

**« Scolarisation de masse des  
garçons et des filles.  
Financement public de l'instruction primaire  
et croissance économique en France au  
XIXème siècle »**

Auteurs

**Claude Diebolt, Magali Jaoul-Grammare, Faustine Perrin**

Document de Travail n° 2020 – 51

*Novembre 2020*

Bureau d'Économie  
Théorique et Appliquée  
BETA

[www.beta-umr7522.fr](http://www.beta-umr7522.fr)

 @beta\_economics

Contact :  
[jaoulgrammare@beta-cnrs.unistra.fr](mailto:jaoulgrammare@beta-cnrs.unistra.fr)

# Scolarisation de masse des garçons et des filles

## Financement public de l'instruction primaire et croissance économique en France au XIX<sup>ème</sup> siècle

Claude Diebolt, Magali Jaoul-Grammare, Faustine Perrin\*

[15 novembre 2020]

**Résumé.** - L'objectif de cet article est d'étudier dans quelle mesure la centralisation du système éducatif français a profité à l'éducation de masse. Pour ce faire, nous utilisons des informations sur les financements liés à l'éducation alloués par l'État, les départements, les régions, les ménages et les entreprises de 1820 à 1913. Pour mener notre analyse, nous procédons en deux étapes. Tout d'abord, nous étudions l'évolution des différents types de financement dans le temps et nous nous appuyons sur la méthodologie des points atypiques pour détecter l'existence d'éventuelles ruptures dans les séries. Ensuite, nous analysons la causalité entre les différents types de financement et les taux de scolarisation des garçons et des filles. Les résultats originaux que nous obtenons révèlent que la scolarisation de masse en France au XIX<sup>ème</sup> siècle, a été davantage rendue possible par la volonté politique que par nécessité économique.

**Mots clés :** Instruction primaire ; Financement ; XIX<sup>ème</sup> siècle ; France

**Codes JEL :** C22 ; C81 ; C82 ; H52 ; I24 ; N33

---

\* BETA/CNRS (UMR 7522), Université de Strasbourg, 61 avenue de la Forêt Noire, France  
Email : [cdiebolt@unistra.fr](mailto:cdiebolt@unistra.fr) ; [jaoulgrammare@beta-cnrs.unistra.fr](mailto:jaoulgrammare@beta-cnrs.unistra.fr) ; [faustine.perrin@unistra.fr](mailto:faustine.perrin@unistra.fr).

## 1. Introduction

L'accumulation de capital humain est largement reconnue comme un moteur essentiel de la croissance économique. Une question demeure néanmoins largement ouverte, celle de savoir si l'éducation a d'abord été un investissement d'accompagnement ou, au contraire, si elle s'est imposée quasi instantanément comme l'une des causes majeures de la croissance économique, si ce n'est son facteur déterminant.

Pour mieux appréhender la relation de cause à effet entre la formation de capital humain et la croissance économique, nous pensons qu'il est avant tout nécessaire de mieux comprendre le phénomène de la scolarisation de masse et ses origines. Le système scolaire français par exemple, a considérablement évolué tout au long du XIX<sup>ème</sup> siècle. Le siècle a été marqué par la mise en œuvre d'un large éventail de lois par les pouvoirs publics qui ont remodelé l'organisation des écoles et ont conduit à la centralisation du système qui a finalement favorisé l'enseignement public.

Les données suggèrent que la scolarisation des garçons dans l'enseignement primaire était déjà élevée au début du XIX<sup>ème</sup> siècle et n'a cessé de croître pendant le reste du siècle. Si l'expansion de l'éducation de masse pour les garçons a été bien étudiée, la scolarisation des filles reste négligée. Bien que les différences entre les sexes aient persisté tout au long du XIX<sup>ème</sup> siècle, l'étude des taux de scolarisation montre des améliorations significatives dans l'accès des filles à l'enseignement primaire formel et un rattrapage du niveau des garçons au tournant du XX<sup>ème</sup> siècle. L'application des lois nationales et l'augmentation des financements publics de l'État ont probablement favorisé ce processus. En l'état actuel des connaissances, des recherches plus approfondies nous paraissent toutefois incontournables afin de mieux saisir l'évolution du système de financement des écoles et comment il a contribué à l'expansion de l'éducation de masse pour les garçons et les filles dans la France du XIX<sup>ème</sup> siècle.

L'objectif de cet article est, dès lors, d'étudier dans quelle mesure la centralisation du système éducatif français a profité à l'éducation de masse. Pour ce faire, nous utilisons des informations sur les financements liés à l'éducation alloués par l'État, les départements, les régions, les ménages et les entreprises au cours de la période 1820-1913. Pour mener notre analyse, nous procédons en deux étapes. Tout d'abord, nous analysons l'évolution des différents types de financement dans le temps et nous nous appuyons sur la méthodologie des points atypiques pour détecter l'existence d'éventuelles ruptures dans les séries. Ensuite, nous étudions la causalité entre les différents types de financement et les taux de scolarisation.

La détection des *outliers* dans les diverses dépenses d'éducation pour le niveau primaire met en évidence des points de rupture dans la structure des dépenses et souligne l'importance des lois Ferry (1881 et 1882) dans l'évolution du financement du primaire en France : l'État remplace les ménages et les communes dans le financement de l'éducation primaire. Au-delà d'une modification de la structure de financement, les diverses dépenses ont également impacté les taux de scolarisation. L'analyse de la causalité montre ainsi qu'une augmentation des dépenses des communes ou de l'État favorise le développement de la scolarisation. En revanche, il n'existe aucune relation de causalité entre la richesse produite (estimée ici à partir du PIB par habitant) et le nombre d'élèves inscrits dans l'éducation primaire, ni entre le PIB par habitant et les dépenses d'éducation. Nos résultats semblent ainsi suggérer que la scolarisation de masse ayant eu lieu en France au XIX<sup>ème</sup> siècle a été davantage rendue possible par la volonté politique que par nécessité économique.

## 2. La Perspective Cliométrique

Il existe encore aujourd'hui un fossé considérable entre l'avancée conceptuelle et la mesure quantitative des concepts ainsi produits. Certes, la mesure est difficile et les instruments à notre disposition sont loin de correspondre aux besoins. On ne peut cependant en tirer argument pour refuser une approche quantitative du réel. Un concept est, par nature, une image abstraite de la réalité, une construction de l'esprit ; seule la mesure le relie au réel. Sans elle, potentielle ou effective, il ne peut que rester stérile, sans possibilité de se transformer en action, en avancée vers de nouveaux champs de la connaissance.

Ainsi, après la Seconde Guerre mondiale, les développements de la statistique ont fait progresser le champ de la mesure et ont incontestablement permis de nouveaux progrès de la théorie. Néanmoins, les progrès de l'observation n'ont, bien trop souvent, concernés que le temps présent. L'analyse du très long terme, qui n'a pas entraîné la production statistique correspondante, a dû se contenter des indicateurs existants, souvent très éloignés des concepts que l'on se proposait d'approcher. La production statistique de l'après Seconde Guerre mondiale offre, bien évidemment, un recul non négligeable, mais insuffisant si l'on veut couvrir un champ supérieur à une cinquantaine d'années. Pour ce qui concerne l'analyse historique des relations de causalités entre éducation et économie, par exemple, elle donne des instruments très puissants pour étudier la période allant de 1950 à nos jours, mais elle ne permet pas d'analyser avec précision les périodes

précédentes et surtout les ruptures provoquées par les guerres, les changements institutionnels ou encore les crises, qu'elles soient économiques, démographiques, sanitaires, etc. Mais, puisque nous interrogeons le passé, la construction statistique nouvelle ne peut se faire qu'avec des matériaux que nous lègue l'histoire. Or, ces matériaux ont été construits pour répondre à des besoins de connaissance et à des interrogations qui, par définition, n'ont rien à voir avec les questions que nous nous posons aujourd'hui.

C'est l'objet de la cliométrie de construire, sur cette base, les observations nouvelles dont nous avons besoin (Diebolt, 2016 ; Diebolt, 1997 ; Diebolt *et. al.*, 2017 ; Diebolt et Hauptert, 2019). La méthode en elle-même implique des limites qui sont celles de l'existence de traces transmises par les générations passées. Toutefois, le matériau est considérable et la question est davantage celle des moyens de les exploiter. En fait, la cliométrie s'appuie sur les méthodes de la comptabilité nationale rétrospective (élaboration de comptes satellites notamment). Elle vise à représenter l'économie d'un pays sous une forme simplifiée. D'apparence intangible, cette démarche ambitionne tout d'abord d'observer et de mesurer des faits socio-économiques. Elle permet, ensuite, de briser l'ensemble complexe des phénomènes dont se compose l'activité économique et sociale. Elle permet, enfin, de rendre les faits socio-économiques comparables, afin de les classer dans un nombre limité de catégories, de manière à pouvoir les étudier comme des éléments d'un ensemble homogène, c'est-à-dire de façon agrégée.

Ce faisant, production de savoir organisé, la cliométrie engendre son propre champ théorique, celui de la double adéquation de la mesure au concept qui la fonde et au réel dont elle doit rendre compte. Lorsque nous souhaitons observer une réalité historique dans sa dimension de longue période, il faut bien admettre que les instruments dont nous disposons, notre savoir-faire, sont considérablement en retard par rapport aux questions que nous nous posons. Que signifie, par exemple, la juxtaposition de mesures instantanées en des séries temporelles, quand l'objet observé lui-même évolue ? La comparaison des niveaux à deux dates quelque peu éloignées n'a certainement pas grande signification. En revanche, on doit certainement accorder plus de valeur au mouvement qui est ainsi décrit. Ce sont donc des tendances que nous mesurons, des cycles, des chocs.

A ce jour, cette démarche de recherche, appliquée à l'analyse comparée des dépenses d'éducation de la France, de l'Allemagne, du Royaume-Uni et de l'Espagne, nous a permis de formuler, à partir des ensembles statistiques reconstitués et à l'aide des méthodes traditionnelles de traitement statistiques et mathématiques des séries chronologiques, une nouvelle explication

théorique des relations dynamiques entre éducation et croissance économique aux XIX<sup>ème</sup> et XX<sup>ème</sup> siècles. Ce travail nous a conduits à poser l'hypothèse d'une transformation structurelle périodique du système social articulée sur celle du système économique. En fait, tout se passe comme si après avoir été un investissement d'accompagnement, formalisé par un modèle mathématique d'inspiration keynésienne (Diebolt, 1999a), l'éducation devenait, depuis la fin de la Seconde Guerre mondiale, au moins pour les pays les plus développés, l'une des causes, sinon le facteur déterminant de la croissance économique, synonyme d'un nouveau mode de régulation du système économique et social ou l'accumulation de capital physique demeure importante, mais deviendrait subsidiaire par rapport à l'accumulation de capital humain (Diebolt, 2000 ; Diebolt et Fontvieille, 2001).

Cet article s'inscrit dans cette dynamique. Il vise à étudier dans quelle mesure la centralisation du système éducatif français a profité à l'éducation de masse et a favorisé la croissance économique. Pour ce faire, nous utilisons les comptes nationaux rétrospectifs du financement de l'enseignement primaire alloué par l'État, les départements, les communes, les ménages, et autres administrations. Pour mener notre analyse, nous procédons en trois étapes. Un préalable à l'analyse du financement de l'enseignement primaire est le rappel de son fonctionnement et des principales évolutions intervenues au cours de la période analysée. Ensuite, nous analysons l'évolution des différents types de financement en utilisant la méthodologie des points atypiques pour détecter l'existence d'éventuelles ruptures dans les séries. Enfin, nous étudions les relations de cause à effet entre les différents types de financement et les taux de scolarisation des garçons et des filles. Ce faisant, nous testons également les liens de causalité avec la croissance économique.

Dans cette optique, notre ambition est de mieux comprendre quand et comment la scolarisation s'est développée en France, quelles en ont été les principales étapes et comment il convient de les expliquer.

### **3. Contexte et données**

Par arrêté du 15 mars 1876, une commission de statistiques de l'enseignement primaire est, pour la première fois, instituée au ministère de l'Instruction publique en France. Son objectif est de publier, tous les cinq ans, une statistique détaillée de l'enseignement primaire. Dès son origine, elle est présidée par Émile Levasseur et composée de hauts fonctionnaires du ministère de l'Instruction publique et de représentants du Conseil supérieur de statistique.

