

# Application de la régression quantile et ses tests statistiques sur les données PISA en Suisse romande

Eric Eboulet  
Alina Matei




# Application de la régression quantile et ses tests statistiques sur les données PISA en Suisse romande

Eric Eboulet  
Université de Lausanne

Alina Matei  
Institut de recherche et de documentation pédagogique et  
Université de Neuchâtel



IRD  
Faubourg de l'Hôpital 43  
Case postale 556  
CH-2002 Neuchâtel



Tél. ++41 (0) 32 889 86 18  
Fax ++41 (0) 32 889 69 71

E-mail: [documentation@irdp.ch](mailto:documentation@irdp.ch)  
<http://www.irdp.ch>

Cette publication est également disponible sur le site de l'IRDP  
<http://publications.irdp.relation.ch/publications>

Cette publication de l'IRDP est un document de travail. La diffusion de ce document est restreinte et toute reproduction, même partielle, ne peut se faire sans l'accord de son(ses) auteur(s).

Photo de couverture : Maurice Bettex - IRDP

# Table des matières

<b>Résumé</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
<b>2 La régression linéaire et la régression quantile</b>	<b>5</b>
2.1 La régression linéaire . . . . .	5
2.2 La régression quantile . . . . .	5
2.2.1 La fonction quantile et quantile conditionnelle . . . . .	5
2.2.2 La fonction quantile empirique . . . . .	6
2.2.3 Définition de la méthode de régression quantile . . . . .	6
2.2.4 Estimation des paramètres . . . . .	7
2.2.5 Le pseudo $R^2$ : $R^1$ . . . . .	8
<b>3 Tests sur les quantiles</b>	<b>9</b>
3.1 Test d'égalité des quantiles . . . . .	9
3.2 Tests en régression quantile . . . . .	9
3.2.1 Test de qualité de prédiction . . . . .	10
3.2.2 Test de stabilité des paramètres . . . . .	11
<b>4 Application</b>	<b>13</b>
4.1 Présentation des données . . . . .	13
4.1.1 Origine des données . . . . .	13
4.1.2 Description des variables d'étude . . . . .	13
4.2 Analyses et résultats . . . . .	14
4.2.1 Coefficients de régressions . . . . .	14
4.2.2 Performance de modèles - tests de qualité de modèles . . . . .	18
4.2.3 Différences et similitudes cantonales . . . . .	19
<b>5 Conclusions</b>	<b>23</b>
<b>Références</b>	<b>25</b>
<b>Annexes</b>	<b>27</b>



## Résumé

Dans le cadre des enquêtes du *Programme international pour le suivi des acquis des élèves* (PISA), on s'intéresse habituellement à des moyennes de scores des élèves. Une démarche plus informative est de considérer non pas les moyennes de ces scores, mais leurs quantiles. On définit le  $q^e$  quantile des scores observés comme la valeur  $x_q$  qui a la propriété suivante :  $q\%$  des scores sont plus petits que  $x_q$  et  $(1 - q)\%$  sont plus grands que  $x_q$  sur l'échelle des scores ordonnés. Les performances des élèves faibles sont ainsi associées aux quantiles de petit ordre de la distribution des scores et celles des élèves performants aux quantiles de grand ordre.

Les scores calculés dans le cadre de l'enquête PISA sont en général modélisés à l'aide d'une régression linéaire basée sur la méthode des moindres carrés ordinaires (MCO). Le but est d'obtenir un modèle capable d'expliquer une partie de la variance des scores moyens, étant donné certaines caractéristiques des élèves ou de leurs écoles. Cependant, dans l'établissement d'un modèle MCO, les quantiles de la distribution empirique des scores, et en particulier les scores extrêmes, sont ignorés. Le choix d'une telle méthode a donc des conséquences sur la modélisation des scores extrêmes, caractérisant les élèves faibles ou performants. Une alternative est l'utilisation de la régression quantile que nous employons ici. Notre démarche est déterminée par les développements récents des tests statistiques basés sur les quantiles (Furno, 2011; Li *et al.*, 2012). Ces tests permettent de construire des modèles de régression non seulement au centre de la distribution des scores, mais également aux extrémités, et d'étudier ainsi les scores des élèves faibles ou performants. À l'aide de ces tests statistiques, nous mettons en évidence également des différences cantonales au niveau de certains quantiles de la distribution empirique des scores et des similitudes à d'autres quantiles. Notre exemple s'appuie sur les données PISA Suisse romande en 2006.

**Mots clés** : scores PISA extrêmes, régression quantile, tests statistiques sur les quantiles.



## 1 Introduction

Le Programme international pour le suivi des acquis des élèves (PISA) évalue les compétences des élèves de 15 ans en mathématiques, en sciences et en lecture. Ce programme, établi par l'Organisation de coopération et de développement économique (OCDE), a été lancé en 2000 et ses enquêtes se déroulent tous les trois ans.

La Suisse a participé à toutes les vagues de l'enquête internationale PISA. En plus de cette enquête internationale, la Suisse a interrogé en 2000, 2003, 2006, 2009 et 2012, un échantillon complémentaire d'élèves fréquentant la dernière année d'école obligatoire. La Suisse romande a été chaque fois bien représentée dans cet échantillon complémentaire. Étant donné l'hétérogénéité du système éducationnel suisse, où chaque canton a son propre système, des disparités cantonales sont présentes également dans les scores PISA dans toutes les vagues de l'enquête.

Dans une démarche explicative, on associe souvent aux scores des caractéristiques des élèves ou de leurs écoles. Des modèles de régression (dont les paramètres sont estimés d'habitude par la méthode des moindres carrés ordinaires - MCO) sont couramment utilisés pour mettre en lien le score moyen obtenu et certaines caractéristiques, comme le sexe, la langue parlée à la maison ou le niveau socio-économique des parents (voir, par exemple, Nidegger, 2008). Cependant, lorsqu'on s'intéresse aux scores faibles ou performants des élèves, ces types de modèles ne sont plus utilisables. Ainsi, du fait de leur modélisation unique des scores moyens, le recours à la régression quantile (RQ) devient alors nécessaire. Plusieurs auteurs tels Bratti *et al.* (2007), Hao et Naiman (2007), Davino et Goiricelaya (2008), Fryer et Levitt (2010), Furno (2011) s'en réfèrent lorsque leurs intérêts portent sur des études de scores PISA extrêmes.

Nous présentons ici une méthodologie globale, basée sur les quantiles, comprenant des modèles RQ et leurs tests statistiques (Furno, 2011; Li *et al.*, 2012). Ces tests statistiques sont plus informatifs que ceux de type MCO. En plus, ils permettent de construire des modèles de régression non seulement au centre de la distribution des scores, mais également aux extrémités, et d'étudier ainsi les scores des élèves faibles ou performants. À l'aide de ces tests statistiques, nous mettons en évidence également des différences cantonales au niveau de certains quantiles de la distribution empirique des scores et des similitudes à d'autres quantiles.

Le présent document fait appel à des outils statistiques encore peu utilisés dans les sciences de l'éducation. Nous mettons en garde les lecteurs : il s'agit d'un traitement assez technique. Toutefois, la section 4 offre une présentation appliquée de ces outils sur l'échantillon complémentaire PISA Suisse romande 2006. Le document est organisé comme suit. Après cette section, la section 2 est consacrée à la théorie sur les méthodes d'estimation des paramètres des modèles MCO et RQ. Dans la section 3, les tests sur les quantiles sont présentés. La section 4 se concentre sur la mise en œuvre des méthodes et tests décrits précédemment, en utilisant les données de l'enquête PISA Suisse romande 2006. Elle met en exergue les contributions des diverses caractéristiques des élèves afin d'expliquer à différentes échelles les scores PISA en mathématiques, en sciences et en lecture, révélant ainsi des différences et des similitudes cantonales. Enfin, la section 5 conclut. Pour une lecture moins technique nous suggérons aux lecteurs de passer directement à la section 4.





## 2 La régression linéaire et la régression quantile

### 2.1 La régression linéaire

Dans les études empiriques, lorsqu'on cherche à analyser l'effet d'un ensemble de  $p$  variables indépendantes  $X_1, \dots, X_p$  sur une variable dépendante  $Y$ , une approche standard consiste à spécifier un modèle linéaire, puis à estimer ses paramètres inconnus :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi} + \varepsilon_i, \quad i = 1, \dots, n,$$

où  $\beta_0, \dots, \beta_p$  sont les paramètres inconnus à estimer et  $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ ,  $i = 1, \dots, n$ , les termes d'erreurs supposés indépendants.

Le plus souvent, ces paramètres sont estimés par la *méthode des moindres carrés ordinaires* (MCO), sous certaines hypothèses (voir, par exemple, Draper et Smith, 1998, chapitre 5). La démarche avec cette méthode consiste à les calculer en approximant la fonction moyenne de la distribution conditionnelle de la variable dépendante  $Y$ , étant donné les variables  $X_j$  ( $j = 1, \dots, p$ ). Néanmoins, si au moins une de ces hypothèses, comme par exemple la normalité de la distribution des termes d'erreurs n'est pas confirmée ou encore que les termes d'erreurs sont hétéroscédastiques, les hypothèses du modèle sont invalidées, et les estimations des paramètres, les intervalles de confiances, et d'autres statistiques calculées deviennent peu fiables. Cette inefficacité de la méthode est d'ailleurs en liaison avec la faible robustesse de la moyenne en tant qu'estimateur de position.

Pour palier cet inconvénient, Koenker et Bassett (1978) introduisent la *régression quantile* (voir également Koenker, 2005). C'est une des méthodes de régression plus robuste que la méthode MCO, qui définit une nouvelle procédure d'estimation des paramètres  $\beta_0, \dots, \beta_p$ , qui soit à la fois :

- peu sensible à la présence des points aberrants ;
- efficace pour une grande variété des termes erreurs  $\varepsilon_i$  ;
- aussi efficace que la méthode MCO lorsque les termes d'erreurs ont une distribution normale.

Cette approche étend le modèle de régression standard aux quantiles de la distribution conditionnelle de la variable dépendante. Pour une introduction en français de la régression quantile, voir Givord et D'Haultfoeuille (2013).

### 2.2 La régression quantile

#### 2.2.1 La fonction quantile et quantile conditionnelle

Soient  $Y$ , une variable aléatoire (discrète ou continue) avec une fonction de distribution de probabilité  $F_Y(y) = P(Y \leq y)$  et  $\tau$ , l'ordre de quantile de  $Y$  ( $0 < \tau < 1$ ). Le  $\tau$ -ème quantile de  $Y$ , noté  $Q_Y(\tau)$ , est défini comme le plus petit  $y$  avec  $F_Y(y) \geq \tau$ , c'est-à-dire :

$$Q_Y(\tau) = F_Y^{-1}(\tau) = \inf\{y : F_Y(y) \geq \tau\}, \quad 0 < \tau < 1. \quad (1)$$

Ce sont donc des seuils en-dessous desquels se trouvent  $(100 * \tau)\%$  de la donnée qui leur sont inférieurs.

