



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

depp Direction de l'évaluation,
de la prospective
et de la performance

Le parcours scolaire, les compétences et l'insertion professionnelle des élèves entrés en 6e en 1995

Fabrice MURAT

Série Etudes

Document de travail n° 2024-E26
Décembre 2024

**Le parcours scolaire,
les compétences et
l'insertion
professionnelle des
élèves entrés en 6e en
1995**



Direction de l'évaluation, de la prospective et de la performance

61-65, rue Dutot
75732 Paris Cedex 15

Directrice de la publication

Magda Tomasini

Auteur :

Fabrice MURAT

e-ISSN 2779-3532

SOMMAIRE



➤ Introduction.....	7
➤ Le parcours des élèves	7
Niveau d'études et diplôme.....	8
Age de fin d'études	11
➤ Les compétences.....	13
Les mesures de compétences dans le panel 1995	13
Corrélations entre les évaluations	14
➤ Le lien entre compétences et parcours scolaire.....	16
Le parcours scolaire dépend beaucoup du niveau initial	16
Différentes modélisations du lien entre compétences et parcours scolaires.....	16
Le parcours des bacheliers.....	20
➤ L'insertion professionnelle	21
L'accès à l'emploi	22
Le salaire.....	23
La profession	25
➤ Lien entre parcours scolaire et situation professionnelle.	25
Différentes façons de prendre en compte le parcours scolaire.....	25
Résultats détaillés des modèles	26
Différentes synthèses des évaluations.....	28
Variantes de modélisation du lien entre compétences et insertion professionnelle	29
➤ Comparaison avec les enquêtes Génération et Emploi.....	31
Le plus haut diplôme	31
La profession	32
L'accès à l'emploi	33
Le salaire.....	34
➤ Annexes – le traitement des données	36

Les différentes sources	36
Suivi par la DEPP	36
Suivi par le SIES	36
Suivi par l'Insee	37
Fin des études.....	39
Observation de la situation professionnelle et fin des études	41
Le niveau d'études.....	44
Le diplôme.....	45
Confrontation des niveaux d'études et des diplômes	46
Evaluation de compétences.....	47
Pondération	50
Références des publications de la DEPP	53

➤ Introduction

Ce document de travail apporte un certain nombre d'informations complémentaires par rapport à l'article paru dans *Education & formations*, consacré à l'insertion professionnelle des jeunes, à partir du panel d'élèves entrés en 6^e en 1995¹. Il y a à la fois des analyses qui ne pouvaient entrer dans les limites de l'article, mais aussi des informations concernant les traitements effectués sur les données.

Ce travail s'attache à montrer l'intérêt du panel pour étudier un moment clef de la vie des individus : la fin des études et le début de l'insertion professionnelle. Il s'agit en effet d'une époque importante pour les jeunes. Ils mettent un terme à une période en grande partie consacrée à l'acquisition de connaissances et de compétences. Ils cessent de fréquenter des établissements scolaires ou d'enseignement supérieur et vont développer d'autres formes de sociabilité. La plupart vont entrer sur le marché du travail. Durant cette période, ils vont aussi quitter le domicile parental et éventuellement se mettre en couple. Tous ces changements interviennent dans des calendriers différents d'un jeune à l'autre, sans solution de continuité : les jeunes sortants du système éducatif vont développer d'autres compétences, par leurs activités personnelles et professionnelles ou par la formation continue ; ils ont souvent eu des expériences professionnelles, durant leurs vacances, pour financer leurs études ou dans le cadre de l'apprentissage ; certains ont déjà dû déménager car leur formation se déroulait loin du logement parental, tandis que d'autres, après la fin de leurs études, attendent d'avoir des moyens financiers suffisants pour prendre leur propre logement.

Pour étudier tous ces changements et comprendre l'entrée dans la vie adulte, il est fondamental de disposer d'informations longitudinales. Nous allons montrer l'intérêt du panel d'élèves entrés en 6^e en 1995, suivi par la DEPP (Direction de l'évaluation, de la prospective et de la performance) dans l'enseignement secondaire, par le SIES (Sous-direction des systèmes d'information et des études statistiques) dans l'enseignement supérieur, puis par l'Insee (Institut national de la statistique et des études économiques) dans leur vie adulte, notamment sur le plan professionnel, avec l'enquête EVA (Entrée dans la vie adulte). Cette source fournit des informations riches sur le parcours scolaire et l'insertion professionnelle, mais aussi des facteurs potentiellement explicatifs, comme le milieu social et les compétences à l'entrée de l'enseignement secondaire. Avec ces données, des études ont pu être menées sur la décohabitation (Despalins et Saint-Pol, 2012) ou la santé (Degorre *et al.*, 2009).

Nous allons ici nous centrer sur trois types d'information : des indicateurs de synthèse sur le parcours des élèves, les mesures de compétences disponibles et les données sur l'insertion professionnelle. Nous laisserons en particulier de côté toutes les informations sur le milieu social d'origine des élèves, qui feront l'objet d'analyses ultérieures, les confrontant avec les caractéristiques ici étudiées. Dans le corps du texte, nous donnerons des résultats complémentaires à l'article, les aspects plus techniques, en particulier sur le traitement des données, étant renvoyés en annexe.

Rappelons que le suivi des élèves entrés en 6^e en 2007, avec une organisation proche de celui de 1995, s'est achevé en 2023 et que les traitements et résultats présentés ici pourront donc être en grande partie reproduits sur ces données plus récentes.

➤ Le parcours des élèves

L'article publié est centré sur l'insertion professionnelle, si bien que la description du parcours scolaire y est assez limitée. Il nous semble cependant utile de donner ici plus d'informations, car le panel apporte des informations intéressantes par rapport aux connaissances existantes.

Les analyses d'ensemble sur le parcours scolaire, montrant l'allongement notable des études en France au cours du XX^e siècle, en particulier pour les filles, et la persistance des inégalités sociales se sont appuyées sur des sources diverses. Les statistiques de l'éducation nationale fournissent naturellement des informations sur la part des élèves aux différents de formation et l'âge de fin d'études (Emin et Esquieu, 1998 ; Defresne et Krop, 2016). La série des enquêtes FQP (Formation et Qualification Professionnelle) a aussi été beaucoup utilisée pour étudier le niveau d'études de la

¹ Murat F. (2024) L'insertion professionnelle des jeunes : influence du parcours scolaire et des compétences générales. *Education & formations*, n° 107.

population, en particulier pour mesurer l'évolution des inégalités sociales (Goux et Maurin, 1997 ; Thélot et Vallet, 2000). Sur une période plus récente, de nombreuses études ont été menées à partir des enquêtes Générations du Cereq (Cereq 2008, 2014 ; Barret et al., 2014, pour reprendre des travaux portant sur une période proche de celle du panel 1995) et du bilan Formation-Emploi fondé sur l'enquête Emploi (Insee, 2009 ; DEPP 2010). Ces deux sources ont l'intérêt de confronter le parcours scolaire et l'insertion professionnelle.

Par rapport à ces sources, le panel 1995 a l'avantage de fournir une information très riche sur le parcours scolaire (y compris dans l'enseignement supérieur) et de comporter des évaluations de compétences, donnant une image plus précise de la réussite scolaire. La description du milieu social, obtenue en interrogeant directement les familles, est aussi particulièrement détaillée. Les panels, en particulier celui de 1995, ont donné lieu à de nombreuses études décrivant le parcours des élèves (Caille, 2017). Cependant, les problématiques de ces études sont très marquées par la césure importante qu'est le baccalauréat dans le parcours scolaire. La DEPP publie très régulièrement des travaux sur le parcours dans l'enseignement secondaire, mais très peu de choses sur le parcours dans l'enseignement supérieur. A l'inverse, le SIES se focalise sur le parcours des étudiants, ce qui conduit d'ailleurs à construire un échantillon de bacheliers complémentaire au panel de la DEPP. Pour les élèves entrés en 6^e en 1995, l'arrivée à l'heure en terminale devait se faire à la rentrée 2001, mais cela ne représente qu'une fraction de la cohorte, beaucoup arrivant plus tard. Le SIES a complété l'échantillon avec des bacheliers 2002 en retard, pour constituer une cohorte de bacheliers, plus simple à analyser, sur laquelle a porté une grande part de ses analyses. Sur le panel 1995 proprement dit, une note d'information du SIES fait la synthèse du parcours dans l'enseignement supérieur (Lemaire, 2012) en donnant quelques éléments sur l'ensemble des élèves entrés en 6^e avant de se centrer sur les bacheliers. Cette note prolonge une étude similaire sur le panel 1989 (Lemaire, 2006). La note sur le panel 1995 a été reprise et complétée pour *France Portrait Social 2012* (Lefebvre, 2012).

La plupart des nombreuses publications de la DEPP à partir du panel 1995 se limitent au parcours dans l'enseignement secondaire, au mieux jusqu'au baccalauréat. Quelques fiches dans *RERS (Repères et références statistiques)* incluent l'entrée dans l'enseignement supérieur dans l'analyse : la fiche 4.26 de *RERS 2012*, la plus complète, présente le détail des diplômes des élèves entrés en 6^e en 1995, du DNB (Diplôme national du brevet) au bac +5... Les différentes éditions de *RERS* comportent aussi de nombreuses fiches sur le parcours de l'ensemble de la cohorte dans l'enseignement secondaire et des bacheliers 2002 dans l'enseignement supérieur. L'Insee Première publié pour la première année de suivi de l'enquête EVA (Lemaire et al., 2007) donne quelques éléments à la fois sur la scolarité dans l'enseignement secondaire et dans le début de l'enseignement supérieur, mais à cette date, il n'était pas possible de faire un bilan d'ensemble sur ce deuxième aspect.

Il nous semble donc intéressant de donner une image plus précise de la situation finale en termes scolaires, de l'ensemble de la cohorte. Ce bilan va par ailleurs bénéficier de l'information provenant de l'enquête EVA, permettant de reconstruire le parcours des jeunes que la DEPP et le SIES ont perdus. Cela donnera une image du parcours scolaire des jeunes sortis durant les années 2000.

Pour décrire le parcours scolaire, trois types d'indicateurs peuvent être envisagées :

- le plus haut niveau de formation atteint (rappelons que ce n'est pas forcément le dernier) ;
- le plus haut diplôme obtenu ;
- l'âge de fin d'études.

Ces informations se recoupent fortement, mais donnent des images un peu différentes : un niveau d'études a pu être atteint sans être validé par un diplôme ; le plus haut niveau de diplôme ne donne qu'une vision partielle de la qualification du jeune, qui a pu passer d'autres examens ; un jeune peut finir ses études à un âge avancé, sans avoir obtenu un diplôme élevé, du fait de redoublements et de réorientations. Les données du panel permettent de rendre compte de la variété de ces parcours.

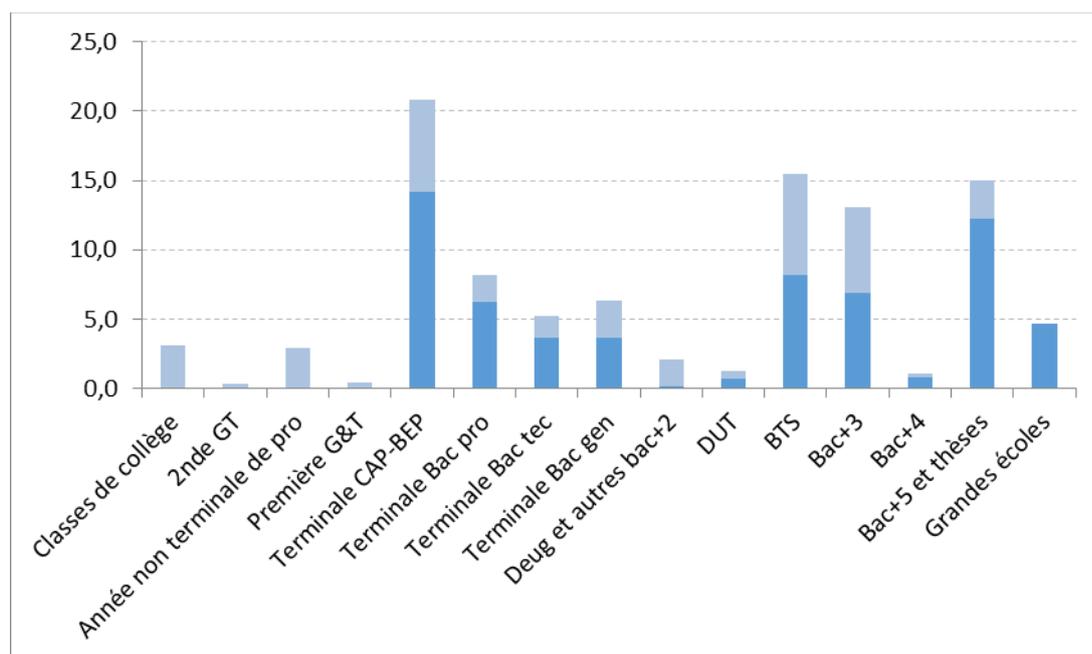
Niveau d'études et diplôme

Environ 7 % des jeunes sortent avant une année de terminale de l'enseignement du secondaire (FIGURE 1 •)². Dans la moitié des cas, cette sortie a eu lieu au cours du collège, les sorties dans l'enseignement professionnel étant aussi assez fréquentes. Par construction, ces jeunes n'ont obtenu

² Cette partie va se fonder sur les 16 246 jeunes (sur les 17 830 suivis au départ) que la DEPP a pu suivre dans l'intégralité de leur cursus secondaire ou qui ont répondu à l'enquête EVA. Une pondération, dont la construction est présentée en annexe, corrige l'attrition.

aucun diplôme (l'obtention du DNB n'est pas prise en compte). Les sorties juste après une terminale sont plus nombreuses, en particulier dans l'enseignement professionnel, qui a effectivement vocation à préparer à une insertion assez rapide sur le marché du travail : 21 % des élèves s'arrêtent au niveau CAP-BEP, 8 % au niveau bac professionnel ; environ 11 % s'arrêtent après une terminale générale ou technologique (à peu près à égalité entre les deux filières). Le taux d'obtention du diplôme est plus élevé pour les sortants d'une terminale professionnelle (77 %), que pour les sortants d'un niveau CAP-BEP, qui sont sans doute des élèves plus en difficulté (68 %). Les taux sont aussi assez bas pour les sortants d'une terminale générale ou technologique (70 % pour les terminales technologiques et 57 % pour les terminales générales), car la vocation naturelle de ces baccalauréats, en particulier du baccalauréat général, est la poursuite d'études dans l'enseignement supérieur. Le fait de sortir à ce niveau n'est pas attendu et s'explique donc souvent par un échec à l'examen.

FIGURE 1 • Plus haut niveau d'études atteint et obtention du diplôme correspondant



Champ : France hors DROM, jeunes entrés en sixième en 1995.

Lecture : chaque barre bleue indique la proportion d'élèves ayant comme plus haut niveau d'études la formation indiquée ; la partie foncée donne la proportion ayant validé le diplôme correspondant.

Source : DEPP-SIES-Insee, enquête panel 1995.

Le même phénomène explique aussi le très faible taux de réussite des jeunes sortant après un DEUG (5 %). Ces jeunes sont d'ailleurs assez peu nombreux, ayant commencé à prendre en compte la réforme LMD (2 %). On compte aussi peu de sortants après un DUT (1 %) et cette sortie peut s'expliquer par un échec (dans 43 % des cas). En revanche, 15 % des jeunes sortent après un BTS, mais celui-ci n'est validé que dans 53 % des cas.

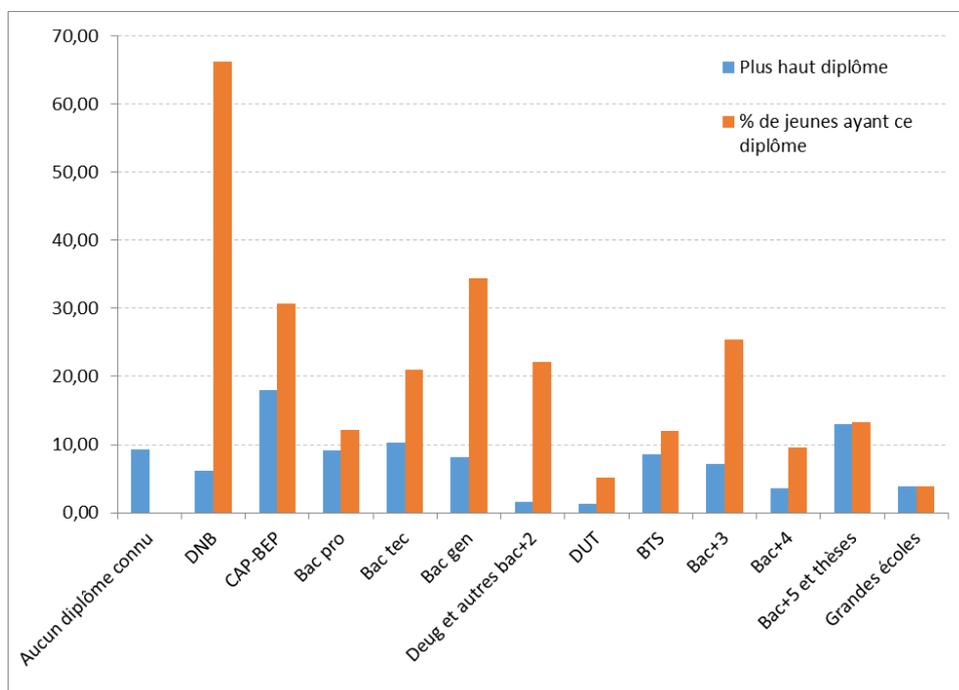
Toujours en cohérence avec la réforme LMD, 13 % des sorties ont lieu après un bac +3 (validé dans 53 % des cas) et seulement 1 % après une maîtrise. Les cursus universitaires longs (bac +5 et thèse) sont assez fréquents (15 % des jeunes) et validés dans 81 % des situations. Il y a aussi 5 % des jeunes qui intègrent une grande école (tous obtiennent un diplôme de ce niveau, soit de leur grande école, soit d'un niveau master).

Même avec le taux de validation, cette approche par niveau ne rend pas compte de la répartition par diplôme : les jeunes dont le plus haut diplôme est le baccalauréat général sont ceux qui s'arrêtent à ce niveau en réussissant l'examen (leur nombre peut être déduit des données précédentes), mais aussi ceux qui ont commencé un DEUG, un DUT ou un STS sans l'obtenir (que l'on ne peut obtenir avec les données précédentes, car les jeunes échouant à ces niveaux, peuvent avoir eu un baccalauréat technologique ou professionnel). Il convient donc de répartir les jeunes selon le plus haut diplôme obtenu (FIGURE 2 •).

Parmi les élèves entrés en 6^e en 1995, 9 % n'ont obtenu aucun diplôme et 6 % n'ont eu que le DNB. Un peu plus de deux jeunes sur cinq ont pour plus haut diplôme un examen de l'enseignement secondaire : 18 % un CAP-BEP, 9 % un baccalauréat professionnel, 10 % un baccalauréat

technologique et 8 % pour le baccalauréat général. Environ un dixième des élèves finissent leurs études avec un diplôme de niveau bac +2 : 5 % un DEUG, 1 % un DUT et 8 % un BTS. Les diplômes de niveau supérieur regroupent 7 % des jeunes pour les bac + 3, 3 % pour les bac + 4, 13 % pour les bac + 5 et 4 % de jeunes passés par une grande école.

FIGURE 2 • Résultats par diplôme



Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Cette approche par le plus haut diplôme masque des cumuls : les diplômés de l'enseignement supérieur ont logiquement un baccalauréat ; un bac +2 est généralement nécessaire pour avoir un bac +3. À l'inverse, si seulement 1 % des jeunes ont un DUT pour plus haut diplôme, cela tient au fait que les titulaires de ce diplôme poursuivent souvent leurs études et obtiennent un diplôme plus élevé. Afin d'affiner l'analyse, pour chaque diplôme, nous avons déterminé les élèves l'ayant préparé : passés par une terminale pour le CAP-BEP et le baccalauréat, entrés dans la filière pour les examens du supérieur (les jeunes débutant un BTS et abandonnant à la fin de la première année seront pris en compte et considérés comme en échec). Pour chaque groupe, on calcule alors la réussite à ce niveau (ainsi, un élève ayant commencé puis abandonné un BTS, pour en recommencer un dans une autre spécialité, qu'il aura validé, sera considéré comme en réussite). En multipliant la proportion d'élèves ayant préparé un diplôme par le taux de réussite, on obtient la proportion d'élèves ayant eu ce diplôme, même si cela n'est pas le plus élevé.

Le cas du DNB est un peu particulier : presque tous les élèves sont censés le préparer (les élèves de 3^e Segpa faisaient alors notablement exception). Cependant, comme cet examen n'est pas indispensable pour la suite du parcours, certains élèves ne le présentent pas ; de plus, pour les élèves perdus par la DEPP, mais retrouvés par l'Insee, le DNB n'est pas forcément le plus diplôme le plus élevé, le seul qu'il doit indiquer ; il est possible aussi que les remontées d'information concernant le DNB aient connu des difficultés. Cela explique sans doute le taux d'obtention assez bas, de 66 % (rappelons que le taux de réussite en 1999, principale année de passation, était de 75 %). Cela reste toutefois le diplôme le plus fréquent de la cohorte.

