RESTAURANTS AND YOU

Thibault NONON - Jérôme CREVATIN

Econometrics I - ENAC

21.05.2014

Table of contents

Introduction

- Estimate the monthly budget to eat in restaurants for a person
- Build a model with collected datas

Method to get datas

- Use a Google poll spread by Facebook to get datas
- n=199 answers obtained in 5 days

Target

People who go to restaurants

Variables and expected effects

- k=20 variables

Variables	Effects
Age	+
Car	+
Children	?
Composition of the meal	?
Couple	+
Loyalty card	-
Gender	?
Income	+
Job	+

Variables and expected effects

Variables	Effects
Local product	+
Menu	+
Number of times	+
Evening or midday	+
Occupancy rate	?
City/Countryside	=
Tips	+
Type	+
Waiting time before seated	?
Waiting time before served	+
Weather	?

First Model

Our model is ruled with the following equation :

$$PAY = X\beta + u$$

PAY is the parameter that we want to estimate
X is the matrix of our variables defined before
β is the vector of coefficients
u is the vector of unobservable parameter

Coding the datas

2 possible answers question

Coded with a '1' or a '0'

Multiple possible answers question

- Each possibilty of answers observed
- Loss of information
- Coded with k=0,1,...,n

Correlation Matrix

	AGE	CAR	CHILDREN	COMPMEAL	COUPLE	FIDELITY	GENDER	INCOME	JOB	LOCALPRC	MENU	NBRTIME	NIGHT	OBS	OCCUPATI	PAY	PLACE	TIPS	TYPEREST	WAIT	WAITMEAL	WEATHER
AGE	1,000000	0,120858	0,880629	-0,086193	0,121365	-0,128883	0,033943	0,603614	-0,290466	0,040249	-0,092633	0,174056	-0,154143	0,170636	-0,082277	0,247140	-0,143938	0,070847	0,089610	0,096122	-0,014611	0,031716
CAR	0,120858	1,000000	0,111386	0,045892	0,046221	0,021678	-0,021828	0,182013	-0,154279	0,086123	0,103326	0,010738	-0,086120	-0,071663	-0,000428	0,097170	-0,035606	0,143123	-0,143643	-0,016432	0,070689	0,009638
CHILDREN	0,880629	0,111386	1,000000	-0,063386	0,126230	-0,121356	0,062388	0,651161	-0,230634	0,021957	-0,047157	0,111703	-0,156960	0,140865	-0,033248	0,233772	-0,152372	0,063632	0,044296	0,049218	0,027460	0,063691
COMPMEAL	-0,086193	0,045892	-0,063386	1,000000	-0,026901	0,028037	-0,080234	-0,080870	-0,052467	0,028568	-0,018073	-0,254488	0,118780	-0,024932	-0,073059	-0,181132	0,077932	-0,071864	-0,074561	-0,053131	-0,037712	-0,099291
COUPLE	0,121365	0,046221	0,126230	-0,026901	1,000000	-0,046101	-0,094584	0,172424	-0,134209	0,043813	-0,003103	0,115297	0,079150	-0,007583	0,027803	0,147793	-0,053368	-0,077863	-0,050863	0,063181	0,097287	-0,012338
FIDELITY	-0,128883	0,021678	-0,121356	0,028037	-0,046101	1,000000	0,079438	-0,149220	0,135290	-0,026921	0,073625	0,074129	-0,094459	-0,089878	0,043894	-0,008508	0,019677	0,050673	-0,082064	-0,114778	0,084123	-0,012521
GENDER	0,033943	-0,021828	0,062388	-0,080234	-0,094584	0,079438	1,000000	0,057115	0,142088	-0,062738	-0,075269	0,061520	-0,183813	-0,009818	-0,061726	0,048300	-0,077001	-0,059058	-0,026732	-0,192554	-0,119810	0,041157
INCOME	0,603614	0,182013	0,651161	-0,080870	0,172424	-0,149220	0,057115	1,000000	-0,537255	0,139350	0,029612	0,328608	-0,254903	0,128584	0,071509	0,568078	-0,050594	0,241653	-0,054722	0,005125	-0,003729	0,131195
JOB	-0,290466	-0,154279	-0,230634	-0,052467	-0,134209	0,135290	0,142088	-0,537255	1,000000	-0,054523	-0,023400	-0,237509	0,008691	-0,132421	0,001427	-0,296767	0,032183	-0,087072	-0,027130	-0,098617	0,065831	-0,063643