Si, d'autre part,  $Y$  a une fonction de distribution de probabilité conditionnelle à  $X$ , notée par  $F_{Y|X}(y)$ , où  $X$  est une variable aléatoire ; sous la condition  $X = x$ , il s'ensuit que le quantile

conditionnel d'ordre  $\tau$  de la variable  $Y$  sachant  $X = x$  s'écrit :

$$Q_Y(\tau|X) = F_{Y|X}^{-1}(y) = \inf\{y : F_{Y|X}(y) \geq \tau\}, 0 < \tau < 1.$$

### 2.2.2 La fonction quantile empirique

La fonction quantile empirique est un estimateur de la fonction quantile  $Q_Y(\tau)$  lorsque la fonction de distribution de probabilité  $F_Y$  de la variable  $Y$  est inconnue. Considérons  $Y_1, Y_2, \dots, Y_n$  un échantillon de  $n$  observations dont la distribution de probabilité est inconnue. La fonction de distribution empirique, notée  $\hat{F}_Y$ , se détermine par le quotient entre le nombre des observations inférieures ou égales à une quantité spécifiée  $y$  et le nombre total des observations :

$$\hat{F}_Y(y) = \frac{\sum_{i=1}^n 1_{\{Y_i \leq y\}}}{n}.$$

De manière analogue à l'équation (1), on peut en déduire la fonction quantile empirique :

$$\hat{Q}_Y(\tau) = \hat{F}_Y^{-1}(y) = \inf \left\{ y : \frac{\sum_{i=1}^n 1_{\{Y_i \leq y\}}}{n} \geq \tau \right\}, 0 < \tau < 1.$$

### 2.2.3 Définition de la méthode de régression quantile

La méthode RQ a été formalisée par Koenker et Bassett (1978) comme une extension de la relation entre une ou plusieurs variables indépendantes  $X_j$  ( $j = 1, \dots, p$ ) et la moyenne conditionnelle de variable dépendante  $Y$  sachant  $X_j = x_{ji}$  (régression standard) au lien entre ces variables  $X_j$  et les quantiles conditionnels de  $Y$  sachant  $X_j = x_{ji}$ . Cette dernière citée, comme décrit par Buchinsky (1998), est le  $\tau$ -ème quantile de  $Y$  sachant  $X_1, X_2, \dots, X_p$ , donné par :

$$Q_{y_i}(\tau|X_1 = x_{1i}, \dots, X_p = x_{pi}) = \beta_0(\tau) + \beta_1(\tau)x_{1i} + \dots + \beta_p(\tau)x_{pi} + \tilde{\epsilon}_i(\tau),$$

où  $0 < \tau < 1$  et les termes d'erreurs  $\tilde{\epsilon}_i(\tau)$ ,  $i = 1, \dots, n$ , sont indépendants et identiquement distribués. La seule hypothèse nécessaire ici est la nullité du  $\tau$ -ème quantile conditionnel du terme d'erreur :

$$Q_{\tilde{\epsilon}_i(\tau)}(\tau|X_1 = x_{1i}, \dots, X_p = x_{pi}) = 0.$$

Ainsi, la méthode RQ va au-delà de la méthode MCO en permettant la détermination de l'effet des variables indépendantes sur la forme et l'ampleur de la distribution de la variable d'intérêt, avec l'utilisation de plusieurs différentes valeurs de  $\tau$  et non seulement de la moyenne de celle-ci. Ceci étant, les coefficients de régression quantile  $\beta_j(\tau)$ ,  $j = 1, \dots, p$ , représentent alors la manière dont les quantiles spécifiés se modifient lors d'un changement d'une unité d'une variable  $X_j$ ,  $j = 1, \dots, p$ , (Koenker, 2005).

## 2.2.4 Estimation des paramètres

Comme la régression standard estime les paramètres du modèle par la fonction moyenne conditionnelle, la méthode RQ offre une démarche plus riche pour estimer les paramètres des modèles en utilisant l'ensemble des quantiles conditionnels. Celle-ci consiste à estimer les coefficients de régression en minimisant la somme pondérée des valeurs des termes d'erreurs positifs et négatifs, respectivement par le  $\tau$  d'intérêt et son complémentaire  $(1 - \tau)$ . Dans le cadre de la méthode RQ, un coefficient de régression est donc une fonction de  $\tau$ , alors que dans la régression standard, le coefficient de régression est une valeur unique pour toute la distribution.

Soient  $\beta(\tau)$ , le vecteur des coefficients de régression à estimer,  $\beta(\tau) = (\beta_0(\tau), \beta_1(\tau), \dots, \beta_p(\tau))'$  et  $\mathbf{b}(\tau)$  un vecteur d'estimation de  $\beta(\tau)$ ,  $\mathbf{b}(\tau) = (b_0(\tau), b_1(\tau), \dots, b_p(\tau))'$ . Les paramètres à estimer correspondent à la meilleure estimation  $\mathbf{b}(\tau)$  du vecteur  $\beta(\tau)$ , déterminée en minimisant la somme pondérée des valeurs positives et négatives des différences entre la valeur observée  $y_i$  de  $Y_i$  et la valeur prédite  $b_0(\tau) + b_1(\tau)x_{1i} + \dots + b_p(\tau)x_{pi}$  de  $y_i$ , notée  $\hat{y}_i$  ( $i = 1, \dots, n$ ), respectivement par le  $\tau$  d'intérêt et son complémentaire  $(1 - \tau)$ . La somme pondérée est donnée par :

$$\sum_{i \in \{i | y_i \geq \mathbf{x}'_i \mathbf{b}(\tau)\}} \tau |y_i - \mathbf{x}'_i \mathbf{b}(\tau)| + \sum_{i \in \{i | y_i < \mathbf{x}'_i \mathbf{b}(\tau)\}} (1 - \tau) |y_i - \mathbf{x}'_i \mathbf{b}(\tau)|,$$

où  $\mathbf{x}_i = (1, x_{1i}, x_{2i}, \dots, x_{pi})'$ ,  $i = 1, \dots, n$ .

Formellement, cette somme peut être écrite comme :

$$\sum_{i=1}^n \rho_\tau(\tilde{\epsilon}_i),$$

où  $\rho_\tau(\tilde{\epsilon}_i)$  est la fonction des termes d'erreurs définie par :

$$\rho_\tau(\tilde{\epsilon}_i) = \tilde{\epsilon}_i(\tau - 1_{\{\tilde{\epsilon}_i < 0\}}),$$

$0 < \tau < 1$ ,  $1_{\{\tilde{\epsilon}_i < 0\}} = 1$  si  $\tilde{\epsilon}_i < 0$  et  $0$  si  $\tilde{\epsilon}_i \geq 0$ .

Le principe de la méthode RQ est simple. Toutefois sa réalisation numérique reste complexe. Koenker et Bassett (1982) conçoivent l'algorithme pour estimer les coefficients de régression à partir des hypothèses ci-dessous :

1. le modèle de régression définit un hyperplan de  $p$  dimensions qui passe par au moins  $p + 1$  observations ;
2. il y a tout au plus  $n\tau$  observations en dessous et au moins  $(n - p - 1)\tau$  au-dessus de n'importe quel hyperplan de la  $\tau$ -ème régression quantile ;

et qui conduisent aux divers implémentations d'algorithmes itératives des logiciels *R* et *SAS*.

La meilleure estimation de  $\beta(\tau)$ ,  $\mathbf{b}(\tau)$ , a la propriété suivante :

$$\mathbf{b}(\tau) = \underset{\beta(\tau) \in \mathbb{R}^{p+1}}{\operatorname{argmin}} V(\beta(\tau); \tau),$$

où

$$V(\beta(\tau); \tau) = \sum_{i=1}^n \rho_{\tau}(y_i - \mathbf{x}'_i \beta(\tau)) \quad (2)$$

est la fonction objective du modèle avec  $p + 1$  paramètres à estimer.

### 2.2.5 Le pseudo $R^2 : R^1$

Le critère de qualité d'ajustement du modèle établi avec la méthode RQ est analogue à celui avec la méthode MCO, basée sur la mesure  $R^2$ . Le pseudo  $R^2$ , noté  $R^1$ , mesure la qualité du modèle avec la méthode RQ pour le quantile spécifié, en termes de somme pondérée des valeurs absolues des résidus. Comme  $R^2$ ,  $0 < R^1 < 1$ . Toutefois, contrairement à  $R^2$  qui mesure une qualité globale sur l'ensemble de la distribution conditionnelle,  $R^1$  mesure une qualité locale pour chaque quantile.  $R^1$  est donc en mesure d'explorer plus d'informations à différents niveaux de la distribution conditionnelle de la variable d'intérêt. Comme l'ont proposé Koenker et Machado (1999),  $R^1$  est motivé par la mesure de qualité pour la méthode MCO,  $R^2$ , où la somme des carrés des résidus est remplacée par la fonction objective du modèle *non-constraint* (avec  $p$  variables indépendantes) et la somme des carrés totaux par la fonction objective du modèle *constraint* (avec uniquement la constante).

Cela peut se traduire par les hypothèses ci-dessous, sur les coefficients excepté la constante  $\beta_0(\tau)$  :

$$H_0 \quad : \quad \beta_1(\tau) = \dots = \beta_p(\tau) = 0,$$

$$H_1 \quad : \quad \text{Au moins un des } \beta_j(\tau) \text{ est différent de 0.}$$

$R^1$  indique alors le rapport de la différence entre les fonctions objectives issues de chacune des hypothèses et la fonction objective sous l'hypothèse  $H_0$  :

$$R^1 = \frac{\tilde{V}(\tau) - \hat{V}(\tau)}{\tilde{V}(\tau)}, \quad (3)$$

où  $\tilde{V}(\tau)$  et  $\hat{V}(\tau)$  sont respectivement les fonctions objectives des modèles *constraint* et *non-constraint* ; voir l'expression (2). De plus, à partir de l'expression (3),  $R^1$  peut être utilisé pour comparer plusieurs modèles emboîtés.

### 3 Tests sur les quantiles

#### 3.1 Test d'égalité des quantiles

Le test décrit par Li *et al.* (2012) est assimilable au test ANOVA. Cependant, au lieu de comparer les moyennes de  $k \geq 2$  populations indépendantes et de distribution normale  $\mathcal{N}(\mu_i, \sigma_i^2)$ ,  $i = 1, \dots, k$ , il compare les quantiles de ces  $k$  populations, en confrontant les hypothèses ci-dessous :

$$H_0 \quad : \quad \tau_{l1} = \dots = \tau_{lk},$$

$$H_1 \quad : \quad \tau_{li} \neq \tau_{lj}, \quad i \neq j, \quad i, j \in \{1, \dots, k\},$$

où  $\tau_{li} = \mu_i + \sigma_i z_l$  est le  $l$ -ème quantile de la  $i$ -ème population et  $z_l$  le  $l$ -ème quantile de la distribution normale standard.

Plus formellement, les hypothèses s'écrivent :

$$H_0 \quad : \quad H\theta = 0,$$

$$H_1 \quad : \quad H\theta \neq 0,$$

avec

$$H_{(k-1) \times k} = \begin{pmatrix} 1 & 0 & \dots & 0 & -1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -1 \end{pmatrix}, \quad \theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \vdots \\ \theta_k \end{pmatrix}, \quad \theta_i = \tau_{li}, \quad i = 1, \dots, k.$$

La statistique de test  $T$  est donnée par :

$$T = \tilde{n}(H\hat{\theta} - H\theta)'(HAH')^{-1}(H\hat{\theta} - H\theta) \sim \chi_{k-1}^2,$$

où  $\tilde{n}$  est la plus grande taille d'échantillons tirés parmi les  $k$  populations,

$$A = \text{diag} \left( \frac{\sigma_1^2}{\lambda_1} \left(1 + \frac{z_l^2}{2}\right), \dots, \frac{\sigma_k^2}{\lambda_k} \left(1 + \frac{z_l^2}{2}\right) \right),$$

$\lambda_i = \lim_{\tilde{n} \rightarrow +\infty} \left(\frac{n_i}{\tilde{n}}\right)$ ,  $0 < \lambda_i \leq 1$ ,  $i = 1, \dots, k$ ,  $n_i$  est la taille de l'échantillon  $i$  et  $\chi_{k-1}^2$  est la loi de Khi-carré à  $(k - 1)$  degrés de liberté.