On compte 40 % des élèves passés par une terminale de CAP-BEP et un taux de réussite de 77 %, si bien que 31 % des élèves entrés en 6^e ont ce diplôme : comparé à 18 % de jeunes pour lesquels il s'agit du plus haut diplôme, cela signifie que 13 % des titulaires d'un CAP-BEP obtiennent un diplôme plus élevé, un baccalauréat professionnel le plus souvent. Sur l'ensemble des élèves de terminale, et pas seulement sur ceux qui s'arrêtent à ce niveau, les taux de validation sont plus élevés que dans le graphique 2 : 80 % pour le baccalauréat professionnel, 89 % pour le baccalauréat technologique, 90 % pour le baccalauréat général, si bien que respectivement 12 %, 21 % et 34 % des jeunes ont les diplômes correspondants (les cas de cumuls entre deux des trois filières sont très marginaux). À

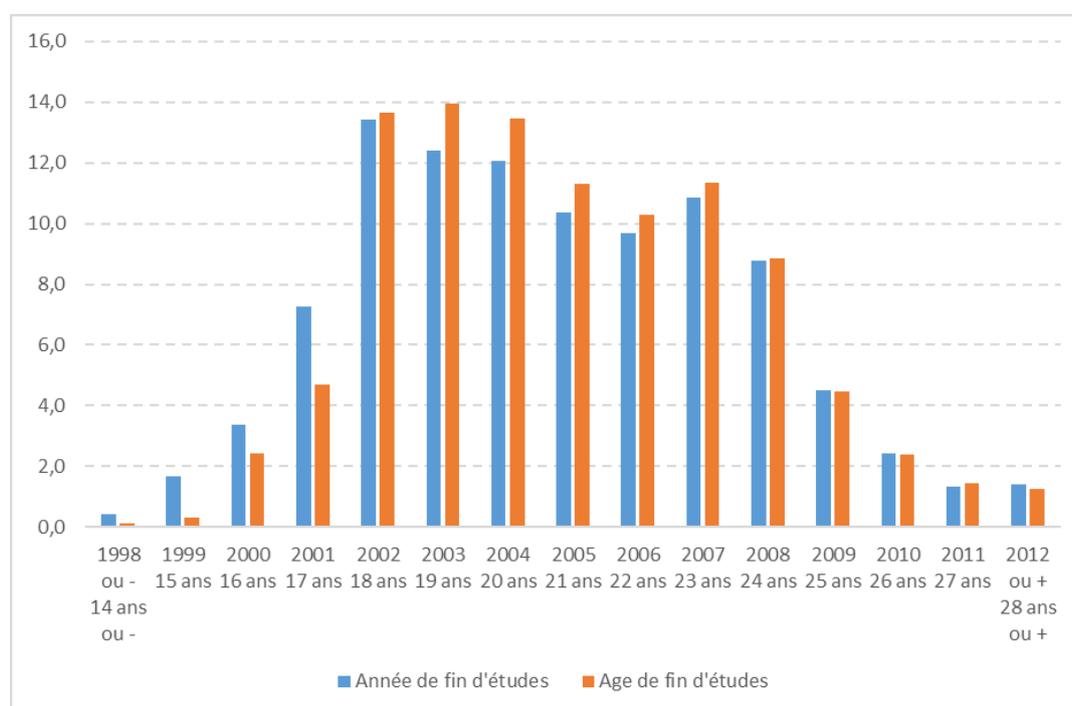
l'entrée de l'enseignement supérieur, on trouve 39 % de jeunes passés par un DEUG ou un autre Bac+2, 7 % par un DUT et 21 % par BTS (les cumuls sont là tout à fait possibles) ; les taux de réussite sont respectivement de 57 % (nettement au-dessus des 5 % du graphique 2, car on prend en compte ceux qui poursuivent leurs études après le DEUG : cette valeur est sans doute plus proche du taux de réussite habituel à cet examen), 71 % et 56 %. Au total, 22 % des jeunes ont un DEUG, 5 % un DUT et 12 % un BTS. L'écart est encore assez net pour les bac +3 et les bac +4 : alors que le bac +3 est le plus haut diplôme de 7 % des jeunes, 25 % ont validé un diplôme de ce niveau (25 - 7 = 18 % ont donc poursuivi leurs études ensuite) ; les proportions sont respectivement 4 % et 10 % pour le bac +4. En revanche, pour les bac +5 et les grandes écoles, les différences sont minimales (elles peuvent venir d'étudiants de grande école validant un master).

Âge de fin d'études

L'âge de fin d'études peut fournir une indication de l'investissement scolaire, qu'il semble simple de comparer dans le temps et entre pays. Sur une génération, il y a équivalence avec l'année de fin d'études ; dans le cas du panel 1995, il faut tenir compte du retard (ou parfois de l'avance) de certains élèves par rapport à l'âge normal à l'entrée en 6^e (par convention, 11 ans au 1^{er} janvier appartenant à l'année scolaire considérée).

Un peu moins de 10 % des élèves entrés en sixième en 1995 sont sortis du système éducatif avant 18 ans (FIGURE 3 •). A l'inverse, presque 10 % des élèves ont poursuivi leurs études au moins jusqu'à 25 ans (1,3 % étaient en fait encore en études initiales lors de la dernière interrogation). La répartition par année d'études est légèrement décalée vers la gauche : il y a plus de jeunes qui ont fini leurs études en 2001 ou avant (12,7 %) que de jeunes qui ont fini leurs études à 17 ans ou moins. Cela tient au fait que ces jeunes arrêtant l'école précocement sont généralement entrés en 6^e avec un an de retard et plus, ce qui majore donc leur âge, quand ils quittent le système éducatif.

FIGURE 3 • Année et âge de fin des études



Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

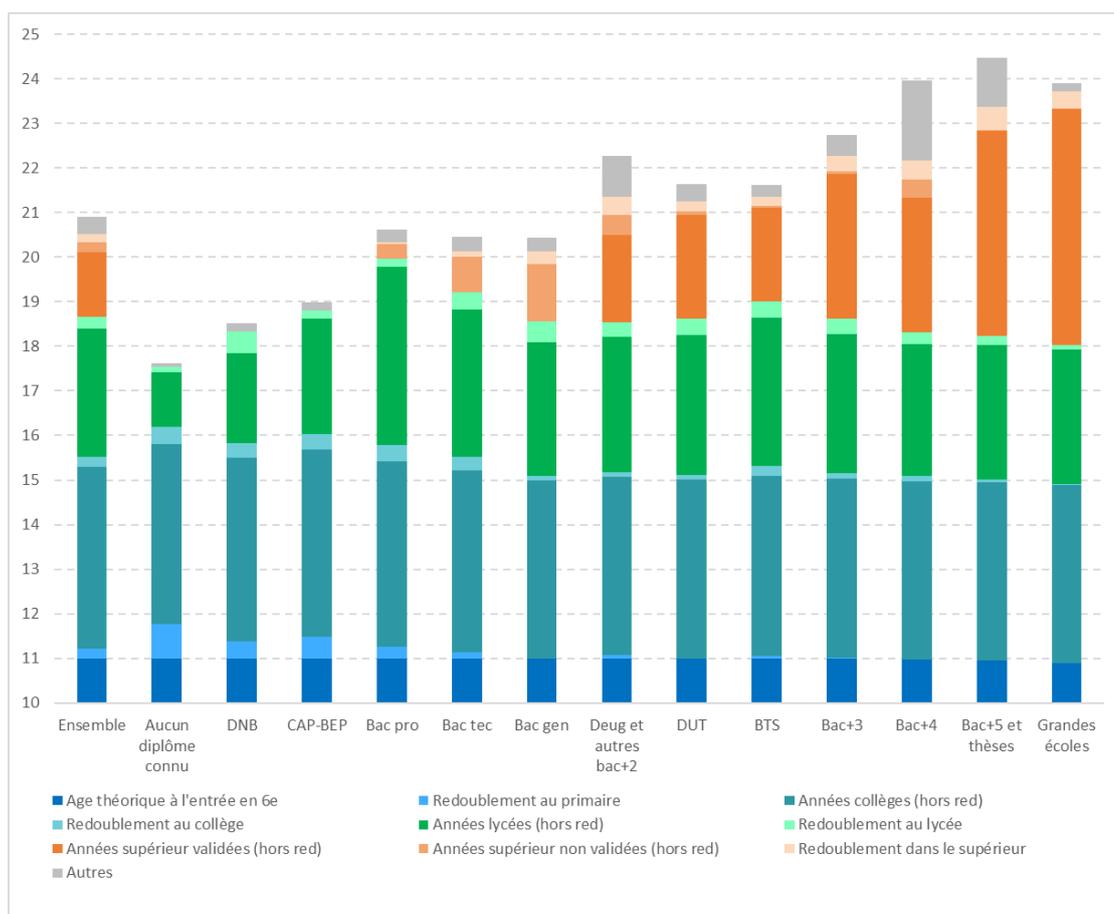
La durée des études peut ainsi rendre compte de l'investissement scolaire de l'élève, mais dépend aussi de certains aléas dans son parcours (redoublements, réorientations). C'est pourquoi il peut être intéressant de confronter l'âge de fin d'études avec le plus haut diplôme atteint et le parcours scolaire. L'âge de fin d'études peut être décomposé en plusieurs parties :

- L'âge théorique à l'entrée en 6^e, qui est fixé à 11 ans pour l'ensemble des élèves (mais dans le cas où l'élève est en avance, nous avons pris l'âge observé) ;

- Le retard pris avant l'entrée en 6^e que l'on obtient en comparant l'âge observé à l'entrée en 6^e avec l'âge théorique ;
- Les redoublements en cours de collège, déterminés par le fait qu'une année de formation au collège est identique à l'année précédente ;
- Le nombre d'années de collèges autres que des années de redoublement
- Les redoublements en cours de lycée, déterminés par le fait qu'une année de formation au lycée est identique à l'année précédente ;
- Le nombre d'années au lycée autres que des années de redoublement
- Les redoublements dans l'enseignement supérieur, déterminés par le fait qu'une année de formation dans l'enseignement supérieur est identique à l'année précédente ;
- Le nombre d'années dans l'enseignement supérieur autres que des années de redoublement, qui ont été validées par un diplôme du niveau correspondant (il peut s'agir de réorientation, par exemple pour une personne qui suit deux BTS différents à la suite) ;
- Le nombre des années dans l'enseignement supérieur non validées (logiquement, toutes si aucun diplôme n'a été obtenu après le baccalauréat) ;
- Les autres années de formations, qui correspondent à des cas très spécifiques (scolarisation en section médico-social, préparation des concours enseignants), mais aussi d'années de formations non connues (par exemple, si le jeune a été perdu par la DEPP et le SIES et récupéré par l'Insee avec un bac + 5 : il n'est pas possible de bien décrire son parcours).

Les écarts d'âge de fin d'études selon le plus haut diplôme obtenu sont assez élevés (FIGURE 4 •). En moyenne, les jeunes finissent leurs études à un peu moins de 21 ans : outre les 11 ans passés avant l'entrée en 6^e, ils ont passé 10 ans après cette date, dont en moyenne un an de redoublement (répartis de façon à peu près équilibré entre les 4 niveaux repérés).

FIGURE 4 • Âge de fin d'études selon le plus haut diplôme



Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Pour les élèves n'ayant obtenu aucun diplôme, la durée des études est logiquement beaucoup plus basse : ils finissent leurs études à un peu moins de 18 ans, dont un peu plus d'une année passée en redoublement, le plus souvent au primaire ou au collège. À l'opposé, les jeunes ayant obtenu un

bac+5 ont en moyenne 24,5 ans à la fin de leurs études et redoublé une année, généralement au lycée ou dans l'enseignement supérieur.

Le cas des jeunes ayant pour plus haut diplôme un baccalauréat général est intéressant à détailler. Alors que l'âge normal à l'issue d'un bac est 18 ans, l'âge de fin d'études observé pour cette population est de 20,5 ans. Un peu plus de trois quarts d'année s'explique par des redoublements et le reste essentiellement par le parcours non validé dans l'enseignement supérieur. Les jeunes sortis avec un DEUG ont en moyenne perdu une demi-année en redoublement et fait une demi-année en bac +3 sans la valider (ces cas sont moins fréquents en STS et IUT). Les élèves de bac +4 passent aussi un peu plus d'un an en redoublements ou années non validées de bac +5. Il s'agit plutôt de minorants, car les « autres années de formation » peuvent aussi comporter des années de ce type. Malgré cette dernière limite, il a semblé intéressant de calculer un âge de fin d'études « corrigé », retirant de l'âge de fin d'études les années de redoublements et les années d'enseignement supérieur non validées, pour mieux rendre compte du cursus des jeunes. Il pourra fournir une meilleure image du parcours scolaire que l'âge observé.

↳ Les compétences

La mesure des compétences des élèves, notamment en mathématiques et en français, fournit des indicateurs de la réussite scolaire complémentaires au parcours dans les différentes formations et à la réussite aux examens. Les panels de la DEPP comportent de plus en plus de données de ce type, le panel 1995 se trouvant dans une situation encore transitoire.

Trois types d'utilisation des évaluations de compétences peuvent être faits, auxquelles les données disponibles dans le panel 1995 vont plus ou moins bien se prêter :

- Le développement des compétences est un objectif en soi. Les scores de compétences peuvent donc être directement considérés comme des indicateurs de réussite scolaire. La place de l'enquête PISA de l'OCDE dans le débat public montre l'importance qu'a prise cette perspective pour mesurer l'efficacité des systèmes éducatifs. Il existe aussi de nombreuses études à partir des évaluations de la DEPP où les résultats aux évaluations sont comparés dans le temps et entre différents groupes (selon le sexe ou l'origine sociale). Cette perspective n'est pas très fréquente avec les panels de la DEPP, qui visent plutôt l'étude de parcours : (Caille et Vallet, 1996) fournit un exemple sur le panel 1989, avec les évaluations nationales à l'entrée en 6^e, disponibles aussi dans le panel 1995, pour étudier la situation des enfants d'immigrés ; (Ben Ali et Vourc'h, 2015) fournit un autre exemple sur le panel 2007, où les compétences en fin de 3^e ont été comparées à celles en fin de 6^e (sur des tests comparables), pour montrer l'impact du milieu social sur la progression des élèves.
- En fait, dans les panels, les évaluations de compétences sont plutôt utilisées comme des éléments déterminants dans l'explication des parcours scolaires. Depuis le panel 1989 et jusqu'au panel 2007, l'importance des compétences à l'entrée en 6^e pour expliquer le parcours scolaire a été montrée dans de nombreuses études (voir [Barhoumi, Caille, 2020] pour un exemple récent). Dans ce cadre, les différences de compétences initiales entre groupes sociaux expliquent en partie les différences d'accès au baccalauréat ou de sorties sans qualification. Cependant, l'explication n'est que partielle et les différences « secondaires », à compétences données, sont intéressantes à analyser, signes de différences d'aspiration ou de confiance en soi.
- Une dernière perspective, proche de la précédente, mais moins fréquente, est de prendre les compétences comme déterminants de phénomènes non scolaires. Les compétences peuvent fournir une mesure du capital humain, complémentaire du diplôme dans l'analyse du marché du travail. Ce sont plutôt les évaluations de compétences auprès des adultes, comme IVQ (Information et Vie quotidienne) ou PIAAC (*Programme for the International Assessment of Adult Competencies*) qui vont se prêter à ce type d'analyses, encore assez rares en France. Sur le panel 1995, on peut toutefois citer le travail de (Le Rhun et Monso, 2015) contrôlant les résultats aux évaluations en 6^e pour étudier l'insertion professionnelle des élèves de BTS (voir l'article pour plus de précisions sur le sujet).

Les mesures de compétences dans le panel 1995

Pour répondre à ces attentes, le panel 1995 comporte plusieurs types d'évaluation :

- L'appréciation des chefs d'établissement sur le niveau en mathématiques et en français à l'entrée en 6^e, sur une échelle de 0 à 10 (signalons que le premier panel, celui de 1962, organisé par l'INED,

comportait déjà une information de ce type, certes subjective, mais, on va le voir, non déconnectée de la réalité);

- Les résultats aux évaluations nationales à l'entrée en 6^e, en mathématiques et en français (ces évaluations passées par tous les élèves à l'entrée en 6^e ont été mises en place en 1989 et arrêtées en 2007 ; elles ont aussi été remontées dans les panels 1989 et 2007) ;
- Les notes de contrôle continu en mathématiques, en français et en langue vivante en 3^e : la collecte s'étant faite assez tôt dans l'année scolaire, les notes à l'examen du brevet n'ont pas été récupérées ; celles-ci sont par contre présentes dans le panel 2007 ;
- Les notes au baccalauréat dans les différentes matières (pour cet examen, comme pour le DNB, ce sont les résultats à la dernière session passée qui ont été pris en compte, en cas de redoublement) ;
- Les résultats à des épreuves spécifiques au panel 1995 en « civisme » (El Jammal *et al.*, 1996) et en « méthodes de travail » (Héraud et Stéfanou, 1997) : les épreuves en « civisme » donne lieu à 3 sous-scores, l'un en connaissances civiques (signification du 14 juillet, procédure de vote des lois...), un deuxième sur le comportement dans le milieu scolaire (opinions sur les délégués, sur le fait de sécher un cours) et un troisième sur des opinions générales (opinions sur la façon dont les lois sont appliquées).

Pour faciliter la comparaison, toutes ces évaluations, sur des échelles très différentes, ont été standardisées à une moyenne 0 sur l'ensemble de la population et un écart-type de 1. Il n'y a donc pas de possibilité d'étudier la progression absolue des élèves sur une même échelle de compétences, au fil de la scolarité. De plus, il y a de nombreuses non-réponses, très souvent non aléatoires : une partie seulement de la population passe le DNB, une part encore plus faible passe le baccalauréat et parmi les candidats au baccalauréat, il y a des différences de niveau entre séries, dont ne rendent pas compte les notes au baccalauréat. C'est pourquoi un important travail d'imputation et de standardisation fine a été fait que l'on va décrire rapidement (plus de détail est donné en annexe) :

- Les non-réponses aux évaluations nationales à l'entrée en 6^e ont été imputées à l'aide des appréciations des chefs d'établissement, disponibles pour tous les élèves : pour les non-répondants aux évaluations nationales en mathématiques (ou en français), on a pris directement l'appréciation en mathématiques (standardisée globalement sur une moyenne de 0 et un écart-type de 1) ; pour les répondants à l'évaluation nationale, le score a été restandardisé sur la moyenne et l'écart-type de l'appréciation observés sur cette population. Le score après imputation a alors bien une moyenne de 0 et un écart-type de 1 sur toute la population.
- La même procédure a été utilisée pour les notes en 3^e (mathématiques, français, LV1) en utilisant les scores aux évaluations nationales dans la même discipline (éventuellement imputés).
- Pour les notes aux baccalauréats, la procédure est un peu plus compliquée :
 - pour la note globale (avant oral de rattrapage), en français, en LV1 et en mathématiques, le recalage sur les non-répondants, se fait sur les notes en 3^e ; pour les répondants, le recalage se fait pour chaque série indépendamment (on prend en compte le fait que les élèves de terminale S ont en moyenne en 3^e des notes plus élevées que les élèves de terminale professionnelle, de façon plus nette que ce que suggèrent les notes au baccalauréat) ;
 - pour les autres notes (histoire-géographie, philosophie, physique-chimie, sciences de la vie et de la Terre, économie) afin de ne pas apporter d'informations redondantes, la valeur pour les non-répondants est la même pour tous (la moyenne de cette population sur la note totale, en mathématiques ou en français, selon la discipline).

Corrélations entre les évaluations

La comparaison entre les évaluations doit donc être faite avec prudence (les unes ont servies à renseigner les autres) et il sera surtout intéressant d'avoir une image globale des corrélations et de comparer les coefficients entre les mathématiques et le français (qui n'ont pas fait l'objet d'imputations croisées).

Les corrélations entre les différentes évaluations apparaissent assez élevées (de l'ordre de 0,7 pour une même discipline, entre différents moments de mesure), mais cela tient souvent en grande partie au processus d'imputation (TABLEAU 1 •).

À l'entrée en 6^e, la corrélation entre l'évaluation des chefs d'établissement et l'évaluation nationale est de 0,77 en mathématiques et de 0,74 en français (enlever les imputations fait baisser de seulement 0,02 ces valeurs). Les évaluations en français et en mathématiques sont aussi assez fortement liées entre elles : 0,75 pour les appréciations des chefs d'établissement, 0,74 pour les évaluations nationales. Ainsi, même si les appréciations des chefs d'établissement, comme les notes

de contrôle continu, dépendent sans doute du contexte social et scolaire du collège et de la classe, elles donnent une image des compétences des élèves assez cohérente avec les évaluations standardisées. Ceci dit, les évaluations à l'entrée en 6^e avaient justement pour objectif de faire un diagnostic pour l'enseignant des forces et des faiblesses des élèves à l'entrée du collège et ont pu être utilisées pour remplir les appréciations.

Les corrélations avec les notes de contrôle continu de 3^e paraissent aussi assez nettes : 0,75 entre l'évaluation nationale et la note en mathématiques ; 0,76 en français. Cependant, cela tient en partie au processus d'imputation : en enlevant les imputations, les corrélations passent respectivement à 0,60 et 0,61, ce qui demeure assez correct, étant donné la réputation de subjectivité qu'ont les notes de contrôle continu. Manifestement, elles ont tout de même une bonne cohérence avec les évaluations standardisées, ce qui est d'autant plus notable que 3 années d'apprentissage les séparent.

TABLEAU 1 • Corrélations entre les différentes évaluations

			1	2	3	4	5	6	7	8	9	10	11	12	13
Evaluations chef d'étab	M	1	1	0,75	0,76	0,65	0,55	0,49	0,20	0,21	0,25	0,30	0,23	0,31	0,30
	F	2	0,75	1	0,64	0,72	0,47	0,58	0,15	0,27	0,25	0,30	0,27	0,31	0,32
Evaluations 6e	M	3	0,77	0,65	1	0,76	0,60	0,52	0,24	0,23	0,28	0,35	0,26	0,34	0,31
	F	4	0,66	0,74	0,76	1	0,49	0,61	0,18	0,31	0,26	0,34	0,30	0,35	0,32
Notes de 3e	M	5	0,65	0,56	0,75	0,63	1	0,67	0,27	0,23	0,30	0,22	0,16	0,24	0,20
	F	6	0,58	0,67	0,65	0,76	0,74	1	0,19	0,36	0,30	0,23	0,22	0,26	0,26
Notes au bac	M	7	0,61	0,53	0,71	0,60	0,84	0,66	1	0,19	0,58	0,07	0,03	0,06	0,02
	F	8	0,54	0,61	0,61	0,71	0,67	0,84	0,66	1	0,44	0,14	0,09	0,14	0,11
	T	9	0,62	0,62	0,70	0,70	0,79	0,79	0,85	0,81	1	0,14	0,08	0,13	0,09
Evaluations spécifique	CV1	10	0,30	0,30	0,35	0,34	0,29	0,29	0,29	0,30	0,33	1	0,21	0,30	0,21
	CV2	11	0,23	0,27	0,26	0,30	0,23	0,28	0,22	0,26	0,26	0,21	1	0,32	0,38
	CV3	12	0,31	0,31	0,34	0,35	0,29	0,31	0,28	0,30	0,32	0,30	0,32	1	0,22
	MT	13	0,30	0,32	0,31	0,33	0,28	0,32	0,25	0,29	0,30	0,21	0,38	0,22	1

Note : la partie du bas donne les corrélations entre les variables après imputation et standardisation ; celle du haut donne les corrélations avec les variables brutes.