Avant la Première Guerre mondiale, huit volumes sont publiés. Le premier relatif à l'année scolaire 1876-1877 se situe explicitement en prolongement du recensement quinquennal de la population, afin de faciliter la comparaison entre le nombre des enfants à instruire et celui des enfants inscrits ou présents dans les écoles. Le second volume, de loin le plus important (avec plus de 700 pages), présente la *Statistique comparée de l'Enseignement Primaire de 1829 à 1877*. Il fait état de la quasi-totalité des statistiques disponibles se rapportant à cette période. Les autres volumes sont relatifs aux années scolaires 1881-1882, 1886-1887, 1891-1892, 1896-1897, 1901-1902, 1906-1907. Le sixième volume (1896-1897) contient un résumé des progrès de l'instruction primaire sous la Troisième République. Le dernier volume a paru à la fin de l'année 1909. Les documents destinés à figurer dans le neuvième volume (1911-1912) étaient rassemblés à la veille de la Grande Guerre, mais la publication ne put avoir lieu. À partir de cette date, la Commission de statistique ne fonctionne plus.

Après la guerre de 1914-1918, la Statistique générale de la France procède tous les ans au relevé des données les plus importantes relatives à l'instruction primaire, à partir des états originaux du ministère de l'Éducation nationale (nombre d'écoles, de classes, de maîtres et d'élèves par département, etc.). Inutile toutefois, en comparaison avec les anciennes publications quinquennales, de souligner l'insuffisance d'une telle documentation.

D'une manière générale, les statistiques concernant le système éducatif en France aux XIX<sup>ème</sup> et XX<sup>ème</sup> siècles peuvent être classées sous deux grandes rubriques. D'une part, il s'agit de données statistiques issues d'enquêtes ponctuelles réalisées à l'initiative d'un ministre, d'une commission spéciale ou d'un groupe de parlementaires pour examiner certains aspects du fonctionnement des établissements scolaires (Charmasson, 1986). Par exemple, elles comprennent des statistiques sur l'enseignement primaire, secondaire et supérieur. Elles comprennent également des informations recueillies dans le cadre de la gestion quotidienne des établissements d'enseignement ou de l'éducation dans son ensemble et dont l'objectif principal (et parfois unique) était administratif. Il s'agit principalement des données publiées dans les annuaires statistiques de la SGF (Statistique Générale de la France, base principale de la plupart des études globales sur l'évolution de la fréquentation scolaire en France).

L'information statistique utilisée dans cet article est issue, tout d'abord, du rapport du ministre de l'Éducation publique et des Beaux-Arts, Jules Ferry, au président de la République en date du 25 janvier 1880. Une commission permanente de statistique avait été installée par le ministre précédent afin de mieux connaître tout ce qui concernait l'éducation populaire en disposant des

données les plus fiables possibles. Rappelons à cet effet que les premiers relevés officiels concernant l'éducation ne remontent qu'à l'année 1829, que les statistiques étaient disparates et que les situations communales ou départementales pouvaient être fort différentes vis-à-vis à la fois de la situation réelle, existence ou non d'écoles de garçons et de filles, et du recensement statistique.

On peut ainsi citer une phrase du président de la Commission ayant réalisé le rapport de 1879 sur l'enseignement, É. Levasseur, « *Incorrectes sur beaucoup de points quand on les examine en détail, ces statistiques donnent par leur ensemble une idée juste des grands changements qui se sont produits, depuis un demi-siècle, dans nos écoles, elles montrent le progrès accompli dans la suite des années et laissent apercevoir les principales influences qui l'ont déterminé.* »

De ce fait, tout en sachant que les chiffres publiés sont les plus exacts possibles, des modifications ont été effectuées ponctuellement pour corriger des erreurs ou l'absence d'informations, mises en évidence par la confrontation de l'ensemble des statistiques. Ces réserves concernent tout particulièrement l'analyse détaillée au niveau d'une année ou d'un département donné et ce d'autant plus que la période concernée est perturbée par des variations du territoire français. Le département des Alpes-Maritimes est créé en 1860 avec le comté de Nice et une partie du département du Var. A la même époque sont créés les départements de Savoie et Haute-Savoie. En 1871, une partie des départements de la Meurthe et de la Moselle est annexée par l'Allemagne, la partie restée française devient la Meurthe-et-Moselle. La même année, le département du Bas-Rhin est annexé par l'Allemagne, de même qu'une partie du département du Haut-Rhin, la partie restée française prenant le nom de Territoire de Belfort.<sup>2</sup>

Dans son préambule au rapport de 1880, Jules Ferry met en exergue trois statistiques illustrant l'immense progression de l'éducation primaire au cours des quarante années précédentes. Ainsi, de 1837 à 1877, le nombre des élèves s'est accru de 82 %. Le nombre des écoles s'est accru de 36 %, de 75 % pour les seules écoles publiques ; les écoles de filles ont été multipliées par quatre.

---

<sup>2</sup> L'analyse de la croissance démographique doit être faite à territoire constant, et l'évolution des ratios de certains départements perd toute signification du fait de leur apparition, disparition ou transformation en cours de période. La statistique démographique utilisée est issue, elle aussi, du rapport Ferry de 1880 et peut comporter de légères différences avec les statistiques publiées par l'Institut National de la Statistique et des Etudes Economiques (INSEE) dans l'annuaire rétrospectif de 1966 par exemple. Cependant, que l'on compare la variation démographique à territoire constant de 1851 ou à territoire actuel, la croissance 1851-1876 est semblable : 5,5% dans le premier cas, 5,4% dans le second cas (38 437 592 habitants en 1876 contre 36 472 000 en 1851). Dans ce contexte de croissance démographique, la population des 5-15 ans reste relativement stable à l'intérieur des nouvelles frontières, perdant même quelques dixièmes de point entre 1856 et 1861.

Le nombre des maîtres s'est élevé de 85 % avec une amélioration du ratio, élèves/maître, qui passe de 53 à 48.<sup>3</sup>

Les données pour l'instruction primaire utilisées dans cet article proviennent également de [Carry \(1995, 1999\)](#) et de [Diebolt \(1995, 1999a, 1999b\)](#). Ces ensembles statistiques ont pour but de rassembler dans une série de tableaux cohérents, l'ensemble des flux financiers intervenant dans le domaine de l'éducation afin d'évaluer ce que coûte l'éducation à l'ensemble de la collectivité nationale, à étudier son financement et à déterminer les coûts des différents niveaux d'enseignement et des différentes activités. Le domaine de l'éducation est défini par une liste d'activités d'enseignement, d'activités annexes et de biens et services liés à l'éducation. Le compte de l'éducation retient deux catégories d'agents économiques intervenant dans le domaine. Les unités de financement qui supportent la dépense d'éducation et les unités de production qui réalisent les activités d'enseignement ou annexes à partir des ressources qui sont mises à leur disposition par les unités de financement. On peut également mentionner les comptes de l'État concernant l'éducation, tels qu'établis par [Fontvieille \(1976, 1982\)](#) à partir des registres financiers du Compte Général de l'Administration des Finances, et les recherches de [Grew et Harrigan \(1991\)](#). Enfin, il convient de mentionner les travaux fondateurs publiés par l'Institut National de Recherche Pédagogique ([Briand et al., 1987](#) ; [Luc, 1985](#)) et celles produites par la Direction de l'Évaluation et de la Prospective du Ministère de l'Éducation Nationale sous forme de comptes satellites ([DEP/INSEE, 1990](#)).

#### **4. Les évolutions de l'instruction primaire publique durant le XIX<sup>ème</sup> siècle**

##### **4.1. Les écoles et les effectifs**

Le nombre d'écoles progresse très rapidement sur la période 1837-1901 comme le montre notre Figure 1 présentant l'évolution du nombre des écoles mixtes, des écoles de garçons et des écoles de filles. Les nombreuses fondations d'écoles – de la loi Guizot à la loi d'avril 1867 – sont le fruit d'initiatives locales/municipales, qui permettent de réduire de manière substantielle le nombre de communes sans écoles. De ce fait, si en 1837 on comptait encore 5 667 communes n'ayant aucune école, elles n'étaient plus que 312 en 1876, et 53 en 1906.

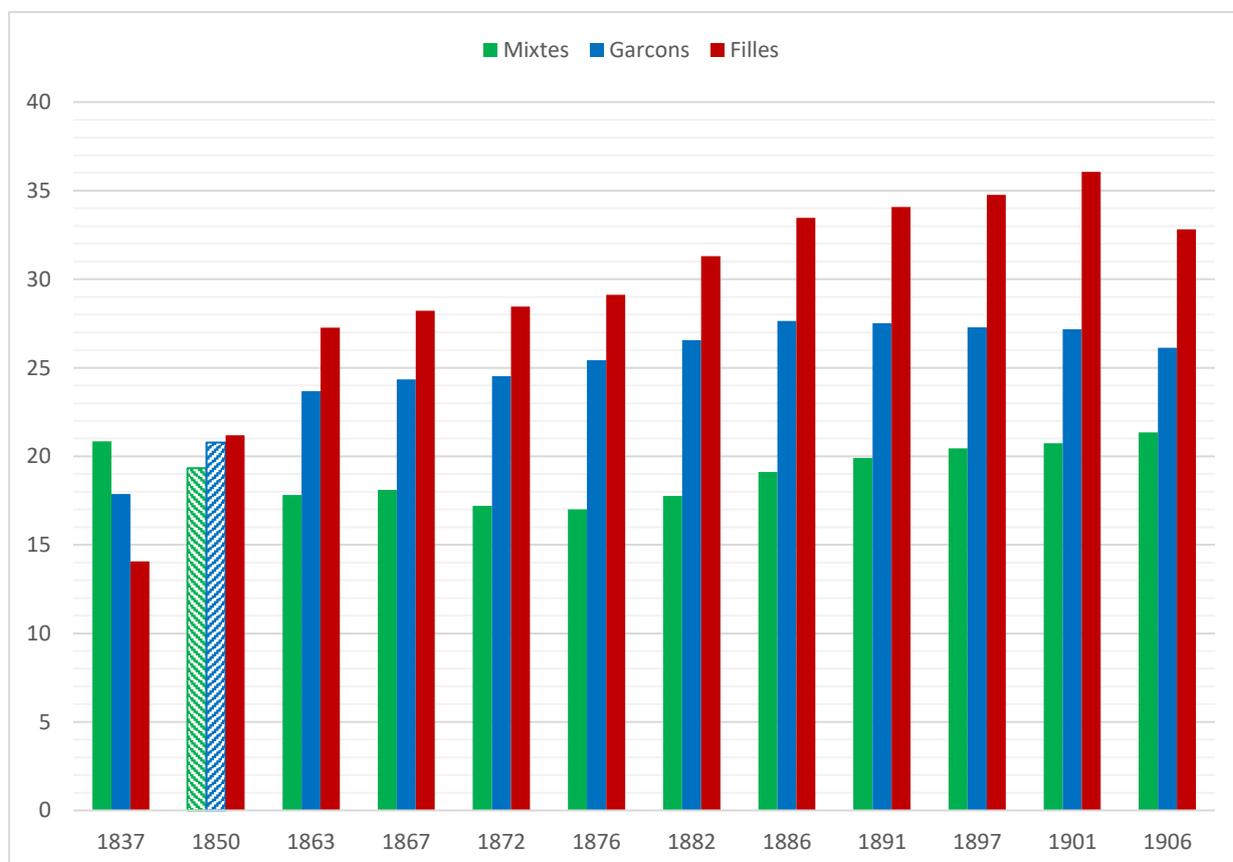
---

<sup>3</sup> La fin de l'intervention du ministre est constituée par une mise en garde sur les possibles doubles-comptes à certaines périodes ou les éventuelles erreurs statistiques pouvant entacher des comparaisons fines départementales ou annuelles.

L'évolution du nombre d'écoles publiques est encore plus spectaculaire en distinguant les écoles de filles des écoles de garçons. En effet, le nombre d'écoles publiques de filles est multiplié par plus de 4, passant de 5 453 en 1837 à 23 104 en 1906, soit une variation de plus de 300%. L'étude de l'évolution du nombre d'écoles par type sur la période 1830-1906, nous permet d'identifier trois mouvements significatifs : (i) une forte augmentation du nombre des écoles publiques; (ii) un accroissement du nombre des écoles de filles ; et (iii) une augmentation des établissements laïques (après 1876).

Cet accroissement du nombre d'écoles se traduit en moyenne nationale par une nette augmentation du ratio nombre d'écoles pour 1000 enfants de 5-15 ans, qui passe de 9,4 en 1850 à 10,7 en 1867 et 11,2 en 1876. Les filles bénéficient plus particulièrement de ces évolutions puisque leurs écoles qui représentaient 26% de l'ensemble en 1837 (contre 34% d'écoles de garçons, et 40% d'écoles mixtes) dépassent les 42% en 1901 (contre 33% d'écoles de garçons et 24% d'écoles mixtes).

