram

Maximum number of explanatory variables

- Education rate
- Median family income (US \$)
- % of Individuals in Poverty
- % of women in the Labor Force
- % living in Metropolitan areas
- % of Christian and Jewish Adherents in the Population
- Number of years since the non-fault law was enacted
- % Minority
- Crime Rate
- % Dependence on or Abuse of Illicit Drugs or Alcohol

Expected effect on the divorce rate

Independent Variables	Expected effect on the divorce rate
EDUCATION	Negative
INCOME	Positive
POVERTY	Ambiguous
LABOR	Positive
LAW	Negative
METROPOLITAN	Positive
RELIGIOUS	Negative
MINORITY	Ambiguous
CRIME	Positive
ILLICIT	Positive

First equation

Divorce rate function in 2010 using ordinary least squares (EQ01)

$$\begin{aligned} \text{DIVORCE}_i = & \beta_0 + \beta_1 \text{EDUCATION}_i + \beta_2 \text{CRIME}_i + \beta_3 \text{ILLICIT}_i + \beta_4 \text{INCOME}_i \\ & + \beta_5 \text{LABOR}_i + \beta_6 \text{LAW}_i + \beta_7 \text{METROPOLITAN}_i + \beta_8 \text{MINORITY}_i \\ & + \beta_9 \text{POVERTY}_i + \beta_{10} \text{RELIGIOUS}_i \end{aligned} \quad (1)$$

First equation

Linear regression output (EQ01)

Dependent Variable: DIVORCE2010

Method: Least Squares

Date: 05/05/12 Time: 17:29

Sample: 150

Included observations: 42

	Coefficient	Std. Error	t-Statistic	Prob.
C	39.62674	42.62191	0.929727	0.3597
CRIME	-0.000889	0.000991	-0.897014	0.3766
EDUCATION	-0.088474	0.046864	-1.887902	0.0684
ILLICIT	0.106130	0.168467	0.629976	0.5333
INCOME	-0.012297	0.040013	-0.307311	0.7607
LABOUR	-0.059963	0.051475	-1.164895	0.2529
LAW	-0.013635	0.021056	-0.647557	0.5220
METROPOLITAN	-0.009147	0.009543	-0.958453	0.3453
MINORITY	0.013081	0.013834	0.945606	0.3517
POVERTY	-0.098861	0.101725	-0.971843	0.3387
RELIGIOUS	-0.024079	0.011167	-2.156277	0.0389
R-squared	0.484561	Mean dependent var		3.873810
Adjusted R-squared	0.318291	S.D. dependent var		0.832000
S.E. of regression	0.686947	Akaike info criterion		2.307007
Sum squared resid	14.62877	Schwarz criterion		2.762111
Log likelihood	-37.44715	Hannan-Quinn criter.		2.473821
F-statistic	2.914292	Durbin-Watson stat		2.150593
Prob(F-statistic)	0.010842			

Figure : Equation 1

First equation

Linear regression output (EQ01)

$$\widehat{DIVORCE}_i = 39.63 - 8.85 \cdot 10^{-2} EDUCATION_i - 8.89 \cdot 10^{-4} CRIME_i \\ + 1.06 \cdot 10^{-1} ILLICIT_i - 1.23 \cdot 10^{-2} INCOME_i - 6 \cdot 10^{-2} LABOR_i \\ - 1.36 \cdot 10^{-2} LAW_i - 9.1 \cdot 10^{-3} METROPOLITAN_i + 1.31 \cdot 10^{-2} MINORITY_i \\ - 9.89 \cdot 10^{-2} POVERTY_i - 2.41 \cdot 10^{-2} RELIGIOUS_i \quad (2)$$

Comments on coefficient signs: unexpected results

- When INCOME ↗, DIVORCE ↘
- When POVERTY ↗, DIVORCE ↘
- When MINORITY ↗, DIVORCE ↗
- When LABOR ↗, DIVORCE ↘
- When METROPOLITAN ↗, DIVORCE ↘
- No impact of CRIME