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Les corrélations pour les notes au baccalauréat sont beaucoup plus sensibles au processus de standardisation : les corrélations peuvent sembler élevées (0,84 par exemple entre la note de contrôle continu de 3^e et la note à l'examen au baccalauréat, en mathématiques), mais cela tient en très grande partie aux imputations et aux standardisations. La corrélation tombe ainsi à 0,27 entre la note de contrôle continu et la note au baccalauréat pour les mathématiques et à 0,36 pour le français, quand on revient aux données brutes. Ceci dit, si l'on raisonne à série fixée, par exemple pour les élèves de terminale S, la corrélation remonte à 0,42 en mathématiques et 0,44 en français. Sans être parfait, le lien entre la notation en 3^e et celle au baccalauréat paraît non négligeable, quand on compare les élèves ayant passé des épreuves comparables (il faudrait aussi tenir compte du fait que les élèves arrivent au baccalauréat sur plusieurs sessions). Toutefois, il faut être conscient que dans nos données, du fait des imputations et du recalage par série, ces deux évaluations vont être très redondantes, de façon donc un peu mécanique.

La corrélation entre les scores non cognitifs et les évaluations en français et en mathématiques ne dépasse pas 0,34 (avec les évaluations à l'entrée en 6^e faites à la même date). Ceci dit, les corrélations entre ces scores, même au sein du domaine civisme ne sont non plus très élevées (0,3). Ce sont des domaines particuliers, difficile à évaluer, avec donc une erreur de mesure dégradant les corrélations.

Comment utiliser ces mesures, dans le cadre des trois objectifs décrits plus haut ?

– En ce qui concerne l'utilisation des évaluations comme mesure de la réussite scolaire « à expliquer », par exemple en fonction du milieu social, il est bien sûr possible de considérer chacune d'entre elles comme une image intéressante des compétences, mais dans notre perspective qui est de décrire la situation des jeunes à la fin de leurs études, elles ont l'inconvénient de ne pas dépasser le baccalauréat, les notes au baccalauréat ayant l'inconvénient d'être très hétérogènes (du fait de la non passation de certains jeunes, des différences entre séries, du fait que les élèves l'ont passé au cours de différentes sessions). Pour se faire une idée des compétences des jeunes à

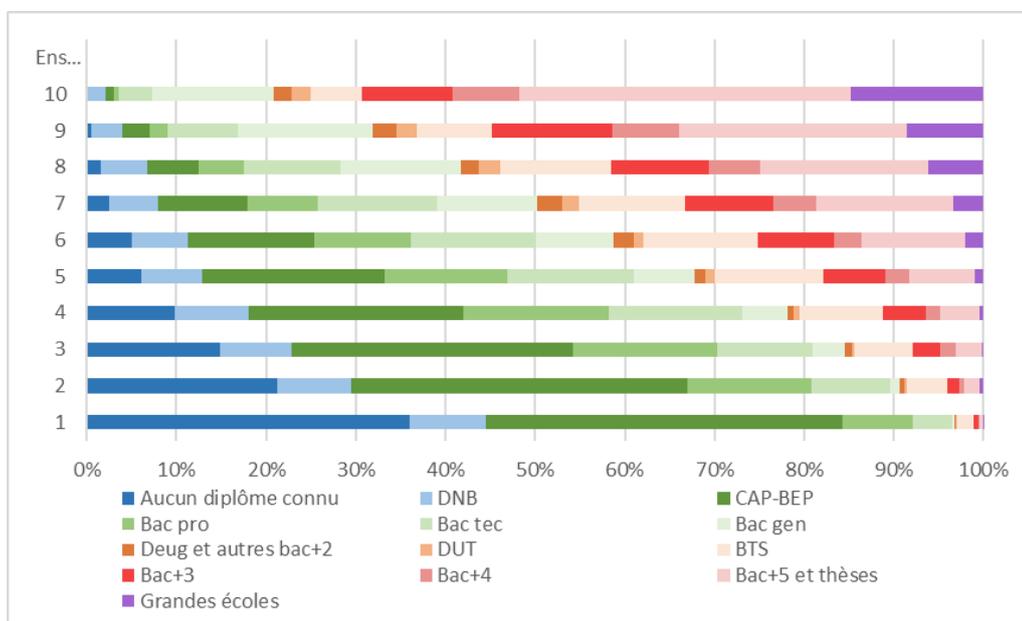
la fin de leurs études, il est préférable d'utiliser les évaluations de compétences des adultes, IVQ et PIAAC, en se restreignant aux sortants (Murat, 2021 ; Murat, 2022).

- Dans l'idée d'étudier le parcours scolaire, le plus pertinent est de se restreindre aux mesures de compétences à l'entrée en 6^e (y compris les appréciations des chefs d'établissement). En effet, celles en 3^e et surtout au baccalauréat dépendent de ce parcours, ce qui rend la corrélation plus compliquée à interpréter. Dans certains cas, les évaluations pourront être synthétisées en un score global.
- Dans l'idée d'étudier l'insertion professionnelle, il semble souhaitable de prendre la mesure des compétences la plus tardive possible, car la plus proche de cette insertion. Les notes au baccalauréat seraient alors les plus indiquées. Cependant, nous avons indiqué les limites de ces notes. En combinant des sources multiples, on atténue les problèmes d'erreurs de mesure et on fiabilise la mesure, même si cela a l'inconvénient de mélanger des évaluations de natures diverses et à des dates différentes.

↳ Le lien entre compétences et parcours scolaire

Le parcours scolaire dépend beaucoup du niveau initial

FIGURE 5 • Plus haut diplôme en fonction des compétences à l'entrée en 6^e



Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

En accord avec les nombreuses études qui ont montré que le parcours dans l'enseignement secondaire varie nettement selon les résultats aux évaluations à l'entrée en 6^e, le lien entre les compétences et le plus haut diplôme obtenu est fort. Les élèves ont été classés en 10 groupes d'ordre croissant de réussite selon leurs compétences à l'entrée en 6^e (en utilisant les appréciations des chefs d'établissement et les scores aux évaluations nationales, en français et en mathématiques) et leur plus haut diplôme a été confronté à ce classement (FIGURE 5 •). Les jeunes qui se trouvaient dans le premier décile de compétences à l'entrée en 6^e sont 36 % à n'obtenir aucun diplôme et un peu moins de 10 % n'obtiennent que le DNB ; dans 40 % des cas, ils ont eu un CAP ou un BEP ; les baccalauréats professionnels et technologiques représentent respectivement 8 % et 4 % ; ils sont très peu nombreux à obtenir un baccalauréat général ou à avoir validé des études supérieures. Les jeunes qui se trouvaient dans le meilleur décile de compétences à l'entrée en 6^e ont un destin scolaire très différent : 0,1 % n'obtiennent aucun diplôme et moins de 2 % s'arrêtent avec uniquement le DNB ; l'entrée dans l'enseignement supérieur concerne plus de 80 % d'entre eux et est validée par un Bac+5 dans 37 % des cas et par un diplôme de grande école dans 15 % des cas.

Différentes modélisations du lien entre compétences et parcours scolaires

Une façon plus globale de mesurer le lien entre compétences et parcours scolaire consiste à modéliser le niveau d'études en fonction des compétences, pour voir dans quelle mesure celles-ci

prédisent le plus haut niveau atteint. Pour cela, nous avons retenu plusieurs indicateurs synthétiques du parcours scolaire :

- L'âge de fin d'études
- L'âge de fin d'études corrigé, c'est-à-dire en excluant les années de redoublement et les années non validées par un diplôme
- Le diplôme, que l'on considérera sous trois formes :
 - la répartition dans 11 catégories (les 3 diplômes de niveau bac +2 ont été regroupés ; en revanche, on a maintenu la distinction des 3 baccalauréats en les hiérarchisant selon les chances qu'ils donnent de faire des études longues : professionnel, technologique, général) ;
 - une version « quantifiée » : à chaque diplôme est attribuée l'âge de fin d'études théorique (16 ans pour une absence de diplôme, 17 ans pour le DNB, le CAP et le BEP, 18 ans pour le baccalauréat généraux et technologique, 19 ans pour le baccalauréat professionnelle, 20 ans pour les bac +2 et ainsi de suite dans l'enseignement supérieur) ; utiliser l'âge moyen de fin d'études corrigé observé donne des résultats très proches (divergence de moins d'un point des indicateurs) ;
 - une simple distinction entre les bacheliers et les non-bacheliers (conforme à la plupart des analyses menées sur le panel).

Pour les deux premiers indicateurs et la version « quantifiée » du diplôme, c'est le R^2 de la régression de ces indicateurs par les évaluations qui est utilisé pour mesurer l'ampleur de la corrélation ; pour la version brute du diplôme, c'est le pseudo- R^2 de Cow & Snell de la régression logistique polytomique ordonnée. Il s'agit bien sûr d'une mesure approximative de la corrélation réelle entre les compétences et le parcours scolaire : les erreurs de mesure sur les évaluations (et probablement aussi sur les indicateurs de parcours scolaire) atténuent les corrélations ; à l'inverse, il existe des variables à la fois corrélées avec les compétences et avec le parcours, dont la non-prise en compte biaise à la hausse les corrélations. Ainsi, les compétences à l'entrée en 6^e sont liées au retard scolaire, car les élèves ayant redoublé à l'école primaire n'ont pas atteint le niveau des autres élèves ; or, ce retard peut en lui-même avoir un impact sur le parcours scolaire, à niveau de compétences donné ; cet effet se trouve en partie intégré dans les R^2 de nos modèles. Le même raisonnement peut être fait avec l'origine sociale, liée aux compétences initiales, mais ayant aussi une influence sur les parcours, à compétences fixées. L'objectif de ce travail est donc d'avoir une vision descriptive de la relation entre compétences et parcours scolaire (dont des travaux ultérieurs préciseront l'ampleur réelle) et surtout de comparer différentes spécifications pour trouver la plus efficace et la plus parcimonieuse.

Les résultats à l'entrée en 6^e, en français et en mathématiques, estimés par les chefs d'établissement et mesurés par l'évaluation nationale, avec les évaluations spécifiques en civisme et méthodes de travail, permettent de prédire 21,0 % de la variance de l'âge de fin d'études brut, 27,5 % de l'âge corrigé et 34,5 % du diplôme quantifié (0). C'est un niveau assez élevé, sachant que la réussite au diplôme dépend des compétences dans d'autres disciplines. La correction de l'âge de fin d'étude paraît pertinente puisqu'elle augmente sensiblement la corrélation avec les scores initiaux. Le diplôme quantifié paraît encore mieux rendre compte des compétences que cet âge corrigé et, dans la mesure où le pseudo- R^2 peut être comparé au R^2 , le diplôme dans le détail des 13 groupes est encore mieux corrélé aux compétences initiales (41,0 %) ³. Ce modèle permet de la même façon de bien prédire l'obtention d'un baccalauréat (41,0 %), ce qui cohérent avec les nombreuses études ayant montré l'importance du niveau initial dans la scolarité secondaire.

³ Le diplôme quantifié a été modélisé sous forme de régression polytomique ordonnée pour vérifier ce point : les pseudo- R^2 ainsi obtenus sont de 2 à 3 points supérieurs au R^2 (37,7 % par exemple pour le premier modèle) ; une partie de l'avantage du diplôme détaillé s'explique par l'indicateur de corrélation retenu, mais le reste peut être interprété comme une plus grande pertinence de cette mesure.

TABLEAU 2 • Corrélations entre la qualification finale et les évaluations (R² ou pseudo-R²)

	Age de fin d'études	Age de fin d'études corrigé	Diplôme quantifié	Diplôme	Bachelier	Diplôme quantifié
	R ²			Pseudo-R ²		
Evaluations en 6 ^e	21,0	27,5	34,5	41,4	41,0	37,7
Evaluations en 6 et 3 ^e	24,9	32,2	39,5	47,3	46,7	43,2
Evaluations en 6 ^e , 3 ^e et bac	37,4	43,5	54,3	62,5	70,4	55,2
Score global en 6 ^e	20,0	26,0	32,3	38,5	38,6	35,0
Score global en 6 ^e (specif. quad.)	20,1	26,4	33,1	38,7	38,7	35,0
Note globale standardisée au bac	29,5	35,7	44,5	53,6	56,6	49,6
Note glob stand. au bac (specif. quad.)	30,2	37,6	47,0	54,5	57,4	50,2
Note moyenne standardisée au bac	36,3	42,3	52,5	60,4	67,0	52,7
Note moy. stand. au bac (specif. quad.)	37,9	43,0	53,5	62,4	67,0	54,3
Moyenne de toutes les évaluations	30,5	37,9	46,9	54,8	55,5	49,1
Moy. toutes évaluations (specif. quad.)	30,6	38,4	47,4	54,9	55,5	49,2
Note glob. stand. au bac (1 ^{re} session)	28,4	36,1	44,4	52,7	53,8	49,0
Note glob. std. au bac (après rattrapage)	29,8	35,4	44,1	53,9	58,0	49,5
Note brute au bac	4,9	10,7	17,4	17,7	67,4	22,0
Note brute au bac (après rattrapage)	5,5	10,4	17,3	19,7	73,9	22,6

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

La prise en compte des notes en 3^e permet encore d'améliorer la qualité des modèles (39,5 % pour le diplôme « quantifié »). Enfin, si l'on utilise l'ensemble des évaluations, les corrélations sont encore plus forte (54,3 % pour le diplôme « quantifié »). Dans ce cas, cependant, comme la mesure intègre des évaluations de compétences en cours de parcours, l'interprétation est plus délicate (logiquement, cette mesure fondée en grande partie sur les notes au baccalauréat prédit très bien l'obtention de ce diplôme, avec un R² de 70 %, le décalage s'expliquant par exemple par la non prise en compte des notes de rattrapage) et il s'agit plutôt de montrer dans quelle mesure, lors des études sur l'insertion professionnelle, les évaluations de compétences vont fournir un indicateur du capital humain, complémentaire du diplôme tout en lui étant fortement lié⁴.

Nous avons proposé différentes synthèses des évaluations :

- La moyenne de toutes les évaluations à l'entrée en 6^e (qui a servi dans la FIGURE 5 •) prédit presque aussi bien le parcours scolaire que le modèle avec le détail des résultats : le R² vaut 32,3 % pour le diplôme quantifié contre 34,5 % et une spécification quadratique améliore encore les choses (33,1 %).
- Parmi les différentes synthèses intégrant les évaluations jusqu'au baccalauréat, c'est la moyenne des notes dans les différentes matières du baccalauréat qui est la plus efficace : pour le diplôme quantifié, le R² vaut 52,5 % en spécification linéaire et 53,5 % en spécification quadratique, soit proche du 54,3 % du modèle utilisant le détail de toutes les évaluations. C'est mieux que la moyenne « officielle » avant rattrapage (R²=47 % en spécification quadratique) ou la moyenne de

⁴ A titre de comparaison, un travail similaire ayant croisé les scores à des tests en lecture et en calcul et le plus haut diplôme obtenu, pour les jeunes venant de finir leurs études, à partir de l'enquête IVQ, a donné une part de variance commune de 33 % (Murat, 2021). C'est moins que dans notre travail, alors que les deux indicateurs sont recueillis de façon simultanée, sans doute parce que les tests d'IVQ sont assez courts et ne permettent pas une mesure aussi fine des compétences que les données des panels.

toutes les évaluations disponibles ($R^2=47,4$ % en spécification quadratique). La note moyenne que nous privilégierons dans la suite a l'avantage de se centrer sur les disciplines les plus importantes du baccalauréat à un moment important pour expliquer le parcours dans l'enseignement ; elle bénéficie aussi de l'information par les évaluations en 6^e et en 3^e, par la standardisation des séries et l'imputation pour les non-candidats au baccalauréat.

- Nous avons aussi testé la sensibilité des résultats à certains choix lors de la construction des données : en cas de redoublement, retenir les résultats de la première session et non la dernière change très peu les résultats et plutôt à la baisse ($R^2=44,4$ % pour le diplôme quantifié en spécification linéaire contre 44,5 % pour la note à la dernière session) ; de même, prendre les notes après l'oral de rattrapage dégrade plutôt le modèle ($R^2=44,1$ %), à part pour prédire la réussite au baccalauréat ($R^2=58$ %) ; enfin, les notes brutes au baccalauréat, non standardisées, ont un pouvoir explicatif nettement plus faible (R^2 un peu supérieur à 17 %), là encore, à l'exception logique de la réussite au baccalauréat.

TABLEAU 3 • Modélisation du diplôme quantifié selon les évaluations à l'entrée en 6^e

Variable		Coefficient
R ²		34,5%
Constante		19,08***
Appréciation du chef d'établissement	Lecture	0,11***
	Français écrit	0,15***
	Français oral	0,06**
	Mathématiques	0,24***
Évaluations nationales	Mathématiques	0,46***
	Français	0,36***
Évaluations spécifiques	Connaissances civiques	0,13***
	Vie scolaire	0,01
	Opinions	0,12***
	Méthodes de travail	0,1***
R ²		32,3%
Constante		19,08***
Score global en 6 ^e		1,32***
R ²		33,1%
Constante		18,93***
Score global en 6 ^e		1,38***
Score global en 6 ^e au carré		0,15***

Champ : France hors DROM, jeunes entrés en sixième en 1995.

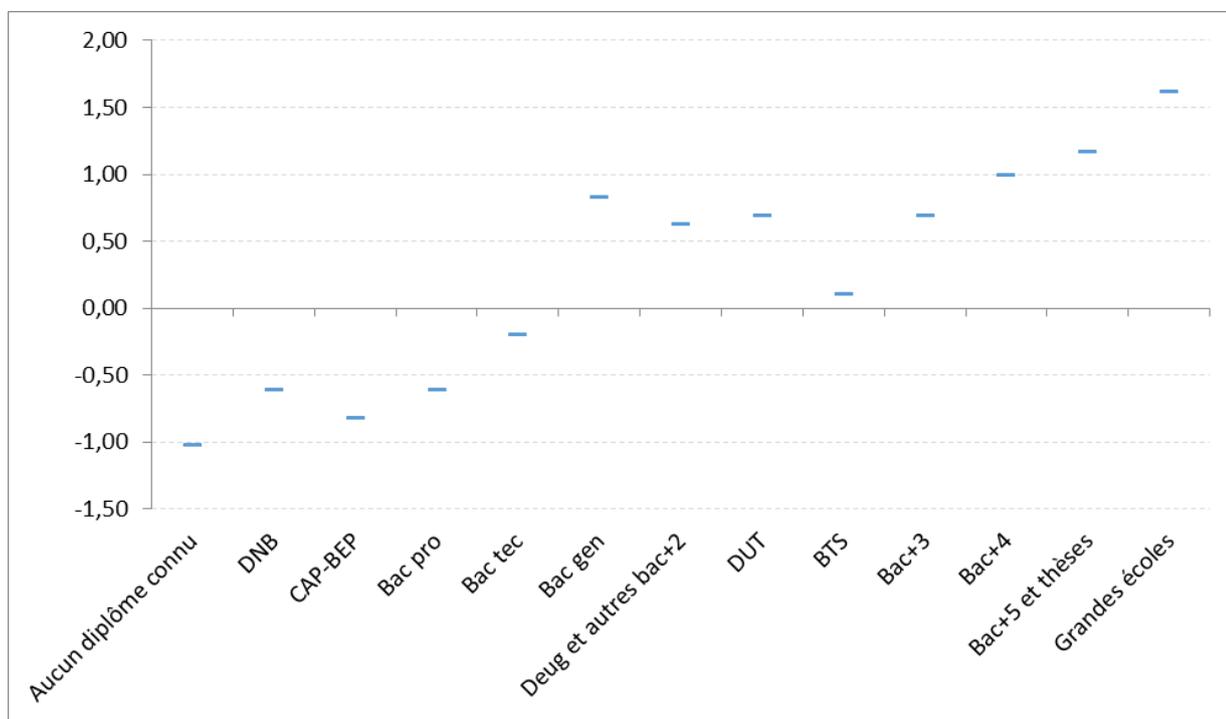
Source : DEPP-SIES-Insee, enquête panel 1995.

Le R^2 ne présente qu'une vision globale de la relation entre les compétences et le parcours scolaires. Les coefficients des modèles donnent une information plus fine sur la forme de cette relation (TABLEAU 3 •). Le modèle linéaire reliant le diplôme quantifié avec le score global à l'entrée en 6^e donne un coefficient de 1,32 : un écart-type de score (rappelons qu'un écart-type de note sur 20 vaut entre 4 et 5 selon les disciplines) implique des études plus longues (en tout cas en termes de validation) de 1,32 année. Le coefficient positif du score élevé au carré signifie que l'écart est encore plus grand pour les élèves les plus compétents au départ.

Si l'on détaille les résultats pour les différentes évaluations disponibles à l'entrée en 6^e, les résultats aux évaluations nationales se voient affectés de coefficients plus forts que les appréciations par les chefs d'établissement et les mathématiques plus que le français : le coefficient pour le score à l'évaluation nationale en mathématiques est de 0,46 contre 0,24 pour l'appréciation en

mathématiques et 0,36 pour le score à l'évaluation nationale en français. Les résultats aux épreuves spécifiques du panel ont aussi un impact non négligeable sur la scolarité (avec des coefficients de l'ordre de 0,1 année par écart-type). Les deux disciplines les moins prédictives sont l'appréciation en français oral et le score de vie scolaire.

FIGURE 6 • Score global moyen aux évaluations selon le plus haut diplôme obtenu



Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Il peut être intéressant de présenter les écarts de score global en fin d'études secondaires (moyenne des notes standardisées au baccalauréat) entre diplômes. Ainsi, les jeunes qui n'obtiennent aucun diplôme ont un score global moyen à plus d'un écart-type en-dessous de la moyenne (- 1,02) (FIGURE 6 •). Les scores sont supérieurs pour les élèves avec un CAP-BEP (- 0,82) et pour ceux ayant eu le DNB uniquement ou un baccalauréat professionnel (environ - 0,6). Les bacheliers technologiques n'ayant pas poursuivi dans l'enseignement supérieur se situent à la moyenne globale et les bacheliers généraux dans la même situation à + 0,83. Une hiérarchie assez nette se dessine aussi dans l'enseignement supérieur : les BTS sont à + 0,11, les DEUG à + 0,63 et les BAC + 3 à + 0,69, les DUT à + 0,69, les bac + 4 à + 0,99 et les bac + 5 à + 1,17, les sortants de grande école se trouvant à + 1,62. Ces résultats sont assez conformes à la hiérarchie attendue des diplômes : le point le plus troublant est le score élevé des bacheliers généraux, par rapport aux bac + 2 (plus élevés que les titulaires d'un BTS notamment). Nous avons donc vérifié que la même hiérarchie était observée sur les scores à l'entrée en 6^e. Les sortants avec uniquement un baccalauréat général semble bien une population atypique.

Le parcours des bacheliers

Par ailleurs, il est possible de reprendre une analyse « prédictive » avec les scores en fin de scolarité secondaire, si l'on se restreint aux bacheliers et que l'on étudie leur scolarité ultérieure (TABLEAU 4 •). Cela pourra être relié aux études sur le parcours des bacheliers suivis par le SIES, dans lesquelles la mention au baccalauréat est souvent utilisée.