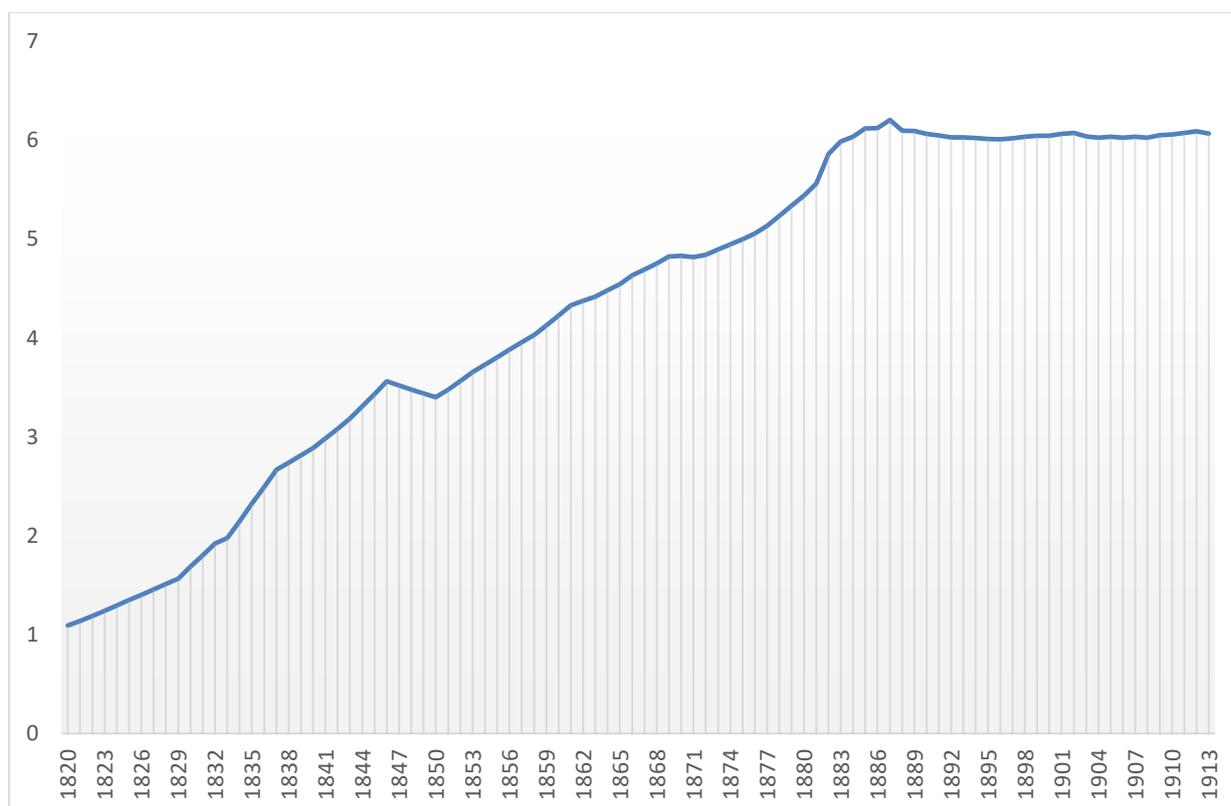
Figure 1 : Évolution du nombre d'écoles par catégorie, 1837-1906 (en milliers)



Dans un contexte de relative stabilité de la population potentielle de l'enseignement primaire, l'on assiste à une très forte augmentation du nombre d'inscrits dans les écoles (Flora, 1983). Ce phénomène est observé dans tous les pays européens. Au cours de la période 1850-1876, l'on constate une très forte augmentation des élèves inscrits en primaire. On compte, en 1850 en France, 3 321 423 inscrits, soit 51,5 % de la population des 5-15 ans, en 1867, 4 515 967 élèves étaient recensés (soit 68,6 %) et leur nombre s'élevait à 4 716 935 en 1876 soit 73,6 % de la classe d'âge.

La variation du nombre d'inscrits est encore plus impressionnante si l'on se réfère au recensement des élèves de 1837, premier recensement complet disponible, puisque l'on comptait à cette époque seulement 2 690 035 écoliers, et donc, un accroissement de 75 % en l'espace de quarante années (Figure 2). Cette forte augmentation de la scolarisation est vraisemblablement due à une très forte augmentation des moyens matériels et humains mis à la disposition de l'enseignement primaire.

**Figure 2 : Evolution du nombre d'enfants scolarisés, 1820-1913 (en millions)**



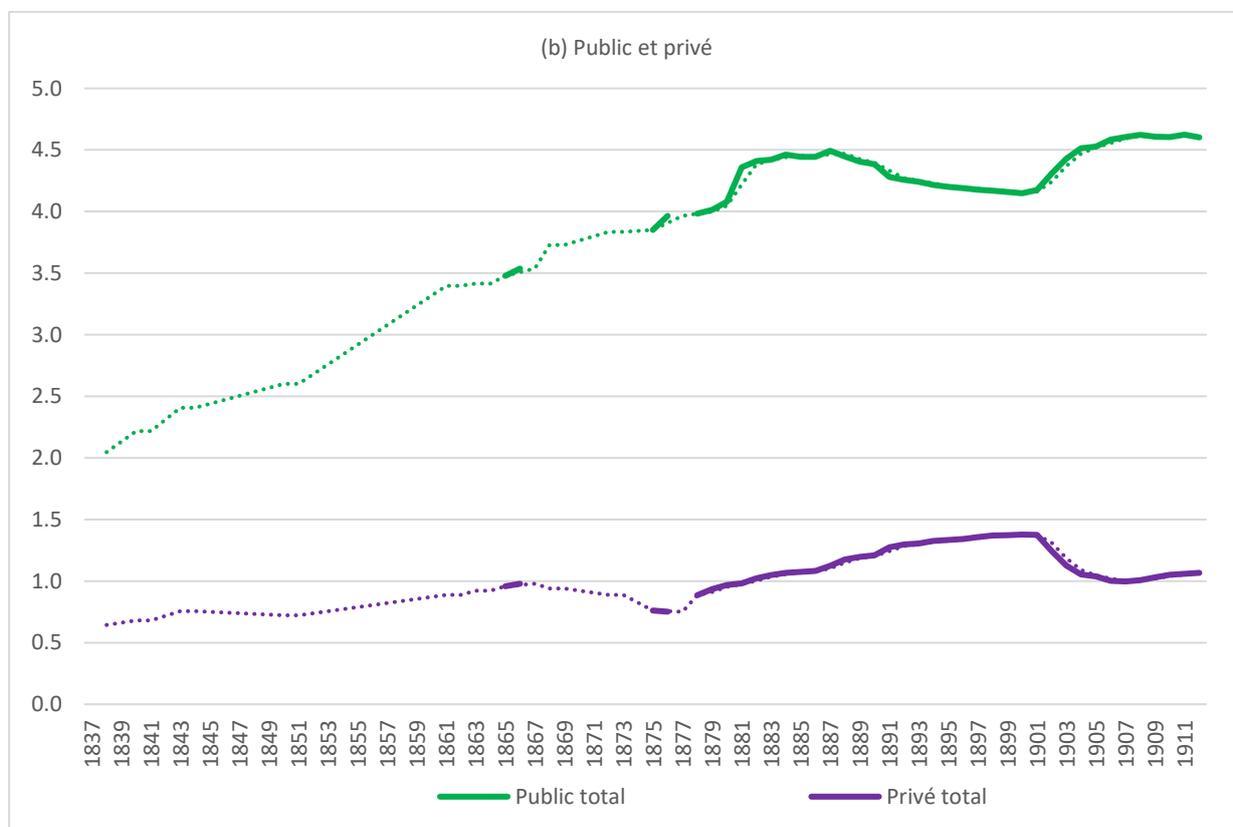
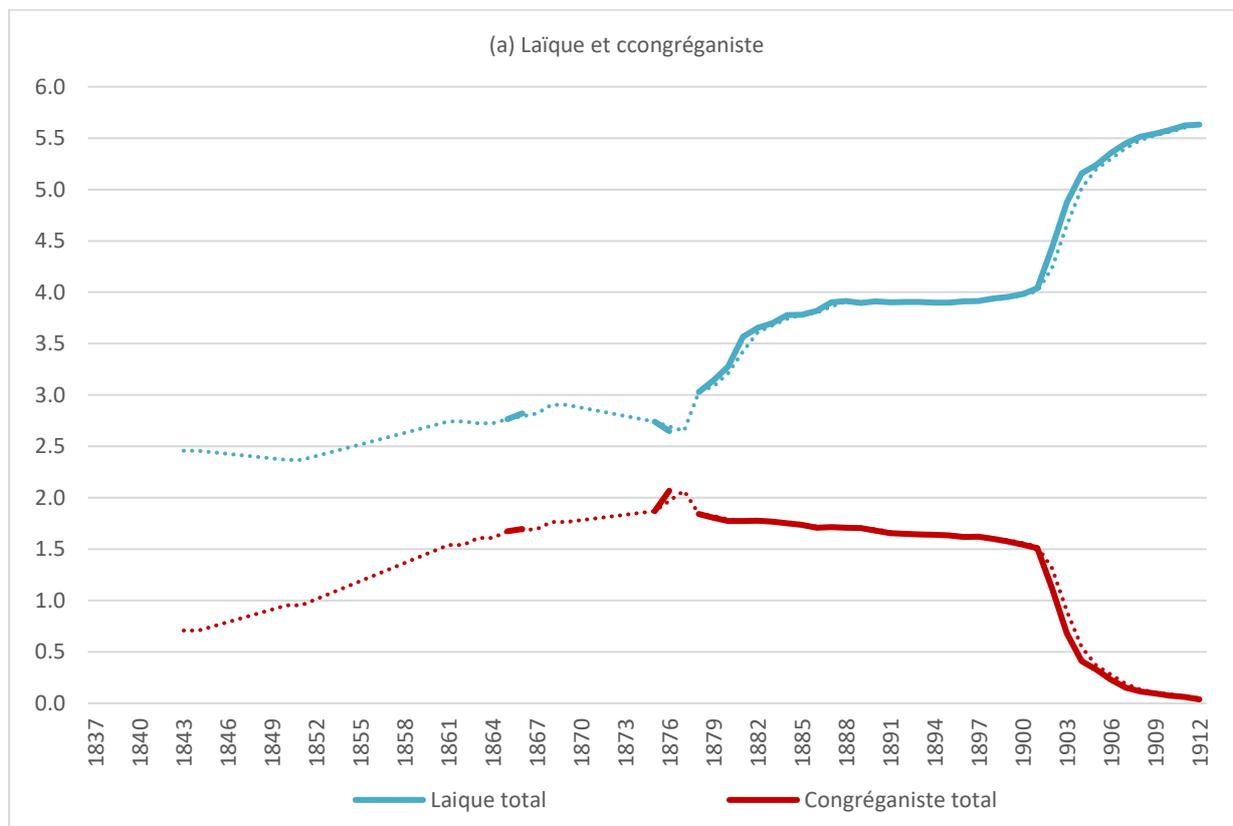
Dès 1876, presque autant de filles que de garçons étaient scolarisés, respectivement 2.316 millions et 2.4 millions. Il est cependant nécessaire d'introduire une certaine réserve au sujet de la scolarisation des filles en 1837 en raison de la possible sous-estimation des taux de scolarisation dans les écoles mixtes. Cette précaution ne nous empêche pas néanmoins de pouvoir conclure que l'évolution des taux de scolarisation a été significative pour les garçons mais aussi pour les filles.

La Figure 3 présente l'évolution du nombre d'enfants scolarisés par type d'établissement : (a) laïque ou congréganiste ; (b) public ou privé. Le graphique reflète un accroissement quasi constant des scolarisés laïques et publics. Une lente diminution des effectifs congréganistes est observée à partir de l'année 1876. S'en suit une chute brutale du nombre d'enfants scolarisés dans ce type d'établissement à partir des lois de 1901 et 1904 au profit des écoles laïques. La période étudiée se caractérise donc par une augmentation quasi constante de la scolarisation des filles et des garçons dans les écoles publiques laïques, avec un rattrapage très clair des filles sur les garçons. La laïcisation quasi complète des écoles publiques de garçons s'effectue dans les années 1890, alors qu'il faudra attendre les lois de 1901 et 1904 pour observer la laïcisation complète des écoles publiques de filles. Les lois Ferry semblent, en revanche, n'avoir eu que très peu d'effet sur les taux de scolarisation.

Les moyens en enseignants progressent encore plus rapidement que le nombre d'écoles. L'accroissement des moyens mis à la disposition de l'instruction primaire a également permis de répondre aux besoins d'enseignement sans augmenter considérablement les effectifs moyens par école, de 55 enfants par établissement en 1850, on passe à 64 en 1867 et à 66 en 1876. Cette constatation est naturellement liée au fait que la spectaculaire variation des effectifs tient autant à la croissance des effectifs dans les écoles existantes qu'à la création de nouvelles écoles.

