

Projet Économétrie:
Étude des dynamiques régissant la criminalité du
Grand Paris

A.Gardahaut J.Mpoy

M1 MAS

Université de Rennes

14 Janvier 2019

À l'attention de Mme.Thelen

Contents

1	Introduction	3
1.1	Problématique et Contextualisation	3
1.2	Données utilisées	4
2	Démarche	5
2.1	Objectif	5
2.2	Modélisation	5
3	Analyse descriptive	5
3.1	Variables	5
3.2	Typologie de communes	7
4	Modélisation et Interprétation pour le Grand Paris	8
4.1	Spécification	8
4.1.1	Transformation en log de la variable taux_delit	8
4.1.2	Transformation logarithmique des variables explicatives	11
5	Choix des variables du modèle	12
6	Présence de multicollinéarité dans le modèle	13
6.1	La multicollinéarité et ses conséquences	13
6.2	Modification du modèle	15
7	Test de Spécification du modèle : Variable omise	18
7.1	Étude de l'Hétéroscédasticité	18
7.2	Interprétation de la régression	24
8	Interprétation du modèle log-log	24
8.1	Validité globale du modèle	24
8.2	Qualité d'ajustement	24
8.3	Significativité des variables explicatives	24
8.4	Coefficients estimés	24
8.5	Outliers	25
8.6	Critiques du modèle	27
9	Étude de Paris et du reste du Grand Paris	28
9.1	Test de Chow et Interprétation	30
10	Étude d'une typologie de crime dans le Grand Paris	34
10.1	taux de cambriolages (CRL) et de vols d'automobiles (VA)	34
11	Conclusion	36

1 Introduction

1.1 Problématique et Contextualisation

Si Paris est la ville préférée mondiale des touristes (Travelers's Choice 2018 de TripAdvisor), la région parisienne reste une région associée à la criminalité et la délinquance dans l'imaginaire collectif (français).

D'après Durkheim, une société sans crimes ne doit être considérée comme normale (1966).

Il est donc plutôt normal d'observer des crimes et d'avoir des criminels dans une société.

L'important est de déterminer les variables qui causent la criminalité dans une société, les éléments qui provoquent une baisse ou une hausse du taux de crime et ce qui pousse un individu au crime.

En effet la criminalité a un effet négatif sur l'activité économique et sur la qualité des citoyens qui se sentent moins en sécurité chez eux ou dans l'espace public.

De nombreuses études ont été réalisées sur le lien entre la criminalité et des variables économiques en particulier le taux de chômage, cependant les études faites dans le cadre de différentes sociétés semblent ne pas parvenir à un consensus.

Dans le cadre de ce projet nous souhaitons étudier les dynamiques qui régissent la criminalité du Grand Paris.

Pour ce faire, nous nous sommes appuyés sur le travail de Alison Oliver : "The Economics of Crime: An Analysis of Crime Rates in America".

L'étude cherche à déterminer les facteurs qui déterminent de manière significative les fluctuations du taux de crime au USA au cours du temps.

L'étude porte sur des données nationale sur la période 1960-1998.

Elle se base la théorie du "Rational Choice" de Becker qui exprime ce qui pousse l'individu au crime.

En effet, la criminalité est la somme de crime individuels.

Selon Becker, le bénéfice d'un crime est égal à :

$$b - (pc + o)$$

où b le gain lié au crime, p la probabilité de se faire prendre, c le coût de la punition et o le coût d'opportunité.

L'individu étant rationnel, un crime sera commis seulement si :

$$b - (pc + o) > 0$$

L'idée centrale est donc d'étudier les variables qui influent sur b, p, c et o.

La théorie nous pousse donc à penser que les variables favorisant b vont entraîner une hausse du taux de crime et inversement pour p, c et o.

'Crime reduction can occur through reducing the benefits of crime, raising the probability of being caught, or increasing the costs of punishment conditional upon being caught' (A.Oliver)

Ehrlich (1973) énonce que l'éducation influe sur la commission des crimes dans la mesure où un individu mieux éduqué saura mieux évaluer les bénéfices à tirer des activités légales et illégales.

Un investissement de l'individu dans l'éducation peut être associé avec un plus grand accès à des revenus légaux donc le coût d'opportunité o aurait tendance à être plus grand. On s'attend donc à ce que le taux de population éduquée évolue dans le sens inverse à celui du taux de crime.

Il exprime également que le taux de chômage peut être vu comme un indicateur des opportunités de revenu disponibles dans le marché du travail légal.

Donc si le taux de chômage croît, les opportunités dans le marché du travail traditionnel diminuent, donc plus susceptible de générer du revenu via des activités illégales, donc o diminue.

Ainsi d'après la théorie et les recherches existantes, on s'attend à ce qu'une variation du taux de chômage entraîne une variation du taux de crime dans le même sens.

Ehrlich exprima également que la médiane du revenu moyen dans une communauté a un impact sur le taux de crime, en particulier sur le taux de crime à la propriété, car il mesure les bénéfices à commettre un crime : b ('opportunities provided by potential victims of crime').

Selon les criminologues, la distribution en âge de la population a une grande influence sur le taux de crime.

Il semble que le taux de crime est impacté principalement par la proportion de jeunes hommes.

On s'attend donc à ce que le taux de jeunes soit en relation directe (positive) avec le taux de crime.

Les travaux de Becker (1968) et Ehrlich (1973) montrent également l'effet négatif sur le taux de crime de la présence

policière, du nombre d'arrestations, et de la sévérité des condamnations.

Des variables sociétales comme la structure des ménages ou bien l'appartenance à certaines communautés semblent également avoir un impact sur le taux de crime mais ce ne sera pas étudié ici.

Dans l'étude de A.Oliver, la variable dépendante est le taux de crime national et les variables explicatives sont des variables économiques, sociales et liés à la répression (Deterrence) au niveau national.

Plusieurs modèles sont construits avec plus ou moins de variables.

Peu d'entre elles semblent significatives.

Cependant cette étude montre que l'impact du taux de jeunes et de la répression semble significatif sur le taux de crime.

L'auteur nous invite également à étudier les crimes par genres pour étudier plus finement les dynamiques en place.

1.2 Données utilisées

Dans le cadre de ce projet, nous avons récoltés les données sur les communes du Grand Paris pour l'année 2015.

Nous avons réalisé une jointure entre les dossiers complets l'INSEE pour les différentes communes et arrondissements de 2015, le dossier complet concernant les ménages français en 2015 (issu du recensement de la population) et les données de l'étude "Crimes et délits commis dans les communes du Grand Paris en 2015" de l'ONDRP (Observatoire National de la Délinquance et des réponses pénales).

Nous avons dû supprimer certaines villes où nous n'avions pas toutes les données.

Nous avons récoltés nos données pour 135 communes/individus de la région parisienne correspondant à la zone du Grand Paris.

Nous avons récolté les variables suivantes:

- Nom: Nom de la commune ou arrondissement
- Pop: Population de la commune/arrondissement
- NB_VPNC: nb de faits de violences physiques non crapuleuses
 - VPC: violences physiques crapuleuses
 - VSP : vols simples contre les particuliers
- Log : nb de logements
- med15: niveau de vie médian pour 2015 (en k)
 - Paris: indicatrice Paris
 - CR: nb Cambriolages
- CRP: cambriolages en résidences principales
- CRL: cambriolages (principale+secondaire)
 - VA : vols d'automobiles
- DD: destructions et dégradations
- ILS: infractions à la législation sur les stupéfiants
 - y_i: somme de tous les crimes confondus
- taux_delit: taux tous crimes confondus pour 1000 habitants
- taux_etabl_soc: nb d'établissements sociaux pour 1000 habitants
 - taux_chom: taux de chômage (en %)
 - taux_sco: % des actifs ayant au moins le BEPC
 - taux_pauv: taux de pauvreté des habitants de la commune
- taux15_29: % des 15-29 ans dans la population de la commune

L'indicateur des violences physiques non crapuleuses (**VNPC**) est composé : des homicides (hors homicides pour voler et règlements de compte), des coups et blessures volontaires mortels et non mortels, des violences à dépositaire de l'autorité, des mauvais traitements à enfants et des séquestrations prises d'otages.

L'indicateur des violences physiques crapuleuses (**VPC**) est composé: des homicides pour vol et des règlements de compte, des vols avec violences avec arme (à feu ou blanche) et les vols avec violences commis sans arme.

L'indicateur med15 correspond à la médiane du niveau de vie .À titre d'informations , le niveau de vie correspond au revenu disponible divisé par le nombre d'unité de consommation

2 Démarche

2.1 Objectif

Notre objectif dans le cadre de ce projet est d'étudier les variables qui déterminent la criminalité du Grand Paris. Pour ce faire, nous allons tenter de répondre à 3 questions:

1. Quels sont les variables qui déterminent de manière significative la criminalité du Grand Paris?
2. Il y a t'il une différence significative entre Paris et le reste du Grand Paris ?
3. Les différents types de crimes sont-ils déterminés par les mêmes variables?

2.2 Modélisation

Pour débiter, comme dans l'article de A.Oliver, nous prendrons le taux de crime (tous types confondus) de chaque commune pour variable dépendante et les variables explicatives économiques et sociales, justifiées via le cadre de la théorie de Becker et Ehrlich, que nous avons pu récolter, en s'approchant dans la mesure du possible des variables discutées dans l'introduction.

Ces variables sont mesurées pour 2015 et sont :

1. le taux de scolarisation
2. le niveau de vie médian
3. le taux de chômage
4. le taux de jeunes (15-29)
5. le taux d'établissements sociaux (ETOQ15 INSEE)
6. le taux de pauvreté

Nous avons été contraints de choisir des alternatives en raison de la difficulté à accéder à certaines données.

Notamment nous n'avons pas trouvé les données voulues sur la répression des crimes.

Nous avons décidés d'intégrer le taux d'établissements sociaux pour essayer de mesurer l'investissement dans l'action publique et sociale.

Nous avons été forcés de prendre le taux de 15-29 ans au lieu du 15-25ans souhaité car nos données ne permettaient pas un découpage plus fin.

Nous avons pris le taux de pauvreté afin de mesurer les inégalités de revenu (pour remplacer l'indice GINI utilisé dans l'étude de A.Oliver).

3 Analyse descriptive

3.1 Variables

On étudie les agrégateurs classiques en statistiques.

taux_delit	etabl_soc	taux_chom	taux_sco
Min. : 17.33	Min. : 33.0	Min. : 0.0736	Min. : 0.4620
1st Qu.: 27.49	1st Qu.: 222.5	1st Qu.: 0.0995	1st Qu.: 0.6613
Median : 34.25	Median : 361.5	Median : 0.1187	Median : 0.7395
Mean : 43.31	Mean : 737.2	Mean : 0.1351	Mean : 0.7269
3rd Qu.: 44.57	3rd Qu.: 843.0	3rd Qu.: 0.1646	3rd Qu.: 0.8205
Max. : 309.19	Max. : 4629.0	Max. : 0.2695	Max. : 0.8880
taux_pauv	tauxl5_29	SUPERF	MED15
Min. : 0.0500	Min. : 0.1520	Min. : 1.000	Min. : 12.91
1st Qu.: 0.0970	1st Qu.: 0.1870	1st Qu.: 3.000	1st Qu.: 18.51
Median : 0.1485	Median : 0.1995	Median : 4.000	Median : 22.90
Mean : 0.1731	Mean : 0.2017	Mean : 5.423	Mean : 23.83
3rd Qu.: 0.2225	3rd Qu.: 0.2137	3rd Qu.: 7.000	3rd Qu.: 28.54
Max. : 0.4550	Max. : 0.2940	Max. : 22.000	Max. : 43.35
Paris	taux_etabl_soc	densite	
Min. : 0.0000	Min. : 5.095	Min. : 1175	
1st Qu.: 0.0000	1st Qu.: 8.295	1st Qu.: 5105	
Median : 0.0000	Median : 10.764	Median : 8500	
Mean : 0.1538	Mean : 13.364	Mean : 10872	
3rd Qu.: 0.0000	3rd Qu.: 14.762	3rd Qu.: 14208	
Max. : 1.0000	Max. : 62.310	Max. : 38365	

Figure 1: summary(variables)

Pop	NB_VPNC	taux_VPNC	VPC	taux_VPC
4.495581e+04	2.764134e+02	3.102574e+00	3.698352e+02	5.013161e+00
VSP	taux_VSP	Log	CR	CRP
1.263041e+03	2.217147e+01	2.652239e+04	1.960132e+02	1.629038e+00
CRL	VA	taux_VA	DD	taux_DD
4.039442e+00	8.226702e+01	NA	3.074401e+02	2.945929e+00
ILS	taux_ILS	y_i	taux_delit	etabl_soc
2.496327e+02	4.042928e+00	2.502590e+03	3.429701e+01	9.282061e+02
taux_chom	taux_sco	taux_pauv	tauxl5_29	SUPERF
4.648170e-02	1.069785e-01	9.570816e-02	2.462925e-02	3.718517e+00
MED15	Paris	taux_etabl_soc	densite	
6.663877e+00	3.621970e-01	8.613791e+00	8.385538e+03	

Figure 2: Écarts types

	taux_delit	etabl_soc	taux_chom	taux_sco	taux_pauv	taux15_29	SUPERF	MED15	Paris
taux_delit	1.00000000	0.2440622	0.08683512	0.07608215	0.12556091	0.43400531	-0.09293958	0.11228282	0.5589980
etabl_soc	0.24406215	1.00000000	-0.16675758	0.38579036	-0.11200622	0.49144476	0.22943394	0.33796492	0.8188825
taux_chom	0.08683512	-0.1667576	1.00000000	-0.89212354	0.95847718	0.22138222	0.22356349	-0.82968152	-0.1873930
taux_sco	0.07608215	0.3857904	-0.89212354	1.00000000	-0.87568393	0.01225343	-0.26723738	0.90528747	0.4213964
taux_pauv	0.12556091	-0.1120062	0.95847718	-0.87568393	1.00000000	0.26146752	0.19139653	-0.83743880	-0.1335202
taux15_29	0.43400531	0.4914448	0.22138222	0.01225343	0.26146752	1.00000000	-0.01277778	-0.03446046	0.6455919
SUPERF	-0.09293958	0.2294339	0.22356349	-0.26723738	0.19139653	-0.01277778	1.00000000	-0.21071391	-0.0141678
MED15	0.11228282	0.3379649	-0.82968152	0.90528747	-0.83743880	-0.03446046	-0.21071391	1.00000000	0.3874836
Paris	0.55899805	0.8188825	-0.18739296	0.42139641	-0.13352015	0.64559188	-0.01416780	0.38748364	1.00000000
taux_etabl_soc	0.71175884	0.5266744	-0.43971004	0.64751532	-0.41401972	0.44423008	-0.18476590	0.69002867	0.7568226
densite	0.19281896	0.6740992	-0.13055741	0.38769469	-0.06376039	0.47968514	-0.30119993	0.25049034	0.6283672
	taux_etabl_soc	densite							
taux_delit	0.7117588	0.19281896							
etabl_soc	0.5266744	0.67409916							
taux_chom	-0.4397100	-0.13055741							
taux_sco	0.6475153	0.38769469							
taux_pauv	-0.4140197	-0.06376039							
taux15_29	0.4442301	0.47968514							
SUPERF	-0.1847659	-0.30119993							
MED15	0.6900287	0.25049034							
Paris	0.7568226	0.62836715							
taux_etabl_soc	1.0000000	0.44028619							
densite	0.4402862	1.00000000							

Figure 3: Matrice de corrélation

On observe de fortes variations pour le taux de délit sur l'ensemble du Grand Paris.

Tandis que 7 observations possèdent un taux de délit supérieur à 100 délits pour 1000 habitants, 6 sont des arrondissements de Paris, avec un maximum à 309 crimes par habitant pour le 1er arrondissement.

De l'autre côté du spectre, 60% des communes possèdent un taux de délit inférieur à 40 délits pour 1000 habitants.

Ce constat appuie notre intuition d'une possible discrimination entre Paris et le reste du Grand Paris.

De même on observe, de très fortes disparités au niveau de l'éducation, on observe une fort taux d'éducation (72%) sur la majeure partie du Grand Paris mais certaines communes accusent un retard certain.

Ces mêmes communes accusent généralement un taux de pauvreté et de chômage élevé.

Cependant, il ne semblerait pas que cela ait un impact sur le taux de jeunes.

Une forte densité semble généralement associée avec une population jeune.

On observe des écarts qui vont du quitte au double (de 15% à 30%).

Les communes (hors Paris) comptabilisant les plus hauts niveau de vie médian (med15) sont bien moins denses que les autres communes.

Ce ne sont pas les communes ou arrondissements les plus sujets aux crimes et possèdent de faibles taux de jeunes.

Sans surprise, on observe une forte corrélation entre le taux de chômage et le taux de pauvreté, il faudra en prendre compte dans le modèle.

Le taux de délit n'est pas fortement corrélé avec les autres variables à l'exception du taux d'établissements sociaux, ceci sera discuté plus loin.

La densité est fortement corrélée avec les établissements sociaux et le taux de jeunes, on observe surtout pour ces points l'influence du statut particulier de PARIS.

On observe une forte corrélation entre le niveau de vie médian, il faudra étudier si il y a colinéarité.

On observe une très faible corrélation entre le taux de jeunes et le taux de scolarisation.

On observe une corrélation négative entre le taux de pauvreté et le taux d'établissements sociaux.

Enfin Paris est très corrélée avec le taux de scolarisation et le taux d'établissements sociaux.

Il est intéressant d'observer que les différents types de crimes n'opèrent pas dans les mêmes typologies de communes.

3.2 Typologie de communes

Il pourrait être intéressant d'étudier la distance à Paris des communes.

La majorité des crimes commis à Paris, qui concentre les plus hauts taux de délits de la région étudiée sont des VSP (vols simples de particuliers), ce qui pourrait ne pas être sans lien avec la densité, le niveau de vie (social et économique) et le statut particulier de la ville de Paris.

Par ailleurs, les communes ayant un taux élevé de VSP (50-60%) sont très proches de Paris.

Les villes de densité plus faibles et moins denses semblent beaucoup moins sujettes aux VSP (10-20%).

Les crimes liés aux stupéfiants (ILS), aux dégradations (DD) sont davantage localisés en périphérie (nord) de Paris,

principalement dans les communes avec des taux de chômage élevés et une scolarisation faible, même si certaines communes et arrondissements frontaliers en concentrent également.

Les cambriolages (CRL) sont davantage recensé dans les villes aux revenus moyens, de faibles taux de chômage et quelque peu éloignées de Paris.

