

« Bref retour cliométrique sur 50 ans de performances scolaires en lecture et en mathématique en France : 1970-2020 »

Auteurs


Nadir ALTINOK, Claude DIEBOLT

Document de Travail n° 2023 – 04

Janvier 2023

Bureau d'Économie
Théorique et Appliquée
BETA

www.beta-umr7522.fr

 @beta_economics

Contact :
jaoulgrammare@beta-cnrs.unistra.fr

Bref retour cliométrique sur 50 ans de performances scolaires en lecture et en mathématique en France : 1970-2020

Nadir ALTINOK¹ et Claude DIEBOLT²

Résumé

Cette contribution vise à retracer l'évolution de longue période des compétences scolaires des élèves français de 1970 à 2020. Partant des résultats fournis par les enquêtes internationales sur les acquis des élèves, nous présentons, à partir d'une base de données inédite, l'originalité de la trajectoire nationale vis-à-vis des autres pays de l'OCDE. Au cours des 50 dernières années, nous montrons que, pour le cas de la France, les performances en lecture et en mathématiques diminuent, les scores moyens étant significativement faibles. Par ailleurs, nous nous interrogeons sur la trajectoire française en termes d'acquis scolaires qui, pas à pas, tend à diverger de celle des autres pays de l'OCDE. Ce faisant, nos résultats enrichissent les conclusions des travaux issus des enquêtes nationales tout en livrant de nouveaux éléments de preuve associés à des éclairages historiques et comparatifs renouvelés.

Mots-clés : Qualité de l'éducation, Enquêtes internationales sur les acquis des élèves, Cliométrie, PISA, TIMSS, PIRLS

Classification JEL : C8, I2, J24, N3, O15

¹ Nadir ALTINOK, BETA/CNRS, Université de Lorraine & Association Française de Cliométrie. Email: nadir.altinok@univ-lorraine.fr (auteur correspondant)

² Claude DIEBOLT, BETA/CNRS, Université de Strasbourg & Association Française de Cliométrie. Email: cdiebolt@unistra.fr

1. Introduction

En 1989, Christian Baudelot et Roger Establet publiaient *Le niveau monte*. Ils y affirmaient qu'il ne faut pas s'alarmer quant à la capacité du système éducatif français à faire progresser les élèves (Baudelot et Establet, 1989). Près de vingt ans plus tard, les conclusions des auteurs sont plus timorées et soulignent l'importance accordée aux acquis scolaires, notamment aux résultats de l'évaluation PISA (*Programme for International Student Assessment*), peu glorieux pour l'hexagone (Baudelot et Establet, 2009).

Certes, entre 1970 et 2020, le taux d'achèvement scolaire dans le secondaire est passé de 63% à près de 100%. La scolarisation universelle est atteinte pour quasiment tous les jeunes entre 6 et 16 ans. Il serait cependant trompeur de se focaliser sur la seule expansion scolaire. La scolarisation n'équivaut pas à l'apprentissage (Pritchett, 2013). Qui plus est, depuis plusieurs décennies, il est convenu d'accepter qu'une année de scolarisation n'apporte pas le même niveau de compétences, quel que soit le pays. Le gain en termes de compétences scolaires est le plus souvent basé sur des évaluations standardisées sur les acquis des élèves. La plus connue, PISA, a mis en avant les faiblesses du système éducatif français. Après quelques années de critiques vis-à-vis d'évaluations supposément non adaptées au contexte français (Mons et Pons, 2013, Cédelle, 2015), force est de constater que l'école française n'est pas la meilleure école du monde. Quant à la désaffection des candidats aux fonctions d'enseignement, quel que soit le niveau scolaire, elle témoignerait d'un malaise au sein du système éducatif français.

Mais l'expansion scolaire a-t-elle vraiment conduit à une diminution des acquis scolaires en France ?

Récemment, un rapport de la DEPP (2022) montre une baisse continue, depuis 30 ans, du niveau en mathématiques. Il reprend et standardise les scores recueillis dans l'ensemble des enquêtes qui évaluent le niveau en mathématiques par le biais d'une échelle de scores. Martin *et al.* (2022) reprennent les résultats de cette étude et l'approfondissent pour souligner un déclassement de la France vis-à-vis des systèmes éducatifs d'autres pays de l'OCDE, avec un niveau général qui semble diminuer, ou tout du moins se stabiliser, et une montée des inégalités à travers un tassement dans les résultats des moins performants.

Notre contribution se situe en prolongement de ces travaux. En recourant à une base de données inédite, construite à partir des enquêtes internationales sur les acquis des élèves, nous comparons l'évolution de la performance, depuis 1970, du niveau des élèves français en mathématiques, sciences et en lecture.

2. Données et méthodologie

L'originalité de notre travail réside dans l'actualisation et dans l'emploi d'une base de données contenant des scores comparables dans plusieurs domaines de compétences scolaires, à différents niveaux et sur une période couvrant environ cinq décennies (1970-2020)³.

³ Altinok N., Diebolt C.: "Cliometrics of Learning-Adjusted Years of Schooling: Evidence from a New Dataset", *AFC Working Papers*, 2, 2023, pp. 1-78 (https://www.cliometrie.org/images/wp/AFC_WP_02_2023.pdf).

Pour y parvenir, nous avons dû surmonter deux difficultés majeures : la faible participation des pays à des enquêtes sur les acquis des élèves – notamment avant les années 1990 – mais aussi la grande hétérogénéité des données disponibles depuis les années 1970 (cf. Filmer *et al.*, 2020, Angrist *et al.*, 2021).

2.1. Les évaluations sur les acquis des élèves

Notre base de données internationale sur les acquis des élèves mobilise les résultats issus de diverses enquêtes (TIMSS, PIRLS et PISA⁴). D'autres évaluations se focalisent sur les compétences des adultes (IALS, PIAAC⁵).

Dès 1960, sous l'égide de l'Association Internationale pour l'Évaluation du Rendement scolaire (IEA) une enquête pilote est menée pour 12 pays dont la France, ceci afin d'évaluer les compétences en mathématiques, lecture, géographie, sciences et compétences non verbales (Foshay *et al.*, 1962). Par la suite, plusieurs évaluations ont testé les élèves du primaire et du secondaire en mathématiques et en sciences (*First International Mathematics Study*, *Second International Mathematics Study*, *First International Science Study* et *Second international Science Study*). L'évaluation TIMSS débute en 1995 et se focalise sur les mathématiques et les sciences. Depuis cette date, plusieurs vagues de tests se sont déroulés avec une périodicité de 4 ans⁶. En complément, l'IEA a aussi évalué les compétences en lecture des élèves et ce dès les années 1970. La première étude principale a été menée en 1970 et a inclus 14 pays (*Six Subject Survey: Reading Comprehension*). C'est toutefois l'enquête PIRLS qui a mis en avant les compétences des jeunes en lecture. L'évaluation mesure les compétences des élèves du grade 4 (équivalent CM1 en France) tous les 5 ans depuis 2001⁷. Parallèlement aux enquêtes de l'IEA, c'est l'évaluation PISA qui a suscité un grand écho médiatique, notamment en Europe ; PISA étant une évaluation développée par l'OCDE. Elle a démarré en 2000. Testant les élèves de 15 ans dans trois domaines de compétences (mathématiques, sciences et lecture), PISA est devenue l'enquête recouvrant le plus de pays dans le monde. Organisée tous les 3 ans, sept différentes vagues ont été effectuées jusqu'en 2018 avec plus de 70 pays participants lors de la dernière vague. En parallèle des évaluations sur les acquis des élèves, les enquêtes sur les compétences des adultes (IALS et PIAAC) peuvent aussi s'avérer être très utiles pour mieux évaluer les compétences des élèves. En utilisant une méthodologie spécifique, certains auteurs ont, en effet, obtenu des données de panel relatifs au niveau de compétences des jeunes adultes (Schwerdt et Wiederhold, 2019).

