Car prices of the IENAC

Christophe Brunner - Clément Chauvin

ENAC

May 21, 2013

Introduction

- Econometrics study
- Variable of interest: car price of the IENAC

Poll

- Target: IENAC98 to 12
- Spreading: emails
- 91 workable data

- **Dependent variable**
  - Car price
- **Explanatory variables**
  - Gender
  - Age
  - Civil/Fonctionnaire
  - Promo
  - Track T/L/S
  - Geographical origin North/South
  - Rank whithin class
  - Associative involvment + level of responsability
  - New/used car
  - Year of purchase
  - Age of the car when purchasing
  - French model or not
  - Wage

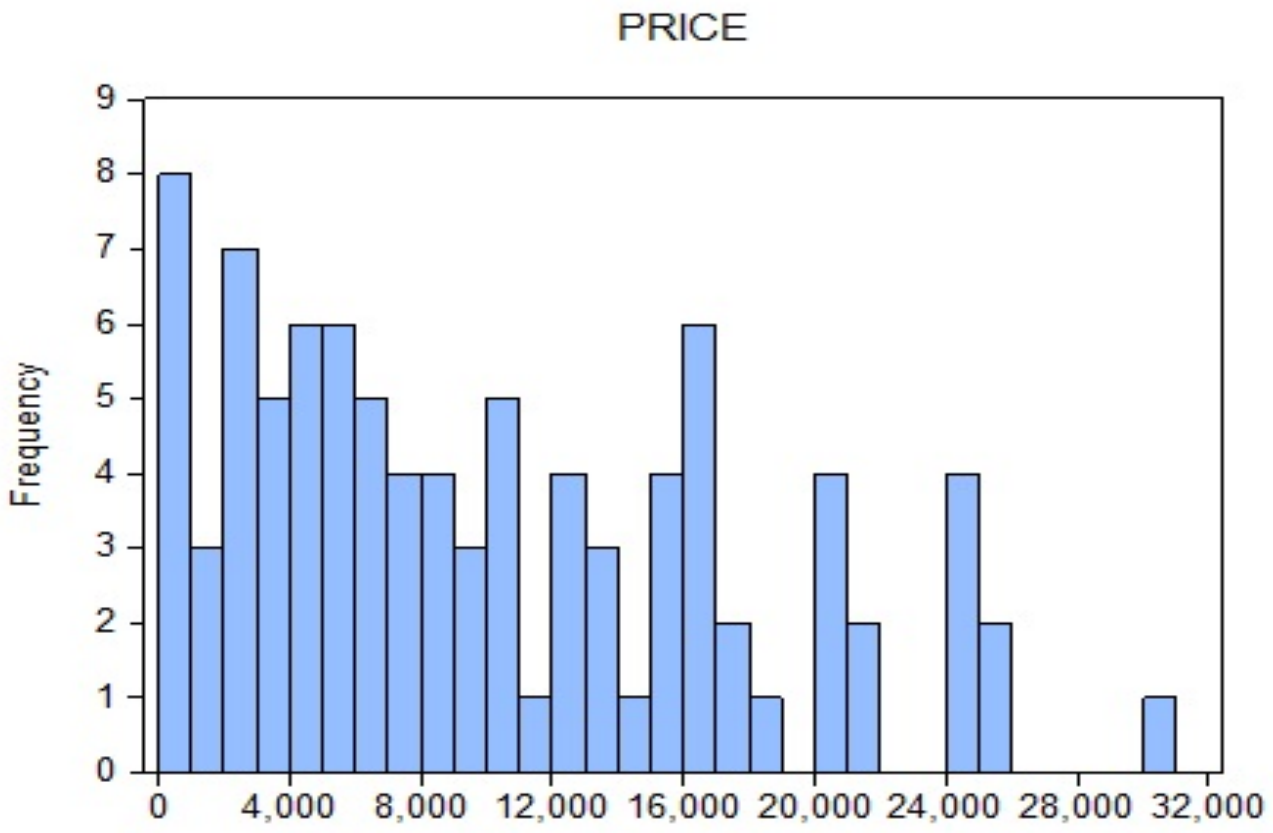

Figure : Car price histogram

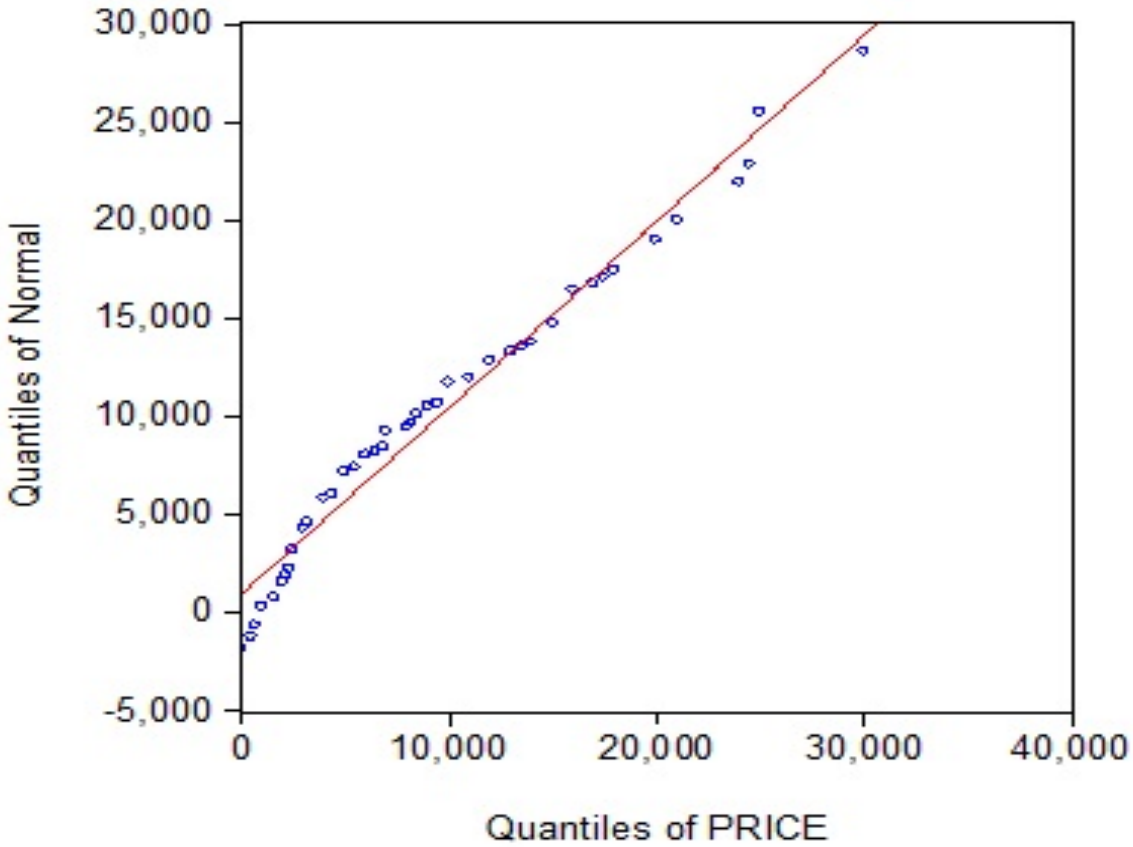

Figure : Quantiles of price vs normal

Expected effects on car price

Independent variable	Expected effect
Male	?
Age	+
North	?
Civil	-
Promo	-
Rank	-
Respo	+
T	?
New	+
Year of purchase	+
French model	-
Age of car	-
Wage	+

Car price function (EQ01)

$$PRICE_i = \beta_0 + \beta_1 \cdot MALE_i + \beta_2 \cdot AGE_i + \beta_3 \cdot NORTH_i + \beta_4 \cdot CIVIL_i + \beta_5 \cdot PROMO_i + \beta_6 \cdot RANK_i + \beta_7 \cdot RESPO_i + \beta_8 \cdot T_i + \beta_9 \cdot NEW_i + \beta_{10} \cdot YEAR\_PURCHASE_i + \beta_{11} \cdot FRENCH\_MODEL_i + \beta_{12} \cdot CAR\_AGE_i + \beta_{13} \cdot WAGE_i + u_i \quad (1)$$

Dummy variables

MALE, NORTH, CIVIL, T, NEW, FRENCH_MODEL are dummy variables

Linear Regression Output

Dependent Variable: PRICE
Method: Least Squares
Date: 05/18/13 Time: 16:13
Sample: 1 91
Included observations: 86