On constate une très nette féminisation du corps enseignant (36% d'institutrices en 1840, 54% en 1863). Ces mouvements sont naturellement induits par l'évolution de la création d'établissements d'enseignement primaire, forte croissance du nombre d'écoles et plus particulièrement d'écoles de filles (7026 sur les 10093 écoles créées de 1850 à 1867) et d'écoles congréganistes. La féminisation du corps enseignant dépasse les 53% en 1876 avec cependant des pourcentages très différents selon que l'on considère l'enseignement laïque (34% de femmes) ou l'enseignement congréganiste (79,7%). Les hommes qui représentaient 64 % du personnel en 1840 deviennent minoritaires en 1876 en ne représentant plus que 47 % du personnel enseignant. L'essentiel de l'accroissement du nombre des enseignants, près de 80 %, est le fait d'un recrutement féminin. Alors que l'on comptait 0,8 enseignant titulaire par école vers 1840, cette moyenne s'élevait à 1,5 en 1876.

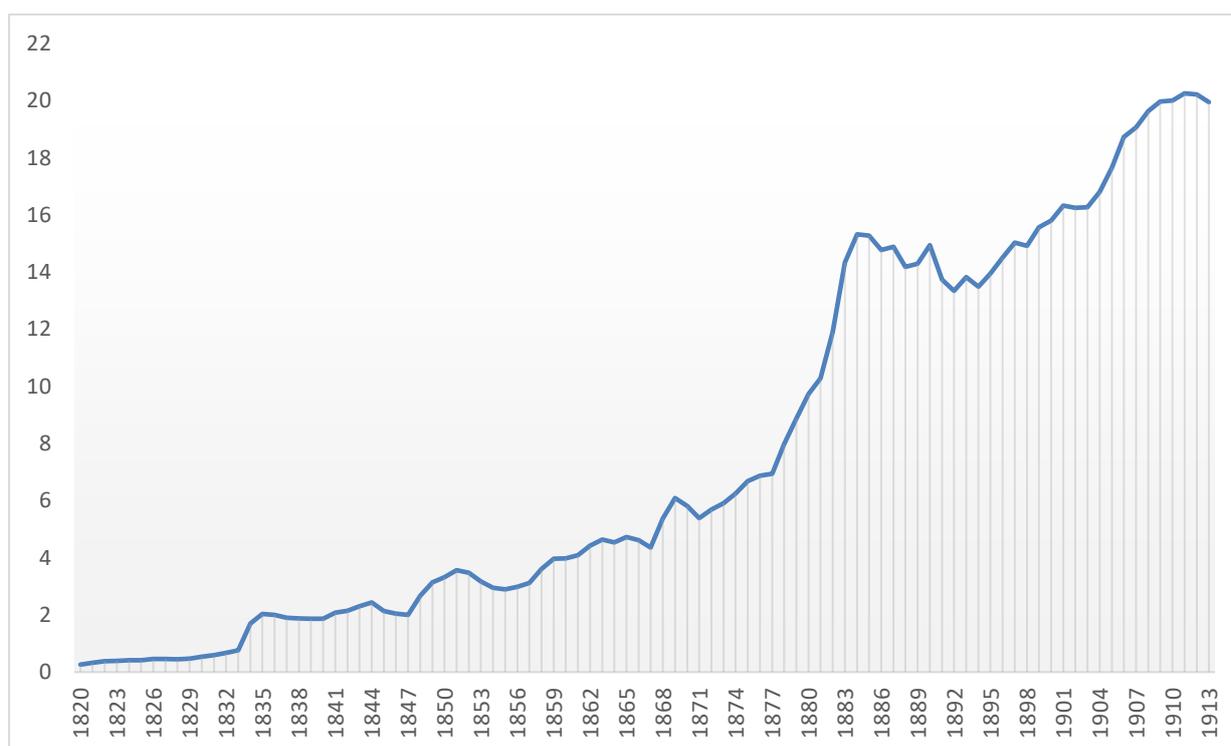
Figure 3 : Enfants scolarisés par type d'école, 1837-1913 (%)



## 4.2. Le financement de l'enseignement primaire

L'analyse du financement de l'enseignement primaire nécessite au préalable un rappel de son fonctionnement et des principales transformations intervenues sur la période analysée. Les coûts de fonctionnement sont essentiellement représentés par la fourniture par la commune d'un logement convenable à l'instituteur (lois du 28 juin 1833 et du 15 mars 1850) et le traitement de cet instituteur théoriquement à la charge de la Nation (décrets du 28 octobre 1793 et 19 novembre 1793). Néanmoins, la faiblesse des ressources rendait quasiment impossible l'application de ces décrets. De fait, les instituteurs devaient se contenter du logement fourni par la commune et d'une rétribution payée par les parents.

Figure 4 : Financement de l'enseignement primaire, 1820-1913 (millions de francs constants)

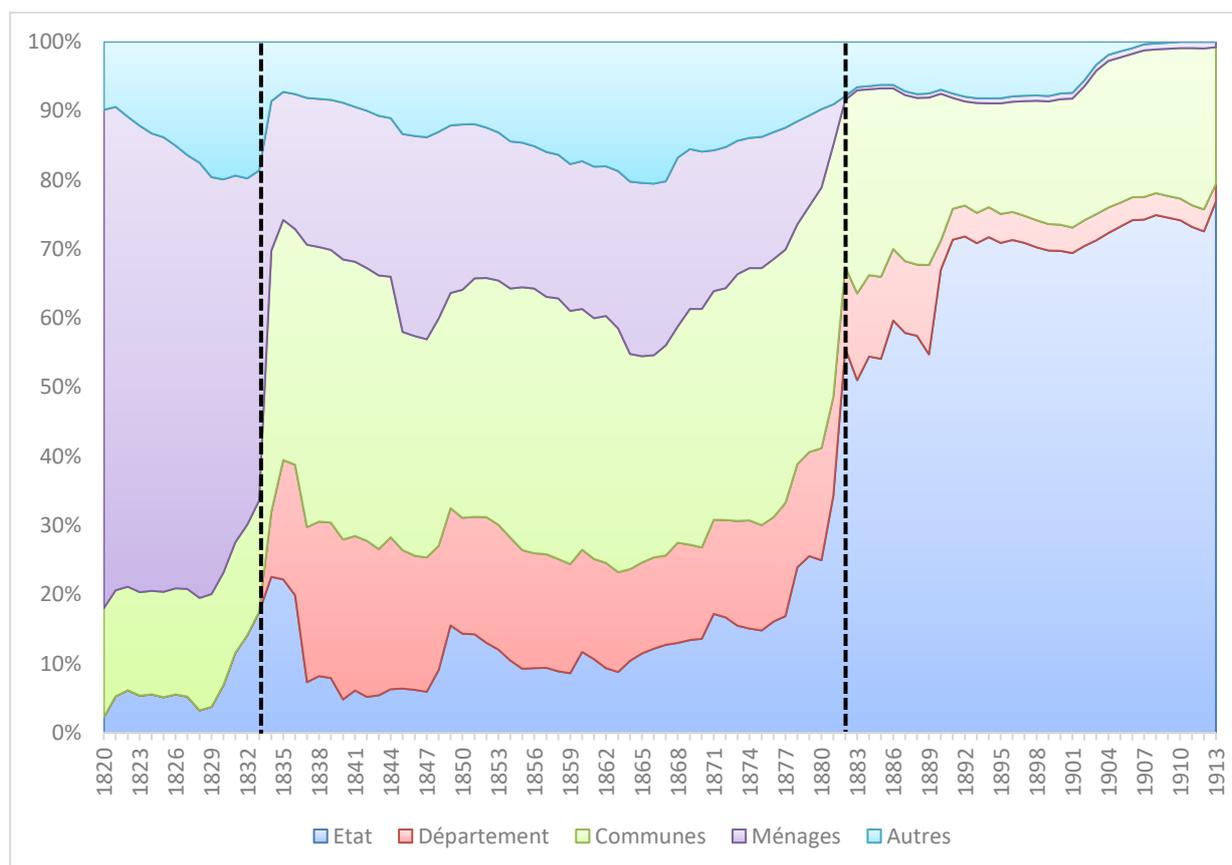


La loi du 28 juin 1833 apporte une certaine réglementation en déterminant un minimum de traitement fixe des instituteurs de 200 francs et une procédure de financement. Ainsi les communes ne disposant pas de ressources ordinaires suffisantes devaient voter trois centimes additionnels s'ajoutant aux trois contributions locales (foncière, personnelle et mobilière). Les départements devaient également pallier l'éventuelle insuffisance des ressources ordinaires communales en votant

deux centimes additionnels sur les mêmes taxes. Si les ressources communales et départementales étaient insuffisantes, l'État devait concourir au paiement du traitement fixe des instituteurs.

De fait, en 1850, l'essentiel de la charge pesait directement sur les familles. Ainsi, sur 12 307 076 francs de coûts, les familles payaient 8 745 408, les communes et départements 2 765 668 et l'État n'intervenait que pour 796 000 francs. De par ce financement direct ou indirect des parents, les plus forts taux de fréquentation des écoles se retrouvent dans les zones de développement économique rapide et d'une certaine richesse, à la fois par nécessité de disposer d'une main-d'œuvre ayant quelque instruction et par l'obligation de devoir financer les coûts de cette instruction.

**Figure 5 : Structure du financement de l'enseignement primaire, 1820-1913 (%)**



La priorité donnée à l'enseignement primaire va transformer ce financement à la fois en volume (Figure 4) et en structure (Figure 5). La croissance très importante des moyens mis à la disposition de l'enseignement primaire s'est naturellement accompagnée d'une très forte

augmentation des coûts de fonctionnement, multipliés par 2,7 en l'espace de vingt ans, soit un taux moyen annuel d'un peu moins de 5 % (Tableau 1). Les coûts de fonctionnement de l'enseignement primaire étaient financés principalement par les communes et les ménages, mais également par l'État et les départements et pour un très faible montant par des dons et legs. Ces financements ont largement évolué sur la période 1820-1913 comme le montre la Figure 5.

Plusieurs chocs peuvent être observés dans les données. A partir de 1877, on constate une très forte et rapide augmentation des financements et de la participation l'État. L'accroissement des investissements publics se produit alors dans la période qui précède les lois Ferry. L'année 1880 marque une chute brutale et drastique du financement des ménages ; l'année 1889 marque une chute similaire du financement des départements et des communes. L'année 1894 quant à elle se caractérise à l'inverse par une augmentation forte et continue du financement par les communes.

**Tableau 1 : Coût de fonctionnement annuel par élève (en francs)**

|                     | <b>1856</b> | <b>1867</b> | <b>1876</b> |
|---------------------|-------------|-------------|-------------|
| <b>État</b>         | 0,80        | 0,79        | 2,23        |
| <b>Département</b>  | 0,83        | 0,64        | 1,45        |
| <b>Commune</b>      | 2,79        | 2,73        | 6,57        |
| <b>Ménages</b>      | 2,50        | 3,27        | 4,00        |
| <b>Legs et dons</b> | 0,05        | 0,08        | 0,20        |
| <b>Total</b>        | 6,98        | 7,51        | 14,45       |

En francs par écolier inscrit, les variations 1856-1876 sont très importantes avec un doublement des coûts, ceux-ci passant globalement de 6,98 francs à 14,45 francs (Tableau 1). L'accélération se fait presque uniquement sur la période 1867-1876. L'année 1867 se situe à 7,51 francs, soit à un niveau peu différent de celui connu en 1856. Les lois votées à partir de l'année 1867 ont modifié, très largement, le financement par inscrit comme le montre le Tableau 1.

## 5. Méthodologie

En référence à ces repères statistiques, institutionnels et juridiques, notre démarche vise, à partir de la démarche méthodologique développée par [Darné et Diebolt \(2004\)](#), à montrer que des événements rares, des chocs, peuvent engendrer des effets variés sur les séries temporelles des dépenses pour l'instruction primaire en France. Cette approche sera complétée par une analyse des causalités.

### 5.1. Points atypiques

De manière générale, lorsqu'en histoire économique l'on s'intéresse à l'analyse des chocs, deux méthodologies économétriques sont envisageables. Soit, à l'instar de la démarche traditionnelle, étudier les chocs sous la forme de fonctions de réponses impulsionnelles. Dans ce cas, l'analyse est fondée sur l'estimation d'un modèle VAR, et s'inscrit essentiellement dans une démarche analytique et prévisionniste puisque les chocs envisagés sont simulés et donc fictifs. Soit, à l'instar des travaux d'économétrie historique les plus actuels, analyser les chocs sous la forme de points atypiques ou d'*outliers*. Dans ce cas, l'analyse des chocs s'inscrit dans une démarche analytique et historique, les chocs étant effectifs. Notre démarche de recherche s'inscrit dans cette seconde voie de recherche. En d'autres termes, nous avons recours ici à la méthode des points atypiques.

En théorie statistique, lorsqu'une observation s'écarte fortement de sa valeur moyenne ou de sa tendance, elle est considérée comme exceptionnelle. Elle se définit donc par une valeur particulière, non représentative, et leur nombre n'excède généralement pas 1 % de la longueur de la chronique. Cependant, la définition de ces valeurs fondées uniquement sur leur taille et leur rareté n'est pas opérationnelle. Elle est trop vague et nécessite de fixer, au préalable, des seuils de taille et de fréquence à partir desquels une valeur peut être alors qualifiée d'exceptionnelle.