⁴ Respectivement *Trends on International Mathematics and Science Study*, *Progress on International Reading Literacy Study* et *Programme for International Student Assessment*.

⁵ Respectivement *International Adult Literacy Survey* et *Programme for the International Assessment of Adult Competencies*.

⁶ La deuxième vague date de 1999, puis 2003, 2007... La dernière vague avec des données disponibles date de 2019.

⁷ Les différentes vagues de tests sont 2001, 2006, 2011, 2016 et 2021. Les données de PIRLS 2021 n'étaient pas disponibles lors de la rédaction de cet article.

2.2. Méthodologie générale

Notre recherche repose sur un regroupement des différentes enquêtes internationales (présentées dans la section précédente) sur des échelles comparables. Puisque ces enquêtes sont de nature diverses, il nous a fallu élaborer une méthodologie permettant de les ajuster entre elles, ceci afin d'éviter tout biais potentiel. En effet, inutile de mentionner ici que le score obtenu, en données brutes, pour un pays aurait aisément pu être plus élevé dans une enquête par rapport à une autre, du simple fait que la première enquête, par exemple, aurait eu un niveau de difficulté supérieur. Ce faisant, si l'on comparait naïvement les résultats entre les deux enquêtes, il serait hautement probable que l'interprétation première aurait conduit à une baisse du niveau des élèves. Nous avons donc procédé à un ajustement des enquêtes, ceci afin d'envisager des comparaisons dans le temps et dans l'espace. L'idée générale qui a servi de fil conducteur à notre démarche consiste à utiliser les résultats des pays ayant participé simultanément à plusieurs enquêtes à la même période (Angrist *et al.*, 2021). En effet, en combinant l'ensemble des pays ayant participé à plusieurs évaluations, il est possible de calculer des taux de conversion, comme on le ferait pour un taux de change entre des devises par exemple. Cette méthode est utilisée pour ajuster les scores des enquêtes réalisées à partir des années 1990 et a déjà été mise en œuvre par de précédentes recherches (Angrist *et al.*, 2013, Angrist *et al.*, 2021, Gust *et al.*, 2022). Une approche additionnelle a aussi été implémentée pour ancrer les enquêtes entre elles dans le temps, et surtout avant 1990. Suivant l'approche pionnière de Hanushek et Kimko (2000), il est ainsi possible d'ancrer les données des enquêtes internationales en se référant aux résultats des Etats-Unis dans l'enquête nationale sur les acquis des élèves, à savoir le *National Assessment of Educational Progress* (NAEP)⁸.

L'évaluation NAEP a débuté dès les années 1970. Désormais, elle permet une comparaison des scores en lecture, mathématiques et sciences sur près de 50 ans. Jusqu'en 1990, les enquêtes de l'IEA ont été ancrées, en réajustant les scores, de telle sorte à correspondre aux scores des Etats-Unis dans l'enquête NAEP. Ainsi, si le score des Etats-Unis augmente de 1% entre 1970 et 1980, l'enquête de l'IEA de 1980 sera ajustée de telle manière que le score des Etats-Unis soit égal à cette même variation. Ce travail d'ancrage est toutefois facilité depuis les années 1990 puisque le recours à l'utilisation de la méthode IRT (*Item Response Theory*) permet dorénavant d'avoir des scores comparables entre les différentes séries d'évaluation au sein des enquêtes. Aussi, il est possible de comparer les résultats de PISA 2009 avec ceux de PISA 2015 sans ajustement spécifique. Ce faisant, l'ajustement via l'enquête NAEP devient quasi inutile pour la plupart des enquêtes réalisées après 1990.

⁸ Le détail de la méthodologie pour l'obtention d'une base de données regroupant plus de 130 pays est présentée dans Altinok et Diebolt (2023). Le présent article se focalise sur les résultats de la France, en comparant l'hexagone à un échantillon de 20 pays de l'OCDE. Pour mémoire, rappelons ici que la base de données créée contient bien sûr un échantillon de données plus larges et plus détaillées.

Quant à PISA et TIMSS, le calcul d'un taux de conversion général suffit pour obtenir une base de données comparable sur le plan des acquis scolaires⁹. Une nouveauté réside toutefois dans l'implémentation de la méthode d'imputation utilisée dans Lim *et al.* (2018). En recourant à des données fortement corrélées avec les acquis scolaires, notre recherche permet ainsi d'élargir à la fois la période couverte par les évaluations sur les acquis, mais aussi de prédire les scores pour des pays n'ayant pas participé à toutes ces évaluations. Cette procédure d'imputation utilise le paquet "Amelia II" (Honaker *et al.*, 2011) et s'avère utile pour obtenir des données de panel sur près d'un demi-siècle (1970-2020). En effet, la base de données préparée par Angrist *et al.* (2021) ne couvre que la période 2010-2020. Le recours à de l'imputation multiple et aux enquêtes les plus anciennes, en utilisant l'ancrage par le taux de conversion et par le biais de l'évaluation NAEP, nous permet donc d'obtenir des scores comparables pour les principaux pays de l'OCDE entre 1970 et 2020.

De plus, afin de combiner quantité et qualité de l'éducation, nous proposons d'associer le nombre moyen d'années de scolarisation avec les scores des élèves aux évaluations sur les acquis. Filmer *et al.* (2020) ont été les premiers à effectuer une telle conversion. Cette combinaison apporte ainsi une mesure relative de la quantité d'éducation, en prenant en compte le différentiel de qualité éducative entre les pays. L'hypothèse principale se base sur l'idée qu'une année d'éducation dans un pays donné n'apporte pas nécessairement le même rendement scolaire en termes d'acquis dans un autre pays. Certains pays s'avèrent être plus efficaces que d'autres, comme le soulignent les évaluations telles que PISA. Dans cette perspective, l'indicateur hybride que nous mettons en œuvre se nomme "*Learning-Adjusted Years of Schooling*" (Filmer *et al.*, 2020). Afin de calculer cet indicateur relatif, nous divisons le score obtenu par 700 (étant considéré comme un plafond théorique), ce qui nous donne une valeur comprise entre 0 et 1. Puis, ce ratio est multiplié par le nombre moyen d'années de scolarisation, issu de la base de données de Barro et Lee (2013). La nouvelle base de données ainsi obtenue fournit des informations au sujet de trois variables complémentaires : la qualité des systèmes éducatifs (via les scores aux différentes enquêtes sur les acquis et les seuils de performance), la quantité d'éducation (via le nombre moyen d'années de scolarisation) et le nombre d'années de scolarisation ajusté par la qualité de l'éducation (en anglais "*Learning-Adjusted Years of Schooling*" ou LAYS).