LOCALPRC	0,040249	0,086123	0,021957	0,028568	0,043813	-0,026921	-0,062738	0,139350	-0,054523	1,000000	0,008128	0,068319	0,020372	0,014046	-0,061049	0,180049	-0,115718	0,157612	0,011282	-0,005295	0,105150	0,002586
MENU	-0,092633	0,103326	-0,047157	-0,018073	-0,003103	0,073625	-0,075269	0,029612	-0,023400	0,008128	1,000000	0,166961	0,082202	-0,094549	0,099121	0,157768	0,003745	0,106191	-0,039441	0,098665	0,091259	0,030431
NBRTIME	0,174056	0,010738	0,111703	-0,254488	0,115297	0,074129	0,061520	0,328608	-0,237509	0,068319	0,166961	1,000000	-0,332442	-0,018981	0,049703	0,701537	0,035013	0,118900	0,138291	-0,002514	0,005479	-0,008469
NIGHT	-0,154143	-0,086120	-0,156960	0,118780	0,079150	-0,094459	-0,183813	-0,254903	0,008691	0,020372	0,082202	-0,332442	1,000000	0,018948	-0,032460	-0,159093	-0,066515	-0,025019	-0,088766	0,089411	0,026809	-0,053523
OBS	0,170636	-0,071663	0,140865	-0,024932	-0,007583	-0,089878	-0,009818	0,128584	-0,132421	0,014046	-0,094549	-0,018981	0,018948	1,000000	0,010058	0,062070	-0,000255	0,092338	0,047460	0,103133	0,085293	-0,034870
OCCUPATI	-0,082277	-0,000428	-0,033248	-0,073059	0,027803	0,043894	-0,061726	0,071509	0,001427	-0,061049	0,099121	0,049703	-0,032460	0,010058	1,000000	0,160137	0,257684	0,031366	-0,192890	0,093640	0,101369	-0,016038
PAY	0,247140	0,097170	0,233772	-0,181132	0,147793	-0,008508	0,048300	0,568078	-0,296767	0,180049	0,157768	0,701537	-0,159093	0,062070	0,160137	1,000000	0,008592	0,201104	0,006767	0,045112	0,023694	0,047535
PLACE	-0,143938	-0,035606	-0,152372	0,077932	-0,053368	0,019677	-0,077001	-0,050594	0,032183	-0,115718	0,003745	0,035013	-0,066515	-0,000255	0,257684	0,008592	1,000000	-0,090414	0,015220	-0,026601	-0,013755	-0,000705
TIPS	0,070847	0,143123	0,063632	-0,071864	-0,077863	0,050673	-0,059058	0,241653	-0,087072	0,157612	0,106191	0,118900	-0,025019	0,092338	0,031366	0,201104	-0,090414	1,000000	0,034245	0,094491	0,127067	0,084715
TYPEREST	0,089610	-0,143643	0,044296	-0,074561	-0,050863	-0,082064	-0,026732	-0,054722	-0,027130	0,011282	-0,039441	0,138291	-0,088766	0,047460	-0,192890	0,006767	0,015220	0,034245	1,000000	-0,180095	-0,126970	-0,024807
WAIT	0,096122	-0,016432	0,049218	-0,053131	0,063181	-0,114778	-0,192554	0,005125	-0,098617	-0,005295	0,098665	-0,002514	0,089411	0,103133	0,093640	0,045112	-0,026601	0,094491	-0,180095	1,000000	0,197487	-0,041555
WAITMEAL	-0,014611	0,070689	0,027460	-0,037712	0,097287	0,084123	-0,119810	-0,003729	0,065831	0,105150	0,091259	0,005479	0,026809	0,085293	0,101369	0,023694	-0,013755	0,127067	-0,126970	0,197487	1,000000	-0,041103
WEATHER	0,031716	0,009638	0,063691	-0,099291	-0,012338	-0,012521	0,041157	0,131195	-0,063643	0,002586	0,030431	-0,008469	-0,053523	-0,034870	-0,016038	0,047535	-0,000705	0,084715	-0,024807	-0,041555	-0,041103	1,000000

Correlation Matrix

High correlation between :

- Children and Age : 0,88
- Children and Income : 0,65
- Age and Income : 0,60

We delete Children and Age due to their high p-value

Regression

Results computed with E-views

R^2

- $R^2 = 0{,}67$
- Value quite high

Comments

- We can maybe improve the model

Dependent Variable: PAY
Method: Least Squares
Date: 05/20/14 Time: 17:44
Sample: 1 199
Included observations: 199

