

# Fonctions de production éducationnelle: le cas de la Suisse

MURIEL MEUNIER\*

JEL-Classification: I20, J16, C2

Keywords: fonction de production éducationnelle, transformation Box-Cox, genre, PISA

## 1. Introduction

Il aura fallu attendre que les économistes investissent le secteur de l'éducation pour qu'apparaissent les fonctions de production éducationnelle. En soulevant l'idée que les écoles peuvent être considérées comme des firmes spécialisées dans la production d'éducation, la relation estimée est devenue connue en tant que «educational production function» (EPF) plutôt que de simples analyses d'input-output (SCHULTZ, 1963; HANUSHEK, 1979). S'il existe un relatif consensus concernant le lien entre les caractéristiques de l'élève ou de sa famille et ses résultats scolaires (LEIBOWITZ, 1974; HAVEMAN et WOLFE, 1995), les avis sont partagés lorsqu'il s'agit de la relation entre les inputs de l'école et la performance scolaire (HANUSHEK, 1986). On doit au «rapport Coleman» (1966) d'ouvrir les hostilités en concluant que les différences entre les écoles ont peu à voir avec les différences de performances entre les élèves. Depuis, les EPF ont été massivement estimées pour tenter (notamment) d'infirmer ou confirmer ce résultat. Le débat est toujours d'actualité (PSACHAROULOULOS, 2000; HANUSHEK, 2003).

La revue de littérature présentée par HANUSHEK (1986) sur la base de 147 études abonde dans le sens du rapport Coleman. L'effet de la taille de la classe

\* Département d'économie politique, Université de Genève, Boulevard du Pont d'Arve 40, 1211 Genève 4. Tél: 022 379 8918, Fax: 022 379 8293, Email: Muriel.Meunier@ecopo.unige.ch. Je remercie tout particulièrement Andrea Baranzini, Giovanni Ferro-Luzzi, Yves Flückiger, Tobias Müller, José Ramirez et un rapporteur anonyme pour leurs commentaires et suggestions. Je tiens également à remercier Jean-Marc Falter, les participants du Meeting Annuel 2005 de la Swiss Society of Economics and Statistics et un rapporteur anonyme pour leurs remarques sur une version antérieure de ce papier. Je suis reconnaissante envers Mme Huguette Mc Cluskey (responsable du projet PISA.ch) et Mme Claudia Zahner (Section SCHUL) de l'Office fédéral de la Statistique pour leurs précieuses explications concernant les données PISA 2000. Je reste seule responsable des erreurs et omissions qui pourraient encore subsister.

(ratio professeur/élèves) sur les performances n'est statistiquement significatif que dans 20% des études; celui du niveau de formation des enseignants dans 10% des études et celui de l'expérience des enseignants dans 37% des études. En revanche, GOLDHABER et BREWER (1997) estiment qu'il y a de bonnes raisons de croire que beaucoup d'études sur les EPF, particulièrement celles réalisées dans les années 1970, souffrent de déficiences majeures au niveau des données et de la méthodologie. D'une part, les investigations concernant l'effet potentiel des écoles n'auraient mené à rien à cause du degré d'agrégation des données (SUMMERS et WOLFE, 1977). D'autre part, une des raisons pouvant expliquer pourquoi les recherches sur le sujet aboutissent à des conclusions aussi diverses concerne la forme fonctionnelle utilisée (FIGLIO, 1999). Les études plus récentes ont plutôt tendance à conclure que les professeurs, mais également les écoles, comptent beaucoup pour les acquis des élèves (RIVKIN, HANUSHEK et KAIN, 2005).

La première contribution de ce papier est de participer au débat concernant l'impact des caractéristiques de l'école sur les performances des élèves. Ces estimations s'inscrivent dans le courant des recherches effectuées en Suisse par WOLTER et CORADI VELLACOTT (2002) qui se sont intéressés à l'effet de l'ordre dans la fratrie sur les résultats en lecture ou FISCHER (2005) qui a étudié l'impact de la démocratie directe sur l'éducation publique. Les résultats de l'estimation des EPF à partir des données de PISA 2000 montrent que le niveau de formation de l'enseignant influence très significativement les scores des élèves. Si celui-ci correspond au premier cycle de l'enseignement supérieur en pédagogie, l'élève obtiendra un score supérieur en lecture (+11%), en mathématiques (+12.7%) et en sciences (+12.9%). Les estimations par genre permettent d'affiner ce résultat puisqu'il apparaît que la formation de l'enseignant influence davantage le score des garçons que celui des filles.

La seconde contribution de ce papier est de proposer une forme fonctionnelle plus flexible et de tester si l'utilisation majoritairement répandue de la forme fonctionnelle linéaire est la plus appropriée. Le rejet des formes fonctionnelles linéaire, log-linéaire ou inverse tend à rejeter cette hypothèse et à encourager de plus amples recherches dans ce sens (MONTMARQUETTE et MAHSEREDJAN, 1985; FIGLIO, 1999).

Le reste de l'article est organisé de la manière suivante. La prochaine section présente le modèle Box-Cox ainsi que la méthode delta permettant d'interpréter les coefficients obtenus en valeur absolue ou relative. La section 3 présente la base de données et le choix des variables et la section 4, les résultats de l'analyse empirique. Des estimations par genre sont également reproduites à la section 5. La dernière section conclut avec les implications pour les choix de politiques publiques et les suites possibles de recherche.

## 2. Modélisation

La fonction de production éducationnelle ci-dessous permet d'estimer un modèle simple d'acquisition (TODD et WOLPIN, 2003). De manière très générale, elle peut être modélisée de la façon suivante:

$$A_i = f(X_i, F_i, S_i) \quad (1)$$

Dans cette équation,  $A$  représente une mesure de l'output du système scolaire (par exemple, le score obtenu à un test) et la lettre  $i$  est un indice pour les individus. Le vecteur  $X_i$  contient les variables relatives aux caractéristiques de l'élève, le vecteur  $F_i$  celles relatives aux caractéristiques de sa famille et le vecteur  $S_i$  celles relatives aux caractéristiques de son école.

Nous partons de l'hypothèse de base que la fonction est linéaire et que  $X_i$ ,  $F_i$  et  $S_i$  sont exogènes. Si l'on suppose que dans le modèle

$$A_i = \beta_1 + \beta_2 X_i + \beta_3 F_i + \beta_4 S_i + \varepsilon_i,$$

les  $A_i$  sont normalement et indépendamment distribués avec l'espérance

$$= \beta_1 + \beta_2 X_i + \beta_3 F_i + \beta_4 S_i$$

et la variance  $= \sigma^2$  alors la fonction de densité de probabilité jointe des  $A_i$  (avec  $i = 1$  à  $n$ ), étant donné la moyenne et la variance ci-dessus, peut être écrite de la manière suivante:

$$f(A_1, A_2, \dots, A_n | \beta_1 + \beta_2 X_i + \beta_3 F_i + \beta_4 S_i, \sigma^2) \quad (2)$$

Si les hypothèses de base sont correctes, l'équation est correctement spécifiée et l'estimation de cette fonction de production par maximum de vraisemblance donnera des estimateurs convergents des coefficients des caractéristiques de l'élève, de sa famille et de son école. Or, rien ne permet a priori de savoir si ces hypothèses sont respectées. Box et Cox (1964) proposent alors de «travailler avec une famille paramétrique de transformations de  $A$  à  $A^{(\theta)}$ , le paramètre  $\theta$  (...) définissant une transformation particulière». Ce que l'on retrouve désormais dans la littérature sous le nom de transformation Box-Cox propose donc d'estimer le modèle suivant:

$$A_i^{(\theta)} = \beta_1 + \beta_2 X_i + \beta_3 F_i + \beta_4 S_i + \varepsilon_i \quad (3)$$

avec

$$\begin{aligned} A_i^{(\theta)} &= \frac{A^\theta - 1}{\theta} \text{ si } \theta \neq 0 \\ A_i^{(\theta)} &= \ln(A) \text{ si } \theta = 0 \end{aligned} \quad (4)$$

La transformation tient pour  $A_i > 0$ . Etant donné que la majorité des variables explicatives sont des variables muettes, nous conservons la forme linéaire à droite du signe d'égalisation.

Supposons désormais que pour  $\theta$  non connu, les observations transformées satisfont toutes les hypothèses normales théoriques. La densité de probabilité pour les observations non transformées est obtenue en multipliant la densité normale par le Jacobien de transformation (Box et Cox, p. 215). On obtient alors la fonction de vraisemblance du modèle Box-Cox:

$$\ln L_c = -\frac{n}{2} [\ln(\tilde{\sigma}^2) + \ln(2\pi) + 1] + (\theta - 1) \sum_{i=1}^n \ln(A_i) \quad (5)$$

où

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \quad (6)$$

Afin d'être en mesure d'interpréter les coefficients obtenus, nous utilisons la procédure communément appelée la méthode delta issue du théorème de Taylor. Cette transformation nous permet d'obtenir deux effets qualitatifs distincts:

$$\text{la différence absolue: } A_1 - A_0 \cong \beta_k A_0^{1-\theta} \quad (7)$$

$$\text{la différence relative: } \frac{A_1 - A_0}{A_0} \cong \beta_k A_0^{-\theta} \quad (8)$$

où  $A_0$  est le score moyen dans l'une des trois disciplines et  $\beta_k$  le coefficient estimé de la  $k^{\text{ième}}$  variable<sup>1</sup>.

1 Par exemple, l'effet du genre (1: une fille, 0: un garçon) sur le score en lecture sera déterminé de la manière suivante: étant donné que le score moyen de l'échantillon total est  $A_0 = 510.823$ , le coefficient estimé est  $\beta = 458.733$  et le thêta estimé est  $\theta = 1.517731$ , la différence absolue correspond à:  $A_1 - A_0 = 458.733 \times 510.823^{1-1.517731} = 18.17$  (points).

### 3. Données et variables

#### 3.1 L'enquête PISA 2000

Les données utilisées proviennent de l'échantillon national de l'enquête PISA (Programme International pour le Suivi des Acquis des élèves) réalisée en 2000 par l'OCDE (OCDE, 2002). L'objectif de cette enquête était de tester les compétences des élèves dans le but de comparer les acquis des jeunes. Ceux-ci ont été appréhendés dans trois domaines: en lecture, en mathématiques et en sciences (avec en 2000, une priorité pour la lecture). La population de référence a été définie en fonction de l'année scolaire. Il s'agit de la 9<sup>e</sup> année du système scolaire suisse qui correspond à la dernière année de la scolarité obligatoire<sup>2</sup>.

Concernant l'échantillonnage, la Direction nationale du projet (i.e. la Confédération et les cantons) a tiré des échantillons représentatifs des trois grandes régions linguistiques<sup>3</sup>. Dans une première étape, les écoles ont été tirées au sort sur la base des données scolaires de 1998/99 de l'Office fédéral de la statistique (OFS). Dans une seconde étape, les élèves ont été sélectionnés de manière aléatoire parmi les écoles retenues. Les élèves n'ont donc pas tous participé à PISA et seuls certains élèves de certaines écoles ont répondu aux tests. Chaque élève et chaque école possèdent un numéro d'identification mais il n'est malheureusement pas possible de savoir si les élèves d'une même école sont ou non dans la même classe. De plus, on ne connaît pas les caractéristiques de la classe des élèves sélectionnés. Les considérations d'effets de paire au niveau de la classe ne peuvent donc pas être examinées.