De manière très concrète, les données sur la qualité de l'éducation sont réparties sur plusieurs niveaux scolaires (primaire, secondaire) et domaines de compétences (mathématiques, lecture, sciences) et disponibles entre 1970 et 2020 pour la plupart des pays de l'OCDE. En moyenne sur la période étudiée, la performance de la France est inférieure sur le plan de la qualité de l'éducation (20 points de différence), mais aussi sur la scolarisation : le nombre moyen d'années de scolarisation est de 8,6 années contre 10,2 pour la moyenne des pays de l'OCDE-20 (Tableau 1). Ces deux retards entraînent, à leur tour, un déficit d'environ 1,3 année de scolarisation ajustée par la qualité de l'éducation (indicateur LAYS).

⁹ Supposons que la moyenne des scores des pays participant à TIMSS 1999 et PISA 2000 en mathématiques soit égale à 520 dans TIMSS et 495 dans PISA, alors le taux de conversion entre PISA et TIMSS sera de $525/495 = 1,06$. Les scores de PISA seront ainsi réévalués à la hausse d'environ 6,06 %.

Tableau 1. Base de données sur l'éducation (1970-2020)

	Moyenne	Ecart-type	Minimum	Maximum
OCDE-20				
Qualité de l'éducation	508	26	430	560
Années de scolarisation	10.2	2.0	3.2	13.8
Indicateur mixte (LAYS)	7.4	1.7	2.0	10.4
France				
Qualité de l'éducation	488	26	442	514
Années de scolarisation	8.6	2.1	5.5	11.3
Indicateur mixte (LAYS)	6.1	1.7	3.4	8.2

Notes : Les données regroupent toutes les observations disponibles entre 1970 et 2020 pour une sélection de 20 pays de l'OCDE. Au total 220 observations sont disponibles pour l'échantillon complet et 11 observations pour la France (données disponibles tous les 5 ans). Source : Altinok et Diebolt (2023)

3. Une évolution des acquis scolaires contrastée

Comme indiqué dans l'introduction de cet article, la principale innovation de notre travail réside dans la mise à disposition de données comparables pour la qualité de l'éducation sur près d'un demi-siècle (1970-2020). Notre méthode est comparative et se focalise ici sur des pays similaires à la France, les pays de l'OCDE et plus précisément les 20 pays qui, *ceteris paribus*, font le cœur de l'OCDE¹⁰ et que nous nommons OCDE-20. Un avantage de se restreindre à ces pays est la disponibilité des données sur le long terme et le fait de placer la France autour de pays qui ont quasiment tous achevé la scolarisation primaire et secondaire universelle dès le début des années 2000. On peut ainsi éviter d'inclure des pays qui observent une sélection de leur public scolaire (comme le Mexique ou la Turquie), ce qui aurait tendance à surestimer leur performance dans les différents domaines de compétences¹¹.

3.1. Une performance moindre par rapport à son niveau économique

Non commençons par comparer la qualité de l'éducation en France en coupe instantanée pour l'année 2020. Le niveau de compétences que nous relevons pour le cas français est l'un des plus faibles des pays de l'OCDE-20, notamment pour le niveau primaire, que ce soit en lecture ou en mathématiques (Figure 1). Alors que le score moyen de la France est de 483 points en mathématiques au niveau primaire, il avoisine 550 points en Irlande. En supposant qu'une année d'acquis est similaire à environ 30 points (Avvisati, 2021), cela signifierait que la France est en retard de près de trois ans sur l'Irlande. Les

¹⁰ Nous qualifions ces pays de "membres historiques de l'OCDE" et sélectionnons principalement des pays ayant de fortes similitudes avec la France : Allemagne, Australie, Autriche, Belgique, Canada, Danemark, Espagne, États-Unis, Finlande, France, Grèce, Irlande, Italie, Nouvelle-Zélande, Norvège, Pays-Bas, Portugal, Royaume-Uni, Suède et Suisse. Les pays asiatiques ont des performances plus élevées que les pays occidentaux aux enquêtes sur les acquis. Cependant, certaines spécificités culturelles pouvant expliquer ces différences (voir notamment Boman, 2022). Dans le cadre de cet article, nous avons donc préféré centrer notre attention sur les pays non asiatiques de l'OCDE.

¹¹ A titre d'exemple, Spaul (2018) a montré que les résultats de la Turquie dans l'enquête PISA semblaient surestimés au début des années 2000 du fait d'une scolarisation des jeunes de 15 ans plus faible que dans les autres pays de l'OCDE.

résultats sont plus élevés au niveau secondaire, la France réalisant une performance qui est proche de la moyenne des pays de l'OCDE.

Le Tableau 2 présente l'évolution de la performance pour chaque domaine de compétence (mathématiques, lecture et sciences) et chaque niveau scolaire analysé (primaire et secondaire). La période étudiée s'étale sur 50 ans, de 1970 à 2020. Partant d'une moyenne internationale en 2000 qui est de 500 et d'un écart-type de 100, il apparaît que, depuis 1970, la performance scolaire en France a augmenté. Le niveau en mathématiques pour le primaire est passé de 424 points à 482 points entre 1970 et 2020. Cette hausse représente environ 60 points, soit 12 points par décennie ou l'équivalent de 2 années scolaires d'acquis sur un demi-siècle. Si l'on observe la colonne relative à la variation standardisée de long terme (colonne 7), les gains sont tous positifs sur le long terme : l'école française a donc globalement amélioré sa capacité à apprendre aux élèves depuis 1970. Les gains sont les plus importants en lecture, tant au niveau primaire qu'au niveau secondaire ; ils sont moindres en mathématiques. Signalons ici que les résultats de l'enquête PISA ont, de manière générale, souligné une baisse de niveau des élèves depuis quelques années (OCDE, 2019).

Par-delà l'analyse de long terme (50 ans), nous avons également calculé les variations sur le court terme (sur les 20 dernières années). Ici, les résultats sont clairement plus contrastés et soulignent une baisse absolue de la performance dans tous les domaines et niveaux depuis 2000 (colonne 8). La performance en lecture a diminué d'environ 11 points par décennie sur les 20 dernières années, tandis que la diminution est d'environ 8 points par décennie en mathématiques dans le secondaire. Si l'on rapporte cette variation à l'équivalence en termes d'années d'acquisition, les effets sont assez importants. Alors que, sur le long terme, la France a gagné l'équivalent de plus de 2 années d'acquis en lecture, une perte d'une année d'acquis scolaires a été observée sur la période 2000-2020. La baisse est moindre en mathématiques (respectivement 0,38 et 0,56 pour les niveaux primaire et secondaire).