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-46.76410	24.46847	-1.911199	0.0576
CAR	3.760918	5.518304	0.681535	0.4964
COMPMEAL	0.248759	2.398267	0.103724	0.9175
COUPLE	-0.320196	5.377203	-0.059547	0.9526
FIDELITY	1.124047	5.635072	0.199473	0.8421
GENDER	2.421527	5.460805	0.443438	0.6580
INCOME	0.024281	0.003421	7.098429	0.0000
JOB	10.13502	6.323588	1.602732	0.1107
LOCALPRO	10.79460	5.624650	1.919159	0.0565
MENU	1.529899	5.409549	0.282815	0.7776
NBRTIME	8.688840	0.724386	11.99476	0.0000
NIGHT	27.85207	8.630241	3.227265	0.0015
OCCUPATION	12.65796	5.614540	2.254496	0.0254
PLACE	0.483530	7.864436	0.061483	0.9510
TIPS	1.454642	4.960075	0.293270	0.7697
TYPEREST	-0.001269	0.005807	-0.218434	0.8273
WAIT	2.474021	3.277305	0.754895	0.4513
WAITMEAL	-1.853592	4.551838	-0.407218	0.6843
WEATHER	2.143875	8.261094	0.259515	0.7955

R-squared	0.666324	Mean dependent var	69.12563
Adjusted R-squared	0.632956	S.D. dependent var	58.88663
S.E. of regression	35.67596	Akaike info criterion	10.07744
Sum squared resid	229099.3	Schwarz criterion	10.39187
Log likelihood	-983.7051	Hannan-Quinn criter.	10.20470
F-statistic	19.96918	Durbin-Watson stat	2.134750
Prob(F-statistic)	0.000000		

Testing our model

Test the model to see the validity

- Homoscedasticity
- Normality
  - Jarque Bera Test
  - Quantile plot

Homoscedasticity

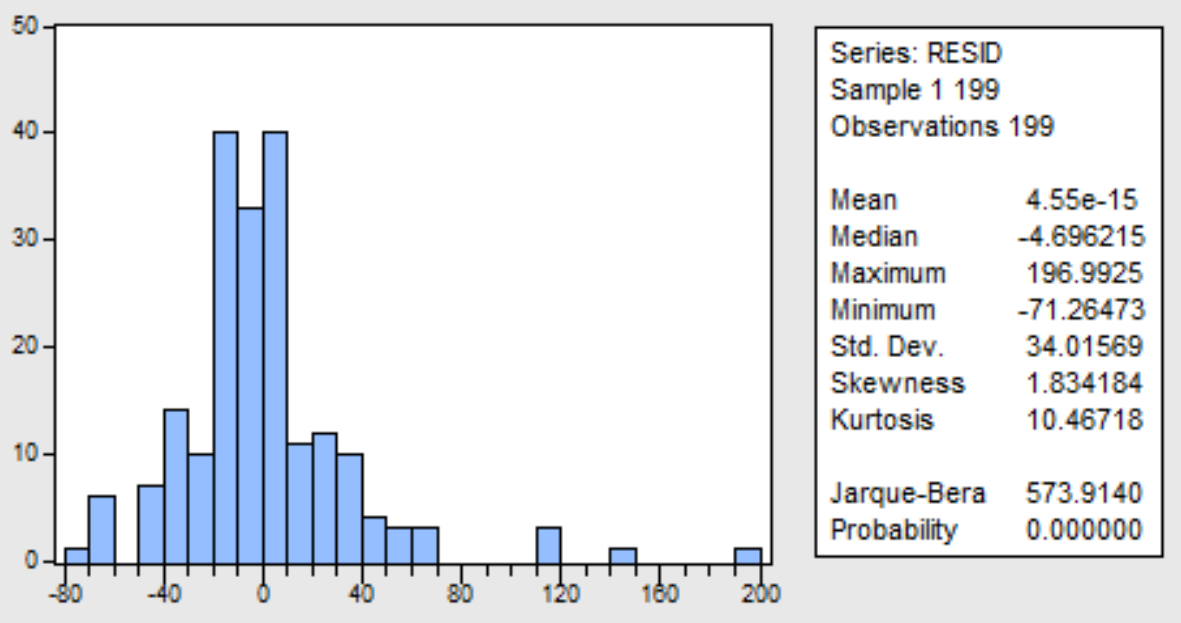
- The null of homoscedasticity should NOT BE rejected
- i.e. $Var(X_i) = \sigma^2$