Concernant le déroulement de l'enquête, chaque élève a rempli un *questionnaire élève* portant sur des données contextuelles en supplément du cahier de tests. La direction de chaque école a également rempli un *questionnaire établissement* contenant des questions sur les caractéristiques de l'école et de son personnel.

2 La scolarité obligatoire dure 9 ans et comprend les degrés primaire et secondaire I. Dans la majorité des cantons, le primaire dure 6 ans (de 6/7 ans à 11/12 ans) et le secondaire I, 3 ans (de 12/13 ans à 14/15 ans). Pour plus de détails sur le système éducatif suisse, voir le site internet de la CDIP ([http://www.ides.ch/umfrage2003/mainUmfrage\\_F.html](http://www.ides.ch/umfrage2003/mainUmfrage_F.html)).

3 L'échantillon national d'origine compte 7997 élèves de 9<sup>e</sup> provenant de 243 établissements dont 5 236 suisses alémaniques, 1 815 suisses romands et 946 suisses italiens.

### 3.2 La sélection des variables

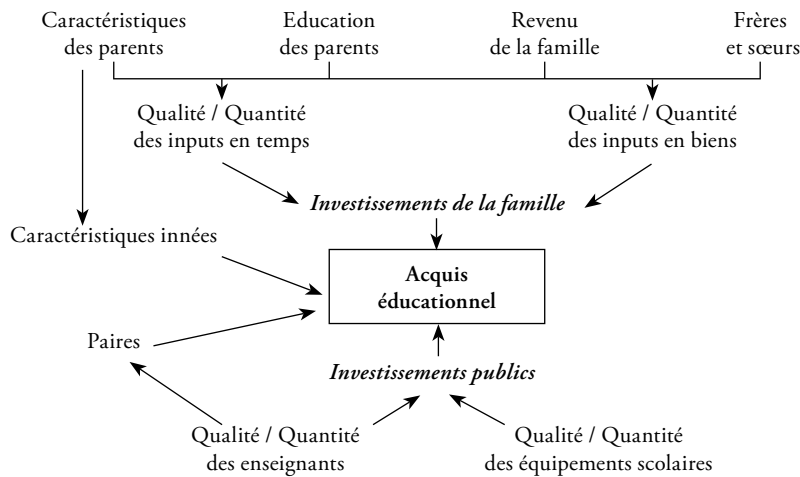
Le cadre conceptuel général décrit les acquis d'un élève donné à un instant  $t$  comme une fonction cumulative des inputs de la famille, des pairs et autres élèves, de l'école et des enseignants (HANUSHEK, 1986). Tel que le présente la figure 1, les inputs éducatifs qui influencent l'apprentissage proviennent essentiellement de trois types d'agents: la société (les investissements publics), la famille (les investissements de la famille) et les enfants (à travers les choix qu'ils effectuent étant donné les investissements et opportunités qui leurs sont disponibles).

Un des objectifs de la recherche empirique est de comprendre la technologie qui combine les inputs éducatifs afin de créer un output cognitif. Le choix des variables explicatives est alors issu de l'analogie qui est faite entre le processus d'acquisition du savoir d'un individu et le processus de production d'une firme. Les variables sélectionnées doivent représenter au mieux les inputs éducatifs qui ont un impact dans le processus éducationnel. Cette sélection est cependant contrainte par la disponibilité des informations dans la base de données.

En effet, dans la mesure où le processus d'acquisition du savoir est un processus cumulatif, l'idéal serait d'avoir accès à l'ensemble du parcours scolaire de l'élève ainsi qu'à toutes les expériences présentes et antérieures de la famille et de l'école. Étant donné que de telles données sont rares voire inexistantes, l'utilisation de variables proxy est l'une des solutions permettant de palier le problème des données manquantes. Celles-ci ne sont pas considérées comme des inputs ayant un impact direct dans le processus de production mais sont insérées en faisant l'hypothèse que leur inclusion va diminuer le biais de variables omises parce qu'elles sont corrélées avec les inputs omis. Les variables d'origine ethnique et de revenu de la famille peuvent être considérées comme de telles variables proxy (TODD et WOLPIN, 2003).

Les caractéristiques de l'élève retenues incluent le genre car il s'agit d'une caractéristique sociale qui influence l'apprentissage (SUMMERS et WOLFE, 1977). Elles comprennent également la nationalité. La revue de littérature présentée par HAVEMAN et WOLFE (1995) indique que les études qui s'intéressent aux impacts des déterminants de la famille et du voisinage sur les diplômés à l'école secondaire et sur le nombre d'année d'études utilisent souvent l'origine ethnique comme une des caractéristiques des antécédents de l'élève. Par ailleurs, cette variable est intéressante dans le contexte particulier de la Suisse qui a connu plusieurs vagues d'immigration distinctes. Tenir compte de la nationalité permet de prendre en considération le degré d'intégration socio-économique potentiel des élèves lié au fait qu'il s'agit d'individus provenant d'une vague d'immigration plus ou moins

Figure 1: Les déterminants des acquis éducationnels  
(adaptée de HAVEMAN et WOLFE, 1995)



récente<sup>4</sup>. L'âge enfin, permet de capter les effets liés à l'expérience (les élèves nés dans la première moitié de l'année peuvent mieux réussir du fait qu'ils sont plus matures). Il permet également de tenir compte d'éventuelles stratégies éducationnelles de la part des parents et des enfants dans la mesure où le redoublement donne lieu à deux orientations possibles en Suisse. Dans un premier cas, les élèves qui ont échoué doivent recommencer le même programme. Dans le second, les élèves passent par un autre type d'enseignement et entrent souvent dans un programme plus ambitieux leur ouvrant de plus larges perspectives (OFS, 1999).

Les choix effectués par les parents vont également agir sur le capital humain que les enfants accumuleront au cours de leur vie, du moins dans les premières années. En agissant comme une unité de production qui emploie des inputs afin de générer de l'utilité à ses membres, la famille va influencer les acquis (BECKER et TOMES, 1986). Les décisions prises au sein du ménage vont affecter la structure familiale ou le nombre d'enfants qui vont à leur tour modifier l'environnement affectif ou financier de l'élève (suite à un divorce par exemple).

4 Les systèmes scolaires des 41 pays qui ont participé aux tests PISA 2003 se différencient considérablement du point de vue du devenir qu'ils assurent aux enfants immigrés (OCDE, 2006). Dans de nombreux pays, les enfants issus de familles immigrées sont désavantagés dès le départ.

Les dotations génétiques que les parents transmettent à leurs enfants constituent également une source d'inégalité dans le processus d'acquisition (LEIBOWITZ, 1974). Tous les enfants ne naissent pas avec les mêmes capacités innées. Étant donné que celles-ci sont inobservables, il est commun d'utiliser le niveau d'éducation de la mère et du père comme mesure des capacités intellectuelles des parents. De plus, les caractéristiques génétiques des parents et leur éducation déterminent le temps qu'ils vont investir pour l'éducation de leurs enfants d'un point de vue quantitatif et qualitatif.

Le statut des parents sur le marché du travail permet également d'approximer la quantité des inputs allouée aux enfants en temps et en biens. Une meilleure mesure du niveau économique des ressources dévouées aux enfants par les parents aurait été le niveau de revenu de la famille dans laquelle les enfants grandissent<sup>5</sup>. Bien que l'importance de cette variable ait été de nombreuses fois mise en évidence dans la littérature (HANUSHEK, 1986), elle n'était malheureusement pas disponible dans PISA 2000<sup>6</sup>. Aussi, à défaut de disposer du revenu, le statut des parents sur le marché du travail peut être considéré comme une variable proxy des ressources financières de la famille bien qu'elle ne capture pas entièrement l'effet du revenu (le type d'emploi aurait été une caractéristique plus déterminante pour expliquer les différences de revenus). Précisons toutefois que le revenu n'est pas une mesure parfaite des ressources de la famille. D'une part, parce qu'il est généralement reporté pour une seule année et représente donc une mauvaise mesure du revenu permanent. D'autre part, parce que le revenu en tant que tel renseigne peu sur la manière dont il est alloué aux enfants et échoue à capturer d'autres ressources économiques qui leur sont dévouées telle que l'allocation en temps. L'investissement des parents dans l'éducation de leurs enfants est ici approximé par la variable qui mesure l'aide de la mère ou du père en ce qui concerne les devoirs.

- 5 L'hypothèse sous-jacente à l'inclusion de cette variable est que les familles avec des revenus plus élevés dépenseront davantage pour les biens qui ont un impact sur le processus d'acquisition. De plus, le revenu peut permettre de compenser pour les données manquantes des inputs de la famille (TODD et WOLPIN, 2003).
- 6 L'indice ISEI (International Socio-Economic Index of Occupational Status) aurait également pu servir de variable proxy du revenu des parents. Il s'agit d'un indice regroupant les attributs des professions qui permet de convertir la formation des parents en revenus (pour plus de détails, voir le rapport technique de PISA 2000 (OCDE, 2002)). Les résultats restent stables lorsque la variable ISEI est utilisée à la place des variables d'éducation des parents et de statut professionnel (afin d'éviter tout problème de colinéarité). Le coefficient de la variable ISEI est alors positif et significatif.

Les caractéristiques de l'école retenues sont la taille de l'école, une variable proxy de la taille de la classe (le ratio entre la taille de l'école et le nombre de professeurs) et le niveau de qualification des enseignants (approximé par la proportion du personnel enseignant de l'établissement qui a participé à un programme de formation continue au cours des trois derniers mois, la proportion de professeurs avec un niveau ISCED 5 en pédagogie<sup>7</sup> et la proportion de professeurs complètement certifiés). Il s'agit d'inputs éducatifs qui affectent le processus d'acquisition (RIVKIN et al., 2005). La localisation de l'école (la région ainsi que le fait que l'école soit située dans un environnement plus ou moins urbanisé) est notamment utilisée afin de tenir compte des différences d'investissements publics qui peuvent affecter les inputs scolaires.

#### 4. Résultats empiriques

Les fonctions de production éducationnelle ont été estimées par maximum de vraisemblance. Les coefficients ont ensuite été transformés à l'aide de la méthode delta pour obtenir des variations relatives et absolues. Les résultats sont reportés séparément: le tableau 1 présente les coefficients relatifs aux caractéristiques de l'élève, le tableau 2 ceux relatifs aux caractéristiques de la famille et le tableau 3 enfin, ceux relatifs aux caractéristiques de l'école. Les variables dépendantes sont le score en lecture (colonne 1), mathématiques (colonne 2) et sciences (colonne 3)<sup>8</sup>. L'échantillon compte 4 771 élèves pour la lecture, 2 637 pour les mathématiques et 2 669 pour les sciences (cf. tableau A1 pour les statistiques descriptives et la définition des variables). Les scores moyens de notre échantillon sont de 511 en lecture, 546 en mathématiques et 509 en sciences<sup>9</sup>.

Des estimations séparées ont préalablement été réalisées afin de vérifier la stabilité des coefficients. Le fait d'introduire les caractéristiques de l'individu, puis de

7 L'ISCED 5 correspond au premier cycle de l'enseignement supérieur.

8 Parmi tous les élèves de l'échantillon d'origine ayant participé au test (7 997), seuls 2 653 ont répondu à la fois au questionnaire en lecture et en sciences, 2 647 ont répondu à la fois au questionnaire en lecture et en mathématiques, 1 804 ont répondu à la fois au questionnaire en lecture, mathématiques et sciences et 893 ont seulement répondu au questionnaire en lecture. C'est ce qui explique la taille plus modeste des échantillons pour ces estimations.