First equation

	CRIME	DIVORCE2 010	EDUCATION	ILLICIT	INCOME	LABOUR	LAW	METROPOLITAN	MINORITY	POVERTY	RELIGIOUS
CRIME	1.000000	0.115385	-0.203508	-0.253542	-0.057040	-0.450333	-0.279061	0.382744	0.683167	0.233924	-0.183815
DIVORCE2 010		1.000000	-0.530949	-0.088518	-0.436373	-0.354821	-0.052127	-0.306412	-0.005189	0.320173	-0.370788
EDUCATION			1.000000	0.210066	0.834498	0.526983	-0.184013	0.422042	0.073228	-0.687504	-0.024796
ILLICIT				1.000000	0.086209	0.609141	0.133210	-0.265020	-0.298826	-0.330638	0.085483
INCOME					1.000000	0.526090	-0.242428	0.506121	0.097806	-0.846879	-0.095582
LABOUR						1.000000	-0.044961	-0.177598	-0.406863	-0.722750	0.163196
LAW							1.000000	-0.107651	-0.288307	0.090866	0.278248
METROPOLITAN								1.000000	0.542832	-0.224269	-0.050688
MINORITY									1.000000	0.236999	-0.080757
POVERTY										1.000000	0.092261
RELIGIOUS											1.000000

Figure : Correlation Matrix

First equation

The correlation between some independent variables is very high :

- Income \Leftrightarrow Education : 0.83
- Income \Leftrightarrow Poverty : -0.85
- Labour \Leftrightarrow Poverty : -0.72
- Crime \Leftrightarrow Minority : 0.68
- Poverty \Leftrightarrow Education : -0.69

First equation

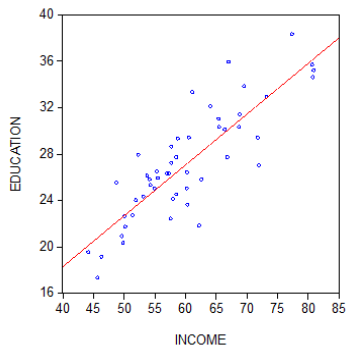


Figure : Education against income scatter plot

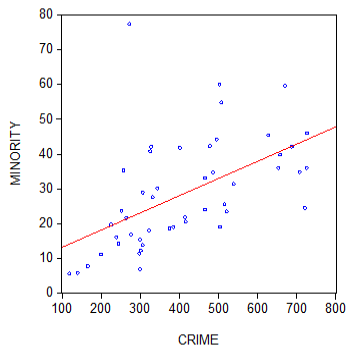


Figure : Minority against crime scatter plot

Equations 1A & 1B

Dependent Variable: DIVORCE2010
 Method: Least Squares
 Date: 05/07/12 Time: 16:25
 Sample: 1 50
 Included observations: 43

	Coefficient	Std. Error	t-Statistic	Prob.
C	36.37129	37.23794	0.976727	0.3354
EDUCATION	-0.075585	0.033308	-2.269290	0.0295
ILLICIT	0.086637	0.151683	0.571174	0.5715
LABOR	-0.028111	0.039736	-0.707437	0.4840
LAW	-0.014069	0.018815	-0.747728	0.4596
METROPOLITAN	-0.007974	0.008177	-0.975196	0.3362
MINORITY	0.001035	0.010539	0.098207	0.9223
RELIGIOUS	-0.026556	0.010302	-2.577783	0.0143
R-squared	0.469659	Mean dependent var		3.844186
Adjusted R-squared	0.363591	S.D. dependent var		0.844676
S.E. of regression	0.673842	Akaike info criterion		2.214600
Sum squared resid	15.89222	Schwarz criterion		2.542265
Log likelihood	-39.61389	Hannan-Quinn criter.		2.335432
F-statistic	4.427898	Durbin-Watson stat		2.127952
Prob(F-statistic)	0.001303			

Result

- Adjusted $R^2 = 0.364$

Figure : EQ1A regression output

Equations 1A & 1B

EQUATION 1-B

Dependent Variable: DIVORCE2010
 Method: Least Squares
 Date: 05/07/12 Time: 16:26
 Sample: 1 50
 Included observations: 42

	Coefficient	Std. Error	t-Statistic	Prob.
C	34.11990	39.52806	0.863182	0.3941
CRIME	-0.000371	0.000807	-0.459951	0.6485
EDUCATION	-0.076948	0.034919	-2.203639	0.0344
ILLICIT	0.102277	0.152543	0.670477	0.5071
LABOR	-0.034742	0.038120	-0.911390	0.3685
LAW	-0.012769	0.019944	-0.640238	0.5263
METROPOLITAN	-0.005466	0.008260	-0.661815	0.5126
RELIGIOUS	-0.026817	0.010236	-2.619948	0.0130
R-squared	0.458542	Mean dependent var		3.873810
Adjusted R-squared	0.347065	S.D. dependent var		0.832000
S.E. of regression	0.672293	Akaike info criterion		2.213397
Sum squared resid	15.36723	Schwarz criterion		2.544382
Log likelihood	-38.48134	Hannan-Quinn criter.		2.334716
F-statistic	4.113342	Durbin-Watson stat		2.028015
Prob(F-statistic)	0.002270			