Heteroskedasticity Test: White

F-statistic	0.929033	Prob. F(18,180)	0.5445
Obs*R-squared	16.91619	Prob. Chi-Square(18)	0.5289
Scaled explained SS	65.51370	Prob. Chi-Square(18)	0.0000

Normality

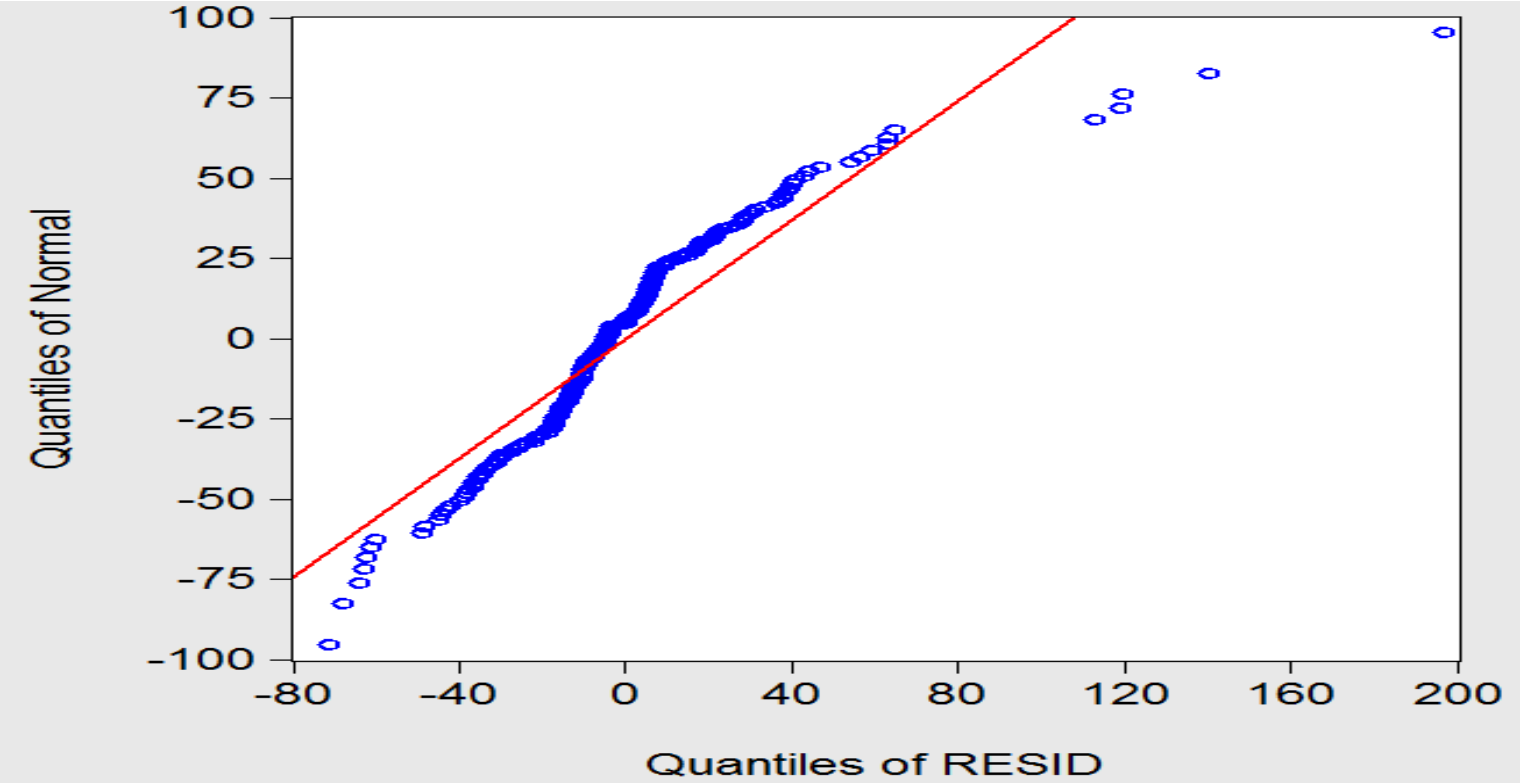
- Jarque Bera Test



Hypothesis of normal law is rejected