Dans le cadre de cet article, après avoir spécifié l'échelle de mesure et la période de référence, nous considérons qu'une observation revêt un caractère exceptionnel lorsque sa valeur (positive ou négative) est très élevée et lorsque sa fréquence est petite. Bien que subjective d'un point de vue littéral, cette définition nous permet de classer ces valeurs en deux catégories : événements rares ou extrêmes. Sur un plan purement statistique, un événement extrême représente une valeur de deux à trois fois l'écart-type de la série alors qu'une valeur atypique est définie par une valeur nettement supérieure à trois fois l'écart-type de la série. Quant aux observations rares ou atypiques, autres qu'extrêmes, elles se définissent comme telles relativement à une distribution ou à un modèle. Supposons une distribution gaussienne : une valeur aberrante est, dans ce cas, détectée par la distance qui la sépare du reste des données. Les observations qui ne sont pas compatibles avec l'hypothèse de normalité de la distribution peuvent être assimilées à des valeurs atypiques. Le

processus inhérent aux données peut également exclure certaines valeurs considérées alors comme exceptionnelles.

Les évènements rares se décomposent en deux groupes. Ces observations singulières peuvent être liées aux erreurs dans l'enregistrement ou dans la saisie des données. Elles peuvent également être associées à des évènements climatiques, naturels, politiques, économiques ou financiers. En ce qui nous concerne, le terme d'évènements rares sera toujours associé à celui de valeurs atypiques, aberrantes ou exceptionnelles, les valeurs extrêmes ne faisant pas l'objet de nos investigations cliométriques. L'analyse des points atypiques permet ainsi d'associer des observations atypiques à des évènements économiques, politiques, financiers, etc. Les faits économiques, financiers et/ou politiques qui affectent les observations n'adoptent pas tous un comportement identique. Ces évènements peuvent avoir des effets variés sur les séries temporelles : ponctuel, rémanent ou durable dans le temps.

Box et Tiao (1975) sont les premiers à s'interroger sur les possibles dynamiques de ces évènements. Pour cela, ils s'intéressent à la réduction du niveau de pollution dans le centre-ville de Los Angeles suite à l'ouverture d'une voie de contournement et à la mise en application de la loi 1963. Ces deux mesures, selon les autorités, doivent avoir pour effet une diminution plus ou moins immédiate du degré de pollution. Les auteurs s'interrogent alors sur l'impact de cette décision. Ils analysent différentes structures dynamiques susceptibles d'expliquer le comportement du niveau de pollution suite aux précédentes mesures.

Les auteurs retiennent deux types de comportements principaux, qu'ils associent à des dynamiques particulières, à savoir les *Additive outliers* (AO) et les *Innovative outliers* (IO) pour décrire l'évolution du niveau de pollution. Ces deux comportements renvoient à l'étude de Fox (1972). En effet, Fox est le premier à proposer une définition et une classification des *outliers* reposant sur une modélisation autorégressive de la série. Il met ainsi en évidence l'existence de deux types de points : le type I, qualifié par la suite d'*Additive outlier*, et le type II, qualifié par la suite d'*Innovative outlier*. Ces deux types de points sont les plus étudiés dans la littérature.

Tsay (1988) définit d'autres types d'*outliers* dont les *Level shifts* (LS) et les *Temporary changes* (TC). Ces types de points ne sont pas exactement considérés comme des valeurs aberrantes, mais plutôt comme des changements structurels. Un *Level shift* modifie le niveau de la série. Il peut être considéré comme une succession d'*Additive outliers* de même amplitude, dont l'effet est permanent. Notons que les *Level shifts* perturbent très fortement l'estimation des modèles et donc les résidus estimés. La modélisation de ces quatre valeurs atypiques (AO, IO, LS et TC) est initialement présentée sous forme ARMA, puis étendue aux modèles ARIMA. Ces quatre types de points sont aujourd'hui les plus utilisés dans la littérature.

Dans un premier temps, nous testons l'existence de points de rupture dans les séries temporelles des dépenses de l'instruction primaire, telles que financées par l'État, les départements, les communes, les ménages, et autres administrations publiques<sup>4</sup>. Dans un second temps, nous déterminons les points de rupture pour les séries relatives à la scolarisation. Enfin, nous nous intéressons aux relations de causalité liant financement de l'instruction primaire et scolarisation d'une part, puis scolarisation et croissance économique ensuite.

## 5.2. Causalité

L'analyse de la causalité à la Granger s'inscrit dans le cadre des modèles VAR (*vector autoregressive*) non structurels, introduit dans la recherche historique par [Eckstein et al. \(1984\)](#)<sup>5</sup>. L'un des avantages de ces modèles sur les modèles traditionnels est qu'ils permettent une meilleure analyse dynamique des systèmes, en prenant d'une part en considération la structure intrinsèque des séries<sup>6</sup> et d'autre part, en tenant compte d'éventuels effets dynamiques entre les variables. De plus, dans un modèle VAR, toutes les variables sont à la fois exogènes et endogènes ce qui offre la possibilité d'envisager toutes les relations possibles entre les variables sans *a priori* sur leur éventuelle endogénéité.<sup>7</sup> Malgré leur opposition historique, les modèles structurels et non-structurels sont liés mathématiquement si bien que l'on peut aisément passer de l'un à l'autre ([Monfort et Rabemanajara, 1990](#); [Hendry et Mizon, 1993](#)).

Dans les modèles non structurels, chaque équation décrit l'évolution d'une variable en fonction de ses valeurs passées et des valeurs passées des autres variables du système :

$$X_t = \sum_{i=1}^p A_i X_{t-1} + \varepsilon_t \quad \Leftrightarrow \quad \Phi(L)X_t = \varepsilon_t, \quad (1)$$

---

4 Les administrations publiques sont classées en trois catégories : les administrations publiques centrales (Etat, organismes divers d'administrations centrales contrôlés par l'Etat), les administrations publiques locales (collectivités territoriales, groupements de communes à fiscalité propre et les « organismes divers d'administration locale ») et les autres administrations publiques. C'est dans ce dernier groupe que l'on trouve de nos jours les organismes de sécurité sociale ; l'assurance chômage...

5 "The methodology of vector autoregression appears useful for studying historical series on climatic, economic and demographic variables where we do not yet have a sufficient theoretical foundation for specifying and estimating structural models", p. 295.

6 La structure intrinsèque de la série fait référence à son identification au sein de la classification ARIMA ([Box and Jenkins, 1976](#)).

7 Les modèles VAR non structurels sont parfois critiqués sur deux points : tout d'abord à cause de la possibilité d'inclure au modèle un grand nombre de variables qui pose le problème des degrés de liberté qui s'évanouissent lors de l'estimation ; ensuite on leur reproche souvent le peu de théorie auxquels ils font référence et qui leur vaut souvent le qualificatif de modèles a-théoriques.

Où  $X_t$  est un vecteur ( $n \times 1$ ) de variables,  $A_i$  est une matrice ( $n \times n$ ),  $p$  est le retard optimal du modèle et  $\varepsilon_t$  est un vecteur ( $n \times 1$ ) des innovations /résidus.  $E(\varepsilon_t \varepsilon_t') = \Sigma$  est une matrice symétrique positive ( $n \times n$ ).  $\Phi(L)$  est un polynôme d'opérateur retard tel que  $L^k X_k = X_{t-k}$ .

L'utilisation de ces modèles nécessite au préalable de tester un certain nombre d'hypothèses (Tableau B1, Annexe B). Tout d'abord, il est nécessaire de travailler avec des séries stationnaires<sup>8</sup>. Pour cela, nous utilisons les tests de racines unitaires de Elliott, Rothenberg et Stock (1996), plus robustes (Salanié, 1999) que les tests classiques de Dickey-Fuller (1979). Une fois que les variables sont stationnaires, il convient de déterminer le nombre optimal de retards<sup>9</sup> (Tableau B2, Annexe B). Ensuite, il faut tester la présence de cointégration Engle et Granger, 1987; 1991) entre les variables<sup>10</sup> et si nécessaire de la corriger (Modèle Vectoriel à Correction d'Erreur) afin d'éviter les problèmes de régressions fallacieuses (Granger and Newbold, 1976)<sup>11</sup>. Pour cela, nous utilisons le test de Johansen (1988). Les résultats de nos tests (Tableau B3, Annexe B) montrent l'absence de cointégration quel que soit le modèle étudié.

A partir de là, l'analyse peut prendre deux directions non exclusives : l'étude des relations de causalité (relation de court terme) et celle de la dynamique du modèle. Il existe deux approches de la causalité Granger, 1969; Sims, 1980). Bien que celles-ci soient équivalentes (Bruneau, 1996),<sup>12</sup> nous optons ici pour un test de Granger (1969).<sup>13</sup>

Afin d'étudier la direction et le signe de la causalité, nous analysons comment réagit une variable lorsqu'un changement se produit sur la seconde. Si l'on considère un modèle à deux variables, l'équation (1) s'écrit :

---

8 Une variable  $X_t$  est dite stationnaire si tous ses moments sont invariants avec l'origine du temps. Il existe deux types de processus non-stationnaires : les processus TS (Trend Stationary Processes) qui présentent une non-stationnarité de type déterministe et les processus DS (Difference Stationary Processes) pour lesquels la non stationnarité est de type aléatoire. Ces processus sont respectivement stationnarisés par un écart à la tendance déterministe et par un filtre aux différences. Dans ce dernier cas, le nombre de filtres appliqués donne l'ordre d'intégration des variables. Ainsi, une variable est intégrée d'ordre  $d$  s'il est nécessaire de la différencier  $d$  fois pour la rendre stationnaire.

9 Il existe différents critères pour déterminer le retard optimal d'un modèle VAR (Akaike, 1974; Schwartz, 1978; Hannan-Quinn, 1979), basés sur la maximisation du logarithme de la fonction de log-vraisemblance. Ici, le retard retenu est celui qui satisfait le plus grand nombre de ces critères.

10 On dit que deux variables sont cointégrées si elles présentent une relation de long terme stable autrement dit si elles partagent une tendance commune. Une condition nécessaire à la cointégration est que les variables soient intégrées de même ordre.

11 Si deux variables sont cointégrées, la bonne qualité statistique du modèle provient de la présence d'une tendance commune stable dans le long terme. Il est donc nécessaire de corriger cette relation avant l'analyse causale et dynamique.

12 La causalité au sens de Granger (1969) concerne la propagation d'impulsions déterministes interprétables comme des modifications liées à des changements structurels. Sims (1980), au contraire, fonde son analyse sur la propagation d'impulsions stochastiques représentatives de « surprises ».

13 La définition de la causalité est donnée par Granger (1969) : « Une variable  $X$  cause une variable  $Y$  si la prédictibilité de  $Y$  est améliorée lorsqu'on introduit de l'information relative à  $X$  dans l'analyse.

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \dots + \begin{bmatrix} A_p & B_p \\ C_p & D_p \end{bmatrix} \begin{bmatrix} X_{t-p} \\ Y_{t-p} \end{bmatrix} + [\varepsilon_t] \quad (2)$$

Ainsi, pour une relation de causalité allant de X vers Y, le signe de la relation est déterminé par le signe du ratio :

$$\sigma_{X \rightarrow Y} = \frac{\sum_{i=1}^p C_i}{1 - \sum_{i=1}^p D_i}, \quad (3)$$

Le test de causalité pourrait être interprété comme le fait d'évaluer si le décalage d'une autre variable apporte, ou non, une contribution incrémentielle nette et significative au mouvement d'une variable dépendante, une fois que la propre corrélation de la variable dépendante est prise en compte.

L'analyse dynamique qui consiste à étudier les effets des variations d'une variable sur les variations de l'autre variable. Pour cela, on considère les innovations/résidus comme des variables exogènes. La simulation de chocs sur ces dernières nous permet de comprendre comment (fonction de réponse impulsionnelle) et dans quelle proportion (décomposition de la variance) les autres variables du système sont impactées. En d'autres termes, on observe comment des simulations de chocs sur les innovations de la variable X, vont affecter la variable Y.

Nous testons ainsi dans un premier temps les relations de causalité à la Granger entre les différentes séries des dépenses de l'instruction primaire et la scolarisation. Puis, dans un second temps nous étudions les relations de causalité entre : (i) le PIB par habitant et les dépenses ; et (ii) le PIB par habitant et la scolarisation, de manière à pouvoir évaluer si les effets que nous observons découlent d'une réelle volonté d'investir dans l'éducation indépendamment des richesses créées ou si les effets observés découlent directement de l'accroissement du PIB par habitant au cours de la période observée.

