Fertility rate

Fransisco Sanchez & Pol Queffeulou

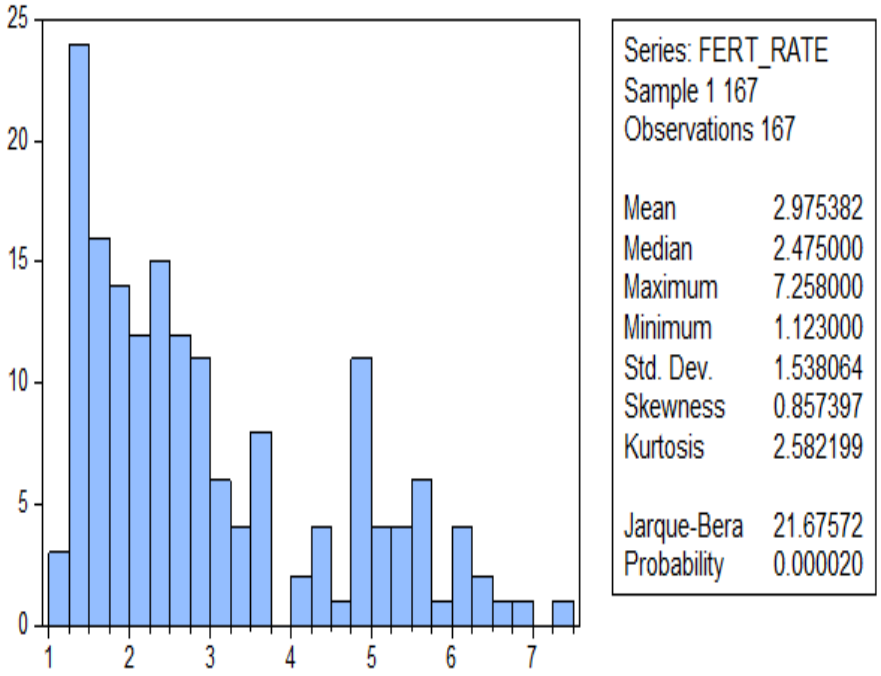
ENAC

May 11, 2012

Summary

- Dependent and explanatory variables
- First equation (EQ01)
- Second equation (EQ02)
- Third equation (EQ03)

Dependent variable



- Around the median : European countries
- 4 to 7.3 : African countries

FIGURE : Fertility rate histogram

Explanatory variables

- Contraceptive Prevalence (% of women ages 15-49)
- GDP (current US$)
- Pregnant women receiving prenatal care (%)
- Births attended by skilled health staff (% of total)
- Prevalence of HIV, total (% of population ages 15-49)
- Antiretroviral therapy coverage (% of people with advanced HIV infection)
- Neonatal mortality rate
- Lifetime risk of maternal death
- Population Growth (%annual)