De même pour les vols d'automobiles (VA), ce sont les villes situées à la périphérie du Grand Paris qui sont les plus touchées.

La plus faible densité, donc une plus faible congestion et l'éloignement à Paris peut laisser penser que cela rend ce type de crime plus aisé à réaliser.

4 Modélisation et Interprétation pour le Grand Paris

4.1 Spécification

4.1.1 Transformation en log de la variable `taux_delit`

Nous disposons de la variable continue `taux_delit` ,pour que notre étude puisse se dérouler , il nous est nécessaire de savoir si l'on doit effectuer une transformation logarithmique de la variable `taux_delit` ou non. Pour décider de la transformation (ou non) de la variable `tauxdelit` , nous avons décidé de suivre une démarche visuelle. Dans un premier temps ,il serait intéressant d'obtenir une visualisation de la distribution de cette variable continue.

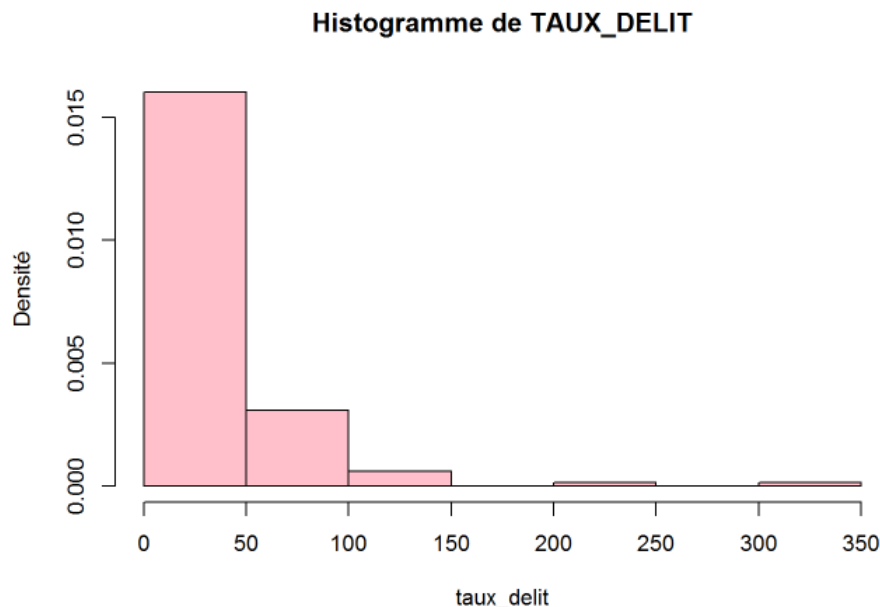


Figure 4: Histogramme de la distribution de `taux_delit`

On peut donc constater que cette variable continue a une longue queue de distribution ,de ce fait, il peut s'avérer difficile d'observer des liens entre les variables explicatives du modèle et `taux_delit` .

Ainsi,il semble nécessaire d'effectuer une transformation en log de notre variable dépendante ,la distribution de la nouvelle variable "ressemblera" plus à la courbe de densité d'une variable aléatoire gaussienne à valeurs dans \mathbb{R} .

En conséquence ,il sera plus aisé d'observer des liens entre une variable explicative et `taux_delit`.

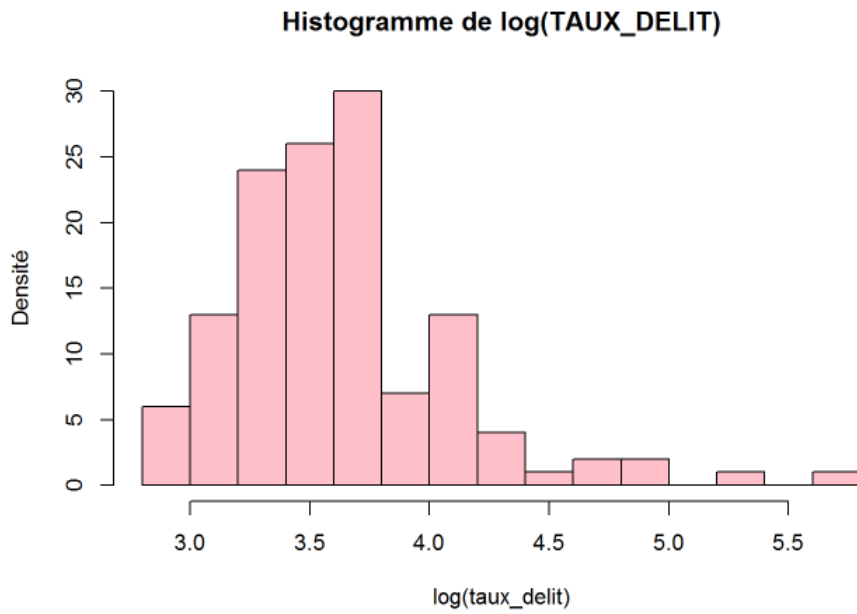
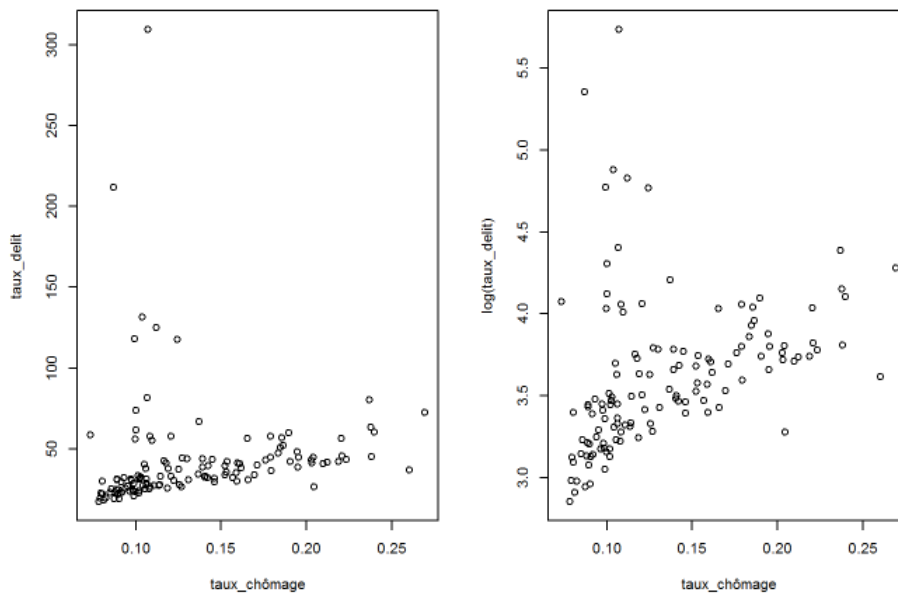


Figure 5: Histogramme de la distribution de $\log(\text{taux_delit})$

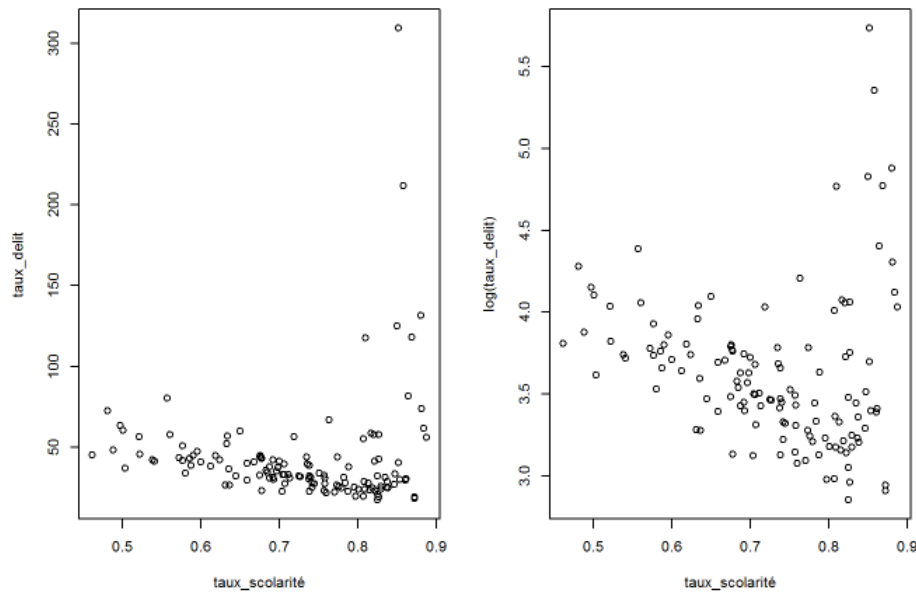
Pour montrer l'impact de la transformation logarithmique, pour chaque variable explicative, nous allons représenter les couples $(x_{ij}, y_i) \in \mathbb{R}^2, i = 1, \dots, 130, j$ fixé, y_i est égale à taux_delit_i .

Comparaison selon la variable taux_chom



On peut déjà apercevoir une relation linéaire entre la variable taux_delit et taux_chom , cependant, on peut constater que la relation linéaire entre taux_chom et $\log(\text{taux_delit})$ semble plus évidente à voir.

Comparaison selon la variable taux_sco



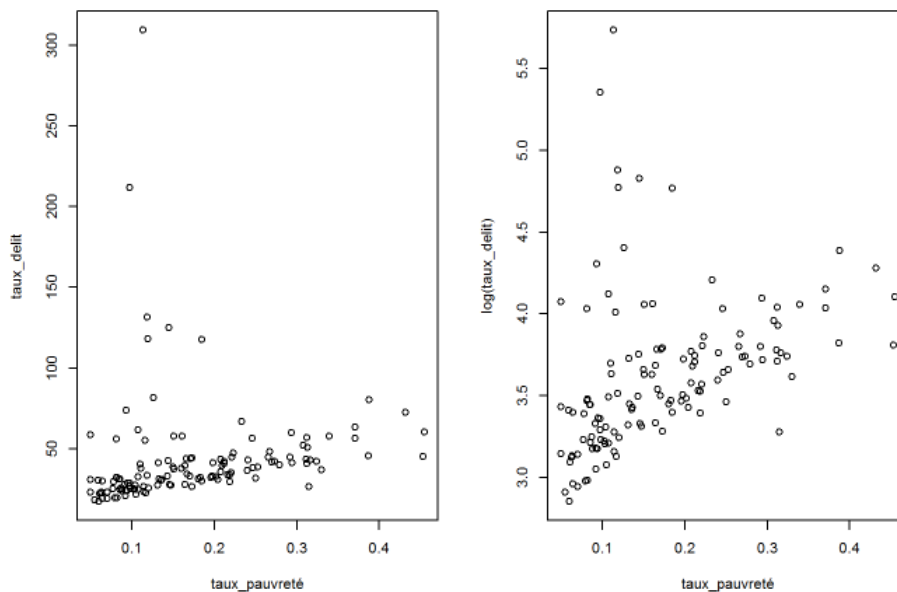
Ici , ce n'est pas vraiment le cas, la distinction n'est plus si évidente , la relation linéaire entre taux_sco et $\log(\text{taux_delit})$ ne semble pas plus intense que celle entre taux_sco et taux_delit .

Sur le 2e graphique , on peut voir plus aisément que des communes se distinguent des autres , ce sont des communes qui ont taux de criminalité élevé alors que le taux de scolarisation l'est aussi.

Nous reviendrons sur ce phénomène plus en détails par la suite.

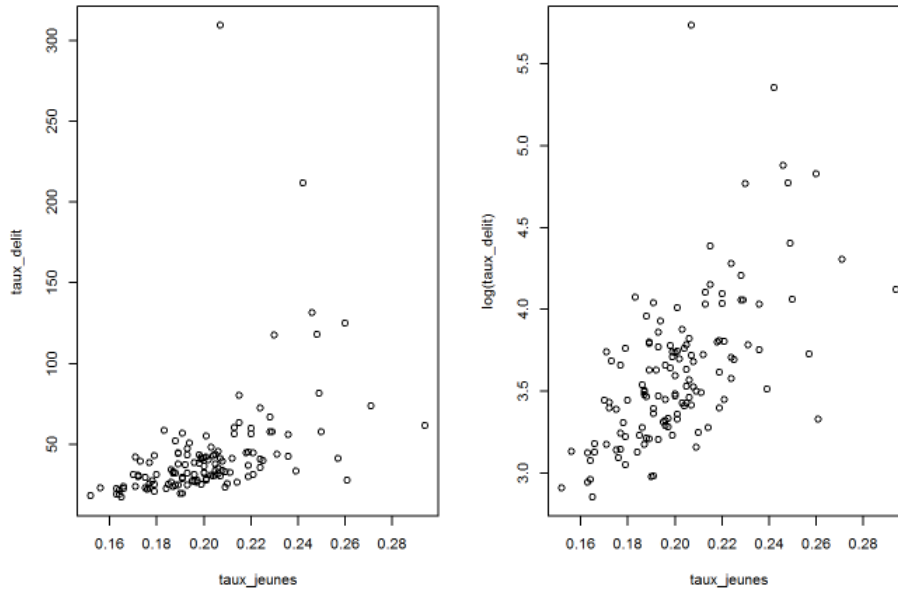
Cependant avec la raison qui a été donnée au préalable (longue queue de distribution), nous décidons de conserver tout de même la transformation logarithmique de taux_delit .

Comparaison selon la variable taux_pauv



On pouvait déjà remarquer une relation linéaire entre taux_delit et taux_pauv sur le premier nuage de points mais la relation linéaire (entre taux_pauv et $\log(\text{taux_delit})$) du deuxième nuage de points est bien plus évidente.

Comparaison selon la variable taux15_29



On constate le même résultat pour la variable `taux15__29`, la relation linéaire entre `taux15__29` et `log(taux_delit)` semble plus évident à voir que la relation linéaire entre `taux15__29` et `taux_delit`.

Ainsi , dans toute la suite de notre étude , grâce à l'analyse visuelle de la distribution d'une variable continue et de nuage de points , nous optons pour une transformation logarithmique de la variable continue `taux_delit`.

4.1.2 Transformation logarithmique des variables explicatives

Maintenant que nous avons effectué , une transformation logarithmique de notre variable dépendante soit `taux_delit` , il nous reste à savoir si l'on doit (oui ou non) effectuer la même transformation sur les variables explicatives.

On a donc à notre disposition deux modèles antagonistes :

$$\log(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \epsilon_i , i = 1, \dots, 130 \quad (1)$$

, ϵ_i un bruit aléatoire.

$$\log(y_i) = \beta_0 + \beta_1 \log(x_{i1}) + \beta_2 \log(x_{i2}) + \beta_3 \log(x_{i3}) + \beta_4 \log(x_{i4}) + \mu_i , i = 1, \dots, 130 \quad (2)$$

, μ_i un bruit aléatoire.

Dans ces modèles, `taux_chom` sera égale à X_1 , `taux_sco` à X_2 , `taux_pauv` par X_3 , `taux15__29` par X_4 et `taux_delit` par Y .

Pour choisir entre ces deux modèles , nous allons comparer le R^2 de chaque modèle , on choisira celui aura qui R^2 le plus grand, ceci est possible car les modèles ont été construits sur la base des mêmes observations , de plus , la variable à expliquer des modèles est identique (Daniel Dranove ,Northwestern University).

Le but n'est que de comparer les R^2 (aucune interprétation ne sera faite que ce soit par rapport au test de significativité et de validité globale ou par rapport aux coefficients estimés car le but ici est d'user de la méthode proposée par David Dranove).

```
##
## Call:
## lm(formula = log(taux_delit) ~ taux_chom + taux_sco + taux_pauv +
##      taux15_29, data = data_base)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.74839 -0.19995 -0.03875  0.11808  2.00489
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.08969    0.64691  -0.139  0.88996
## taux_chom    0.99876    2.53701   0.394  0.69449
## taux_sco     1.95437    0.74697   2.616  0.00998 **
## taux_pauv    2.67439    1.18510   2.257  0.02576 *
## taux15_29    8.42776    1.51360   5.568  1.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3489 on 125 degrees of freedom
## Multiple R-squared:  0.4584, Adjusted R-squared:  0.4411
## F-statistic: 26.45 on 4 and 125 DF,  p-value: 6.703e-16
```

Figure 6: Sorties du modèle (1)

```
##
## Call:
## lm(formula = log(taux_delit) ~ log(taux_chom) + log(taux_sco) +
##      log(taux_pauv) + log(taux15_29), data = data_base)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8301 -0.2277 -0.0427  0.1306  2.0272
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    7.8457    0.6948  11.292 < 2e-16 ***
## log(taux_chom)  0.2378    0.3739   0.636  0.526
## log(taux_sco)   0.6504    0.4685   1.388  0.168
## log(taux_pauv)  0.1950    0.1935   1.008  0.315
## log(taux15_29)  1.9543    0.3279   5.960  2.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3593 on 125 degrees of freedom
## Multiple R-squared:  0.4257, Adjusted R-squared:  0.4074
## F-statistic: 23.17 on 4 and 125 DF,  p-value: 2.43e-14
```

Figure 7: Sorties du modèle (2)

On peut donc constater que le R^2 du modèle (1) est plus grand que le R^2 du modèle (2) (le R^2 du premier modèle vaut 0.46 alors que celui du deuxième modèle vaut 0.42). De ce fait, selon la règle évoquée au préalable, on choisira la spécification du modèle (1) soit log-niveau.

5 Choix des variables du modèle

Les variables de notre modèle ont été choisies en se basant principalement les travaux de Becker et d'Ehrlich et plus précisément sur le bénéfice du crime que nous avons évoqué au préalable soit :

$$b - (pc + o)$$

,où b représente le gain (financier ou psychologique) lié au crime , p la probabilité de se faire capturé , c le coût de la punition et o le coût d'opportunité du crime.

Nous supposons que chaque variable (hormis le taux de jeunes)du modèle influe sur au moins une des 4 quantités soit b, p, c ou o .

Le taux de chômage

Dans ses travaux de 1973 , Ehrlich évoque le fait que le taux de chômage puisse avoir un impact sur la commission de crimes .Selon lui , le taux de chômage a une influence sur le coût d'opportunité lié au crime : si le taux de chômage est important alors cela réduit les gains espérés provenant d'une activité légale , de ce fait le coût d'opportunité du crime diminue et donc le bénéfice lié au crime ($b - (pc + o)$) augmente.

Il devient donc de plus en plus intéressant de s'adonner à une activité illégale ce qui implique une hausse du crime.

Au niveau des résultats de notre étude , on s'attendra donc à ce que `taux_chom` soit en relation directe (positive) avec `taux_delit`.

taux de scolarisation

Lors de la même année (1973),Ehrlich considère le rôle de l'éducation dans la commission de crimes.Selon lui , l'éducation peut amener 2 conséquences ,en effet , grâce à l'éducation ,les individus peuvent obtenir (en général) un meilleur salaire ,ainsi le coût d'opportunité du crime augmente ,cela fera donc diminuer le bénéfice lié au crime , la criminalité diminuera.