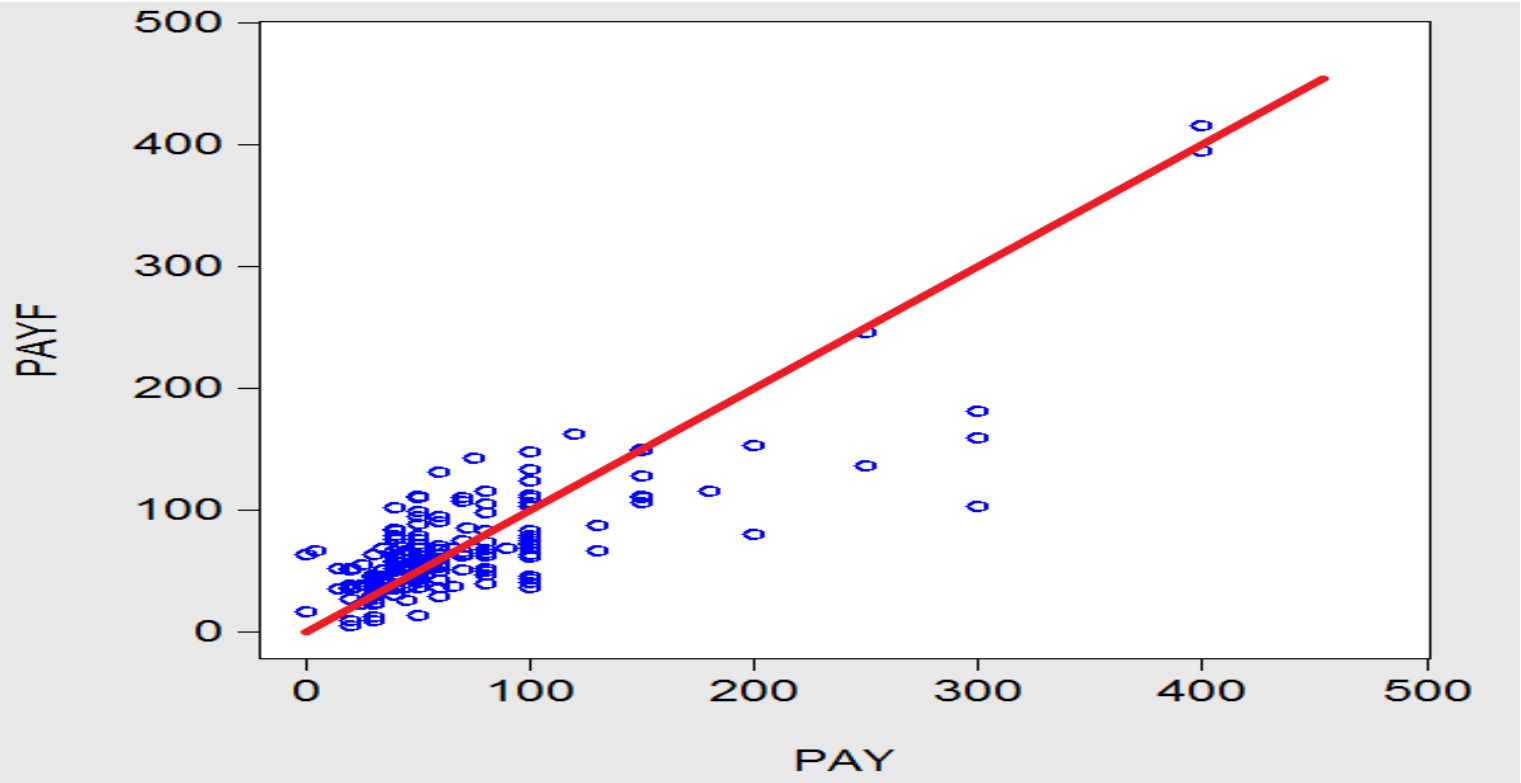
Normality

- Quantile plot



Forecast

- Linear model quiet good
- Delete observations which are not following the model



First Conclusion

- Estimation could be much better
- Some problems with high values
- Normality of errors non verified
- R^2 good