	Coefficient	Std. Error	t-Statistic	Prob.
C	1243487.	1133039.	1.097480	0.2761
MALE	3136.037	1277.706	2.454428	0.0165
AGE	-598.0613	485.1967	-1.232616	0.2217
NORTH	-1219.219	1163.966	-1.047470	0.2984
CIVIL	-2019.368	1977.071	-1.021393	0.3105
PROMO	-1016.227	496.3171	-2.047535	0.0443
RANK	-657.5162	598.9597	-1.097764	0.2760
RESPO	579.4049	634.0948	0.913751	0.3639
T	656.9236	1155.413	0.568562	0.5714
NEW	6986.757	1667.087	4.190996	0.0001
YEAR_PURCHASE	408.8092	211.8654	1.929570	0.0576
FRENCH_MODEL	-1494.449	1191.342	-1.254424	0.2137
CAR_AGE	-420.7166	110.5477	-3.805748	0.0003
WAGE	0.038506	0.045534	0.845648	0.4006

R-squared	0.633880	Mean dependent var	9401.163
Adjusted R-squared	0.567775	S.D. dependent var	7411.271
S.E. of regression	4872.453	Akaike info criterion	19.96848
Sum squared resid	1.71E+09	Schwarz criterion	20.36803
Log likelihood	-844.6448	Hannan-Quinn criter.	20.12928
F-statistic	9.588984	Durbin-Watson stat	1.983473
Prob(F-statistic)	0.000000		

Figure : First equation

$R^2 = 0.63 \rightarrow$ quite strong model

Real effects on car price

Independent variable	Expected effect	Real effect
Male	?	+
Age	+	-
North	?	-
Civil	-	-
Promo	-	-
Rank	-	-
Respo	+	+
T	?	+
New	+	+
Year of purchase	+	+
French model	-	-
Age of car	-	-
Wage	+	+

Covariance Analysis: Ordinary
Date: 05/18/13 Time: 16:16
Sample: 1 91
Included observations: 86
Balanced sample (listwise missing value deletion)

Correlation	MALE	AGE	NORTH	CIVIL	PROMO	RANK	RESPO	T	NEW	YEAR_PU...	FRENCH...	CAR_AGE	WAGE
MALE	1.000000												
AGE	0.198430	1.000000											
NORTH	0.068592	0.059388	1.000000										
CIVIL	0.108634	0.214061	-0.120081	1.000000									
PROMO	-0.172432	-0.963566	-0.118466	-0.260949	1.000000								
RANK	0.011079	0.050249	0.005205	0.050391	-0.053537	1.000000							
RESPO	-0.058805	-0.287244	0.104368	0.160676	0.289292	0.179515	1.000000						
T	-0.212225	0.085910	0.049932	-0.016890	-0.041157	-0.077967	-0.010202	1.000000					
NEW	0.070034	0.214327	0.060006	0.166728	-0.244124	0.001520	-0.172387	0.174781	1.000000				
YEAR_PURCHASE	-0.001086	-0.273327	-0.089244	-0.031491	0.235936	-0.081428	-0.070604	-0.061982	-0.150747	1.000000			
FRENCH_MODEL	0.005553	-0.165081	0.178488	0.051115	0.119405	-0.139825	0.246167	-0.158016	-0.033824	-0.146413	1.000000		
CAR_AGE	-0.022274	-0.269052	-0.025848	0.047168	0.305667	-0.020325	0.132185	-0.094330	-0.438425	-0.032406	0.030252	1.000000	
WAGE	0.138693	0.753553	0.089245	-0.003854	-0.736948	-0.085848	-0.337740	0.054882	0.172216	-0.087664	-0.163625	-0.248978	1.000000

Figure : Correlation Matrix

Correlation

High correlation between AGE, PROMO and WAGE, we choose to remove AGE and WAGE

Wald test on T coefficient

High p-value so we do not reject the null hypothesis. Thus we consider: $\beta_8 = 0$

Wald Test:
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	0.323262	(1, 72)	0.5714
Chi-square	0.323262	1	0.5697

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(9)	656.9236	1155.413

Restrictions are linear in coefficients.