Les résultats à l'entrée en 6^e (en incluant les évaluations spécifiques) permettent de prédire une part significative, mais relativement modeste, du diplôme « quantifié » des bacheliers (13,6 %). Le modèle s'améliore assez sensiblement ajoutant les notes en 3^e (17,7 %), mais la prise en compte de la seule note globale au baccalauréat standardisée permet de faire mieux (25,3 %). Cela illustre aussi l'importance de la standardisation, puisque la note brute n'a un pouvoir explicatif que de 7,4 %. Or dans l'étude des parcours dans l'enseignement supérieur ou de l'insertion professionnelle, c'est

souvent la mention, découlant de la note brute, qui est utilisée pour affiner l'analyse. La standardisation ici proposée a donc un intérêt appréciable. Il est vrai que l'introduction de la série comme variable de contrôle a à peu près le même effet que la standardisation. Parmi les différents indicateurs de synthèse, les différences sont assez minces (on est proche de 25 % et du niveau avec le détail des évaluations), que l'on utilise une spécification linéaire ou quadratique.

TABLEAU 4 • Corrélations entre la qualification finale et les évaluations pour les bacheliers (R² ou pseudo-R²)

	Age de fin d'études	Age de fin d'études corrigé	Diplôme quantifié par âge théorique	Diplôme
	R ²		Pseudo-R ²	
Évaluations en 6 ^e	3,1	9,3	13,6	23,3
Évaluations en 6 ^e et 3 ^e	4,7	12,8	17,7	29,5
Évaluations en 6 ^e , 3 ^e et bac	12,6	20,7	26,5	43,2
Score global en 6 ^e	2,8	8,5	12,1	21,0
Score global en 6 ^e (specif. quad.)	3,4	9,2	12,6	21,0
Note globale standardisée au bac	9,6	19,6	25,3	38,6
Note glob stand. au bac (specif. quad.)	9,7	19,6	25,3	40,0
Note moyenne standardisée au bac	11,4	19,7	24,8	40,8
Note moy. stand. au bac (specif. quad.)	11,4	19,7	24,8	42,4
Moyenne de toutes les évaluations	7,2	16,3	21,9	36,8
Moy. toutes évaluations (specif. quad.)	8,5	18,2	23,3	37,0
Note glob. stand. au bac (1 ^{re} session)	8,1	20,3	25,6	37,7
Note glob. std. au bac (après rattrapage)	10,1	19,1	24,3	39,0
Note brute au bac	1,3	6,5	7,4	5,0
Note brute au bac (après rattrapage)	1,5	6,2	6,9	5,4

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

La synthèse par la moyenne de toutes les évaluations est la moins efficace (21,9 % en linéaire, 23,9 % en quadratique), sans doute parce qu'elle intègre des évaluations (en 6^e et en 3^e) trop éloignée de la scolarité dans l'enseignement supérieur.

Enfin, on notera que les modélisations de l'âge de fin d'études sont nettement moins bonnes (12,6 % avec l'ensemble des évaluations contre 26,5 %) suggérant que cette mesure simple donne une image assez bruitée de la valeur scolaire des jeunes.

➤ L'insertion professionnelle

Nous allons présenter ici quelques statistiques de cadrage sur la situation professionnelle des jeunes. Cela oblige à se restreindre aux seuls répondants à ces questions, ce qui représente 12 307 individus. Il paraît en particulier utile de rappeler que l'enquête EVA est aussi un panel et que la situation des jeunes est observée plusieurs fois, ce qui permettrait d'utiliser des techniques d'analyse longitudinale.

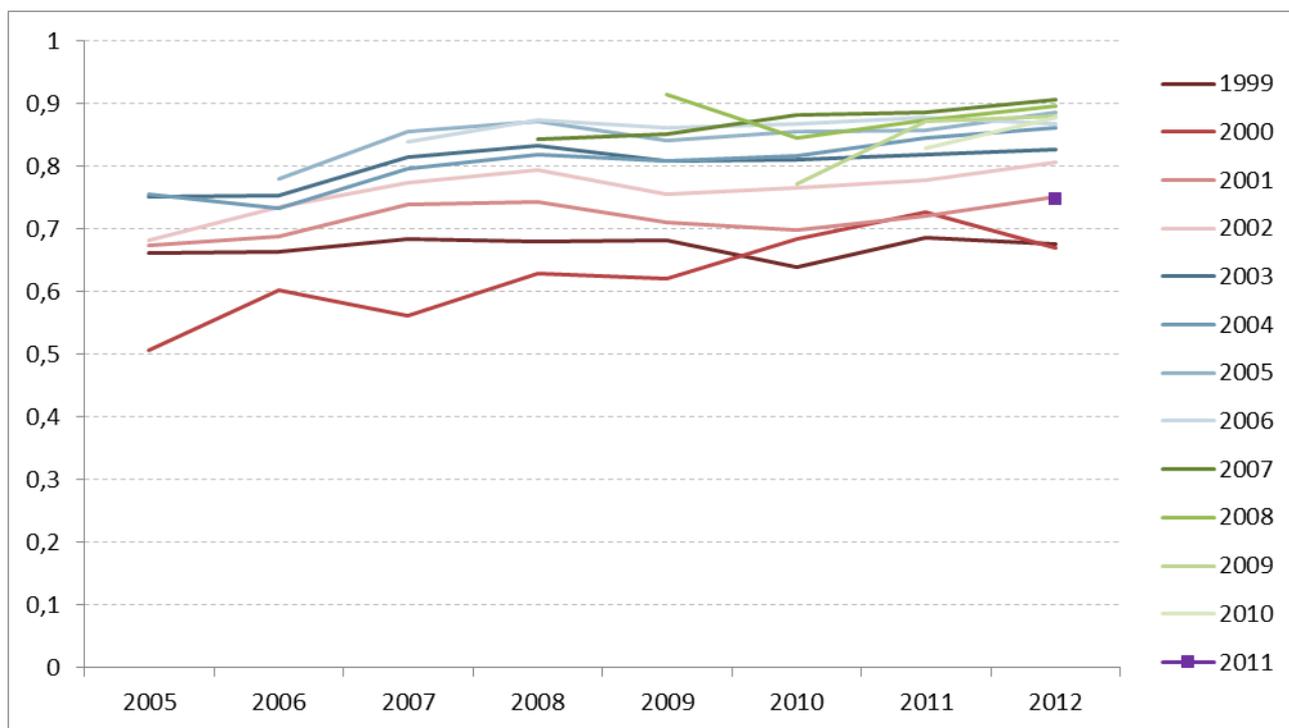
Les données d'EVA sur la situation professionnelle sont a priori délicates à utiliser, comparées à ce qui est fait sur les enquêtes Génération et le bilan Formation-Emploi. En effet, contrairement à ces deux dispositifs, qui ciblent une cohorte de jeunes « sortants », venant de finir leurs études, le panel EVA porte sur une cohorte d'entrants en 6^e à la même date et finissant leurs études à des dates très éloignées (plus de 10 ans entre le premier sorti et le dernier sorti, sachant que 1 % des jeunes étaient

même encore en études lors de la dernière interrogation d'EVA), dans des conjonctures économiques différentes. Une autre difficulté vient du fait que le suivi EVA n'a commencé qu'en mars 2005, alors que certains jeunes avaient fini leurs études et commencé à travailler depuis plusieurs années. Il n'est donc pas possible d'avoir véritablement pour eux leur première situation professionnelle. Enfin, la transmission des données entre la DEPP, le SIES et l'Insee privilégie le suivi scolaire : on veut être certain qu'un jeune a fini ses études avant de le faire interroger par EVA, si bien qu'après sa première déclaration de fin d'études, il a généralement été réinterrogé par le SIES l'année suivante pour valider cet arrêt, décalant d'autant le suivi de la situation professionnelle fine dans EVA (par « fine », nous entendons notamment l'information sur le salaire ou la profession ; le statut d'activité est connue dans l'enquête du SIES).

L'accès à l'emploi

Comme les sortants précoces sont observés à partir de 2005, à un moment plus avancé de leur carrière professionnelle, leur situation est sans doute plus favorable qu'à la fin de leurs études, moment où l'on observe les jeunes ayant fait des études longues. En termes de conjoncture, les sortants précoces sont aussi avantagés par le fait d'être sortis avant la crise économique de 2007, en comparaison des plus diplômés. Ceci dit, les analyses sur les enquêtes Génération et le bilan Formation Emploi ont montré que ces derniers avaient été peu touchés par la crise, au contraire des jeunes sortant assez jeunes du système éducatif. Cela signifie que le panel 1995 donne sans doute une bonne image des écarts d'insertion professionnelle selon le niveau d'études au début des années 2000, avec une légère sous-estimation due au fait que les sortants précoces sont observés plus tardivement dans leur carrière professionnelle, mais un aspect conjoncturel atténué par le fait que la situation des plus diplômés sortis autour de 2010 était assez proche de celle des plus diplômés sortis autour de 2000. Les écarts selon le niveau d'études ont sans doute été encore plus importants autour de 2010, car les moins diplômés sortant en 2010 ont connu une insertion professionnelle bien moins bonne qu'en 2000. Rappelons enfin que les jeunes sortis précocement ont moins souvent répondu à EVA : une pondération corrige ce problème, mais il est possible que les jeunes dans ce cas qui répondent aient un profil particulier, peut-être plus favorable que les autres, ce qui biaiserait encore l'observation en faveur de cette population. Il est donc possible que les écarts, déjà importants, que nous allons montrer en termes de situation professionnelle en fonction du parcours scolaire ou des compétences, soient en fait sous-estimés.

FIGURE 7 • Taux d'emploi selon la cohorte de sortie et l'année d'observation



Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Ceci dit, nous allons montrer que ces biais ne doivent pas être trop importants, car les résultats sont assez plausibles et, on le verra plus loin, cohérents avec ceux des enquêtes Génération et Emploi.

Ainsi, calculés sur les répondants aux différentes années d'observation, l'évolution du taux d'emploi (proportion de jeunes en emploi parmi l'ensemble des jeunes, incluant donc les chômeurs et les inactifs) paraît assez naturelle (FIGURE 7 •) : il tend à augmenter avec les années d'observation, ce qui est le signe d'une insertion professionnelle grandissante. Chaque cohorte définie par l'année de sortie démarre généralement à un niveau plus élevé que la cohorte précédente, ce qui peut s'expliquer par le fait qu'elle a un plus haut niveau de qualification, garantissant un meilleur accès à l'emploi. Cela n'est plus vrai à partir de la cohorte 2009, peut-être à cause de la crise économique dont les conséquences se font sentir, mais la taille des populations concernées peut aussi fragiliser les résultats (la cohorte 2011 observée seulement en mars 2012 ne comporte que 190 individus ; la faible taille de la cohorte 1999, avec 107 personnes au départ, amène aussi à considérer avec prudence les données la concernant).

Globalement, sur les 12 307 jeunes observés au moins une fois en termes de situation professionnelle, le taux d'emploi est de 75 % lors de la première observation (qui n'est pas forcément, rappelons-le, la première situation professionnelle après la fin des études) et de 81 % lors de la dernière observation (qui peut être la première si le jeune n'a répondu qu'une fois).

Le salaire

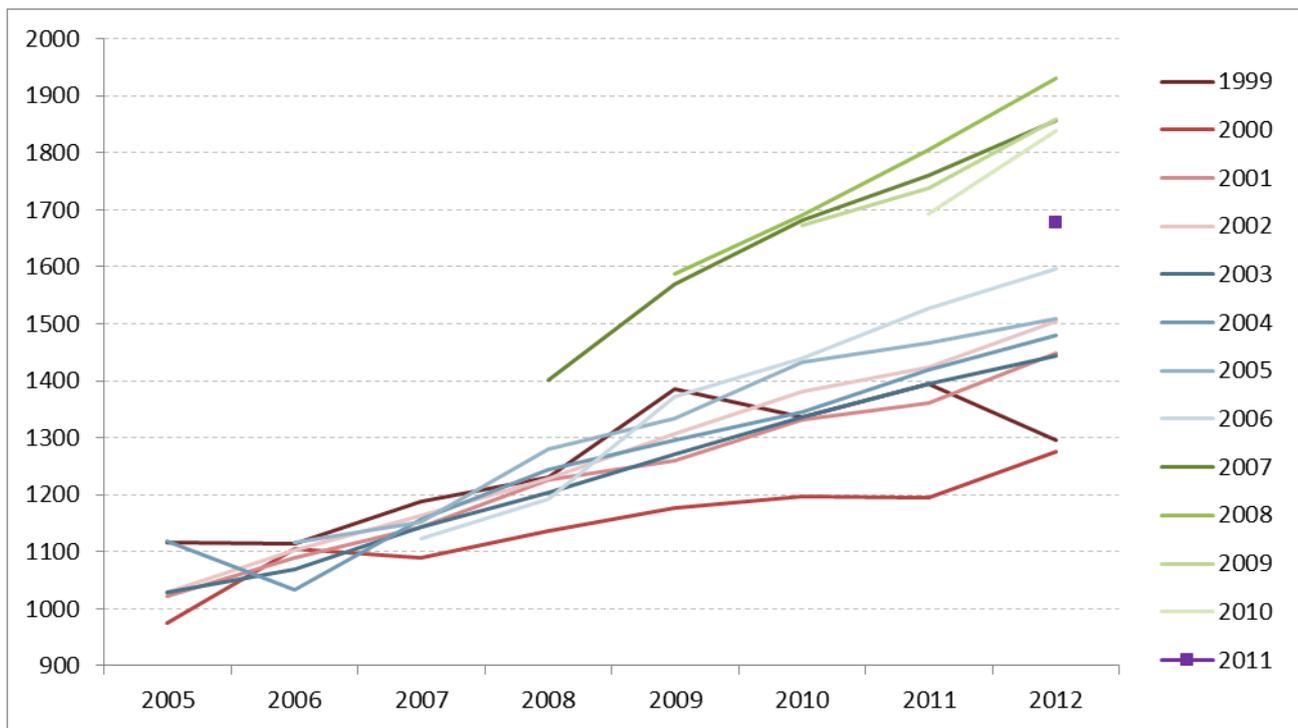
En termes de salaires (niveau de départ et évolution), deux groupes se détachent (FIGURE 8 •). Les cohortes antérieures à 2007 partent d'un niveau compris entre 1 000 euros⁵ (pour les cohortes 2001-2003) et 1 100 euros (pour les cohortes 2004-2006) pour atteindre entre 1 450 et 1 500 euros en mars 2012 (la cohorte 2006 atteint 1 600 euros). Les cohortes 1999 et 2000 ont un profil atypique (baisse brutale en mars 2012 pour la cohorte 1999 ; stagnation sur la période pour la cohorte 2000), mais elles sont représentées par peu d'individus (21 pour la cohorte 1999 en mars 2012 et moins de 100 pour la cohorte 2000 à partir de mars 2006). Le deuxième groupe est celui des cohortes à partir de 2007 : les premiers salaires observés sont nettement plus haut (1 400 euros pour la cohorte 2007 ; entre 1 600 et 1 700 euros pour les cohortes suivantes) et progressent fortement lors des années suivantes (en tout cas sur la courte période d'observation les concernant). Ceci tient à une évolution très nette du niveau d'études entre la cohorte 2006 et les suivantes : pour les sortants 2006, le niveau atteint est généralement au mieux la licence (37 %) ou un bac + 2 (43 %) et presque jamais un bac + 5 (6 %), la situation étant encore moins favorable pour les cohortes précédentes ; pour les sortants 2007, la proportion de sortants au niveau bac + 5 (en incluant les Grandes Ecoles) est de 48 % et elle augmente encore pour les cohortes suivantes. Connaissant le lien, que nous allons encore présenter plus loin, entre salaire et niveau d'études, il est normal de voir évoluer nettement le salaire entre ces cohortes.

Le premier salaire observé (qui n'est pas toujours le premier salaire perçu) est assez dispersé (FIGURE 9 •). Il a une valeur moyenne de 1350 euros (la médiane est à 1300 euros), mais 10 % des jeunes touchent moins de 750 euros, tandis que les 10 % les mieux payés dépassent 1950 euros. En mettant le seuil à 25 % de la population, l'écart passe à 1050 euros pour le premier quartile contre 1550 euros pour le troisième quartile. Le dernier salaire observé (qui est observé un nombre variable d'années après le premier : assez élevé pour les sortants avant 2004, très limité pour les jeunes sortis juste avant 2012) donne à peu près la même image en termes de dispersion, à un niveau un peu plus élevé, puisque la moyenne est alors de 1500 euros.

Pour caractériser la situation professionnelle du jeune, le salaire peut être complété par la catégorie sociale (pour information, le type de contrat, le type d'employeur et la durée de travail sont aussi disponibles). Au moment où le premier emploi est observé, un peu plus d'un tiers des jeunes sont employés, un quart, ouvriers, un autre quart exercent une profession intermédiaire, le statut de cadre supérieur étant attribué à un peu plus de 10 % de la population (TABLEAU 5 •). Entre cette première observation et la dernière, les jeunes grimpent un peu dans la hiérarchie sociale : la part des cadres et la part des professions intermédiaires augmentent chacune de deux points.

⁵ Les salaires de la FIGURE 7 • sont en euros courants. En revanche, le premier et le dernier salaires observés utilisés ensuite ont été ramenés en euros 2012 en utilisant l'indice des prix à la consommation pour les déflater. Dans la suite, on arrondira dans le texte les salaires à 50 euros près.

FIGURE 8 • Salaire moyen selon la cohorte de sortie et l'année d'observation

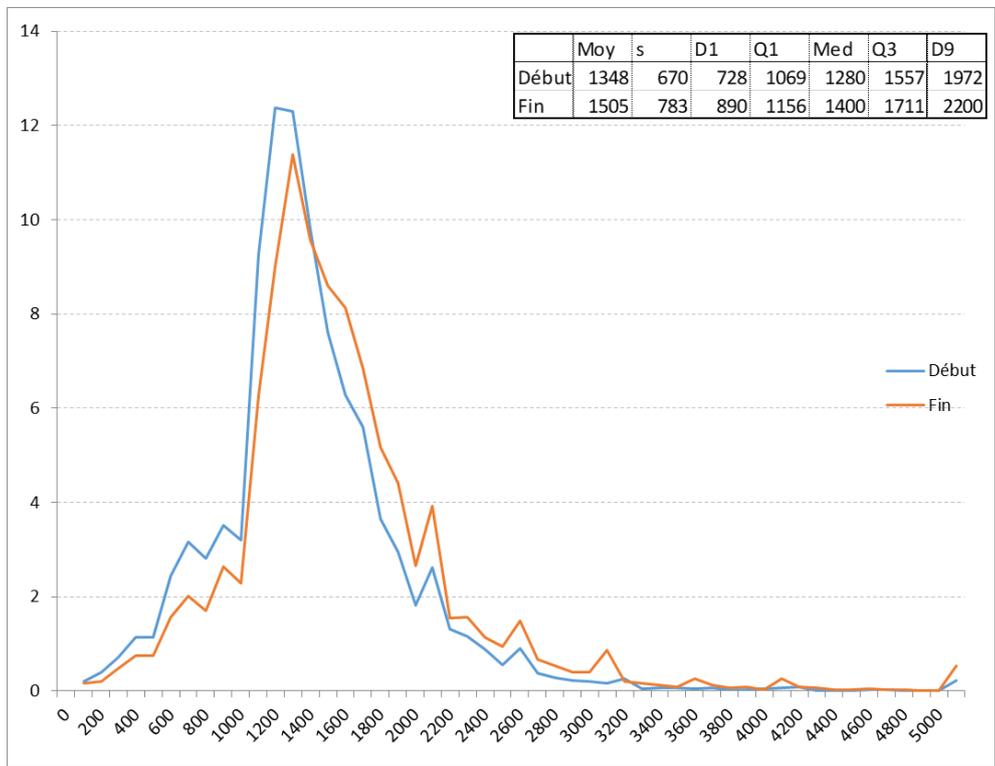


Lecture : les jeunes ayant fini leurs études en juin 2006 et occupant un emploi salarié, ont, en mars 2007, un salaire mensuel moyen de 1 200 euros.

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi percevant un salaire.

Source : DEPP-SIES-Insee, enquête panel 1995.

FIGURE 9 • Distribution du premier et du dernier salaire



Lecture : le graphique donne la répartition par tranches de 100 euros du premier et du dernier salaire observé.

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi percevant un salaire.

Source : DEPP-SIES-Insee, enquête panel 1995.

La profession

Les écarts de salaires selon la catégorie sociale sont importants : pour le premier salaire, les jeunes employés sont les moins bien payés avec un salaire de 1100 euros ; les ouvriers gagnent 1250 euros et les professions intermédiaires 1450 euros ; les cadres sont, comme attendu, les mieux payés avec 2000 euros. Les écarts sont du même ordre pour le dernier salaire, à un niveau plus élevé.

TABLEAU 5 • Professions exercée par le jeune

	Début		Fin	
	%	Salaire	%	Salaire
Agriculteur	0,3	1247	0,6	1233
Artisan-commerçant	1,6	1450	3,2	1476
Cadre supérieur	11,6	1994	13,7	2201
Profession intermédiaire	25,2	1439	27,0	1586
Employé	34,5	1119	32,3	1232
Ouvrier	24,3	1269	22,3	1390
Sans profession	2,4	1324	0,8	1274

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi (et percevant un salaire pour la colonne Salaire).

Source : DEPP-SIES-Insee, enquête panel 1995.

↳ Lien entre parcours scolaire et situation professionnelle

Nous allons présenter ici quelques résultats complémentaires sur le lien entre parcours scolaire et situation professionnelle, par rapport à l'article consacré au sujet.

Différentes façons de prendre en compte le parcours scolaire

Dans la première partie de ce document, nous avons montré que le panel permettait d'avoir une image très précise du parcours scolaire des jeunes dans les enseignements secondaire et supérieur, et de construire des indicateurs de synthèse variés sur le sujet (bien d'autres informations, comme le passage par le secteur privé ou l'éducation prioritaire, l'établissement ou la classe fréquentée, le fait d'avoir redoublé, pourraient aussi être utilisées). Nous allons ici relier la situation professionnelle à ces différents indicateurs.

TABLEAU 6 • Corrélation entre le parcours scolaire et la situation professionnelle (R² ou pseudo-R²)

	Accès à un emploi	Salaire	Log du salaire	Accès à un statut de cadre ou PI
Plus haut diplôme	8,1	28,8	20,3	41,7
Plus haut niveau d'études	4,6	26,1	17,4	39,6
Plus hauts diplôme et niveau d'études	9	29,5	21,0	43
Niveau d'études et indicatrice de réussite	7,5	28,1	19,5	41,7
Niveau d'études et indic. de réussite et spécialité	11,3	31,9	23,5	43,9
Âge de fin d'études	4,6	19,5	14,9	30,5
Âge de fin d'études corrigé	5,7	23,5	17,8	34,6
Diplôme quantifié en âge théorique	3,9	24,9	18,1	37,4
Diplôme quantifié en âge théorique (spécif. quad.)	5,3	25,3	18,2	37,6

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi (et percevant un salaire pour la colonne Salaire).