Nous savons que le niveau des compétences scolaires est étroitement lié au niveau économique des nations (Hanushek et Woessmann, 2015). Même si la relation est moins évidente, on pourrait donc logiquement attendre de meilleurs résultats au sein des écoles les mieux dotées en ressources financières (Hanushek, 2019). Qui plus est, afin de tester dans quelle mesure la performance de la France se situe en-dessous de ce qu'on pourrait attendre au vu de son niveau économique ou de l'ampleur de ses dépenses d'éducation, nous opérons différents tests. Nous effectuons, tout d'abord, une analyse graphique des corrélations entre chaque domaine de compétence et niveau scolaire et l'indicateur de richesse économique. Plus concrètement, nous établissons une corrélation entre les performances scolaires et le PIB par habitant en dollars ajustés par la méthode des parités de pouvoir d'achat. Nous observons une corrélation assez forte, avec un coefficient proche ou supérieur à 60% dans la plupart des cas. En même temps, ce calcul met en lumière l'écart de performance de la France, notamment au niveau primaire, aussi bien en lecture qu'en mathématiques.

La Figure 2 illustre l'écart significatif de la France vis-à-vis des autres pays en mathématiques au niveau primaire¹². Alors que le score escompté serait d'environ 515 points, le score réel n'est que de 482 points.

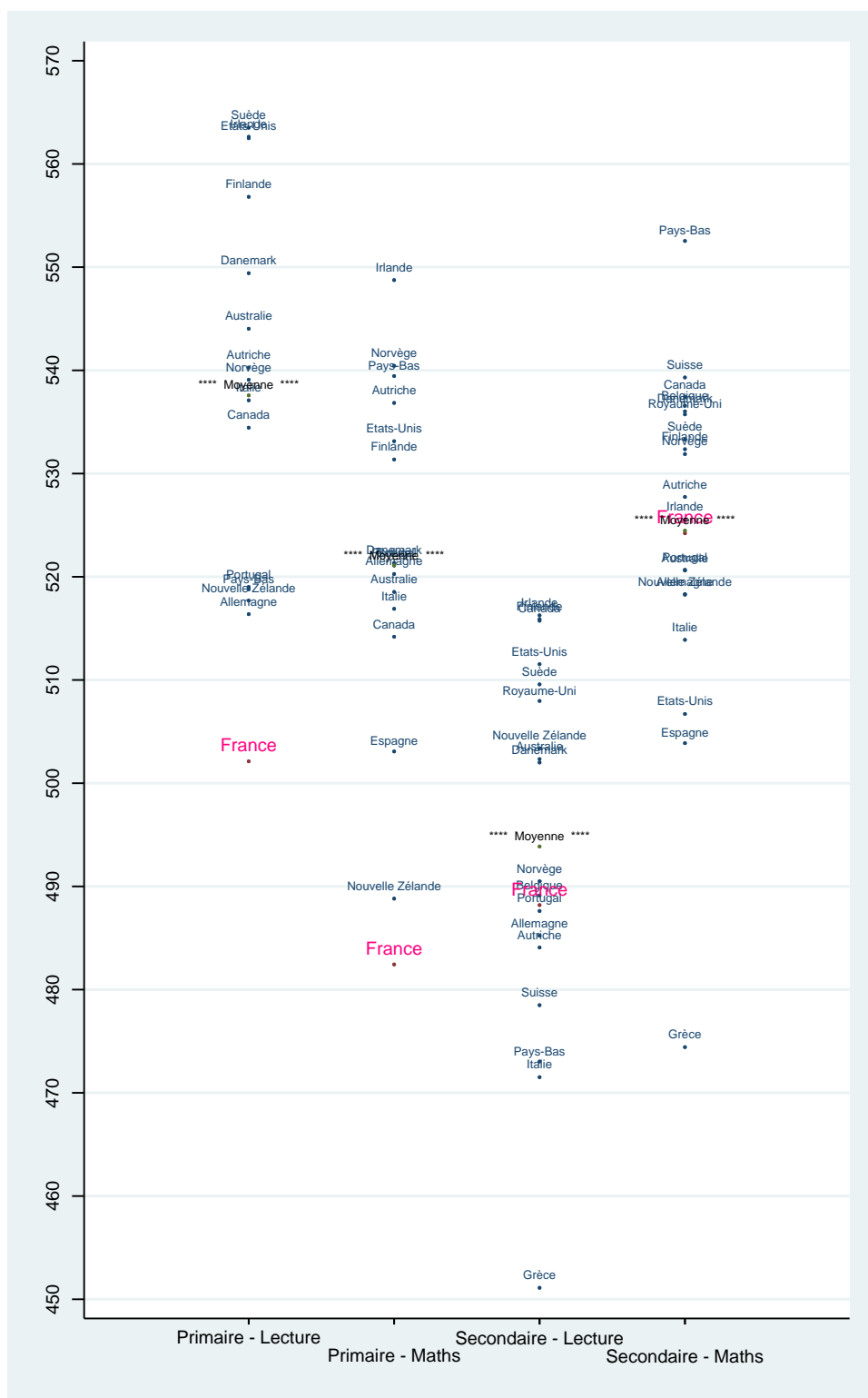
Tableau 2. Evolution de la performance scolaire et effets de productivité

	1970	1980	1990	2000	2010	2020	Variation standardisée		Effets de productivité			
							Long terme	Court terme	Long terme		Court terme	
									Limite faible	Limite haute	Limite faible	Limite haute
Moyenne	442	457	495	514	508	501	11.82	-6.79	1.97	2.95	-0.45	-0.68
Primaire	423	443	500	503	499	490	13.48	-6.43	2.25	3.37	-0.43	-0.64
Lecture	424	455	522	525	520	502	15.68	-11.29	2.61	3.92	-0.75	-1.13
Mathématiques	424	437	489	494	484	482	11.66	-5.76	1.94	2.92	-0.38	-0.58
Sciences	421	437	490	491	492	487	13.11	-2.24	2.18	3.28	-0.15	-0.22
Secondaire	460	471	491	526	517	511	10.16	-7.14	1.69	2.54	-0.48	-0.71
Lecture	420	439	469	505	499	488	13.63	-8.28	2.27	3.41	-0.55	-0.83
Mathématiques	489	491	517	541	523	524	7.11	-8.36	1.18	1.78	-0.56	-0.84
Sciences	473	485	485	531	529	521	9.73	-4.79	1.62	2.43	-0.32	-0.48

Notes : Les six premières colonnes indiquent les scores standardisés de la France dans notre base de données. La variation standardisée représente ainsi la variation absolue de la performance pour chaque dimension sur une base décennale. Les données de long terme concernent la période 1970-2020, tandis que les données de court terme sont relatifs à la période 2000-2020. Les quatre dernières colonnes reprennent les résultats de l'OCDE sur les gains potentiels obtenus par une année de scolarisation. Nous considérons deux effets extrêmes, en utilisant les résultats de l'Allemagne. La limite faible est calculée à partir d'un effet de 20 points pour chaque année de scolarisation, tandis que la limite haute est basée sur une hypothèse de 30 points. Voir Avvisati *et al.* (2021) pour les analyses de productivité. Source : Altinok et Diebolt (2023)