Cependant,le niveau d'éducation peut aussi accroître la criminalité en augmentant les opportunités pour les délinquants : si le niveau d'éducation augmente, alors les particuliers auront de meilleurs salaires mais cela implique une hausse du gain que le délinquant peut obtenir à travers un cambriolage ou un vol.

Cette dernière situation fait augmenter le bénéfice financier du criminel c'est-à-dire b et donc la criminalité augmente Lorsque le niveau d'éducation est élevé , cela traduit souvent par une économie plus prospère , au niveau des résultats de notre étude ,on s'attend à ce que `taux_sco` soit en relation indirecte (négative) avec `taux_delit`.

Ces 2 alternatives sont bien illustrées sur le nuage de points ,où l'on constate que plus le taux de scolarisation augmente plus le taux délit baisse mais aussi que si le taux de scolarisation est très élevé alors le taux de criminalité le sera aussi.

taux de jeunes

De manière générale , les criminologues considèrent que le taux de jeunes (dans une population donnée) est l'une des variables qui a la plus d'influence sur la criminalité. Il semble de difficile de préciser si il existe au moins une quantité ((b, p, c, o)) qui soit impacté par le taux de jeunes.Les 2 nuages de points qui ont été visionnés précédemment confirment le jugement des criminologues , en effet, sur les nuage de points , on constate que plus le taux de jeunes , plus le taux de criminalité augmente.

On s'attend à ce que `taux15_29` soit en relation directe avec `taux_delit`,en observant aux nuages de points ,on serait amené à penser que `taux15_29` et `taux_delit` sont en relation directe.

Dans les résultats de notre modèle , on souhaitera que le taux de jeunes influence de manière directe le taux de délit.

taux de pauvreté

En ce qui concerne le taux de pauvreté ,Isaac Ehrlich (en 1973) a mis en évidence une forte corrélation positive entre le taux de pauvreté et la criminalité (cela concernait surtout des crimes contre la propriété).

Bien évidemment , on ne peut pas en retirer une causalité .

Cependant , il serait intéressant de se demander si le taux de pauvreté influence directement le taux de criminalité c'est-à-dire que si le taux de pauvreté augmente alors cela va provoquer une augmentation de la criminalité par le biais d'une hausse de crimes contre la propriété.

En effet ,supposons qu'il y a une hausse du taux de pauvreté ,donc cela impliquerait une baisse des salaires et par la même occasion celle du coût d'opportunité du crime .

Cela fera augmenter le bénéfice lié au crime :il y aura donc une hausse du taux de criminalité.

Cependant , tout ce qui vient d'être dit n'est qu'hypothétique ,on ne peut déduire de causalité à partir des résultats d'Ehrlich (corrélation positive)

On s'attendra donc à ce que `taux_pauv` soit en relation directe avec `taux_delit`.

Notre hypothèse est justifiée par le nuage de point qui a été fait précédemment , on constate que plus le taux de pauvreté augmente , plus le taux de délit augmente.

6 Présence de multicollinéarité dans le modèle

6.1 La multicollinéarité et ses conséquences

Matrice des corrélations

Analyser la matrice des corrélations peut nous permettre de savoir si (peut-être) il existe des liaisons entre les variables explicatives de notre modèle.De plus , elle peut nous aider à quantifier l'intensité de la liaison entre 2 variables. En aucun cas , on ne cherchera dans cette section à évaluer l'influence d'une variable sur une autre (déduction de causalité).

```
##          taux_chom    taux_sco    taux_pauv    taux15_29
## taux_chom  1.0000000 -0.89082788  0.9584772  0.22138222
## taux_sco  -0.8908279  1.00000000 -0.8724155  0.01279429
## taux_pauv  0.9584772 -0.87241555  1.0000000  0.26146752
## taux15_29  0.2213822  0.01279429  0.2614675  1.00000000
```

Figure 8: Matrice des corrélations

On constate qu'il existe une forte corrélation linéaire positive, en effet, la variable `taux_pauv` est fortement corrélée (positivement) avec le `taux_chom`, la corrélation linéaire entre ces 2 variables est égale à 0.96 (ce qui est très élevée). Il existe aussi de fortes corrélations linéaires négatives entre nos variables explicatives, en effet, la corrélation linéaire entre `taux_sco` et `taux_chom` vaut -0.89 , celle entre `taux_pauv` et `taux_sco` est égale quant à elle à -0.87 .

On peut aussi remarquer que la variable `taux15_29` n'est fortement corrélée (linéairement) avec aucune variable explicative du modèle, la plus grande corrélation linéaire (en valeur absolue) entre `taux15_29` et les autres variables explicatives vaut 0.26 : ce n'est pas une valeur élevée.

De plus, on peut remarquer que `taux15_29` est faiblement corrélée (négativement) avec `taux_sco` avec une corrélation linéaire égale à -0.01 .

Multicolinéarité des variables explicatives

Précédemment, dans notre étude, nous avons pu obtenir les résultats du modèle (1) que nous avons décidé de garder. Parmi ces résultats, nous avons obtenu les coefficients estimés $\hat{\beta}_j, j = 0, \dots, 4$, ainsi que les tests de significativité de chaque paramètre $\beta_j, j = 0, \dots, 4$.

Chaque test de significativité (apparaissant dans ces sorties) prend en considération 2 hypothèses (nulle et alternative) qui s'écrivent de la manière suivante :

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

avec $j \in \{0, \dots, 4\}$.

L'hypothèse nulle sera rejetée si et seulement si la probabilité critique de chaque test est strictement supérieure à α .

Cependant, ce qui est le plus interpellant, ce sont les coefficients estimés du modèle, en effet, on constate que $\hat{\beta}_4$ est égale à 8.42, cela reviendrait donc à dire que si le taux de jeunes augmente de 1 point de % alors le taux de criminalité augmentera de 842% : cela semble contre-intuitif.

De plus, on constate que `taux_chom` n'est pas significatif (avec une p-value égale à 0.70), cela peut sembler assez surprenant au vu de la liaison linéaire que l'on a pu observer à l'aide des nuages de points.

Ces 2 problèmes que nous venons d'évoquer ne sont pas rares en présence de multicolinéarité entre les variables explicatives, de ce fait, on pourrait être amené à penser qu'il existe de la multicolinéarité (ou colinéarité) entre nos variables explicatives.

De plus, la multicolinéarité provoque des problèmes d'estimation, en effet, lorsque l'on se retrouve dans une situation où il y a une multicolinéarité, les variances des estimateurs $\hat{\beta}_j$ augmentent et donc les estimations deviennent instables.

Ainsi, il arrive (en cas de multicolinéarité) que les coefficients estimés $\hat{\beta}_j, j = 0, \dots, 4$ soient très éloignés des paramètres B_j du modèle (1).

Il aurait été pratique, de se baser sur la matrice de corrélation pour savoir quelles sont les variables qui semblent être colinéaires, cependant une corrélation linéaire forte entre 2 variables n'implique pas forcément qu'elles sont colinéaires. De plus, en regardant, seulement la matrice de corrélation, on ne pourrait pas détecter la présence de multicolinéarité. De ce fait, pour détecter la présence de colinéarité et de multicolinéarité entre les variables explicatives, nous allons calculer le facteur d'inflation de la variance pour chaque variable explicative $X_j, j = 1, \dots, 4$ (c'est une méthode qui permet de détecter les variables qui engendrent la présence de multicolinéarité).

Principe:

Pour chaque variable X_j , on réalise une régression avec les autres variables explicatives, on note R_j^2 le coefficient de détermination associé à cette régression.

Pour chaque variable X_j , le VIF_j est définie par :

$$\frac{1}{1 - R_j^2}$$

La valeur du VIF sera d'autant plus forte que X_j est combinaison linéaire des autres variables explicatives.

Cette valeur représente le degré auquel on peut dire qu'il existe une combinaison linéaire qui relie X_j et les autres variables explicatives.

Si $VIF_j > 10$, on suspectera X_j d'engendrer de la multicolinéarité dans notre modèle.

Calcul des VIF

```
## taux_chom  taux_sco  taux_pauv  taux15_29
## 14.737632  6.709983 13.634075  1.472797
```

Figure 9: Facteur d'inflation de la variance pour chaque variable explicative

Au vue des résultats , on peut donc supposer qu'il y a de la multicollinéarité dans notre modèle, en effet, il y a de fortes chances que `taux_pauv` soit combinaison linéaire des variables explicatives restantes (sans la compter) , il en est de même pour `taux_chom` . Ainsi , on se doit de modifier notre modèle à travers des suppressions et/ou des remplacements de variables explicatives pour éviter ce problème de multicollinéarité (ou de colinéarité).

6.2 Modification du modèle

Nous avons pu constater que $VIF_3 > 10$, cela nous laisse supposer que `taux_pauv` peut être combinaison linéaire des autres variables , nous allons donc remplacer `taux_pauv` par `med15` qui vaudra donc X_4 .

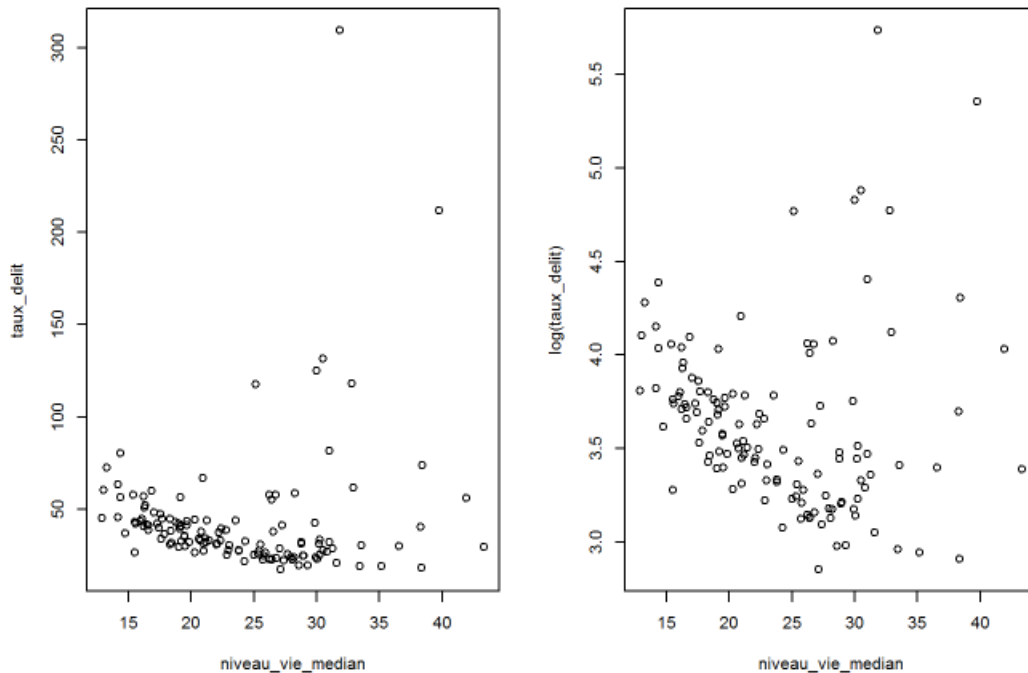
Cette variable influence aussi le bénéfice du crime , en effet , supposons que l'on ait une augmentation du niveau de vie médian (provenant d'une augmentation des revenus ayant pour origine une activité licite) alors le coût d'opportunité du crime augmente, il sera ainsi donc moins rentable de s'engager dans une activité illégale . On observera donc une baisse de la criminalité. Dans un autre sens , elle peut aussi contribuer à la hausse du crime (vols, cambriolages etc...) et ce dès lors que le niveau de vie est très élevé car le gain espéré par les délinquants augmentent ce qui fait augmenter b et donc le bénéfice lié au crime : la criminalité augmente.

On se doit donc de revoir la spécification du modèle , pour ce faire , il nous faut comparer les nuages de points des couples (x_{i4}, y_i) avec le nuage de points des couples $(x_{i4}, \log(y_i))$ pour $i = 1, \dots, 130$.

On préservera la même règle qu'au préalable : il y aura transformation en log de la variable Y soit de `taux_delit` .

Dans un premier , nous allons comparer les nuages de points des couples (x_{i4}, y_i) avec le nuage de points des couples $(x_{i4}, \log(y_i))$ pour $i = 1, \dots, 130$.

Comparaison avec la variable `med15`



Grâce aux observations , on constate que la relation linéaire entre `taux_delit` et `med15` est évidente , cependant elle l'est moins par rapport à la relation linéaire entre `med15` et $\log(\text{taux_delit})$. Ces nuages de points confirment la théorie qui a été évoquée au préalable.

Comme nous l'avions fait précédemment, nous allons comparer 2 modèles antagonistes :

$$\log(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \alpha_i, \quad i = 1, \dots, 130 \quad (3)$$

avec α_i un bruit aléatoire.

$$\log(y_i) = \beta_0 + \beta_1 \log(x_{i1}) + \beta_2 \log(x_{i2}) + \beta_3 \log(x_{i3}) + \beta_4 \log(x_{i4}) + \varsigma_i, i = 1, \dots, 130 \quad (4)$$

avec ς_i un bruit aléatoire.

Il nous est nécessaire de comprendre qu'ici med15 est identifié par X_4 . Nous serons amenés à choisir notre modèle en comparant R^2 , on peut effectuer cette comparaison grâce aux mêmes raisons que nous avons données pour comparer le modèle (1) et le modèle (2) (Daniel Dranove).

```
##
## Call:
## lm(formula = log(taux_delit) ~ (MED15) + (taux_sco) + (taux_chom) +
##      (taux15_29), data = data_base)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.80244 -0.18065 -0.06918  0.15459  1.99109
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.10307     0.66065   0.156  0.87628
## MED15        0.01851     0.01084   1.708  0.09019 .
## taux_sco     0.57716     0.91003   0.634  0.52710
## taux_chom    5.54175     1.71602   3.229  0.00158 **
## taux15_29    9.49580     1.46611   6.477 1.94e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3518 on 125 degrees of freedom
## Multiple R-squared:  0.4492, Adjusted R-squared:  0.4316
## F-statistic: 25.49 on 4 and 125 DF,  p-value: 1.887e-15
```

Figure 10: Sorties du modèle (3)


```
##
## Call:
## lm(formula = log(taux_delit) ~ log(MED15) + log(taux_sco) + log(taux_chom) +
##     log(taux15_29), data = data_base)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.80057 -0.20024 -0.07993  0.17271  1.95193
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.8832     1.1160   5.272 5.74e-07 ***
## log(MED15)        0.8276     0.3383   2.446 0.01582 *
## log(taux_sco)    -0.1801     0.5628  -0.320 0.74945
## log(taux_chom)    0.8436     0.2639   3.196 0.00176 **
## log(taux15_29)    1.9738     0.3060   6.450 2.21e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3524 on 125 degrees of freedom
## Multiple R-squared:  0.4475, Adjusted R-squared:  0.4298
## F-statistic: 25.31 on 4 and 125 DF,  p-value: 2.274e-15
```

Figure 11: Sorties du modèle (4)

On constate que le modèle (4) possède un R^2 égale à 0.4475 alors que le modèle (3) en possède un égale à 0.4492 . Ainsi ,si l'on use de la méthode que nous avons utilisée précédemment , on serait amené à choisir le modèle (3). Les coefficients estimés $\hat{\beta}_j$ du modèle (3), en effet, $\hat{B}_4 = 9.49$, cela voudrait donc dire que si le taux de jeunes augmente de 1 point de % alors cela provoquerait une augmentation du taux délit de 949% : cela n'est pas intuitif et dénoué de tout sens . Il en est de même pour le coefficient $\hat{\beta}_3$, ce n'est pas une "bonne" estimation de β_3 . On optera donc pour le modèle (4) qui possède un R^2 proche du modèle (3) tout en ayant des coefficients $\hat{\beta}_j$ qui semblent moins contre-intuitifs en vue d'une interprétation .

Remarque:

Observons les corrélations de log(med15) avec les 3 autres variables du modèle.

```
## log(taux_chom) log(taux_sco) log(taux15_29)
##      -0.9007668      0.9243308      -0.1080210
```

On constate que log(med15) est fortement corrélée positivement avec log(taux_sco) (0.924) ,cette nouvelle variable est aussi fortement corrélée négativement avec log(taux_chom):-0.90.

De plus , sa corrélation linéaire avec log(taux15_29) est faible (-0.1) : c'est la seule variable explicative du modèle avec laquelle sa liaison est faible en valeur absolue.

Même si la corrélation n'implique pas forcément la colinéarité , il est important de savoir s'il y a présence de colinéarité (certaines corrélations dépassent 0.90 en valeur absolue).

Pour ce faire , on va utiliser la même méthode que précédemment: on va calculer le facteur d'inflation de la variance de chaque variable explicative.

```
##      log(MED15) log(taux_sco) log(taux_chom) log(taux15_29)
##      9.220025   8.035637   7.585161   1.371693
```

Figure 12: Les VIF du modèle (4)

Si on garde encore un seuil de 10 , on peut donc supposer qu'il n'y a pas de présence de multicolinéarité dans notre modèle d'étude.

7 Test de Spécification du modèle : Variable omise

Le fait de recourir à une spécification linéaire peut être questionné , oublier certaines caractéristiques importantes (pour notre modèle) pourrait avoir d'importantes conséquences.

En effet, le fait d'oublier une variable explicative est lourd de conséquences dès lors que cette variable omise est corrélée avec l'une des variables explicatives du modèle : les estimateurs des paramètres β_j seront biaisés .

Dans notre étude , nous allons utiliser le Ramsey RESET test qui est un test général d'erreur de spécification.

Pour cette étude , notre modèle est le suivant :

$$y_i = \beta_0 + \beta_1 \log(x_{i1}) + \beta_2 \log(x_{i2}) + \beta_3 \log(x_{i3}) + \beta_4 \log(x_{i4}) + \epsilon_i, i = 1, \dots, 130 \quad (a)$$

Le principe consiste à ré estimer le modèle en ajoutant les valeurs ajustées au carré \hat{y}_i^2 au modèle :

$$y_i = \beta_0 + \beta_1 \log(x_{i1}) + \beta_2 \log(x_{i2}) + \beta_3 \log(x_{i3}) + \beta_4 \log(x_{i4}) + \gamma \hat{y}_i^2 + \epsilon_i, i = 1, \dots, 130 \quad (b)$$

Si le modèle (a) est le vrai modèle (on dispose donc de la bonne spécification) alors cela implique que la variable \hat{Y}^2 du modèle (b) n'aidera pas à expliquer Y , de ce fait ,seules les variables explicatives du modèle (a) permettent d'expliquer la variable dépendante Y . Si tel est le cas , cela revient à dire que $\gamma = 0$.