Figure : Wald test on: $C(9) = \beta_8 = 0$

Other Wald tests

We can also remove every variable with $p \geq 0.2$

Linear Regression Output

Dependent Variable: PRICE
Method: Least Squares
Date: 05/18/13 Time: 16:29
Sample: 1 91
Included observations: 87

	Coefficient	Std. Error	t-Statistic	Prob.
C	-52479.54	406327.7	-0.129156	0.8976
MALE	2549.784	1184.407	2.152794	0.0343
PROMO	-481.7927	124.2828	-3.876584	0.0002
NEW	6831.726	1548.557	4.411671	0.0000
YEAR_PURCHASE	511.3240	193.0191	2.649085	0.0097
CAR_AGE	-451.1506	104.9347	-4.299345	0.0000

R-squared	0.601555	Mean dependent var	9373.563
Adjusted R-squared	0.576960	S.D. dependent var	7372.552
S.E. of regression	4795.222	Akaike info criterion	19.85510
Sum squared resid	1.86E+09	Schwarz criterion	20.02516
Log likelihood	-857.6968	Hannan-Quinn criter.	19.92358
F-statistic	24.45806	Durbin-Watson stat	2.071623
Prob(F-statistic)	0.000000		

Figure : Optimized equation

$R^2 = 0.60$ and Adjusted-$R^2 = 0.58$

Optimized car price function (EQ02)

$$PRICE_i = \beta_0 + \beta_1 \cdot MALE_i + \beta_2 \cdot PROMO_i + \beta_3 \cdot NEW_i + \beta_4 \cdot YEAR\_PURCHASE_i + \beta_5 \cdot CAR\_AGE_i + u_i \quad (2)$$

Heteroskedasticity

Heteroskedasticity Test: White

F-statistic	1.739256	Prob. F(14,72)	0.0662
Obs*R-squared	21.98675	Prob. Chi-Square(14)	0.0789
Scaled explained SS	24.49072	Prob. Chi-Square(14)	0.0399

Figure : Heteroskedasticity test

We reject the null hypothesis at the 90% level of significance

Normality and exogeneity

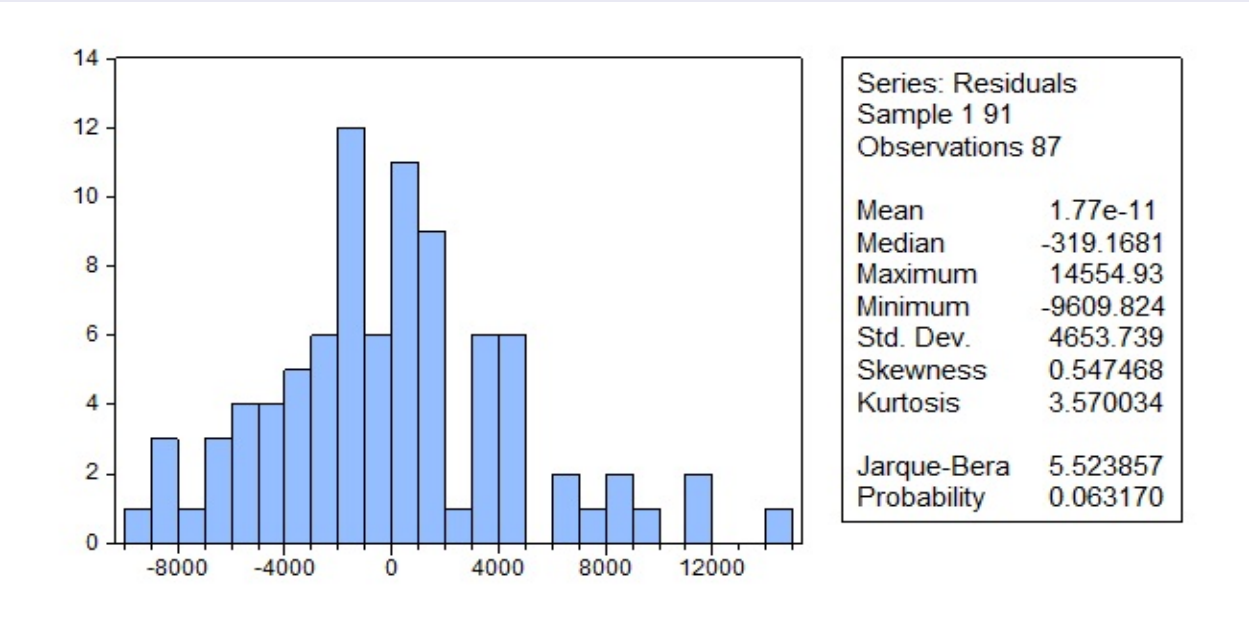

Figure : Normality and exogeneity test

$Mean = 0$ so we have exogeneity.
The normality of residuals is rejected at the 90% level of significance

Conclusion

- Strong influence of promo, age of the car, new/used, but quite obvious parameters!
- No influence of tracks T/L/S, and level of responsability
- Surprising result: negative impact of age