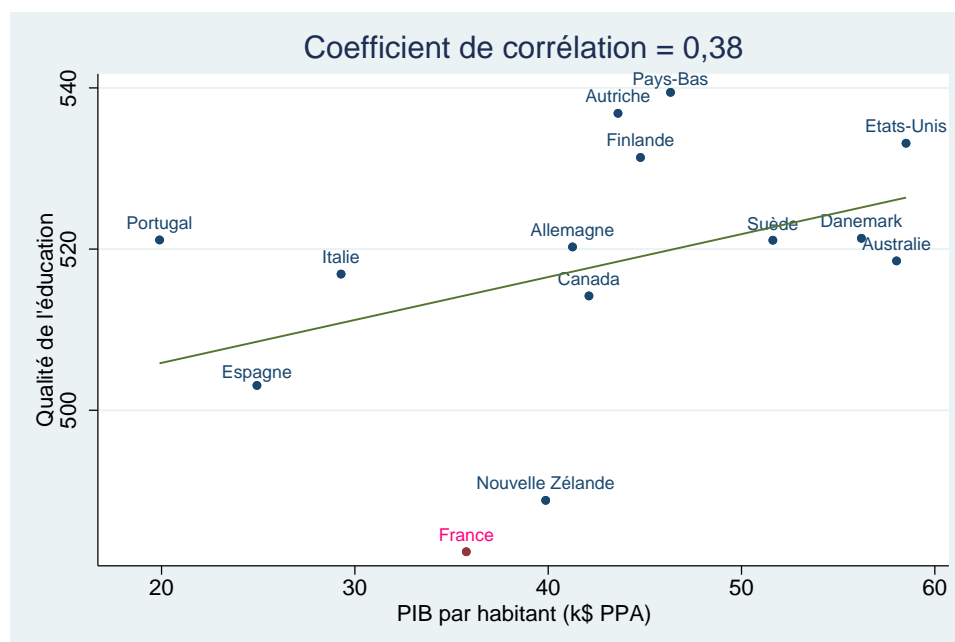
¹² Cette analyse a été effectuée pour chaque domaine de compétence et chaque niveau scolaire. L'ensemble des résultats est disponible sur demande auprès des auteurs.

Figure 1. Performance moyenne des pays de l'OCDE en 2020



Notes : Seuls 20 pays de l'OCDE sont utilisés pour établir ce classement. Les scores représentent une moyenne des résultats des pays à chaque domaine de compétence (mathématiques et lecture) pour l'année 2020. La méthodologie utilisée pour construire les données est présentée en section 2. Pour des raisons de clarté, les résultats obtenus pour le Japon sont omis. Source : Altinok et Diebolt (2023)

Figure 2. Niveau économique et scores moyens en mathématiques au primaire



Notes : Cette figure relie le niveau économique des pays (mesuré par le PIB par habitant, exprimé en milliers de dollars des Etats-Unis, ajusté par la méthode des parités de pouvoir d'achat [PPA]) avec la qualité de l'éducation, mesurée à partir des résultats standardisés en mathématiques, au niveau primaire. Source : Altinok et Diebolt (2023).

Afin de mieux appréhender ce potentiel écart de performance, nous calculons une régression de la performance sur le niveau économique des pays. Au-delà du coefficient associé à cette dernière variable, c'est le résidu attaché à chaque pays qui retient notre attention (Tableau 3).

Dans chacune des régressions, nous associons une dimension particulière des compétences éducatives avec le PIB par habitant. Puisque les données sont en panel, cela nous permet d'obtenir des effets fixes pays qui ne sont pas expliqués par l'indicateur de richesse économique. Sur les 20 pays inclus dans l'échantillon¹³, le classement de la France est souvent en dernière position, soulignant le score anormalement faible lorsque l'on prend en compte son niveau économique. En effet, rappelons ici que le classement des résidus permet de mesurer ce qui n'est pas pris en compte dans le modèle, ici le niveau économique.

Dans presque toutes les régressions, c'est la Finlande qui présente le résidu le plus élevé. Ce pays, souvent mis en lumière dans les études PISA, affiche en effet une performance bien supérieure à ce qu'il devrait avoir en tenant compte de son niveau économique. Pour la France, c'est le cas inverse qui se produit. A titre d'illustration, les résultats de la régression entre les mathématiques et le niveau économique laissent supposer une sous-performance d'environ 17 points (colonne 3), une fois le niveau économique pris en compte. Comme souligné précédemment (Avvisati, 2021), cette différence renvoie à l'équivalent de la moitié d'année de scolarisation perdue pour la France. Notons toutefois que même si la « sous-performance » de la France n'est apparue que dans les années 2000, elle semble être latente

¹³ Pour certaines régressions, le nombre de pays est inférieur à 20.

depuis au moins 1970. En effet, le classement des résidus révèle des résultats assez proches si l'on compare le court terme et le long terme (voir les lignes "Période 1970-2020" et "Période 2000-2020"). Notons, par ailleurs, que ces différences de résidus sont basées sur des effets fixes et contrôlées par le niveau économique. Il nous semble donc hautement probable que les explications de l'écart de la France reposent sur des causes plus structurelles que conjoncturelles.

Une approche similaire avec un indicateur alternatif – les dépenses éducatives exprimées en pourcentage du PIB par habitant – permet d'aboutir à des conclusions similaires. En effet, lorsque l'on compare des pays selon leur niveau de dépenses d'éducation, la France accuse un retard vis-à-vis des autres pays de l'OCDE dans le niveau primaire, tandis que la situation est proche du score escompté au niveau secondaire¹⁴.

Tableau 3. Régression de la performance scolaire avec le niveau économique des pays

	(1)	(2)	(3)	(4)	(5)	(5)
PIB/hab	18.067*** (3.303)	23.904*** (2.165)	25.869*** (3.397)	22.803*** (3.398)	16.935*** (3.009)	21.312*** (2.128)
Constante	407.224*** (10.912)	420.160*** (7.737)	403.921*** (11.268)	413.957*** (11.058)	411.378*** (10.640)	413.516*** (7.185)
Niveau	Primaire	Secondaire	Tous	Tous	Tous	Tous
Matière	Toutes	Toutes	Maths	Sciences	Lecture	Tous
R ²	0.403	0.566	0.357	0.347	0.661	0.372
Observations	540	659	397	397	405	1199
Nb pays	18	20	20	20	20	20
Période 1970-2020						
Classement	16	18	18	18	18	18
Résidu	-15.69964	-16.9256	-17.1515	-23.73801	-7.824101	-16.44279
Meilleur classement	Finlande	Finlande	Finlande	Finlande	Royaume Uni	Finlande
Période 2000-2020						
Classement	16	19	17	17	18	18
Résidu	-16.80939	-9.705521	-11.60537	-16.92644	-9.4867	-12.90707
Meilleur classement	Finlande	Finlande	Finlande	Finlande	Finlande	Finlande

Notes : Nous effectuons une régression multiple du niveau économique des pays sur la performance scolaire en données de panel, en incluant des variables indicatrices pour les années. Les résidus pour chaque pays sont ensuite calculés et ordonnés. Le classement des résidus est effectué selon deux périodes distinctes : 1970-2020 et 2000-2020. * significatif à 10%, ** significatif à 5%, *** significatif à 1%. Les erreurs-types sont présentées entre parenthèses.