Explanatory variables

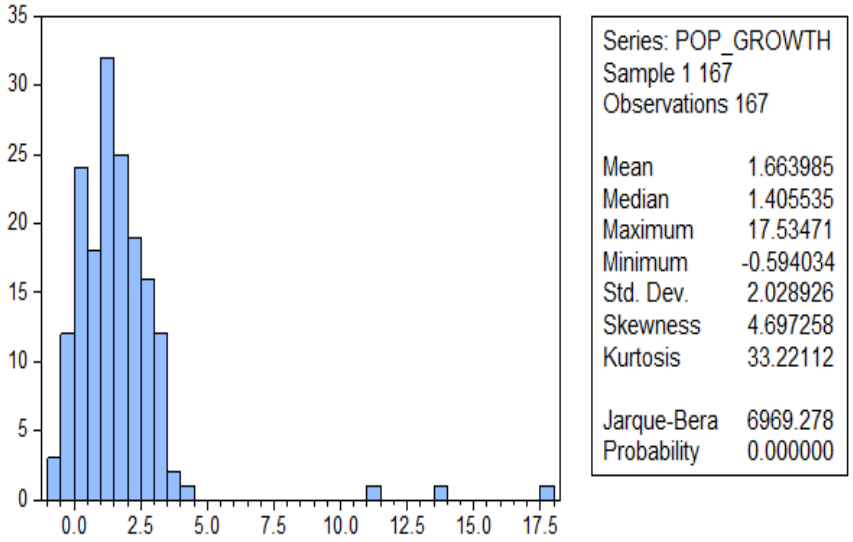


FIGURE : Population growth histogram

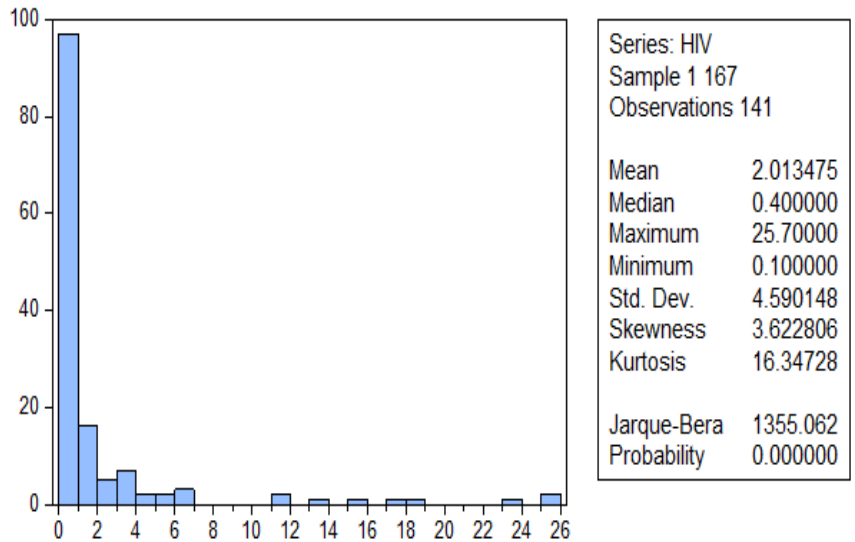


FIGURE : Prevalence of HIV histogram

- Negative values : Eastern Europe countries
- Highest values : Qatar, Bahrain, Arab Emirates

- 25.7 : Swaziland
- 24 : Botswana

Regression linear output (eq01)

Dependent Variable: FERT_RATE
Method: Least Squares
Date: 05/08/12 Time: 18:24
Sample: 1 62
Included observations: 62

	Coefficient	Std. Error	t-Statistic	Prob.
C	1.316597	0.732040	1.798532	0.0779
GDP	-1.40E-14	4.74E-13	-0.029610	0.9765
HIV	0.025063	0.014999	1.670976	0.1007
MATERNAL_DEATH	0.345192	0.058702	5.880379	0.0000
NEO_MORT_RA	0.023136	0.012523	1.847465	0.0704
POP_GROWTH	0.476817	0.083973	5.678187	0.0000
PRENATAL_CARE	0.007713	0.006121	1.260064	0.2133
CONTRACEPT	-0.008318	0.005418	-1.535181	0.1308
THERAPY_HIV	-0.000687	0.003281	-0.209256	0.8351
HEALTH_STAFF	-0.001187	0.004877	-0.243478	0.8086

R-squared	0.923065	Mean dependent var	3.490968
Adjusted R-squared	0.909749	S.D. dependent var	1.520468
S.E. of regression	0.456776	Akaike info criterion	1.417442
Sum squared resid	10.84950	Schwarz criterion	1.760528
Log likelihood	-33.94071	Hannan-Quinn criter.	1.552146
F-statistic	69.32144	Durbin-Watson stat	2.531613
Prob(F-statistic)	0.000000		

FIGURE : eq01 : Regression linear output

p-values

- some are really high : GDP (0.98), Therapy_HIV (0.83) and Health_Staff (0.81)
- Another one is quite significant : Prenatal_Care(0.21)

$\Rightarrow$ Joint test

Joint test

Wald Test:
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	0.563802	(4, 52)	0.6900
Chi-square	2.255209	4	0.6889

FIGURE : Joint test
c(2)=c(7)=c(9)=c(10)=0

Wald Test:
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	0.032349	(3, 52)	0.9921
Chi-square	0.097047	3	0.9922

FIGURE : Joint test
c(2)=c(9)=c(10)=0

Conclusion of the joint test

- 0.99 is higher than 0.69
- Decision : drop only 3 variables

$\Rightarrow$ eq02

Regression linear output (eq02)

Dependent Variable: FERT_RATE
Method: Least Squares
Date: 05/09/12 Time: 13:03
Sample: 1 62
Included observations: 62