Source : DEPP-SIES-Insee, enquête panel 1995.

Ces indicateurs vont être confrontés à l'accès à l'emploi, le salaire (en brut et sous forme logarithmique) et l'accès à un statut de cadre supérieur ou une profession intermédiaire. Pour le salaire, c'est le R^2 de la régression qui est utilisé pour mesurer l'ampleur de la corrélation ; pour les variables d'accès, c'est le pseudo- R^2 de Cow & Snell.

Le plus haut diplôme dans une nomenclature en 13 groupes fournit une information décisive pour comprendre l'insertion professionnelle (TABLEAU 6 •) : la corrélation est modérée pour l'accès à l'emploi, avec un pseudo R^2 de 8,1 % (mais nous avons indiqué plus haut que le panel n'était pas une source parfaite sur cet aspect et qu'en particulier il devait donner une image un peu trop favorable des jeunes ayant fini tôt leurs études) ; elle est plus nette pour le salaire, avec un R^2 de 28,8 % (les modèles concernant le logarithme du salaire sont systématiquement moins bons : cela se tient sans doute au fait que l'hypothèse de distribution lognormale des salaires, justifiant l'usage du logarithme dans la modélisation, n'est, au vu de la FIGURE 9 •, pas vraiment vérifiée, une hypothèse de simple normale étant acceptable, sinon parfaitement adaptée) et pour l'accès à un statut de cadre, avec un pseudo- R^2 de 41,7 %.

Retenir le niveau d'études, sans prendre en compte sa validation par un diplôme, donne des corrélations un peu moins bonnes ($R^2=26,1\%$ pour le salaire) mais apporte tout de même une information supplémentaire puisqu'on atteint un $R^2=29,5\%$ en combinant plus haut diplôme et plus haut niveau d'études (cela rend compte du fait qu'un titulaire du baccalauréat gagne un peu plus s'il a entrepris des études supérieures, mêmes sans les valider). Nous avons privilégié une façon plus condensée de croiser les deux variables : le niveau d'études et une indicatrice de validation par un diplôme (voir en annexe, la construction de cette variable et globalement les liens complexes entre ces informations) : le R^2 est de 28,1 %, ce qui indique un gain de 2 points par rapport au modèle avec uniquement le niveau d'études, estimation de la valeur propre de la certification, comme ont pu le faire Le Rhun et Monso (2015) sur les BTS. Cette possibilité d'interprétation nous semble justifier le choix de cette spécification, même si elle est un peu moins « performante » que la simple prise en compte du diplôme (R^2 de 28,1 % contre 28,8 %).

La spécialité de la formation apporte une information très pertinente pour étudier l'insertion professionnelle, le R^2 pour le salaire gagnant presque 4 points ($R^2=31,9$). Nous détaillerons plus loin l'effet de cette variable

Enfin, nous avons testé différentes présentations quantitatives du parcours scolaire : la plus naturelle, l'âge de fin d'études donne des résultats corrects, mais sensiblement en-dessous des spécifications qualitatives : le R^2 vaut 19,5 % pour le salaire (contre 28,8 % avec le diplôme en 13 groupes). Retirer les années redoublées ou non validées par un diplôme améliore assez nettement la corrélation (le R^2 passe à 23,5 %) mais la quantification du diplôme par un âge théorique nécessaire pour l'atteindre (18 ans pour un bac, 20 ans pour un bac+2... cet indicateur peut différer du précédent, par exemple pour un jeune ayant validé deux BTS, mais aussi quand le parcours scolaire n'a pas été bien renseigné, compliquant le calcul de l'indicateur précédent) donne d'encore meilleurs résultats ($R^2=24,9\%$), en particulier en spécification quadratique ($R^2=25,3\%$ rendant compte du fait que les écarts de salaires bruts augmentent avec le nombre d'années d'études).

Résultats détaillés des modèles

L'un des objectifs de l'article est d'étudier le lien entre compétences et insertion professionnelle, en particulier en contrôlant le parcours scolaire. Nous avons donc retenu une spécification assez fine du parcours scolaire (compte tenu de la richesse du panel sur ce point, il aurait été possible d'aller encore plus loin), mais non destinée à décrire lisiblement le lien entre parcours scolaire et insertion professionnelle. Cette difficulté va surtout concerner la prise en compte de la spécialité (TABLEAU 7 •).

En effet, dans les modèles ci-dessous, confrontant le parcours scolaire (et le sexe) et les trois variables décrivant l'insertion professionnelle (accès à un emploi, salaire et accès à un poste qualifié), les coefficients associés au plus haut niveau d'études et à sa validation par un diplôme donnent des résultats assez classiques : la hiérarchie des niveaux de diplôme est globalement bien respectée pour le salaire et l'accès à un poste qualifié : les bac+2 sont à 130 euros au-dessus des sortants précoces, les bac+5 à 300 euros, les sortants d'une grande école à 650 euros (pour les salaires, on peut cependant noter le décrochage des jeunes s'arrêtant après un baccalauréat général, public assez particulier et des jeunes sortis après un BTS, qui ne se distinguent pas des sortants précoces). La

hiérarchie est moins nette pour l'accès à l'emploi, mais cette donnée est dans le panel 1995 assez bruitée, compte tenu du protocole de collecte assez complexe définissant le passage entre le SIES et l'Insee au moment de la fin des études. L'indicatrice repérant l'obtention du diplôme dans le niveau le plus haut suivi est dans les trois modèles affectée d'un coefficient positif (130 euros en plus pour le salaire par exemple). Ce résultat est cohérent avec l'étude de (Le Rhun et Monso, 2015).

La description de la spécialité est plus délicate : cette information est normalement disponible pour les élèves passés par l'enseignement professionnel ou par l'enseignement supérieur et nous considérons la série du baccalauréat comme une spécialité en terminale générale et technologique. Nous avons créé des indicatrices pour ces trois enseignements : pour un jeune ayant fait un BTS, après un baccalauréat technologique ayant suivi un CAP, il est donc possible d'avoir les valeurs « BTS Services ou DUT tertiaire », « Bac STT » et « Echanges et gestions ». Signalons de plus que la référence « Pas de spécialité » correspond aux cas où le jeune n'est pas concerné (il n'a pas fait d'études supérieures) et aux cas où l'information n'a pas pu être récupérée.

Ceci rend l'interprétation des coefficients un peu difficile : un sortant de grande école a normalement une spécialité dans l'enseignement supérieur, par exemple Sciences (avec un surcroît de salaire de 200 euros) et la différence avec les sortants précoces de 650 euros est donc un minorant de l'écart réel. Cependant, dans la partie de l'article concernée ayant pour objectif d'étudier le lien entre compétences et insertion professionnelle, sous contrôle du parcours scolaire, le fait que la spécification du parcours scolaire soit peu lisible ne nous semble pas poser problème.

Dans la dernière partie de l'article, où c'est l'objectif complémentaire, étudier le lien entre parcours scolaire et insertion professionnelle, en contrôlant les compétences, qui est traité, on a justement simplifié la spécification du parcours scolaire en le réduisant à l'âge de fin d'études (ce qui a l'intérêt supplémentaire de rapprocher l'analyse des modèles de Mincer).

TABLEAU 7 • Situation professionnelle selon le parcours scolaire

		Accès à l'emploi	Salaire	Accès à un statut de cadre
R ² ou Pseudo-R ²		11,5%	32,3%	44,1%
Constante		0,5***	1086***	-2,78***
Garçons		0,08	182***	0,13**
Niveau d'études (Ref=Classe de collège)	2nde GT	1,09**	-15	-9,98
	Année non terminale de pro	-0,33*	49	0,43
	Première G&T	0,92***	25	1,39**
	Terminale CAP-BEP	0,4***	0	0,11
	Terminale Bac pro	0,92***	81**	0,94**
	Terminale Bac tec	1,03***	33	1,25***
	Terminale Bac gen	0,66***	-15	1,54***
	Deug et autres bac + 2	0,66**	134**	2,57***
	DUT	0,47*	148**	2,3***
	BTS	0,66***	45	2***
	Bac + 3	0,54***	107**	2,72***
	Bac + 4	0,99***	317***	3,31***
	Bac + 5 et thèses	0,33*	304***	3,23***
Grandes écoles	0,73***	645***	4,18***	
Diplôme obtenu	Oui	0,65***	128***	0,69***
Spécialité du supérieur	Sciences	0,86***	203***	0,29**

(Ref=Pas de spécialité)	Économie-gestion	0,94***	259***	0,05
	Lettres-Arts	0,59***	-63*	0,14
	Paramédical	1,83***	373***	1,32***
	BTS Prod. ou DUT Secondaire	0,83***	210***	0,75***
	BTS Services ou DUT Tertiaire	0,97***	171***	0,15
	Droit	0,54***	142***	0,17
	Langues	0,74***	36	-0,32**
	Sciences humaines	0,77***	-36	0,03
	STAPS	1,72***	-43	-0,12
Spécialité du baccalauréat GT (Ref=Pas de spécialité)	Bac S	-0,22*	106***	0,39***
	Bac L	-0,37***	-41	0,18
	Bac ES	-0,16	13	0,13
	Bac STI	-0,06	3	0,12
	Bac STT	-0,51***	-27	-0,17
	Autre bac technologique	-0,07	17	0,24*
Spécialité du baccalauréat Professionnel (Ref=Pas de spécialité)	Spé. pluri-techno de la product.	0,55	73	0,23
	Agric., pêche, forêt, espaces verts	0,8***	4	-0,45**
	Transformations	0,4***	49*	-0,41*
	Échanges et gestion	-0,18*	-51**	-0,15
	Services aux personnes	-0,44***	-127***	-0,26
	Communication et information	0,53***	50**	-0,1
	Génie civil, construction, bois	-0,25	-136***	-1,06**
	Matériaux souples	32,2%	44**	-23,0%
	Méca., électricité, électronique	0,33***	-39*	-0,45***
Services à la collectivité	-0,74***	-75	0	

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi (et percevant un salaire pour la colonne Salaire).
Source : DEPP-SIES-Insee, enquête panel 1995.

Différentes synthèses des évaluations

Dans l'article, nous avons utilisé comme mesure synthétique des compétences, donnant une image à la fin des études secondaires, la moyenne des principales notes standardisées au baccalauréat. Rappelons que la standardisation (décrite en annexe) comportait une utilisation des évaluations en 6^e et des notes en 3^e à la fois pour imputer une valeur aux personnes ne passant pas le baccalauréat et un calage entre les séries pour tenir compte de leur sélectivité différente.

Nous allons confronter ici l'insertion professionnelle avec d'autres indicateurs synthétiques envisageables (TABLEAU 8 •). Une première possibilité est de prendre la note « officielle » (avant rattrapage et dernière session passée pour les redoublants) avec la standardisation évoquée ci-dessus. Cette note a un lien moins bon avec la situation professionnelle ($R^2=17,3\%$ par exemple pour l'équation de salaire sans contrôle du parcours scolaire, contre $21,4\%$ pour la moyenne des notes). Cela tient soit à la prise en compte de notes dans les « petites » matières, qui dégrade la mesure ou à l'utilisation de pondérations non optimales.

Nous avons aussi examiné l'impact de nos choix dans la construction des notes, en prenant la note « officielle » après rattrapage ou celle de la première session passée pour les redoublants (toujours en utilisant la standardisation). Les différences sont minimales : les modèles sont très légèrement moins

bons avec la note après oral de rattrapage ($R^2=17,1\%$ pour le salaire contre $17,3\%$ avant) et très légèrement meilleurs en prenant la première session ($R^2=17,4\%$).

TABLEAU 8 • Corrélation entre le parcours scolaire et la situation professionnelle (R^2 ou pseudo- R^2)

	Accès à un emploi		Salaire		Accès à un statut de cadre ou PI	
	Sans contrôle	Avec contrôle	Sans contrôle	Avec contrôle	Sans contrôle	Avec contrôle
Note moyenne au bac	2,3	11,5	21,4	32,7	31,4	44,3
Note globale standardisée au bac	3,5	11,8	17,3	32,3	28,1	44,2
Moyenne de toutes les évaluations	2,3	11,5	17,7	32,4	29,7	44,3
Note glob. stand. au bac après rattrap.	3,5	11,8	17,1	32,3	28,1	44,2
Note glob. stand. au bac (1 ^{re} session)	3,5	11,9	17,4	32,4	27,8	44,2

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi (et percevant un salaire pour la colonne Salaire).
Source : DEPP-SIES-Insee, enquête panel 1995.

Variantes de modélisation du lien entre compétences et insertion professionnelle

L'article met en évidence une nette hétérogénéité du lien entre compétences et insertion professionnelle selon le sexe et le niveau scolaire. Pour le sexe, ce sont les compétences en français qui sont l'objet principal de l'hétérogénéité. La robustesse de ce résultat, établi avec les évaluations à l'entrée en 6^e, peut être assurée en utilisant les autres évaluations, en 3^e et au baccalauréat. Dans les deux cas, on modifie non seulement les mesures des compétences, mais aussi le champ d'analyse, en ne retenant que les candidats au brevet pour les notes en 3^e et les candidats au baccalauréat pour les notes à cet examen (TABLEAU 9 •).

TABLEAU 9 • Situation professionnelle selon les compétences (variantes de l'article sur les effets hétérogènes)

Modèle	Accès à l'emploi	Salaire	Accès à un statut de cadre
Sexe+Éducation+Note 3 ^e +effets hétérogènes (R^2)	11,4%	33,0%	42,0%
Note 3 ^e (Filles sans études supérieures)	0,08	22*	0,07
Note 3 ^e (Δ Garçons)	-0,19***	-26**	-0,3***
Note 3 ^e (Δ Etudes supérieures)	0,3***	36**	0,24***
Sexe+Education+Note Bac+effets hétérogènes (R^2)	10,8%	32,2%	34,4%
Note Bac (Filles sans études supérieures)	-0,31***	21	0,08
Note Bac (Δ Garçons)	-0,23***	-8	-0,3***
Note Bac (Δ Etudes supérieures)	0,62***	95***	0,34***

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi (et percevant un salaire pour la colonne Salaire).
Source : DEPP-SIES-Insee, enquête panel 1995.

Avec les notes en 3^e, le lien entre compétences et insertion sociale pour les filles sans études supérieures est moins net qu'avec le score à l'entrée en 6^e : seul le coefficient dans la modélisation du salaire est significativement positif au seuil de 10 % (les deux autres sont cependant du bon signe). En revanche, l'hétérogénéité selon le sexe et le niveau d'études est bien confirmée : pour les jeunes ayant fait des études supérieures, le lien est nettement plus fort que pour ceux qui n'ont pas dépassé la terminale ; à l'inverse, ce lien paraît moins fort, voire négatif pour les garçons, surtout s'ils n'ont pas fait d'études supérieures.

Pour les candidats au baccalauréat, les résultats sont légèrement différents : pour les filles n'ayant pas fait d'études supérieures (il est vrai peu nombreuses parmi les candidats au baccalauréat et au

profil particulier), la note à l'examen a un lien négatif significatif avec l'accès à l'emploi et non significatif pour le salaire et l'accès à un poste qualifié (mais au moins de signe positif). En revanche, l'hétérogénéité des coefficients selon le sexe et le niveau scolaire est confirmée (seul le coefficient pour l'effet croisé selon le sexe dans la modélisation du salaire n'est pas significatif).

Dans l'article, les analyses par discipline ont utilisé les résultats à l'évaluation nationale à l'entrée en 6^e pour les mathématiques et le français (TABLEAU 10 •). D'autres mesures sont disponibles et des variantes de modélisation ont été effectuées en les substituant à ces évaluations de référence (faire une modélisation globale avec toutes les évaluations en même temps entraînait de trop fortes colinéarités). L'utilisation de l'appréciation des principaux sur le niveau en mathématiques à la place des évaluations en 6^e dégrade la qualité de la modélisation pour l'accès à l'emploi et le salaire (le coefficient n'est même plus significatif pour le salaire). Les appréciations en français, que ce soit la maîtrise de lecture, du français écrit ou du français oral donnent des résultats similaires à l'évaluation nationale.

TABLEAU 10 • Situation professionnelle selon les compétences (variantes de l'article sur le détail par compétences)

	NAT F NAT M	NAT F A M	A Ecrit NAT M	A Lecture NAT M	A Oral NAT M	3 ^e F 3 ^e M	Bac F Bac M
Accès à un emploi	12,1%	11,9%	12,1%	11,9%	11,6%	10,7%	10,7%
Français (Fille)	0,08*	0,16***	0,04	0,07*	0,03	0	0,03
Français (Δ Garçons)	-0,18***	-0,18***	-0,15***	-0,12**	-0,19***	-0,11	-0,1
Mathématiques	0,18***	0,07**	0,19***	0,17***	0,23***	0,11**	0,1**
Connaiss. civiques	-0,02	-0,02	-0,02	-0,02	-0,01	0	0
Vie scolaire	0,03	0,03	0,03	0,03	0,02	0,05	0,05
Opinions	-0,07***	-0,06**	-0,07***	-0,07***	-0,05*	-0,07*	-0,07*
Méthodes de travail	0,01	0,01	0,01	0,01	0,01	0	0
Salaires	32,8%	32,8%	32,7%	32,4%	33,3%	31,6%	31,7%
Français (Fille)	15*	23***	19**	26***	6	37***	23**
Français (Δ Garçons)	-42***	-43***	-26***	-24***	-26**	-22	9
Mathématiques	20***	9	13**	17***	31***	38***	39***
Connaiss. civiques	6	6	5	6	11**	14**	14**
Vie scolaire	-11**	-11**	-12**	-12***	-17***	-16**	-16**
Opinions	8*	9*	8*	8*	12**	9	8
Méthodes de travail	25***	25***	24***	24***	27***	34***	34***
Accès poste qualif.	44,3%	44,3%	44,3%	44,3%	42,0%	30,4%	30,3%
Français (Fille)	0,1*	0,1**	0,14***	0,19***	0,15**	0,16***	0,12***
Français (Δ Garçons)	-0,19***	-0,19***	-0,18***	-0,2***	-0,21***	-0,29***	-0,13**
Mathématiques	0,05	0,06*	0,03	0,06*	0,05	0,15***	0,15***
Connaiss. civiques	0,09***	0,09***	0,09***	0,09***	0,1***	0,07**	0,07**
Vie scolaire	0,02	0,02	0,02	0,02	0,02	0,02	0,02
Opinions	0,02	0,02	0,02	0,02	0,03	0,04	0,04
Méthodes de travail	0,04	0,04	0,03	0,04	0,03	0,04	0,04

Champ : France hors DROM, jeunes entrés en sixième en 1995 en emploi (et percevant un salaire pour la colonne Salaires).

Source : DEPP-SIES-Insee, enquête panel 1995.

Utiliser les notes de 3^e en mathématiques et en français, uniquement sur les candidats au brevet, diminue un peu l'ampleur des coefficients pour l'accès à l'emploi, mais les augmentent pour le salaire

et l'accès à un poste qualifié. C'est aussi le cas pour les notes au baccalauréat pour les candidats à cet examen. L'hétérogénéité du coefficient en français selon le sexe est toutefois moins nette et non significative pour l'accès à l'emploi et le salaire.

Des analyses plus précises avec l'ensemble des notes, en langue vivante en 3^e ou dans les autres disciplines du baccalauréat peuvent être envisagées, mais il faudrait tenir compte d'une non-réponse assez importante (peu de jeunes ont une note en physique) et de la colinéarité entre ces variables.

➤ Comparaison avec les enquêtes Génération et Emploi

Nous avons évoqué plus haut les difficultés que posent les panels DEPP pour construire des indicateurs de synthèse sur toute la trajectoire des jeunes, en particulier en termes d'insertion professionnelle. C'est pourquoi il paraît utile de confronter quelques résultats globaux avec des données équivalentes issues des enquêtes faisant référence sur l'insertion des jeunes, les enquêtes Génération et Emploi.

Le plus haut diplôme

TABLEAU 11 • Année et âge de fin d'études selon le diplôme des parents

	Génération		Enquête Emploi		Panel 1995
	2004	2010	2002-2004	2008-2010	
Aucun diplôme			11	9	9
DNB			6	8	6
Non-diplômés	18	16	17	17	15
CAP, BEP, mention complémentaire	17	14	17	16	18
Bac professionnel ou technologique	15	19		17	19
Bac général	8	9		8	8
Total bac	23	28	24	25	28
BTS, DUT et autres bac + 2	14	11	15	12	11
DEUG			1		2
BTS			12		9
DUT			2		1
Bac + 2/3 Santé social	4	4	3	3	
Total bac + 2	18	15	18	15	11
Licence générale (L3) et autres bac + 3	5	5			
Licence professionnelle	1	4			
Total bac + 3	6	9	6	11	7
M1 et autres bac + 4	4	1	5		4
M2 et autres bac + 5	7	10	5		
Doctorat	2	2	2		
Total bac + 5 et +	9	12	7	11	13
Écoles de commerce	2	2	3		
Écoles d'ingénieurs	3	3	3		
Total Grandes Écoles	5	5	6	5	4

Champ : France hors DROM - primo-sortants du système éducatif en 2004 et 2010 - sortis depuis 1 à 4 ans de formation initiale - jeunes entrés en sixième en 1995.
Source : Cereq, enquêtes Génération 2004 et Génération 2010 - Insee, enquête Emploi 2002-2004 et 2008-2010 - DEPP-SIES-Insee, enquête panel 1995.

Dans un premier temps, nous avons confronté ces sources sur le plus haut niveau de diplôme. Le panel fournit a priori des données de qualité sur le sujet, mais il existe une attrition non négligeable, qui peut la dégrader et la synthèse entre les informations de la DEPP, du SIES et de l’Insee (comme décrite en annexe) est loin d’être simple. La confrontation avec d’autres sources est donc utile.

Parmi les élèves entrés en 6^e en 1995, 9 % n’ont obtenu aucun diplôme et 6 % n’ont eu que le DNB (TABLEAU 11 •). Un peu moins de deux jeunes sur quatre ont pour plus haut diplôme un examen de l’enseignement secondaire : 18 % un CAP-BEP, 19 % un baccalauréat technologique ou professionnel et 8 % pour le baccalauréat général. Environ un dixième des élèves finissent leurs études avec un diplôme de niveau bac+2 : 1 % un DEUG, 2 % un DUT et 9 % un BTS. Les diplômes de niveau supérieur regroupent 7 % des jeunes pour les Bac+3, 4 % pour les Bac+4, 13 % pour les Bac+5 et 4 % de jeunes passés par une Grande École.