Build a more precise model

Second Regression

Results computed with E-views

R^2

- $R^2 = 0{,}75$
- Increase of 0,08. Better

Dependent Variable: PAY
Method: Least Squares
Date: 05/20/14 Time: 19:45
Sample (adjusted): 1 194
Included observations: 194 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-29.21282	17.89054	-1.632864	0.1043
CAR	3.630528	4.039553	0.898745	0.3700
COMPMEAL	0.041336	1.753717	0.023571	0.9812
COUPLE	2.836317	3.953223	0.717470	0.4740
FIDELITY	2.457430	4.098537	0.599587	0.5496
GENDER	-2.555333	4.005545	-0.637949	0.5243
INCOME	0.022418	0.002488	9.009974	0.0000
JOB	8.485826	4.682879	1.812096	0.0717
LOCALPRO	4.310320	4.082130	1.055900	0.2925
MENU	-1.773886	3.962013	-0.447723	0.6549
NBRTIME	7.631318	0.542587	14.06469	0.0000
NIGHT	15.86780	6.287427	2.523736	0.0125
OCCUPATION	9.300705	4.087070	2.275641	0.0241
PLACE	-4.453580	5.684277	-0.783491	0.4344
TIPS	2.905295	3.627807	0.800841	0.4243
TYPEREST	-0.001032	0.004260	-0.242182	0.8089
WAIT	1.462968	2.392645	0.611444	0.5417
WAITMEAL	3.025372	3.313539	0.913033	0.3625
WEATHER	-4.232394	6.129238	-0.690525	0.4908

R-squared	0.754957	Mean dependent var	63.94845
Adjusted R-squared	0.729753	S.D. dependent var	49.43408
S.E. of regression	25.69845	Akaike info criterion	9.423542
Sum squared resid	115571.8	Schwarz criterion	9.743590
Log likelihood	-895.0836	Hannan-Quinn criter.	9.553139
F-statistic	29.95342	Durbin-Watson stat	2.205587
Prob(F-statistic)	0.000000		

Homoscedasticity

- The null of homoscedasticity should NOT BE rejected
- White cross terms test added
- i.e. $Var(X_i) = \sigma^2$

Heteroskedasticity Test: White

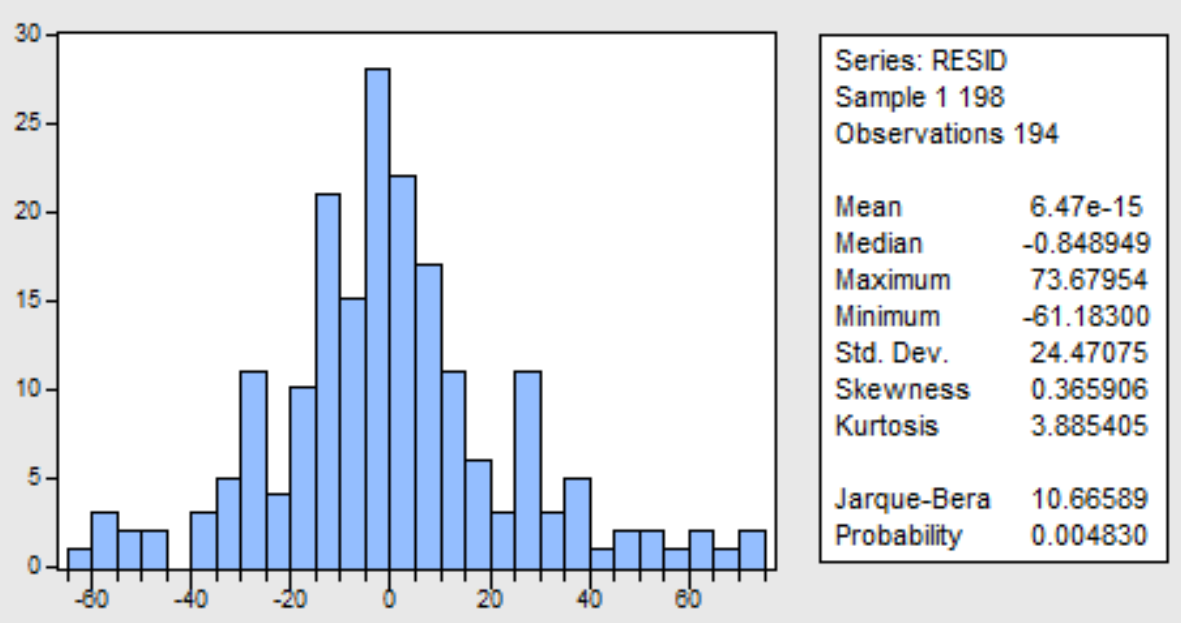
F-statistic	1.576678	Prob. F(179,14)	0.1660
Obs*R-squared	184.8313	Prob. Chi-Square(179)	0.3669
Scaled explained SS	216.9826	Prob. Chi-Square(179)	0.0277