9 En comparaison internationale, la Suisse est 17<sup>e</sup> en lecture (avec un score moyen de 494, la Finlande est 1<sup>ère</sup> avec 547), 7<sup>e</sup> en mathématiques (avec un score moyen de 529, le Japon est 1<sup>er</sup> avec 557) et 18<sup>e</sup> en sciences (avec un score moyen de 495, la Corée du Sud est 1<sup>ère</sup> avec 552). Dans les trois matières, le score moyen des pays de l'OCDE est de 500 (OFS, 2002).

rajouter celles de la famille et enfin celles de l'établissement entraînant une attrition des données relativement conséquente. Il a donc été nécessaire de contrôler que cette attrition n'impliquait pas de variation de la répartition des scores par région et par taille d'école. Nous avons ainsi pu nous prémunir du risque que l'attrition des données ne soit pas aléatoire entre les régions ou entre les écoles (par exemple, une corrélation entre la taille de l'école et le taux de réponses manquantes). Les résultats sont stables d'une spécification à l'autre.

#### *4.1 Les caractéristiques de l'élève*

Les résultats du tableau 1 indiquent clairement que le genre et la nationalité sont des variables explicatives importantes<sup>10</sup>. Par rapport à un garçon, une fille obtient un score supérieur en lecture (+3.6%) mais inférieur en mathématiques (-5.1%) et en sciences (-3.1%). Le sexe de l'élève a donc bien un impact sur l'apprentissage (SUMMERS et WOLFE, 1977) et explique peut-être les différences d'orientation selon le genre que l'on observe en Suisse. Précisons cependant que pour pouvoir se prononcer sur la question il aurait été utile de contrôler pour l'accès des femmes aux études supérieures et au marché du travail, information dont on ne dispose malheureusement pas. En effet, BAKER et PERKINS JONES (1993) trouvent que dans les pays qui offrent des opportunités équivalentes aux hommes et aux femmes, les différences de performances des élèves en mathématiques dues au genre sont plus petites.

Depuis le début des années 1980, le nombre des élèves étrangers ou de langue étrangère a nettement augmenté en Suisse<sup>11</sup> (OFS, 1999). Nos résultats montrent que par rapport aux élèves nés en Suisse, ceux nés en Italie, Espagne, Portugal ou autres obtiennent un score inférieur (-2.5% en lecture, -4.3% en mathématiques et -4% en sciences), et que ceux nés en Ex-Yougoslavie, Albanie, Kosovo ou Turquie obtiennent un score encore plus faible (-7.6% en lecture, -8% en mathématiques et -8.5% en sciences). Ces résultats méritent d'être soulignés et suggèrent les analyses suivantes. D'une part, on peut envisager que le fort impact de la nationalité sur les scores des élèves reflète l'intégration sociale, culturelle et économique des immigrants (la nationalité capte également sans doute une partie de l'effet de la variable revenu omise). D'autre part, les effets négatifs de

10 SUMMERS et WOLFE (1977) contrôlent également pour le quotient intellectuel de l'élève, une variable non disponible dans PISA.

11 La part des classes très hétérogènes (au moins 30% d'élèves de nationalité étrangère et/ou parlant une autre langue que celle enseignée à l'école) est passée de 20% en 1980 à 34% en 1998.

**Tableau 1: Estimation Box-Cox des fonctions de production éducationnelle: les caractéristiques de l'élève**

Variables	(1) Lecture		(2) Mathématiques		(3) Sciences	
	relative (%)	absolue	relative (%)	absolue	relative (%)	absolue
genre	3.56***	18.17	-5.07***	-27.69	-3.07***	-15.62
national1	réf.	réf.	réf.	réf.	réf.	réf.
national2	1.07	5.47	-2.47	-13.47	0.01	0.07
national3	-2.47**	-12.59	-4.34***	-23.66	-4.00***	-20.36
national4	-7.65***	-39.07	-8.04***	-43.90	-8.53***	-43.47
âge	-0.31***	-1.60	-0.30***	-1.65	-0.31***	-1.55

Source: PISA 2000. Notes: \*  $\chi^2 < 10\%$ , \*\*  $\chi^2 < 5\%$ , \*\*\*  $\chi^2 < 1\%$ . Observations: 4771 en lecture, 2637 en mathématiques et 2669 en sciences.

l'immigration semblent s'estomper avec le temps: le premier groupe appartenant à la première vague traditionnelle d'immigration en Suisse et le second, à une vague d'immigration beaucoup plus récente.

L'âge enfin, a peu d'impact sur les scores. La moyenne de cette variable est étonnamment élevée alors que les élèves sont censés avoir tous entre 14 et 15 ans (entre 168 et 180 mois) en dernière année de la scolarité obligatoire. On peut alors soupçonner un nombre conséquent d'élèves ayant redoublé (c'est-à-dire avec un âge supérieur à la moyenne). Le fait que cette variable soit statistiquement significative mais ait peu d'impact peut être le résultat d'un effet ambigu: d'un côté, les élèves qui profitent de leur redoublement pour améliorer leurs résultats (effet positif sur le score); de l'autre, ceux qui se contentent de redoubler avec leurs lacunes (effet négatif sur le score).

#### 4.2 Les caractéristiques de la famille

En Suisse, plusieurs langues nationales scindent naturellement le système éducatif en trois grandes régions linguistiques. Il s'agit de l'allemand (parlé par environ 64% de la population), du français (20%) et de l'italien (7%). Les individus ne parlant pas une des langues nationales représentent un peu moins de 10% (OFS, 2004). Les résultats du tableau 2 indiquent que si la langue parlée à la maison correspond à celle du test, l'élève obtiendra un score plus élevé (+5.7% en lecture, +4.7% en mathématiques et +5.9% en sciences). L'amélioration des

performances scolaires de la Suisse passe donc sans doute par l'intégration linguistique de ses immigrants.

Au sujet de la structure familiale, les élèves qui déclarent ne pas vivre dans une famille nucléaire obtiennent des résultats moins bons (en mathématiques par exemple:  $-2.6\%$  s'il s'agit d'une famille célibataire,  $-3.3\%$  pour une famille mixte et  $-5.5\%$  pour la catégorie autre). Comme dans beaucoup d'autres pays, la composition de la population selon l'état civil a beaucoup évolué en Suisse. En 30 ans, la proportion des personnes divorcées a triplé, ce qui a profondément modifié la structure familiale<sup>12</sup>. Les modifications de l'environnement affectif des élèves ne sont donc vraisemblablement pas sans effets sur leurs résultats scolaires. On peut considérer que la structure familiale capte également les différences de ressources entre les familles liées au fait que le niveau de revenu d'une famille monoparentale est a priori plus faible que celui d'une famille nucléaire ou recomposée.

LEIBOWITZ (1974) ou BECKER et TOMES (1976) présentent la famille comme une unité de production dont les choix auront des répercussions sur le développement des enfants. Le nombre de frères et sœurs (variable proxy des investissements de la famille) a pourtant un impact faible sur les scores. Etant donné que l'évolution de la fécondité depuis la fin du 19<sup>e</sup> siècle est caractérisée par une baisse à long terme du nombre moyen d'enfants par femme entraînant une moindre hétérogénéité entre les familles (OFS, 2003), le rang dans la fratrie aurait sans doute présenté davantage d'intérêt (WOLTER et CORADI VELLACOT, 2002).

La littérature sur la mobilité (scolaire ou sociale) intergénérationnelle a largement mis en évidence le fait que les enfants ne partent pas tous avec les mêmes atouts (HAVEMAN et WOLFE, 1995). En Suisse, si l'on rapporte le niveau d'éducation du fils ou de la fille à celui de son père ou de sa mère (en l'occurrence celui des deux parents qui possède le titre scolaire le plus élevé), il apparaît que l'héritage scolaire l'emporte sur la mobilité au sein de chaque niveau d'éducation<sup>13</sup> (OFS, 1997). Nos résultats confirment ceci puisque l'éducation secondaire des parents a un impact significatif sur le score<sup>14</sup>. Les élèves dont les parents ont un niveau

12 En 2002, la population résidante permanente se composait de 42% de célibataires, 46% de personnes mariées, 6% de divorcées et 6% de personnes veuves (OFS, 2003). Ces statistiques sont celles de la population résidente totale alors que notre échantillon comprend une tranche d'âge particulière, celle des parents en âge d'avoir des enfants en 9<sup>e</sup> année.

13 Les tables de mobilités ont été générées à partir des données de l'Enquête sur la famille en Suisse collectées en 1994-1995. Les hommes sont plus mobiles et aussi plus souvent en mobilité ascendante que les femmes.

14 Il est vraisemblable que les variables d'éducation des parents captent également une partie de l'effet de la variable revenu omise. En effet, on peut considérer que les parents ayant un niveau d'éducation élevé ont une position sur le marché du travail associée à un salaire élevé.

**Tableau 2: Estimation Box-Cox des fonctions de production éducationnelle:  
les caractéristiques de la famille**

Variables	(1) Lecture		(2) Mathématiques		(3) Sciences	
	relative (%)	absolue	relative (%)	absolue	relative (%)	absolue
langue	5.74***	29.30	4.72***	25.74	5.92***	30.16
célibataire	-1.82***	-9.31	-2.65***	-14.46	-3.38***	-17.22
nucléaire	réf.	réf.	réf.	réf.	réf.	réf.
mixte	-1.49*	-7.62	-3.28***	-17.89	-1.85	-9.40
autre	-4.74***	-24.23	-5.47***	-29.84	-4.79***	-24.40
nsib	-0.93***	-4.75	-0.23	-1.28	-0.58**	-2.95
med1	-2.29**	-11.72	-2.27*	-12.40	-3.34**	-17.01
med2	réf.	réf.	réf.	réf.	réf.	réf.
med3	3.40***	17.39	2.83***	15.46	4.11***	20.95
med4	6.35***	32.41	5.99***	32.68	8.01***	40.83
mteduc	-1.35**	-6.91	-0.36	-1.94	-2.24***	-11.39
fed1	-2.11**	-10.76	-1.68	-9.19	-0.83	-4.21
fed2	réf.	réf.	réf.	réf.	réf.	réf.
fed3	3.57***	18.26	2.27***	12.36	3.84***	19.58
fed4	4.78***	24.40	4.49***	24.50	4.51***	22.98
fteduc	1.23**	6.27	0.45	2.47	1.89***	9.63
mwork1	réf.	réf.	réf.	réf.	réf.	réf.
mwork2	1.65***	8.45	2.16***	11.78	1.50**	7.63
mwork3	-1.57	-8.01	-1.18	-6.42	-0.71	-3.59
mwork4	0.97*	4.96	1.42**	7.73	1.22	6.24
fwork1	réf.	réf.	réf.	réf.	réf.	réf.
fwork2	-1.76*	-8.98	-3.41***	-18.62	1.38	7.04
fwork3	1.04	5.32	-2.30	-12.54	-5.15**	-26.23
fwork4	-0.65	-3.30	-0.61	-3.32	1.55	7.89
devoirm	-1.78***	-9.10	-2.98***	-16.25	-1.31*	-6.70
devoirp	-1.21**	-6.20	-1.42**	-7.77	-2.72***	-13.83
devoirfs	-1.38***	-7.05	-0.82	-4.46	-1.49**	-7.61

Source: PISA 2000. Notes: \* $\chi^2 < 10\%$ , \*\* $\chi^2 < 5\%$ , \*\*\* $\chi^2 < 1\%$ . Observations: 4771 en lecture, 2637 en mathématiques et 2669 en sciences.

d'éducation supérieur à la scolarité obligatoire (catégorie de référence) obtiennent des scores supérieurs. D'une part, l'impact est plus important s'il s'agit de la mère et d'autre part, si celle-ci a terminé avec succès une école de niveau maturité (+6.3% en lecture, +6% en mathématiques et +8% en sciences).