	Coefficient	Std. Error	t-Statistic	Prob.
C	1.251172	0.670081	1.867194	0.0672
HIV	0.024486	0.014084	1.738602	0.0877
MATERNAL_DEATH	0.347779	0.055399	6.277700	0.0000
NEO_MORT_RA	0.024783	0.010954	2.262390	0.0276
POP_GROWTH	0.476869	0.081499	5.851186	0.0000
PRENATAL_CARE	0.006783	0.004494	1.509444	0.1369
CONTRACEPT	-0.008292	0.005090	-1.629189	0.1090

R-squared	0.922921	Mean dependent var	3.490968
Adjusted R-squared	0.914513	S.D. dependent var	1.520468
S.E. of regression	0.444558	Akaike info criterion	1.322533
Sum squared resid	10.86975	Schwarz criterion	1.562693
Log likelihood	-33.99851	Hannan-Quinn criter.	1.416826
F-statistic	109.7592	Durbin-Watson stat	2.526492
Prob(F-statistic)	0.000000		

FIGURE : eq02 : Regression linear output

Observations

- All explanatory variables are highly individually significant.
- Adjusted R-squared is higher than in eq01

$\Rightarrow$ Is there a correlation between explanatory variables ?

Correlation matrix & Heteroskedasticity

Correlation

	CONTRACEPT	FERT_RATE	HIV	MATERNAL_...	NEO_MORT...	POP_GROW...	PRENATAL_...
CONTRACEPT	1.000000	-0.827319	-0.059900	-0.764153	-0.867991	-0.651602	0.428840
FERT_RATE	-0.827319	1.000000	0.205232	0.899350	0.853849	0.814483	-0.437052
HIV	-0.059900	0.205232	1.000000	0.163279	0.173426	0.032715	0.114253
MATERNAL_...	-0.764153	0.899350	0.163279	1.000000	0.801829	0.687046	-0.524377
NEO_MORT...	-0.867991	0.853849	0.173426	0.801829	1.000000	0.650362	-0.529244
POP_GROW...	-0.651602	0.814483	0.032715	0.687046	0.650362	1.000000	-0.378717
PRENATAL_...	0.428840	-0.437052	0.114253	-0.524377	-0.529244	-0.378717	1.000000

FIGURE : eq02 : Correlation matrix

White test

No homoskedasticity in this model

Observation

High correlation (0.86, 0.76)

$\Rightarrow$ What are the relations ?