L'hypothèse  $H_0$  est rejetée lorsque la *p*valeur correspondant au test, à savoir *p*valeur =  $P(T \geq t_{obs}) = 1 - F_{\chi_{k-1}^2}(t_{obs})$ , est inférieure au niveau de signification  $\alpha$ , où  $t_{obs}$  est la valeur calculée de  $T$  et  $F_{\chi_{k-1}^2}(\cdot)$  est la fonction de distribution cumulative de la distribution de Khi-carré à  $(k - 1)$  degrés de liberté .

#### 3.2 Tests en régression quantile

Les tests présentés dans cette section ont été développés par Furno (2011). Ils portent globalement sur l'hypothèse d'invariance des coefficients de régressions comme décrit par Johnston et Di Nardo (1997), en mettant en confrontation un modèle de régression *contrainte* et un ou plusieurs autres modèles de régression *non-contrainte*. Chow (1960) utilise la même démarche lorsqu'il analyse des données temporelles. Nous présentons ci-dessous ces tests.

### 3.2.1 Test de qualité de prédiction

Soit  $n$  la taille de l'échantillon d'étude. L'évaluation de la qualité de prédiction d'un modèle construit est déterminante lorsqu'on veut prédire une réponse correspondant aux variables indépendantes en présence dans le modèle. Le test de qualité de prédiction consiste d'abord à estimer les paramètres du modèle sous les  $n_1$  ( $n_1 \approx \frac{n}{2}$ ) premières observations de l'échantillon d'étude, puis d'utiliser le modèle construit pour prédire les valeurs de la variable dépendante des données restantes. Cette démarche permet de comparer la constance des coefficients estimés entre les modèles de régression *contrainte* (le modèle est identique pour toutes les  $n$  observations) et *non-contrainte* (le modèle reste identique pour seulement  $n_1$  observations) définis dans le tableau 1.

Modèles de régression	
Contrainte	$Q_{y_i}(\tau x_i) = \beta_0(\tau) + \beta_1(\tau)x_{1i} + \dots + \beta_p(\tau)x_{pi}, i = 1, \dots, n$
Non-contrainte	$Q_{y_i}(\tau x_i) = \beta'_0(\tau) + \beta'_1(\tau)x_{1i} + \dots + \beta'_p(\tau)x_{pi}, i = 1, \dots, n_1$ ( $n_1 \approx \frac{n}{2}$ )

TABLE 1 – Modèles comparés pour le test de qualité de prédiction

Autrement dit, cela permet d'apprécier la stabilité des paramètres à estimer sans tenir compte de la taille d'échantillon utilisée. Ainsi, on peut appréhender l'homogénéité des paramètres sous forme d'un test d'hypothèse :

$$H_0 : \begin{pmatrix} \beta_0(\tau) \\ \vdots \\ \beta_p(\tau) \end{pmatrix} = \begin{pmatrix} \beta'_0(\tau) \\ \vdots \\ \beta'_p(\tau) \end{pmatrix},$$

$$H_1 : \text{Au moins un des } \beta_j(\tau) \text{ est différent de } \beta'_j(\tau), j = 0, \dots, p.$$

La statistique de test  $C^1$  est donnée par :

$$C^1 = \frac{(\tilde{V}(\tau) - \hat{V}(\tau)) / \Delta_{(ddl_{H_0}, ddl_{H_1})}}{\hat{V}(\tau) / ddl_{H_1}} \sim F(n_2, n_1 - (p + 1)),$$

où  $\Delta_{(ddl_{H_0}, ddl_{H_1})} = n - (p + 1) - (n_1 - (p + 1)) = n_2$ ,  $ddl_{H_1} = n_1 - (p + 1)$  et  $F(n_2, n_1 - (p + 1))$  est la loi de Fisher à  $n_2$  et  $(n_1 - (p + 1))$  degrés de liberté,  $\tilde{V}(\tau)$  et  $\hat{V}(\tau)$  sont les fonctions objectives des modèles *contraint* et *non-contraint* (donnés dans le Tableau 1), respectivement (voir l'expression (2)). La règle de décision est de rejeter  $H_0$  quand  $C^1 > F_{1-\alpha}(n_2, n_1 - (p + 1))$ , avec  $F_{1-\alpha}(n_2, n_1 - (p + 1))$  le quantile d'ordre  $(1 - \alpha)$  de la loi de Fisher avec  $n_2$  et  $n_1 - (p + 1)$  degrés de liberté.

#### Principe d'application du test

1. Estimer les paramètres du modèle en utilisant la méthode RQ pour :
  - (a) l'échantillon complet ( $n$  observations) ;
  - (b) le sous-échantillon de  $n_1$  premières observations ;
2. Évaluer la fonction objective pour chacun des deux modèles *contraint* et *non-contraint* ;
3. Calculer la statistique de test  $C^1$  ;
4. Rejeter l'hypothèse nulle au seuil  $\alpha$  si la valeur calculée de  $C^1 > F_{1-\alpha}(n_2, n_1 - (p + 1))$ .

### 3.2.2 Test de stabilité des paramètres

Le test de stabilité sur les paramètres ou encore le test de changement structurel vise à constater statistiquement des modifications de comportements dans l'échantillon soumis à l'étude. Pour ce faire, on doit caractériser au mieux ce qui permet de définir les sous-échantillons à confronter (en utilisant des informations externes ou une variable indépendante disponible dans la donnée, comme variable de segmentation) et déceler la nature du changement survenu (modification ou non des coefficients relatifs aux variables indépendantes).

En considérant  $G$  groupes<sup>1</sup> de  $n_i$  observations ( $n_i > p + 1$ ), le test de stabilité repose sur la comparaison des modèles de régression *contrainte* (les coefficients sont les mêmes quel que soit le groupe étudié) et *non-contrainte* (les coefficients ne sont pas nécessairement identiques sur les différents groupes) pour déterminer si, sur ces groupes étudiés, les coefficients sont identiques. Le tableau 2 présente ces différents modèles.

Modèles de régression	
Contrainte	$Q_{y_i}(\tau x_i) = \beta_0(\tau) + \beta_1(\tau)x_{1i} + \dots + \beta_p(\tau)x_{pi}, i = 1, \dots, n$ $Q_{y_i}(\tau x_i) = \beta_0^1(\tau) + \beta_1^1(\tau)x_{1i} + \dots + \beta_p^1(\tau)x_{pi}, i = 1, \dots, n_1$
Non-contrainte	$\vdots$ $Q_{y_i}(\tau x_i) = \beta_0^G(\tau) + \beta_1^G(\tau)x_{1i} + \dots + \beta_p^G(\tau)x_{pi}, i = 1, \dots, n_G$

TABLE 2 – Modèles comparés pour le test de stabilité des paramètres. Les différents sous-groupes ont été constitués à partir d'une variable de segmentation

Cette comparaison informe sur la nécessité ou non de distinguer les modèles de régression à  $(p + 1)$  paramètres dans les  $G$  groupes, en appréhendant la question sous les tests d'hypothèses suivants :

$$H_0 : \begin{pmatrix} \beta_0(\tau) \\ \vdots \\ \beta_p(\tau) \end{pmatrix} = \begin{pmatrix} \beta_0^1(\tau) \\ \vdots \\ \beta_p^1(\tau) \end{pmatrix} = \dots = \begin{pmatrix} \beta_0^G(\tau) \\ \vdots \\ \beta_p^G(\tau) \end{pmatrix},$$

$$H_1 : \text{Au moins un des coefficients diffère des autres.}$$

La statistique de test s'appuie sur les sommes pondérées des valeurs absolues de résidus des modèles de régression *contrainte* et *non-contrainte*. Elle s'écrit :

$$\tilde{C}^1 = \frac{[\tilde{V}(\tau) - \sum_{i=1}^G \hat{V}_i(\tau)] / \Delta_{(ddl_{H_0}, ddl_{H_1})}}{\sum_{i=1}^G \hat{V}_i(\tau) / ddl_{H_1}} \sim F((G-1)(p+1), n - G(p+1)),$$

$$\text{où } \Delta_{(ddl_{H_0}, ddl_{H_1})} = n - (p+1) - \sum_{i=1}^G (n_i - (p+1)) = \sum_{i=1}^{G-1} (p+1) = (G-1)(p+1) \text{ et } ddl_{H_1} = \sum_{i=1}^G (n_i -$$

1. Un groupe est un sous-échantillon de taille  $n_i$  obtenu du tri des données selon la variable de segmentation, telle que  $\sum_{i=1}^G n_i = n$ .



$(p + 1) = n - G(p + 1)$  et  $\tilde{V}(\tau), \hat{V}_i(\tau)$  sont les fonctions objectives des modèles de régression *contrainte* et *non-contrainte* (donnés dans le tableau 2), respectivement (voir l'expression (2)). Sous  $H_0$ , la statistique  $\tilde{C}^1$  suit la loi de Fisher à  $(G - 1)(p + 1)$  et  $(n - G(p + 1))$  degrés de liberté. La région critique (*R.C.*) du test s'écrit :

$$R.C. : \tilde{C}^1 > F_{1-\alpha}((G - 1)(p + 1), n - G(p + 1)),$$

où  $F_{1-\alpha}((G - 1)(p + 1), n - G(p + 1))$  est le quantile d'ordre  $(1 - \alpha)$  de la loi de Fisher à  $(G - 1)(p + 1)$  et  $(n - G(p + 1))$  degrés de liberté.

### Principe d'application du test

1. Regrouper les données selon la variable de segmentation ;
2. Effectuer la régression avec la méthode RQ sur :
  - (a) l'échantillon complet ( $n$  observations) ;
  - (b) chacun des  $G$  groupes ;
3. Évaluer la fonction objective pour toutes les  $(G + 1)$  régressions ;
4. Calculer la statistique de test  $\tilde{C}^1$  ;
5. Conclure avec *R.C.* en rejetant l'hypothèse nulle ( $H_0$ ) au seuil  $\alpha$ , si la valeur calculée de  $\tilde{C}^1 > F_{1-\alpha}((G - 1)(p + 1), n - G(p + 1))$  (c'est-à-dire que l'on ne peut pas affirmer l'homogénéité des paramètres inter-groupes).

### Remarque

Les tests présentés dans les sections 3.2.1 et 3.2.2 restent valables lorsque les paramètres du modèle de régression sont estimés par la méthode MCO. Cependant, les statistiques de test  $C^1, \tilde{C}^1$  sont notées  $C$  et  $\tilde{C}$  respectivement et les fonctions objectives sont remplacées par les sommes des carrés des résidus.

## 4 Application

L'étude présentée tout au long de cette section a pour objectif de mettre en œuvre la régression RQ et les tests décrits précédemment. Cette approche permet globalement de déterminer les paramètres du meilleur modèle RQ contextuel à différentes échelles de la distribution conditionnelle des scores PISA par rapport à des variables indépendantes et d'apprécier leur stabilité sur différents sous-échantillons.