Result

- Adjusted $R^2 = 0.347$
 \Rightarrow We keep EQ1A

Figure : EQ1B regression output

Equation 1A

Wald Test:
Equation: EQ01A

Test Statistic	Value	df	Probability
F-statistic	0.009645	(1, 35)	0.9223
Chi-square	0.009645	1	0.9218

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(7)	0.001035	0.010539

Restrictions are linear in coefficients.

Interpretation

- High P-values \Rightarrow Joint tests
- We drop MINORITY

Figure : Joint test for equation 1A

Last equation

Linear regression output (EQ02)

Dependent Variable: DIVORCE2010

Method: Least Squares

Date: 05/07/12 Time: 16:38

Sample: 1 50

Included observations: 43

	Coefficient	Std. Error	t-Statistic	Prob.
C	37.58120	34.65414	1.084465	0.2854
EDUCATION	-0.075400	0.032794	-2.299211	0.0274
ILLCIT	0.088421	0.148507	0.595397	0.5553
LABOR	-0.029491	0.036652	-0.804627	0.4263
LAW	-0.014653	0.017604	-0.832350	0.4107
METROPOLITAN	-0.007629	0.007279	-1.048089	0.3016
RELIGIOUS	-0.026445	0.010099	-2.618727	0.0128
R-squared	0.469513	Mean dependent var		3.844186
Adjusted R-squared	0.381098	S.D. dependent var		0.844676
S.E. of regression	0.664509	Akaike info criterion		2.168364
Sum squared resid	15.89660	Schwarz criterion		2.455071
Log likelihood	-39.61982	Hannan-Quinn criter.		2.274092
F-statistic	5.310361	Durbin-Watson stat		2.117654
Prob(F-statistic)	0.000522			

Figure : Equation 2

Tests

Heteroskedasticity Test: White

F-statistic	1.553047	Prob. F(26,16)	0.1809
Obs*R-squared	30.79692	Prob. Chi-Square(26)	0.2359
Scaled explained SS	20.06225	Prob. Chi-Square(26)	0.7886

Figure : Test for heteroskedasticity

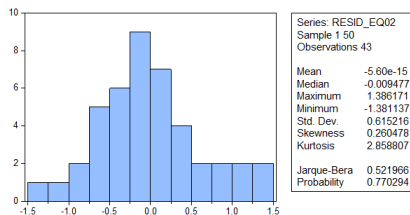
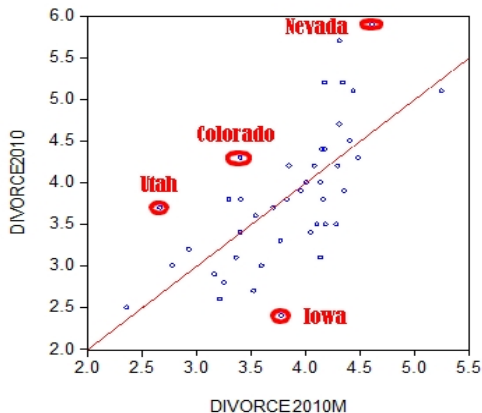


Figure : Normal and Exogeneity test

Modelization against real results



Interpretation

- Quite good model
 - overestimation for Iowa
 - underestimation for Utah, Colorado and Nevada
- ⇒ Limits for our model

Figure : Divorce rate model versus real results

Equation 2 output

$$\begin{aligned}\widehat{DIVORCE}_i = & 37.58 - 7.54 \cdot 10^{-2} EDUCATION_i + 8.84 \cdot 10^{-2} ILLICIT_i - 2.95 \cdot 10^{-2} LABOR_i \\ & - 1.47 \cdot 10^{-2} LAW_i - 7.63 \cdot 10^{-3} METROPOLITAN_i \\ & - 2.64 \cdot 10^{-2} RELIGIOUS_i\end{aligned}\quad (3)$$

Questions ?