Correlation between explanatory variables

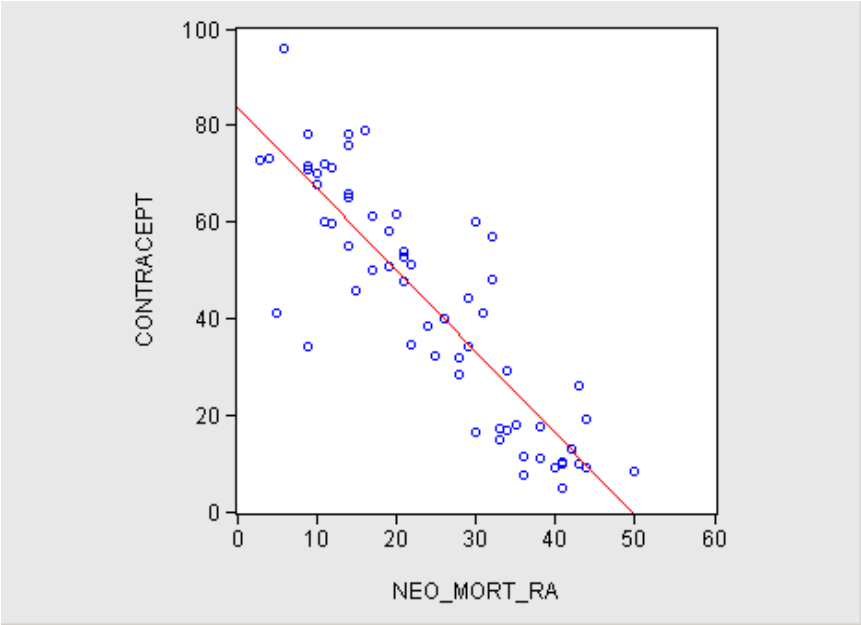


FIGURE : Contraception against neonatal mortality rate

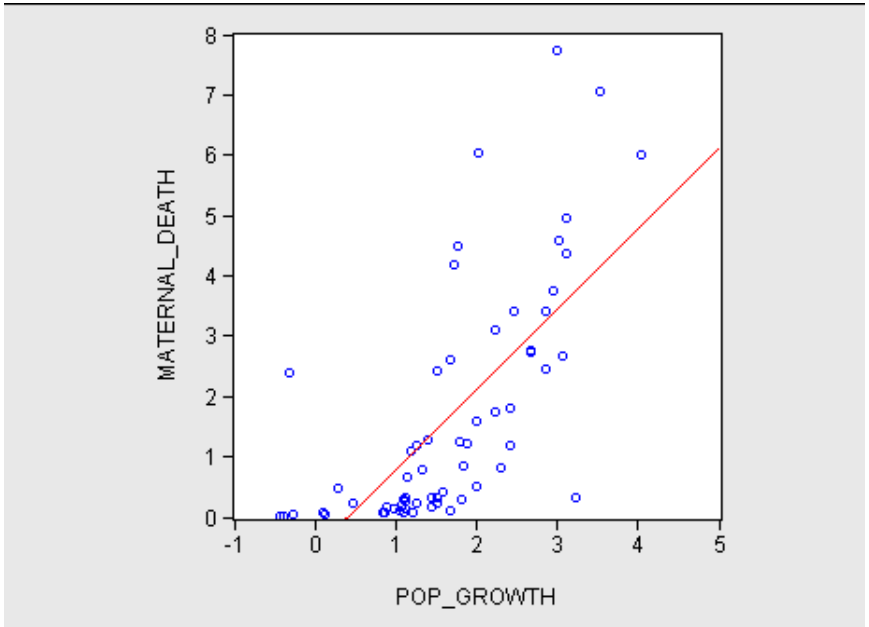


FIGURE : Maternal death against population growth

$\Rightarrow$ Drop 2 others explanatory variables for eq03 : Maternal death and neonatal mortality

Regression linear output (eq03)

Dependent Variable: FERT_RATE
Method: Least Squares
Date: 05/09/12 Time: 18:16
Sample: 1 62
Included observations: 62

	Coefficient	Std. Error	t-Statistic	Prob.
C	2.915064	0.697928	4.176738	0.0001
HIV	0.069508	0.024089	2.885508	0.0055
POP_GROWTH	1.117532	0.113288	9.864495	0.0000
PRENATAL_CARE	-0.017038	0.007144	-2.384929	0.0204

R-squared	0.722521	Mean dependent var	3.490968
Adjusted R-squared	0.708168	S.D. dependent var	1.520468
S.E. of regression	0.821379	Akaike info criterion	2.506676
Sum squared resid	39.13047	Schwarz criterion	2.643911
Log likelihood	-73.70696	Hannan-Quinn criter.	2.560558
F-statistic	50.34151	Durbin-Watson stat	2.035969
Prob(F-statistic)	0.000000		

FIGURE : eq03 : Regression linear output

Observations

- Adjusted R-squared is lower than in eq02
- Increase in HIV $\Rightarrow$ Increase in Fertility rate
- Increase in population growth $\Rightarrow$ Increase in Fertility rate
- Increase in prenatal care $\Rightarrow$ Decrease in Fertility rate

Heteroskedasticity & Residuals

Heteroskedasticity Test: White

F-statistic	1.631213	Prob. F(9,52)	0.1308
Obs*R-squared	13.65033	Prob. Chi-Square(9)	0.1353
Scaled explained SS	11.28273	Prob. Chi-Square(9)	0.2568