Ces données ont été confrontées avec les résultats des enquêtes Génération 2004 et 2010 (Barret et al., 2014) et avec ceux tirés des enquêtes Emploi (*État de l’École* 2006 et 2012). La cohérence est assez bonne, avec quelques différences, qui peuvent s’expliquer soit par un décalage de générations (en particulier avec l’impact de la réforme LMD) soit par le mode de collecte. La part de jeunes ayant au mieux le DNB est un peu moins élevée dans le panel 1995 (15 % contre 16 à 18 % dans les enquêtes Génération et Emploi). La proportion de jeunes au niveau Bac + 2 est aussi plus faible (11 % dans le panel contre 15 % dans Génération 2010 et les enquêtes Emploi de la même époque). La comparaison pour les bac + 3 et bac + 4 est rendue aussi compliquée par la mise en place de la réforme LMD : les proportions respectives de 7 % et 4 % à ces niveaux sont assez proches des valeurs de Génération 2004 (6 % et 4 %), mais moins cohérentes avec Génération 2010 (9 % et 1 %). Les divergences sont beaucoup moins fortes pour les CAP-BEP et les baccalauréats, ainsi que pour les bac + 5 et les Grandes Écoles.

La profession

Au moment de la première situation professionnelle observée de façon détaillée, un peu plus d’un tiers des jeunes occupent un poste d’employé, un quart un emploi d’ouvrier, un autre quart une profession intermédiaire, le statut de cadre supérieur étant attribué à un peu plus de 10 % de la population (TABLEAU 12 •). Entre cette première observation et la dernière, les jeunes grimpent un peu dans la hiérarchie sociale : la part des cadres et la part des professions intermédiaires augmentent chacune de deux points.

Par rapport aux résultats obtenus à partir des enquêtes Génération et Emploi (Cereq, 2008 ; Cereq, 2014, Insee, 2009), on observe un léger déficit de cadres et de professions intermédiaires : en effet, selon les enquêtes et les points d’observation (situation initiale ou situation au bout de 3 ans de vie professionnelle), la part des cadres varie de 14 % à 20 % et celle des professions intermédiaires de 28 % à 36 %. Cela peut tenir à l’extension de l’accès au statut de cadre (entre Génération 2004 et Génération 2010, la part de cadres 3 ans après la fin des études passe de 14 % à 18 %). Il se peut aussi que l’observation des derniers sortants du panel, très diplômés, ne soit pas assez profonde pour repérer leur accès au statut de cadre.

TABLEAU 12 • Catégorie sociale des jeunes dans les enquêtes Génération et Emploi

	Génération 2004	Génération 2010		Enquête Emploi 2007	Panel 1995	
	À 3 ans	Initiale	À 3 ans	Sortants 1-5 ans	Début	Fin
Agriculteur-Artisan	2	2	2	2	2	4
Cadre supérieur	14	18	20	16	11	14
Prof. intermédiaire	36	28	31	29	25	27
Employé	24	26	28	32	36	34
Ouvrier	24	26	19	23	26	24

Champ : France hors DROM - primo-sortants du système éducatif en 2004 et 2010 - personnes sorties depuis 1 à 4 ans de formation initiale - jeunes entrés en sixième en 1995.

Source : Cereq, enquêtes Génération 2004 et Génération 2010 - Insee, enquête Emploi 2007 - DEPP-SIES-Insee, enquête panel 1995.

L'accès à l'emploi

Globalement, sur les 12 307 jeunes observés au moins une fois en termes de situation professionnelle, le taux d'emploi est de 75 % lors de la première observation (qui n'est pas forcément, rappelons-le, la première situation professionnelle après la fin des études) et de 81 % lors de la dernière observation (qui peut être la première si le jeune n'a répondu qu'une fois) (TABLEAU 13 •). À titre de comparaison, le taux d'emploi à 3 mois est de 65 % dans Génération 2004 et 62 % dans Génération 2010 et le taux d'emploi à 3 ans est de 77 % dans Génération 2004 et 69 % dans Génération 2010 (Barret *et al.*, 2014). À partir de l'enquête Emploi, pour les sortants depuis 1 à 4 ans, le taux d'emploi en 2007 était de 75 % (Insee, 2009). Les taux d'emploi dans le panel sont assez cohérents avec ces valeurs, en tenant compte du fait qu'il s'agit souvent d'une insertion quelques temps après la fin des études et non immédiate. De plus, la plupart des jeunes du panel ont arrêté leurs études avant 2010, sans être trop touchés par la crise économique de la fin des années 2000.

L'accès à l'emploi dépend fortement du niveau d'études atteint. Quand les jeunes sont sortis sans diplôme ou avec uniquement le DNB, le taux d'emploi est respectivement de 51 % et 58 % ; à l'opposé pour les sortants d'une grande école, ce taux approche de 90 %. Les autres diplômes se classent de façon intermédiaires : autour de 70 % pour les jeunes ayant fini un cursus de second cycle dans l'enseignement secondaire, en le validant ou non (on dépasse 80 % s'il s'agit d'un bac professionnel) ; autour de 80 % pour les jeunes ayant fait des études supérieures.

TABLEAU 13 • Situation professionnelle en fonction du diplôme (panel 1995)

	Taux d'emploi	Premier salaire
Aucun diplôme connu	50,7	1027
DNB	58,2	984
CAP-BEP	74,5	1046
Bac pro	81,6	1155
Bac tec	74,9	1094
Bac gen	67,5	1156
DEUG et autres bac +2	76,7	1155
DUT	83,7	1458
BTS	83,6	1295
Bac +3	87,3	1472
Bac +4	67,4	1515
Bac +5 et thèses	80,6	1594
Grandes écoles	89,9	2142

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Dans Génération 2004, le taux d'emploi à 3 mois pour les sortants ayant au mieux le DNB est de 45 % et monte à 57 % à 3 ans, contre 50 % pour les sans diplôme et 58 % pour les titulaires d'un DNB, dans le panel (TABLEAU 14 •). Pour les autres diplômes, la concordance est toujours meilleure pour le taux à 3 ans : 76 % d'emploi pour les CAP-BEP dans Génération 2004 contre 75 % dans le panel ; 78 % pour les baccalauréats professionnels et technologique contre 81 % pour les baccalauréats professionnels et 75 % pour les baccalauréats technologiques dans le panel ; 62 % pour les baccalauréats généraux contre 67 % ; environ 85 % pour les BTS-DUT dans les deux sources... La concordance est moins bonne pour les bac + 5 : 90 % dans Génération 2004 contre 81 % dans le panel, sans doute parce que la durée d'observation après la fin des études est pour eux trop courte dans le panel (le taux dans le panel se trouve à mi-chemin entre ce taux de 81 % à 3 ans et le taux de 68 % à trois mois). Notons qu'en étant cohérents avec les taux de l'enquête Génération 2004, les taux d'emploi divergent logiquement de

ceux de Génération 2010, qui enregistrent une baisse marquée (15 points) des taux pour les moins diplômés, tandis que les taux évoluent moins pour les autres niveaux de diplômes.

TABLEAU 14 • Situation professionnelle en fonction du diplôme (enquêtes Génération)

	2004		2010	
	À 3 mois	À 3 ans	À 3 mois	À 3 ans
Aucun diplôme ou DNB	45	57	32	41
CAP, BEP, mention complémentaire	64	76	54	61
Bac professionnel ou technologique	71	78	67	70
Bac général	63	62	59	55
BTS, DUT et autres bac+2	70	85	70	79
Bac +2/3 Santé social	93	97	95	96
Licence générale (L3) et autres bac +3	75	81	70	70
Licence professionnelle	78	91	75	85
M1 et autres bac +4	68	83	70	79
M2 et autres bac +5	68	90	74	84
Doctorat	80	91	85	92
Écoles de commerce	71	94	71	90
Écoles d'ingénieurs	64	93	81	94

Champ : France hors DROM - primo-sortants du système éducatif en 2004 et 2010.

Source : Cereq, enquêtes Génération 2004 et Génération 2010.

Le salaire

Le premier salaire observé dans le panel a une valeur moyenne de 1350 euros (la médiane est à 1300 euros) Le dernier salaire observé (qui est recueilli un nombre variable d'années après le premier : assez élevé pour les sortants avant 2004, très limité pour les jeunes sortis juste avant 2012) est de 1500 euros.

A titre de comparaison, les enquêtes du Cereq donnent un premier salaire de 1270 euros et un salaire à trois ans de 1430 euros pour les sortants de 2004 et respectivement de 1340 euros et 1450 euros pour les sortants de 2010 (Barret *et al.*, 2014). Dans l'enquête Emploi, on observe un salaire de 1270 pour les personnes sorties depuis moins de 5 ans du système éducatif en 2007 et de 1410 euros pour celles sorties entre 5 et 10 ans (Insee, 2009). Le premier salaire observé paraît un peu plus élevé dans le panel que ces valeurs (rappelons qu'il est en euros 2012 et qu'il ne s'agit pas forcément du premier salaire réel), mais l'écart est assez limité.

Pour Génération 2004, les sortants d'une école d'ingénieur touchent un premier salaire de 2070 euros (TABLEAU 15 •), qui augmente de plus de 200 euros en 3 ans ; les bac +5 sont à 1610 euros en premier salaire, qui augmente de plus de 300 euros en 3 ans (dans le panel, le salaire est de 1620 euros) ; les titulaires d'un baccalauréat général sans diplôme supérieur touchent un premier salaire de 1140 euros et les bacheliers technologiques ou professionnels 1180 (à comparer à 1210 euros dans le panel pour les bacheliers généraux et 1180/1240 pour les bacheliers technologiques/professionnels) ; enfin, les jeunes sans diplômes perçoivent 1150 euros en premier salaire dans Génération 2004 (contre 1120 euros dans le panel). Les écarts entre diplômes sont très proches dans le panel 1995 de ceux que l'on observe dans Génération 2004 (et le constat est le même avec Génération 2010).

TABLEAU 15 • Salaire en euros 2010 en fonction du diplôme (enquêtes Génération)

	2004		2010	
	À 3 mois	À 3 ans	À 3 mois	À A 3 ans
Aucun diplôme ou DNB	1150	1220	1130	1160
CAP, BEP, mention complémentaire	1220	1310	1200	1260
Bac professionnel ou technologique	1180	1310	1170	1300
Bac général	1140	1320	1110	1260
BTS, DUT et autres bac +2	1270	1460	1310	1460
Bac +2/3 Santé social	1610	1720	1630	1700
Licence générale (L3) et autres bac +3	1330	1540	1300	1450
Licence professionnelle	1430	1660	1460	1600
M1 et autres bac +4	1380	1650	1470	1620
M2 et autres bac +5	1610	1970	1680	2020
Doctorat	1980	2340	2100	2350
Écoles de commerce	NS	NS	1990	2290
Écoles d'ingénieurs	2070	2340	2120	2350

Champ : France hors DROM - primo-sortants du système éducatif en 2004 et 2010.
Source : Cereq, enquêtes Génération 2004 et Génération 2010.

➤ Annexes – le traitement des données

Cette annexe présente les traitements, souvent complexes, qui ont été appliqués aux données afin de faire la synthèse des nombreuses informations disponibles, sur le parcours scolaire et professionnel des jeunes du panel 1995.

Les différentes sources

Pour étudier le parcours des élèves entrés en 6^e en 1995, trois sources de données sont à mobiliser :

- Le suivi de la DEPP, qui comporte les données de l'enquête de recrutement, les résultats aux évaluations 6^e, les résultats aux évaluations spécifiques en civisme et méthodes de travail, le suivi dans le second degré (incluant les résultats aux examens), l'enquête d'orientation en fin de 3^e, l'enquête Famille de 1998, les réponses à l'enquête Jeunes de 2002
- Le suivi du SIES (à l'époque le bureau en charges des statistiques sur l'enseignement supérieur faisait encore partie de la DEPP, mais par commodité, nous le désignerons par la structure qu'il a intégrée par la suite), qui décrit le parcours dans l'enseignement supérieur des bacheliers (à partir de la rentrée 2002). Le fichier est structuré par année après le bac. Pour travailler en millésime, il faut tenir compte de l'année d'obtention du baccalauréat. Pour chaque année, le jeune décrit la formation suivie et les diplômes éventuellement obtenus. Les réorientations sont aussi abordées. Le suivi est normalement arrêté après deux années où le jeune n'étudie plus : le jeune passe alors dans le dispositif EVA de l'Insee. Il y a aussi quelques questions sur la situation professionnelle et les conditions de vie (situation familiale, revenus, opinion sur l'avenir...).
- Le suivi de l'Insee, qui pose des questions sur la scolarité actuelle (quand le jeune perdu dans les dispositifs précédents est en fait encore en études initiales ou quand il les a reprises) et passée, sur l'insertion professionnelle et sur les conditions de vie. Il y a eu des interrogations spécifiques, notamment sur la santé par la DREES (Direction de la recherche, des études, de l'évaluation et des statistiques).

Cette collecte par trois voies est assez complexe et lors de la mise en place de l'enquête EVA par l'Insee, il y a eu quelques problèmes de transitions, des jeunes interrogés par plusieurs voies, d'autres par aucune. De plus, la première année, certains jeunes censés avoir arrêté leurs études sont apparus encore scolarisés et rendus à la DEPP ou au SIES. Ce retour, entraînant de nouvelles complications, n'a pas été repris les années suivantes.

Suivi par la DEPP

La DEPP a perdu la trace de 2 024 élèves au cours de son suivi (TABLEAU 16 •). Ces élèves ont été transférés à partir de 2004 à l'Insee pour qu'il les retrouve et les interroge dans le cadre de l'enquête EVA. Pour les 15 826 autres élèves du panel, la DEPP a repéré leur sortie, généralement entre 2001 et 2006. Dans certains cas, la sortie peut être considérée comme définitive : décès, problèmes de maladie grave ou départ vers l'étranger. Dans les autres cas, ces élèves ont été transmis au SIES (s'ils étaient bacheliers) ou à l'Insee. Ainsi, 46 % de l'ensemble des élèves du panel ont été suivis par le SIES et l'Insee (il s'agit forcément de bacheliers), 6 % n'ont répondu qu'au SIES, 24 % n'ont répondu qu'à l'Insee et autant n'ont pu être suivis après leurs études secondaires.

Suivi par le SIES

Normalement, le suivi par le SIES commençait à la rentrée suivant l'obtention du baccalauréat et se faisait jusqu'à ce que le jeune déclare 2 fois de suite ne plus être en études et son cas était alors transmis à l'Insee, mais les transmissions ont pu aussi se faire après des non-réponses. De plus, dans quelques cas, des jeunes perdus par la DEPP ont été repérés en études supérieures par l'Insee et ont alors été retransmis au SIES, qui a suivi leur parcours. Comme le SIES a réintégré les données collectées par l'Insee dans son suivi, nous les considérerons comme suivis par le SIES. Pour les 9 197 jeunes suivis par le SIES, l'année jusqu'à laquelle s'est fait le suivi, est présentée en fonction de l'année d'obtention du baccalauréat (TABLEAU 17 •). Ici, comme lorsque nous travaillerons sur EVA, on entend par « suivi » le fait d'avoir répondu à l'enquête. Un élève interrogé par le SIES, mais non répondant n'est pas considéré comme suivi, en tout cas l'année considérée (mais il peut être considéré comme suivi l'année suivante, s'il est réinterrogé et répondant). Ainsi, parmi les 4 601 bacheliers 2002, 327 ont été suivis uniquement en 2002-2003, tandis que 90 l'ont été jusqu'en 2010-2011 (certains parmi ces derniers ont pu ne pas répondre une année donnée).

La détermination de l'année de sortie va dépendre de la réponse aux questions sur les études en cours, qui sont abordées plus loin.

TABLEAU 16 • Fin de suivi par la DEPP

Sorties DEPP (rentrée)	Perdus	Fin d'études secondaires	Total	Suivis SIES et EVA	Suivis SIES uniquement	Suivis EVA uniquement	Non suivis
				%	%	%	%
1996	31	51	82	0	0	5	95
1997	176	35	211	0	0	24	76
1998	123	78	201	0	0	23	77
1999	241	297	538	0	0	29	71
2000	193	554	747	0	0	36	64
2001	258	1140	1398	0	0	48	52
2002	262	6614	6876	60	7	17	16
2003	283	4251	4534	56	7	21	16
2004	350	2104	2454	49	7	29	16
2005	87	567	654	43	7	36	14
2006	0	135	135	35	5	41	19
Total	2004	15826	17830	46	6	24	24

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

TABLEAU 17 • Première et dernière années de suivi par le SIES

Année du bac	Dernière année de suivi SIES (rentrée)									
	2002	2003	2004	2005	2006	2007	2008	2009	2010	Total
2002	327	267	365	431	452	1557	884	228	90	4601
2003	0	296	489	272	507	668	341	217	88	2878
2004	0	0	124	357	145	503	126	70	21	1346
2005	0	0	0	28	133	96	45	16	10	328
2006	0	0	0	0	5	23	11	4	1	44
Total	327	563	978	1088	1242	2847	1407	535	210	9197

Champ : France hors DROM, jeunes entrés en sixième en 1995 ayant eu le baccalauréat.
Source : DEPP-SIES-Insee, enquête panel 1995.

Suivi par l'Insee

L'Insee a pu suivre le parcours de 12 494 jeunes (les autres n'ont pu être retrouvés ou sont décédés, partis à l'étranger ou n'ont jamais accepté de répondre à l'enquête). Parmi les 3 253 jeunes que l'Insee a commencé à suivre à partir de 2004-2005, plus précisément à partir de mars 2005, 447 n'ont répondu qu'à cette première interrogation, tandis que 1 555 ont répondu jusqu'en mars 2012 (TABLEAU 18 •). Là encore, parmi ces jeunes, un nombre important n'ont pas répondu tous les ans. Cependant, pour cette cohorte, comme pour les jeunes que l'Insee a commencé à suivre plus tard, la majorité ont répondu plusieurs années, souvent jusqu'à la fin du dispositif.

Si l'on considère la dernière date d'observation par la DEPP et/ou le SIES, comme définie plus haut, les réponses à l'Insee sont peu fréquentes pour les jeunes sortis très tôt du suivi scolaire (souvent perdus par la DEPP), mais la proportion de jeunes non retrouvés par EVA est aussi non négligeable, même pour les sortis après 2005, donc issus généralement du suivi SIES : le taux oscille entre 3 % et 15 % (TABLEAU 19 •).

TABLEAU 18 • Première et dernière années de suivi par EVA

Première année de suivi EVA (mars)	Dernière année de suivi EVA (mars)									
	2005	2006	2007	2008	2009	2010	2011	2012	Total	% de parcours complet
2005	447	250	237	223	254	148	139	1555	3253	75,8
2006	0	248	121	127	122	69	71	865	1623	87,0
2007	0	0	175	126	161	55	81	1049	1647	92,1
2008	0	0	0	120	111	48	57	861	1197	92,8
2009	0	0	0	0	221	100	138	2160	2619	95,2
2010	0	0	0	0	0	46	72	1247	1365	96,4
2011	0	0	0	0	0	0	67	647	714	100,0
2012	0	0	0	0	0	0	0	76	76	100,0
Total	447	498	533	596	869	466	625	8460	12494	88,9

Champ : France hors DROM, jeunes entrés en sixième en 1995, suivis par EVA.

Source : DEPP-SIES-Insee, enquête panel 1995.

TABLEAU 19 • Début de suivi par EVA selon la dernière année d'observation de la DEPP et le SIES

Dernière année DEPP SIES (rentrée)	Début de suivi par EVA (mars) (%)									
	Non suivi	2005	2006	2007	2008	2009	2010	2011	2012	Total
1995	95	4	1	0	0	0	0	0	0	82
1996	76	22	2	0	0	0	0	0	0	211
1997	77	20	3	0	0	0	0	0	0	201
1998	71	26	3	0	0	0	0	0	0	538
1999	64	33	4	0	0	0	0	0	0	747
2000	52	41	4	2	0	0	0	0	0	1398
2001	49	42	6	3	1	0	0	0	0	2290
2002	42	40	12	6	0	0	0	0	0	1994
2003	32	2	52	14	0	0	0	0	0	1657
2004	21	12	21	42	5	0	0	0	0	1302
2005	15	4	1	55	23	2	0	0	0	1169
2006	13	4	0	0	67	15	1	0	0	1242
2007	6	3	0	0	0	85	5	0	1	2847
2008	3	3	0	0	0	0	85	6	2	1407
2009	8	5	0	0	0	0	0	82	5	535
2010	6	3	1	0	0	0	0	88	2	210
Total	30	18	9	9	7	15	8	4	0	17830

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Par ailleurs, le décalage entre la dernière observation par la DEPP et/ou le SIES et celle par EVA est, nous l'avons dit, très net pour les jeunes sortis avant 2000 (qui n'ont été interrogés au mieux qu'en mars 2005), mais il n'est pas non plus inexistant pour les jeunes suivis par le SIES. Ainsi, pour les jeunes observés pour la dernière fois par la DEPP et/ou le SIES en 2004-2005, 12,1 % l'ont été par EVA la même année : il ne s'agit pas généralement d'une double interrogation (il y en a cependant parfois eu), mais de jeunes en études repérés par EVA, réintégré dans le suivi du SIES l'année suivante et qui

n'ont pas répondu, d'où une année d'arrêt de suivi SIES en 2004 (en revanche, dans la colonne 2005, il y a de nombreux jeunes renvoyés de la DEPP au SIES, que celui-ci a pu suivre, à partir de 2005-2006, parfois jusqu'en 2010-2011, avant de les retransmettre à EVA ; leur dernière date d'observation DEPP-SIES est alors postérieure à 2004). Parmi les jeunes observés par la DEPP et/ou le SIES une dernière fois en 2004, 20,5 % l'ont été, plus logiquement l'année suivante, 42,5 % seulement en 2006 et presque 5 % à partir de 2007. Le décalage diminue ensuite, mais il faut attendre, du fait d'un changement de règle de transmission, les jeunes observés pour la dernière fois en 2007 par la DEPP ou le SIES pour noter une observation par EVA dès l'année suivante dans la très grande majorité des cas (plus de 80 %).

Fin des études

On a évoqué plus haut la façon de déterminer la date de sortie des études secondaires. Pour l'enseignement supérieur, il faut confronter la date de dernière observation et la date de dernière observation en études. Celle-ci va être déterminée en utilisant à la fois la situation au 31 octobre et la question au 1^{er} mars. Si le jeune se déclare en études à l'une de ces deux questions, on le considère en études l'année scolaire considérée. Il est ainsi possible de déterminer la dernière année d'études observée (on n'a pas vérifié que des années de non études se sont insérées dans le suivi : on suppose que le maintien dans le suivi SIES signifie que l'élève est encore en études initiales).