Concernant l'éducation tertiaire des parents, il est intéressant de constater que l'effet sur le score (significatif en lecture et en sciences mais relativement faible) est négatif s'il s'agit de la mère et positif s'il s'agit du père. Ce résultat qui semble contre-intuitif suggère l'interprétation suivante: si les mères éduquées travaillent davantage et occupent des postes ne leur permettant pas de s'occuper suffisamment de leurs enfants, ce manque d'attention peut avoir des répercussions négatives sur les résultats scolaires. Un homme en revanche, occupera un poste associé à un salaire plus élevé ce qui peut avoir des répercussions positives sur l'enfant.

L'investissement des parents dans l'éducation de leurs enfants (approximé par la variable qui mesure l'aide de la mère ou du père en ce qui concerne les devoirs) a un impact négatif sur le score des élèves. Ce résultat révèle peut être le fait que les parents qui aident leurs enfants le font sans doute parce que ceux-ci rencontrent des difficultés à l'école.

#### 4.3 Les caractéristiques de l'école

Etant donné que l'instruction publique est du ressort des cantons, la Suisse possède autant de systèmes scolaires que de cantons (à savoir vingt-six). Si cette spécificité présente l'avantage de pouvoir travailler avec une expérience naturelle unique, elle rend le travail d'analyse et de compréhension parfois complexe<sup>15</sup>. Les résultats relatifs à la localisation de l'école montrent que les élèves qui habitent les régions nord-ouest, centrale et orientale réussissent mieux que les suisses romands<sup>16</sup> (voir tableau 3). Il est cependant difficile d'analyser ce résultat avec les informations contenues dans PISA. On peut néanmoins se demander si ces différences sont le résultat de modalités de fonctionnement différentes selon les

15 L'OFS compile et convertit les 26 fichiers de données cantonales afin d'obtenir une statistique scolaire suisse. Cette harmonisation est effectuée à l'aide de la Classification des statistiques suisses de l'éducation qui repose sur l'ISCED (International Standard Classification of Education) de l'UNESCO.

16 Ce résultat est particulièrement intéressant puisqu'il est en contradiction avec les résultats annoncés dans le rapport national de l'OFS/CDIP (2002) sur PISA 2000. A partir de statistiques descriptives, le rapport national concluait notamment que les suisses romands sont meilleurs en lecture que les suisses alémaniques et les suisses italiens (OFS/CDIP, p. 35).

systèmes scolaires ou pourraient provenir des sources publiques de financement qui sont très différentes selon les régions. Nous ne disposons malheureusement pas de cette information dans PISA, variable pourtant très importante (KRUEGER, 1998). La taille de la localité dans laquelle se situe l'école a en revanche peu d'impact sur le score<sup>17</sup>.

Des EPF par région<sup>18</sup> ont également été estimées (résultats non reproduits mais disponibles sur demande). Il apparaît que le pouvoir explicatif des variables varie très fortement d'une région à l'autre. Concernant la formation des enseignants par exemple, l'impact sur le score en lecture est significatif et positif dans les quatre régions mais dans des proportions très différentes. Il est regrettable que la taille de l'échantillon soit si faible car des estimations par canton auraient pu être effectuées afin de tenter d'exploiter l'expérience naturelle de la compétence éducative cantonale.

Les résultats indiquent que le niveau de qualification des professeurs (ISCED 5 en pédagogie) a un impact significatif positif très important sur le score des élèves (+11% en lecture, +12.7% en mathématiques et +12.9% en sciences)<sup>19</sup>. Alors que le rapport Coleman concluait que l'expérience des enseignants ne comptait que marginalement, il apparaît que ce n'est pas le cas avec les données PISA 2000. Nos résultats vont donc plutôt dans le sens de ceux de SUMMERS et WOLFE (1977) ou GOLDHABER et BREWER (1997).

Finalement, la taille des classes est significative mais n'a pas d'impact quantitatif important sur le score. Il est possible que ce résultat provienne du fait que cette variable ait été approximée par le ratio entre la taille de l'école et le nombre de professeurs. En effet, ARIAS, DOUGLAS et WALKER (2004) montrent que de petites tailles de classes ont un impact positif sur la performance des élèves et DUSTMANN, RAJAH et VAN SOEST (2003), que la taille de la classe a un effet significatif sur la décision de rester à l'école (et affecte donc significativement les revenus futurs). SUMMERS et WOLFE (1977) indiquent que la taille des classes a également un impact lorsque celle-ci est explorée sur des sous-échantillons dépendants du niveau d'aptitude. Il est donc surprenant que cette variable n'apparaisse

17 En 2000, 5 villes ont plus de 100 000 habitants: Zürich (350 000), Genève (177 000), Bâle (169 000), Berne (128 000) et Lausanne (121 000); et une dizaine entre 30 000 et 100 000.

18 Le regroupement en 4 grandes régions est basé sur celui proposé par la CDIP (la Conférence des Directeurs cantonaux de l'Instruction Publique). Il présente l'avantage de regrouper les cantons qui ont des systèmes scolaires relativement similaires.

19 L'utilisation de données agrégées au niveau de l'école dans une analyse réalisée au niveau des élèves peut entraîner un biais dans les estimations (MOULTON, 1990).

**Tableau 3: Estimation Box-Cox des fonctions de production éducationnelle: les caractéristiques de l'école**

Variables	(1) Lecture		(2) Mathématiques		(3) Sciences	
	relative (%)	absolue	relative (%)	absolue	relative (%)	absolue
romande	réf.	réf.	réf.	réf.	réf.	réf.
nordw	6.20***	31.68	4.80***	26.20	5.96***	30.38
centrale	5.41***	27.64	4.66***	25.45	4.19***	21.35
orientale	4.24***	21.66	3.75***	20.47	3.80***	19.36
village	-2.01***	-10.24	-1.09	-5.97	-1.15	-5.86
ptville	-0.77	-3.96	-1.32**	-7.19	-0.22	-1.12
ville	réf.	réf.	réf.	réf.	réf.	réf.
gdville	-1.34	-6.83	-1.93*	-10.51	-2.14*	-10.92
schlsize	0.01***	0.04	0.01***	0.04	0.01***	0.05
stratio	-0.33***	-1.67	-0.21***	-1.14	-0.27***	-1.37
sc15q01	-0.02***	-0.08	0.00	-0.03	-0.01	-0.05
propqual	10.98***	56.07	12.67***	69.17	12.84***	65.41
propcert	1.36*	6.96	-0.19	-1.06	1.81*	9.22

Source: PISA 2000. Notes: \* $\chi^2 < 10\%$ , \*\* $\chi^2 < 5\%$ , \*\*\* $\chi^2 < 1\%$ . Observations: 4771 en lecture, 2637 en mathématiques et 2669 en sciences.

pas explicitement dans PISA 2000<sup>20</sup>, surtout quand on connaît le débat virulent qui entoure cette problématique (KRUEGER, 2003).

#### 4.4 La forme fonctionnelle

Alors que les formes fonctionnelles ont souvent été le centre d'intérêt en économie appliquée dans des domaines divers tels que les fonctions de productions manufacturées, il est surprenant qu'il n'y ait pas d'intérêt similaire dans l'étude des fonctions de production en éducation (MONTMARQUETTE et al., 1985). Il n'y

20 L'effectif des classes est fixé en Suisse par les cantons qui définissent légalement les tailles minimale et maximale. La taille dépend de facteurs démographiques et de la politique menée en matière financière et dans le domaine de la formation (OFS, 1999). En 1997, le nombre moyen d'élèves par classe au degré secondaire I était de 20 élèves.

a pas de raisons théoriques d'imposer une forme fonctionnelle spécifique lors de l'estimation de fonctions de production éducationnelle. Tel que le mentionne FIGLIO (1999): «les spécifications traditionnelles des fonctions de production éducationnelle sont excessivement restrictives parce qu'elles échouent à saisir les effets potentiellement non linéaires des ressources de l'école». Ses résultats indiquent que la fonction de production traditionnelle tend à sous-estimer les effets des inputs scolaires sur les acquis des élèves.

Le tableau 4 présente le  $\theta$  estimé par le modèle Box-Cox et nous renseigne sur la forme fonctionnelle en lecture, mathématiques et sciences. Dans les trois matières, ce modèle rejette à la fois les spécifications linéaire ( $\theta = 1$ ), log linéaire ( $\theta = 0$ ) et inverse ( $\theta = -1$ ). Cette modélisation présente plusieurs avantages<sup>21</sup>. Tout d'abord, il n'est pas nécessaire d'imposer une forme fonctionnelle aux données *a priori*. Ce modèle propose de plus une transformation qui peut permettre (si le modèle linéaire est rejeté) d'améliorer l'ajustement du modèle aux données. Il est par ailleurs possible de tester des spécifications fonctionnelles alternatives (linéaire, log-linéaire et inverse). Enfin, la transformation de la variable expliquée peut être utilisée lorsqu'il y a des problèmes de normalité ou d'hétéroscédasticité<sup>22</sup>.

Pourtant, le rejet des formes fonctionnelles linéaire, log-linéaire et inverse n'implique pas nécessairement que l'on doive adopter la forme fonctionnelle proposée par le modèle Box-Cox. Ce rejet peut provenir du fait que le modèle est mal spécifié notamment à cause des variables omises (GODFREY et WICKENS, 1981). C'est pourquoi les fonctions de production éducationnelle ont également été estimées avec un modèle à effets fixes et un modèle à effets aléatoires (tableaux A2, A3 et A4 en annexe). L'intérêt de ces deux modèles est qu'ils permettent de tenir compte du fait que certains inputs de l'école sont manquants dans l'estimation car inobservables. Dans le premier cas, un effet spécifique à chaque école et qui ne varie pas entre les individus est inclut dans l'estimation (modèle à effets fixes). Dans le second, les inputs scolaires omis sont traités comme un terme d'erreur aléatoire spécifique à chaque école (modèle à effets aléatoires)<sup>23</sup>.

21 Une estimation par MCO présente également certains avantages, notamment en présence d'hétéroscédasticité. L'analyse graphique des résidus des estimations linéaires et les tests pour l'hétéroscédasticité (Breusch-Pagan/Cook-Weisberg) indiquent qu'il n'y a pas de problème d'hétéroscédasticité sur les différents échantillons utilisés.

22 Il n'existe pas, à notre connaissance, de recherches en économie de l'éducation qui utilisent une modélisation Box-Cox. En revanche, il s'agit d'une méthodologie utilisée dans d'autres domaines de recherche (citons par exemple BUCHINSKY (1995) ou MACHADO et MATA (2000)).

23 Pour plus de précisions sur la modélisation, voir GOLDHABER et BREWER (1997).

Tableau 4: Estimation Box-Cox de la forme fonctionnelle

	$\theta$ estimé		Test $H_0$		
			$\theta = -1$	$\theta = 0$	$\theta = 1$
(1) Lecture	1.517731	Rest. Log. Lik.	-27 792.22	-27 258.18	-27 017.21
		LR stat chi2	1 608.52	540.45	58.51
		P-Value	0.000	0.000	0.000
(2) Mathématiques	1.347687	Rest. Log. Lik.	-15 122.67	-14 885.69	-14 788.07
		LR stat chi2	682.06	208.10	12.85
		P-Value	0.000	0.000	0.000
(3) Sciences	1.266308	Rest. Log. Lik.	-15 561.42	-15 268.35	-15 158.04
		LR stat chi2	816.12	229.97	9.35
		P-Value	0.000	0.000	0.002

Source: PISA 2000. Notes: DF = 1,  $\alpha = 0.05$ ,  $\chi^2 = 3.84146$ : rejet de  $H_0$ . Observations: 4 771 en lecture, 2 637 en mathématiques et 2 669 en sciences.