### 4.1 Présentation des données

#### 4.1.1 Origine des données

L'échantillon utilisé pour ce travail provient de l'enquête complémentaire PISA réalisée en 2006 (voir [http://www.irdp.ch/recherche/pisa/ocde\\_pisa2000\\_2003.html](http://www.irdp.ch/recherche/pisa/ocde_pisa2000_2003.html)).

Cette enquête intègre le troisième exercice triennal de PISA. Elle concerne en partie les élèves de 9<sup>e</sup> année (indépendamment de leur âge), correspondant à la fin de scolarité obligatoire en Suisse. Son objectif est de fournir des informations permettant de comparer les trois grandes régions linguistiques (de langues française, allemande et italienne), ainsi que les cantons romands.

A cet effet, les élèves ont dû fournir plusieurs renseignements parmi tant d'autres. Ceux-ci sont les suivants (Nidegger, 2008) :

- leurs caractéristiques (âge, sexe) ;
- leurs environnements familiaux (la composition de la famille, le niveau éducationnel et professionnel des parents, les biens à disposition, le pays de provenance, le canton de résidence, les langues parlées à la maison, les activités avec les parents) ;
- leurs environnements scolaires (le programme d'étude suivi, les cours complémentaires ou spéciaux, les devoirs personnels, la grandeur des classes, la vision portée sur les enseignants et l'école) ;
- leurs familiarités avec les nouvelles technologies d'informations, etc.

Les scores PISA obtenus dans les différents domaines d'évaluation, à savoir les mathématiques, les sciences et la lecture ont été également considérés. Nous avons utilisé uniquement les données provenant de la Suisse romande. Les poids d'échantillonnage y associés n'ont pas pu être pris en considération, car l'application des tests statistiques présentés dans la section 3 en présence des poids reste encore à développer dans la littérature. Les résultats présentés dans notre analyse diffèrent donc légèrement de ceux produits par Nidegger (2008).

#### 4.1.2 Description des variables d'étude

Les variables prises en considération dans notre étude concernent les scores en mathématiques, en sciences, en lecture (variables d'intérêts), l'origine de la famille, le canton de résidence, la langue parlée à la maison, le genre, le niveau socio-économique des parents, les niveaux personnels d'appréciation en mathématiques, en sciences et en lecture et le niveau de scolarités souhaité par les parents (voir également le chapitre 7, Nidegger, 2008). Après l'exclusion des élèves n'ayant pas répondu au moins à une des questions, le jeu de données contient 8627 élèves ayant renseigné chacune des informations regroupées ci-dessus. Il est détaillé dans le tableau 3.

Variabes	Définition	Type	Modalités
math	scores PISA en mathématiques	continue	
sciences	scores PISA en sciences	continue	
lecture	scores PISA en lecture	continue	
âge	âge des élèves	continue	
canton	les cantons de Suisse romande	nominale	Berne-fr, Fribourg-fr, Genève, Jura, Neuchâtel, Vaud, Valais-fr
sexe	genre	nominale	filles = 1, garçon = 2
immig	origine de la famille	nominale	Suisse = 1, autres = 2
langue	langue parlée à la maison	nominale	français = 1, autres = 2
NSE	niveau socio-économique des parents	ordinaire	très bas = 1, bas = 2, élevé = 3, très élevé = 4
NSP	niveau de scolarité souhaité par les parents	ordinaire	école obligatoire = 1, école professionnelle = 2, maturité = 3, haute école professionnelle = 5, université = 5
NCM	niveau personnel d'appréciation en mathématiques	ordinaire	bas = 1, élevé = 2
NCS	niveau personnel d'appréciation en sciences	ordinaire	bas = 1, élevé = 2
NCL	niveau personnel d'appréciation en lecture	ordinaire	bas = 1, élevé = 2

TABLE 3 – Description de variables d'étude

## 4.2 Analyses et résultats

### 4.2.1 Coefficients de régressions

Les tableaux 4 - 6 (cf. annexes) résument les effets des caractéristiques des élèves sur les scores PISA en mathématiques, en sciences et en lecture. L'introduction des variables ordinales et nominales présentées dans le tableau 3 dans nos modèles nécessitent le choix des modalités de référence. Ainsi, les coefficients des différents modèles (MCO et RQ) sont interprétés par rapport aux références suivantes : le canton (Fribourg-fr), le genre (filles), l'origine de la famille (Suisse), la langue parlée à la maison (français), le niveau socio-économique (très bas), le niveau de scolarité souhaité par les parents (école obligatoire), le niveau personnel d'appréciation en mathématiques (bas), le niveau personnel d'appréciation en sciences (bas) et le niveau personnel d'appréciation en lecture (bas). Les coefficients présentés dans les tableaux 4 - 6 sont donc associés aux modalités restreintes (les modalités de référence n'ont pas des coefficients associés dans les modèles). Par exemple, on note dans ces tableaux par *sexe*.2, la deuxième modalité de la variable *sexe*, c'est-à-dire les garçons (cf. tableau 3). On suit une notation similaire pour toutes les autres modalités restreintes.

La première colonne dans les tableaux 4 - 6 présente les paramètres estimés du modèle établi avec la méthode MCO et les autres, avec la méthode RQ pour les quantiles de 5%, 25%, 50%, 75% et 95%. Les interprétations de ces effets ne sont que suggestives des causes des réelles performances des élèves.

Les modèles incluent, à partir de l'échantillon initial, respectivement 10 variables explicatives pour le premier intérêt et 9 pour les deux derniers, sélectionnées par les méthodes de sélection automatiques de variables pour MCO (la méthode *stepwise* avec le critère AIC a été appliquée). Ceux-ci sont présentés ci-dessous :

$$\begin{aligned} \text{math}_i &= \beta_0^{\text{math}} + \beta_1^{\text{math}} \text{sexe.2} + \beta_2^{\text{math}} \text{âge} + \beta_3^{\text{math}} \text{NSP.2} + \beta_4^{\text{math}} \text{NSP.3} + \\ &\beta_5^{\text{math}} \text{NSP.4} + \beta_6^{\text{math}} \text{NSP.5} + \beta_7^{\text{math}} \text{NCS.2} + \beta_8^{\text{math}} \text{NCM.2} + \\ &\beta_9^{\text{math}} \text{NCL.2} + \beta_{10}^{\text{math}} \text{Berne\_fr} + \beta_{11}^{\text{math}} \text{Genève} + \beta_{12}^{\text{math}} \text{Jura} + \\ &\beta_{13}^{\text{math}} \text{Neuchâtel} + \beta_{14}^{\text{math}} \text{Vaud} + \beta_{15}^{\text{math}} \text{Valais\_fr} + \beta_{16}^{\text{math}} \text{NSE.2} + \\ &\beta_{17}^{\text{math}} \text{NSE.3} + \beta_{18}^{\text{math}} \text{NSE.4} + \beta_{19}^{\text{math}} \text{immig.2} + \beta_{20}^{\text{math}} \text{langue.2} + \\ &\varepsilon_i^{\text{math}}, \end{aligned}$$

$$\begin{aligned} \text{sciences}_i &= \beta_0^{\text{scie}} + \beta_1^{\text{scie}} \text{sexe.2} + \beta_2^{\text{scie}} \text{âge} + \beta_3^{\text{scie}} \text{NSP.2} + \beta_4^{\text{scie}} \text{NSP.3} + \\ &\beta_5^{\text{scie}} \text{NSP.4} + \beta_6^{\text{scie}} \text{NSP.5} + \beta_7^{\text{scie}} \text{NCS.2} + \beta_8^{\text{scie}} \text{NCL.2} + \\ &\beta_9^{\text{scie}} \text{Berne\_fr} + \beta_{10}^{\text{scie}} \text{Genève} + \beta_{11}^{\text{scie}} \text{Jura} + \beta_{12}^{\text{scie}} \text{Neuchâtel} + \\ &\beta_{13}^{\text{scie}} \text{Vaud} + \beta_{14}^{\text{scie}} \text{Valais\_fr} + \beta_{15}^{\text{scie}} \text{NSE.2} + \beta_{16}^{\text{scie}} \text{NSE.3} + \\ &\beta_{17}^{\text{scie}} \text{NSE.4} + \beta_{18}^{\text{scie}} \text{immig.2} + \beta_{19}^{\text{scie}} \text{langue.2} + \varepsilon_i^{\text{scie}}, \end{aligned}$$

$$\begin{aligned} \text{lecture}_i &= \beta_0^{\text{lect}} + \beta_1^{\text{lect}} \text{sexe.2} + \beta_2^{\text{lect}} \text{âge} + \beta_3^{\text{lect}} \text{NSP.2} + \beta_4^{\text{lect}} \text{NSP.3} + \\ &\beta_5^{\text{lect}} \text{NSP.4} + \beta_6^{\text{lect}} \text{NSP.5} + \beta_7^{\text{lect}} \text{NCS.2} + \beta_8^{\text{lect}} \text{NCL.2} + \\ &\beta_9^{\text{lect}} \text{Berne\_fr} + \beta_{10}^{\text{lect}} \text{Genève} + \beta_{11}^{\text{lect}} \text{Jura} + \beta_{12}^{\text{lect}} \text{Neuchâtel} + \\ &\beta_{13}^{\text{lect}} \text{Vaud} + \beta_{14}^{\text{lect}} \text{Valais\_fr} + \beta_{15}^{\text{lect}} \text{NSE.2} + \beta_{16}^{\text{lect}} \text{NSE.3} + \\ &\beta_{17}^{\text{lect}} \text{NSE.4} + \beta_{18}^{\text{lect}} \text{immig.2} + \beta_{19}^{\text{lect}} \text{langue.2} + \varepsilon_i^{\text{lect}}. \end{aligned}$$

Des modèles similaires sont construits pour RQ au niveau de 5%, 25%, 50%, 75% et 95%, respectivement.

A partir de la modalité de référence et en contrôlant les autres variables restantes dans le modèle, on apprend dans quelle mesure les autres modalités d'une même variable indépendante pourraient expliquer ou prédire les différences entre ces modalités. Ainsi trois possibilités d'interprétation se dégagent pour l'analyse de cette variable indépendante (Field *et al.*, 2012) :

- i - effet insignifiant : après avoir contrôlé les autres variables en présence dans le modèle, il n'y a pas de contribution significative apportée par la différence entre cette modalité restreinte et la modalité de référence. En d'autres termes, le changement dans les scores est le même si on considère un élève de la modalité de référence et un autre de cette modalité restreinte, étant donné les mêmes caractéristiques des élèves ;
- ii - effet positif : un coefficient positif et significatif impliquerait que les élèves associés à la modalité restreinte surpassent ceux de la modalité de référence, lorsque les caractéristiques des élèves sont les mêmes ;
- iii - effet négatif : un coefficient négatif et significatif pourrait révéler que les scores des élèves correspondant à la modalité restreinte sont inférieurs à ceux associés à la modalité de référence, lorsque les caractéristiques des élèves sont les mêmes.