## 6. Dépenses de l'instruction, lois, et généralisation de la scolarisation

### 6.1. Points atypiques

La détection des *outliers* dans les diverses dépenses d'éducation pour le niveau primaire met en évidence des points de rupture dans la structure des dépenses. L'analyse confirme plusieurs constats graphiques discutés dans notre section 4. L'effet de la loi Guizot de 1833 ayant marqué positivement et de manière permanente le financement de l'instruction primaire est ainsi confirmé.

La loi Duruy de 1867 a permis la généralisation de la scolarisation en France et marque un commencement d'organisation de l'enseignement des filles notamment en améliorant la situation des institutrices mais aussi en contribuant à l'augmentation du nombre des écoles spéciales de filles. La mise en application de la loi Duruy s'accompagne d'un accroissement (temporaire) du financement de l'instruction primaire par les ménages.

L'analyse souligne l'importance des lois Ferry (1881 et 1882) dans l'évolution du financement du primaire en France : l'État remplace les ménages et les communes dans le financement de l'éducation primaire (Tableau 2). La loi Goblet de 1886 parachève les lois Ferry en confiant à un personnel exclusivement laïque l'enseignement dans les écoles publiques, remplaçant ainsi les instituteurs et institutrices congréganistes. La mise en application de cette loi engendre un effet négatif sur les dépenses des ménages qui contribuaient au financement de ce type d'instruction.

La loi de 1889 porte quant à elle directement sur les questions de dépenses ordinaires de l'instruction primaire publique et les traitements du personnel de ce service. Les dépenses ordinaires de l'enseignement primaire public sont à la charge de l'État, des départements et des communes, selon les règles édictées par la présente loi. La majorité des dépenses sont avec cette loi à la charge de l'État. Les traitements des personnels sont financés par l'État. On observe alors un transfert des efforts de dépenses des communes vers l'État avec comme objectif de réduire les inégalités au niveau national.

En terme de magnitude, les effets les plus impactants restent ceux causés par la mise en place des lois Ferry, avec le passage d'un financement de l'éducation primaire principalement supporté par les communes à un financement public et national de l'éducation.

**Tableau 2 : Points de rupture dans le financement de l'instruction primaire**

| Année                   | Valeur | Type | Effet                  | Evènement            |
|-------------------------|--------|------|------------------------|----------------------|
| <b>Dépenses totales</b> |        |      |                        |                      |
| 1834                    | 0.68   | LS   | Positif permanent (+)  | Loi Guizot 1833      |
| 1848                    | 0.22   | LS   | Positif permanent (+)  | Révolution de 1848   |
| <b>État</b>             |        |      |                        |                      |
| 1882                    | 2.37   | AO   | Positif ponctuel (+)   | Lois Ferry 1881-1882 |
| <b>Communes</b>         |        |      |                        |                      |
| 1882                    | -1.52  | AO   | Négatif ponctuel (-)   | Lois Ferry 1881-1882 |
| <b>Départements</b>     |        |      |                        |                      |
| 1883                    | 0.28   | TC   | Positif temporaire (+) | Lois Ferry 1881-1882 |
| 1886                    | -0.28  | LS   | Négatif permanent (-)  | Loi Goblet 1886      |
| 1888                    | -0.23  | AO   | Négatif ponctuel (-)   | Loi Goblet 1886      |
| 1890                    | -1.30  | LS   | Négatif permanent (-)  | Loi de 1889          |
| <b>Ménages</b>          |        |      |                        |                      |
| 1868                    | 0.20   | TC   | Positif temporaire (+) | Loi Duruy 1867       |
| 1871                    | -0.10  | AO   | Négatif ponctuel (-)   | Guerre de 1870       |
| 1881                    | -0.49  | LS   | Négatif permanent (-)  | Loi Ferry 1881       |
| 1882                    | -0.56  | LS   | Négatif permanent (-)  | Loi Ferry 1882       |

**Tableau 3 : Points de rupture dans la scolarisation primaire**

| Année                  | Valeur | Type | Effet                 | Evènement            |
|------------------------|--------|------|-----------------------|----------------------|
| <b>Nombre d'élèves</b> |        |      |                       |                      |
| 1882                   | 0.15   | LS   | Positif permanent (+) | Lois Ferry 1881-1882 |
| 1887                   | 0.12   | AO   | Positif ponctuel (+)  | Loi Goblet de 1886   |

*Note* : la série du nombre d'illettrés ne présente pas de points de rupture

Tout comme la structure de financement, du point de vue de la scolarisation, il semble que les lois Ferry aient joué un rôle prépondérant. En revanche, le PIB par habitant ne révèle aucun point de rupture sur la période étudiée.

## 6.2. Causalité

Au-delà d'une modification de la structure de financement, les diverses dépenses ont également impacté les taux de scolarisation (Tableau 4). Ainsi, une augmentation des dépenses des communes ou de l'État favorise le développement de la scolarisation. Lorsqu'un changement opère, le délai de réaction du nombre d'élèves scolarisés dépend du type de dépense. Si une modification des dépenses de l'État entraîne une réaction immédiate en termes de scolarisation, le délai est plus long (3 ans) pour les dépenses décentralisées. *A contrario*, une charge plus importante pour les ménages a tendance à freiner la scolarisation. La réaction suite à une modification des dépenses des ménages se situe dans un délai intermédiaire de 2 ans.

On note également une boucle rétroactive entre les dépenses à la charge du département et la scolarisation : une augmentation des dépenses départementales entraîne une augmentation du nombre d'élèves qui entraîne à son tour une augmentation des dépenses départementales. Dans les 2 cas, les réactions se font dans un délai de 3 ans. En revanche, les changements dans le financement de l'éducation sont sans effet sur l'illettrisme.

L'absence de relation de long terme stable mise en évidence par les tests de cointégration entre la scolarisation, le financement de l'éducation et la croissance économique-, souligne l'importance du contexte historique et économique dans l'évolution des séries, ce dernier générant les événements conditionnant l'évolution de la croissance économique (Darné et Diebolt, 2006).

L'analyse des relations de causalité entre PIB par habitant, divers types de financements de l'instruction primaire, et effectifs scolarisés ne révèle, à notre grande surprise, aucune relation de causalité. Il n'existe aucune relation de causalité entre PIB par habitant et nombre d'élèves inscrits dans l'éducation primaire. L'analyse ne montre aucune relation entre PIB par habitants et divers types de financement de l'instruction primaire. Il semble que l'accroissement des investissements dédiés à l'instruction publique soit le résultat d'une réelle volonté politique de permettre à tous les enfants français, filles comme garçons, d'avoir accès à une instruction primaire.

**Tableau 4 : Relations de causalité entre type de dépenses et scolarisation**

|                    | Nombre d'élèves                                | Délai de réaction* | Pourcentage de dépendance**   |
|--------------------|--|--------------------|---|
| <b>État</b>        | Dépenses $\xrightarrow{+}$ Nombre d'élèves     | Immédiat           | 20%   |
| <b>Département</b> | Dépenses $\leftrightarrow^{+}$ Nombre d'élèves | 3 ans / 3 ans      | $\Delta$ élèves = 5% $\Delta$ dépenses<br>$\Delta$ dépenses = 10% $\Delta$ élèves |
| <b>Communes</b>    | Dépenses $\xrightarrow{+}$ Nombre d'élèves     | 3 ans              | 12%   |
| <b>Ménages</b>     | Dépenses $\xrightarrow{-}$ Nombre d'élèves     | 2 ans              | 16%   |

*Note* : \* Temps de réaction d'une variable suite à une variation de l'autre variable ; \*\* Pourcentage de variation d'une variable dû aux variations de l'autre variable

Nos résultats semblent ainsi suggérer que la scolarisation de masse ayant eu lieu en France au XIX<sup>ème</sup> siècle ait été rendue possible par la volonté politique plus qu'elle ne s'explique par l'accroissement des richesses disponibles dans l'économie au cours de la période étudiée.

## 7. Conclusion

La scolarisation de masse en France est un processus qui s'est fait en plusieurs étapes. La détection des points de rupture dans les diverses dépenses d'éducation pour le niveau primaire souligne l'importance des lois Ferry (1881 et 1882) dans l'évolution du financement du primaire en France : l'État remplace les ménages et les communes dans le financement de l'éducation primaire. Au-delà d'une modification de la structure de financement, les diverses dépenses ont également impacté les taux de scolarisation. L'analyse de la causalité montre ainsi qu'une augmentation des dépenses des communes ou de l'État favorise le développement de la scolarisation. En revanche, il n'existe aucune relation de causalité entre PIB par habitant et nombre d'élèves inscrits dans l'éducation primaire, ni entre PIB par habitant et dépenses d'éducation. Nos résultats semblent ainsi suggérer que la scolarisation de masse ayant eu lieu en France au XIX<sup>ème</sup> siècle ait été rendue

possible par la volonté politique plus qu'elle ne s'explique par l'accroissement des richesses disponibles dans l'économie au cours de la période étudiée.

Nous avons dans cette étude étudié la situation au niveau national pour mieux comprendre la manière dont les institutions formelles ont permis de au processus de massification de s'installer. Bien que n'ayant eu qu'un effet limité sur les taux de scolarisation et sur le processus de scolarisation de masse, les lois Ferry ont eu un profond effet sur la structure des financements, ainsi que sur la réduction des inégalités, en diminuant le poids du financement jusqu'alors principalement supporté par les communes et en centralisant les dépenses au niveau de l'État.

Une grande diversité existe au niveau régional. Il serait intéressant désormais de mieux comprendre les déterminants au niveau désagrégé. Géographiquement, l'on remarque des structures de financement par inscrit très différentes, d'une région à l'autre. Si l'on considère le niveau géographique départemental, on constate des situations extrêmes où la part de l'État peut être inférieure à 1% ou inversement supérieure à 50%. On peut observer une sorte de compensation entre les pourcentages de l'État et des départements dans nombre de cas, ainsi les départements recevant principalement leur financement de l'État figurent parmi les moins dotés par leur assemblée départementale. Quelles sont les régions leaders en terme de financement de l'instruction primaire ? Les disparités s'observent et se maintiennent-elles sur l'ensemble de la période observée ? Existe-t-il au contraire un processus de convergence et de rattrapage au cours du XIXème siècle ?

Toutes ces questions méritent d'être étudiées de manière plus approfondie. L'analyse spatiale est un fantastique outil d'analyse pour mieux comprendre les trajectoires suivies par les départements français, pouvant ainsi nous aider à mieux comprendre les dynamiques et mécanismes en jeu. C'est à cela que s'attacheront nos recherches futures.

## Bibliographie

- Akaike H., 1974, "A new look at the statistical model identification", *IEEE, Transactions on Automatic Control*, 19, pp. 716-723.
- Box G.E.P., Jenkins G.M., 1976, *Time Series Analysis, Forecasting and Control*, Holden Day, San Francisco.
- Box G.E.P., Tiao G.C., 1975, "Intervention Analysis with Application to Economic and Environmental Problems", *Journal of the American Statistical Association*, 70, pp. 70-79.
- Briand J.P., Chapoulie J.M., Huhuet F., Luc J.-N., Prost A., 1987, *L'enseignement primaire et ses extensions. Annuaire statistique, 19ème - 20ème siècles. Ecoles maternelles, primaires, primaires supérieures et professionnelles*, Economica/Inrp, Paris.
- Bruneau C., 1996, "Analyse économétrique de la causalité : un bilan de la littérature", *Revue d'économie politique*, 106, pp. 323-353.
- Carry A., 1995, "Les indicateurs monétaires du volume de l'éducation en France", in : A. Carry, C. Diebolt, L. Fontvieille, C.E. Núñez, G. Tortella (Eds.) : *Education et croissance économique: évolution de longue période et prospective. Une analyse comparée des systèmes éducatifs allemand, espagnol et français*, Stimulation Plan for Economic Science, European Union, Vol. 1, Bruxelles, pp. 207-310.
- Carry A., 1999, Le compte satellite rétrospectif de l'éducation en France (1820-1996), Numéro spécial de la revue *Economies et Sociétés*, Série HEQ, 25.
- Charmasson T. Ed., 1986, *L'histoire de l'enseignement, 19ème - 20ème siècles. Guide du chercheur*, Inrp/Publications de la Sorbonne, Paris.
- Darné O., Diebolt C., 2004, "Unit Roots and Infrequent Large Shocks: New International Evidence on Output", *Journal of Monetary Economics*, 51, pp. 1449-1465.
- Darné, O., Diebolt, C., 2006, "Chocs temporaires et permanents dans le PIB de la France, du Royaume-Uni et des États-Unis", *Revue d'économie politique*, 116(1), pp. 65-78.
- Dep/Insee, 1990, *Le compte de l'éducation. Principes et méthodes*, 7, Les dossiers Education & Formations, Paris.
- Dickey D., Fuller W., 1979, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, 74, pp. 427-431.
- Dickey D., Fuller W., 1981, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49, pp. 1057-1072.
- Diebolt C., 1995, "Les indicateurs physiques du volume de l'éducation en France", in : A. Carry, C. Diebolt, L. Fontvieille, C.E. Núñez, G. Tortella (Eds.) : *Education et croissance économique: évolution de longue période et prospective. Une analyse comparée des systèmes éducatifs allemand, espagnol et français*, Stimulation Plan for Economic Science, European Union, Vol. 1, Bruxelles, pp. 45-206.