Cependant , si le modèle est mal spécifié (au moins une variable est omise) , \hat{Y}^2 améliorera l'ajustement , ainsi on aura alors $\gamma \neq 0$

Pour le RAMSEY RESET TEST , on a 2 donc hypothèses antagonistes :

$$H_0 : \gamma = 0$$

$$H_1 : \gamma \neq 0$$

Sous H_0 , $R \sim \mathcal{F}_{1,124}$ où R est la statistique du RAMSEY RESET TEST.

Règle de décision:

Rejet de H_0 si et seulement si $\mathbb{P}(R > r) \leq \alpha$ avec $\alpha = 0.05$ et $r = \text{RESET}$ la valeur observée de la statistique R .

```
##  
## RESET test  
##  
## data: reg  
## RESET = 0.9159, df1 = 1, df2 = 124, p-value = 0.3404
```

Figure 13: Résultats du RAMSEY RESET TEST

On a que que la p-value est égale donc à $\mathbb{P}(R > 0.9159) = 0.3404$.

De ce fait, on ne rejette pas H_0 , on peut donc supposer qu'il n'y a pas de variable omise.

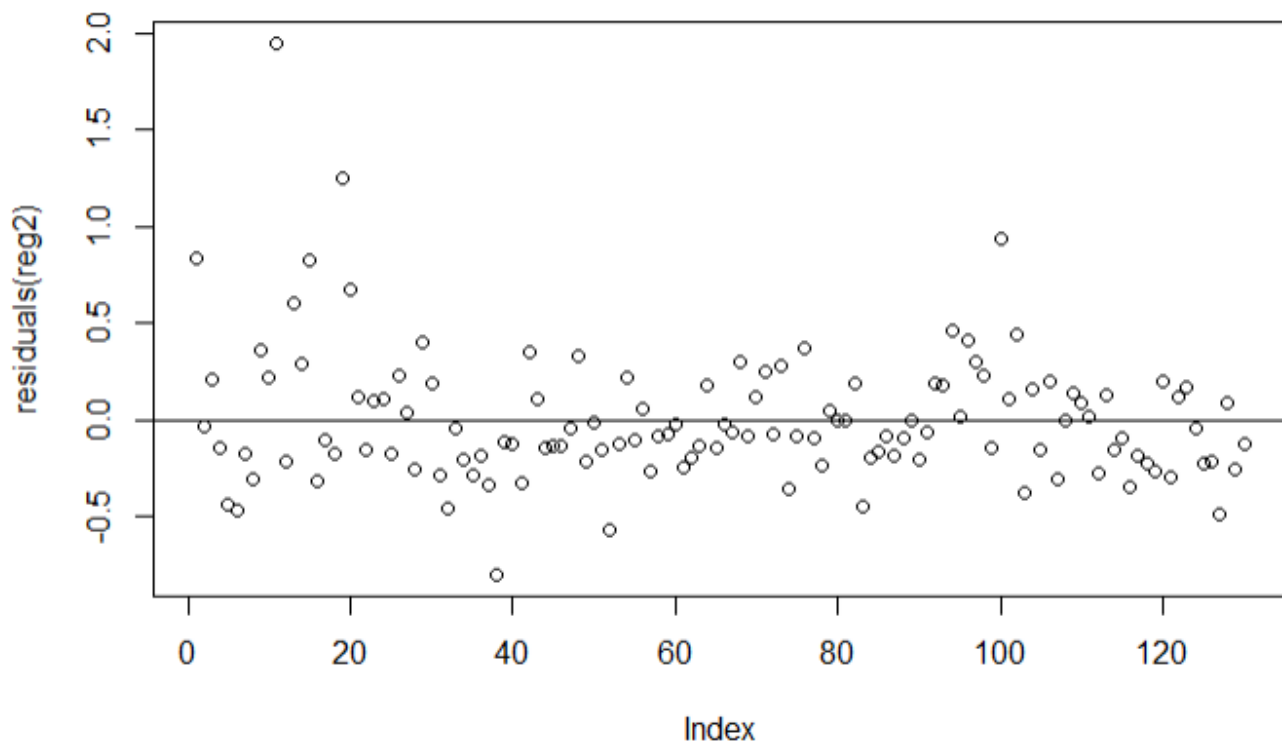
7.1 Étude de l'Hétéroscédasticité

On va étudier l'hétéroscédasticité des résidus de notre modèle. On réalise tout d'abord une étude graphique.

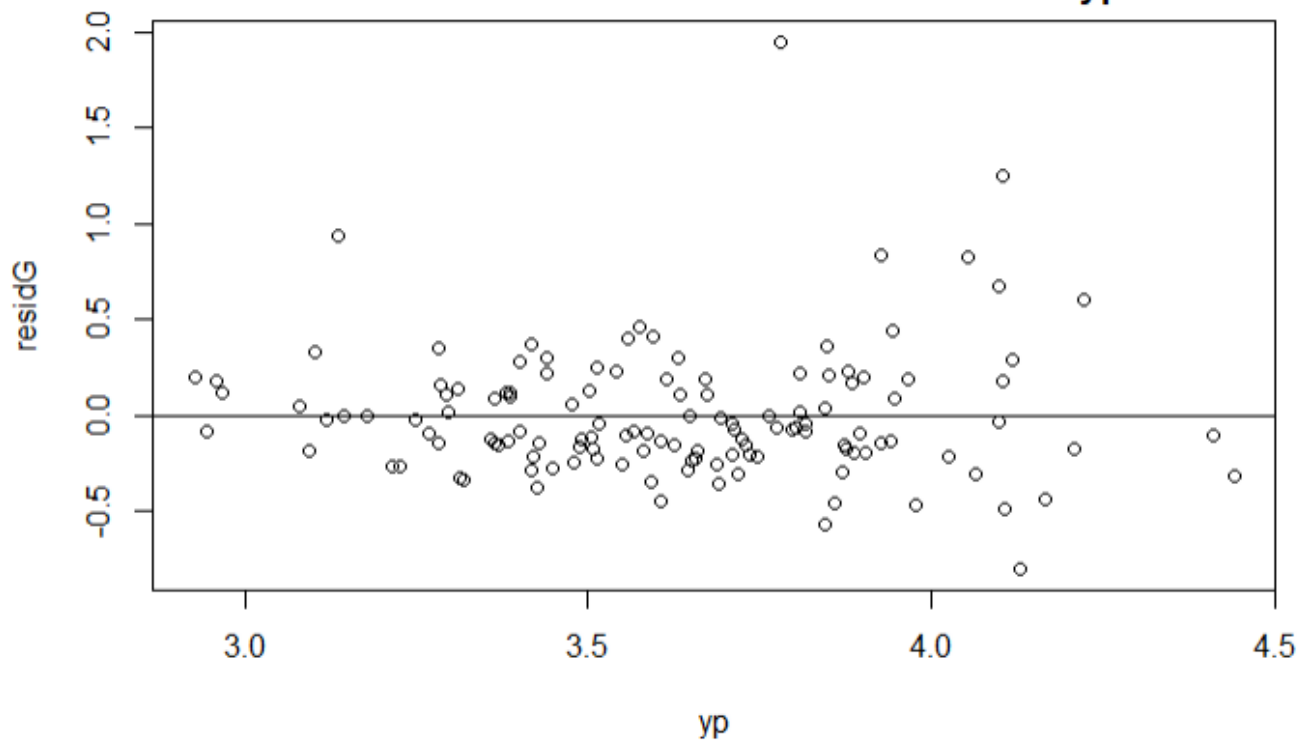
```
reg2=lm(log(taux_delit) ~ log(taux15_29)+log(MED15)+log(taux_sco)+log(taux_chom) , data=tab2)  
  
residG = residuals(reg2)  
residG2 = residG^2  
yp = fitted(reg2)  
  
par(mfrow=c(1,1))  
plot(residuals(reg2))  
abline(a=0,b=0)
```

Figure 14: Code du Logiciel R

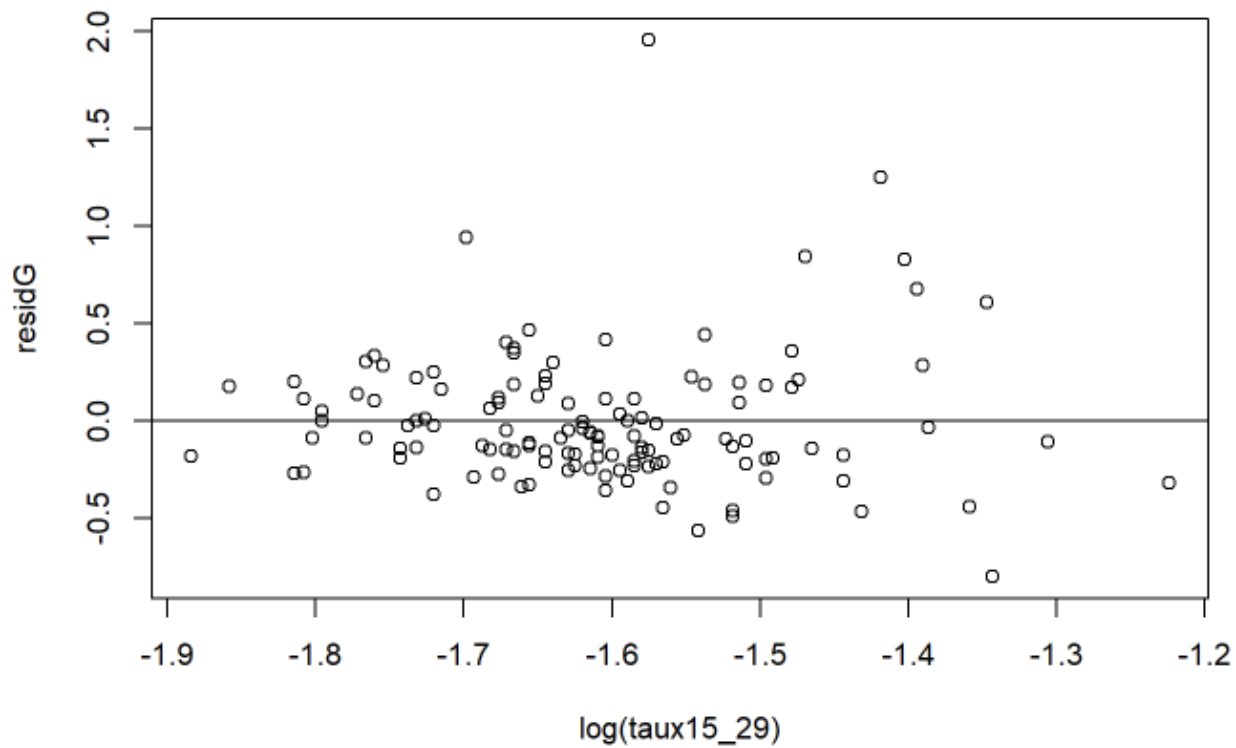
Si on observe quelques résidus importants, le reste semble assez homogène:



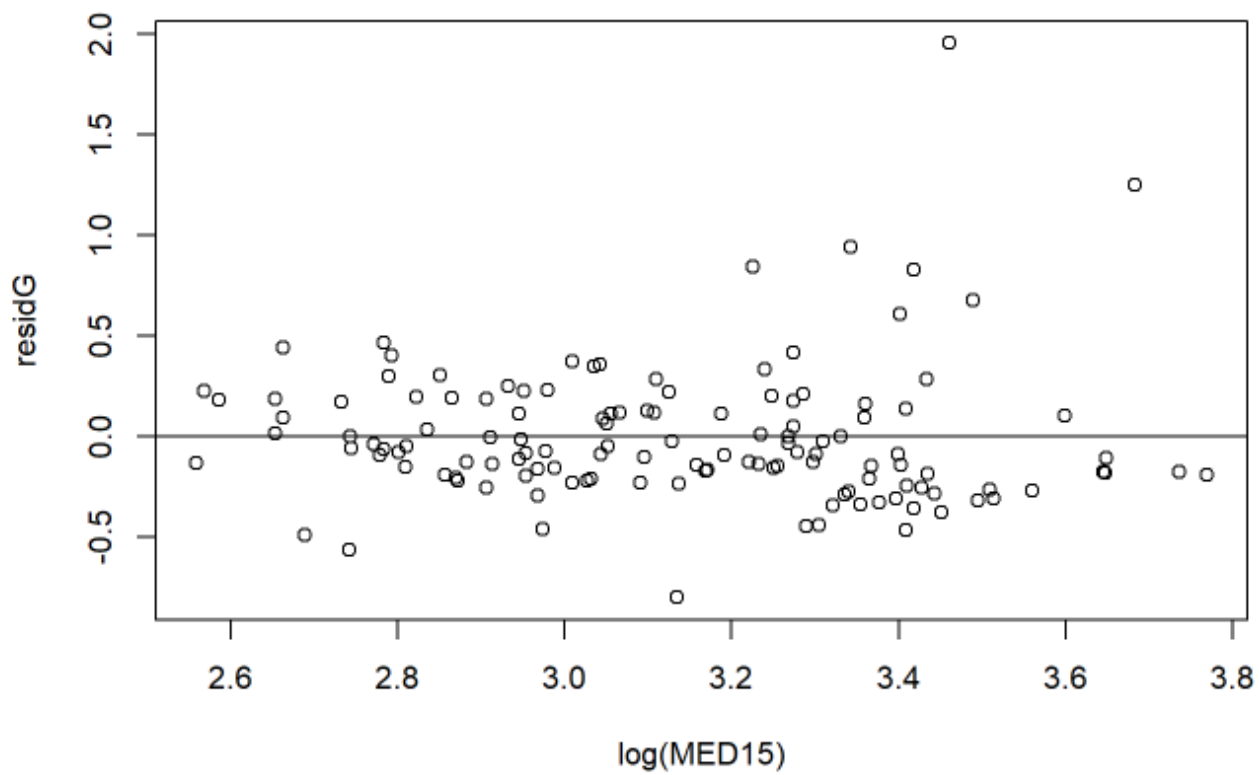
hétéroscédasticité - les résidus fonction des yp?



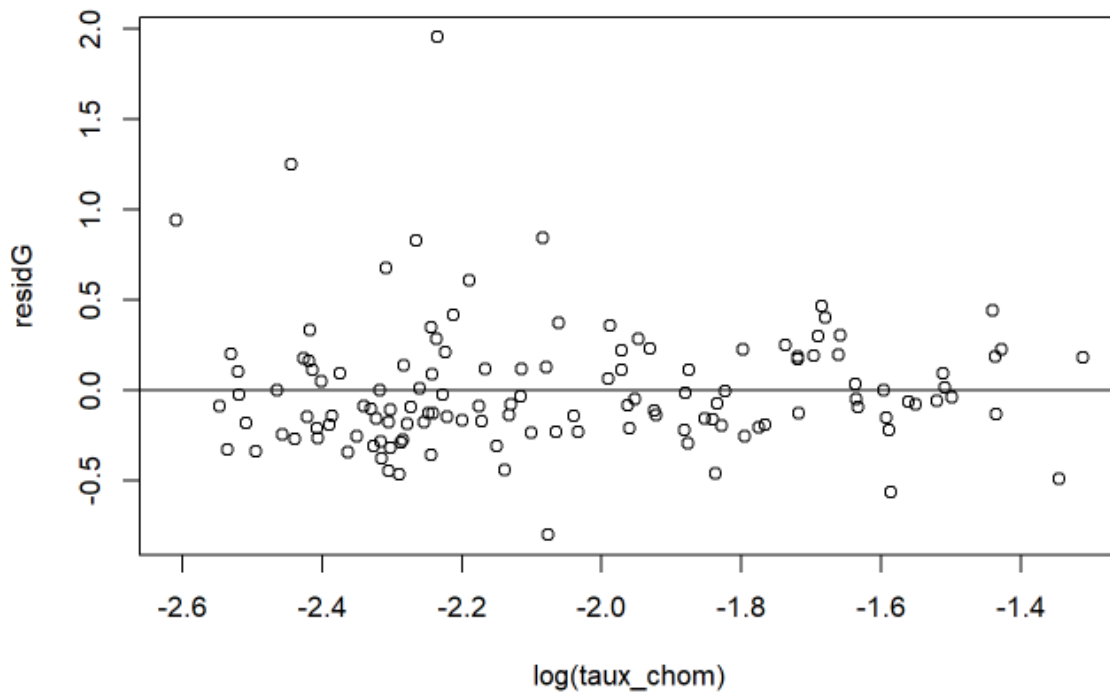
hétéroscédasticité - les résidus fonction du taux de jeune?



hétéroscédasticité - les résidus fonction de MED15?



hétéroscédasticité - les résidus fonction du taux de chômage?

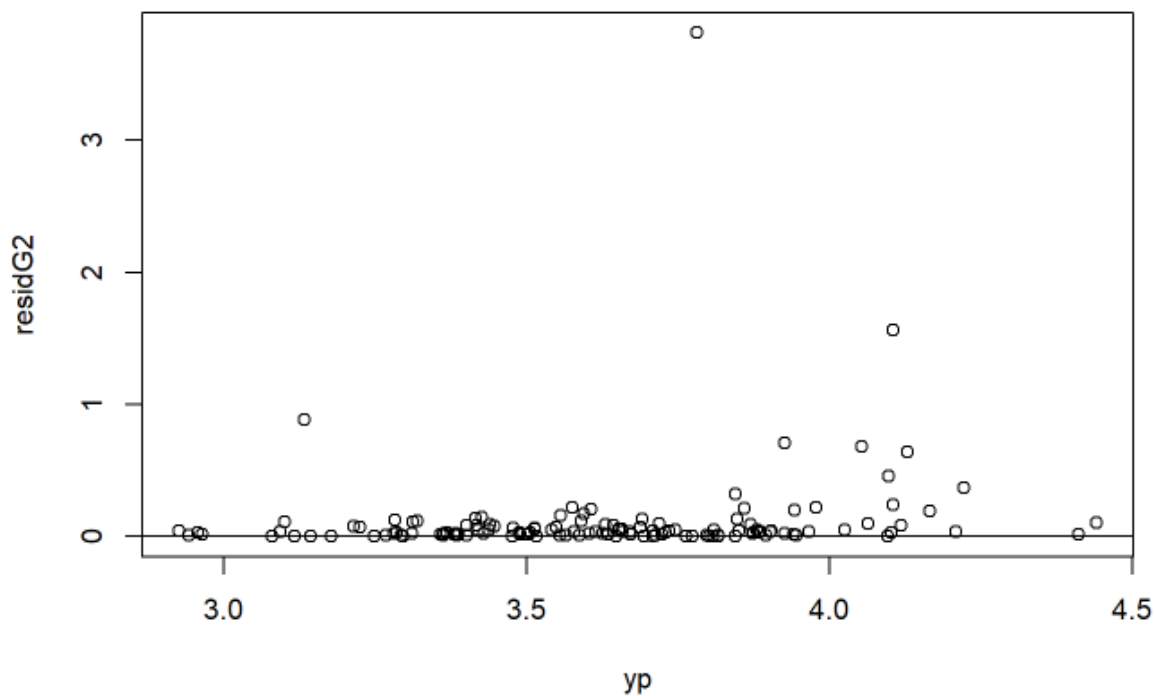


On observe la même chose pour toutes les variables.