## Résultats - méthode MCO

En général, les résultats sont du signe et de la contribution attendus à ceux donnés dans la littérature (voir Nidegger, 2008). Les garçons font mieux que les filles en mathématiques et en sciences (l'inverse se produit en lecture). Conformément à la littérature liée à l'environnement scolaire et à l'importance du milieu familial sur les performances scolaires des élèves, les modèles confirment que l'origine de la famille, le niveau socio-économique des parents, la langue parlée à

la maison, le niveau de scolarité souhaité par les parents et les niveaux personnels d'appréciation pour les matières d'intérêt sont en général tous statistiquement significatifs au seuil de 1% pour expliquer chacun des scores en mathématiques, en sciences et en lecture, respectivement. Les coefficients de la régression MCO sont donnés en annexe ; voir les tableaux 4, 5 et 6.

### Résultats - méthode RQ

L'approche par la méthode RQ évalue les effets des mêmes variables indépendantes sur les différentes échelles de la répartition empirique des scores PISA. Nous avons utilisé pour estimer les paramètres des modèles RQ la librairie R *quantreg* (Koenker, 2013). Les coefficients sont estimés pour les quantiles de 5%, 25%, 50%, 75% et 95%, respectivement. Les coefficients de la régression RQ sont donnés en annexe ; voir les tableaux 4, 5 et 6.

Les tableaux 4, 5 et 6 dans l'annexe montrent des différences entre les paramètres des modèles MCO et celles de modèles RQ pour chaque matière et pour chaque quantile considéré. Ces différences suggèrent l'importance d'utiliser des modèles RQ pour la modélisation des scores. Dans l'approche RQ, l'impact des variables sur les scores de chacune des trois matières est beaucoup plus complexe que dans le cas MCO. Il se reflète dans le changement de signe, contribution et signification des coefficients estimés. Nous donnons ci-dessous un aperçu des coefficients estimés :

- âge des élèves : les paramètres  $\gamma$  associés sont statistiquement significatifs au seuil de 1% dans toutes les modélisations des scores dans les trois matières. Ceux-ci sont de signe négatifs et révèlent les mauvais résultats des élèves plus âgés, pour chacun des quantiles spécifiés. Toutefois, dans les trois matières, on remarque que l'âge a moins d'effet pour les niveaux élevés de la distribution des scores que pour les niveaux faibles. En effet, les paramètres se rétrécissent lorsqu'on évolue des scores faibles vers ceux plus élevés (quantiles de 5% à 95%).
  - cantons : le canton Fribourg-fr est utilisé comme référence de la variable *canton* pour comparer les scores PISA de ce canton et chacun des cantons romands, en contrôlant les autres variables indépendantes utilisées dans la construction des différents modèles. Le choix de Fribourg-fr a été fait par rapport aux performances très bonnes obtenues par les élèves fribourgeois. Les élèves du canton Fribourg-fr réalisent de meilleurs résultats que les autres élèves de Suisse romande dans tous les différents modèles établis pour chacune des trois disciplines (les signes des coefficients associés aux différentes modalités de la variable *canton* sont en général significatifs et négatifs). Dans les trois matières, comme dans le cas de la variable *âge*, on remarque que la différence entre le canton de référence et les autres cantons a moins d'effet pour les niveaux élevés de la distribution des scores que pour les niveaux faibles, car les valeurs absolues des paramètres se rétrécissent lorsqu'on évolue des scores faibles vers ceux plus élevés (quantiles de 5% à 95%). Quelques exceptions à ce constat sont observées (plus visiblement en sciences), où le décalage entre le canton de référence et les autres cantons est plus petit pour les élèves très faibles (quantiles de 5%) que pour des autres élèves (par exemple pour le quantile de 25%).
- En général, les coefficients conservent le même signe négatif et la même significativité statistique au seuil de 1% tout le long de la distribution des scores PISA comme pour MCO. Des exceptions sont présentes dans la modélisation des scores extrêmes supérieurs (quantiles de 95%) : en sciences (Berne-fr, coefficient -9.57, erreur standard 6.40) et en lecture (Berne-fr, coefficient -7.93, erreur standard 5.58 et Vaud, coefficient -4.67, erreur standard

3.49) où ils ne sont pas significatifs (la différence entre le canton de référence et les autres cantons mentionnés ci-dessus n'est pas significative quand les autres caractéristiques présentes dans le modèle sont fixées).

- genre des élèves : les coefficients associés à cette variable indépendante sont statistiquement significatifs au seuil de 1% dans les différents modèles, pour chacune des variables d'intérêts. Les différences tout le long de la distribution des scores des garçons par rapport aux filles sont positives dans la modélisation des scores en mathématiques et en sciences, mais négatives dans celle des scores en lecture. Elles augmentent au fur et à mesure qu'on se déplace des scores inférieurs vers les scores supérieurs en mathématiques et en sciences (l'écart des scores entre filles et garçons est plus grand pour les élèves performants que pour les élèves faibles). Cet effet est contraire dans les scores en lecture où ces différences diminuent quand on se rapproche des scores supérieurs.
- origine de la famille : les différences des élèves non-natifs par rapport aux élèves natifs sont toutes statistiquement significatives au seuil de 1% et de signe négatif dans chacun des modèles établis dans les trois matières et pour tous les quantiles. Ce qui implique que les élèves non-natifs font de moins bons résultats que les élèves natifs dans les trois matières et à tous les niveaux. Toutefois, on ne peut pas constater une décroissance monotone des effets (du quantile de 5% vers le quantile de 95%) comme dans le cas de la variable *âge*.
- langue parlée à la maison : l'impact de cette variable est presque similaire à la précédente. Les coefficients s'y référant sont statistiquement significatifs au seuil de 1% aux différents quantiles spécifiés des scores PISA, excepté ceux de 5% et de 95% en mathématiques. Ils sont tous de signe négatif (à l'exception du coefficient pour le quantile de 95% en mathématiques), ce qui montre que les élèves allophones de Suisse romande sont moins performants que les francophones. Ces écarts sont plus grands chez les élèves faibles (quantile de 5%) en lecture (coefficient -17.27, erreur standard 4.96). Une exception importante est à signaler : le coefficient pour le quantile de 95% en mathématiques n'est pas significatif ce qui montre que la langue parlée à la maison n'influence pas les performances des élèves très performants (étant donné les autres caractéristiques).
- niveau socio-économique des parents : le niveau socio-économique très défavorable a été considéré comme la modalité de référence. Généralement, l'évolution du niveau socio-économique des parents est proportionnelle aux compétences des élèves dans les études des scores PISA. Plus le niveau socio-économique est élevé, plus les performances dans les scores PISA sont supérieures. Ceci est confirmé dans les différents modèles établis. Les coefficients sont en majorité tous statistiquement significatifs au seuil de 1% dans chacun des scores considérés. On constate une exception importante pour les élèves très performants (quantile de 95%) et dans les trois matières : les différences entre les élèves ayant un niveau socio-économique favorable (niveaux 3 et 4) et ceux avec un niveau socio-économique très défavorable (niveau 1) ne sont pas statistiquement significatives. Ce qui montre que le niveau socio-économique des parents ne joue pas un rôle déterminant dans les performances de ces élèves.
- niveau de scolarité souhaité par les parents : le souhait des parents de l'arrêt de leurs enfants juste après l'école obligatoire est la modalité de référence. Les différences entre la modalité de référence et les autres modalités sont en grande partie toutes statistiquement significatives au seuil de 1% le long de la distribution des scores PISA. Toutefois, l'interprétation de cette variable dans les modèles reste complexe.
- niveaux personnels d'appréciation des trois matières : le niveau personnel d'appréciation en

mathématiques a été exclu dans la sélection automatique des variables pour les modèles MCO dont l'intérêt porte sur les scores en sciences et en lecture, mais il est inclus dans la construction des modèles pour les scores en mathématiques. Il est également statistiquement significatif (au seuil de 1%) et de signe positif dans le modèle RQ de 95% en mathématiques (coefficient 6.90, erreur standard 2.00), mais non-significatif dans les modèles RQ aux quantiles de 5%, 25%, 50% et 75%. Ce qui montre des différences considérables entre les modèles MCO et RQ construits aux quantiles spécifiés. Les élèves très performants (quantile de 95%) avec une meilleure appréciation personnelle en mathématiques (niveau 2) obtiennent des meilleurs scores dans cette matière par rapport à ceux qui ont un niveau d'appréciation moins bon.

Dans chacune des trois disciplines, les différences de scores des élèves ayant des niveaux personnels d'appréciation élevé et faible en sciences sont de signe positif et statistiquement significatifs au seuil de 1% pour tous les modèles RQ.

Les écarts dans les appréciations personnelles en lecture diffèrent des précédents. Ils ne sont ni tous significatifs, ni tous de signe positif dans les différents modèles pour chacun des trois domaines d'intérêt. Dans certains modèles, comme par exemple celui des scores faibles (quantile de 5%) en sciences, les élèves ayant un niveau supérieur d'appréciation personnelle en lecture surpassent leurs camarades ayant un faible niveau personnel d'appréciation. L'effet contraire se produit dans les scores supérieurs en sciences pour les mêmes catégories d'élèves, montrant qu'un score plus élevé en sciences est obtenu par des élèves avec un niveau d'appréciation faible en lecture. Toutefois, comme dans le cas de la variable *niveau de scolarité souhaité par les parents*, l'interprétation de cette variable dans les modèles reste complexe.

#### 4.2.2 Performance de modèles - tests de qualité de modèles

La figure 1 illustre les mesures de qualités  $R^2$  et  $R^1$  (à gauche) et les valeurs des statistiques de tests  $C$  et  $C^1$  (à droite ; voir la section 3.2.1) pour les seuils de signification de 1% et de 5% pour l'appréciation de la qualité des modèles construits, à partir de la modélisation des scores PISA en mathématiques, en sciences et en lecture, avec la méthode MCO et la méthode RQ.

Les modèles établis avec la méthode RQ pour les scores PISA dans les trois matières, étant donné les caractéristiques sélectionnées des élèves, présentent des courbes de mesures de qualité d'ajustement plutôt concaves. Ces mesures ( $R^1$ ) augmentent du quantile 5% jusqu'à celui de 50% avant de décroître jusqu'au quantile 95%.

Dans la figure 1, la droite horizontale passant par la mesure  $R^2$  est largement au-dessus des courbes définies par  $R^1$ . Il ne faut pas toutefois faire une comparaison directe entre les mesures  $R^2$  et  $R^1$ , car elles sont construites à partir de métriques différentes, même si les deux peuvent prendre des valeurs entre 0 et 1. Comme les scores PISA ont une distribution normale, on peut considérer que les modèles MCO et les modèles RQ pour le quantile de 50% ont plus ou moins la même qualité d'ajustement, même si la valeur de  $R^1$  est en dessous de celle de  $R^2$ . Les modèles RQ pour les quantiles extrêmes montrent un moins bon ajustement, en suggérant que, pour expliquer les scores extrêmes, d'autres variables (que celles d'habitude utilisées pour la moyenne) sont à ajouter dans les modèles.

Les courbes déterminées par les valeurs  $C^1$  sont presque toutes convexes et se situent en dessous du seuil de signification (1%). L'hypothèse nulle du test de qualité de prédiction (section 3.2.1) est rejetée au seuil de 1% pour tous les quantiles considérés. On conclut que les modèles

RQ ont un bon pouvoir prédictif. Bien que significatif à tous les quantiles spécifiés, ces modèles suggèrent une meilleure prédiction au centre par rapport aux extrémités de la distribution des scores, car la valeur de  $C^1$  pour le quantile de 50% a la plus petite valeur. La droite horizontale passant par la valeur de la statistique de test  $C$  (méthode MCO) se trouve en dessous des courbes de  $C^1$ . Toutefois, comme dans le cas des mesures  $R^1$  et  $R^2$ , une comparaison directe entre les valeurs de  $C$  et  $C^1$  n'est pas possible.