3.2. Une convergence de la performance éducative inachevée

Les divers résultats des évaluations nationales laissent toutes supposer une diminution de la performance en mathématiques depuis près d'une décennie (Chabanon et Pastor, 2019). Afin de confirmer ou d'infirmer cette tendance, nous nous sommes focalisés sur l'évolution de la performance en mathématiques au niveau primaire, depuis 1970, pour les principaux pays de l'OCDE.

¹⁴ Ces résultats ne sont pas publiés dans cet article, mais disponibles sur demande auprès des auteurs.

Nous remarquons qu'un net décrochage semble effectivement s'opérer en France depuis le début des années 2000, alors que son classement était déjà faible pour les périodes antérieures (voir Tableau 1). Une baisse de la performance est également observée au niveau secondaire sur les vingt dernières années après une hausse assez forte depuis les années 1970. De manière plus significative, il apparaît que la France était nettement au-dessus de la moyenne de l'OCDE-20 en 2000, tandis qu'en 2020, son classement retrouve un niveau plus proche de la moyenne.

Précédemment, nous avons observé une hausse tendancielle sur le long terme de la performance scolaire en France, mais aussi une baisse sur les deux dernières décennies pour tous les domaines de compétence (section 3.1). Il nous semble important de questionner ces variations au regard des autres pays de l'OCDE-20 (Tableau 4). Pour ce faire, nous calculons pour chaque domaine de compétences, ainsi que pour chaque année, l'écart entre la performance de la France et celle de la moyenne des 20 pays de l'OCDE inclus dans notre échantillon. La différence est standardisée et exprimée en termes d'écart-type. Puisque notre base de données a un écart-type de 100, le calcul effectué pour la standardisation s'effectue en divisant l'écart de performance par 100. Alors que la France a un score de 434 en mathématiques au primaire en 1970, contre une moyenne de 482 points pour l'OCDE, l'écart de performance est de 48 points, soit 0,48 écart-type en défaveur de la France. De manière quasi-systématique, la performance de la France est presque toujours inférieure à celle de la moyenne de l'OCDE, sauf en mathématiques au niveau secondaire pour les décennies 1980 et 1990.

Nous calculons également la variation de la performance pour le court (20 années) et le long terme (50 années). Les quatre dernières colonnes de notre Tableau 4 soulignent que la variation de la performance française est assez proche de la moyenne de l'OCDE sur le long terme (deux dernières colonnes), même si les progrès sont légèrement moindres en mathématiques. Sur le court terme, la différence de trajectoire est toutefois nette : alors que la performance diminue dans tous les domaines en France, la variation est plus contrastée au sein des pays de l'OCDE. Le niveau primaire s'améliore en moyenne pour les pays de l'OCDE d'environ 5 points pour décennie, alors qu'une baisse de 8 points est observée en France.

Dans la Figure 2, nous affinons la variation de la performance scolaire. Nous posons l'hypothèse que les pays de l'OCDE ayant un déficit de compétence initial auront une hausse plus soutenue de leur performance dans les décennies suivantes. Pour vérifier cette intuition, nous croisons le niveau initial des mesures éducatives avec la variation de celles-ci sous forme standardisée (i.e. exprimée en écart-types). L'hypothèse de convergence est plutôt confirmée, la pente de la droite de régression étant négative dans la totalité des quadrants. De manière plus précise, rappelons que nous testons ici la convergence de trois indicateurs de l'éducation : un indicateur quantitatif de l'éducation, à savoir le nombre d'années de scolarisation ; l'indicateur d'acquis scolaires et enfin l'indicateur d'années de scolarisation ajusté par le niveau des acquis scolaires (que nous appelons LAYS, voir section 2.2).

Tableau 4. Evolution comparative de la performance scolaire

	1970	1980	1990	2000	2010	2020	Valeur initiale		Valeur finale		Variation court terme		Variation long terme	
							France	OCDE	France	OCDE	France	OCDE	France	OCDE
							Moyenne	-0.36	-0.35	-0.12	-0.08	-0.17	-0.18	442
Primaire	-0.44	-0.43	-0.02	-0.13	-0.25	-0.36	423	467	490	526	-6	5	13	12
Lecture	-0.27	-0.31	0.08	-0.07	-0.17	-0.35	424	450	502	538	-11	3	16	17
Mathématiques	-0.52	-0.45	-0.03	-0.11	-0.30	-0.39	424	476	482	521	-6	8	12	9
Sciences	-0.51	-0.53	-0.16	-0.23	-0.29	-0.35	421	473	487	521	-2	4	13	10
Secondaire	-0.21	-0.20	-0.18	0.02	-0.06	-0.02	460	482	511	513	-7	-5	10	6
Lecture	-0.21	-0.25	-0.18	-0.02	-0.02	-0.06	420	441	488	494	-8	-7	14	11
Mathématiques	-0.11	-0.09	0.02	0.10	-0.06	0.00	489	499	524	524	-8	-3	7	5
Sciences	-0.32	-0.24	-0.38	-0.03	-0.11	-0.01	473	505	521	522	-5	-6	10	3