Les résultats des estimations avec effets fixes par école indiquent que les coefficients des variables spécifiques à l'élève (le genre, la nationalité et l'âge) ainsi que les variables de langue parlée à la maison, structure familiale et taille de la fratrie restent stables (aussi bien au niveau des coefficients que du degré de significativité). Les variables d'éducation des parents et de statut sur le marché du travail sont en revanche sensibles à l'inclusion des effets fixes. Si l'on considère que les élèves d'une même école ont certaines caractéristiques communes (et inobservées), il est possible que l'ajout des effets fixes par école capture une partie des effets qui étaient auparavant (dans l'estimation MCO) captés par les variables d'éducation des parents et de statut sur le marché du travail<sup>24</sup>.

Concernant les résultats des estimations avec effets aléatoires, les conclusions sont sensiblement les mêmes à savoir que les coefficients des variables genre, nationalité, âge, langue, structure familiale et fratrie restent relativement stables. Précisons que la formation des enseignants (propqual) demeure une variable ayant un impact fort et significatif sur le score des élèves quelle que soit la modélisation utilisée.

24 Des résultats similaires sont obtenus lorsque l'on estime les EPF avec le modèle Box-Cox et des variables muettes pour les écoles. Les spécifications linéaire, log-linéaire et inverse sont rejetées.

Finalement, les résultats obtenus par MCO diffèrent légèrement de ceux obtenus avec le modèle Box-Cox. Ces différences peuvent provenir du fait que le modèle Box-Cox propose une transformation permettant (étant donné que le modèle linéaire est rejeté) d'améliorer l'ajustement du modèle aux données. Il serait intéressant de réaliser des estimations Box-Cox sur des sous-échantillons définis selon différents niveaux de performance (par exemple: élèves faibles, élèves moyens, bons élèves) afin de tester si la non linéarité des fonctions de production éducationnelle est plus ou moins avérée selon les groupes.

## 5. Estimations par genre

D'un point de vue méthodologique, «établir que les différences entre les genres existent dans le processus éducationnel implique que les études actuelles sur la discrimination du marché peuvent être biaisées d'une manière ou d'une autre» (POLACHEK, 1978). Une attention particulière est donc accordée au genre. Repré- nous le modèle de base qui a été estimé:

$$A_i^{(\theta)} = \beta_1 + \beta'_2 X_i + \beta'_3 F_i + \beta'_4 S_i + \varepsilon_i \quad (3)$$

et considérons les deux groupes de données suivants: le premier pour les filles et le second pour les garçons. On estime alors ce modèle pour les deux groupes séparément:

$$A_1^{(\theta)} = \beta_{11} + \beta'_{21} X_1 + \beta'_{31} F_1 + \beta'_{41} S_1 + \varepsilon_1 \quad \text{pour le groupe 1: filles} \quad (3a)$$

$$A_2^{(\theta)} = \beta_{12} + \beta'_{22} X_2 + \beta'_{32} F_2 + \beta'_{42} S_2 + \varepsilon_2 \quad \text{pour le groupe 2: garçons} \quad (3b)$$

Etant donné que nous supposons qu'il est pertinent de réaliser des estimations séparées pour les garçons et les filles, nous testons l'hypothèse nulle que les coefficients de la fonction de production éducationnelle pour les filles sont égaux aux coefficients de la fonction de production éducationnelle pour les garçons. L'hypothèse nulle peut être spécifiée de la manière suivante:

$$H_0 : \beta_{11} = \beta_{12}, \beta_{21} = \beta_{22}, \beta_{31} = \beta_{32} \text{ et } \beta_{41} = \beta_{42}$$

Le test de Chow (1960) effectué dans les trois domaines nous permet dans chacun des cas (lecture, mathématiques et sciences) de rejeter l'hypothèse nulle et de

conclure que les coefficients sont instables, autrement dit qu'il est pertinent de réaliser des estimations par genre<sup>25</sup>.

Le tableau 5A (en annexe) présente les résultats des estimations des EPF en lecture par genre. L'échantillon en lecture comprend 4771 élèves dont 50% de filles (2 383 observations) et 50% de garçons (2 388 observations). Par rapport aux résultats de l'estimation sur l'échantillon total, il est intéressant de constater que le pays de naissance de l'élève a un impact différent selon le genre. Celui-ci est significatif uniquement pour les garçons nés en Italie, Espagne, Portugal ou autres (-3.5%) alors que le coefficient est significatif à la fois pour les filles (-9.9%) et les garçons (-6.1%) nés en Ex-Yougoslavie, Albanie, Kosovo ou Turquie. On peut envisager que ces résultats traduisent des différences culturelles concernant le rôle des enfants selon le genre au sein de la famille.

Concernant les caractéristiques de la famille, il apparaît que la structure familiale a un impact négatif plus fort pour les filles que pour les garçons. Les filles issues d'une famille célibataire ont de moins bons résultats (-2.3%) que celles vivant dans une famille nucléaire alors que les résultats scolaires des garçons ne semblent pas être affectés par cette structure familiale particulière. En cas de séparation des parents, les traditions familiales font sans doute que ce sont les filles qui sont davantage sollicitées que les garçons en ce qui concerne les tâches ménagères. L'impact est également plus important pour les catégories mixte et autres.

Autre résultat intéressant, si la mère possède un niveau d'éducation inférieur à la scolarité obligatoire, l'impact sur le score est significativement négatif uniquement pour les filles (-5.3%) alors que si la mère possède un niveau d'éducation supérieur à la scolarité obligatoire, l'impact est positif à la fois pour les filles (+3.4% et +6.2%) et les garçons (+3.4% et +6.5%). En revanche, si le père possède un niveau d'éducation inférieur à la scolarité obligatoire, l'impact est significativement négatif uniquement pour les garçons (-2.7%). Si le niveau d'éducation du père est supérieur à la scolarité obligatoire, l'effet sur le score est positif (environ +4% pour les filles et les garçons).

Par rapport à une activité à plein temps, le fait que la mère travaille à temps partiel a un impact positif pour les garçons et les filles. La catégorie autre (par exemple: ménagère ou retraitée) a un impact significatif et positif uniquement pour les garçons. L'activité actuelle du père en revanche ne semble pas être une

25 Etant donné que l'on obtient des  $\theta$  différents selon le genre, nous estimons les deux modèles en imposant un  $\theta$  unique (celui de l'estimation sur l'échantillon total). Le test de Chow donne les résultats suivants: 3.68 en lecture, 4.25 en mathématiques et 2.13 en science. Ici,  $F_{lu}(40, \infty) = 1.39$  (à 5%) et 1.59 (à 1%). Etant donné que  $F_{calculé} > F_{lu}$ , l'hypothèse nulle est rejetée.

variable ayant un pouvoir explicatif fort sur les résultats, ceci s'expliquant certainement par une faible variabilité dans les données (étant donné que la majorité des pères travaille à temps plein).

Au sujet des caractéristiques de l'école, celles qui ont le plus d'impact sur les scores sont la localisation et le niveau de formation des enseignants. Par rapport à la Suisse romande, les élèves du nord-ouest, de Suisse centrale et de Suisse orientale réussissent mieux. Il existe des différences selon le genre mais celles-ci sont difficilement interprétables. Enfin, la proportion de professeur avec un niveau isced 5 en pédagogie est une variable fortement significative et positive. Il est intéressant de constater que le coefficient est beaucoup plus fort pour les garçons (+12.9%) que pour les filles (+9.3%). Ces estimations ont également été produites en mathématiques et en sciences et on retrouve des résultats identiques (les échantillons sont plus petits).

## 6. Conclusion

Ce papier présente les estimations d'EPF pour la Suisse à partir de l'échantillon national de PISA 2000. En utilisant la spécification économétrique favorisée par le modèle Box-Cox, les résultats mettent en évidence l'importance de:

- *l'intégration de l'élève dans son environnement scolaire.* L'impact de la nationalité de l'élève et de la langue parlée à la maison sur les résultats scolaires suggère que l'amélioration des performances de la Suisse passe sans doute par l'intégration de ses immigrants;
- *la situation économique de la famille.* La structure familiale, l'éducation des parents ainsi que la situation de la mère sur le marché du travail sont des variables qui ont un impact significatif sur les performances des élèves. Les résultats indiquent que les situations pouvant affecter l'économie de la famille, comme par exemple la séparation des parents, ont un impact négatif sur les performances de l'élève;
- *la spécificité de la Suisse du point de vue de l'organisation de son système éducatif.* Grâce à la compétence éducative cantonale, vingt-six systèmes scolaires différents coexistent et il apparaît que certaines régions performant mieux que d'autres. Malheureusement, il est difficile avec les données PISA de pousser d'avantage l'analyse, notamment parce que les sous-échantillons cantonaux sont trop petits. On ne peut que souhaiter que les pouvoirs publics deviennent attentifs au fait que l'expérience naturelle unique de la Suisse devrait être davantage exploitée;

- la *formation des enseignants*, contribuant ainsi de manière significative au débat sur l'impact des inputs scolaires sur les performances des élèves. Durant les années 1990, la Suisse a été amenée à modifier la répartition de ses dépenses publiques totales, contrainte par des restrictions budgétaires. La forte augmentation des dépenses engagées dans le poste de la prévoyance sociale s'est faite au détriment des dépenses publiques d'éducation (OFS, 2000). Or s'il faut à l'avenir se questionner sur la meilleure manière d'allouer les ressources en éducation, la formation des enseignants apparaît être un poste clé en terme de performance. Précisons cependant qu'il aurait été nécessaire d'accorder davantage d'importance à ce que l'on retrouve dans la littérature sous le nom d'effets de paires. Malheureusement les données disponibles ne nous permettent pas de déterminer les caractéristiques de la classe de l'élève. Etant donné que les élèves ont été tirés au sort dans chaque école, rien ne nous permet de conclure que ceux sélectionnés sont représentatifs des autres élèves présents dans la classe de chaque individu. Il serait vraiment utile à l'avenir, de pouvoir disposer de cette information;
- la *forme fonctionnelle*. Par rapport à la littérature traditionnelle, la forme fonctionnelle linéaire n'est pas imposée aux données a priori mais déterminée à l'aide du modèle Box-Cox (1964). Les résultats obtenus indiquent que l'utilisation largement répandue de la forme linéaire n'est pas appropriée. Ils indiquent également que les coefficients obtenus par les moindres carrés ordinaires diffèrent légèrement de ceux obtenus avec le modèle Box-Cox. Ces différences peuvent provenir du fait que le modèle Box-Cox propose une transformation qui permet (étant donné que le modèle linéaire est rejeté) d'améliorer l'ajustement du modèle aux données. Par ailleurs, des estimations Box-Cox effectuées sur des sous-échantillons définis selon différents niveaux de performance (par exemple: élèves faibles, élèves moyens, élèves bons) permettraient de tester si la non linéarité des fonctions de production éducationnelle est plus ou moins avérée selon les groupes;
- la *caractéristique genre dans les EPF*. Les estimations par genre mettent en évidence des différences significatives entre les filles et les garçons. Il apparaît que les filles et les garçons ne sont pas sensibles aux mêmes caractéristiques que ce soit au niveau de la famille ou de l'école: les filles semblent marginalement plus sensibles à l'environnement familial et les garçons, plus sensibles à l'environnement scolaire.

