

## ANNEXE À LA TROISIÈME PARTIE : LA LECTURE D'UNE SORTIE D'ORDINATEUR

La présente section jette un coup d'oeil dans les coulisses de l'analyse de régression. On y présente trois extraits de sorties d'ordinateur dont sont tirés les résultats de Lemelin et Polèse (1995). Le logiciel utilisé est SAS. Le premier extrait contient les résultats relatifs à l'équation (6) de Lemelin et Polèse (1995) :

$$\ln PURB_i = a + b \ln PTOT_i + c \ln GNPC_i + d (\ln GNPC_i)^2 \quad (6)$$

Les résultats du second extrait ne sont pas mentionnés dans Lemelin et Polèse (1995). Il s'agit du modèle tronqué

$$\ln PURB_i = a + b \ln PTOT_i + c \ln GNPC_i$$

Le troisième extrait enfin contient les résultats relatifs à l'équation (1) de Lemelin et Polèse (1995) :

$$\ln PLAR_i = \ln K + h \ln PURB_i \quad (1)$$

Parmi les éléments que l'on peut repérer dans les sorties d'ordinateur, mentionnons :

- les valeurs estimées des paramètres
- leur erreur d'échantillonnage type
- la valeur du  $t$  de Student associé à l'hypothèse de coefficient nul
- la probabilité critique (test bilatéral)
- les sommes de carrés  $SSM$ ,  $SSR$  et  $SST$
- le nombre de degrés de liberté associé à chaque somme de carrés
- le coefficient de détermination multiple  $R^2$
- le coefficient de détermination multiple ajusté  $\bar{R}^2$
- la valeur de la variable-test  $F$  pour l'hypothèse que tous les coefficients sauf la constante sont nuls
- la probabilité critique correspondante

PRIMATE moins HK ; UN POP DIV et GNPC 15  
 18:19 Tuesday, August 29, 1995

Model: EQPURB  
 Dependent Variable: LPURB

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	88.81851	29.60617	409.878	0.0001
Error	60	4.33390	0.07223		
C Total	63	93.15241			

Root MSE	0.26876	R-square	0.9535
Dep Mean	9.47138	Adj R-sq	0.9511
C.V.	2.83760		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >  T
INTERCEP	1	-4.602824	0.87667347	-5.250	0.0001
LPTOT	1	0.971663	0.02793210	34.787	0.0001
LGNPC	1	0.915958	0.20762625	4.412	0.0001
SQLGNPC	1	-0.045267	0.01344728	-3.366	0.0013

.....  
 PRIMATE moins HK ; UN POP DIV et GNPC 16  
 18:19 Tuesday, August 29, 1995

Dependent Variable: LPURB

Test: PTOTEG1 Numerator: 0.0743 DF: 1 F value: 1.0292  
 Denominator: 0.072232 DF: 60 Prob>F: 0.3144

PRIMATE moins HK ; UN POP DIV et GNPC 17  
18:19 Tuesday, August 29, 1995

Model: EQPURB1  
Dependent Variable: LPURB

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	87.99999	43.99999	520.920	0.0001
Error	61	5.15242	0.08447		
C Total	63	93.15241			
Root MSE	0.29063	R-square	0.9447		
Dep Mean	9.47138	Adj R-sq	0.9429		
C.V.	3.06851				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >  T
INTERCEP	1	-1.901520	0.38173324	-4.981	0.0001
LPTOT	1	0.957929	0.02988116	32.058	0.0001
LGNPC	1	0.221163	0.02438139	9.071	0.0001

.....  
PRIMATE moins HK ; UN POP DIV et GNPC 18  
18:19 Tuesday, August 29, 1995

Dependent Variable: LPURB

Test: PTOTEG1 Numerator: 0.1674 DF: 1 F value: 1.9823  
Denominator: 0.084466 DF: 61 Prob>F: 0.1642

PRIMATE moins HK ; UN POP DIV et GNPC 24  
 18:19 Tuesday, August 29, 1995

Model: EQPLAR1C  
 Dependent Variable: LPLAR

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	37.70040	37.70040	222.803	0.0001
Error	62	10.49099	0.16921		
C Total	63	48.19138			

Root MSE	0.41135	R-square	0.7823
Dep Mean	8.09242	Adj R-sq	0.7788
C.V.	5.08316		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >  T
INTERCEP	1	2.066973	0.40693370	5.079	0.0001
LPURB	1	0.636174	0.04262018	14.927	0.0001

### **DIGRESSION : LA FORME DE LA RELATION ENTRE LA POPULATION URBAINE ET LE PIB PER CAPITA**

L'équation (6) est une relation quadratique entre les logarithmes des variables. Mais quelle est réellement la forme de la relation entre la population urbaine et le PIB per capita ?

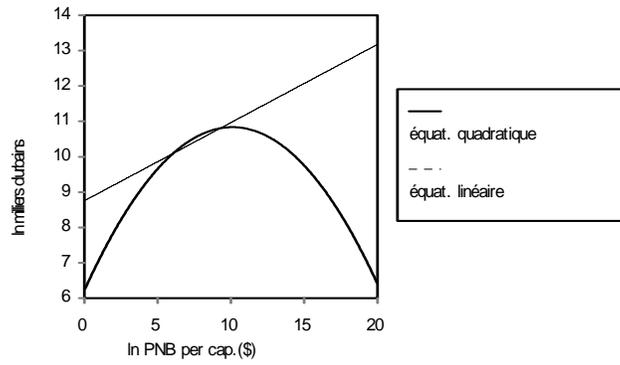
Avant l'ère des chiffriers électroniques, lorsqu'on voulait connaître la forme d'une fonction sans avoir à calculer un grand nombre de valeurs, on avait recours au calcul différentiel et à l'algèbre. Il fallait trouver les dérivées premières et secondes de la fonction, puis résoudre les équations permettant de déterminer ses points d'intersection avec l'axe des abscisses, ses maxima et minima, ses points d'inflexion, et enfin déterminer le signe de la fonction et de ses dérivées entre ces repères. Aujourd'hui, il suffit de simuler à l'aide d'un chiffrier électronique.