Heteroskedasticity Test: White

F-statistic	1.177821	Prob. F(18,175)	0.2840
Obs*R-squared	20.96297	Prob. Chi-Square(18)	0.2813
Scaled explained SS	24.60946	Prob. Chi-Square(18)	0.1361

Normality

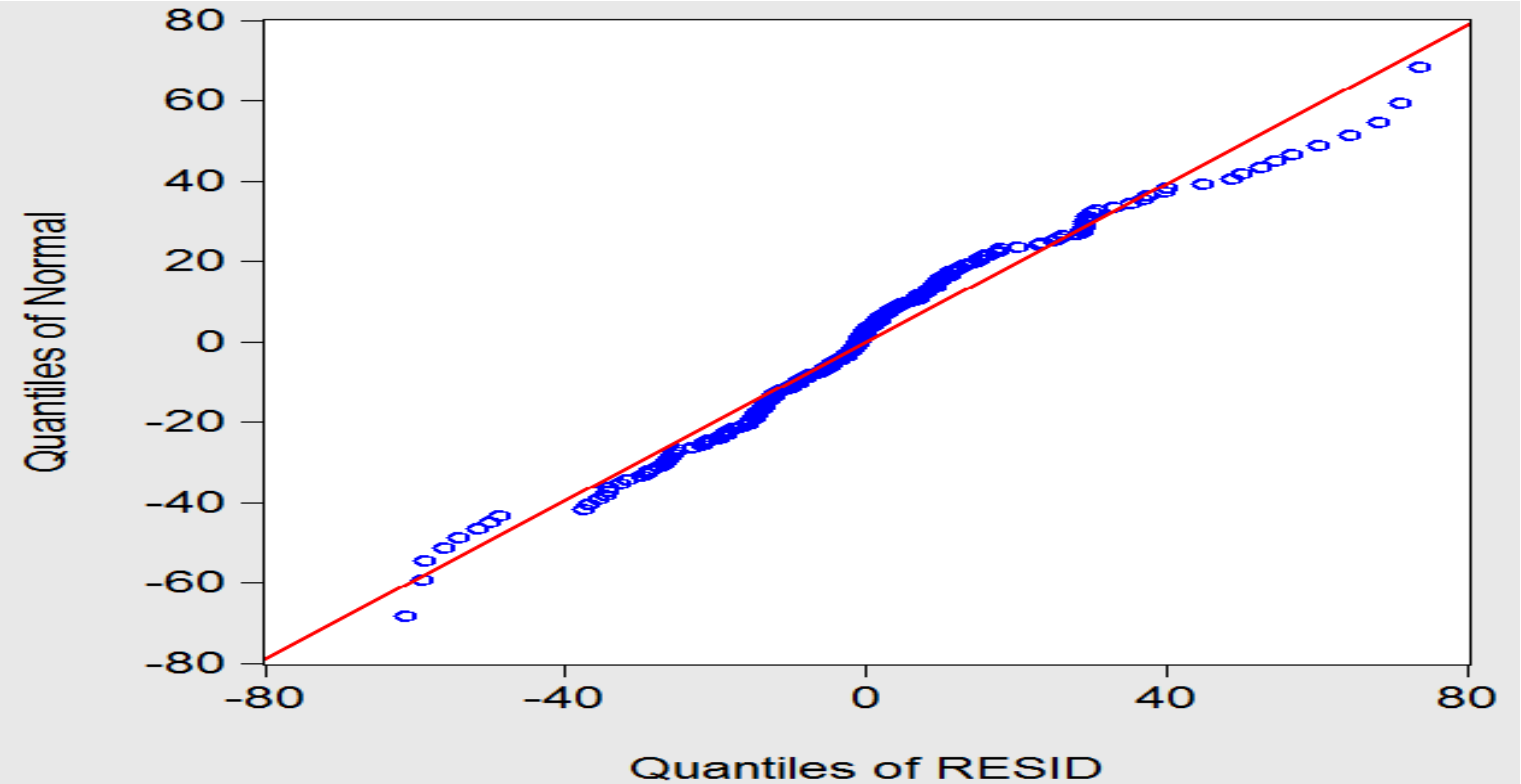
- Jarque Bera Test



A normal law seems to appear. Test result is quite better.