### 4.2.3 Différences et similitudes cantonales

Le test d'égalité des quantiles (section 3.1) a été appliqué pour tester l'égalité des quantiles des scores dans les différents cantons (indépendamment des caractéristiques des élèves). Le test rejette l'hypothèse nulle au seuil de 1% en montrant des différences significatives entre les cantons à tous les quantiles considérés (5%, 25%, 50%, 75% et 95%) et pour toutes les matières. Le même test a été appliqué ci-dessous pour des paires de cantons, en utilisant la correction de Bonferroni, en plus de celui de stabilité des coefficients.

Considérons le test de stabilité des coefficients (section 3.2.2) pour les méthodes MCO et RQ, en prenant en compte comme variable de segmentation, la variable *canton*. Nous utilisons ce test pour des comparaisons de paires de cantons ( $G = 2$  pour le test donné dans la section 3.2.2). Le choix de  $G$  a été déterminé pour mettre en évidence des différences ou des similitudes entre les cantons, étant donné les variables indépendantes considérées. Le tableau 7 (cf. annexes) résume toutes les comparaisons cantonales par paires. Le nombre de comparaisons étant grand ( $C^2_7$  comparaisons possibles), nous discutons ici quelques résultats, les plus intéressants pour la mise en évidence des différences et des similitudes cantonales.

Pour deux cantons quelconques A et B, le test de stabilité des coefficients compare les coefficients des deux modèles de régression, un construit sur les données du canton A et l'autre sur celles du canton B. Un rejet de l'hypothèse nulle d'égalité des coefficients au seuil de 1% ou de 5% pour la méthode MCO met en évidence deux modèles MCO différents pour le canton A et le canton B. En général, les résultats de l'application du test sont significatifs au seuil de 5%. Ce qui induit l'existence des deux modèles MCO différents pour chaque canton. Pour les scores en sciences et les cantons Jura et Vaud, le test n'est pas significatif.

En utilisant les tests de stabilité des coefficients dans la régression quantile, on a plus d'information sur la performance des modèles à différents quantiles. Ainsi, ces tests peuvent mettre en évidence des différences ou des similitudes cantonales au niveau des quantiles extrêmes (c'est-à-dire au niveau des élèves faibles ou performants), ce qui ne peut pas être réalisé avec les tests basés sur la méthode MCO qui ne considère que des scores moyens. Quelques exemples sont concluants. La partition des données selon la variable de segmentation *canton* - Berne-fr (694 observations), Jura (689 observations) et Neuchâtel (1403 observations) - permet d'estimer séparément le même modèle dans chacun des cantons. Ce qui permet également de comparer la stabilité des différentes contributions des caractéristiques des élèves dans la modélisation des scores extrêmes en mathématiques et en sciences, entre les cantons Berne-fr et Jura, puis celle des scores en lecture entre les cantons Jura et Neuchâtel.

Sur la figure 2, sont présentés simultanément la ligne indiquant le quantile  $F_{0.99}((p+1), n-2(p+1))$ , la ligne pour le quantile  $F_{0.95}((p+1), n-2(p+1))$ , les statistiques de test pour la stabilité des coefficients avec la méthode MCO ( $\tilde{C}$ ) et celles avec la méthode RQ ( $\tilde{C}^1$ ) aux quantiles spécifiés et pour les exemples considérés.

Le premier graphique dans la figure 2 concerne les cantons Berne-fr et Jura et les scores



en mathématiques. Le test de stabilité des coefficients pour la régression quantile de 5% rejette l'hypothèse nulle au seuil de 5% pour ces deux cantons, montrant l'existence de deux modèles RQ différents. Le même test appliqué avec la méthode MCO rejette aussi l'hypothèse nulle au seuil de 5%. Pour le même seuil, les tests de stabilité des coefficients de type RQ donnent plus ou moins les mêmes résultats que le test de type MCO pour tous les quantiles spécifiés et différents de 5% et de 95%. Au seuil de 1%, les deux tests diffèrent. Le test avec la méthode RQ pour le quantile de 5% rejette l'hypothèse nulle (montrant l'existence de deux modèles différents) contrairement à celui avec la méthode MCO. Autrement dit, conditionnellement aux variables indépendantes, le modèle de régression quantile utilisé dans chaque canton (Berne-fr et Jura) ne peut pas prédire de la même manière les scores des élèves faibles dans les deux cantons. Cette conclusion reste la même lorsqu'on s'intéresse aux élèves performants en mathématiques de ces cantons (quantile de 95%).

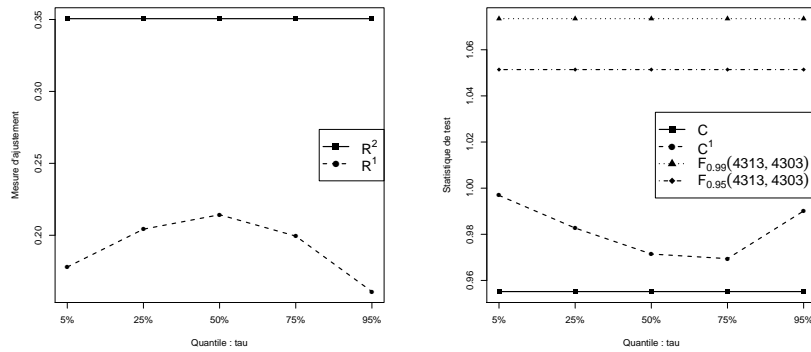
Par ailleurs, en faisant fi des variables indépendantes, on peut tester directement l'égalité des quantiles de 5% des scores pour ces deux cantons (section 3.1). L'application du test d'égalité des quantiles de 5% pour ces deux cantons est fait au seuil de  $0.05/7=0.71\%$  (dû à la correction de Bonferroni). Le résultat du test montre que les scores des élèves faibles sont significativement différents dans les deux cantons, indépendamment de toutes les autres caractéristiques des élèves, comme l'âge, le sexe, etc. Toutefois, le résultat de ce test est non significatif lorsque l'intérêt porte sur les scores des élèves très performants (quantile de 95%).

Un exemple similaire concerne les mêmes cantons et les scores en sciences (le deuxième graphique dans la figure 2).

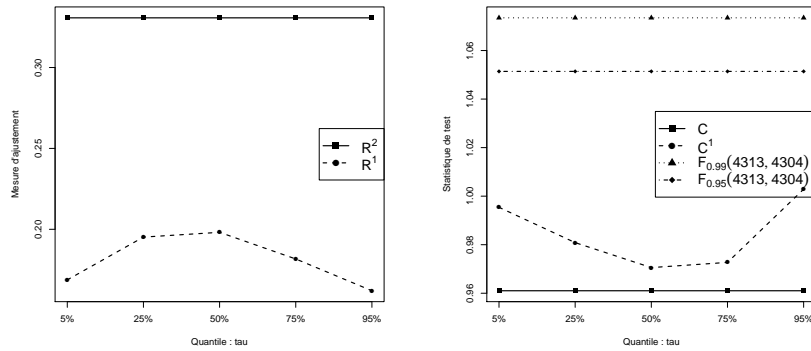
Un autre exemple concerne les modèles RQ pour sur les quantiles de 5% et de 95%, les cantons Jura et Neuchâtel et les scores en lecture (le troisième graphique dans la figure 2). On constate que des modèles différents sont à construire dans les deux cantons pour les élèves faibles ou très performants, respectivement. Par contre, pour le même intérêt, le test d'égalité des quantiles de 5% et de 95%, respectivement, n'est pas significatif au seuil de 0.71%. En d'autres termes, les scores des élèves faibles et très performants en lecture ne sont pas significativement différents dans les deux cantons, indépendamment de toutes les autres caractéristiques des élèves. En conclusion, en prenant en compte les caractéristiques des élèves, deux modèles différents sont à construire (un pour chaque canton). Sans prendre en compte leurs caractéristiques, on n'a pas d'évidence statistique de la différence des scores dans les deux cantons.

Les analyses précédentes suggèrent l'existence de différences ou non entre les scores cantonaux faibles et performants, indépendamment ou non des caractéristiques sélectionnées des élèves. En outre, si une étude doit être menée sur les élèves avec un niveau moyen de compétences en mathématiques et en sciences étant donné les mêmes caractéristiques sélectionnées dans les cantons Berne-fr et Jura, elle pourrait être réduite à un seul canton dont les conclusions pourraient être prises en compte pour l'autre canton (le modèle de prédiction est le même). Il en est de même pour les cantons Jura et Neuchâtel lorsque l'intérêt porte sur les scores en lecture et s'intéresse aux compétences moyennes des élèves étant donné les caractéristiques sélectionnées. Cependant, lorsque l'intérêt concerne les élèves avec un bas ou un haut niveau de performances, cette démarche devient impossible et des modèles différents de prédiction des scores doivent être construits. Les mêmes constats (l'analyse n'est pas montrée ici) s'appliquent aux scores en mathématiques et aux cantons Berne-fr et Vaud, aux scores en sciences et aux cantons Berne-fr et Vaud, Jura et Neuchâtel, Jura et Vaud, respectivement.

Modèles établis en mathématiques



Modèles établis en sciences



Modèles établis en lecture

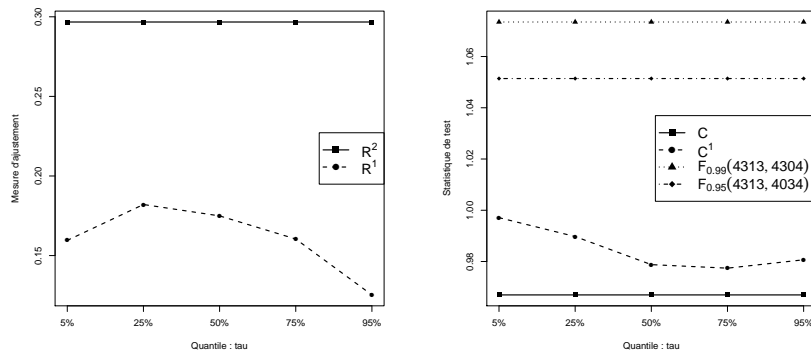


FIGURE 1 – Mesure d'ajustement (à gauche) et qualité de prédiction des scores PISA (à droite), en mathématiques, en sciences et en lecture, respectivement