Trois situations peuvent alors se présenter :

- Pour 813 jeunes, aucune année d'études supérieures n'est observée : ces jeunes ont été interrogés un ou deux fois par le SIES après leur baccalauréat et n'ont jamais indiqué être en études (parfois ils n'ont pas répondu à l'enquête). On utilisera donc les données DEPP pour déterminer leur année de fin d'études.
- Pour 5 080 jeunes, la situation finale repérée par le SIES est une absence d'études : on considèrera que l'année suivant la dernière année d'études observée est l'année de fin d'études (ce n'est pas forcément la dernière année d'interrogation par le SIES, puisqu'au début du dispositif, le SIES réinterrogeait les jeunes ne se déclarant plus en études une première fois, pour être sûr de l'arrêt, avant de les envoyer à l'Insee).
- Pour 3 304 jeunes, la dernière situation observée est une année d'études et ils n'ont plus répondu au SIES ensuite : il y a donc un doute sur le fait que ce soit la dernière année d'études. Les données EVA vont servir pour valider ou non ce choix.

Les règles de synthèse des informations de la DEPP et du SIES sont alors les suivantes :

- Si le jeune a été observé comme non en études par le SIES, on le considère comme sorti l'année suivant sa dernière année d'études.
- Si le jeune n'a jamais été repéré comme en études par le SIES et a été repéré comme sorti par la DEPP, il est considéré comme sorti l'année suivant sa dernière année d'études observée par la DEPP.
- Si le jeune a été observé en études par le SIES, sans que le SIES ne l'observe hors études, il est considéré comme perdu l'année suivant sa dernière année d'études.
- Si le jeune n'a jamais été repéré comme en études par le SIES et a été repéré comme perdu par la DEPP, il est considéré comme perdu l'année suivant sa dernière année d'études observée par la DEPP.

Avec ces règles, on dispose d'une information assez fiable sur la date de sortie pour 12 644 individus ; en revanche, pour 5 186 jeunes, la date de sortie n'est pas connue avec certitude. C'est pour l'année scolaire 2002-2003 que l'on observe le plus de sorties (2 093) : il s'agit généralement de jeunes ayant suivi un CAP, même s'il y a aussi quelques bacheliers à l'heure qui ne poursuivent pas d'études supérieures.

Dans le suivi EVA, la date de fin des « études observées dans EVA » peut être déterminée de deux manières, *a priori* incompatibles :

- la date déclarative de fin d'études demandée lors de la première interrogation si le jeune dit qu'il n'est plus en études (cette question ne sera plus posée dans les interrogations suivantes) ; si par la suite, les années suivantes, il se déclare en études, on considèrera cela comme une reprise d'études ou de la formation continue et ce ne sera pas pris en compte.
- la date suivant la dernière année d'études, quand le jeune s'est déclaré en études lors de la première interrogation et n'a jamais dit avoir arrêté ses études entre temps. Notons que nous ne savons pas si à la première interrogation, il s'agit d'études initiales, puisqu'il y a un décalage entre le moment où la DEPP et le SIES transmettent les données à l'Insee et la date d'interrogation.

Dans le deuxième cas, on peut distinguer deux situations :

- celle où l'année suivant la dernière année d'études observée, le jeune a répondu, en disant donc qu'il n'était pas en études : cela valide la sortie
- celle où l'année suivant la dernière année d'études observée, le jeune n'a pas répondu et sa sortie est alors douteuse (il est seulement « perdu »). Cependant, comme il n'est pas alors possible de la valider par une autre source, on la considèrera bien cette situation comme une sortie.

Ainsi, pour 9 798 jeunes, une date de sortie va pouvoir être construite avec EVA à partir de la date déclarative à la première interrogation ; pour 1 389 jeunes, c'est la série des réponses à la question sur le suivi des études, qui va permettre de repérer l'arrêt, grâce à une observation hors études ; dans 624 cas, de la première à la dernière année d'observation d'EVA, le jeune a été en études : on ne peut donc savoir quand il les a finies (c'est bien sûr la situation des 247 jeunes interrogés par EVA pour la première fois en mars 2012, qui étaient encore en études et n'ont pas été interrogés par la suite). Il existe par ailleurs 683 jeunes ayant répondu à EVA, mais pour lesquels la date de sortie ne peut pas du tout être déterminée : il s'agit généralement de jeunes qui n'étaient pas en études lors de la première interrogation, mais qui n'ont pas voulu répondre à la question sur l'année de fin des études.

La date de sortie calculée sur les données scolaires a ensuite été confrontée avec celle sur les données EVA. Concernant les 12 644 jeunes pour lesquels la date de sortie scolaire est validée par une année de non études, environ un tiers n'a pas de date dans EVA, généralement parce qu'ils n'ont pas répondu à l'enquête (ou comme on l'a vu plus haut, à la question sur la date de fin des études) et cela concerne logiquement plutôt les cohortes sorties précocement. Pour les autres, les années de sortie dans les deux sources se correspondent exactement pour les deux tiers des cas ; on atteint 76 % en tenant compte du mois de sortie déclaré dans EVA : une date de fin d'études dans EVA « novembre 2003 » est compatible avec une dernière année d'études en 2003-2004 et donc une sortie en 2004-2005 ; de même une date de fin d'études dans EVA « janvier 2003 » est compatible avec une sortie scolaire en 2002-2003, si cette année scolaire interrompue n'a pas été prise en compte dans le suivi DEPP-SIES, ce qui pourrait se comprendre.

Il y a un peu plus de 10 % de cas où EVA indique une date de sortie plus ancienne que la sortie scolaire. Pour les sorties précoces, le décalage est souvent de seulement un an, l'éloignement entre la date de sortie et l'année d'interrogation dans EVA peut expliquer ce décalage (petite erreur de mémoire sur un événement éloigné et peut-être peu agréable à évoquer). On compte 4 % de cas où le décalage est d'au moins deux ans. Les premières analyses suggèrent des cas de réorientation vers un niveau inférieur : un jeune ayant un fait un baccalauréat professionnel après une licence va parfois donner comme date de fin d'études celle de sa licence plutôt que son baccalauréat professionnel, considéré plutôt comme une reconversion professionnelle.

A l'opposé, il y a environ 14 % de cas où EVA donne une date de fin d'études postérieure aux sources scolaires : il s'agit souvent de jeunes qui ont été repérés en études dès la première interrogation par EVA (et le retour vers le SIES, effectué pour l'année 2004, complexe à organiser, n'a pas été effectué pour les interrogations suivantes). EVA fournit alors peut-être un complément intéressant au suivi de leurs études par rapport au SIES, même s'il peut aussi s'agir de reprises d'études.

Le même travail peut être mené pour les 5 186 jeunes perdus par la DEPP ou le SIES. La proportion de cas où les données EVA ne sont pas disponibles est un peu plus forte (37 %). Les cohérences sont logiquement moins fréquentes : 38 % en incluant les cas où la prise en compte du mois de sortie valide la confrontation. Si quelques jeunes indiquent dans EVA une date de fin d'études antérieure à leur date de sortie selon les sources scolaires (un peu moins de 10 % des cas), l'écart est généralement dans l'autre sens, ce qui signifie que les sorties douteuses sont souvent infirmées par EVA (53 % des cas), qui fournit donc sans doute une vision plus complète de la scolarité.

Pour établir une synthèse de ces informations, nous proposons les règles suivantes :

- si la sortie a été validée par la DEPP ou le SIES, c'est la date correspondante qui est prise en compte ;
- si la sortie n'est pas validée, c'est-à-dire si l'élève a été perdu par la DEPP où que la dernière année d'observation par le SIES est une année d'études, on prend la date de sortie par EVA si elle est connue ;
- si la date de sortie DEPP/SIES n'est pas validée et que la date de sortie n'est pas connue dans EVA, on prend la date de perte DEPP/SIES comme date de sortie. Ce dernier cas est bien sûr le plus délicat, en particulier pour les jeunes que la DEPP a perdus, parfois très tôt. Dans ce cas, leur date

de fin d'études est incertaine et il sera préférable de les exclure de l'analyse utilisant cette information (en utilisant une pondération pour corriger la perte).

TABLEAU 20 • Date de sortie (première rentrée où l'élève n'est plus en études) combinant l'ensemble des sources

Rentré	Effectif	%
1996	80	0,4
1997	165	0,9
1998	173	1,0
1999	491	2,8
2000	734	4,1
2001	1359	7,6
2002	2308	12,9
2003	2183	12,2
2004	2105	11,8
2005	1730	9,7
2006	1596	9,0
2007	1820	10,2
2008	1467	8,2
2009	751	4,2
2010	409	2,3
2011	221	1,2
2012	238	1,3

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Avec cette définition, les dates de sorties (pouvant inclure décès ou départ à l'étranger) varient entre 1996 et 2012, avec une concentration sur 2001 et 2008 (TABLEAU 20 •).

Observation de la situation professionnelle et fin des études

Pour repérer la situation professionnelle, plusieurs données sont disponibles :

- Dans le suivi DEPP, on dispose d'un motif de sortie.
- Dans le suivi SIES, on dispose ;
 - d'une question sur la situation au 31 octobre, incluant l'emploi, la recherche d'emploi, l'inactivité, à côté du fait d'être en études (les réponses sont en théorie multiples) ;
 - d'une question plus précise sur la situation professionnelle (incluant le type de contrat) pour les personnes ne se déclarant plus en études, à partir de la deuxième année d'interrogation SIES (2003 pour les bacheliers 2002, 2004 pour les bacheliers 2003, etc.).
- Dans le suivi EVA, on dispose de plusieurs questions sur la situation professionnelle, posées à tous.
- Dans tous les questionnaires, il y a aussi des questions sur les revenus (le fait d'en recevoir par un travail régulier), qui pourraient servir à caractériser la situation professionnelle.

Le but du travail à venir est d'étudier l'insertion professionnelle et non le travail étudiant. On a donc privilégié les questions posées aux personnes ayant fini leurs études dans le questionnaire du SIES. On a aussi introduit un filtre pour ne retenir dans EVA que des situations après la date de fin d'études calculées plus haut. On n'a pas tenu compte des informations apportées par les questions sur les revenus.

Logiquement et malheureusement, cette information sera rarement disponible pour les jeunes ayant fini leurs études avant 2000, car ils n'ont généralement pas été retrouvés et n'ont pas répondu à EVA : le taux de non-réponse est de 65,1 % par exemple pour les jeunes ayant fini leurs études en 2000 ou perdus à cette date et il est encore moins bon avant (TABLEAU 21 •). Les quelques jeunes (238) encore

en études en mars 2012 (et dont nous avons mis arbitrairement la date de sortie à la rentrée 2012), dernière d'observation EVA ne pourront pas non plus être caractérisés en termes d'insertion professionnelle. Par ailleurs, pour les jeunes sortis précocement, la première date d'observation (2004 au mieux) sera très postérieure à leur année de fin d'études et à la date réelle d'insertion sur le marché du travail. En revanche, entre 2005 et 2011, pour environ 90 % des sortants, la situation professionnelle peut être repérée, le plus souvent l'année de leur sortie. Sinon, grâce aux questions du SIES sur la situation professionnelle, il y a peu de décalage entre l'année de sortie et la première interrogation sur le sujet.

TABLEAU 21 • Début de l'observation de la situation professionnelle selon la date de fin des études (%)

Année de sortie (rentrée)	Première année d'observation de la situation professionnelle (hors études initiales) en mars...									
	Aucune observation	2005	2006	2007	2008	2009	2010	2011	2012	Total
1996	98	100	0	0	0	0	0	0	0	80
1997	98	100	0	0	0	0	0	0	0	165
1998	89	95	5	0	0	0	0	0	0	173
1999	78	83	13	3	1	0	0	0	0	491
2000	65	86	11	4	0	0	0	0	0	734
2001	54	89	10	2	0	0	0	0	0	1359
2002	49	84	12	3	1	0	0	0	0	2308
2003	39	81	10	8	1	0	0	0	0	2183
2004	25	35	50	15	1	0	0	0	0	2105
2005	16	0	71	25	4	0	0	0	0	1730
2006	10	0	0	93	6	0	0	0	0	1596
2007	7	0	0	0	97	3	0	0	0	1820
2008	7	0	0	0	0	81	17	1	1	1467
2009	9	0	0	0	0	0	94	5	1	751
2010	11	0	0	0	0	0	0	94	6	409
2011	14	0	0	0	0	0	0	0	100	221
2012	100									238
Total	31	28	18	17	15	10	7	3	2	17830

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Dans la suite, nous allons nous intéresser uniquement au 12 307 jeunes pour lesquels on observe une situation professionnelle. Une part importante (8 210) a pu être suivie jusqu'à la dernière année d'observation, dont beaucoup ont été répondants en 2004 (1 739 sur 3500), même s'il est possible qu'ils aient été non répondants lors d'une interrogation intermédiaire (TABLEAU 22 •).

Pour 787 jeunes, aucune situation d'emploi n'est observée sur la période 2004-2011. Dans les autres cas, le jeune a répondu au moins une fois être en emploi, le plus souvent la première année observée de son insertion sur le marché du travail (qui peut ne pas être sa première année réelle), mais pas toujours.

Sur les 12 307 jeunes ayant répondu à la question sur la situation professionnelle, 11 520 ont donc déclaré au moins une fois être en emploi. Dans 909 cas, il n'y a pas de salaire observé (TABLEAU 23 •). Cela peut venir du fait que le jeune était non salarié, qu'il n'a pas souhaité répondre à la question ou qu'elle ne lui a pas été posée (notamment s'il n'a été observé en emploi hors études initiales que dans le suivi SIES, qui ne comportait pas d'informations détaillées sur l'emploi). Dans les autres cas, le salaire n'est pas forcément observé la première année de l'emploi, pour les mêmes raisons qu'au-

dessus. Ainsi, pour les 2 161 jeunes observés pour la première fois en emploi en mars 2007, le salaire n'est connu à cette date que pour 901, un an plus tard pour 197 et deux ans plus tard pour 742.

TABLEAU 22 • Première et dernière année de suivi de la situation professionnelle

Première année de situation professionnelle	Dernière année d'observation de la situation professionnelle								
	2004	2005	2006	2007	2008	2009	2010	2011	Total
2004	397	278	257	242	282	150	155	1739	3500
2005	0	245	139	168	171	85	89	1319	2216
2006	0	0	167	137	197	78	107	1421	2107
2007	0	0	0	134	143	74	104	1349	1804
2008	0	0	0	0	61	33	50	1036	1180
2009	0	0	0	0	0	53	54	781	888
2010	0	0	0	0	0	0	47	341	388
2011	0	0	0	0	0	0	0	224	224
Total	397	523	563	681	854	473	606	8210	12307

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

TABLEAU 23 • Date du premier salaire observé et date du premier emploi observé dans EVA

Première année d'emploi observée	Première année de salaire observé (mars)									
	Aucune observation	2005	2006	2007	2008	2009	2010	2011	2012	Total
2005	138	1673	274	315	55	17	11	1	8	2492
2006	179	0	864	249	602	68	28	11	8	2009
2007	198	0	0	901	197	742	83	28	12	2161
2008	157	0	0	0	540	995	150	33	21	1896
2009	83	0	0	0	0	530	609	82	48	1352
2010	69	0	0	0	0	0	482	222	41	814
2011	50	0	0	0	0	0	0	434	21	505
2012	35	0	0	0	0	0	0	0	256	291
Total	909	1673	1138	1465	1394	2352	1363	811	415	11520

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Bien sûr, le problème est encore plus sensible pour les jeunes sortis précocement, qui ont quitté l'école et ont commencé à s'insérer professionnellement, bien avant la première observation d'EVA. Ainsi, même si 1 673 jeunes sur 2 492 observés pour la première fois en emploi en 2004 ont un salaire renseigné pour cette année, rien n'assure qu'il s'agit du premier salaire perçu.

Quand la personne a exercé un emploi, on récupérera un certain nombre d'informations sur celui-ci (PCS, type d'employeur, type de contrat, temps partiel) pour le premier emploi et pour le dernier. On récupérera aussi ces informations pour le premier emploi où le salaire est observé et pour dernier où il est observé. Enfin, pour les salaires, on raisonnera en euros 2012 en actualisant les salaires antérieurs grâce à l'indice des prix.

mars-12	mars-11	mars-10	mars-09	mars-08	mars-07	mars-06	mars-05
1	1,021	1,037	1,038	1,067	1,083	1,101	1,120

Le niveau d'études

Nous allons maintenant voir comment décrire la scolarité et en particulier déterminer le plus haut niveau d'études atteint et le plus haut diplôme obtenu. D'autres variables peuvent être construites : le dernier niveau de formation suivie, des indicatrices de passages par certains niveaux de formation (pour calculer des taux d'obtention des diplômes correspondants), des indicatrices de redoublements et des nombres d'années d'études passées dans un cursus.

TABLEAU 24 • Plus haut niveau d'études dans les différentes sources

Code	Formation	DEPP	SIES	DEPP+ SIES	EVA	DEP+ SIES+ EVA	Lié au dip
	Non connue dans SIES-EVA		53,0		29,0		
01	Classes de collège	8,4		8,4	1,0	6,8	6,8
02	2 ^{de} GT	1,1		1,1	0,2	0,8	0,8
03	Année non terminale de pro	4,5		4,5	2,8	3,8	3,6
04	Première G&T	1,0		0,9	0,5	0,7	0,7
05	Terminale CAP-BEP	21,2		21,1	9,9	20,0	20,2
06	Terminale bac pro	10,3		7,3	6,3	7,8	7,8
07	Terminale bac tec	19,9		6,6	1,8	5,0	5,1
08	Terminale bac gen	33,6		5,6	3,3	6,1	6,1
11	DEUG et autres bac +2		6,9	6,9	0,7	2,0	2,0
12	DUT		2,3	2,3	1,3	1,2	1,1
13	BTS		11,6	11,6	14,3	14,4	14,2
21	Bac +3		8,2	8,2	14,8	12,0	12,2
31	Bac +4		1,2	1,2		0,9	0,9
41	Bac +5 et thèses		9,8	9,8	14,2	14,3	14,3
42	Grandes écoles		4,4	4,4		4,4	4,4
99	Autres formations		2,6				

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Dans un premier temps, on construit un plus haut niveau d'études à partir des données de la DEPP : les variables CLAS1995-CLAS2005 donnent de l'année scolaire 1995-1996 à l'année scolaire 2005-2006 la classe fréquentée, dans une nomenclature très détaillées (TABLEAU 24 •).

Cette information peut être complétée pour les terminales générales et les terminales technologiques en précisant la série (TABLEAU 25 •).

Pour les données du SIES, trois séries de variables sont utilisées à partir de la rentrée 2002 et pendant 8 ou 9 ans :

- FORM1-FORM9 donnent la formation principale dans laquelle l'étudiant est inscrit au 31/10 passé.
- FORMPAR1-FORMPA9 donnent une autre formation dans laquelle l'étudiant a été inscrit.
- FORMARSB1, FORMARS2-FORMARS8 donnent la formation dans laquelle l'étudiant est inscrit au 01/03 après l'abandon de la formation du 31/10.

Les « autres formations » correspondent essentiellement à des formations de niveau IV et V. Ce code n'est attribué que si aucun des codes 12 et à 42 n'est valable pour une formation suivie.

TABLEAU 25 • Type de baccalauréat

Terminale	Fréquence	Pourcentage
S	2961	16,6
L	1146	6,4
ES	1891	10,6
Pas de terminale générale	11832	66,4
STI	865	4,9
STT	1790	10,0
Autre terminale technologique	888	5,0
Pas de terminale technologique	14287	80,1

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Une synthèse des niveaux d'études issus des données de la DEPP et du SIES est faite en retenant en priorité cette dernière source d'informations.

Dans les enquêtes EVA, deux séries de questions, posées de 2005 à 2012, portent sur le niveau d'études :

- FORMAGR2004-FORMAGR2011 donnent, si la personne se déclare en études, la formation suivie. Il n'est pas impossible que dans quelques cas, il s'agisse en fait d'une reprise d'études. Comme elles se produiraient peu de temps après la fin des études, cela ne nous semble pas trop délicat de les prendre en compte.
- CLASARR2004-CLASS2011 donnent le niveau de la dernière formation suivie si la personne a fini ses études.

Une synthèse de ces deux variables a pu être effectuée pour avoir un niveau de source EVA.

Le niveau d'études issu des sources DEPP et SIES a été confronté avec l'information venant de l'enquête EVA et quand cette deuxième source d'informations donnait un niveau supérieur, elle a été retenue.

Le diplôme

Dans le secondaire, les diplômes passés entre 1997-1998 et 2005-2006 sont donnés dans les variables EXAM1_1998-EXAM1_2006 et EXAM2_1998-EXAM2_2006 pour un éventuel deuxième diplôme préparé. Il faut utiliser les variables RESULT1_1998-RESULT1_2006 et RESULT2_1998-RESULT2_2006 pour savoir si le diplôme a été obtenu. On obtient ainsi un plus haut diplôme obtenu dans les études secondaires (TABLEAU 26 •).

Dans le suivi du SIES, deux séries de variables permettent de repérer l'obtention de diplôme à partir de la troisième année après le bac : TYPDIF3-TYPDIF9 et TYPDIP3B-TYPDIP8B (l'obtention éventuelle d'un deuxième diplôme n'a pas été faite la dernière année de suivi).

La fusion des données du SIES et de la DEPP a été effectuée.

Dans l'enquête EVA, une série de questions est posée en première interrogation sur le plus haut diplôme obtenu (DIPD2004-DIPD2011). Il y a aussi des questions pour les personnes qui ont repris des études, mais il nous semble préférable de ne pas les prendre en compte, car les formulations visent plutôt la formation continue.

Dans EVA, il n'est pas possible de distinguer les différents diplômes de niveau bac + 2, ni ceux au-delà du bac + 3. Dans la suite, l'attribution du diplôme se fera en fonction du plus haut niveau d'études : si une personne a fait des études en IUT et dit avoir eu un diplôme de niveau bac+2 dans EVA, on lui attribuera un DUT.

Cette information a ensuite été confrontée aux diplômes issus des sources DEPP et SIES et on retient le diplôme le plus élevé.