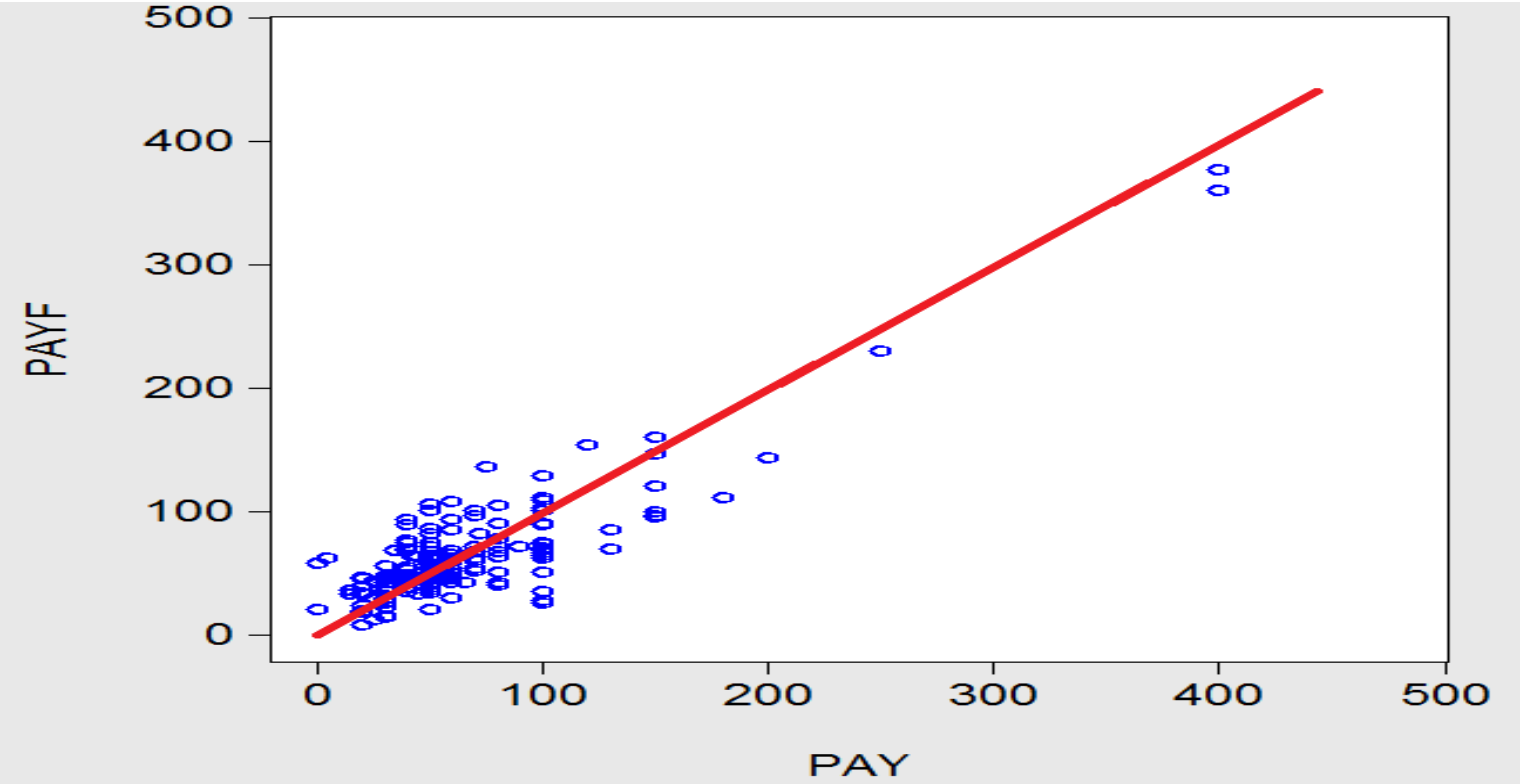
Normality

- Quantile plot



Forecast

- Linear model better
- Marginal results for PAY around 0



Second Conclusion

- Better estimation
- R^2 higher : 0,75
- Normality of errors improved
- Homoscedasticity of residuals

Second model is better thant the first one

Main effects

OCCUPANCY

- $\beta = 9{,}3$
- p-value $= 0{,}0241$

EVENING (or MIDDAY)

- $\beta = 15{,}86$
- p-value $= 0{,}0125$

CITY (or COUNTRYSIDE)

- $\beta =$ -4,45
- p-value $= 0{,}4344$

General Conclusion

Model

- Precise
- Errors shaping a normal law
- R^2 high
- Could be improved

Why this is not a perfect model

- People targetted (all students)
- Y may have be changed
- Multiple answers variables have a high p-value
- Number of answers should have been higher

General Conclusion

- Thank you for your attention
- Some questions ?