Pour ce qui est de l'équation (6), les résultats présentés ci-haut montrent que le coefficient  $b$  est approximativement égal à 1. Cela implique que, toutes choses étant égales par ailleurs (c'est-à-dire, notamment, le PIB per capita), la population urbaine est à peu près proportionnelle à la population totale. Nous allons donc concentrer notre attention sur la forme de la relation entre le PIB per capita et la population urbaine, pour une population totale donnée. Nous avons fixé le chiffre de la population à sa valeur moyenne dans l'échantillon (67,6 millions d'habitants) et fait varier le PIB per capita entre zéro et la valeur invraisemblable de 100 000 \$. Pour chaque valeur, nous avons calculé quelle était la population urbaine prédite par le modèle. Les résultats sont reproduits ci-après sous forme graphique.

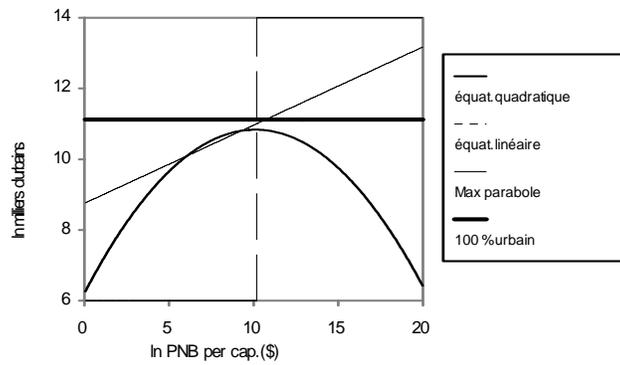
Les points à faire ressortir sont notamment :

- Le modèle représenté par l'équation (6) est un modèle descriptif, qui n'est valide qu'à l'intérieur d'un certain domaine de variation.
- L'aspect visuel de la courbe dépend du choix des échelles (naturelle ou logarithmique)

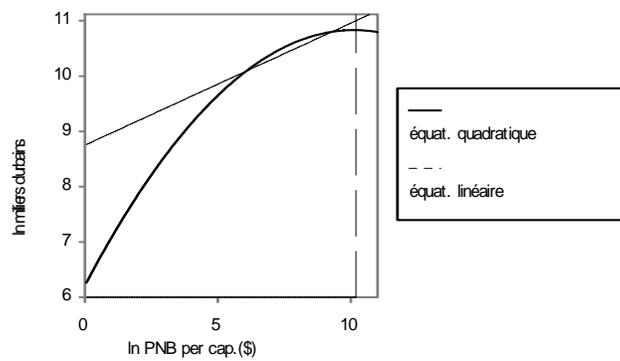
### Population urbaine calculée pour une pop. totale de 67,6 millions



### Population urbaine calculée pour une pop. totale de 67,6 millions

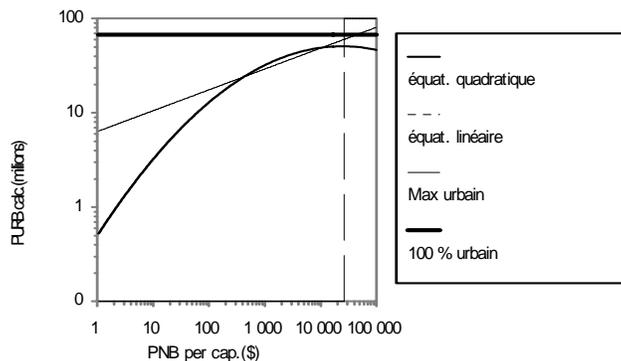


### Population urbaine calculée pour une pop. totale de 67,6 millions



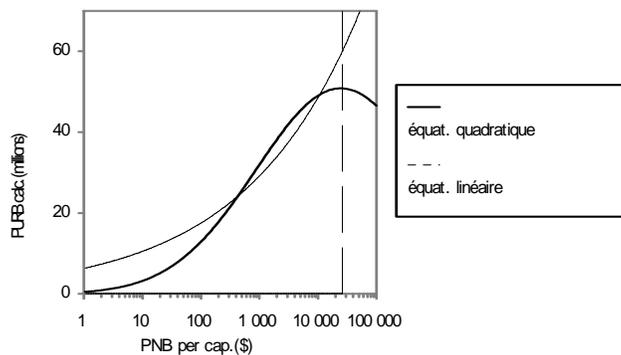
### Population urbaine calculée

pour une pop. totale de 67,6 millions



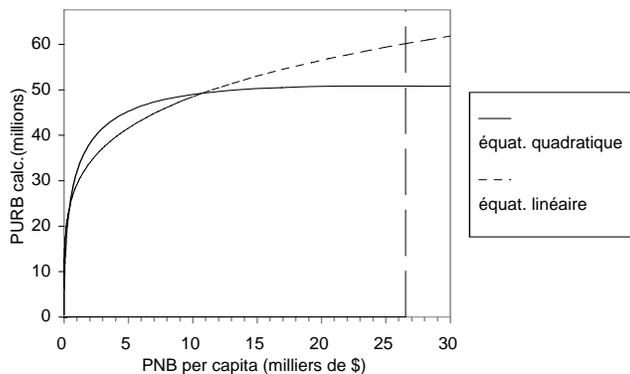
### Population urbaine calculée

pour une pop. totale de 67,6 millions



### Population urbaine calculée

pour une pop. totale de 67,6 millions



# Urbanisation observée/calculée