L'importante question des ressources scolaires n'a malheureusement pas pu être abordée dans cette étude. L'estimation de l'efficacité du système d'éducation suisse en utilisant une base de données qui combinerait à la fois les données de

l'échantillon national de PISA 2000 et les dépenses publiques d'éducation permettrait de poursuivre cette recherche.

## Annexe

Tableau A1: Statistiques descriptives et définitions des variables

Variables	N	Moy.	Std. Dev.	Définitions
Lecture	4771 (182)	510.823 (502.87)	85.2542 (55.8682)	Score en lecture (min = 199.988; max = 740.526)
Math	2637 (181)	545.812 (536.84)	82.4107 (51.3786)	Score en mathématiques (min = 254.594; max = 785.802)
Sciences	2669 (181)	509.410 (501.76)	87.7586 (59.0120)	Score en sciences (min = 175.306; max = 766.680)
CARACTÉRISTIQUES DE L'ÉLÈVE				
genre	4771 (182)	0.4995 (0.4871)	0.5000 (0.1369)	1 = si l'élève est une fille; 0 = si l'élève est un garçon
national1	4771 (182)	0.8833 (0.8745)	0.3212 (0.1202)	1 = si Suisse; 0 = sinon
national2	4771 (182)	0.0109 (0.0100)	0.1038 (0.0241)	1 = si Allemagne ou Autriche, France ou Belgique; 0 = sinon
national3	4771 (182)	0.0497 (0.0493)	0.2173 (0.0673)	1 = si Italie, Espagne, Portugal ou autres pays; 0 = sinon
national4	4771 (182)	0.0561 (0.0661)	0.2303 (0.0940)	1 = si Ex-Yougoslavie, Albanie ou Kosovo, Turquie; 0 = sinon
âge	4771 (182)	188 (188)	7.5865 (4.1881)	Age de l'élève en mois (min = 168; max = 228)
CARACTÉRISTIQUES DE LA FAMILLE				
langue	4771 (182)	0.8386 (0.8278)	0.3679 (0.1541)	1 = si la langue la plus souvent parlée à la maison est celle du test; 0 = sinon
célibataire	4771 (182)	0.0998 (0.1010)	0.2997 (0.1010)	1 = si célibataire (les élèves ont déclaré vivre avec la mère ou le père, ou une tutrice ou un tuteur); 0 = sinon
nucléaire	4771 (182)	0.8181 (0.8110)	0.3858 (0.1139)	1 = si nucléaire (les élèves ont déclaré vivre avec leur père et leur mère); 0 = sinon
mixte	4771 (182)	0.0543 (0.0544)	0.2266 (0.0538)	1 = si mixte (les élèves ont déclaré vivre avec leur père et une autre femme ou leur mère et un autre homme ou avec deux tuteurs); 0 = sinon

Variables	N	Moy.	Std. Dev.	Définitions
autre	4 771 (182)	0.0279 (0.0306)	0.1646 (0.0572)	1 = si autre (les élèves ont déclaré une autre combinaison); 0 = sinon
nsib	4 771 (182)	1.6940 (1.6930)	1.0531 (0.3405)	Nombre de frère(s) et sœur(s) (min = 0; max = 12)
med1	4 771 (182)	0.0717 (0.0784)	0.2580 (0.0902)	1 = si la mère n'a pas été à l'école ou a terminé l'école primaire (4-6 ans de scolarité); 0 = sinon
med2	4 771 (182)	0.3607 (0.3798)	0.4803 (0.1712)	1 = si la mère a terminé la scolarité obligatoire (7-9 ans de scolarité); 0 = sinon
med3	4 771 (182)	0.3748 (0.3570)	0.4841 (0.1322)	1 = si la mère a terminé un apprentissage, une école professionnelle ou une école de culture générale de niveau diplôme; 0 = sinon
med4	4 771 (182)	0.1928 (0.1847)	0.3946 (0.1440)	1 = si la mère a terminé avec succès une école du niveau de la maturité (gymnase, collège, école normale); 0 = sinon
mteduc	4 771 (182)	0.1532 (0.1544)	0.3602 (0.1237)	1 = si la mère a terminé une formation supérieure (université, école polytechnique, école professionnelle supérieure); 0 = sinon
fed1	4 771 (182)	0.0651 (0.0674)	0.2469 (0.0770)	1 = si le père n'a pas été à l'école ou a terminé l'école primaire (4-6 ans de scolarité); 0 = sinon
fed2	4 771 (182)	0.3182 (0.3410)	0.4658 (0.1717)	1 = si le père a terminé la scolarité obligatoire (7-9 ans de scolarité); 0 = sinon
fed3	4 771 (182)	0.4058 (0.3983)	0.4911 (0.1573)	1 = si le père a terminé un apprentissage, une école professionnelle ou une école de culture générale de niveau diplôme; 0 = sinon
fed4	4 771 (182)	0.2109 (0.1933)	0.4080 (0.1515)	1 = si le père a terminé avec succès une école du niveau de la maturité (gymnase, collège, école normale); 0 = sinon
fteduc	4 771 (182)	0.3167 (0.3108)	0.4652 (0.1683)	1 = si le père a terminé une formation supérieure (université, école polytechnique, école professionnelle supérieure); 0 = sinon
mwork1	4 771 (182)	0.2310 (0.2343)	0.4215 (0.1167)	1 = si la mère travaille à temps plein; 0 = sinon
mwork2	4 771 (182)	0.4112 (0.4077)	0.4921 (0.1423)	1 = si la mère travaille à temps partiel; 0 = sinon
mwork3	4 771 (182)	0.0243 (0.0231)	0.1540 (0.0366)	1 = si la mère ne travaille pas mais cherche un emploi; 0 = sinon
mwork4	4 771 (182)	0.3335 (0.3348)	0.4715 (0.1249)	1 = si autre (ex: tâches ménagères, retraitée); 0 = sinon

Variables	N	Moy.	Std. Dev.	Définitions
fwork1	4 771 (182)	0.9122 (0.9052)	0.2831 (0.0919)	1 = si le père travaille à temps plein; 0 = sinon
fwork2	4 771 (182)	0.0367 (0.0388)	0.1880 (0.0467)	1 = si le père travaille à temps partiel; 0 = sinon
fwork3	4 771 (182)	0.0115 (0.0117)	0.1068 (0.0268)	1 = si le père ne travaille pas mais cherche un emploi; 0 = sinon
fwork4	4 771 (182)	0.0396 (0.0442)	0.1951 (0.0661)	1 = si autre (ex: tâches ménagères, retraité); 0 = sinon
devoirm	4 771 (182)	0.3261 (0.3326)	0.4688 (0.1376)	1 = si la mère aide <i>plusieurs fois par mois</i> ou <i>plusieurs fois par semaine</i> aux devoirs; 0 = si la mère n'aide <i>jamais</i> ou <i>presque jamais, quelques fois par an</i> ou <i>environ une fois par mois</i> aux devoirs
devoirp	4 771 (182)	0.2345 (0.2307)	0.4238 (0.1070)	1 = si le père aide <i>plusieurs fois par mois</i> ou <i>plusieurs fois par semaine</i> aux devoirs; 0 = si le père n'aide <i>jamais</i> ou <i>presque jamais, quelques fois par an</i> ou <i>environ une fois par mois</i> aux devoirs
devoirfs	4 771 (182)	0.1821 (0.1796)	0.3860 (0.0972)	1 = si les frère(s) et sœur(s) aident <i>plusieurs fois par mois</i> ou <i>plusieurs fois par semaine</i> aux devoirs; 0 = si les frère(s) et sœur(s) n'aident <i>jamais</i> ou <i>presque jamais, quelques fois par an</i> ou <i>environ une fois par mois</i> aux devoirs
<b>CARACTÉRISTIQUES DE L'ÉCOLE</b>				
romande	4 771 (182)	0.3282 (0.2857)	0.4696 (0.4530)	1 = Fribourg, Genève, Jura, Neuchâtel, Valais, Vaud ou Tessin; 0 = sinon
nordw	4 771 (182)	0.2532 (0.2912)	0.4349 (0.4556)	1 = Berne, Soleure, Bâle-Campagne, Bâle-Ville ou Argovie; 0 = sinon
centrale	4 771 (182)	0.0796 (0.0769)	0.2708 (0.2672)	1 = Lucerne, Schwytz, Obwald ou Zoug; 0 = sinon
orientale	4 771 (182)	0.3389 (0.3461)	0.4734 (0.4771)	1 = Zurich, Glaris, Shaffhouse, Appenzell Rh.-Ext., Saint-Gall, Grisons ou Thurgovie; 0 = sinon
village	4 771 (182)	0.1258 (0.1758)	0.3316 (0.3817)	1 = si le nombre d'habitants est inférieur à 3 000; 0 = sinon
ptville	4 771 (182)	0.5615 (0.5440)	0.4962 (0.4994)	1 = si le nombre d'habitants est entre 3 000 et 15 000; 0 = sinon
ville	4 771 (182)	0.2362 (0.1923)	0.4248 (0.3952)	1 = si le nombre d'habitants est entre 15 000 et 100 000; 0 = sinon
gdville	4 771 (182)	0.0765 (0.0879)	0.2658 (0.2839)	1 = si le nombre d'habitants est entre 100 000 et un million environ; 0 = sinon



Variables	(1) MCO		(2) Effets fixes		(3) Effets aléatoires	
	Coeff.	t	Coeff.	t	Coeff.	z
med3	17.705	(6.175)***	15.066	(5.723)***	15.666	(5.933)***
med4	32.426	(8.853)***	23.700	(6.995)***	25.456	(7.493)***
mteduc	-7.027	(2.103)**	-5.791	(1.931)*	-5.917	(1.947)*
fed1	-13.236	(2.315)**	-7.528	(1.502)	-8.914	(1.751)*
fed2	réf.	réf.	réf.	réf.	réf.	réf.
fed3	18.343	(6.176)***	15.731	(5.756)***	16.194	(5.917)***
fed4	24.388	(6.375)***	15.810	(4.462)***	17.500	(4.950)***
fteduc	6.144	(2.270)**	1.961	(0.795)	2.778	(1.118)
mwork1	réf.	réf.	réf.	réf.	réf.	réf.
mwork2	9.123	(3.127)***	5.058	(1.937)*	5.767	(2.178)**
mwork3	-8.859	(1.195)	-4.490	(0.707)	-5.769	(0.874)
mwork4	6.170	(2.012)**	3.043	(1.115)	3.578	(1.288)
fwork1	réf.	réf.	réf.	réf.	réf.	réf.
fwork2	-9.368	(1.490)	-9.187	(1.703)*	-9.430	(1.674)*
fwork3	2.429	(0.219)	-1.210	(0.124)	0.145	(0.015)
fwork4	-5.382	(0.911)	-5.180	(0.992)	-5.281	(0.996)
devoirm	-8.206	(3.153)***	-6.799	(2.905)***	-6.991	(2.957)***
devoirp	-6.325	(2.193)**	-3.532	(1.341)	-4.164	(1.575)
devoirfs	-6.940	(2.476)**	-7.397	(2.914)***	-7.243	(2.829)***
romande	réf.	réf.			réf.	réf.
nordw	29.620	(9.402)***			21.254	(3.309)***
centrale	28.150	(6.195)***			20.885	(2.212)**
orientale	18.273	(6.038)***			6.451	(1.055)
village	-9.970	(2.402)**			-16.729	(1.855)*
ptville	-4.591	(1.690)*			-7.542	(1.300)
ville	réf.	réf.			réf.	réf.
gdville	-7.725	(1.684)*			-16.190	(1.735)*
schlsize	0.034	(7.178)***			0.028	(3.000)***
stratio	-1.564	(4.583)***			-1.583	(2.060)**
scl5q01	-0.084	(2.688)***			-0.047	(0.719)
propqual	57.414	(16.515)***			61.937	(8.937)***

Variables	(1) MCO		(2) Effets fixes		(3) Effets aléatoires	
	Coeff.	t	Coeff.	t	Coeff.	z
propcert	10.574	(2.553)**			-0.457	(0.047)
constante	726.117	(25.187)***	783.492	(28.656)***	753.598	(26.15)***
N	4 566		4 566 (159 écoles)		4 566 (159 écoles)	
R <sup>2</sup>	0.3298		<i>within:</i>	0.2021	<i>within:</i>	0.2018
			<i>between:</i>	0.4120	<i>between:</i>	0.5601
			<i>overall:</i>	0.2237	<i>overall:</i>	0.3189

Source: PISA 2000. Notes: \* significatif à 10%; \*\* significatif à 5%; \*\*\* significatif à 1%. Statistiques t et z robustes entre parenthèses. L'échantillon en lecture est plus petit car seules les écoles dont au moins 20 élèves ont participé aux tests PISA ont été conservées.