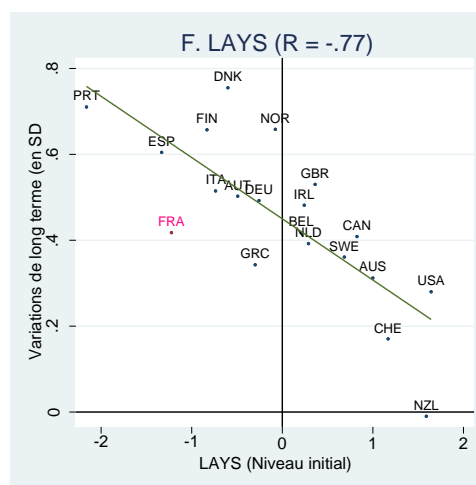
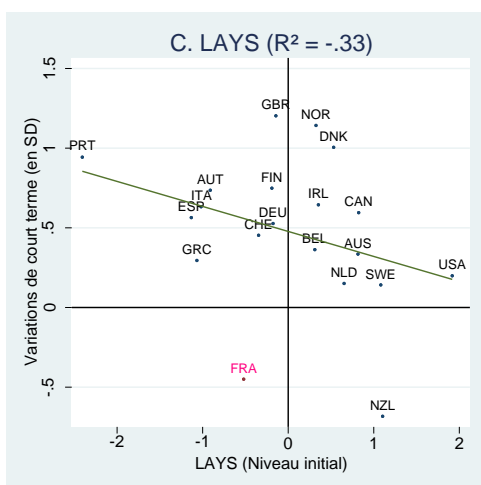
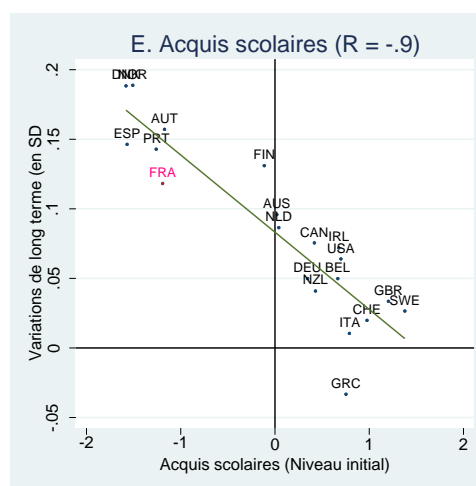
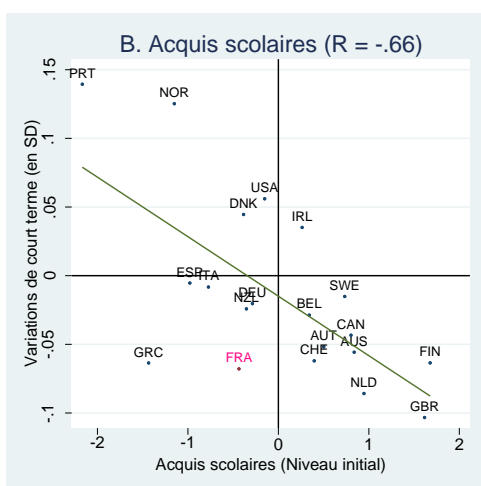
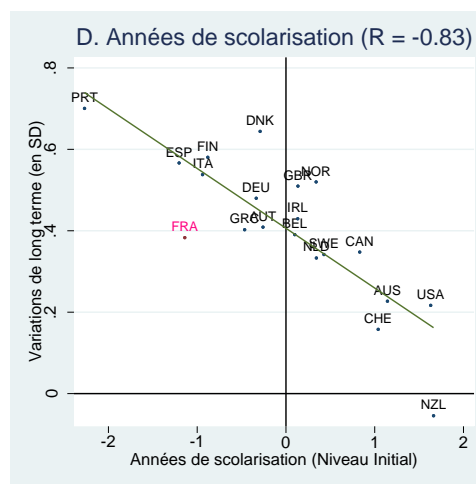
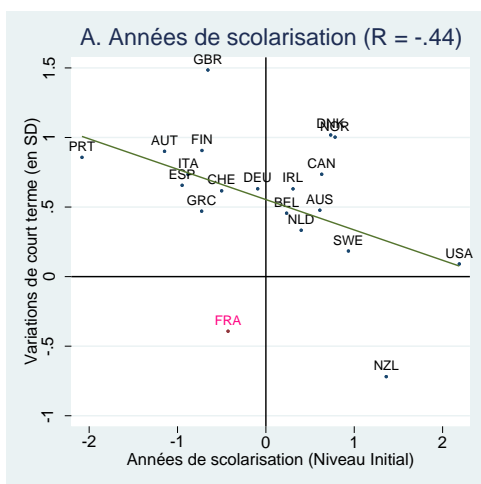
Notes : L'évolution est exprimée en termes d'écart-type de différence de la performance de la France par rapport à la moyenne des 20 pays de l'OCDE inclus dans l'échantillon. On constate qu'en 1970 la performance de la France était de 26 points inférieure à celle des 20 pays de l'échantillon en lecture au niveau primaire (424 contre 450 points). Comme l'écart-type est de 100 points, cela représente une différence de performance d'environ 0,26 écart-type.

La France est marquée par un retard initial dans toutes les dimensions (elle se situe dans la partie négative de l'axe des abscisses). Sous l'hypothèse de convergence, un tel retard devrait conduire à la hausse la performance éducative. Or, cela n'est pas systématiquement le cas, surtout pour les vingt dernières années (2000-2020). A titre d'exemple, le retard français en ce qui concerne les années de scolarisation est loin d'être comblé entre 2000 et 2020, puisque la scolarisation semble diminuer d'environ 0,4 écart-type en France, si l'on prend en compte la variation moyenne des pays de l'OCDE (quadrant A). La performance quantitative ne semble donc pas être à la hauteur de celle observée dans les autres pays de l'OCDE. De manière additionnelle, le retard observé dans le niveau des acquis scolaires en début de période ne semble se résorber que si l'on utilise les données de long terme (valeur des ordonnées positive dans les quadrants D, E et F). Qui plus est, ce retard relatif dans les acquis scolaires semble se renforcer dans le temps depuis les années 2000 (ordonnée négative dans le quadrant B). Notre comparaison entre niveau initial et variation au sujet de l'indicateur hybride LAYS renforce ces résultats : le faible niveau initial est loin de se résorber et a même tendance à se creuser sur le court terme (quadrant C).

Figure 3. Variation de la performance des indicateurs d'éducation

Court terme (2000-2020)

Long terme (1970-2020)



Notes : Le niveau initial représente la valeur de l'indicateur au début de la période indiquée (1970 ou 2000). Les variations sont exprimées en écart-type (*Standard Deviation* = SD) calculés à partir de l'hypothèse que l'écart-type de la base de données est égal à 100 points. Ainsi, un écart de 50 points est équivalent à 0,5 d'écart-type. Source : Altinok et Diebolt (2023).

4. Conclusion

Les décisions d'allocation de ressources à l'éducation reposent, presque toujours, sur des choix politiques tirillés entre des exigences socio-culturelles d'une part, et des exigences économiques d'autre part. Ces choix composent avec d'autres choix portant sur d'autres affectations des ressources nationales. Le caractère concurrentiel de ces choix politiques permet de penser, qu'à un moment ou à un autre, les charges économiques liées à l'éducation rencontrent un plafond. Par ailleurs, au plan économique, on peut admettre, presque intuitivement, l'existence d'une limite au-delà de laquelle développer l'éducation risque de se faire au détriment des possibilités mêmes de la croissance économique. De nos jours, cette limite économique fait, plus que jamais peut-être, figure d'évidence. Cela dit, il n'est pas certain qu'elle ait été pleinement assimilée. Quoiqu'il en soit, nous soutenons l'hypothèse que l'éducation joue un rôle important et positif dans l'activité économique de la France. Ce faisant, l'éducation est elle-même un processus utilisateur de ressources économiques. Ainsi, les pressions démographiques et socio-politiques qui, depuis la fin de la Seconde Guerre mondiale, ont entraîné un développement considérable des systèmes éducatifs ont également atteint un niveau tel que l'éducation figure en bonne place parmi les parties prenantes aux ressources nationales. Dès lors, l'éducation est soumise à la contrainte économique majeure, celle de la rareté relative des ressources susceptibles d'emplois alternatifs. En fait, l'éducation est soumise à une exigence d'efficacité. Cela dit, il ne s'agit nullement d'ériger en principe le caractère utilitariste de l'éducation. Nous pensons uniquement que quelles que soient les finalités assignées à l'éducation, il convient d'utiliser au mieux les ressources disponibles pour les satisfaire.