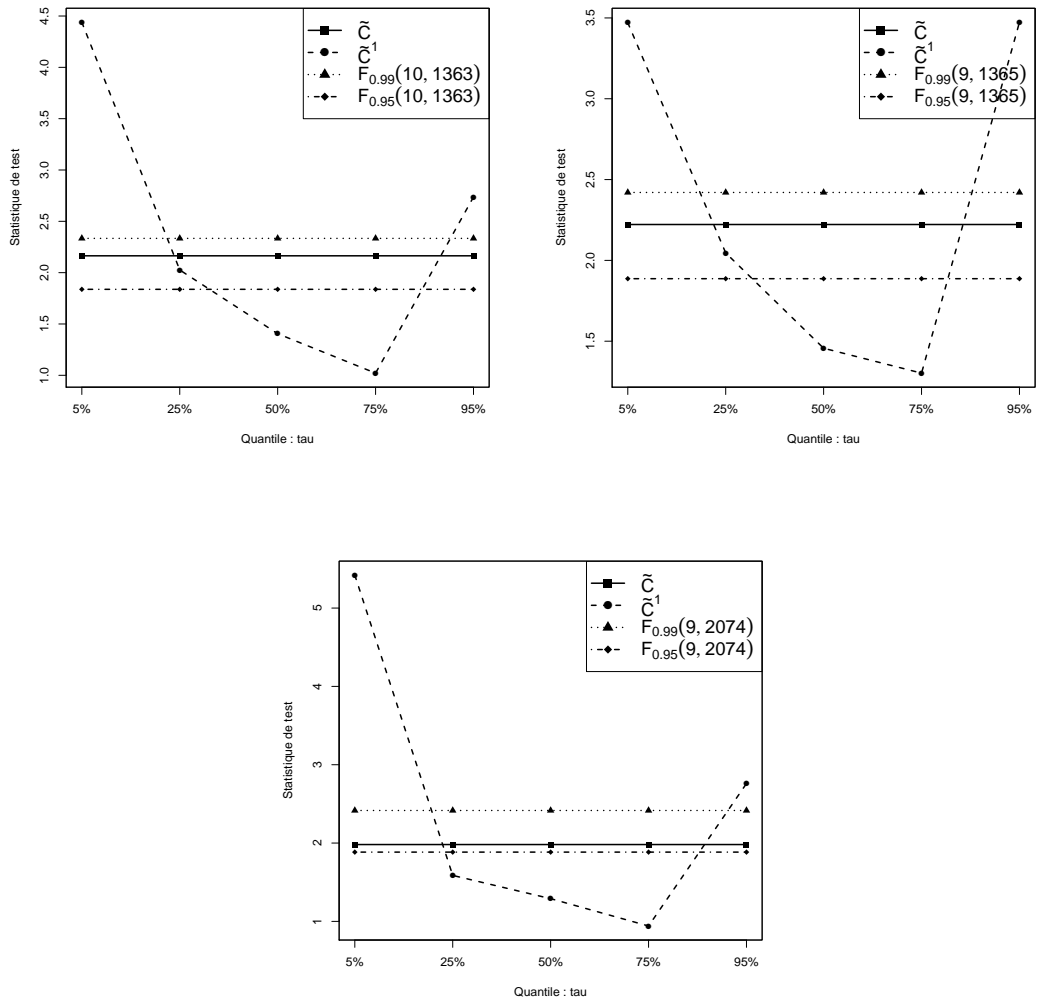


FIGURE 2 – Test de stabilité des coefficients : comparaisons inter-cantoniales des scores en mathématiques entre les cantons Berne-fr et Jura (graphique I), comparaisons inter-cantoniales des scores en sciences entre les cantons Berne-fr et Jura (graphique II) et comparaisons inter-cantoniales des scores en lecture entre les cantons Jura et Neuchâtel (graphique III)

## 5 Conclusions

L'intérêt porté de ce document est la présentation d'une méthodologie globale pour étudier les scores PISA à différents niveaux (comme les scores des élèves faibles ou performants). Celle-ci est réalisée à l'aide des quantiles de la distribution empirique des scores. Comme dans le cas de la régression linéaire (basée sur la méthode MCO), des caractéristiques des élèves (comme le sexe, la langue parlée à la maison, etc.) sont mises en relation avec leurs performances en mathématiques, en lecture et en sciences, respectivement. Le but reste le même, c'est-à-dire de construire des modèles statistiques capables d'expliquer les performances des élèves en fonction de leurs caractéristiques. Cependant, la régression linéaire s'intéresse uniquement à la moyenne des scores. La régression quantile s'intéresse à l'ensemble de la distribution des scores, permettant ainsi une modélisation plus ample des scores.

Nous avons étudié l'effet des facteurs familiaux, socio-économiques et scolaires sur les scores PISA, puis certaines différences et similitudes entre les cantons de Suisse romande. À l'aide des données de l'enquête PISA Suisse romande 2006 et de la régression quantile, nous avons identifié plusieurs conclusions intéressantes que des recherches reposant sur la méthode MCO ne puissent révéler. Nous avons constaté ainsi une hétérogénéité des coefficients dans les modèles statistiques en mathématiques, en sciences et en lecture, au fur et à mesure qu'on se déplace le long de la distribution empirique des scores. En d'autres mots, un seul modèle (comme celui construit avec la méthode MCO) ne peut pas être utilisé pour expliquer toute la diversité des scores PISA. Nous avons constaté, par exemple, à l'aide de la régression RQ, que le niveau socio-économique des parents n'est plus un facteur significatif pour expliquer les scores des élèves très performants dans toutes les trois matières. Le même constat est valable pour la langue parlée à la maison et les élèves très performants en mathématiques.

En prenant en compte des caractéristiques des élèves, nous avons constaté également que des modèles différents sont à construire en mathématiques et dans les cantons Berne-fr et Jura, Berne-fr et Vaud, lorsqu'on s'intéresse aux élèves faibles ou performants. Il en est de même en sciences pour ces élèves et les cantons Berne-fr et Jura, Berne-fr et Vaud, Jura et Neuchâtel, Jura et Vaud. Pour les scores en lecture et les élèves faibles ou performants, le constat précédent reste valable dans les cantons Jura et Neuchâtel.

Les résultats obtenus suggèrent alors que toute étude sur de telles données devrait tenir compte des quantiles de la distribution empirique des scores PISA, et pas seulement du centre de la distribution. En d'autres termes, si on s'intéresse aux élèves de niveau moyen, la régression MCO peut être utilisée pour créer des modèles statistiques. Cependant, si l'intérêt est d'étudier les scores des élèves faibles ou performants, le recours à la régression quantile et ses tests statistiques constitue une solution tout à fait valide.



## Références

- BRATTI, M., CHECCHI, D. et FILIPPIN, A. (2007). Geographical differences in Italian student's mathematical competencies : evidence from PISA 2003. *Giornale degli Economisti e Annali di Economia*, 66(3):299–333.
- BUCHINSKY, M. (1998). Recent advances in quantile regression models : A practical guideline for empirical research. *The Journal of Human Resources*, 33(1):88–126.
- CHOW, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3):591–605.
- DAVINO, C. et GOIRICELAYA, S. (2008). Quantile regression for the evaluation of student satisfaction. *Statistica Applicata*, 20:179–196.
- DRAPER, N. et SMITH, H. (1998). *Applied Regression Analysis*. Wiley, New York, third édition.
- FIELD, A., MILES, J. et FIELD, Z. (2012). *Discovering statistics using R*. Sage, Thousand Oaks.
- FRYER, R. et LEVITT, R. (2010). An empirical analysis of the gender gap in mathematics. *American Economics Journal : Applied Economics*, 2(2):210–240.
- FURNO, M. (2011). Goodness of fit and misspecification in quantile regressions. *Journal of Educational and Behavior Statistics*, 36(1):105–131.
- GIVORD, P. et D'HAULTFOEUILLE, X. (2013). La régression quantile en pratique. Rapport technique, Institut national de la statistique et des études économiques, Paris, France.  
[http://www.insee.fr/fr/publications-et-services/docs\\_doc\\_travail/doc\\_regression\\_quantile.pdf](http://www.insee.fr/fr/publications-et-services/docs_doc_travail/doc_regression_quantile.pdf).
- HAO, L. et NAIMAN, D. Q. (2007). *Quantile regression*. Quantitative applications in the social sciences. Sage, Thousand Oaks.
- JOHNSTON, J. et DI NARDO, J. (1997). *Econometric Methods*. McGraw-Hill, New York.
- KOENKER, R. (2005). *Quantile regression*. Econometric Society Monographs. Cambridge University Press, Cambridge.
- KOENKER, R. (2013). *quantreg : Quantile Regression*, R package version 5.05.  
<http://CRAN.R-project.org/package=quantreg>.
- KOENKER, R. et BASSETT, G. J. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.
- KOENKER, R. et BASSETT, G. J. (1982). An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association*, 77(378):407–415.
- KOENKER, R. et MACHADO, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94(448):1296–1310.
- LI, X., TIAN, L., WANG, J. et MUINDI, J. R. (2012). Comparison of quantiles for several normal populations. *Computational Statistics & Data Analysis*, 56(6):2129–2138.

NIDEGGER, C., éditeur (2008). *PISA 2006 : compétences des jeunes romands, résultats de la troisième enquête auprès des élèves de 9e année*. Institut de recherche et de documentation pédagogique (IRDp), Neuchâtel, Suisse.  
[http://www.irdp.ch/recherche/pisa/rapport\\_regional\\_pisa2006.pdf](http://www.irdp.ch/recherche/pisa/rapport_regional_pisa2006.pdf).

## **Annexes**



Variables	MCO	Quantiles				
		5%	25%	50%	75%	95%
constante	950.67** (19.63)	874.44** (34.24)	920.37** (29.45)	955.40** (24.86)	957.68** (28.85)	979.86** (29.69)
sexe.2	26.25** (1.31)	21.51** (2.44)	25.56** (1.87)	28.10** (1.67)	29.29** (1.82)	26.02** (2.21)
âge	-24.82** (1.24)	-26.26** (2.12)	-25.37** (1.87)	-24.98** (1.57)	-22.74** (1.83)	-20.77** (1.86)
NSP.2	49.80** (1.83)	44.08** (4.16)	54.71** (2.61)	54.33** (2.35)	51.76** (2.73)	35.46** (3.16)
NSP.3	-2.68 (1.64)	3.79 (3.81)	-2.93 (2.33)	-4.94* (2.09)	-4.35 (2.40)	-0.15 (3.06)
NSP.4	10.03** (1.99)	23.14** (5.29)	10.29** (2.90)	12.57** (2.56)	9.05** (3.18)	2.36 (3.67)
NSP.5	23.09** (1.75)	28.81** (4.61)	23.53** (2.52)	25.52** (2.17)	21.54** (2.64)	18.45** (3.46)
NCS.2	11.96** (0.97)	10.84** (1.76)	11.09** (1.38)	11.55** (1.21)	12.78** (1.36)	13.37** (1.64)
NCM.2	5.87* (3.00)	7.18 (9.79)	1.98 (4.20)	3.62 (3.86)	6.80 (4.12)	6.90** (2.00)
NCL.2	-6.62* (2.99)	-1.41 (8.62)	-2.87 (4.22)	-7.73 (4.06)	-10.34** (3.58)	-13.23** (3.11)
Berne-fr	-26.79** (2.78)	-49.99** (6.81)	-30.48** (3.50)	-27.94** (3.80)	-20.98** (4.29)	-12.06** (3.48)
Genève	-70.89** (2.40)	-66.48** (3.20)	-74.05** (3.27)	-73.77** (3.40)	-69.87** (3.19)	-59.60** (4.34)
Jura	-19.04** (2.78)	-25.34** (5.19)	-18.53** (3.84)	-20.36** (3.61)	-15.40** (3.19)	-13.25** (4.39)
Neuchâtel	-44.12** (2.28)	-54.19** (5.72)	-43.72** (3.27)	-43.00** (2.74)	-42.66** (3.11)	-29.79** (3.51)
Vaud	-31.72** (2.28)	-41.94** (4.59)	-34.07** (3.43)	-30.49** (2.96)	-29.69** (3.07)	-17.79** (4.30)
Valais-fr	-26.01** (2.25)	-27.85** (3.86)	-26.55** (3.13)	-27.18** (2.87)	-22.19** (3.03)	-19.22** (3.24)
NSE.2	16.46** (1.45)	14.11** (2.87)	16.60** (2.09)	15.24** (1.84)	16.58** (2.16)	15.16** (2.58)
NSE.3	0.52 (1.32)	-1.89 (2.67)	0.67 (1.89)	0.60** (1.66)	0.16 (1.89)	4.48 (2.26)
NSE.4	3.85** (1.28)	4.00 (2.64)	4.21* (1.79)	4.38** (1.58)	4.66** (1.74)	-0.52 (2.04)
immig.2	-30.50** (1.86)	-32.46** (4.12)	-26.97** (2.67)	-30.60** (2.26)	-31.26** (2.59)	-34.10** (3.46)
langue.2	-9.21** (2.20)	-5.74 (4.94)	-10.65** (3.31)	-13.05** (2.62)	-9.36** (3.52)	0.15 (3.19)
$R^2, R^1$	0.351	0.178	0.204	0.214	0.199	0.161
$C, C^1$	0.955	0.997	0.983	0.971	0.969	0.990

Les erreurs standards sont données entre parenthèses sous les coefficients.