Tableau A3: Estimation des fonctions de production éducationnelle en mathématiques

Variables	(1) MCO		(2) Effets fixes		(3) Effets aléatoires	
	Coeff.	t	Coeff.	t	Coeff.	z
genre	-27.703	(10.325)***	-29.160	(11.51)***	-28.625	(11.27)***
national1	réf.	réf.	réf.	réf.	réf.	réf.
national2	-13.437	(1.210)	-13.345	(1.167)	-14.047	(1.342)
national3	-25.668	(3.537)***	-24.025	(3.721)***	-24.619	(3.631)***
national4	-47.025	(5.731)***	-44.495	(5.934)***	-45.797	(5.908)***
âge	-1.660	(8.369)***	-1.912	(9.895)***	-1.814	(9.502)***
langue	24.686	(5.026)***	18.524	(4.078)***	21.172	(4.568)***
célibataire	-14.733	(3.141)***	-9.759	(2.250)**	-11.667	(2.633)***
nucléaire	réf.	réf.	réf.	réf.	réf.	réf.
mixte	-18.258	(3.024)***	-16.069	(2.864)***	-16.997	(3.020)***
autre	-34.025	(3.721)***	-33.723	(4.303)***	-33.309	(3.982)***
nsib	-1.883	(1.388)	-0.624	(0.491)	-0.997	(0.775)
med1	-11.398	(1.613)	-11.173	(1.714)*	-11.307	(1.708)*
med2	réf.	réf.	réf.	réf.	réf.	réf.
med3	14.939	(4.026)***	11.498	(3.272)***	12.922	(3.678)***
med4	32.738	(7.113)***	24.467	(5.515)***	27.500	(6.244)***
mteeduc	-1.331	(0.312)	-2.510	(0.618)	-1.963	(0.483)

Variables	(1) MCO		(2) Effets fixes		(3) Effets aléatoires	
	Coeff.	t	Coeff.	t	Coeff.	z
fed1	-10.970	(1.604)	-5.838	(0.941)	-8.084	(1.274)
fed2	réf.	réf.	réf.	réf.	réf.	réf.
fed3	12.716	(3.355)***	11.946	(3.352)***	12.123	(3.391)***
fed4	25.361	(5.248)***	16.666	(3.660)***	19.551	(4.274)***
fteduc	1.735	(0.507)	-0.278	(0.086)	0.624	(0.193)
mwork1	réf.	réf.	réf.	réf.	réf.	réf.
mwork2	12.754	(3.415)***	6.925	(1.954)*	8.896	(2.500)**
mwork3	-6.333	(0.658)	-5.460	(0.634)	-6.374	(0.708)
mwork4	9.635	(2.448)**	6.093	(1.656)*	7.234	(1.943)*
fwork1	réf.	réf.	réf.	réf.	réf.	réf.
fwork2	-15.978	(1.919)*	-15.229	(2.078)**	-15.937	(2.074)**
fwork3	-13.427	(0.891)	-19.652	(1.399)	-17.168	(1.219)
fwork4	-4.284	(0.616)	-4.165	(0.639)	-4.331	(0.653)
devoirm	-15.703	(4.742)***	-12.942	(4.073)***	-13.859	(4.411)***
devoirp	-7.002	(1.888)*	-5.173	(1.443)	-5.963	(1.680)*
devoirfs	-4.875	(1.375)	-4.106	(1.213)	-4.235	(1.252)
romande	réf.	réf.			réf.	réf.
nordw	25.839	(6.309)***			22.149	(3.407)***
centrale	27.371	(4.914)***			24.384	(2.813)***
orientale	19.288	(5.030)***			13.561	(2.271)**
village	-4.630	(0.913)			-8.914	(1.054)
ptville	-7.587	(2.218)**			-9.304	(1.625)
ville	réf.	réf.			réf.	réf.
gdville	-11.625	(2.062)**			-16.415	(1.831)*
schlsize	0.035	(5.985)***			0.033	(3.774)***
stratio	-1.042	(2.544)**			-1.047	(1.484)
scl5q01	-0.025	(0.610)			-0.005	(0.071)
propqual	70.206	(15.938)***			73.439	(10.62)***
propcert	0.379	(0.072)			-3.923	(0.400)
constante	788.718	(21.163)***	898.970	(24.40)***	830.438	(22.45)***

Variables	(1) MCO		(2) Effets fixes		(3) Effets aléatoires	
	Coeff.	t	Coeff.	t	Coeff.	z
N	2525		2525 (159 écoles)		2525 (159 écoles)	
R <sup>2</sup>	0.3550		<i>within:</i> 0.2360		<i>within:</i> 0.2349	
			<i>between:</i> 0.2306		<i>between:</i> 0.6015	
			<i>overall:</i> 0.2186		<i>overall:</i> 0.3505	

Source: PISA 2000. Notes: \* significatif à 10%; \*\* significatif à 5%; \*\*\* significatif à 1%. Statistiques t et z robustes entre parenthèses. L'échantillon en mathématiques est plus petit car seules les écoles dont au moins 20 élèves ont participé aux tests PISA ont été conservées.

Tableau A4: Estimation des fonctions de production éducationnelle en sciences

Variables	(1) MCO		(2) Effets fixes		(3) Effets aléatoires	
	Coeff.	t	Coeff.	t	Coeff.	z
genre	-14.955	(5.202)***	-16.862	(6.335)***	-16.379	(6.134)***
national1	réf.	réf.	réf.	réf.	réf.	réf.
national2	2.445	(0.202)	-1.158	(0.102)	0.360	(0.032)
national3	-21.102	(2.878)***	-22.811	(3.694)***	-22.888	(3.481)***
national4	-43.534	(5.091)***	-39.538	(5.366)***	-41.753	(5.436)***
âge	-1.633	(7.858)***	-1.598	(8.235)***	-1.600	(8.215)***
langue	30.732	(5.599)***	26.955	(5.541)***	28.216	(5.612)***
célibataire	-19.193	(3.819)***	-16.786	(3.732)***	-17.114	(3.747)***
nucléaire	réf.	réf.	réf.	réf.	réf.	réf.
mixte	-11.012	(1.620)	-1.674	(0.275)	-4.520	(0.732)
autre	-26.631	(2.551)**	-18.546	(1.995)**	-20.619	(2.152)**
nsib	-3.495	(1.791)*	-1.445	(0.783)	-2.003	(1.077)
med1	-15.494	(2.038)**	-8.078	(1.207)	-9.840	(1.431)
med2	réf.	réf.	réf.	réf.	réf.	réf.
med3	21.470	(5.449)***	18.247	(4.904)***	18.984	(5.167)***
med4	41.814	(8.308)***	31.994	(6.818)***	34.750	(7.434)***
mteduc	-12.226	(2.625)***	-8.735	(2.091)**	-9.814	(2.306)**
fed1	-3.903	(0.500)	4.029	(0.570)	1.076	(0.151)
fed2	réf.	réf.	réf.	réf.	réf.	réf.
fed3	19.202	(4.759)***	14.981	(3.973)***	16.291	(4.354)***
fed4	22.046	(4.329)***	15.008	(3.060)***	17.228	(3.592)***

Variables	(1) MCO		(2) Effets fixes		(3) Effets aléatoires	
	Coeff.	t	Coeff.	t	Coeff.	z
fteduc	9.297	(2.550)**	2.898	(0.856)	4.835	(1.433)
mwork1	réf.	réf.	réf.	réf.	réf.	réf.
mwork2	8.650	(2.212)**	6.030	(1.726)*	6.554	(1.835)*
mwork3	-3.533	(0.284)	2.029	(0.194)	-0.339	(0.030)
mwork4	7.299	(1.785)*	3.854	(1.039)	4.728	(1.259)
fwork1	réf.	réf.	réf.	réf.	réf.	réf.
fwork2	4.797	(0.580)	3.601	(0.468)	3.970	(0.509)
fwork3	-28.416	(2.236)**	-25.136	(2.008)**	-24.848	(2.122)**
fwork4	4.735	(0.491)	4.875	(0.556)	4.301	(0.483)
devoirm	-6.197	(1.710)*	-5.845	(1.762)*	-5.863	(1.754)*
devoirp	-13.618	(3.445)***	-9.592	(2.630)***	-10.664	(2.920)***
devoirfs	-7.360	(1.918)*	-8.397	(2.367)**	-8.183	(2.301)**
romande	réf.	réf.			réf.	réf.
nordw	30.492	(6.869)***			22.375	(2.937)***
centrale	23.084	(3.772)***			15.462	(1.480)
orientale	18.421	(4.420)***			7.612	(1.071)
village	-6.182	(1.097)			-13.225	(1.306)
ptville	-1.391	(0.377)			-5.458	(0.838)
ville	réf.	réf.			réf.	réf.
gdville	-9.352	(1.502)			-17.849	(1.690)*
schlsize	0.045	(7.223)***			0.038	(3.637)***
stratio	-1.206	(2.772)***			-1.159	(1.522)
scl5q01	-0.061	(1.447)			-0.048	(0.647)
propqual	65.761	(13.86)***			69.314	(8.688)***
propcert	11.008	(1.855)*			1.781	(0.153)
constante	721.124	(18.31)***	784.478	(21.05)***	738.212	(19.05)***
N	2 552		2 552 (159 écoles)		2 552 (159 écoles)	
R <sup>2</sup>	0.3458		<i>within:</i> 0.2117		<i>within:</i> 0.2110	
			<i>between:</i> 0.3715		<i>between:</i> 0.5655	
			<i>overall:</i> 0.2236		<i>overall:</i> 0.3369	

Source: PISA 2000. Notes: \* significatif à 10%; \*\* significatif à 5%; \*\*\* significatif à 1%. Statistiques t et z robustes entre parenthèses. L'échantillon en sciences est plus petit car seules les écoles dont au moins 20 élèves ont participé aux tests PISA ont été conservées.