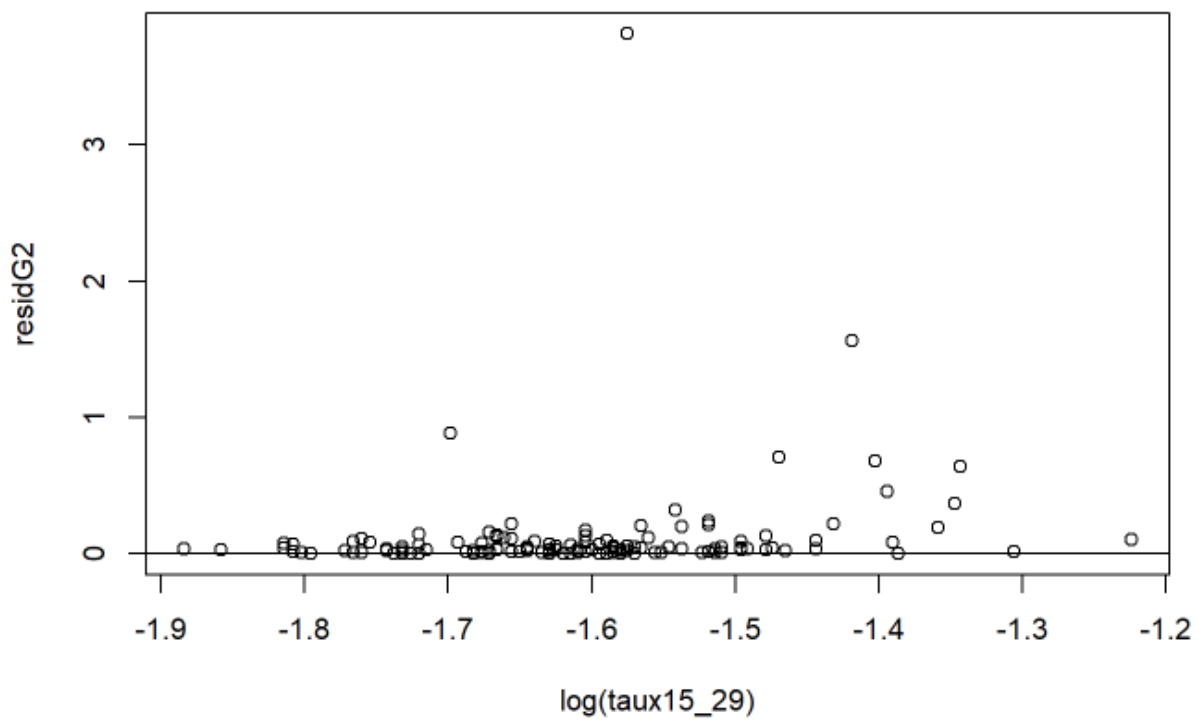
En dehors des quelques individus/résidus plus élevés, on peut supposer une homoscédasticité.

A t'on la variance en fonction d'une variable explicative?

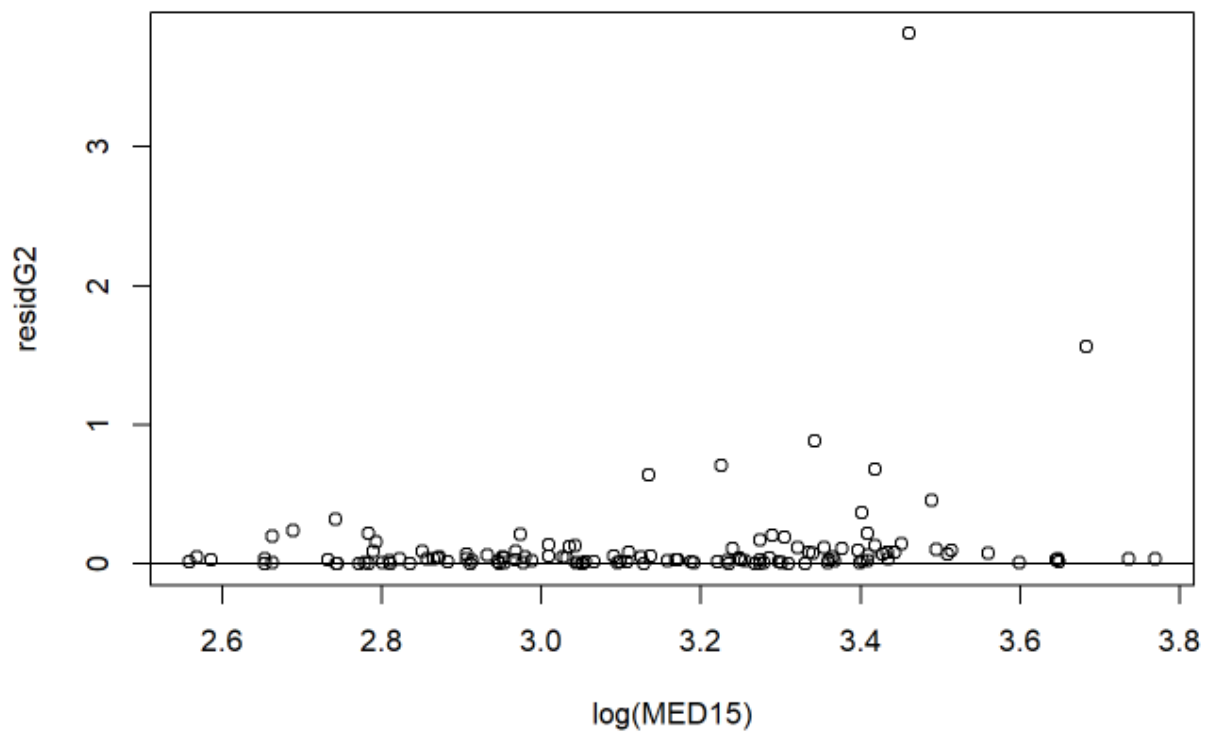
hétéroscédasticité- la variance est-elle fonction du yp



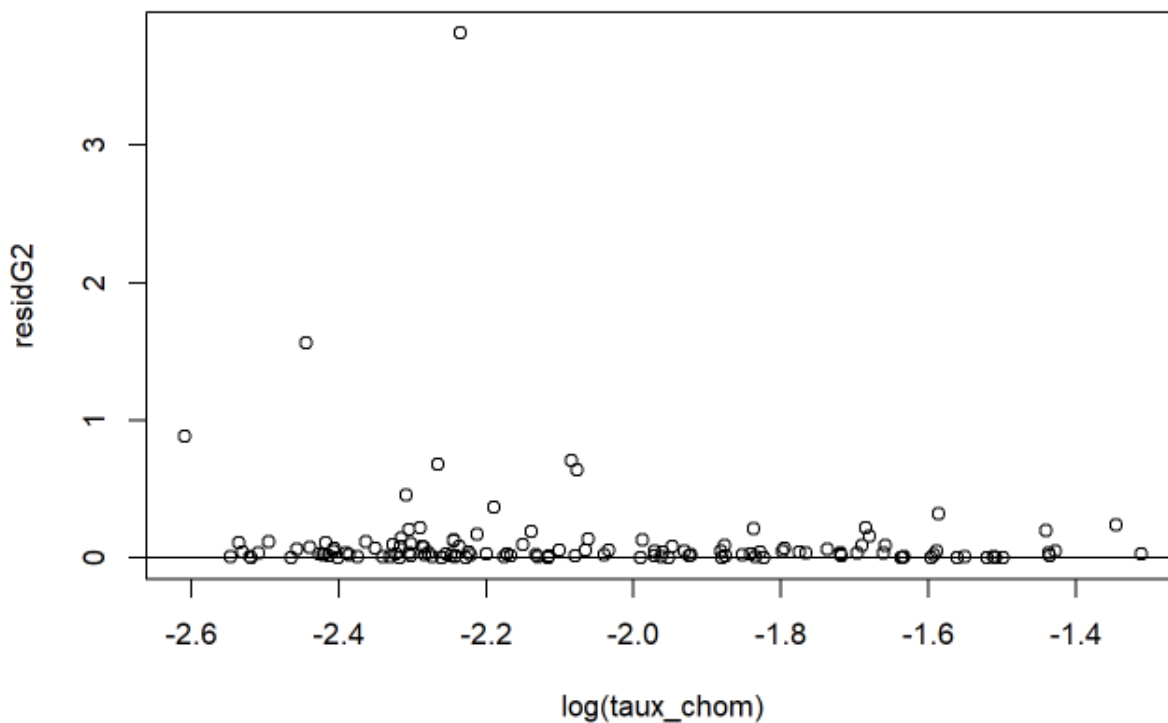
hétéroscédasticité- la variance est-elle fonction du taux de jeunes ?



hétéroscédasticité- la variance est-elle fonction du MED15 ?



hétéroscédasticité- la variance est-elle fonction du taux de chômage?



Il ne semble pas que la variance soit fonction des variables explicatives.

On ne va donc pas étudier l'hétéroscédasticité via le test de Goldfeld-Quandt.

Nous allons donc réaliser un test de White.

L'hypothèse H_0 est l'homoscédasticité des résidus:

On estime notre modèle et on récupère les résidus

On estime ensuite les résidus au carrés par les variables explicatives X_i , les X_i^2 et les produits croisés $X_i \times X_j$ de notre modèle et on récupère le R^2 .

Sous H_0 ,

$$N \times R_{\epsilon^2}^2 \simeq \chi^2(p)$$

où N =le nombre d'individus, p =le nombre de paramètres de la seconde régression

```
bptest(reg2,~log(taux15_29)+log(MED15)+log(taux_sco)+log(taux_chom)+log(taux15_29)^2+log(MED15)^2+log(taux_chom)^2+log(taux15_29)*log(MED15)+log(taux_sco)*log(taux_chom)+log(taux15_29)*log(taux_chom)+log(taux_sco)*log(taux15_29)+log(taux15_29)*log(taux_sco)+log(MED15)*log(taux_chom),data=tab2)
```

```
##
## studentized Breusch-Pagan test
##
## data: reg2
## BP = 12.911, df = 10, p-value = 0.2287
```

Nous observons une p-value de 0.22, nous décidons donc de ne pas rejeter l'hypothèse nulle, on suppose donc l'homoscédasticité des résidus.

Nous n'avons donc pas besoin de procéder à une correction de l'estimateur des MCO via les MCQG.

7.2 Interprétation de la régression

8 Interprétation du modèle log-log

Il est important de rappeler que nous avons décidé de conserver le modèle (4).

8.1 Validité globale du modèle

Le test de validité globale est associée aux 2 hypothèses antagonistes suivantes :

$$H_0 : \forall j \in \{1, 2, 3, 4\}, \beta_j = 0$$

$$H_1 : \exists j \in \{1, 2, 3, 4\}, \beta_j \neq 0$$

Soit F la statistique du test égale à $\frac{125}{4} \times \frac{R^2}{1-R^2}$ sous H_0 , $F \sim \mathcal{F}_{4,125}$.

La valeur observée pour la statistique F est égale $f = 25.31$.

Si notre test est de niveau $\alpha = 0.05$ alors on rejette H_0 car $\mathbb{P}(F > f) \leq \alpha = 0.05$ Le modèle est donc globalement satisfaisant.

8.2 Qualité d'ajustement

Malgré le fait qu'il n'y ait pas un consensus sur le seuil, on constate que la qualité d'ajustement n'est pas complètement médiocre mais n'est pas excellente : R^2 ajustée est égale à 0.4298, de plus le modèle permet d'expliquer 45% de la variance de la variable `taux_delit`.

8.3 Significativité des variables explicatives

En ce qui concerne le test de significativité, un rappel a déjà fait lors de notre étude.

Notre modèle peut se réécrire de la manière suivante :

$$Y = X\beta + \epsilon$$

, avec $\epsilon \in \mathbb{R}^{130}$, un vecteur gaussien centrée et de matrice de variance-covariance $\sigma^2 I_n$, $\sigma > 0$ inconnu. La matrice X est la matrice de design tel que $X = [\mathbf{1}_n, \log(X_1), \log(X_2), \log(X_3), \log(X_4)]$ et $\beta = [\beta_0, \beta_1, \beta_2, \beta_3, \beta_4]^t$, $n = 130$.

Pour un $j \in \{0, \dots, 4\}$ fixé, on a que la statistique T du test associé, sous H_0 , est égale à :

$$\frac{\hat{\beta}_j}{\hat{\sigma} \sqrt{(X^t X)^{-1}_{j+1}}}$$

Sous H_0 , on a que $T = \frac{\hat{\beta}_j}{s_j} \sim \mathcal{T}_{125}$, notons t_j la valeur prise par cette statistique T , on rejettera H_0 si et seulement si $\mathbb{P}(|T| > t_j) \leq \alpha$. On a que $\alpha = 0.05$.

On constate donc que $\log(X_1)$ est significative car la probabilité critique du test associé aux hypothèses $H_0 : \beta_1 = 0, H_1 : \beta_1 \neq 0$ est inférieure à 0.01 mais supérieure à 0.001 d'où son code de significativité *.

Il en est de même pour $\log(X_3)$ la variable est significative car la probabilité critique du test associé aux hypothèses $H_0 : \beta_3 = 0, H_1 : \beta_3 \neq 0$ est inférieure à 0.001 mais $p\text{-value} \gg 0$ d'où son code de significativité **.

La variable $\log(X_4)$ est significative car la $p\text{-value}$ du test associé aux hypothèses $H_0 : \beta_4 = 0, H_1 : \beta_4 \neq 0$ est très proche de 0 d'où son code de significativité ***.

De plus, la variable $\mathbf{1}_n$ est aussi significative car la $p\text{-value}$ du test associé aux hypothèses $H_0 : \beta_0 = 0, H_1 : \beta_0 \neq 0$ est inférieure à 0.01 mais supérieure à 0.001 d'où son code de significativité *.

Ainsi $\log(\text{med15})$ est une variable significative, il en est de même pour $\log(\text{taux_chom})$ ainsi que $\log(\text{taux15_29})$.

8.4 Coefficients estimés

Les coefficients estimés pour $\beta_1, \beta_3, \beta_4$ sont respectivement 0.8276, 0.8436, 1.9738.

En termes d'interprétation cela veut dire qu'une augmentation de 1% du niveau de vie médian fera augmenter le taux de délit de 0.827% toute chose égale par ailleurs.

Une augmentation de 1% du taux de chômage fera augmenter le taux de délit de 0.84%, toute chose égale par ailleurs.

.

De plus, une augmentation de 1% du taux de jeunes fera augmenter le taux de délit de 1.97%, toute chose égale par ailleurs.

Si l'on base sur les travaux d'Ehrlich et de Becker, on constate que ces résultats ne sont pas surprenants. Cependant, nous nous attendions à ce que le niveau de vie médian et le taux de délit soit en relation indirecte hormis celle-ci nous obtenons les relations souhaitées.

```
##
## Call:
## lm(formula = log(taux_delit) ~ log(MED15) + log(taux_sco) + log(taux_chom) +
##     log(taux15_29), data = data_base)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.80057 -0.20024 -0.07993  0.17271  1.95193
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.8832      1.1160   5.272 5.74e-07 ***
## log(MED15)        0.8276      0.3383   2.446 0.01582 *
## log(taux_sco)    -0.1801      0.5628  -0.320 0.74945
## log(taux_chom)    0.8436      0.2639   3.196 0.00176 **
## log(taux15_29)    1.9738      0.3060   6.450 2.21e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3524 on 125 degrees of freedom
## Multiple R-squared:  0.4475, Adjusted R-squared:  0.4298
## F-statistic: 25.31 on 4 and 125 DF,  p-value: 2.274e-15
```

Figure 15: Sorties R du modèle étudié

Remarque: Nous avons aussi pu obtenir une estimation de σ soit $\hat{\sigma} = 0.3524$.