Un coefficient statistiquement significatif au seuil de 5% ou de 1% sous un test bilatéral est signalé avec (\*) ou (\*\*), respectivement.

TABLE 4 – Coefficients de régressions - Mathématiques

Variables	MCO	Quantiles				
		5%	25%	50%	75%	95%
constante	819.12** (20.91)	746.27** (30.84)	834.61** (29.07)	834.20** (26.34)	791.55** (29.44)	846.74** (33.60)
sexe.2	19.60** (1.39)	13.36** (2.56)	16.93** (1.96)	19.13** (1.68)	24.11** (1.82)	22.36** (2.38)
âge	-18.18** (1.32)	-21.09** (1.90)	-21.71** (1.83)	-18.93** (1.67)	-13.72** (1.87)	-13.55** (2.12)
NSP.2	54.73** (1.95)	53.12** (3.65)	58.81** (3.13)	60.07** (2.30)	51.76** (3.13)	42.61** (3.59)
NSP.3	-0.88 (1.75)	-0.42 (3.42)	-0.27 (2.75)	-4.65* (2.10)	1.63 (2.70)	1.85 (3.10)
NSP.4	10.51** (2.13)	14.03** (4.77)	10.06** (3.96)	10.90** (2.16)	9.29** (2.99)	2.22 (4.76)
NSP.5	24.11** (1.86)	20.15** (4.27)	22.09** (3.25)	25.73** (2.05)	26.57** (2.37)	19.36** (3.82)
NCS.2	15.15** (1.03)	14.26** (1.87)	14.39** (1.44)	14.00** (1.24)	16.61** (1.32)	15.74** (1.71)
NCL.2	-4.67 (2.45)	14.88** (5.76)	-3.78 (4.60)	-7.89** (2.88)	-6.33* (3.29)	-11.22** (2.86)
Berne-fr	-21.67** (2.96)	-22.32** (5.09)	-28.93** (4.40)	-19.62** (3.72)	-18.04** (4.46)	-9.57 (6.40)
Genève	-46.89** (2.56)	-40.84** (4.88)	-51.43** (3.30)	-50.34** (3.02)	-45.46** (3.30)	-33.64** (4.21)
Jura	-22.28** (2.97)	-18.28** (5.88)	-21.83** (4.47)	-22.06** (3.22)	-22.58** (4.14)	-18.93** (3.99)
Neuchâtel	-31.11** (2.43)	-33.54** (4.70)	-32.52** (3.77)	-31.33** (2.85)	-27.98** (3.33)	-16.43** (4.93)
Vaud	-23.39** (2.43)	-26.87** (3.52)	-28.20** (3.48)	-23.27** (2.89)	-21.23** (2.98)	-12.24* (4.84)
Valais-fr	-15.29** (2.40)	-5.80 (4.31)	-15.68** (3.22)	-14.71** (2.71)	-16.84** (2.83)	-9.30* (4.09)
NSE.2	18.32** (1.54)	17.93** (2.96)	18.73** (2.11)	18.96** (1.84)	16.19** (2.09)	18.88** (2.39)
NSE.3	0.09 (1.41)	0.49 (2.79)	-2.64 (1.97)	-1.16 (1.71)	1.27 (1.84)	3.80 (2.34)
NSE.4	5.41** (1.36)	5.93* (2.84)	5.53** (1.97)	5.42** (1.68)	7.35** (1.70)	3.76 (2.42)
immig.2	-36.38** (1.98)	-34.36** (3.46)	-34.14** (2.65)	-37.59** (2.48)	-37.62** (2.69)	-35.69** (3.13)
langue.2	-16.64** (2.34)	-14.59** (4.10)	-18.07** (3.31)	-17.50** (2.90)	-15.98** (3.10)	-13.07** (3.89)
$R^2, R^1$	0.331	0.169	0.195	0.198	0.182	0.162
$C, C^1$	0.961	0.995	0.981	0.971	0.973	1.003

Les erreurs standards sont données entre parenthèses sous les coefficients.

Un coefficient statistiquement significatif au seuil de 5% ou de 1% sous un test bilatéral est signalé avec (\*) ou (\*\*), respectivement.

TABLE 5 – Coefficients de régressions - Sciences

Variables	MCO	Quantiles				
		5%	25%	50%	75%	95%
constante	826.07** (19.35)	768.42** (39.84)	846.53** (26.42)	845.89** (23.11)	762.01** (21.36)	824.02** (28.56)
sexe.2	-16.48** (1.29)	-21.18** (2.64)	-17.62** (1.69)	-15.65** (1.45)	-15.48** (1.45)	-13.17** (1.93)
âge	-18.54** (1.22)	-21.50** (2.49)	-22.12** (1.67)	-19.41** (1.47)	-12.01** (1.34)	-13.08** (1.80)
NSP.2	46.69** (1.80)	48.16** (3.98)	52.59** (2.30)	48.66** (1.99)	44.27** (2.15)	34.29** (3.16)
NSP.3	-1.88 (1.62)	-5.36 (3.54)	-2.20 (2.06)	-3.48 (1.82)	-0.82 (1.89)	-1.75 (2.75)
NSP.4	9.75** (1.97)	8.19** (3.17)	10.64** (2.25)	11.00** (2.21)	8.42** (1.76)	4.68 (2.52)
NSP.5	22.89** (1.72)	23.41** (2.91)	24.49** (2.01)	24.97** (1.98)	22.27 (1.54)	16.98 (2.08)
NCS.2	9.64** (0.95)	10.53** (1.93)	10.07** (1.25)	9.95** (1.07)	9.16** (1.05)	8.41** (1.43)
NCL.2	4.03 (2.27)	21.71** (8.73)	6.52 (4.00)	-1.91 (2.13)	1.81 (3.26)	1.13 (2.27)
Berne-fr	-25.18** (2.74)	-41.67** (7.21)	-32.01** (3.93)	-22.03** (2.78)	-14.75** (3.27)	-7.93 (5.58)
Genève	-41.74** (2.37)	-44.19** (5.03)	-48.94** (3.24)	-43.24** (2.50)	-32.76** (2.54)	-30.57** (3.21)
Jura	-29.16** (2.75)	-42.60** (5.36)	-33.05** (3.44)	-26.80** (2.83)	-27.93** (2.54)	-21.04** (4.00)
Neuchâtel	-33.96** (2.25)	-51.67** (4.54)	-39.38** (3.02)	-32.77** (2.54)	-25.74** (2.57)	-24.00** (3.41)
Vaud	-16.98** (2.25)	-24.87** (5.20)	-21.23** (3.29)	-14.86** (2.74)	-11.28** (2.50)	-4.67 (3.49)
Valais-fr	-15.71** (2.22)	-16.00** (4.31)	-16.85** (2.68)	-16.02** (2.21)	-13.62** (2.38)	-9.52** (2.73)
NSE.2	15.30** (1.43)	15.17** (3.01)	15.12** (1.97)	14.36** (1.64)	15.37** (1.69)	15.94** (1.93)
NSE.3	-0.97 (1.30)	1.08 (2.73)	-1.22 (1.74)	-1.36 (1.47)	-0.72 (1.48)	-1.28 (1.99)
NSE.4	4.57** (1.26)	5.40* (2.61)	5.20** (1.59)	5.35** (1.39)	3.55** (1.35)	2.85 (2.14)
immig.2	-26.94** (1.83)	-28.45** (4.13)	-27.49** (2.28)	-28.41** (2.21)	-28.41** (2.08)	-21.55** (2.74)
langue.2	-13.69** (2.17)	-17.27** (4.96)	-13.88** (3.08)	-16.22** (2.64)	-12.72** (2.55)	-14.71** (4.46)
$R^2, R^1$	0.297	0.160	0.182	0.175	0.140	0.125
$C, C^1$	0.967	0.997	0.990	0.979	0.977	0.981

Les erreurs standards sont données entre parenthèses sous les coefficients.

Un coefficient statistiquement significatif au seuil de 5% ou de 1% sous un test bilatéral est signalé avec (\*) ou (\*\*), respectivement.

TABLE 6 – Coefficients de régressions - Lecture

Cantons	Matières	RQ - Rejet $H_0$				MCO	
		Quantile 5%		Quantile 95%		Rejet $H_0$	
		Seuil 1%	Seuil 5%	Seuil 1%	Seuil 5%	Seuil 1%	Seuil 5%
Berne-fr vs Fribourg-fr	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Berne-fr vs Genève	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Berne-fr vs Jura	Math	x	x	x	x		x
	Scie	x	x	x	x		x
	Lect	x	x	x	x	x	x
Berne-fr vs Neuchâtel	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Berne-fr vs Vaud	Math	x	x	x	x		x
	Scie	x	x	x	x		x
	Lect	x	x	x	x	x	x
Berne-fr vs Valais-fr	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Fribourg-fr vs Genève	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Fribourg-fr vs Jura	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Fribourg-fr vs Neuchâtel	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Fribourg-fr vs Vaud	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Fribourg-fr vs Valais-fr	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Genève vs Jura	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Genève vs Neuchâtel	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Genève vs Vaud	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x
Genève vs Valais-fr	Math	x	x	x	x	x	x
	Scie	x	x	x	x	x	x
	Lect	x	x	x	x	x	x

... suite page suivante...

... suite de la page précédente...

Jura vs Neuchâtel	Math	×	×	×	×	×	×
	Scie	×	×	×	×		×
	Lect	×	×	×	×		×
Jura vs Vaud	Math	×	×	×	×	×	×
	Scie	×	×				
	Lect	×	×	×	×	×	×
Jura vs Valais-fr	Math	×	×	×	×	×	×
	Scie	×	×	×	×	×	×
	Lect	×	×	×	×	×	×
Neuchâtel vs Vaud	Math	×	×	×	×	×	×
	Scie		×	×	×		×
	Lect	×	×	×	×	×	×
Neuchâtel vs Valais-fr	Math	×	×	×	×	×	×
	Scie	×	×	×	×	×	×
	Lect	×	×	×	×	×	×
Vaud vs Valais-fr	Math	×	×	×	×	×	×
	Scie	×	×	×	×	×	×
	Lect	×	×	×	×	×	×

TABLE 7: Comparaisons cantonales par paires (tests de stabilité des coefficients) ; × indique le rejet de l'hypothèse  $H_0$  au seuil spécifié