- Diebolt C., 1997, “L'évolution de longue période du système éducatif allemand : 19ème et 20ème siècles”, Numéro spécial de la revue *Economies et Sociétés*, Série AF, 23.
- Diebolt C., 1999a, “Government Expenditure on Education and Economic Cycles in the Nineteenth and Twentieth Centuries. The Case of Spain with Special Reference to France and Germany”, *Historical Social Research. An International Journal for the Application of Formal Methods to History*, 24, pp. 3-31.
- Diebolt, C., 1999b, “Histoire quantitative de l'instruction primaire en France avant les lois Ferry. Les indicateurs monétaires par départements”, *Research Papers*, 7, Humboldt University, Center for Comparative Education, Berlin, 195 pages.
- Diebolt, C., 1999c, “L'évolution de longue période du système éducatif en France: 19ème et 20ème siècles”, *Research Papers*, 8, Humboldt University, Center for Comparative Education, Berlin, 121 pages.
- Diebolt, C., 2000, “Die Erfassung der Bildungsinvestitionen im 19. und 20. Jahrhundert. Deutschland, Frankreich, Großbritannien und Spanien im Vergleich”, *Zeitschrift für Erziehungswissenschaft*, 3, pp. 517-538.
- Diebolt C., 2016, “Cliometrica after 10 Years: Definition and Principles of Cliometric Research”, *Cliometrica*, 10, pp. 1-4.
- Diebolt C., Fontvieille L., 2001, “Dynamic Forces in Educational Development: A Long-Run Comparative View of France and Germany in the 19th and 20th Centuries”, *Compare*, 31, pp. 295-309.
- Diebolt C., Franzmann G., Hippe R., Sensch J., 2017, “The Power of Big Data: Historical Time Series on German Education”, *Journal of Demographic Economics*, 83, pp. 329-376.
- Diebolt C., Hauptert, M. Eds., 2019, *Handbook of Cliometrics*, 2<sup>nd</sup> Edition, Berlin, Springer, Collection "Springer Reference".
- Diebolt C., Jaoul M., San Martino G., 2005, “Le mythe de Ferry : une analyse cliométrique”, *Revue d'Economie Politique*, 115, pp. 471-497.
- Eckstein Z., Schultz T. P., Wolpin K., 1984, “Short-run fluctuations in fertility and mortality in pre-industrial Sweden”, *European Economic Review*, 26(3), pp. 295-317.
- Elliott G., T. J. Rothenberg, J. H. Stock, 1996, “Efficient tests for an autoregressive Unit Root”, *Econometrica*, 64, pp. 813-836.
- Engle R., Granger C., 1987, “Co-integration and Error correction: Representation, Estimation, and Testing”, *Econometrica*, 55, pp. 251-276.
- Engle R., Granger C., 1991, *Long-Run Economic Relationships. Readings in Cointegration*, Oxford University Press, Oxford.
- Flora P., 1983, *State, Economy and Society in Western Europe 1815-1975. A Data Handbook in Two Volumes*, MacMillan Press, London.

- Fontvieille L. 1976, *Evolution et croissance de l'État français : 1815-1969*, Numéro spécial de la revue Economies et Sociétés, Série AF, 13.
- Fontvieille L., 1982, *Evolution et croissance de l'Administration Départementale Française 1815-1974*, Numéro spécial de la revue Economies et Sociétés, Série AF, 14.
- Granger C., 1969, "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica*, 37, pp. 426-438.
- Granger C., P. Newbold, 1976, "Experience with forecasting univariate time series and the combination of forecasts", *Journal of the Royal Statistical Society. Series A (General)*, pp. 131-165.
- Grew R., Harrigan P., 1991, *School, State, and Society. The Growth of Elementary Schooling in Nineteenth-Century France — A Quantitative Analysis*, University of Michigan Press, Ann Arbor.
- Hannan E. J., B. G. Quinn, 1979, "The Determination of the Order of an Autoregression", *Journal of the Royal Statistical Society*, 41, pp.190-195.
- Hendry D., Mizon, G., 1993, "Evaluating dynamic econometric models by encompassing the VAR". In, Phillips, Peter C.B. (ed.) *Models, Methods and Applications of Econometrics*. Oxford, UK. Blackwell, pp. 272-300.
- Johansen S., 1988, "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, pp. 231-254.
- Luc J.-N., 1985, *La statistique de l'enseignement primaire 19ème - 20ème siècles. Politique et mode d'emploi*, Inrp/Cnrs, Paris.
- Monfort A., Rabemananjara R., 1990, "From a VAR model to a structural model, with an application to the wage-price spiral", *Journal of Applied Econometrics*, 5, 203-227.
- Nelson C., Plosser C., 1982, "Trend and Random Walks in Macroeconomic Time Series", *Journal of Monetary Economics*, 10, pp. 139-162.
- Salanié B., 1999, "Guide pratique des séries non-stationnaires", *Économie et prévision*, 137(1), pp. 119-141.
- Schwartz G., 1978, "Estimating the dimension of a model", *The Annals of Statistics*, 6, pp. 461-464.
- Sims C., 1980, "Macroeconomics and Reality", *Econometrica*, 48, pp. 1-48.
- Toutain J.-C., 1987, *Le produit intérieur brut de la France : 1789-1982*, Numéro spécial de la revue Economies et Sociétés, Série AF, 15.
- Toutain J.-C., 1997, "Le produit intérieur brut de la France, 1789-1990", *Economies et Sociétés*, Série HEQ, 1.
- Tsay R.S., 1988, "Outliers, Level Shifts, and Variance Changes in Time Series", *Journal of Forecasting*, 7, pp. 1-20.

## Appendix A – Méthodologie

### a) Outliers

Nous considérons, dans cet article, une série temporelle univariée  $y_t^*$  décrite par un modèle ARIMA(p, d, q) :

$$\alpha(B)\phi(B)y_t^* = \theta(B)a_t \quad t = 1, \dots, T \quad (\text{a.1})$$

où B est l'opérateur de retard,  $\alpha(B)$ ,  $\phi(B)$ ,  $\theta(B)$  sont des polynômes des retards d'ordre respectif d, p, et q, et  $a_t$  un bruit blanc.

Les *outliers* peuvent être modélisés par des polynômes de régression de la manière suivante :

$$y_t = y_t^* + \sum_i \omega_i \nu_i(B) I_t(\tau) \quad i = 1 \dots 4 \quad (\text{a.2})$$

où  $y_t^*$  est un processus ARIMA,  $\nu_i(B)$  est le polynôme caractérisant l'*outlier* qui se produit au temps  $t = \tau$ ,  $\omega_i$  représente son impact sur la série, et  $I_t(\tau)$  est un indicateur prenant la valeur 1 si  $t = \tau$  et zéro sinon.  $\tau$  représente la date d'apparition du point atypique.

Les *outliers* sont classés selon quatre grandes catégories :

- Les *Additive Outliers* (AO) affectent uniquement une observation de la série et non ses valeurs futures. En termes de polynômes de régression, les AO sont modélisés par :  $\nu_i(B) = \nu_1(B) = 1$ .
- Les *Innovative Outliers* (IO) affectent la série de manière temporaire avec la même dynamique qu'une innovation. Le polynôme est alors :  $\nu_i(B) = \theta(B) / \phi(B)$ .
- Les *Level Shifts* (LS) qui augmentent ou diminuent toutes les observations en un certain point de la série par une constante. Ils ont un effet permanent sur le niveau de la série. Dans ce cas le polynôme est :  $\nu_i(B) = 1 / (1 - B)$ .
- Les *Temporary Changes* (TC) augmentent ou diminuent de manière abrupte le niveau de la série qui retourne rapidement à son niveau initial de manière exponentielle. Ils affectent temporairement la série. Leur vitesse retour dépend du paramètre  $\delta$ , dans le polynôme  $\nu_i(B) = 1 / (1 - \delta B)$  avec  $0 < \delta < 1$ .

Nous considérons que les AO et IO sont des points atypiques et que les TC et LS sont plutôt des changements structurels. Les TC représentent un changement éphémère sur le niveau de la série tandis que les LS reflètent plutôt un choc permanent sur le niveau et seulement un effet transitoire sur le taux de croissance. Cependant, les IO auront un effet relativement persistant sur le niveau des séries.

Un modèle ARIMA est ajusté à la série  $y_t^*$  dans l'équation (a.1) et les résidus obtenus sont définis par :

$$\text{où } \pi(B) = \frac{\alpha(B)\phi(B)}{\theta(B)} = 1 - \pi_1 B - \pi_2 B^2 - \dots \quad (\text{a.3})$$

Pour les 4 types d'*outliers* définis précédemment, l'équation (a.3) devient :

$$\text{AO: } \hat{a}_t = a_t + \omega_1 \pi(B) I_t(\tau)$$

$$\text{IO: } \hat{a}_t = a_t + \omega_2 I_t(\tau)$$

$$\text{LS: } \hat{a}_t = a_t + \omega_3 \left[ \frac{\pi(B)}{(1-B)} \right] I_t(\tau)$$

$$\text{TC: } \hat{a}_t = a_t + \omega_2 \left[ \frac{\pi(B)}{(1-\delta B)} \right] I_t(\tau)$$

Ces expressions peuvent alors être vues comme un modèle de régression pour les résidus  $\hat{a}_t$ , c'est à dire :

$$\hat{a}_t = \omega_i x_{i,t} + a_t$$

avec :

- pour tout  $i$  et  $t < \tau$ :  $x_{i,t} = 0$

- pour tout  $i$  et  $t = \tau$ :  $x_{i,t} = 1$

$$x_{1,t+k} = -\pi_k \quad (\text{AO});$$

$$x_{2,t+k} = 0 \quad (\text{IO});$$

- pour  $t > \tau$  et  $k \geq 1$ :  $x_{3,t+k} = 1 - \sum_{j=1}^k \pi_j \quad (\text{LS});$

$$x_{4,t+k} = \delta^k - \sum_{j=1}^{k-1} \delta^{k-j} \pi_j - \pi_k \quad (\text{TC}).$$

Les statistiques de test pour les 4 types d'*outliers* sont données par :

$$\begin{aligned}
\text{AO : } \hat{\tau}_1(\tau) &= [\hat{\omega}_1(\tau)/\hat{\sigma}_a] / \left( \sum_{t=\tau}^n x_{1,y}^2 \right)^{1/2} \\
\text{IO : } \hat{\tau}_2(\tau) &= \hat{\omega}_2(\tau) / \hat{\sigma}_a \\
\text{LS : } \hat{\tau}_3(\tau) &= [\hat{\omega}_3(\tau)/\hat{\sigma}_a] / \left( \sum_{t=\tau}^n x_{3,t}^2 \right)^{1/2} \\
\text{TC : } \hat{\tau}_4(\tau) &= [\hat{\omega}_4(\tau)/\hat{\sigma}_a] / \left( \sum_{t=\tau}^n x_{4,t}^2 \right)^{1/2} \quad \hat{\omega}_i(\tau) = \frac{\sum_{t=\tau}^n \hat{a}_t x_{i,t}}{\sum_{t=\tau}^n x_{i,y}^2} \quad \text{pour } i = 1, 2, 3, 4
\end{aligned}$$

et  $\hat{\omega}_2(\tau) = \hat{a}_t$ , où  $\hat{\omega}_i(\tau) (i = 1 - 4)$  représente l'estimation de l'impact du point atypique au temps  $t = \tau$ , et  $\hat{\sigma}_a$  est une estimation de la variance des résidus.

Les points atypiques sont identifiés lors d'une procédure de détection séquentielle, comprenant une itération interne et une autre externe. Dans l'itération externe, en supposant qu'il n'y a pas de points atypiques, un modèle ARIMA (p, d, q) est estimé, donnant ainsi les résidus. Le modèle est en fait déterminé par une procédure automatique d'identification de modèles fondés sur des contraintes concernant les ordres des polynômes (autorégressif et de moyenne mobile) et sur le critère BIC. Les résultats de l'itération externe sont alors utilisés dans l'itération interne pour identifier les points atypiques.