8.5 Outliers

On observe que certains individus semblent être atypiques:

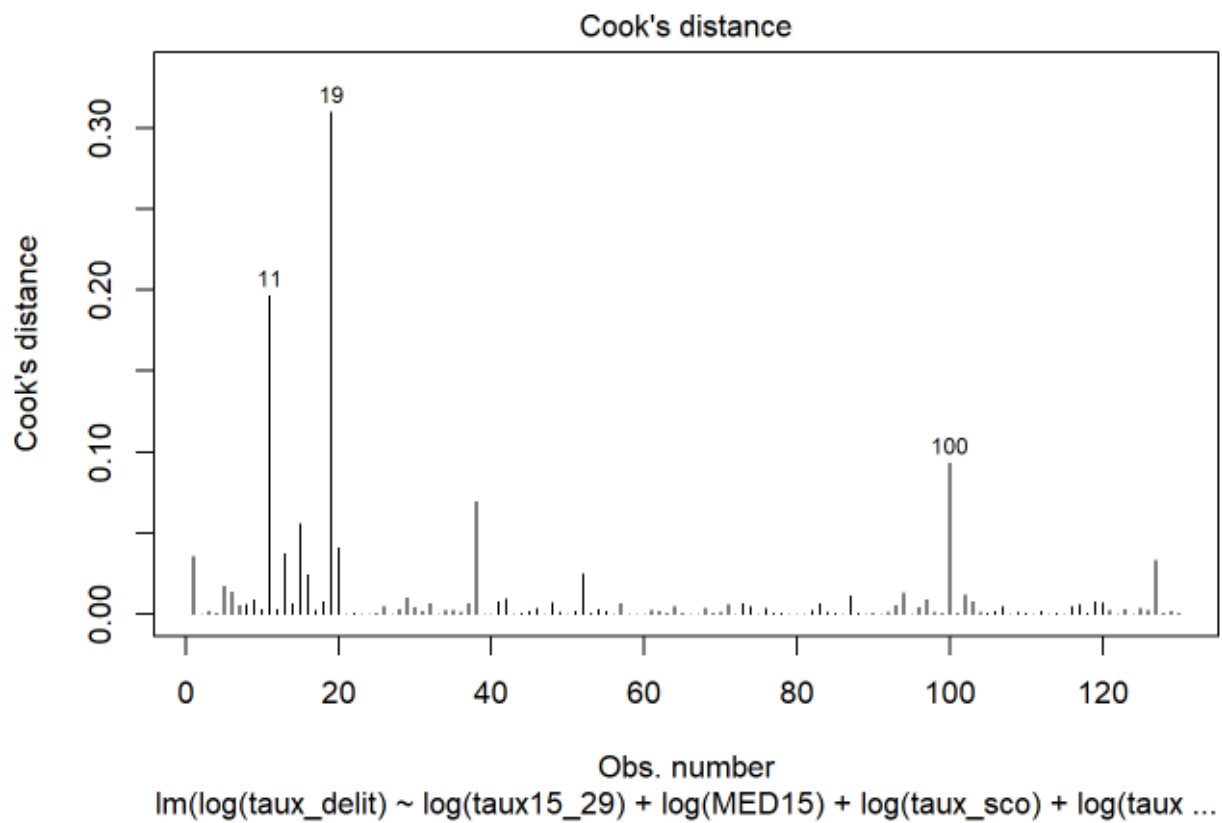


Figure 16: Distances de Cook

```
outliers=influence.measures(reg2)
print(outliers)
```

Figure 17: Détection des Outliers

#	cov.r	cook.d	hat	inf
## 1	0.843	3.60e-02	0.02972	*
## 2	1.090	1.11e-04	0.04540	
## 3	1.055	1.96e-03	0.02688	
## 4	1.061	8.97e-04	0.02539	
## 5	1.025	1.71e-02	0.04945	
## 6	1.004	1.37e-02	0.03631	
## 7	1.133	5.47e-03	0.09096	*
## 8	1.045	6.09e-03	0.03593	
## 9	1.039	9.12e-03	0.04073	
## 10	1.065	3.35e-03	0.03885	
## 11	0.249	1.97e-01	0.03012	*
## 12	1.066	3.34e-03	0.03947	
## 13	0.972	3.77e-02	0.05685	
## 14	1.063	6.83e-03	0.04729	
## 15	0.862	5.58e-02	0.04623	*
## 16	1.132	2.42e-02	0.11446	*
## 17	1.168	2.70e-03	0.11273	*
## 18	1.165	7.70e-03	0.11661	*
## 19	0.638	3.10e-01	0.09972	*
## 20	0.937	4.09e-02	0.05032	
## 21	1.053	3.67e-04	0.01614	
## 22	1.054	8.92e-04	0.02052	
## 23	1.055	2.41e-04	0.01655	
## 24	1.056	3.75e-04	0.01826	
## 25	1.051	9.77e-04	0.01923	
## 26	1.079	4.65e-03	0.05156	
## 27	1.214	2.99e-04	0.14254	*
## 28	1.049	3.34e-03	0.02918	
## 29	1.023	1.01e-02	0.03623	
## 30	1.101	4.07e-03	0.06581	
## 31	1.028	2.15e-03	0.01527	
## 32	0.989	6.79e-03	0.01896	
## 33	1.083	1.11e-04	0.03951	
## 34	1.059	2.22e-03	0.03051	
## 35	1.033	2.61e-03	0.01897	
## 36	1.051	1.26e-03	0.02103	
## 36	1.051	1.26e-03	0.02103	
## 37	1.037	6.84e-03	0.03408	
## 38	0.884	6.96e-02	0.05947	*
## 39	1.051	2.90e-04	0.01380	
## 40	1.051	3.91e-04	0.01460	
## 41	1.045	7.59e-03	0.03964	
## 42	1.047	9.79e-03	0.04566	
## 43	1.058	4.06e-04	0.02037	
## 44	1.046	4.50e-04	0.01226	
## 45	1.087	1.71e-03	0.04875	
## 46	1.150	3.53e-03	0.10076	*
## 47	1.065	9.48e-05	0.02326	
## 48	1.041	7.00e-03	0.03653	
## 49	1.045	1.44e-03	0.01891	
## 50	1.068	1.10e-05	0.02580	
## 51	1.087	2.12e-03	0.04968	
## 52	0.976	2.53e-02	0.04447	
## 53	1.057	5.51e-04	0.02048	
## 54	1.063	3.05e-03	0.03667	
## 55	1.125	1.66e-03	0.07846	*
## 56	1.052	7.29e-05	0.01194	
## 57	1.071	6.44e-03	0.05110	
## 58	1.074	3.73e-04	0.03323	
## 59	1.061	1.89e-04	0.02060	
## 60	1.093	4.78e-05	0.04772	
## 61	1.043	2.30e-03	0.02209	
## 62	1.057	1.73e-03	0.02736	
## 63	1.063	8.69e-04	0.02651	
## 64	1.115	4.59e-03	0.07638	
## 65	1.053	7.06e-04	0.01887	
## 66	1.065	2.97e-05	0.02281	
## 67	1.086	2.91e-04	0.04290	
## 68	1.035	3.60e-03	0.02382	
## 69	1.098	8.03e-04	0.05453	
## 70	1.089	1.09e-03	0.04849	
## 71	1.076	5.85e-03	0.05292	
## 72	1.054	1.46e-04	0.01403	
## 73	1.063	6.57e-03	0.04650	
## 93	1.134	5.60e-03	0.09220	*
## 94	1.003	1.31e-02	0.03504	
## 95	1.086	1.42e-05	0.04182	
## 96	0.998	3.99e-03	0.01397	
##				

Les principaux outliers sont les arrondissements parisiens possédant un (très) fort taux de délit (>80 pour 1000 habitant).

Les 4 principaux sont le 1^{er} arrondissement (11) et 8^{me} arrondissement (19), respectivement les 1^{re} et 2^{me} recensant les plus hauts taux de délits, très supérieurs aux autres communes.

Ces arrondissements, possèdent des taux de criminalité très importants, même comparé aux autres arrondissements de Paris qui généralement possèdent des taux élevés, ceci ne peut être expliqué par le modèle, seulement par la spécificité des dits arrondissements, hauts lieux du tourisme parisien où l'on retrouve sans être exhaustif : le musée du Louvre, le jardin des Tuileries, les boulevards Haussmannniens, les Champs-Élysées, Le palais de l'Élysée, etc ; où les crimes de type VSP (vols simples contre les particuliers) sont légion (62.5 % et 59% des délits commis respectivement).

Cachan (38), concentre un fort taux de jeunes, mais à priori surtout des étudiants, donc moins enclins à la criminalité d'après Ehrlich, ce qui peut expliquer la sur-estimation du taux de délit par le modèle.

Rungis (100), présente une population éduquée, avec un niveau de vie supérieur à la moyenne (med15 dans le Q3 du Grand Paris) et un plein emploi (7% chômage), donc la surestimation du modèle pour cette commune pourrait s'expliquer par la forte activité de l'industrie agro-alimentaire, qui pourrait instaurer un climat moins favorable à la criminalité.

Ces constatations corroborent une potentielle omission de certaines variables explicatives, notamment pour Paris Intra-Muros.

8.6 Critiques du modèle

Plusieurs points sont à évoquer concernant le modèle.

Un certain nombre de problèmes d'endogénéité semblent apparaître. D'une part, le taux de jeunes mesuré est la part de la population âgée de 15ans à 29 ans, cependant les études des criminologues semblent estimer que c'est la part de jeunes hommes âgés de moins de 25 ans qui influe sur le taux de crime.

Certaines études semblent évoquer une agressivité dès 13-14 ans.
On peut considérer ceci comme une légère erreur de mesure.

Le taux d'établissements sociaux, de santé et d'administration publique ne permet pas de mesurer clairement les inégalités mais il s'agit plutôt d'une mesure des investissements sociaux, donc une autre mesure du bénéfice au crime : b dans la théorie de Becker.

Nous étudierons cette donnée un peu plus tard.

Nous n'avons pas pu mesurer des indicateurs de la répression des crimes, donc l'influence du p dans la théorie de Becker, on peut donc supposer qu'il nous manque une portion d'information qui peut expliquer le R^2 pas si élevé obtenu.

Pour la mesure de la scolarisation, nous avons fait un choix, cette mesure n'est peut être pas pertinente ou bien non significative.

D'autre part, il aurait été intéressant d'observer les distance à Paris des communes. Dans le cadre de notre étude , on pourrait remettre en cause la méthode de Daniel Dranove qui consis te à comparer les modèles log-niveau et log-log à travers le R^2 car si on avait suivi cette méthode on aurait obtenu des mauvaises estimations des paramètres β_j du modèle.

9 Étude de Paris et du reste du Grand Paris

Dans le cas dans section 7 , on avait pu constater que selon le RAMSEY RESET TEST , on pouvait supposer qu'il n'y avait pas de variables omises.

Cependant , il serait intéressant d'étudier l'une des variables que nous avons à notre disposition : `taux_etabl_soc`. Malgré notre intérêt pour cette variable,avec les résultats obtenus avec le RAMSEY RESET TEST , cela ne semble pas cohérent d'ajouter `taux_etabl_soc` à notre modèle.

Cette variable permet de mesurer à quelle point une commune effectue des "Incapacitation Acts " ,en français , cela se traduirait "Actes de neutralisation ". Selon les analystes , les "Incapacitation Acts" ne sont pas à confondre avec des actions dissuasifs , en effet , une variable "dissuasive" modifierait le bénéfice du crime ($b - pc$) en influençant les quantités p et/ou c (par exemple par l'augmentation du nombre de policiers ou de la durée d'une peine de prison afin de dissuader les délinquants).

Alors que les entités qui effectuent "Incapacitation Acts" neutralisent les délinquants soit en les châtiant soit en essayant de les accompagner dans leur réinsertion sociale : il y a donc une nuance.

De ce fait , pour être plus précis , on dira que `taux_etabl_soc` mesure les "Incapacitation Acts" d'une commune .

On s'attend à ce que `taux_delit` et `taux_etabl_soc` soit en relation indirecte : si `taux_etabl_soc` augmente alors le `taux_delit` diminuera.

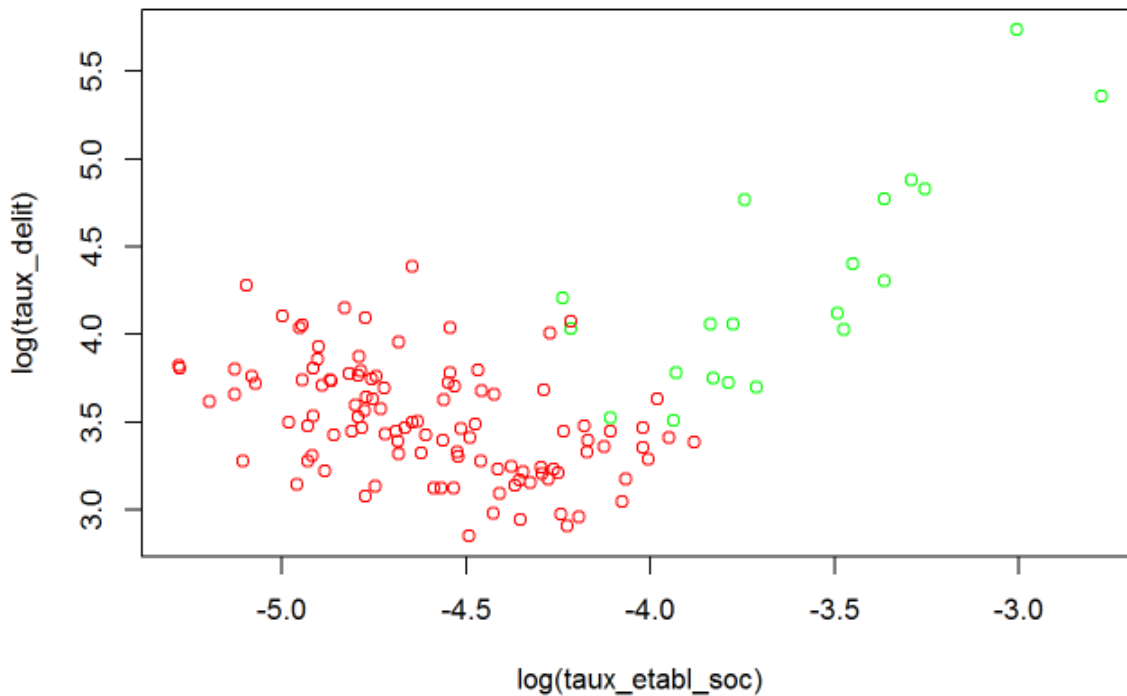


Figure 18: Banlieues(rouge),Paris(vert)

Le résultat est assez étonnant , on peut constater que le nuage de points forme une parabole. Jusqu'à certain seuil , lorsque $\log(\text{taux_etabl_soc})$ augmente le taux_delit diminue .

A partir de $\log(\text{taux_etabl_soc}) = -4.2$ (environ) , on constate que plus le taux_etabl_soc augmente plus le taux de délit augmente : cette tendance peut sembler contre intuitive.

Les communes en vert représentent Paris(il existe d'autres communes qui suivent cette tendance mais d'un point de vue pratique , il semble difficile de les détecter).

Depuis le début de notre étude , on a pu constater que certaines communes ne sont comportaient pas comme les autres: ces communes sont majoritairement à Paris.

En effet , en banlieue , une hausse du taux de chômage fera augmenter la criminalité alors qu'à Paris , une hausse de chômage se traduit par une baisse de la criminalité.

En ce qui concerne le niveau de vie médian ,en banlieue , si il augmente alors la criminalité a tendance à baisser , à Paris ,la relation est inversée.

Si le taux de scolarité augmente en banlieue alors cela se traduit par une baisse du crime , cependant , à Paris , si le taux de scolarité augmente alors on observe une augmentation de la criminalité.

Pour ce qui est du taux de jeunes , les taux de délit que ce soit à Paris ou en banlieue varient dans le même sens : si le taux de jeunes augmente alors le taux de délit augmente.

On pourrait supposer que l'impact de nos variables explicatives sur le taux_delit d'une commune diffère selon la géolocalisation de la commune .

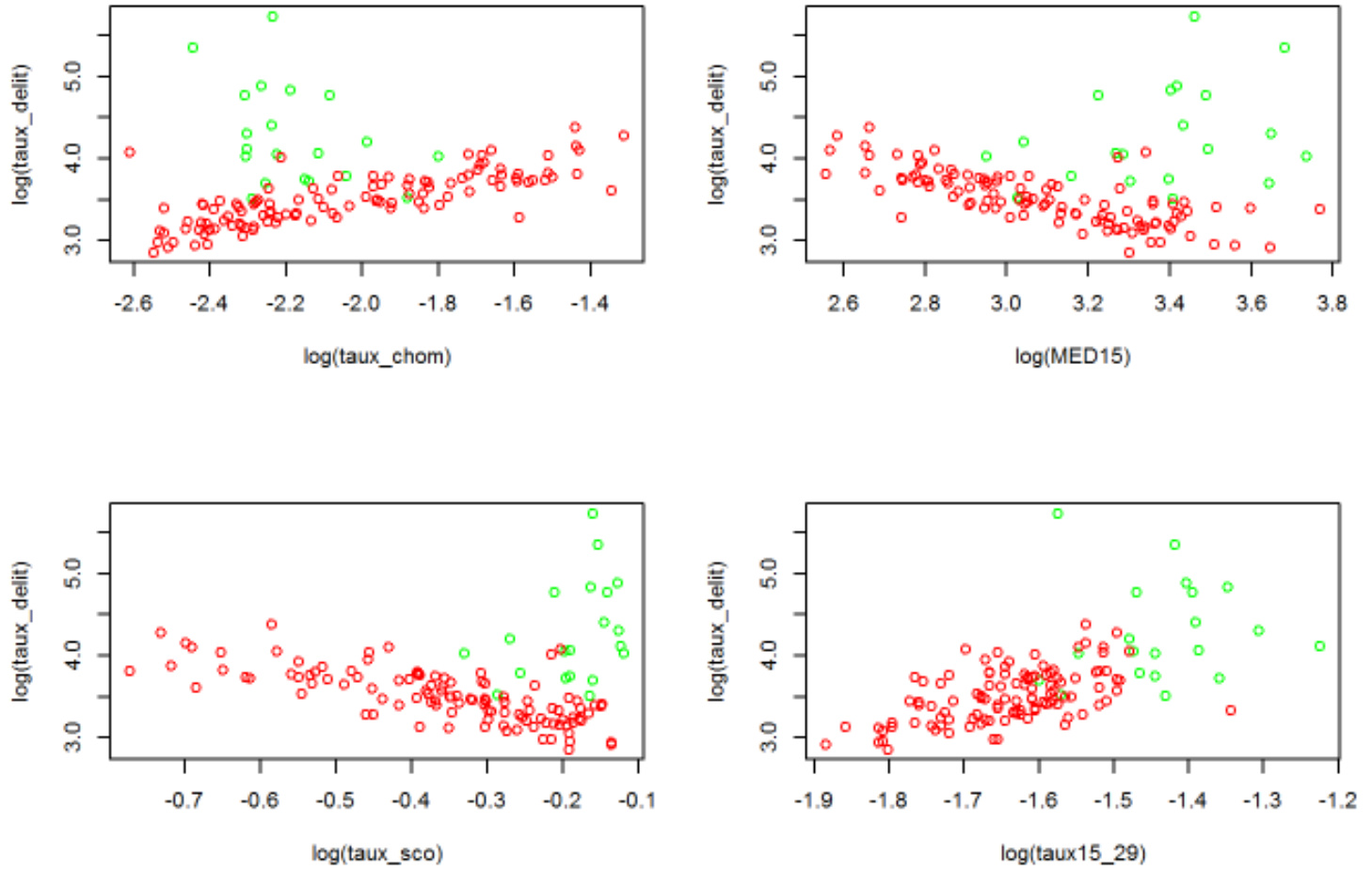


Figure 19: Banlieues(rouge),Paris(vert)

De ce fait, on souhaite savoir, au regard de ce qui précède, si les dynamiques qui régissent la criminalité sont les mêmes pour Paris Intra-Muros et le reste du Grand Paris.
Nous allons donc étudier la stabilité de notre modèle.