TABLEAU 26 • Plus haut diplôme dans les différentes sources

Code	Diplôme	DEPP	SIES	DEPP+SIES	EVA	DEPP+SIES+EVA	Lié au niveau
	Aucun diplôme connu	16,9	70,7	16,9	30,9		
00	Aucun diplôme				3,2	13,0	13,0
01	DNB	8,1		8,1	3,6	7,5	7,4
02	CAP-BEP	16,0		16,0	11,2	17,4	17,4
03	Bac pro	10,4		9,3	6,7	8,6	8,6
04	Bac tec	16,9		9,2	6,8	9,6	9,4
05	Bac gen	31,8		11,3	5,2	8,9	7,6
11	DEUG et autres bac +2		1,2	1,2	12,0	2,2	1,4
12	DUT		1,4	1,4		1,3	1,2
13	BTS		6,4	6,4		8,2	8,0
21	Bac +3		8,8	8,8	20,5	7,0	6,6
31	Bac +4		2,4	2,4		1,1	3,4
41	Bac +5 et thèses		5,5	5,5		11,6	12,3
42	Grandes écoles		3,7	3,7		3,7	3,7

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Confrontation des niveaux d'études et des diplômes

TABLEAU 27 • Indicateur de réussite d'obtention du diplôme correspondant au plus haut niveau d'études

Code	Niveau	Diplôme échoué	Diplôme réussi
01	Classes de collège	100,0	0,0
02	2 ^{de} GT	100,0	0,0
03	Année non terminale de pro	100,0	0,0
04	Première G&T	100,0	0,0
05	Terminale CAP-BEP	32,5	67,5
06	Terminale bac pro	25,0	75,0
07	Terminale bac tec	32,6	67,4
08	Terminale bac gen	43,8	56,2
11	DEUG et autres bac +2	94,9	5,1
12	DUT	42,7	57,4
13	BTS	46,9	53,1
21	Bac +3	47,2	52,8
31	Bac +4	18,9	81,1
41	Bac +5 et thèses	19,1	80,9
42	Grandes écoles	0,0	100,0

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Dans une dernière étape, le niveau d'études a été confronté au diplôme et une mise en cohérence a été faite (voir dernière colonne des TABLEAU 25 • et TABLEAU 26 •) :

- En dehors du DNB, l'obtention d'un diplôme garantit l'atteinte du niveau (donc si pour une personne, le plus haut niveau d'études connu est la terminale, mais qu'elle déclare un BTS, on lui attribue le BTS comme plus haut niveau d'études).
- À l'inverse, on construit aussi un diplôme qui suppose que pour atteindre un certain niveau d'études, il faut avoir obtenu un diplôme au niveau en-dessous (ainsi, une personne déclarant être allé jusqu'en BTS, sans obtenir ce diplôme, se verra attribuer un bac général, même si ce n'est pas dans les données).

Avec un tel niveau de cohérence, il est possible de résumer l'information en retenant le niveau d'études et en ajoutant une indicatrice pour repérer si le diplôme correspondant a été obtenu ou non (avec quelques aménagements : le fait d'avoir un diplôme de bac +5 valide une formation en école d'ingénieur par exemple ; le DNB n'est censé valider aucune formation) (TABLEAU 27 •).

Évaluation de compétences

Le panel 1995 comporte plusieurs types d'évaluation de compétences :

- Des appréciations par les chefs d'établissement (mais aussi par les parents dans l'enquête Famille et par l'élève lui-même dans l'enquête Jeunes) ;
- Des évaluations standardisées spécifiques au panel (en civisme et méthodes de travail) ;
- Des évaluations standardisées exhaustives (à l'entrée en 6^e) ;
- Des notes de contrôle continu (pour le DNB : dans le cas où l'élève a redoublé et passé deux fois le DNB, on prend les notes de la deuxième année) ;
- Des notes à l'examen (pour le baccalauréat : là encore en cas de redoublement, on ne retiendra que les notes de la dernière passation).

Nous allons décrire ici le processus de standardisation et d'imputation quand l'une de ces informations n'est pas disponible.

L'appréciation des compétences en français et en mathématiques par les chefs d'établissement de 6^e, sur une échelle de 1 à 10, est l'information la mieux remplie. Elle manque pour seulement 18 élèves. On leur impute la valeur médiane 5. Ces appréciations ont ensuite été standardisées sur une moyenne de 0 et un écart-type de 1.

Le nombre de non-réponses aux autres évaluations varie sensiblement selon leur nature et le niveau (TABLEAU 28 •) : il reste inférieur à 1 000 pour les évaluations en 6^e (autour de 500 pour les évaluations spécifiques en civisme et méthodes de travail ; un peu au-dessus de 800 pour les évaluations nationales). Ce nombre dépasse 4 000 au DNB, car un certain nombre d'élèves ne le présentent pas et dépasse 7 000 pour le baccalauréat logiquement encore moins souvent passé.

TABLEAU 28 • Nombre de non-répondants dans les différentes évaluations

Evaluation	Nb NR
Évaluation 6 ^e mathématiques	858
Évaluation 6 ^e français	819
Note de contrôle continu mathématiques	4381
Note de contrôle continu français	4379
Note mathématiques bac	7912
Note total bac	7167
Note français bac	8806
Évaluation civisme	477
Évaluation méthodes de travail	481

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Pour les scores en civismes et méthodes de travail, on utilise la valeur modale pour les non-réponses.

Pour les scores aux évaluations 6^e, on utilise les appréciations des chefs d'établissement. On aurait pu avoir recours à une modélisation économétrique ; on préfère utiliser une méthode par ancrage qui préserve mieux la structure des données. Le principe est le suivant :

- On calcule la moyenne et l'écart-type de l'appréciation des chefs d'établissement : sur les répondants à l'évaluation en 6^e, *Mr* et *Er*, et sur les non-répondants, *Mnr* et *Enr*.
- On cale la moyenne et l'écart-type du score à l'évaluation 6^e des répondants à cette évaluation sur les valeurs *Mr* et *Er* ; pour les non-répondants, on prend les appréciations des chefs d'établissement standardisés.
- Le score ainsi imputé préserve ainsi l'information de l'évaluation 6^e pour les répondants, l'ordre de l'appréciation pour les non-répondants et établit un écart entre non-répondants et répondants à l'évaluation nationale, correspond à celui observé en termes d'appréciation.

La même méthode est utilisée pour imputer les notes manquantes de contrôle continu de mathématiques et de français en 3^e en utilisant le score à l'évaluation 6^e dans la même discipline comme variable de référence.

TABLEAU 29 • Corrélation entre les évaluations

		1	2	3	4	5	6	7	8	9
Appréciation maths	1	1	17830	16972	17011	13449	13451	9918	9024	10663
Appréciation français	2	0,74	1	16972	17011	13449	13451	9918	9024	10663
Évaluation 6 ^e maths	3	0,76	0,64	1	16795	12911	12914	9496	8649	10210
Évaluation 6 ^e français	4	0,65	0,72	0,76	1	12938	12941	9498	8647	10210
Note de CC maths	5	0,55	0,47	0,60	0,49	1	13447	8696	7962	9308
Note de CC français	6	0,49	0,58	0,52	0,61	0,67	1	8697	7963	9309
Note maths bac	7	0,21	0,15	0,24	0,18	0,27	0,19	1	8947	9888
Note français bac	8	0,21	0,28	0,23	0,31	0,23	0,36	0,20	1	8992
Note total bac	9	0,25	0,25	0,28	0,26	0,30	0,31	0,58	0,44	1

Champ : France hors DROM, jeunes entrés en sixième en 1995.
Source : DEPP-SIES-Insee, enquête panel 1995.

Pour le baccalauréat, la logique est un peu différente, car on va aussi tenir compte du fait que les notes ne sont pas comparables d'une série à l'autre. On va donc aussi ancrer pour une série donnée les moyennes et écarts-type des notes en français, mathématiques et sur l'ensemble des épreuves sur les valeurs correspondantes de cette population en 3^e (en faisant la moyenne français et mathématiques pour la note totale au baccalauréat).

Il convient bien sûr de vérifier que les corrélations entre les évaluations sont assez bonnes pour procéder à ces imputations avec une telle méthode. Sans être parfaites, elles sont assez correctes (TABLEAU 29 • : la partie basse indique les corrélations ; la partie haute les effectifs d'élèves pour lesquels les deux évaluations comparées sont connues en même temps) :

- Le score aux évaluations 6^e en mathématiques et l'appréciation des chefs d'établissement dans cette matière sont corrélés à 0,76 (0,72 pour le français).
- La note au contrôle continu de 3^e en mathématiques et le score aux évaluations nationales dans cette matière sont corrélés à 0,60 (0,61 en français).
- La note au baccalauréat en mathématiques et la note de contrôle continu dans cette matière en 3^e sont corrélées à 0,27 (0,36 pour le français). La note au baccalauréat sur l'ensemble des matières est corrélée à 0,30 avec la note de contrôle continu en mathématiques en 3^e et 0,31 avec celle en français. Le moindre degré de corrélation tient à l'effet « série » : alors que l'orientation envoie dans les différentes séries des élèves d'un niveau de compétences en mathématiques et en français très différent, les notes au baccalauréat, ajustées aux exigences différentes des séries dans ces matières, vont donner une vision atténuée de ces écarts.

Des imputations ont été effectuées en 6^e et en 3^e : les répondants à l'évaluation 6^e en mathématiques ont une appréciation moyenne par les chefs d'établissement de 0,01 (avec un écart-type de 0,99) ; on va caler la distribution de leur score à l'évaluation 6^e sur ces valeurs ; pour les non-répondants à

cette évaluation, les appréciations sont moins bonnes, avec une moyenne de - 0,18 (et un écart-type de 1,11); ces appréciations sont prises directement pour imputer leur scores aux évaluations 6^e (TABLEAU 30 •). L'écart entre non-répondants et répondants est encore plus net aux notes de contrôle continu en 3^e : les élèves sans note en mathématiques ont un score à l'évaluation nationale dans cette discipline (éventuellement imputé à l'aide des appréciations) de - 0,55 contre 0,18 pour ceux qui ont une note en 3^e. La note de contrôle continu en LV1 a connu le même traitement, avec un calage sur la moyenne des notes de mathématiques et français, en cas de non-réponse.

TABLEAU 30 • Différences de résultats entre répondants/non-répondants

	Répondants		Non-répondants	
	Moyenne	Ec-type	Moyenne	Ec-type
Évaluation 6e maths	0,01	0,99	-0,18	1,11
Évaluation 6e français	0,01	0,99	-0,20	1,14
Note de CC maths	0,18	0,88	-0,54	1,13
Note de CC français	0,18	0,88	-0,55	1,14

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Pour le baccalauréat, il y a à la fois un processus d'imputation par imputation sur les notes de 3^e (qui conduit à attribuer une valeur assez faible aux élèves sans note au baccalauréat : - 0,58 en mathématiques par exemple) et un processus de standardisation (qui va donner une note moyenne en mathématiques de 1,08 aux élèves de S contre 0,31 en terminale L et - 0,29 pour le baccalauréat professionnel).

Pour affiner la mesure des compétences, nous avons pris en compte d'autres disciplines que les mathématiques et le français, dans les résultats au baccalauréat, quand suffisamment d'élèves étaient concernés (TABLEAU 31 •) :

- Français oral (voir plus haut il s'agissait de la note à l'écrit).
- LV1
- Économie-droit
- Sciences (pour bac L ou ES)
- Histoire-géographie
- Philosophie
- Physique
- SVT
- Sciences économiques et sociales

TABLEAU 31 • Différences de résultats selon les séries

	N	Mathématiques		Français		Total	
		Moyenne	Ec-type	Moyenne	Ec-type	Moyenne	Ec-type
Pas de bac	7168	-0,58	0,97	-0,56	0,98	-0,61	0,96
S	2842	1,08	0,59	0,85	0,72	1,03	0,64
ES	1847	0,51	0,62	0,62	0,66	0,61	0,61
L	1111	0,31	0,69	0,80	0,68	0,59	0,66
STI	851	0,35	0,63	-0,05	0,68	0,16	0,60
STG	1749	0,03	0,66	0,12	0,63	0,08	0,57
Autre techno	572	0,07	0,68	0,13	0,63	0,11	0,60
Professionnel	1690	-0,29	0,73	-0,39	0,69	-0,38	0,65

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Le même processus de standardisation que pour le français et les mathématiques a été appliqué, en se calant pour les non-répondants sur la discipline de 3^e la plus proche (les mathématiques par exemple pour la physique). Ceci dit, pour ne pas introduire de trop fortes colinéarités, les valeurs pour les non-répondant ont ensuite été rendues constantes (en fixant au niveau de la moyenne observée). Les notes en physiques ne varient donc que pour les individus ayant passé cette épreuve et n'apporte de l'information que pour eux (le processus d'imputation sert à standardiser proprement la variable).

Pondération

La première partie de ce document décrit le parcours scolaire et est fondé sur les 16 246 élèves que la DEPP a pu suivre ou qui, perdus par la DEPP, ont été retrouvés dans EVA. Cela exclut complètement 386 élèves perdus définitivement (morts ou partis à l'étranger). Il y a donc 1 198 élèves, non-répondants, dont il paraît pertinent de corriger l'absence par une pondération. On a construit un modèle de non-réponse en fonction des données disponibles (PCS des parents, sexe, âge et niveau à l'entrée en 6^e...) pour l'ensemble des jeunes du panel (en excluant l'attrition naturelle, due aux décès et aux départs à l'étranger). La pondération est l'inverse de la probabilité de répondre.

Ce sont les élèves les plus favorisés (enfant de cadres supérieurs, ayant eu de bons résultats à l'entrée en 6^e, de PCS favorisée...) qui tendent à mieux répondre que les autres (TABLEAU 32 •).

TABLEAU 32 • Modélisation de la non-réponse au suivi scolaire (Modèle 1) ou à EVA conditionnellement à un suivi scolaire (Modèle 2)

			Modèle 1	Modèle 2
Constante	Référence	Modalité	-2,55***	-0,2
Sexe	Fille	Garçon	0,16***	0,04
Pays de naissance du père	Inconnu	France	-0,15	-0,1
		Etranger	0,12	0,12
Pays de naissance de la mère	Inconnu	France	-0,08	-0,11*
		Etranger	-0,12	-0,02
Activité du père	Inconnue	En emploi	0	-0,22**
		Sans emploi	-0,19	0,08
Activité de la mère	Inconnue	En emploi	-0,12	-0,18
		Sans emploi	0,37***	0
PCS du père	Inconnue	Agriculteur	-0,86***	-0,49***
		Artisan-commerçant	0,14	0,06
		Cadre supérieur	0,02	0,41***
		Profession intermédiaire	-0,16	0,13**
		Employé	0,09	0,05
		Ouvrier	-0,15**	-0,08
		Sans profession	0,66***	-0,04
PCS de la mère	Inconnue	Agricultrice	-0,91*	-1,02***
		Artisane-commerçante	0,6***	0,1
		Cadre supérieure	-0,09	0,38***
		Profession intermédiaire	0,14	0,24***
		Employée	0,25**	0,14**
		Ouvrière	0,16	0,15*
		Sans profession	-0,33**	0,03
Taille de la fratrie	4 ou +	1	0,02	0,06
		2	-0,05	-0,06*

		3	-0,06	-0,05
Âge à l'entrée en 6 ^e	13 ans ou +	10 ans ou -	-0,12	0,13
		11 ans	-0,14*	0,01
		12 ans	0,15*	-0,03
Score évaluation 6 ^e mathématiques	Tiers +	Tiers -	0,32***	0,08**
		Tiers 0	0,14***	0,01
Score évaluation 6 ^e français	Tiers +	Tiers -	0,27***	-0,01
		Tiers 0	0,08	-0,02
Année d'entrée dans l'élémentaire	Inconnue	1989	0	-0,1*
		1990	-0,09	-0,04
Chg d'établissement	Oui	Non		-0,13***
Année du bac	2006	Pas de bac		-0,14
		2001		-2,13**
		2002		0,61***
		2003		0,41**
		2004		0,12
		2005		0,21
Année de sortie	2012	1996		-0,85
		1997		0,28
		1998		-0,05
		1999		-0,19
		2000		-0,43
		2001		-0,26
		2002		-0,33
		2003		-0,63
		2004		-1,22
		2005		-1,46
		2006		-1,98
		2007		-2,13
		2008		-2,27
		2009		-2,05
2010		-1,65		
2011		-1,43		
Niveau de sortie	Grandes écoles	Classes de collège		0,48***
		2nde GT		0,18
		Année non terminale de pro		0,4***
		Première G&T		0,2
		Terminale CAP-BEP		0,24***
		Terminale Bac pro		-0,36***
		Terminale Bac tec		0,41***
		Terminale Bac gen		0,78***
		Deug et autres bac+2		0,58***
		DUT		0,09

		BTS	-0,62***
		Bac+3	-0,57***
		Bac+4	-0,1
		Bac+5 et thèses	-0,68***
Réponse à l'enquête Famille	Non réponse	Réponse postale	-0,28***
		Réponse téléphonique	-0,01
Réponse à l'enquête jeunes	Non interrogés	Réponse postale	0,12
		Réponse téléphonique	0,1
		Non répondants	1,15***
Élèves boursiers en 6 ^e	Non	Oui	-0,04

Champ : France hors DROM, jeunes entrés en sixième en 1995.

Source : DEPP-SIES-Insee, enquête panel 1995.

Pour les analyses portant sur l'accès à l'emploi ou le salaire, c'est la réponse à EVA qui importe, en particulier les 12 307 pour lesquels on a observé une situation professionnelle. Nous avons procédé en deux étapes :

- Pour les 16 246 élèves suivis scolairement, on construit un modèle de non-réponse avec les mêmes variables qui ont servi à modéliser l'attrition du suivi scolaire et quelques autres variables ayant un sens pour ces élèves répondants scolairement (réponse aux enquêtes précédentes (enquête Famille et enquête Jeunes 2002), date de sortie, niveau d'études...).
- On multiplie cette probabilité de répondre à EVA sachant que l'élève a été suivi scolairement, par la probabilité d'être suivi scolairement, obtenue dans le premier paragraphe, ce qui donne une estimation de la probabilité de répondre à EVA, que l'on inverse pour avoir la deuxième pondération.

Cette procédure en deux étapes, qui utilise le fait que les non-suivis scolairement sont par construction non-répondants à EVA (cela fait partie de leur définition), assure une plus grande cohérence entre les deux pondérations, tout en permettant de tenir compte dans le deuxième modèle de non-réponse d'informations peu fiables pour les non suivis scolairement (la date de sortie notamment). Ainsi les résultats sur le parcours scolaire sur les répondants à EVA avec la deuxième pondération sont proches de ceux obtenus sur tous les suivis scolairement, avec la première pondération.

REFERENCES DES PUBLICATIONS DE LA DEPP



Barret, C., Ryk, F., Volle, N. (2014). Face à la crise, le fossé se creuse entre niveaux de diplôme. *Bref du Cereq*, n°319, Cereq.

Barhoumi, M., Caille, J.-P. (2021). Les élèves sortent de l'enseignement secondaire de plus en plus diplômés, mais au terme de parcours scolaire encore socialement différenciés. *Education et formations*, n°101, 323-3571.

Ben Ali, L., Vourc'h, R. (2015). Evolution des acquis cognitifs au collège au regard de l'environnement de l'élève – Constat et mise en perspective longitudinale. *Éducation & formations*, 86-87, pp. 211–233

Caille J. P., 2017, « Les panels d'élèves de la DEPP : source essentielle pour connaître et évaluer le système éducatif », *Education et formations*, n°95, pp 5-31.

Cereq. (2008). *Quand l'école est finie – Premiers pas dans la vie active de la génération 2004*, Cereq.

Cereq. (2014). *Quand l'école est finie – Premiers pas dans la vie active de la génération 2010*, Cereq.

Defresne, F., Krop, J. (2016). Massification scolaire sous la Ve République : une mise en perspective des statistiques de l'Education nationale (1958-2014). *Éducation & formations*, 91, 5–20

Degorre, A., Guyon, V., Moisy, M. (2009). Sur dix jeunes entrés en 6e en 1995, neuf se déclarent en bonne ou très bonne santé en 2007. *Insee Première*, n°1261.

DEPP. (2010). Le niveau de diplôme. *Etat de l'école*, fiche 26.

Despalins, R., de Saint-Pol, T. (2012). L'entrée dans la vie adulte des bacheliers sous l'angle du logement. *Etudes et résultats*, n°813.

El Jammal, M.-H., Cornali, D., Stefanou, A. (1996). Les méthodes de travail des élèves de sixième. *Note d'information*, n°96.24

Emin, J.-C., Esquieu, P. (1998). Un siècle d'éducation. *Éducation & formations*, 54, 71–80

Goux, D., Maurin, É. (1997). Démocratisation de l'école et persistance des inégalités. *Économie et statistique*, 306, 27–39

Héraud, Q., Stéfanou, A. (1997). Connaissances civiques et comportements en société : robustesse des instruments de mesure. *Éducation & Formations*, 50, 81–89

Insee. (2009). *Formation et emploi*, Insee Référence, Insee

Lefebvre, O. (2012). Les inégalités dans l'accès aux hauts diplômes se jouent surtout avant le bac. *France Portrait Social*, 2012.

Lemaire, S. (2006). Le devenir des bacheliers : parcours après le baccalauréat des élèves entrés en sixième en 1989. *Note d'information*, n°06.01.

Lemaire, S., Guyon, V., Murat, F. (2007). Un élève sur deux entrés en 6^e en 1995 fait des études 10 ans plus tard. *Insee Première*, n°1158

Lemaire, S. (2012). Les parcours dans l'enseignement supérieur : devenir après le baccalauréat des élèves de entrés en sixième en 1995. *Note d'information Enseignement supérieur & Recherche*, n°12.05.

Le Rhun, B., Monso, O. (2015). De l'utilité d'obtenir son diplôme pour s'insérer : l'exemple des brevets de technicien supérieur. *Économie et statistique*, n°478-479-480.

Murat, F. (2021). Les inégalités de compétences à la fin des études. *Économie et Statistique / Economics and Statistics*, n°528-529, Insee.

- Murat, F. (2022). Que savent les jeunes à la fin de leurs études ? Les compétences en lecture, en calcul et en compréhension orale selon le parcours scolaire. *Éducation & formations*, 103, 83–98
- Rocher, T. (2016). Construction d'un indice de position sociale. *Éducation & Formations*, 90, 5–27
- Thélot, C., Vallet, L.-A. (2000). La réduction des inégalités sociales devant l'école depuis le début du siècle. *Economie et Statistique*, 334, 3–32.
- Vallet, L.-A., Caille, J.-P. (1996). Niveau en français et en mathématiques des élèves étrangers ou issus de l'immigration. *Economie et Statistiques*, n°293, 137-153

Retrouvez les travaux de la DEPP sur
education.gouv.fr/etudes-et-statistiques

Publications et archives

Retrouvez toutes les publications et archives de la DEPP sur
archives-statistiques-depp.education.gouv.fr

Jeux de données en open data

Retrouvez tous les jeux de données de la DEPP en open data sur
data.education.gouv.fr