Tableau A5: Estimation Box-Cox des fonctions de production éducationnelle par genre

Variables	(1) Lecture			
	Filles		Garçons	
	relative (%)	absolue	relative (%)	absolue
national1	réf.	réf.	réf.	réf.
national2	0.63	3.26	1.96	9.83
national3	-1.20	-6.28	-3.53**	-17.67
national4	-9.85***	-51.31	-6.06***	-30.32
âge	-0.30***	-1.54	-0.32***	-1.61
langue	4.74***	24.71	6.82***	34.13
célibataire	-2.35***	-12.22	-1.47	-7.34
nucléaire	réf.	réf.	réf.	réf.
mixte	-2.59**	-13.48	-0.61	-3.05
autre	-5.25***	-27.34	-3.87**	-19.36
nsib	-0.46*	-2.37	-1.28***	-6.41
med1	-5.30***	-27.60	0.28	1.39
med2	réf.	réf.	réf.	réf.
med3	3.37***	17.55	3.41***	17.07
med4	6.20***	32.29	6.47***	32.40
mteduc	-1.06	-5.51	-1.86**	-9.33
fed1	-1.00	-5.23	-2.72*	-13.63
fed2	réf.	réf.	réf.	réf.
fed3	3.88***	20.23	3.25***	16.25
fed4	4.39***	22.87	5.05***	25.29
fteduc	0.97	5.08	1.51**	7.57
mwork1	réf.	réf.	réf.	réf.
mwork2	1.31*	6.82	1.96**	9.80
mwork3	-3.22*	-16.75	-0.05	-0.27
mwork4	-0.48	-2.49	2.44***	12.22
fwork1	réf.	réf.	réf.	réf.
fwork2	-0.80	-4.19	-2.45	-12.27
fwork3	1.23	6.39	0.85	4.25
fwork4	-1.59	-8.30	0.44	2.22

Variables	(1) Lecture			
	Filles		Garçons	
	relative (%)	absolue	relative (%)	absolue
devoirm	-3.20***	-16.66	-0.08	-0.39
devoirp	-0.23	-1.18	-2.17***	-10.85
devoirfs	-1.27*	-6.60	-1.75**	-8.75
romande	réf.	réf.	réf.	réf.
nordw	5.75***	29.98	6.67***	33.37
centrale	6.41***	33.37	4.28***	21.44
orientale	3.53***	18.37	4.88***	24.42
village	-1.27	-6.61	-2.53**	-12.65
ptville	-0.67	-3.48	-0.78	-3.92
ville	réf.	réf.	réf.	réf.
gdville	-1.63	-8.51	-1.06	-5.32
schlsize	0.01***	0.04	0.01***	0.04
stratio	-0.28***	-1.46	-0.37***	-1.83
sc15q01	-0.01	-0.05	-0.02**	-0.10
propqual	9.26***	48.24	12.91***	64.63
propcert	0.35	1.85	2.61**	13.06

Source: PISA 2000. Notes: \* $\chi^2 < 10\%$ , \*\* $\chi^2 < 5\%$ , \*\*\* $\chi^2 < 1\%$ . Observations: 2 383 filles et 2 388 garçons.

## Bibliographie

- ARIAS, J. J. and D. M. WALKER (2004), "Additional Evidence on the Relationship between Class Size and Student Performance", *Journal of Economic Education*, pp. 311-329.
- BAKER, D. P. and D. PERKINS (1993), "Creating Gender Equality: Cross-national Gender Stratification and Mathematical Performance", *Sociology of Education*, 66 (2), pp. 91-103.
- BECKER, G. and N. TOMES (1976), "Child Endowments and the Quantity and Quality of Children", *Journal of Political Economy*, 84 (4), pp. 143-162.
- BOX, G. E. P. and D. R. COX (1964), "An Analysis of Transformations", *Journal of the Royal Statistical Society*, 26 (2), pp. 211-252.

- BUCHINSKY, M. (1995), "Quantile Regression, Box-Cox Transformation Model, and the U.S. Wage Structure, 1963-1987", *Journal of Econometrics*, 65, pp. 109-154.
- CHOW, G. C. (1960), "Tests of Equality between Sets of Coefficients in Two Linear Regressions", *Econometrica*, 28, pp. 591-605.
- DUSTMANN, C. (2003), "The Class Size Debate and Educational Mechanisms: Editorial", *Economic Journal*, 113, pp. 1-2.
- FIGLIO, D. N. (1999), "Functional Form and the Estimated Effects of School Resources", *Economics of Education Review*, 18, pp. 241-252.
- FISCHER, J. A. V. (2005), "The Impact of Direct Democracy on Public Education: Performance of Swiss Students in Reading", Discussion paper no. 2005-10, Universitat St. Gallen.
- GODFREY, L. G. and M. R. WICKENS (1981), "Testing Linear and Log-Linear Regressions for Functional Form", *Review of Economic Studies*, XLVIII, pp. 487-496.
- GOLDHABER, D. D. and D. J. BREWER (1997), "Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity", *Journal of Human Resources*, XXXII (3), pp. 505-523.
- HANUSHEK, E. A. (1979), "Conceptual and Empirical Issues in the Estimation of Educational Production Functions", *Journal of Human Resources*, 14 (3), pp. 351-388.
- HANUSHEK, E. A. (1986), "The Economics of Schooling: Production and Efficiency in Public Schools", *Journal of Economic Literature*, 24, pp. 1141-1177.
- HANUSHEK, E. A. (2003), "The Failure of Input-Based Schooling Policies", *Economic Journal*, 113, pp. 64-98.
- HAVEMAN, R. and B. WOLFE (1995), "The Determinants of Children's Attainments: A Review of Methods and Findings", *Journal of Economic Literature*, 33 (4), pp. 1829-1878.
- KRUEGER, A. B. (1998), "Reassessing the View That American Schools Are Broken", Federal Research Bank of New-York Economic Policy Review, 4 (1), pp. 29-46.
- KRUEGER, A. B. (2003), "Economic Considerations and Class Size", *Economic Journal*, 113(February), pp. 34-63.
- LEIBOWITZ, A. (1974), "Home Investment in Children", *Journal of Political Economy*, 82 (2), pp. 111-131.
- MACHADO, J. A. F. and J. MATA (2000), "Box-Cox Quantile Regression and the Distribution of Firm Size", *Journal of Applied Econometrics*, 15, pp. 253-274.
- MONTMARQUETTE, C. and S. MAHSEREDJIAN (1985), "Functional Forms and Educational Production Functions", *Economic Letters*, 19, pp. 291-294.

- MOULTON, B. R. (1990), "An Illustration of a Pitfall in Estimating The Effects of Aggregate Variables on Micro Units", *Review of Economics and Statistics*, 72 (2), pp. 334-338.
- OFFICE FÉDÉRAL DE LA STATISTIQUE (1997), *La mobilité scolaire en Suisse. Une contribution aux rapports sociaux pour la Suisse*, Berne.
- OFFICE FÉDÉRAL DE LA STATISTIQUE (1999), *Les indicateurs de l'enseignement en Suisse, 1999*, Neuchâtel.
- OFFICE FÉDÉRAL DE LA STATISTIQUE (2002), *Dépenses publiques d'éducation. Indicateurs financiers 2000*, Neuchâtel.
- OFFICE FÉDÉRAL DE LA STATISTIQUE et CONFÉRENCE SUISSE DES DIRECTEURS CANTONAUX DE L'INSTRUCTION PUBLIQUE (2002), *Préparés pour la vie? Les compétences des jeunes – Rapport national de l'enquête PISA 2000*, Neuchâtel.
- OFFICE FÉDÉRAL DE LA STATISTIQUE (2003), *Portrait démographique de la Suisse, Edition 2003*, Neuchâtel.
- OFFICE FÉDÉRAL DE LA STATISTIQUE (2004), *Annuaire statistique de la Suisse 2004*.
- ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUE (2001), *Regards sur l'éducation, Les indicateurs de l'OCDE*, Centre pour la recherche et l'innovation dans l'enseignement, Indicateurs des systèmes d'enseignement.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (2002), *Programme for International Student Assessment (PISA) – Manuel for the PISA 2000 Database*, OECD Publishing.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (2002), *PISA 2000 Technical Report*, OECD Publishing.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (2006), *Where Immigrant Students Succeed – A Comparative Review of Performance and Engagement in PISA 2003*, Paris: OECD.
- PSACHAROPOULOS, G. (2000), "Economics of Education à la Euro", *European Journal of Education*, 35(1), pp. 81-95.
- RIVKIN, S. G., E. A. HANUSHEK and J. F. KAIN (2005), "Teachers, Schools, and Academic Achievement", *Econometrica*, 73 (2), pp. 417-458.
- SCHULTZ, T. W. (1963), *The Economic Value of Education*, Columbia University Press.
- SUMMERS, A. and B. WOLFE (1977), "Do Schools Make a Difference?", *American Economic Review*, 67 (4), pp. 639-652.
- TODD, P. E. and K. I. WOLPIN (2003), "On the Specification and Estimation of the Production Function for Cognitive Achievement", *The Economic Journal*, 113, pp. 3-33.
- WOLTER, S. C. and M. CORADI VELLACOTT (2002), "Sibling Rivalry: A Look at Switzerland with PISA Data", IZA Discussion Paper n° 594, IZA-Bonn.

### SUMMARY

This paper estimates a production function for education for Switzerland using the sample from PISA 2000. Based on the Box-Cox model, the evidence shows that the test scores in reading, mathematics and natural sciences are significantly influenced by the school environment, the economic situation of the family, organisational specificities of Switzerland, and the educational formation of the teaching staff. There are also significant differences between girls and boys. Girls seem to respond more sensibly to the family background whereas boys are more responsive to the school environment. The econometric analysis suggests that the wide spread used linear model is not adequate for the sample at hand.

### ZUSAMMENFASSUNG

Der Aufsatz schätzt die Produktionsfunktion für Bildung aus der Erhebung PISA 2000 für die gesamte Schweiz. Unter Verwendung des Box\_Cox Modells wird gezeigt, dass die Integration eines Schülers in sein schulisches Umfeld, die ökonomische Situation der Familie, die Ausbildung des Lehrkörpers sowie die schweizerische Besonderheiten der Organisation der Ausbildung einen signifikanten Einfluss auf die Testergebnisse in Lesen, Mathematik und Naturwissenschaften haben. Weiter konnte festgestellt werden, dass es signifikante Unterschiede zwischen Mädchen und Buben gibt. Insbesondere sind reagieren Mädchen sensibler auf das familiäre Umfeld, während für Buben eher das schulische Umfeld wichtig ist. Ausserdem legen die Ergebnisse nahe, dass die weitverbreitete lineare Spezifikation für den gegebenen Datensatz nicht adäquat erscheint.

### RÉSUMÉ

Ce papier présente l'estimation de fonctions de production éducationnelle à partir de l'échantillon national de PISA 2000 pour la Suisse. En utilisant la spécification économétrique favorisée par le modèle Box-Cox (1964), les résultats montrent que l'intégration de l'élève dans son environnement scolaire, la situation économique de la famille, la spécificité de la Suisse du point de vue de l'organisation de son système éducatif et la formation des enseignants ont un impact significatif sur les scores en lecture, mathématiques et sciences. Les résultats indiquent également que l'utilisation largement répandue de la forme linéaire ne semble pas la

plus appropriée dans le cas des données utilisées. Enfin, les estimations par genre mettent en évidence des différences significatives entre les filles et les garçons. Les filles semblent plus sensibles à l'environnement familial alors que les garçons sont plus sensibles à l'environnement scolaire.