9.1 Test de Chow et Interprétation

On va donc effectuer un test de Chow sur nos 2 sous-populations: Paris Intra-Muros et le reste du Grand Paris.

$$Y = X\beta + \epsilon, \quad i = 1 : 130 \quad \text{modèle contraint}$$

$$Y_i = X_i\beta^1 + \epsilon_i, \quad i = 1 : 20 \quad \text{modèle non-contraint 1}$$

$$Y_i = X_i\beta^2 + \epsilon_i, \quad i = 21 : 130 \quad \text{modèle non-contraint 2}$$

Avec :

$$H0 : \beta^1 = \beta^2 = \beta$$

$$H1 : \beta^1 \neq \beta^2$$

Ce test est basé sur un test de Fisher.

On estime le modèle contraint et on récupère ϵ

On estime les modèles sous-contraints et on récupère ϵ^1 et ϵ^2

On calcule

$$F = \frac{\frac{t_{\epsilon}\epsilon - (t_{\epsilon^1}\epsilon^1 + t_{\epsilon^2}\epsilon^2)}{k}}{\frac{t_{\epsilon^1}\epsilon^1 + t_{\epsilon^2}\epsilon^2}{N - 2k}}$$

Sous H_0 , $F \simeq \mathcal{F}(k, N - 2k)$

```
##test de chow
scrc = sum(reg2$residuals^2)
scrc
fichier1 = tab2[tab2$Paris == 1,] ## selectionne les arrondissements de Paris
#head(fichier1)
EQ1_Paris = lm(log(taux_delit) ~ log(taux_sco)+log(taux15_29)+log(taux_chom)+log(MED15),
data=fichier1)
summary(EQ1_Paris)
scr1 = sum(EQ1_Paris$residuals^2)
scr1
fichier2 = tab2[tab2$Paris == 0,] ## selectionne les autres communes
#head(fichier2)
EQ1_NParis = lm(log(taux_delit) ~ log(taux_sco)+log(taux15_29)+log(taux_chom)+log(MED15),
data=fichier2)
summary(EQ1_NParis)
scr2 = sum(EQ1_NParis$residuals^2)
scr2
|
#degr???s de libert???
ddl_n = (reg2$df.residual - (EQ1_Paris$df.residual + EQ1_NParis$df.residual))
ddl_n
ddl_d = EQ1_Paris$df.residual + EQ1_NParis$df.residual
ddl_d
#F du test de Chow et p-value
Fchow = ((scrc-(scr1+scr2))/ddl_n)/((scr1+scr2)/ddl_d)
Fchow
pvalue = pf(Fchow,ddl_n,ddl_d,lower.tail=FALSE)
pvalue
```

```
#F du test de Chow et p-value
Fchow = ((scrc-(scr1+scr2))/ddl_n)/((scr1+scr2)/ddl_d)
Fchow
```

```
## [1] 14.19402
```

```
pvalue = pf(Fchow,ddl_n,ddl_d,lower.tail=FALSE)
pvalue
```

```
## [1] 6.65356e-11
```

La p-value du test est inférieure à 5% donc on rejette H_0 au niveau 95%.
Les deux sous-populations possèdent donc des caractéristiques différentes.
Étudions les régressions sur les modèles sous-contraints:

```
EQ1_Paris = lm(log(taux_delit) ~log(taux_sco)+log(taux15_29)+log(taux_chom)+log(MED15), data=fichier1)
summary(EQ1_Paris)
```

```
##
## Call:
## lm(formula = log(taux_delit) ~ log(taux_sco) + log(taux15_29) +
##     log(taux_chom) + log(MED15), data = fichier1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.09415 -0.36172 -0.07395  0.30067  1.08087
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.490       7.162   0.348   0.733
## log(taux_sco)      4.520       7.850   0.576   0.573
## log(taux15_29)    -1.879       1.965  -0.956   0.354
## log(taux_chom)    -2.747       2.739  -1.003   0.332
## log(MED15)       -1.795       1.852  -0.969   0.348
##
## Residual standard error: 0.6027 on 15 degrees of freedom
## Multiple R-squared:  0.2227, Adjusted R-squared:  0.01545
## F-statistic: 1.075 on 4 and 15 DF,  p-value: 0.4036
```

On observe qu'aucune variable n'est significative,
 Le modèle n'est pas globalement satisfaisant.
 Le R^2 est très faible donc les variables de notre modèle n'expliquent pas ce qui se passe pour Paris,
Il semble donc que les variables qui déterminent Paris soient différentes de celles qui nous permettent de décrire la région du Grand Paris.

De ce fait, il nous est nécessaire de prendre de nouvelles variables explicatives qui pourrait permettre d'expliquer le taux de délit à Paris.

On partira donc du modèle suivant :

$$z_i = \alpha_0 + \alpha_1 \log(x_{i5}) + \gamma_i, i = 1, \dots, 20 \quad (c)$$

, avec $\gamma_i \sim \mathcal{N}(0, \sigma^2)$, les bruits sont indépendants , de plus le taux_etabl_soc est identifié par X_5 .

On s'attend à ce que la variable X_5 soit en relation indirecte avec le taux de délit.


```
d<-data_base[1:20,]
summary(lm(log(taux_delit)~log(taux_etabl_soc),data=d))

##
## Call:
## lm(formula = log(taux_delit) ~ log(taux_etabl_soc), data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.47537 -0.30706 -0.06621  0.17391  0.68639
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      8.7503     0.8270  10.581 3.72e-09 ***
## log(taux_etabl_soc)  1.2326     0.2266   5.438 3.63e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3839 on 18 degrees of freedom
## Multiple R-squared:  0.6217, Adjusted R-squared:  0.6006
## F-statistic: 29.58 on 1 and 18 DF,  p-value: 3.631e-05
```

Figure 20: Sorties de R pour le modèle (c)

Dans ce modèle, le test de validité globale est associé aux hypothèses antagonistes suivantes :

$$H_0 : \alpha_1 = 0$$

$$H_1 : \alpha_1 \neq 0$$

. La probabilité critique du test est inférieure à 0.05,

On rejette H_0 , de ce fait, on peut supposer que $\alpha_1 \neq 0$. Dans ce modèle, la variable `taux_etabl_soc` est significative. En termes d'interprétation, cela reviendrait à dire que si le taux d'établissement social augmentait de 1% alors le taux de délit augmenterait de 1.23%.

Nous nous attendions pas à cette interprétation , de plus , il nous semblait contre-intuitif de penser que le taux d'établissement social puisse faire augmenter le taux de délit ,les résultats du test confirment ce que l'on a pu observé sur les nuages de points.

Nous ne sommes pas en mesure d'expliquer l'origine de cette liaison , cependant , on peut seulement constater qu'elle existe bel et bien.

```
EQ1_NParis = lm(log(taux_delit) ~ log(taux_sco)+log(taux15_29)+log(taux_chom)+log(MED15), data=fichier2)
summary(EQ1_NParis)
```

```
##
## Call:
## lm(formula = log(taux_delit) ~ log(taux_sco) + log(taux15_29) +
##     log(taux_chom) + log(MED15), data = fichier2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.55761 -0.11585 -0.04907  0.11963  0.93219
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.6169      0.7543   7.447 2.79e-11 ***
## log(taux_sco)    -0.1061      0.3645  -0.291  0.77148
## log(taux15_29)    0.3826      0.2653   1.442  0.15225
## log(taux_chom)    0.5004      0.1611   3.107  0.00243 **
## log(MED15)      -0.1623      0.2506  -0.648  0.51859
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2023 on 105 degrees of freedom
## Multiple R-squared:  0.612, Adjusted R-squared:  0.5973
## F-statistic: 41.41 on 4 and 105 DF, p-value: < 2.2e-16
```

On observe que le modèle est globalement satisfaisant et que la constante et le taux de chômage sont significatifs. On constate que le R^2 (0.59) est supérieur à celui du modèle contraint, ce qui n'est pas étonnant compte tenu du résultat pour Paris Intra-Muros.

L'interprétation de ce modèle est que une variation de 1% du taux de chômage implique une variation dans le même sens de 0.5% du taux de délit.

10 Étude d'une typologie de crime dans le Grand Paris

Nous nous limiterons à l'étude des deux typologies de crimes:

1. Les cambriolages et les vols d'automobiles (CRL et VA)

10.1 taux de cambriolages (CRL) et de vols d'automobiles (VA)

L'analyse descriptive nous donne des candidats naturels pour estimer le taux de cambriolages et le taux de vols d'automobiles: la densité et le taux de chômage.

Nous avons choisi de réunir les cambriolages et les vols d'automobiles du fait de leur nature très similaires.

C'est les analyses précédentes qui nous invitent naturellement à étudier les logarithmes des variables.

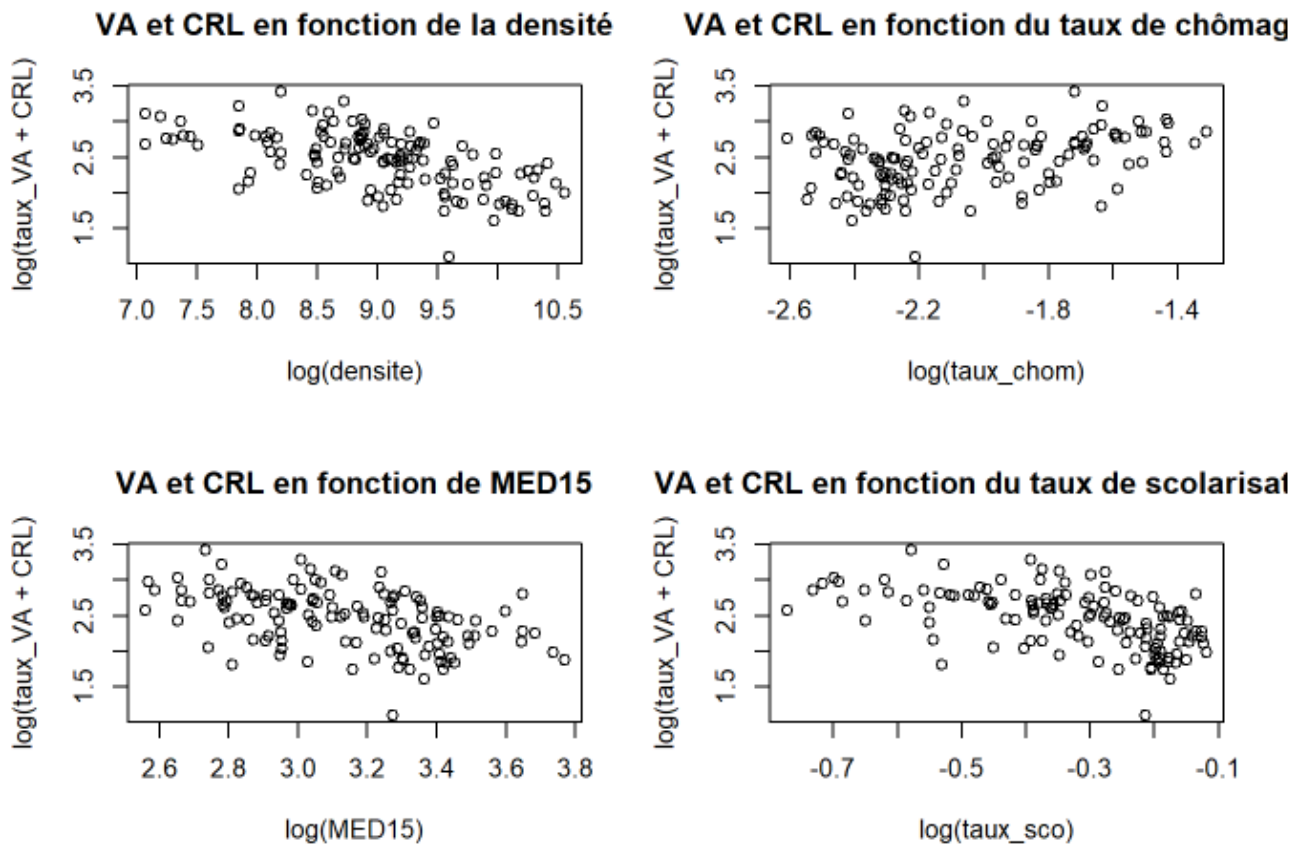


Figure 21: Graphiques $\log(\text{taux_VA} + \text{CRL}) - \log(\text{variables})$

On étudie donc le modèle log-log:

```
{r}
reg2=lm(log(CRL+taux_VA) ~ log(densite)+log(taux_chom) , data=tab2)
summary(reg2)
```

```
Call:
lm(formula = log(CRL + taux_VA) ~ log(densite) + log(taux_chom),
    data = tab2)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-1.11679 -0.18538  0.03221  0.20115  0.75508
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.97304    0.34702  17.212 < 2e-16 ***
log(densite)  -0.29193    0.03338  -8.746 1.19e-14 ***
log(taux_chom) 0.43284    0.08260   5.240 6.55e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.3036 on 126 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.4505,    Adjusted R-squared:  0.4417
F-statistic: 51.64 on 2 and 126 DF,  p-value: < 2.2e-16
```

Le modèle est globalement satisfaisant, la constante et les variables explicatives les sont aussi.

L'interprétation de nos résultats est que :

Une variation de 1% de la densité implique une variation de 0.29% dans le sens opposé.

C'est ce qui était attendu d'après notre analyse, les communes les plus touchées par les cambriolages et les vols d'automobiles sont situées en périphérie du Grand Paris, où la densité est plus faible.

Ainsi une baisse de la densité entraîne bien une hausse de ces crimes particuliers.

Une variation de 1% du taux de chômage implique une variation de 0.43% dans le même sens du taux de cambriolages et vols d'automobiles.

C'est ce qui était attendu d'après notre analyse, un taux de chômage plus élevé signifie un accès moindre au marché du travail légal et ses sources de revenus, ainsi le coût d'opportunité baisse ce qui se traduit bien par une hausse du bénéfice à commettre un crime dans.

On a un R^2 ajusté de 0.44, on est dans le même ordre que pour le modèle tous crimes confondu pour la région parisienne.

On ne parvient pas à expliquer plus d'information que dans le cas général.

11 Conclusion

Nous sommes parvenus à traiter nos trois objectifs, une étude générale du Grand Paris, une distinction entre Paris Intra-Muros et le reste du Grand Paris et une tentative d'étude par typologie de crimes.

En premier lieu, il ressort de notre modèle, malgré les critiques apportées, que certaines variables déterminent de manière significative la criminalité pour le Grand Paris à savoir le **taux de chômage (+)**, le **taux de jeunes(+)** (âgés de 15 à 29 ans) ainsi que le **niveau de vie médian(+)** (med15).

Ce qui corrobore les hypothèses émises en introduction via le cadre de la théorie de Becker.

Cependant, il semble que plusieurs sous-populations coexistent au sein du Grand Paris, l'étude comparative de Paris Intra-Muros et de la banlieue seule le montre.

Les variables qui étaient significatives pour le Grand Paris ne le sont pas, le modèle n'est plus globalement satisfaisant, donc elles ne permettent d'interpréter les dynamiques de la criminalité pour Paris Intra-Muros.

Ce qui met un peu plus en lumière le caractère spécifique de la capitale vis à vis de la criminalité.

En banlieue, seul le taux de chômage est significatif pour expliquer le taux de crime malgré une information expliquée par le modèle plus grande. ($0.597 > 0.43$)

L'intégration de variable contrôlant la répression des crimes et la correction de certaines mesures (taux de jeunes notamment) pourrait nous apporter un surcroît d'information pour déterminer cette criminalité.

Cependant, il faudrait étudier également de près les spécificités des quartiers parisiens, quitte à subdiviser les arrondissements afin de mieux lier certains paramètres explicatifs (à déterminer) des différences inter-arrondissements qui peuvent être fortes et ce même pour des arrondissements présentant (d'après nos données) des caractéristiques similaires (selon nos variables).

De plus, l'intégration d'outils d'économétrie spatiale et le traitement des données panel pourrait nous permettre d'affiner nos modélisations.

L'étude par typologies de crimes n'est pas plus concluante que dans le cas général,

On apporte une information supplémentaire avec la variable densité dans le cas des cambriolages et des vols d'automobiles.

Pour poursuivre cette étude, il serait donc intéressant d'intégrer des variables mesurant la répression des crimes et certaines caractéristiques sociales (les situations des ménages, le taux d'avortement (S.Levitt), caractère touristique de l'environnement), une mesure plus précise des variables (taux de jeunes, inégalités salariales) et des méthodes de traitement des données panel pour étudier l'impact dans le temps de nos variables et d'économétrie spatiale pour étudier l'intuition d'une discrimination par rapport à la distance à la ville de Paris (pour les communes du Grand Paris).

References

- [1] DR. SALIH DURSUN. PROF. DR. SERPIL AYTAÇ. ASSOC. PROF. DR. FERHAT TOPBAŞ : *The Effects of Unemployment and income on Crime: a Panel Data Analysis on Turkey*
- [2] HANS GRÖNQVIST.: *Youth Unemployment and crime: New lessons exploring longitudinal register data*
- [3] PAULINE GIVORD. : *Introduction à l'Econométrie Ecole Centrale de Paris Année 2006-2007*

- [4] DANIEL DRANOVE.: *Practical Regression: Log vs. Linear Specification* Kellogg School of Management-Northwestern University (2002)
- [5] DENIS FOUGÈRE. FRANCIS KRAMARZ. JULIEN POUGET: *Youth Unemployment and Crime in France*, CREST-INSEE
- [6] KERRY L. PAPPS. RAINER WINKELMANN.: *Unemployment and Crime: New evidence for an old question* (1999)
- [7] ONDRP: *Crimes et délits commis dans les communes du Grand Paris en 2015*
- [8] DENIS FOUGÈRE. FRANCIS KRAMARZ. JULIEN POUGET: *Délinquance et mobilité résidentielle*, CREST-INSEE
- [9] EMMANUEL COMBE. SÉBASTIEN DAZIANO: *Lutter contre les vols et cambriolages Une approche économique*, fondapol.org

The Economics of Crime: An Analysis of Crime Rates in America

Alison Oliver

I. Introduction

A growing concern across the nation is the heightened incidence of criminal and violent behavior. There has been a steadily increasing trend in the crime rate over time until the 1990s, where the trend begins to fall off (Figure 1). However, the fluctuations have been a historical pattern since 1960, and the causes of them remain unexplained. Since the publication of Gary Becker (1968), the economics profession has analyzed the determinants of criminal behavior from theoretical and empirical points of view. The growing public awareness is justified because rampant crime and violence may have pernicious effects on economic activity and, more generally, because they directly reduce the quality of life of all citizens who must cope with the reduced sense of personal and proprietary security. Past research is based on models that look at the incentives faced by individuals to commit crimes, and possible causes of the persistence of crime over time (criminal inertia). I will discuss the elements factoring into the individual decision to commit a crime in my theory section. Then I will use the national crime rate over time in order to predict a reduced-form analysis of the country's aggregate individual decisions. The purpose of my paper is to determine what factors are significant in determining fluctuations

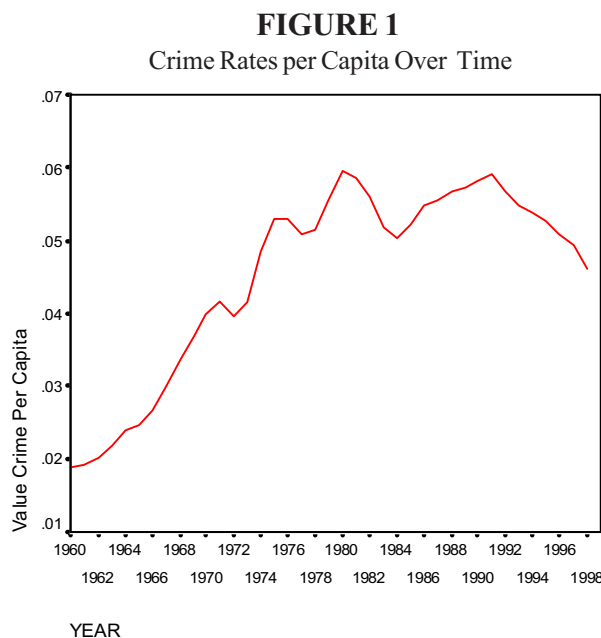
in the crime rate over time.