Les statistiques de test pour les quatre types de points atypiques sont calculées pour chaque observation. La valeur absolue la plus grande de ces statistiques ( $\hat{\tau}_{max} = \max |\hat{\tau}_i(\tau)|$ ) est comparée à une valeur critique prédéfinie. Si la statistique de test est plus grande, alors un *outlier* est identifié au temps  $t = \tau$ . La valeur critique, déterminée par le nombre d'observations de la série, est calculée à partir de simulations. Lorsqu'un point atypique est détecté, l'observation  $Y_t$  au temps  $t = \tau$  est ajustée pour obtenir l'observation corrigée  $Y_t^*$  à partir de l'équation (a.2) en utilisant  $\hat{\omega}_i$ , c'est à dire  $Y_t^* = Y_t - \hat{\omega}_i v_i I_t(\tau)$ . Ce processus est répété jusqu'à ce que plus aucun point ne soit trouvé. Ensuite, on retourne à l'itération externe pour ré-estimer le modèle ARIMA en utilisant les données corrigées, et on recommence l'itération interne. La procédure est répétée jusqu'à ce que plus aucun point atypique ne soit détecté. Finalement, une régression multiple sur la série  $Y_t^*$  est effectuée sur les différents points atypiques identifiés afin de déterminer les points atypiques fallacieux.

## b) Causalité

Une série temporelle  $X_t$  est défini comme faiblement stationnaire ou  $I(0)$  (c'est-à-dire intégrée d'ordre 0) si sa moyenne et sa variance sont invariantes dans le temps et que ses auto-covariances ne dépendent que de l'intervalle de temps séparant les observations. De même, une série est définie comme étant intégrée d'ordre  $d$  s'il est nécessaire de la différencier un nombre  $d$  fois pour obtenir une série  $I(0)$  stationnaire.

Afin de tester la non-stationnarité d'une série  $X_t$  (i.e. test de la présence d'une racine unitaire) nous utilisons le test de Dickey-Fuller augmenté. Cette procédure consiste à estimer l'équation (b.1):

$$(1 - L) X_t = a_0 + a_1 T + b_0 X_{t-1} + \sum_{j=1}^k b_j (1 - L) X_{t-j} + u_t \quad (\text{b.1})$$

où  $L$  est un opérateur de décalage,  $T$  est une tendance et  $u_t$  est un terme de perturbation. Nous rejetons l'hypothèse nulle de non-stationnarité (i.e.  $b_0 = 0$ ) si  $b_0$  est suffisamment négative. La valeur critique est obtenue à partir de [Dickey et Fuller \(1979, 1981\)](#) et Mac Kinnon (see [Engle et Granger 1991](#)). Pour tester que deux séries  $X_t$  et  $Y_t$ , formant le vecteur  $Z_t$  sont cointégrées, nous utilisons la méthodologie développée par [Johansen \(1998, 1992, voir aussi Engle et Granger 1991\)](#). Cette méthode est basée sur une représentation de correction d'erreur d'un modèle  $VAR(p)$  avec un terme d'erreur gaussien:

$$\Delta Z_t \alpha + \sum_{k=1}^{p-1} \beta_k \Delta Z_{t-k} + \delta \Delta Z_t - p + \mu_t \quad (\text{b.2})$$

où  $Z_t$  est un  $m * 1$  vecteur de  $I(0)$  variables (dans notre cas,  $m = 2$ ),  $\beta_k$  and  $\delta$  sont des matrices  $m * m$  matrices de paramètres inconnus, et  $\mu_t$  est un terme d'erreur gaussien. Cette équation est estimée par une procédure de maximum de vraisemblance sous l'hypothèse d'un rang réduit  $r < m$  de  $\delta$ ,

$$H(r): \delta = -\Gamma\Omega' \quad (\text{b.3})$$

où  $\Gamma$  et  $\Omega$  sont des matrices  $m * r$ . Johansen a démontré que dans certaines conditions, cette condition de rang réduit de la matrice implique que  $\Omega Z_t$  est stationnaire.

Enfin, pour tester la causalité entre les séries temporelles  $X_t$  et  $Y_t$ , composantes du vecteur  $Z_t$ , nous suivons les procédures classiques d'[Engle et Granger \(1991\)](#). La méthodologie appliquée

diffère selon que les séries chronologiques sont cointégrées ou non. Si ce n'est pas le cas, nous utilisons la méthodologie standard développée par Granger (1969). Ce test est basé sur l'estimation des relations dynamiques avant les premières variables différenciées (si leurs niveaux ne sont pas stationnaires).

Ces relations sont :

$$(1 - L) X_t = \gamma_0 + \sum_{i=1}^m \lambda_i (1 - L) X_{t-i} + \sum_{k=1}^p \sigma_k (1 - L) Y_{t-k} + \nu_t \quad (\text{b.4})$$

$$(1 - L) Y_t = \eta_0 + \sum_{i=1}^n \varphi_i (1 - L) Y_{t-i} + \sum_{k=1}^q \tau_k (1 - L) X_{t-k} + \mu_t \quad (\text{b.5})$$

où  $(\nu_t, \mu_t)$  est un vecteur aléatoire avec une moyenne de 0 et une matrice de covariance finie.

Pour vérifier la présence d'une (ou plusieurs) relation(s) causale(s), nous devons tester la signification conjointe des variables causales, c'est-à-dire retardées  $Y_t$  dans l'équation (b.4) et retardées  $X_t$  dans l'équation (b.5) au moyen d'un test F classique. Par exemple, si  $\sigma_k \neq 0$  et  $\tau = 0$ , nous concluons que  $Y$  Granger cause  $X$ .

Cependant, si les séries chronologiques semblent cointégrées, la causalité doit être étudiée dans le cadre d'un modèle de correction d'erreur. Ce dernier relie les variations à court terme de la série à l'erreur de déséquilibre (c'est-à-dire l'écart entre le comportement réel et la relation à long terme donnée par le vecteur de cointégration).

Le modèle de correction d'erreur est donné par:

$$(1 - L) Z_t = \alpha_0 + \sum_{i=1}^{p-1} \beta_i (1 - L) Z_{t-i} - \Gamma \Omega' Z_{t-p} + \nu_t \quad (\text{b.6})$$

L'existence d'une relation de cointégration entre les deux variables garantit qu'il existe au moins un lien de causalité entre elles. Tester la causalité équivaut donc à tester la signification conjointe des paramètres sur les variables causales supposées.

## Annexe B – Tests économétriques

Tableau B1 : Tests de stationnarité d'Elliot, Rothenberg et Stock (ERS)

| Variable                       | Hypothèse testée  | Variabes incluses dans l'équation testée | Statistique | Décision     |
|--------------------------------|-------------------|--|-------------|--------------|
| Dépenses totales               | Racine unitaire   | Tendance + constante                     | -1,61***    | Processus DS |
| D (Dépenses totales)           | Racine unitaire   | Constante                                | -5,76***    | Stationnaire |
| Dépenses communes              | Racine unitaire   | Tendance + constante                     | -2,60***    | Processus DS |
| D (Dépenses communes)          | Racine unitaire   | Constante                                | -9,36***    | Stationnaire |
| Dépenses département           | Racine unitaire   | Tendance + constante                     | -1,82***    | Processus DS |
| D (Dépenses départements)      | Racine unitaire   | Constante                                | -11,62***   | Stationnaire |
| Dépenses ménages               | Racine unitaire   | Tendance + constante                     | -1,64***    | Processus DS |
| D (Dépenses ménages)           | Racine unitaire   | Constante                                | -5,75***    | Stationnaire |
| Dépenses État (                | Racine unitaire   | Tendance + constante                     | -0,64***    | Processus DS |
| D (Dépenses État)              | Tendance linéaire | Tendance + constante                     | -7,73***    | Processus TS |
| SD (Dépenses État)             | Racine unitaire   | Constante                                | -7,65***    | Stationnaire |
| Dépenses autres adm. pub.      | Racine unitaire   | Tendance + constante                     | -1,09***    | Processus DS |
| D (Dépenses autres adm. pub.)  | Tendance linéaire | Tendance + constante                     | -5,14***    | Processus TS |
| SD (Dépenses autres adm. pub.) | Racine unitaire   | Constante                                | -4,99***    | Stationnaire |
| PIB par habitant               | Racine unitaire   | Tendance + constante                     | -1,29***    | Processus DS |
| D (PIB par habitant)           | Racine unitaire   | Constante                                | -11,73***   | Stationnaire |
| Nombre d'élèves                | Racine unitaire   | Tendance + constante                     | -0,54***    | Processus DS |
| D (Nombre d'élèves)            | Racine unitaire   | Constante                                | -4,21***    | Stationnaire |
| Nb élèves/population           | Racine unitaire   | Tendance + constante                     | -0,44***    | Processus DS |
| D (Ratio élèves)               | Tendance linéaire | Tendance + constante                     | -5,02***    | Processus TS |
| SD (Ratio élèves)              | Racine unitaire   | Constante                                | -4,33***    | Stationnaire |
| Nombre d'illettrés             | Racine unitaire   | Tendance + constante                     | -1,99***    | Processus DS |
| D (Illettrés)                  | Racine unitaire   | Constante                                | -9,44***    | Stationnaire |
| Nb d'illettrés/population      | Racine unitaire   | Tendance + constante                     | -1,51***    | Processus DS |
| D (Ratio illettrés)            | Racine unitaire   | Constante                                | -9,64***    | Stationnaire |

Notes: \*, \*\*, \*\*\* significatif à 10%, 5%, 1% ; (1) D(variable) signifie que l'on considère les différences premières de la variable; (2) SD(variable) : signifie que l'on considère la variable corrigée d'une tendance linéaire.

Tableau B2 : Sélection du retard optimal

| Variables                                    | Retard Optimal pour chaque critère |     |     |    |    | Retard sélectionné |
|--|------------------------------------|-----|-----|----|----|--------------------|
|  | LR                                 | FPE | AIC | SC | HQ |                    |
| Elèves et Dépenses totales                   | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Elèves et Dépenses autres adm. pub.          | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Elèves et Dépenses communes                  | 2                                  | 2   | 2   | 1  | 2  | 2                  |
| Elèves et Dépenses départements              | 2                                  | 2   | 2   | 1  | 2  | 2                  |
| Elèves et Dépenses État                      | 3                                  | 3   | 3   | 1  | 1  | 3                  |
| Elèves et Dépenses ménages                   | 1                                  | 2   | 2   | 1  | 1  | 1                  |
| Elèves et PIB par habitant                   | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Illettrés et Dépenses totales                | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Illettrés et Dépenses autres adm. pub.       | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Illettrés et Dépenses communes               | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Illettrés et Dépenses départements           | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Illettrés et Dépenses État                   | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Illettrés et Dépenses ménages                | 1                                  | 3   | 3   | 1  | 1  | 1                  |
| Illettrés et PIB par habitant                | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio élèves et Dépenses totales             | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio élèves et Dépenses autres adm. pub.    | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio élèves et Dépenses communes            | 5                                  | 2   | 2   | 1  | 2  | 2                  |
| Ratio élèves et Dépenses départements        | 2                                  | 2   | 2   | 1  | 1  | 2                  |
| Ratio élèves et Dépenses État                | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio élèves et Dépenses ménages             | 2                                  | 2   | 2   | 1  | 1  | 2                  |
| Ratio élèves et PIB par habitant             | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio illettrés et Dépenses totales          | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio illettrés et Dépenses autres adm. pub. | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio illettrés et Dépenses communes         | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio illettrés et Dépenses départements     | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio illettrés et Dépenses État             | 1                                  | 1   | 1   | 1  | 1  | 1                  |
| Ratio illettrés et Dépenses ménages          | 1                                  | 3   | 3   | 1  | 1  | 1                  |
| Ratio illettrés et PIB par habitant          | 1                                  | 1   | 1   | 1  | 1  | 1                  |

**Tableau B3 : Test de Johansen**

| <b>Variables</b>                         | <b>Hypothèse testée :<br/>nb de relations de<br/>cointégration</b> | <b>Valeur<br/>propre</b> | <b>Statistique</b> | <b>Valeur<br/>critique (5%)</b> | <b>Résultat</b>   |
|--|--|--------------------------|--------------------|---------------------------------|-------------------|
| Illettrés et Dépenses totales            | 0  | 0,10                     | 8,63               | 15.41                           | Pas cointégration |
| Illettrés et Dépenses communes           | 0  | 0,04                     | 4,09               | 15.41                           | Pas cointégration |
| Illettrés et Dépenses départements       | 0  | 0,06                     | 7,33               | 15.41                           | Pas cointégration |
| Illettrés et Dépenses ménages            | 0  | 0,09                     | 11,25              | 15.41                           | Pas cointégration |
| Ratio illettrés et Dépenses totales      | 0  | 0,08                     | 8,51               | 15.41                           | Pas cointégration |
| Ratio illettrés et Dépenses communes     | 0  | 0,04                     | 4,96               | 15.41                           | Pas cointégration |
| Ratio illettrés et Dépenses départements | 0  | 0,06                     | 8,55               | 15.41                           | Pas cointégration |
| Ratio illettrés et Dépenses ménages      | 0  | 0,13                     | 14,25              | 15.41                           | Pas cointégration |