Comme tel, l'allocation des ressources n'est que partiellement réalisée par des mécanismes d'ajustements par les prix traduisant la rencontre d'une offre et d'une demande sur un marché. Parallèlement, les décisions de production y sont plus ou moins décentralisées et les objectifs particuliers ne rencontrent pas nécessairement l'objectif collectif. Enfin, les besoins en termes d'éducation sont socialement hétérogènes, ainsi que dans le temps et dans l'espace. En ces termes, l'absence de signal par les prix peut affecter l'efficacité interne (à production donnée les ressources ne sont pas nécessairement affectées au moindre coût réel) et l'efficacité externe (les produits ne sont pas toujours adaptés aux besoins).

Afin d'appréhender correctement production et coût de l'éducation, il est donc nécessaire d'élaborer avec précision des indicateurs permettant de décrire ce que l'on produit et avec quels moyens, ceci afin de définir des indices d'efficacité comparée (dans le temps et dans l'espace). Ceci étant, même si l'on suppose le problème des indicateurs résolu, se pose la question de l'élaboration d'outils d'analyse pertinents, susceptibles de fournir une évaluation qualitative de la productivité des dépenses publiques d'éducation. D'une manière générale, nous pensons qu'il est quasi impossible de définir des indicateurs d'efficacité absolus. En revanche, il paraît tout à fait concevable de définir des indicateurs relatifs mesurant des écarts entre quantité et qualité d'éducation. Ce faisant, le problème n'en est pas pour autant résolu, car il conviendra de développer une méthode d'analyse permettant de distinguer les écarts d'efficacité dus à des phénomènes exogènes, de ceux résultant de comportements stratégiques ou enfin, de ceux imputables à de mauvaises allocations de ressources qui favoriseraient la quantité au dépens de la qualité de l'éducation.

Dans le cadre de cet article, notre recherche, consacrée à l'évolution des acquis scolaires en France, a produit de nouveaux éléments de preuve, sur un demi-siècle et en comparaison avec la plupart des pays de l'OCDE, pour aboutir à la conclusion que les performances éducatives, en lecture et en mathématiques diminuent, les scores moyens étant significativement faibles. Par ailleurs, nous avons montré que la trajectoire française en termes d'acquis scolaires tend à diverger de celle des autres pays de l'OCDE. D'une manière générale, nos résultats enrichissent les conclusions des travaux issus des enquêtes nationales tout en livrant des éclairages historiques et comparatifs renouvelés. Enfin, ils invitent à réinvestir, par-delà les considérations centrées sur des indicateurs de pure quantité, l'importance majeure de la dimension qualitative de l'éducation.

Références

Altinok, N. et C. Diebolt (2023). "Cliometrics of Learning-Adjusted Years of Schooling: Evidence from a New Dataset." *AFC Working Papers*, 2, 1-78.

Angrist, N., S. Djankov, P. K. Goldberg et H. A. Patrinos (2021). "Measuring human capital using global learning data." *Nature*, 592(7854), 403-408.

Angrist, N., H. A. Patrinos et M. Schlotter (2013). "An Expansion of a Global Data Set on Educational Quality: A Focus on Achievement in Developing Countries." *The World Bank: Policy Research Working Papers*.

Avvisati, F. (2021). How much do 15-year-olds learn over one year of schooling?, OECD Publishing.

Barro, R. J. et J. W. Lee (2013). "A new data set of educational attainment in the world, 1950–2010." *Journal of Development Economics*, 104, 184-198.

Baudelot, C. et R. Establet (1989). Le niveau monte. Réfutation d'une vieille idée concernant la prétendue décadence de nos écoles. Paris, Éditions du Seuil.

Baudelot, C. et R. Establet (2009). L'élitisme républicain. L'école française à l'épreuve des comparaisons internationales. Paris, Seuil.

Boman, B. (2022). "Vietnam's exceptional educational achievement: a thematic review of the emerging literature." *Discover Education*, 1(1), 14.

Cédelle, L. (2015). "PISA à travers la presse en France." *Administration & Education*, 145(1), 63-67.

Chabanon, L. et J. M. Pastor (2019). "L'évolution des performances en calcul des élèves de CM2 à trente ans d'intervalle (1987-2017)." *Note de la DEPP*, 19 (08), 1-4.

DEPP (2022). L'état de l'école 2022. Une analyse statistique du système éducatif. Paris, Direction de l'évaluation, de la prospective et de la performance.

Filmer, D., H. Rogers, N. Angrist et S. Sabarwal (2020). "Learning-adjusted years of schooling (LAYS): Defining a new macro measure of education." *Economics of Education Review*, 77, 101971.

Foshay, A. W., R. L. Thorndike, F. Hotyat, D. A. Pidgeon et D. A. Walker (1962). Educational Achievements of Thirteen-year Olds in Twelve Countries: Results of an International Research Project, 1959-61, Unesco.

Gust, S., E. A. Hanushek et L. Woessmann (2022). "Global Universal Basic Skills: Current Deficits and Implications for World Development." *National Bureau of Economic Research Working Paper Series*, 30566.

Hanushek, E. A. (2019). Addressing Cross-National Generalizability in Educational Impact Evaluation, National Bureau of Economic Research.

Hanushek, E. A. et D. D. Kimko (2000). "Schooling, labor-force quality, and the growth of nations." *American Economic Review*, 90(5), 1184-1184.

Hanushek, E. A. et L. Woessmann (2015). The knowledge capital of nations: Education and the economics of growth, MIT Press.

Honaker, J., G. King et M. Blackwell (2011). "Amelia II: A program for missing data." *Journal of Statistical Software*, 45(1), 1-47.

Lim, S. S., R. L. Updike, A. S. Kaldjian, R. M. Barber, K. Cowling, H. York, H. J. Taylor (2018). "Measuring human capital: a systematic analysis of 195 countries and territories, 1990–2016." *The Lancet*, 392(10154), 1217-1234.

Martin, R., T. Renault et B. Roux (2022). Baisse de la productivité en France: échec en «maths»? Focus. Paris, Conseil d'Analyse Economique. 091-2022.

Mons, N. et X. Pons (2013). "Pourquoi n'y a-t-il pas eu de «choc PISA» en France? Sociologie de la réception d'une enquête internationale (2001-2008)." *Revue française de pédagogie* (182), 9-18.

OCDE (2019). PISA 2018 Results (Volume I). Paris, Éditions OCDE.

Pritchett, L. (2013). The rebirth of education: Schooling ain't learning, CGD Books.

Schwerdt, G. et S. Wiederhold (2019). "A Macroeconomic Analysis of Literacy and Economic Performance." *mimeo*.

Spaull, N. (2018). "Who makes it into PISA? Understanding the impact of PISA sample eligibility using Turkey as a case study (PISA 2003–PISA 2012)." *Assessment in Education: Principles, Policy & Practice*, 1-25.