FIGURE : eq03 : White test

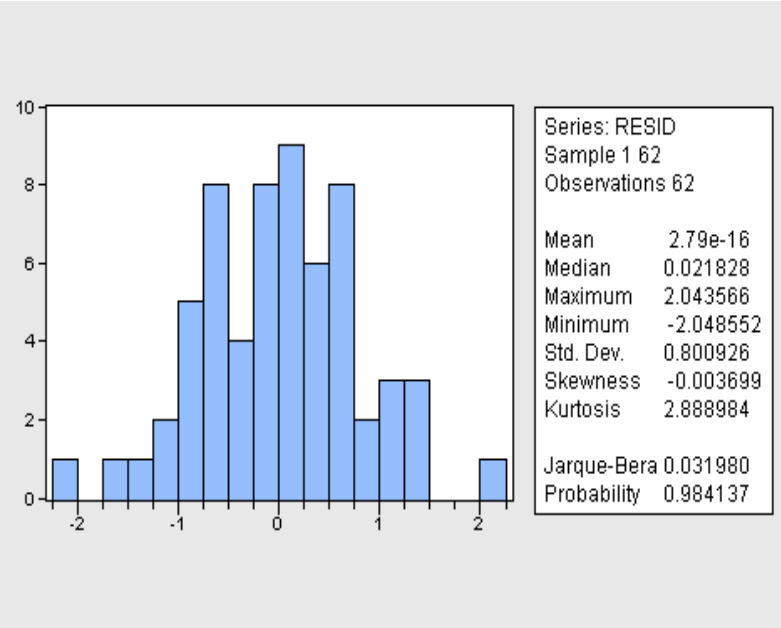


FIGURE : eq03 : Residuals

Observations

No heteroskedasticity

Observations

- Exogeneity
- Normality of errors

Fertility_rate=Fertility_Rate_Model

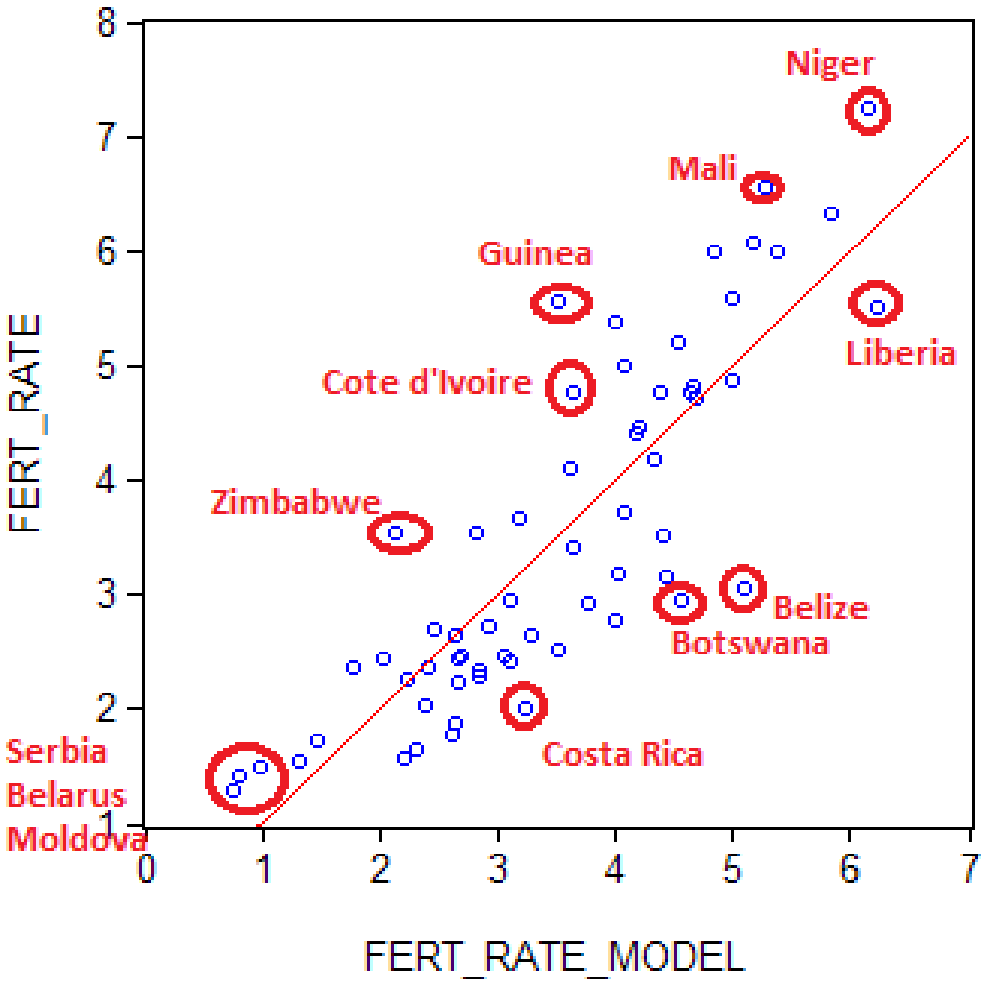


FIGURE : Fertility_rate=Fertility_Rate_Model

Conclusion

- Our final model is quite good BUT we might have forgotten a significant explanatory variable

$\Rightarrow$ Immigration ? Number of hospital ? Religion ?

- Try to include quadratic terms

$\Rightarrow$ Adjusted R-Squared better (with HIV-Squared) BUT model not so different