The rest of the paper will be organized as follows. Section II provides a selective review of theoretical and empirical contributions to the economics literature dealing with criminal behavior. Section III presents the economic model I will be using, and an explanation of the variables and my hypotheses for their effects on the crime rate. Section IV presents the results of the regression and how they are related to the research hypotheses. Finally, Section V indicates the major conclusions of the paper and policy implications derived from my findings.

II. Theory and Literature Review

The theoretical framework behind my paper is based on Becker's Rational Choice model, where an individual's decision to commit a crime is

based on the costs and benefits (1968). In this model, all potential criminals have a benefit of crime (b), includes both the financial and any expected psychological benefits of crime. An individual committing crime faces costs from law-enforcement activities. The severity of the punishment including fines and jail time is one part of the total cost, and the other part is the probability of getting caught. Therefore, the costs will equal



the probability of punishment (p) times the cost of punishment (c). Thus, the net expected returns from crime equal:

$$b - pc$$

Standard differentiation tells us that the number of criminals rises as b rises, and falls as p or c rises. Therefore, the individual decision to commit a crime is conditional upon the following stipulation:

$$(b - pc) > 0 \quad (1)$$

Crime reduction can occur through reducing the benefits of crime, raising the probability of being caught, or increasing the costs of punishment conditional upon being caught.

The theoretical model Becker sets up takes into account benefits, costs, and probability of paying those costs. However, opportunity costs seem to be absent from his model. Isaac Ehrlich analyzes the effects of income levels and distribution on crime in 1973. He argues that pay-offs to crime, especially property crime, depend primarily on the “opportunities provided by potential victims of crime,” as measured by the median income of the families in a given community. He also considers the effect of unemployment on crime rates. He views the unemployment rate in a community as a complementary indicator of income opportunities available in the legal labor market. However, in his empirical study, he finds that unemployment rates are less important determinants of crime than income levels and distribution. These variables are a measure of opportunity cost (o) which could be added to Becker’s model (equation 1): $b - (pc + o)$.

Another factor Ehrlich (1973) discusses, which relates to the effect of economic conditions on the opportunity cost of crime, is the level of education of the population. Education helps to determine the expected rewards from both legal and criminal activities. However, after controlling for income inequality and median income, Ehrlich finds a positive and significant relationship between the average number of school years completed by the adult population (over 25 years) and particularly property crimes commit-

ted across the U.S. in 1960 (1973). He explains this in two different ways. One being that as education increases a criminal can enter into higher paying sectors of crime. However, a more probable theory is explained, especially in connection with property crimes. As a nation’s average educational attainment increases, income will increase and the payoff of property crimes increases, thus increasing criminal’s benefits (b). However, it is expected that without the controls for income, an increase in education could be associated with higher expected legal earnings and thus increase the opportunity cost of committing a crime (o).

Together with the relationship between economic conditions and crime, one of the main issues in the pioneering studies of Becker (1968) and Ehrlich (1973) is the assessment of the effects of police presence, convictions, and the severity of punishments on the level of criminal activity. Individuals who are considering whether to commit crimes are assumed to

evaluate both the risk of being caught and the associated punishment. Their empirical evidence confirms that both factors have a negative effect on crime rates.

Analysts often make a subtle distinction between the “deterrent” effects of policing and convictions, and the “incapacitation” effects of locking-up (or killing, in the case of capital punishment) criminals who may have a

tendency to rejoin the crime industry once they are released. Deterrence essentially aims at modifying the price of crime for all offenders (increasing p and c). While incapacitation acts through the removal of a subset of convicted offenders from the market for offenses either by relocating them in legitimate labor markets, or by excluding them from the social scene for prescribed periods of time.

Steven Levitt (1999) addresses these issues jointly with one of the most recurrent problems in the literature mentioned above. The author attempts to assess whether the apparent negative relationship between crime rates and arrest rates are the product of deterrence effects, incapacitation, or measurement errors associated with the fact that crime tends to go unreported. The author finds that most of this negative relationship in the U.S. is due to deterrence effects (increases in p and c), and not measurement

“Crime reduction can occur through reducing the benefits of crime, raising the probability of being caught, or increasing the costs of punishment conditional upon being caught.”

error or incapacitation, for most types of crime.

The economics literature on crime has progressed from an emphasis on economic conditions (including education) and deterrence effects to more recent considerations of factors that may explain how crime is propagated over time and within communities. In the next section, I will organize some of the ideas from the literature in a simple framework.

III. Empirical Model

For the empirical testing of crime rates, I use the adapted rational choice model (equation 1) to build a model in which the dependent variable is the United States National crime rate (CRIME) and the explanatory variables are a number of national economic and social characteristics. A summary of these variables is presented in Table 1. They are separated into three groups: Economic Variables, Deterrence Variables, and Demographic Variables. The Economic Variables look at the opportunity cost side of the theoretical model, arguing that certain factors contribute to the decision of whether crime is worthwhile. The Deterrence variables test the cost side of the theoretical equation as well as the probability of getting caught. The significance of these variables will have important policy implications on crime prevention. The final group consists of only one variable (AGE), and is referred to as the Demographic Variable. The AGE

variable is the percentage of the population between the ages of 14 and 25. Criminologists often find in their studies that this age group has the highest propensity to commit crimes, so including this variable will give insight to who the prevention policies should be geared towards.

For my research I use a log-linear regression analysis on national time-series data of the United States. In a linear-log model, the dependent variable is unchanged but the independent variable appears in logarithmic form. This model was chosen because the relationship between the Crime Rate per capita and the independent variables is non-linear. Specifically, the marginal effect of the independent variables on the Crime Rate per capita, is increasing at a decreasing rate. This fitted curve can then be extrapolated to generate forecasts of the dependent variable, which is beneficial in policy implications.

A. Economic Variables

The first variable included is education level, which may impact the decision to commit a crime in several ways. The measurement used for education is the percent of the population that graduated from high school. Higher percentages of the population that have received a high school diploma (EDC) may

TABLE 1
Descriptions of Variables

Symbol	Definition	Standard Deviation	Mean	Expected Sign
CRIME	United States National Crime Rate (1960-1998)	.01304	.0452	
Economic Variables				
EDC	Percent of population that completed High School	11.4754	67.3222	?
ENRL	%of population enrolled in secondary school	.02318	.2231	-
GDP	Real GDP per capita	5105.901	21628.33	-
GINI	The Gini Index	.02602	.3788	-
UNEMP	Unemployment Rate	.2552	1.7696	+
Deterrence Variables				
PLC	Police Rate - number of police per capita	.05950	5.3864	-
INCAR	Incarceration population	110.969	263.6364	-
LGRTE	Lagged Crime Rate per capita	.01304	.0452	+
Demographic Variables				
AGE	%of Population under 25	.1100	1.8223	+

Source: FBI Uniform Crime Reports, the Bureau of Economic Analysis, the World Bank, and the National Center for Educational Statistics

be associated with higher expected legal earnings, thus a higher opportunity cost of committing a crime (o). Education through its civic component may also increase the individual's moral stance. On the other hand, education may reduce the costs of committing crimes or open opportunities for an individual to enter higher-paying crime industries. Hence the net effect of education on the individual's decision to commit a crime is ambiguous. It is possible however, that school enrollment (ENRL) alone will reduce the time available for participating in the crime industry, and therefore have an inverse relationship on crime rates. ENRL is specifically the percent of the population enrolled in secondary school. Thus, the empirical model will assess the effect of both secondary enrollment rates and educational attainment on crime rates.

The level and growth of economic activity in society create attractive opportunities for employment in the legal sector (o increases), but since they also improve the wealth of other members of society, the size of the potential loot from crime also rises (b increases). However, based on past research the effect GDP has on employment is higher than increased benefits and therefore, the effect of increased GDP per capita (GDP) on the individual's decision to commit a crime should be negative. The effect of income inequality (GINI) in society should have a negative effect as well, depending on the individual's relative income position. In the case of the poor, an increase in inequality may be crime inducing, because such an increase implies a larger gap between the poor's wages and those of the rich, thus reflecting a larger difference between the income from criminal and legal activities. Therefore, a rise in inequality will have a positive impact on some individual's propensity to commit a crime. The Gini Index is measured by a decimal between 0 and 1, 1 meaning there is total equality. So as inequality rises, the number falls and GINI will have a negative sign.

Ehrlich (1973) considers the effects of unemployment on crime rates. The unemployment rate (UNEMP) can be viewed as a complementary indicator of income opportunities available in the legal labor market. Therefore, when unemployment rates increase, the opportunities in the legal labor sector decrease (o decreases).

B. Deterrence Variables

The strength of the police and the judicial system increases the probability of apprehension and the

punishment for criminal actions (p increases), thus reducing the incentive for an individual to commit a crime. This is the crime deterrence effect. It should also be noted that the past incidence of crime in society might determine an individual's perceived probability of apprehension via systemic interactions, as discussed above. The police rate (PLC), which is measured as the number of police per 100,000 residents, will be used as one of the variables to test the deterrence effect. To account for the "incapacitation" effect on criminals, I use the incarceration rates (INCAR) for prisoners, which is the number incarcerated per 100,000 residents. Incapacitation acts through the removal of a subset of convicted offenders from the market for offenses, by excluding them from the social scene for prescribed periods of time.

The individual's past experience in criminal activities is another important variable that affects in several ways the decision to commit a crime. First, convicts tend to have a limited amount of employment opportunities, and a lower expected income. Second, the theory that an individual learns by doing can apply to criminals as well, which lowers the costs of committing a crime. These arguments strongly suggest the possibility of criminal inertia, that is, present crime incidence explained to some extent by its past incidence. Therefore, I will include the lagged crime rate (LGRTE) variable as a measure of this effect. The lag accounts for the crime rate one year prior.

C. Demographic Variables

The last variable I use is the percentage of the population under the age of 25 (AGE). Criminologists view the change in the population age distribution as the greatest influence on crime trends. As a general rule, the crime rate follows the proportion of young males in the population. The number of juveniles should be increasing over the next decade, and some criminologists fear that this will signal a return to escalating crime rates. Thus, the age variable is expected to have a direct relationship with the crime rate. My empirical model, using the crime rate as a function of the above variables will look like this:

$$CRIME = b_1 + b_2 \ln EDC + b_3 \ln ENRL + b_4 \ln GDP + b_5 \ln GINI + b_6 \ln UNEMP + b_7 \ln INCAR + b_8 \ln PLC + b_9 \ln LGRTE + b_{10} \ln AGE \quad (2)$$

IV. Results

Results of the Linear-log regression using Crime per capita as the dependent variable are summarized in Table 2. Unfortunately, the only variable that served to be significant was the lagged crime rate. Due to the limitation on crime data available, there were only 38 cases recorded. Therefore, a degree of freedom problem is present, and the only way to fix it is to increase the sample size or decrease the number of variables. Due to the lack of available data, a decrease in variable number must occur. The decision on what variables to eliminate is based on another problem with the data. Specifically, EDC, ENRL, GDP, GINI, and INCAR all are considered trend variables. Therefore, a lack of variation within the variables occurs, and there are similar trends between all of them, which means the results will not be clear on the effects each variable has on the crime rate per capita. In addition, the lagged crime rate (LGRTE) may be picking up the influence of the other independent variables.

In order to test for all the effects, I chose one variable from each group, that would most fully represent that effect. The new model is:

$$\text{CRIME} = \beta_1 + \beta_2 \text{UNEMP} + \beta_3 \text{PLC} + \beta_4 \text{AGE} \quad (3)$$

The results of the second regression supported my hypotheses, although explaining less with an R^2 of .331

(see Table 3). The results of the AGE variable supported my hypothesis that an increased population of people between the ages of 14 and 25, increased the crime rate per capita. This means that the marginal effect of AGE on CRIME is increasing as AGE increases. The variable is significant to the .01 level. In the log-linear model the elasticity of Y with respect to X can be calculated using the following equation:

$$B_n/Y \quad (4)$$

Y is the mean value of CRIME, and B_n is the coefficient of the variable being tested. This calculation using the coefficient of AGE from equation 4, results in a 0.0161 elasticity of AGE with respect to CRIME. Meaning, there is a 1.6 percent increase in CRIME (crime per capita) with respect to a 1 percent increase in AGE (the percent of the population between the ages of 14-24).

The PLC variable is significant to the .01 level, and resulted in the expected sign. The elasticity of CRIME with respect to PLC, using equation 4, results in -0.0106. This translates into a 1 percent decrease in crime rates per capita with a 1 percent increase in the police ratio, as hypothesized. The UEMPL results in the expected sign, but unfortunately it is insignificant. The hypothesis is that an increase in unemployment would increase crime, through a decrease in the opportunities to earn money in the legal sector. However, the insignificant finding parallels that of Ehrlich's results.

In this paper, I argue that there are two particularly important areas for research on the economics of crime. First, the actual effect of deterrence on the level of crime continues as a central question for research. Although I was only able to test the effect of police rates on crime, there are other forms of deterrence that take place. Police rates test the probability effect in Becker's Rational Choice (equation 1), while a measure of the intensity of the punishment

TABLE 2
First Log-Linear Regression Results
(Adjusted $R^2 = .770$)

Variable	Coefficient	Significance
<i>Constant</i>	-.612	.103
Economic Variables		
<i>EDC</i>	.0888	.325
<i>ENRL</i>	-.002985	.936
<i>GDP</i>	.0424	.317
<i>GINI</i>	-.003117	.945
<i>UNEMP</i>	.002015	.778
Deterrence Variables		
<i>PLC</i>	.03105	.527
<i>INCAR</i>	-.0382	.038
<i>LGRTE</i>	.05608	.001
Demographic Variables		
<i>AGE</i>	.02547	.161

TABLE 3
Second Log-Linear Regression Results
(Adjusted $R^2 = 0.331$)

Variable	Coefficient	Significance
Constant	0.298	.006
AGE	0.02951	.011
PLC	-0.05721	.009
UNEMPL	0.005196	.211

would test the cost of crime. However, there are different levels of punishment depending on what kind of crime one commits. Future research would benefit from splitting the crimes into groups and testing each individually, so this variable could be included. Second, the levels of crime are not just determined by the level of deterrence. Indeed, in most studies deterrence is important but can explain at best a tiny fraction of the overall level of crime. In principle, social factors can explain significantly greater amount of the variance of crime across space. Due to data limitations, this study was only able to test the significance of being between the ages of 14- 24 as the social factor. But future research could include measures of effects such as divorce, abortion rates, or drug use. Research is increasingly attempting to understand these factors and find ways to test them.

The policy implications of this model are somewhat unclear. The fact that increased police ratios will decrease crime rates obviously implies that more police should be employed, especially in areas with more crime. The significance in the AGE variable implies that there should be policies to keep children between the age of 14-24 off the streets and out of crime. Programs could be set up in schools or communities attracting this group to other hobbies, such as art, music, or sports.

As crime has risen to the top of the theoretical ideas and empirical findings that can be translated into popular discourse and carved into public laws. Gary Becker (1968) crafted what was essentially an early principle-agent model of crime and punishment: criminals are rational, self-interested agents whose behavior is best understood as an optimal response to the incentives set by the government (or principal) via expenditures on law enforcement and corrections. In the years since, others have refined Becker's economic approach to the subject, including my own. My research developed a model in which criminals know that the probability of detection depends both on law enforcement monitoring and on the behavior of the community. The results imply that the decision to commit a crime is a cost benefit analysis, therefore future research should focus in on the population between the ages 14 – 25. Within this group the benefits of crime are the material items received from the act, or feelings of psychological empowerment, which is hard to measure. However, when turning to the costs of crime, one can look at the severity of the punishment, which includes fines and jail time, or the

probability of getting caught. The significance of the two variables on the cost side of the decision to commit a crime has different policy implications and would be important for future research.

References

- Becker, Gary S.** "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76, Mar.-April (1968), 169-217.
- Davis, Michael L.** "Time and Punishment: An Intertemporal Model of Crime." *The Journal of Political Economy*, 96, April 1998, 383-390.
- Ehrlich, Isaac.** "Crime, Punishment, and the Market for Offenses." *Journal of Economic Perspectives*. Vol. 10, Winter 1996, 43-67.
- Ehrlich, Isaac.** "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *The Journal of Political Economy*, Vol. 81, May – Jun., 1973, 521-565.
- Fender, John.** "A General Equilibrium Model of Crime and Punishment." *Journal of Economic Behavior & Organization*. Vol. 39, August 1999, 437-53.
- Jost, Peter J.** "Crime, Coordination, and Punishment: An Economic Analysis." *International Review of Law & Economics*. Vol. 21, March 2001, 23-46.
- Levitt, Steven D.** "Why do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?" *Economic Inquiry*. Vol. 36, 1998, 353-372.
- Martinez-Vazquez, Jorge; Seaman, Bruce A.** "Criminal Behavior in General Equilibrium: Who Benefits from Crime?" *Journal of Socio-economics* Vol. 26, 1997, 303-24.
- Neilson, William S; Winter, Harold.** "On Criminals' Risk Attitudes." *Economics Letters*. Vol. 55, August 1997, 97-102.
- Ramanathan, Ramu.** *Introductory Econometrics with Applications*. San Diego: Harcourt Brace College Publishers, 1998.
- Siegel, Larry.** *Criminology: Theories, Patterns, and Typologies*. Belmont, CA: Wadsworth Thomas Learning, 2001.